

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
INSTITUTO DE INFORMÁTICA
PROGRAMA DE PÓS-GRADUAÇÃO EM COMPUTAÇÃO

DÊNIS FERNANDES

**ONNIS-GI: Uma Rede Neural Oscilatória para
Segmentação de Imagens Implementada em
Arquitetura Maciçamente Paralela**

Tese apresentada como requisito parcial
para obtenção do grau de Doutor em
Ciência da Computação

Prof. Philippe Olivier Alexandre Navaux
Orientador

Porto Alegre, outubro de 2004.

CIP – CATALOGAÇÃO NA PUBLICAÇÃO

Fernandes, Dênis

ONNIS-GI: Uma Rede Neural Oscilatória para Segmentação de Imagens Implementada em Arquitetura Maciçamente Paralela / Dênis Fernandes. – Porto Alegre: Programa de Pós-Graduação em Computação, 2004.

165 f. : il.

Tese (doutorado) – Universidade Federal do Rio Grande do Sul. Programa de Pós-Graduação em Computação, Porto Alegre, BR-RS, 2004. Orientador: Philippe Olivier Alexandre Navaux.

1. Redes neurais oscilatórias. 2. Arquiteturas maciçamente paralelas. 3. Segmentação de imagens. I. Navaux, Philippe Olivier Alexandre. II. Título.

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL

Reitor: Prof. José Carlos Ferraz Hennemann

Vice-Reitor: Prof. Pedro Cezar Dutra Fonseca

Pró-Reitora de Pós-Graduação: Prof^a. Valquiria Link Bassani

Diretor do Instituto de Informática: Prof. Philippe Olivier Alexandre Navaux

Coordenador do PPGC: Prof. Carlos Alberto Heuser

Bibliotecária-Chefe do Instituto de Informática: Beatriz Regina Bastos Haro

“os olhos percebem não só alterações no mundo exterior, que são importantes para a percepção da vida, como também as características dos objetos que os fazem ser escolhidos como objetos de amor - seus encantos”

Sigmund Freud

à minha família, Lu, Vic e Erico, a
verdadeira tese de minha vida

AGRADECIMENTOS

Agradeço ao Professor Philippe Navaux, pelo acolhimento e pelo exemplo de conduta pessoal e profissional. Uma verdadeira orientação, a qual se fez sobretudo com atitudes e não somente com palavras.

Agradeço aos Professores Felipe França, Fernando Osório e Paulo Engel, membros da banca examinadora, pelos questionamentos e sugestões, os quais implicaram evolução significativa da proposta originalmente apresentada.

Agradeço também aos colegas do SEGIME pelo apoio e incentivo, aos funcionários do Instituto de Informática da UFRGS pela prestatividade e competência e aos colegas da PUCRS e amigos que incentivaram e contribuíram de alguma forma para realização deste trabalho.

Finalmente, agradeço à PUCRS pela liberação parcial de carga horária, sem a qual a realização desta tese seria inviável.

SUMÁRIO

LISTA DE ABREVIATURAS E SIGLAS	8
LISTA DE FIGURAS	9
LISTA DE TABELAS	12
RESUMO	13
ABSTRACT	14
1 INTRODUÇÃO	15
2 FUNDAMENTAÇÃO	19
2.1 Segmentação de Imagens	19
2.1.1 Procedimentos para Segmentação de Imagens.....	20
2.1.2 Características Utilizadas em Segmentação de Imagens	21
2.1.3 Segmentação Utilizando Filtros de Gabor	23
2.1.4 Segmentação Baseada em Detecção de Bordas.....	26
2.1.5 Influência do Ruído na Segmentação de Imagens	28
2.1.6 Avaliação da Qualidade da Segmentação	29
2.2 Redes Neurais em Segmentação de Imagens	30
2.2.1 Técnicas de Segmentação de Imagens Baseadas em Redes Neurais.....	31
2.2.2 Redes Neurais Oscilatórias.....	32
2.3 Chips de Visão com Arquiteturas Maciçamente Paralelas	33
2.4 Conclusão	34
3 A REDE NEURAL OSCILATÓRIA LEGION	36
3.1 Neurônio Oscilador de Terman-Wang	36
3.1.1 Modelo do Oscilador de Terman-Wang.....	37
3.1.2 Parâmetros do Oscilador de Terman-Wang	39
3.2 Estrutura da Rede	40
3.2.1 Neurônio Oscilador Básico	41
3.2.2 Sincronismo Entre Osciladores Acoplados	42
3.2.3 Inibição Global	43
3.2.4 Uma Rede LEGION Unidimensional	43
3.2.5 Normalização dos Pesos Dinâmicos	45
3.2.6 Osciladores Líderes.....	46
3.3 Algoritmo para Implementação da Rede	48
3.4 Outros Aspectos Relacionados à Rede LEGION	53
3.4.1 Capacidade de Segmentação	54
3.4.2 Percepção Dúplex	54
3.4.3 O Algoritmo FPS	55
3.5 Aplicações Práticas Desenvolvidas	58
3.5.1 Segmentação de Caracteres de Placas de Automóveis	58
3.5.2 Segmentação de Imagens de Microscopia Eletrônica.....	59
3.5.3 Segmentação de Imagens de Ecocardiografia Fetal	62

3.6	Conclusão.....	63
4	ONNIS-GI: PROPOSTA DE UMA REDE NEURAL OSCILATÓRIA PARA SEGMENTAÇÃO DE IMAGENS	65
4.1	Motivação.....	65
4.2	As Redes ONNIS e ONNIS-GI	65
4.2.1	Elemento de Processamento da Rede ONNIS.....	66
4.2.2	Estruturas de Conexões Excitatórias e Inibitórias	70
4.2.3	EPs Líderes.....	72
4.2.4	Inicialização Global	73
4.2.5	Elemento de Processamento da Rede ONNIS-GI	73
4.2.6	Algoritmo para Implementação da Rede ONNIS-GI.....	75
4.3	Segmentação com Base em Atributos Diversos	76
4.3.1	Segmentação de Imagens em Níveis de Cinza.....	76
4.3.2	Segmentação Utilizando Atributos de Cor.....	77
4.3.3	Segmentação Utilizando Atributos de Textura.....	77
4.3.4	Segmentação Utilizando Filtros de Gabor	78
4.4	Determinação dos Pesos Através de Detecção de Bordas	80
4.4.1	Os Pesos da Rede como Bordas da Imagem	80
4.4.2	Um Procedimento Robusto para Determinação dos Pesos	82
4.5	Conclusão.....	84
5	RESULTADOS OBTIDOS COM A REDE ONNIS-GI	86
5.1	Procedimento de Avaliação da Qualidade da Segmentação	86
5.2	Ambiente de Simulação	89
5.3	Resultados Obtidos na Segmentação de Imagens Artificiais	90
5.4	Resultados Obtidos na Segmentação de Imagens Naturais	91
5.5	Desempenho da Rede ONNIS-GI na Presença de Ruído	92
5.5.1	Resultados Obtidos sem Filtragem Preliminar da Imagem.....	92
5.5.2	Resultados com o Procedimento Robusto Proposto	94
5.6	Efeito da Variação de Iluminação na Imagem de Entrada	96
5.7	Comparação com um Procedimento de Segmentação Clássico	98
5.8	Conclusão.....	99
6	IMPLEMENTAÇÃO EM FPGA	101
6.1	Rede ONNIS	101
6.2	Rede ONNIS-GI.....	106
6.3	Estimativa da Velocidade de Processamento.....	109
6.4	Conclusão.....	110
7	CONCLUSÃO E DESDOBRAMENTO DO TRABALHO	112
7.1	Conclusão Geral	112
7.2	Desdobramento da Pesquisa.....	113
7.3	Artigos a Serem Elaborados a Partir da Tese	114
	REFERÊNCIAS	115
	GLOSSÁRIO	121
	APÊNDICE A APLICATIVOS DESENVOLVIDOS	122
	ANEXO A ARTIGOS PUBLICADOS	129

LISTA DE ABREVIATURAS E SIGLAS

ALN	<i>Adaptive Logic Network</i>
ARMA	<i>Auto Regressive - Moving Average</i>
bit	<i>binary digit</i>
BMP	BitMaP (formato de arquivos de imagens utilizado no sistema Windows)
CCD	<i>Charge-Coupled Device</i>
CSNN	<i>Constraint Satisfaction Neural Network</i>
EP	Elemento de Processamento
FIR	<i>Finite Impulse Response</i>
FPGA	<i>Field Programmable Gate Array</i>
FPS	<i>Feature-Preserving Smoothing</i>
fps	<i>frame per second</i>
LED	<i>Light-Emitting Diode</i>
LEGION	<i>Locally Excitatory Globally Inhibitory Oscillator Network</i>
MLP	<i>MultiLayer Perceptron</i>
OCR	<i>Optical Character Recognition</i>
ONNIS	<i>Oscillatory Neural Network for Image Segmentation</i>
ONNIS-GI	<i>Oscillatory Neural Network for Image Segmentation with Global Inhibition</i>
pixel	<i>picture element</i>
RSR	Relação Sinal/Ruído
SIMD	<i>Single Instruction stream Multiple Data stream</i>
SOFM	<i>Self-Organizing Feature Map</i>
texel	<i>texture element</i>
UFRGS	Universidade Federal do Rio Grande do Sul
VE	Ventrículo Esquerdo

LISTA DE FIGURAS

Figura 1.1: Etapas de um procedimento genérico para análise de imagens	16
Figura 2.1: Componentes real e imaginária da resposta ao impulso de um filtro de Gabor	25
Figura 2.2: Resposta em frequência do filtro de Gabor representado pela resposta ao impulso da Figura 2.1	25
Figura 2.3: Utilização de um filtro de Gabor para detectar texturas com frequência espacial de $u_0=0,1$ e $v_0=0,1$ períodos/pixel. (a) Imagem original. (b) Módulo da imagem filtrada com $\alpha=50$ e $\beta=50$. (c) Módulo da imagem filtrada com $\alpha=400$ e $\beta=400$	26
Figura 2.4: Utilização de um filtro de Gabor para detectar componentes específicas em textura complexas. (a) Módulo da imagem filtrada com $\alpha=400$, $\beta=400$, $u_0=0$ e $v_0=0,1$ períodos/pixel. (b) Soma das imagens (c), (d), (e) e (f).....	26
Figura 2.5: Integração dos sensores e rede de EPs em um <i>chip</i> de visão	34
Figura 3.1: Plano de fase do oscilador de Terman-Wang e <i>nullclines</i> obtidas para $I=1$, $\beta=0,2$ e $\alpha=3$	38
Figura 3.2: Trajetória de fase do oscilador de Terman-Wang para $I=1$, $\beta=0,2$, $\alpha=3$. Exemplos de condições iniciais: A, B e C	38
Figura 3.3: Saída do oscilador de Terman-Wang em função do tempo para $I=1$, $\beta=0,2$, $\alpha=3$, $\varepsilon=0,01$, $x(0)=0$ e $y(0)=0$	39
Figura 3.4: Efeito de I na condição de oscilação oscilador de Terman-Wang com $I=-1$ (a), $I=3$ (b), $\beta=0,2$ e $\alpha=3$	39
Figura 3.5: Efeito da variação de I na saída do oscilador de Terman-Wang. $I=0,5$ (a), $I=1,5$ (b), $\beta=0,2$, $\alpha=3$, $\varepsilon=0,01$, $x(0)=0$ e $y(0)=0$	40
Figura 3.6: Exemplo de topologia de uma rede LEGION bidimensional	40
Figura 3.7: Mecanismo de sincronismo entre dois osciladores acoplados	42
Figura 3.8: Mecanismo de inibição entre dois osciladores não acoplados	43
Figura 3.9: Rede LEGION unidimensional com 4 osciladores e conexão em anel	44
Figura 3.10: Saídas dos osciladores e do inibidor global da rede exemplo com 3 padrões de entrada distintos ($I_1=0,3$, $I_2=0,4$, $I_3=0,5$ e $I_4=0,5$).....	44
Figura 3.11: Saídas dos osciladores e do inibidor global da rede exemplo com 2 padrões de entrada distintos ($I_1=0,3$, $I_2=0,3$, $I_3=0,5$ e $I_4=0,5$).....	45
Figura 3.12: Saídas dos osciladores e do inibidor global da rede exemplo com 2 padrões de entrada distintos ($I_1=0,3$, $I_2=0,5$, $I_3=0,3$ e $I_4=0,5$).....	45
Figura 3.13: Segmentação com rede LEGION. Imagem de entrada (a). Saídas em tempos distintos (b), (c), (d), (e) e (f).....	48
Figura 3.14: Relação não linear entre os pesos e as diferenças entre entradas adjacentes para uma imagem de 8 bits	51

Figura 3.15: Segmentação obtida com a utilização da versão algorítmica da rede LEGION com parâmetros $\theta_p=4$ e $W_z=0,5$	53
Figura 3.16: Fenômeno da percepção dúplex no resultado da segmentação obtido com a rede LEGION	55
Figura 3.17: (a) Imagem original e (b) imagem corrompida por ruído.....	56
Figura 3.18: Imagens filtradas com o algoritmo FPS: (a) 1 iteração, (b) 10 iterações, (c)1000 iterações e (d) 10000 iterações	57
Figura 3.19: Imagem original filtrada com 100 iterações do algoritmo FPS	57
Figura 3.20: Imagem de uma placa de automóvel obtida com uma câmera fotográfica digital	58
Figura 3.21: Segmentos obtidos com a aplicação da imagem da Figura 3.20 a uma rede LEGION.....	59
Figura 3.22: Imagem captada de um TEM representando uma amostra de silício com bolhas de hélio.....	60
Figura 3.23: Região extraída da imagem da Figura 3.22	60
Figura 3.24: Imagem da Figura 3.23 filtrada com o algoritmo FPS	60
Figura 3.25: Resultado da segmentação da imagem da Figura 3.24 com uma rede LEGION	61
Figura 3.26: Segmento relativo ao fundo de silício homogêneo	61
Figura 3.27: Pixels não segmentados pela rede (representados em preto).....	61
Figura 3.28: Segmentação incorreta do VE. (a) Imagem original com indicação da região a ser segmentada. (b) Segmentação SOM/ <i>K-Means</i> . (c) Saída da rede LEGION	62
Figura 3.29: Segmentação correta do VE. (a) Imagem obtida com a segmentação SOM/ <i>K-Means</i> . (b) Resultado da operação de abertura morfológica. (c) Saída da rede LEGION.....	63
Figura 4.1. Filosofia utilizada na concepção do EP original da rede ONNIS	67
Figura 4.2: Estrutura interna do EP da rede ONNIS	68
Figura 4.3: Exemplo da estrutura de conexões excitatórias.....	70
Figura 4.4: Exemplo da estrutura de conexões inibitórias.....	71
Figura 4.5: Estrutura utilizada para determinação dos EPs líderes	72
Figura 4.6: Exemplo da estrutura de conexões de inicialização global.....	73
Figura 4.7: Estrutura interna do EP modificado da rede ONNIS-GI	74
Figura 4.8: Segmentação de imagem em níveis de cinza utilizando a rede ONNIS-GI.....	77
Figura 4.9: Utilização da rede ONNIS-GI para segmentação de uma imagem considerando um atributo de textura.....	78
Figura 4.10: Imagem utilizada para exemplificar a segmentação utilizando filtros de Gabor e a rede ONNIS-GI.....	79
Figura 4.11: Filtragem da imagem da Figura 4.10 utilizando filtros de Gabor.....	79
Figura 4.12: Segmentação da imagem da Figura 4.10 utilizando a rede ONNIS-GI com determinação de pesos baseada em filtros de Gabor	80
Figura 4.13: Pesos determinados com base na semelhança entre as intensidades de pixels adjacentes.....	81
Figura 4.14: Influência do ruído na determinação dos pesos para diversos valores do parâmetro L_w	82
Figura 5.1: Imagem de teste (a). Bordas e valores dos pixels das regiões (b).....	87
Figura 5.2: Segmentos da imagem de teste.....	87
Figura 5.3: Imagem utilizada para avaliação da sensibilidade ao ruído	88
Figura 5.4: Imagem de teste corrompida por ruído gaussiano e $RSR_b=7,96$ dB	88

Figura 5.5: Tela do aplicativo desenvolvido para avaliação do comportamento da rede ONNIS-GI	89
Figura 5.6: Segmentação de uma imagem obtida com a rede ONNIS-GI.....	91
Figura 5.7: Segmentos obtidos com a aplicação da rede ONNIS-GI para segmentação da imagem da Figura 3.24.....	91
Figura 5.8: Pixels não segmentados pela rede	92
Figura 5.9: Segmentos obtidos da imagem de teste com $RSR_b = \infty$ e $L_w = 0$	94
Figura 5.10: Segmentos obtidos da imagem de teste com $RSR_b = 13,98$ dB e $L_w = 7$	94
Figura 5.11: Segmentos obtidos da imagem de teste com $RSR_b = 6,02$ dB e $L_w = 13$	94
Figura 5.12: Exemplo de segmentação utilizando um procedimento robusto para determinação dos pesos ($RSR_b = 13,98$ dB e $L_w = 4$)	96
Figura 5.13: Imagem com variação vertical linear de iluminação.....	97
Figura 5.14: Imagem com variação abrupta de iluminação	97
Figura 5.15: Segmentos obtidos a parti da imagem da Figura 5.14.....	98
Figura 5.16: Segmento obtido através do método de crescimento de regiões por agregação de pixels	99
Figura 6.1: Implementação da rede ONNIS	102
Figura 6.2: Estrutura interna do EP da rede ONNIS	103
Figura 6.3: Lógica interna do EP da rede ONNIS.....	103
Figura 6.4: Etapa para determinação do peso excitatório do EP.....	104
Figura 6.5: Contador interno do EP da rede ONNIS.....	104
Figura 6.6: Imagem de entrada da rede simulada no Max+plus II e implementada em FPGA.....	105
Figura 6.7: Resultado da simulação da rede ONNIS no Max+plus II.....	105
Figura 6.8: Contador interno do EP da rede ONNIS-GI	106
Figura 6.9: Estrutura interna do EP da rede ONNIS-GI.....	106
Figura 6.10: Implementação da rede da rede ONNIS-GI.....	107
Figura 6.11: Resultado da simulação da rede ONNIS-GI.....	108
Figura 6.12: Exemplo típico de segmento com máximo atraso para ativação completa.....	109
Figura APA.1: Aplicativo para simulação do neurônio oscilador de Terman-Wang .	122
Figura APA.2: Aplicativo para simulação de uma rede de quatro neurônios osciladores de Terman-Wang em anel.....	123
Figura APA.3: Aplicativo para simulação de uma rede LEGION com 20x20 osciladores, sem normalização dos pesos dinâmicos.....	124
Figura APA.4: Aplicativo para simulação de uma rede LEGION com 20x20 osciladores e normalização dos pesos dinâmicos	125
Figura APA.5: Aplicativo desenvolvido para simulação da rede LEGION em sua versão algorítmica.....	126
Figura APA.6: Aplicativo desenvolvido para simulação do algoritmo FPS	127
Figura APA.7: Aplicativo desenvolvido para o filtro de Gabor	128

LISTA DE TABELAS

Tabela 5.1: Erro obtido na segmentação da imagem de teste na presença de ruído.....	93
Tabela 5.2: Erro obtido com a utilização do procedimento robusto proposto para determinação dos pesos da rede	95

RESUMO

A presente tese apresenta a concepção de uma rede neural oscilatória e sua realização em arquitetura maciçamente paralela, a qual é adequada à implementação de *chips* de visão digitais para segmentação de imagens. A rede proposta, em sua versão final, foi denominada ONNIS-GI (*Oscillatory Neural Network for Image Segmentation with Global Inhibition*) e foi inspirada em uma rede denominada LEGION (*Locally Excitatory Globally Inhibitory Oscillator Network*), também de concepção recente.

Inicialmente, é apresentada uma introdução aos procedimentos de segmentação de imagens, cujo objetivo é o de situar e enfatizar a importância do tema abordado dentro de um contexto abrangente, o qual inclui aplicações de visão artificial em geral. Outro aspecto abordado diz respeito à utilização de redes neurais artificiais em segmentação de imagens, enfatizando as denominadas redes neurais oscilatórias, as quais têm apresentado resultados estimulantes nesta área. A implementação de *chips* de visão, integrando sensores de imagens e redes maciçamente paralelas de processadores, é também abordada no texto, ressaltando o objetivo prático da nova rede neural proposta.

No estudo da rede LEGION, são apresentados resultados de aplicações originais desenvolvidas em segmentação de imagens, nos quais é verificada sua propriedade de separação temporal dos segmentos. A versão contínua da rede, um arranjo paralelo de neurônios baseados em equações diferenciais, apresenta elevada complexidade computacional para implementação em *hardware* digital e muitos parâmetros, com procedimento de ajuste pouco prático. Por outro lado, sua arquitetura maciçamente paralela apresenta-se particularmente adequada à implementação de *chips* de visão analógicos com capacidade de segmentação de imagens.

Com base nos bons resultados obtidos nas aplicações desenvolvidas, é proposta uma nova rede neural, em duas versões, ONNIS e ONNIS-GI, as quais suplantam a rede LEGION em diversos aspectos relativos à implementação prática. A estrutura dos elementos de processamento das duas versões da rede, sua implementação em arquitetura maciçamente paralela e resultados de simulações e implementações em FPGA são apresentados, demonstrando a viabilidade da proposta.

Como resultado final, conclui-se que a rede ONNIS-GI apresenta maior apelo de ordem prática, sendo uma abordagem inovadora e promissora na solução de problemas de segmentação de imagens, possuindo capacidade para separar temporalmente os segmentos encontrados e facilitando a posterior identificação dos mesmos. Sob o ponto de vista prático, a nova rede pode ser utilizada para implementar *chips* de visão digitais com arquitetura maciçamente paralela, explorando a velocidade de tais topologias e apresentando também flexibilidade para implementação de procedimentos de segmentação de imagens mais sofisticados.

Palavras-Chave: segmentação de imagens, redes neurais oscilatórias, arquiteturas maciçamente paralelas, *chips* de visão.

ONNIS-GI: An Oscillatory Neural Network for Image Segmentation Implemented in Massively Parallel Architecture

ABSTRACT

The present thesis aims at proposing a new oscillatory neural network and at implementing using a massively parallel architecture, which is suitable to the implementation of digital vision chips for image segmentation. The final version of the proposed network called ONNIS-GI (Oscillatory Neural Network for Image Segmentation with Global Inhibition) was inspired on a network called LEGION (Locally Excitatory Globally Inhibitory Oscillator Network), also of recent conception.

Primarily, an introduction to the procedures of image segmentation is presented with the objective to situate and emphasize the importance of the subject in the general context of artificial vision. The thesis also covers the utilization of artificial neural networks in image segmentation, with emphasis on the oscillatory neural networks, which have presented encouraging results in this area. The implementation of vision chips, integrating image sensors and massively parallel processor networks, is also mentioned emphasizing the practical application of the new network proposed.

In the study of the LEGION network, results of original applications in image segmentation are presented, in which its segment temporal separation property is verified. The continuous version of this network, a two-dimensional array of neurons based on differential equations, presents both high computational complexity to implementation using digital hardware and many parameters with little practical setting procedures. On the other hand, its massively parallel topology is particularly adequate to the implementation of analog vision chips suitable to segment images.

Based on the good results obtained with the developed applications, a new oscillatory neural network is proposed in two versions, called ONNIS and ONNIS-GI, supplanting the LEGION network in several aspects related to practical implementation. The structure of the processing elements of the two versions of the proposed network, its implementation in massively parallel architecture and results concerning simulations and FPGA implementation are presented, demonstrating the viability of the proposal.

As a final result, it is concluded that the ONNIS-GI network is more attractive for practical applications, being an innovative approach, efficient in solving image segmentation problems, and possessing the ability to separate the image segments in time for easier identification. Under a practical point of view, the new network can be used for the implementation of vision chips with massively parallel architecture using digital technology, exploring the speed of such implementations and also presenting flexibility to implement more sophisticated segmentation procedures.

Keywords: image segmentation, oscillatory neural networks, massively parallel architectures, vision chips.

1 INTRODUÇÃO

A evolução tecnológica na área da computação, implicando a disponibilidade de máquinas com elevado poder de processamento a preços acessíveis, tem possibilitado a aplicação de procedimentos algorítmicos antes impossível ou extremamente onerosa. As técnicas de processamento digital de imagens (ou também denominadas de processamento de imagens digitais) encontram-se entre as grandes beneficiadas dentro deste quadro tecnológico. Desta forma, a aplicação prática de tais técnicas, bem como o estudo para o desenvolvimento de novas técnicas e de arquiteturas computacionais especiais para implementação eficiente das mesmas são, atualmente, campos férteis de trabalho, cujos resultados têm trazido enormes benefícios nas mais diversas áreas.

As técnicas de processamento digital de imagens podem ser utilizadas basicamente para adequar o conteúdo visual das imagens à interpretação humana ou para implementação de sistemas de visão artificial (ou visão computacional). A diversidade de aplicações e de ferramentas disponíveis faz do processamento digital de imagens uma área de conhecimento bastante ampla e complexa. Apesar de não estabelecido de forma clara, as aplicações de processamento digital de imagens podem ser representadas por um conjunto de classes, incluindo basicamente:

- Aquisição de imagens
- Representação de imagens
- Realce de imagens
- Restauração de imagens
- Análise de imagens
- Reconstrução de imagens e
- Compressão de imagens

Os procedimentos de aquisição de imagens envolvem diversas técnicas para obtenção de uma imagem digital (amostragem espacial, quantização, etc.) a partir de uma cena real. A representação de imagens está relacionada com a forma de caracterização da informação representada pelos pixels da imagem (intensidade de luz refletida, temperatura, frequência espacial, etc.). Já as técnicas utilizadas no realce de imagens têm como objetivo enfatizar determinadas características para visualização humana ou para análise por um sistema de visão artificial, enquanto que as técnicas de restauração de imagens têm como objetivo eliminar ou reduzir distorções e ruídos indesejáveis. Os procedimentos para análise de imagens são utilizados com o objetivo de obter medidas quantitativas de uma imagem, descrições da mesma ou para detectar a presença de determinados objetos. Em reconstrução de imagens, aplicam-se técnicas para obter imagens n-dimensionais a partir de diversas projeções em espaços de menor dimensão. Finalmente, as técnicas de compressão de imagens têm como objetivo a redução da taxa (bits/pixel) necessária para transmitir ou armazenar imagens.

As aplicações envolvendo análise de imagens em sistemas de visão artificial implicam, em geral, a extração de determinados objetos de uma cena para que os mesmos possam ser devidamente descritos e quantificados. Este é o caso, por exemplo, do controle de qualidade de peças produzidas em uma linha de montagem através da utilização de imagens, do reconhecimento automático de caracteres ou ainda de procedimentos de análise e diagnóstico de imagens médicas assistidos por computador.

A Figura 1.1 ilustra genericamente as etapas envolvidas em procedimentos de análise de imagens. A etapa de pré-processamento, logo após a aquisição, implica a restauração, o realce ou simplesmente a representação adequada dos dados da imagem. Nesta etapa, podem ser evidenciadas ou extraídas algumas características desejadas da imagem, como os contornos dos objetos por exemplo, ou pode-se transformar a mesma para algum outro domínio mais adequado à aplicação, como domínio frequência por exemplo. A imagem convenientemente pré-processada pode ser então segmentada através de alguma técnica adequada, com o objetivo de isolar regiões com características semelhantes (segmentos) relevantes dentro da cena. A posterior aplicação de técnicas de classificação adequadas possibilita relacionar rótulos de identificação a cada uma das regiões, utilizando determinadas características das mesmas, segundo uma base de conhecimento prévia. Finalmente, pode então ser elaborada uma descrição, incluindo medidas quantitativas ou não, ou uma interpretação da imagem adequada ao contexto da aplicação.

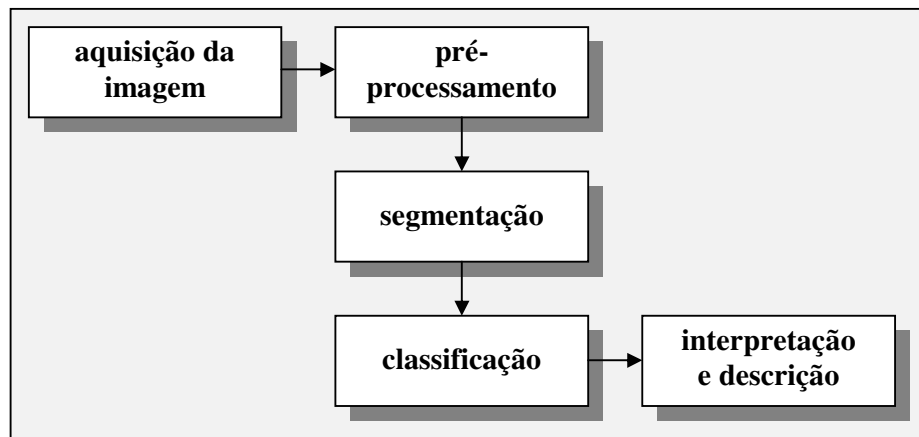


Figura 1.1: Etapas de um procedimento genérico para análise de imagens

A análise de imagens em sistemas de visão artificial envolve, portanto, o estudo de procedimentos de segmentação de imagens, os quais exercem papel fundamental no processo, visto que os diversos objetos que compõem uma cena devem ser separados adequadamente (segmentados) para sua posterior identificação e análise. A segmentação imperfeita de uma imagem conduz, via de regra, a descrições, medidas ou interpretações incorretas dos objetos ou da cena como um todo.

O cérebro humano tem notável facilidade em executar certas tarefas, tais como o reconhecimento de padrões diversos e a identificação individual de objetos em imagens complexas. Tal capacidade, entre outras razões, tem estimulado o estudo de aplicações e de novos modelos de redes neurais artificiais, constituindo-se em uma área atrativa para pesquisas, com apresentação de resultados estimulantes nas mais diversas áreas de aplicação. Recentemente, propostas alternativas de redes neurais artificiais, denominadas redes neurais oscilatórias, as quais possuem inspiração biológica no suposto mecanismo de segmentação executado pelo cérebro humano, têm sido aplicadas

em procedimentos de segmentação de imagens com resultados bastante favoráveis. O estudo destas redes neurais e de aplicações das mesmas, assim como o desenvolvimento de arquiteturas de *hardware* dedicadas a implementações práticas, também constituem um campo de trabalho promissor.

De uma forma geral, técnicas de processamento de imagens são fortes candidatas à aplicação de processamento paralelo. Sob o ponto de vista da implementação prática, a crescente demanda por sistemas de visão artificial que executem algoritmos complexos de forma suficientemente rápida tem justificado o desenvolvimento de circuitos integrados denominados *chips* de visão ou retinas de silício. Nestes *chips*, os fotodetectores são integrados em conjunto com uma rede maciçamente paralela de elementos de processamento (EPs) para execução de operações específicas sobre a imagem captada.

Tendo em vista sua natureza intrinsecamente paralela, as redes neurais oscilatórias apresentam-se como uma solução naturalmente adequada à implementação de arquiteturas de processamento para *chips* de visão dedicados à segmentação de imagens. Tal característica, vinculada à importância prática dos procedimentos de segmentação de imagens, justifica o estudo de arquiteturas maciçamente paralelas envolvendo a utilização de redes neurais oscilatórias.

O presente trabalho tem como objetivo apresentar uma proposta de uma rede neural oscilatória e sua implementação em arquitetura maciçamente paralela, a qual é adequada a aplicações de segmentação de imagens. A rede proposta tem inspiração no mecanismo de segmentação de imagens atribuído ao cérebro humano, possuindo como características atraentes uma topologia adequada à implementação de *chips* de visão em tecnologia digital e a capacidade de separar temporalmente os diversos segmentos encontrados na imagem de entrada.

Inicialmente, no Capítulo 2, uma revisão bibliográfica explora o contexto do trabalho, abordando basicamente procedimentos de segmentação de imagens, aplicação de redes neurais artificiais em segmentação de imagens e *chips* de visão com arquiteturas maciçamente paralelas para processamento de imagens.

No Capítulo 3, é estudada detalhadamente uma proposta recente de rede neural artificial denominada LEGION (*Locally Excitatory Globally Inhibitory Oscillator Network*), inspirada no funcionamento do cérebro humano, a qual tem apresentado resultados estimulantes em segmentação de imagens. A estrutura do neurônio oscilador da rede é apresentada, além de sua topologia e de outras características de funcionamento da rede. O Capítulo 3 também inclui resultados originais obtidos ao longo do desenvolvimento da presente tese com a aplicação desta rede em problemas práticos de segmentação de imagens.

O Capítulo 4 apresenta a proposta da nova rede neural oscilatória em duas versões, denominadas ONNIS (*Oscillatory Neural Network for Image Segmentation*) e ONNIS-GI (*Oscillatory Neural Network for Image Segmentation with Global Inhibition*) as quais são diferenciadas basicamente pela utilização ou não de um dispositivo de inicialização global, e suas implementações em arquiteturas maciçamente paralelas. Para cada uma das versões da rede é apresentado o respectivo elemento de processamento e as estruturas de conexões entre os mesmos. A capacidade da rede proposta para segmentar imagens com base em atributos diversos da mesma é abordada, incluindo a aplicação da rede em conjunto com filtros de Gabor para segmentar regiões com texturas distintas. A determinação dos pesos da rede proposta através de procedimentos robustos é também abordada, tendo em vista a possibilidade de diminuir a sensibilidade da segmentação em relação ao ruído presente na imagem de entrada. Finalmente, de forma análoga à rede LEGION, é também apresentada uma versão

algorítmica desenvolvida para rede ONNIS-GI, possibilitando a implementação da mesma em máquinas seriais.

No Capítulo 5, ilustrando a potencialidade desta nova proposta, são apresentados resultados obtidos com a aplicação da mesma em segmentação de imagens. Os resultados foram obtidos através de um algoritmo implementado em computador PC, o qual simula o funcionamento da rede em sua versão com inicialização global. Além de imagens artificiais e naturais, são também utilizadas imagens com ruído e variação de iluminação para verificar o comportamento da rede em relação a estes aspectos. Neste capítulo é também apresentada uma análise comparativa dos resultados obtidos com a rede ONNIS-GI e com o clássico método de segmentação de imagens denominado Crescimento de Regiões (*Region Growing*).

A implementação das redes ONNIS e ONNIS-GI utilizando dispositivos FPGA é abordada no Capítulo 6. São apresentados os resultados obtidos com protótipos de pequenas dimensões para as duas versões da rede proposta, confirmando a viabilidade de implementação das mesmas em *hardware* digital.

Já o Capítulo 7 apresenta a conclusão e considerações sobre o desdobramento futuro das pesquisas vinculadas ao tema de tese, incluindo uma previsão de artigos a serem elaborados a partir da proposta e dos resultados obtidos.

Finalmente, o Apêndice apresenta os programas aplicativos desenvolvidos para obtenção dos resultados apresentados, enquanto que o Anexo apresenta os artigos publicados durante o desenvolvimento do presente trabalho.

2 FUNDAMENTAÇÃO

Este capítulo apresenta uma revisão bibliográfica que explora o contexto do tema da presente tese, situando e enfatizando a importância do mesmo. São abordados basicamente procedimentos de segmentação de imagens, aplicações de redes neurais artificiais em tais procedimentos e *chips* de visão implementados em arquitetura maciçamente paralela.

Em segmentação de imagens são abordados os diversos procedimentos usualmente utilizados. A aplicação do filtro de Gabor para discriminação de características de textura, assim como de técnicas de detecção de bordas, são enfatizados tendo em vista sua possível utilização em conjunto com o novo procedimento de segmentação proposto. Com o objetivo de estabelecer um critério para comparar os diferentes procedimentos e configurações quanto à qualidade da segmentação das imagens durante o desenvolvimento do presente trabalho, são também relacionados métodos de avaliação encontrados na literatura.

Já na abordagem da utilização de redes neurais artificiais em segmentação de imagens, são relacionados diversos procedimentos encontrados na literatura, baseados em diferentes tipos de redes. Como inspiração do tema da presente tese, é especialmente enfatizada uma rede neural oscilatória denominada rede LEGION, de concepção recente, e sua aplicação em segmentação de imagens.

Finalmente, a abordagem de *chips* de visão com arquiteturas maciçamente paralelas em processamento de imagens, uma tendência atual, enfatiza o aspecto da implementação prática da proposta desenvolvida.

2.1 Segmentação de Imagens

A segmentação de uma imagem é a decomposição da mesma em diversas regiões independentes e homogêneas segundo determinada característica ou conjunto de características, também denominadas feições (*features*). A divisão de uma imagem em partes ou objetos constituintes (GONZALES; WOODS, 2000) ou em suas componentes (JAIN, 1989) são definições encontradas para o termo na literatura. Embora os seres humanos realizem tal procedimento com aparente facilidade na maioria dos casos, a segmentação de imagens apresenta-se como uma etapa fundamental e frequentemente complexa entre a aquisição da imagem e o correto entendimento da mesma através de um sistema de visão artificial.

Formalmente, se F representa o conjunto dos pixels (*picture elements*) de uma imagem, ele pode ser particionado, segundo um critério de homogeneidade aplicado a um grupo de características, em um conjunto de regiões (S_1, S_2, \dots, S_n) tais que as propriedades expressas por (2.1) sejam verificadas. As relações apresentadas em (2.1) indicam que todos os pixels da imagem devem pertencer a algum segmento e que um mesmo pixel não pode pertencer a dois segmentos diferentes (PAL; PAL, 1993).

$$\bigcup_{i=1}^n S_i = F \quad \text{e} \quad S_i \cap S_j = \emptyset, \quad i \neq j \quad (2.1)$$

Diversas técnicas para segmentação de imagens podem ser encontradas na literatura, não existindo entretanto um método que possa ser considerado bom para todos os tipos de imagens e tampouco todos os métodos são igualmente eficientes para um tipo particular de imagem (PAL; PAL, 1993). A seleção da técnica de segmentação a ser aplicada em uma imagem específica pode ser um problema difícil. Por exemplo, métodos que incluem características psico-visuais podem funcionar bem para imagens de intensidade de luz, mas podem ser inadequadas para imagens térmicas, nas quais cada pixel representa a temperatura em um ponto (PAL; PAL, 1993).

2.1.1 Procedimentos para Segmentação de Imagens

A literatura aponta uma grande diversidade de métodos para segmentar imagens. De uma forma geral, os diversos procedimentos empregados podem ser classificados em quatro grandes categorias (MOHAMED; AHMED; FARAG, 1998):

- Métodos considerados clássicos, baseados em limiarização de amplitude (*thresholding*), detecção de bordas ou em crescimento/contração de regiões;
- Métodos estatísticos, tais como o classificador de máxima verossimilhança;
- Métodos baseados em lógica *fuzzy*, como o algoritmo *Fuzzy C-Means* e
- Métodos baseados em redes neurais artificiais com topologias diversas.

A limiarização de amplitude é uma técnica de segmentação adequada quando os diversos objetos em uma imagem ou determinadas características dos mesmos são representados por diferentes faixas de amplitude ou intensidade, como o caso de imagens monocromáticas e de alguns tipos de imagens radiográficas, por exemplo. Nestes métodos, a seleção dos limiares é o ponto mais importante, sendo que uma maneira para determinação dos mesmos é a identificação de vales nos histogramas das imagens. Se forem conhecidos os modelos probabilísticos para cada uma das classes envolvidas, pode ser realizada a determinação de um limiar ou de um conjunto de limiares para minimização da probabilidade de erro de classificação (risco de Bayes). Estas técnicas são particularmente críticas quando a imagem apresenta-se corrompida por ruído ou por alguma outra degradação significativa (PAL; PAL, 1993).

As técnicas baseadas em bordas segmentam objetos através da determinação das fronteiras físicas dos mesmos. Em uma primeira etapa, são detectadas as bordas e os contornos obtidos são completados de forma a apresentarem-se de forma contínua. Após, são extraídas as regiões delimitadas pelos contornos encontrados. Estas técnicas podem apresentar problemas na ocorrência de contato entre objetos diferentes, na presença de sobreposição de objetos, ou ainda na descontinuidade dos contornos devido à presença de ruído na imagem (RUSS, 1995).

Na denominada técnica de crescimento de regiões, a imagem é originalmente dividida em pequenas regiões com amplitudes constantes ou características similares (no caso limite, cada pixel da imagem original isoladamente). Regiões semelhantes são aglutinadas sequencialmente até que as regiões adjacentes tornem-se suficientemente diferentes e não ocorram mais aglutinações. Na técnica de divisão e aglutinação (*split-and-merge*) (JAIN, 1989), os segmentos são identificados através de sucessivas divisões e aglutinações de regiões com características semelhantes, tomando-se como ponto de partida a imagem inteira. Se uma região não é considerada homogênea segundo algum

critério, a mesma é dividida em regiões menores. Estas novas regiões, se homogêneas, são aglutinadas com regiões vizinhas que possuem características semelhantes. O procedimento é realizado até que todas as regiões sejam consideradas homogêneas, não existindo mais condições para divisões ou aglutinações.

Nos métodos estatísticos são utilizadas técnicas baseadas no conhecimento de modelos estatísticos e seus parâmetros. A aplicação de técnicas como a da relaxação probabilística ou do conhecido classificador de máxima verossimilhança são exemplos de métodos de segmentação estatísticos (MOHAMED; AHMED; FARAG, 1998).

Um exemplo da utilização de técnicas baseadas em lógica *fuzzy* para segmentar imagens médicas é encontrado em (MOHAMED; AHMED; FARAG, 1998), onde um algoritmo do tipo *Fuzzy C-Means* modificado é utilizado para implementar um processo automático de segmentação de imagens de tomografia do cérebro humano.

Diversas técnicas de segmentação de imagens baseadas em redes neurais artificiais têm sido desenvolvidas, particularmente utilizando redes MLP, redes de Hopfield ou mapas de Kohonen, com exemplos típicos encontrados respectivamente em (DAHMER, 1998; GHOSH; PAL; PAL, 1992; PICCOLI, 1999). Tais algoritmos têm apresentado bons resultados mesmo na presença de ruído ou distorções na imagem a ser segmentada. Como desvantagem, alguns destes métodos apresentam a necessidade de treinamento, o que pode ser problemático devido ao tempo necessário, ao número de amostras previamente segmentadas disponíveis e ainda à elevada complexidade, em alguns casos, para aplicação em tempo real e implementação em *hardware* não dedicado.

2.1.2 Características Utilizadas em Segmentação de Imagens

Para que as regiões correspondentes aos diversos objetos contidos em uma imagem possam ser corretamente separadas, é necessário que os pixels que formam as mesmas possuam alguma característica ou conjunto de características (atributos ou feições) comuns que os diferencie dos pixels dos demais objetos. A escolha das características associadas aos pixels de uma imagem em procedimentos de segmentação é uma tarefa fundamental e que pode ser bastante complexa, dependendo da imagem a ser segmentada. A escolha de um conjunto de características inadequadas conduz, de uma forma geral, à segmentação incorreta dos objetos da cena, já que os pixels de um mesmo objeto podem não ser convenientemente distinguidos dos pixels de outros objetos.

A característica mais empregada em segmentação de imagens monocromáticas é a intensidade do pixel, a qual não implica normalmente procedimentos elaborados para sua determinação. Imagens monocromáticas possuem uma única intensidade associada a cada pixel, enquanto imagens coloridas possuem um conjunto de intensidades associadas a cada pixel, referentes às amplitudes das cores básicas empregadas. Já as chamadas imagens multiespectrais, tais como imagens de sensoriamento remoto, possuem diversas bandas de frequência, implicando, de forma semelhante às imagens coloridas, um vetor de intensidades associado a cada pixel da imagem.

A disposição espacial de níveis de intensidade dos pixels em uma imagem define uma característica visual denominada textura de um objeto ou de uma superfície. O termo textura geralmente refere-se à repetição de elementos básicos denominados *texels* (*texture elements*) (JAIN, 1989), incluindo também disposições aleatórias com determinadas propriedades estatísticas.

Os diversos problemas relacionados com a análise de textura em imagens podem ser classificados em três categorias (HARALICK, 1986):

- Dada uma determinada região, determinar a qual classe de textura a mesma pertence;

- Dada uma determinada região, determinar uma descrição ou um modelo de textura para a mesma e
- Dada uma imagem composta por diversas regiões com texturas distintas, determinar as fronteiras que separam as regiões.

A primeira categoria mencionada está relacionada com a área de reconhecimento de padrões, tratando-se, mais especificamente, de um problema de classificação. A segunda categoria está relacionada com a obtenção de modelos descritores, os quais podem ser utilizados para caracterizar texturas. Já a terceira categoria está relacionada propriamente com a segmentação de imagens com base nos atributos de textura de suas diversas regiões.

Os procedimentos utilizados para análise de textura em imagens podem ser classificados em dois grandes grupos: os procedimentos estatísticos e os procedimentos estruturais (HARALICK, 1986; JAIN, 1989). Os procedimentos estatísticos utilizam parâmetros para caracterizar as propriedades estocásticas da distribuição espacial dos níveis de intensidade dos pixels da imagem. Já os procedimentos estruturais analisam uma cena em termos de sua organização e das relações entre suas subestruturas, considerando a idéia básica de que um simples elemento de textura pode ser utilizado para formar padrões mais complexos segundo um determinado conjunto de regras (GONZALES; WOODS, 2000). A distinção entre procedimentos estatísticos e estruturais não é absolutamente clara tendo em vista que modelos estruturais de textura podem ser baseados em parâmetros estatísticos obtidos das texturas existentes na imagem.

Entre as técnicas de medida estatística de textura, podem ser citadas a autocorrelação, as transformações ortogonais, a co-ocorrência de tons de cinza, morfologia matemática, análise de gradiente, densidade de extremos relativos, campos de Markov aleatórios discretos e modelos de mosaicos aleatórios (HARALICK, 1986).

O coeficiente de autocorrelação é uma medida do grau de dependência linear entre um pixel e outro pixel deslocado no espaço. Desta forma, a periodicidade espacial dos níveis dos pixels de uma imagem pode ser avaliada em termos da função autocorrelação da mesma. Uma função autocorrelação com picos próximos pode caracterizar uma textura fina, indicando a repetição de características espaciais semelhantes em intervalos pequenos. Por outro lado, uma função autocorrelação com picos distantes pode caracterizar uma textura grossa, indicando a repetição de características espaciais semelhantes em intervalos maiores.

As transformações ortogonais incluem, entre outras, a transformada de Fourier, a qual relaciona a função autocorrelação com o espectro de potência do sinal. Assim como a função autocorrelação, o espectro de potência de uma imagem pode ser utilizado para avaliar a característica de periodicidade e direcionalidade de uma determinada textura, podendo distinguir a mesma de outras texturas (GONZALES; WOODS, 2000).

A determinação de momentos com base no histograma de intensidade ou níveis de cinza de uma imagem apresenta a limitação de não fornecer qualquer informação sobre a posição dos pixels. A utilização de matrizes de co-ocorrências espaciais de níveis de cinza é uma alternativa para tal limitação (GONZALES; WOODS, 2000). A co-ocorrência $P(m,n)$ de níveis de cinza para uma imagem I é definida como o número de pares de pixels que possuem respectivamente as intensidades m e n e que estão separados por uma distância preestabelecida (HARALICK, 1986). A matriz de co-ocorrência de níveis de cinza, na qual cada elemento representa uma estimativa da probabilidade conjunta de ocorrência de dois níveis segundo o critério de posição relativa estabelecido, é obtida através da divisão de cada co-ocorrência pelo número

total de pares de pixels considerados segundo o referido critério (GONZALES; WOODS, 2000). Estatísticas de probabilidade de co-ocorrências normalmente utilizadas incluem energia, entropia, contraste, correlação e homogeneidade (HARALICK, 1986), e também máxima probabilidade, momentos de diferença entre elementos e momentos da diferença inversa entre elementos (GONZALES; WOODS, 2000). Em (HARALICK, 1986) são relacionadas algumas referências que apontam os resultados obtidos com a utilização de estatísticas de co-ocorrência de tons de cinza como sendo superiores aos resultados obtidos com o espectro de potência na discriminação de texturas.

A utilização da morfologia matemática em análise de texturas requer a definição de um elemento estruturante como uma linha, um quadrado ou um disco, por exemplo, e a geração de imagens resultantes de operações de erosão e dilatação binárias (GONZALES; WOODS, 2000) da imagem original com o referido elemento estruturante. Características de textura, como a granulidade, por exemplo, podem ser então obtidas das imagens resultantes do processo (HARALICK, 1986).

A caracterização de texturas em termos da quantidade de bordas por unidade de área pode ser realizada através de métodos de análise do gradiente, incluindo operadores diversos, tais como os de Roberts, Prewitt, Sobel e Laplaciano, por exemplo (JAIN, 1989; HARALICK, 1986). O procedimento de identificação e contagem do número de extremos ou picos por unidade de área, denominado densidade de extremos relativos, pode também ser utilizado para caracterizar texturas (HARALICK, 1986).

Outro procedimento para identificação de texturas são os modelos de campo aleatório de Markov. Nestes modelos, uma determinada textura é relacionada com um conjunto de coeficientes específico adequado a um modelo ARMA (*Auto Regressive-Moving Average*) generalizado (HARALICK, 1986).

Modelos de mosaicos aleatórios são também procedimentos utilizados para caracterizar texturas, os quais são baseados em aproximações estatísticas e estruturais das diversas texturas (HARALICK, 1986).

2.1.3 Segmentação Utilizando Filtros de Gabor

O conceito de filtros de Gabor bidimensionais foi originalmente proposto por John G. Daugman em 1980 como uma ferramenta para compreensão das propriedades perceptivas dos neurônios do córtex visual e também como uma ferramenta para aplicações práticas em problemas de processamento de imagens (DAUGMAN, 1993).

A utilização de filtros de Gabor para caracterização de texturas em segmentação de imagens tem se tornado popular tendo em vista sua apregoada propriedade de modelar certos aspectos do sistema visual humano e de possuir algumas características computacionais atrativas (CAMPBELL; THOMAS; TROSCIANKO, 1996-a). Por exemplo, em (DAUGMAN, 1993) é encontrada uma aplicação do filtro de Gabor para caracterização da íris humana em um sistema de identificação de pessoas. Em (CAMPBELL; THOMAS; TROSCIANKO, 1996-a; CAMPBELL; THOMAS; TROSCIANKO, 1996-b; CAMPBELL; THOMAS; TROSCIANKO, 1997), os filtros de Gabor são associados a mapas de Kohonen para segmentação de imagens naturais. Já em (ANGELO, 2000), filtros de Gabor são utilizados para gerar “bandas de textura” com o objetivo de classificar imagens de sensoriamento remoto utilizando máxima verossimilhança gaussiana.

A equação (2.2) representa a resposta ao impulso da família de filtros de Gabor bidimensionais para o domínio espacial definido pelas dimensões x e y e para o domínio da frequência definido pelas dimensões u e v . Os parâmetros α e β determinam o tamanho efetivo da janela espacial do filtro, implicando diretamente na largura de banda do mesmo no domínio frequência. Já os parâmetros u_0 e v_0 implicam uma componente

harmônica com frequência ω e direção θ definidas respectivamente por (2.3) e (2.4). Uma maior generalização pode ser obtida com a inclusão em (2.2) de um termo relativo a correlação entre x e y , alterando a orientação do envelope gaussiano (DAUGMAN, 1993). Finalmente, a equação (2.5) representa a resposta em frequência do filtro de Gabor obtida através da aplicação da transformada de Fourier sobre a resposta ao impulso do mesmo (2.2).

$$g(x, y) = e^{-\pi(x^2 / \alpha^2 + y^2 / \beta^2)} e^{-j2\pi(xu_0 + yv_0)} \quad (2.2)$$

$$\omega_0 = \sqrt{u_0^2 + v_0^2} \quad (2.3)$$

$$\theta = \tan^{-1}(v_0 / u_0) \quad (2.4)$$

$$G(u, v) = e^{-\pi((u-u_0)^2 \alpha^2 + (v-v_0)^2 \beta^2)} \quad (2.5)$$

onde:

- $g(x,y)$ é a resposta ao impulso do filtro de Gabor nas coordenadas x e y ;
- α e β são os parâmetros do envelope gaussiano bidimensional;
- u_0 e v_0 são componentes da frequência de oscilação da resposta do filtro;
- ω_0 é a frequência de oscilação da resposta do filtro;
- θ é a direção da componente de oscilação e
- $G(u,v)$ é a resposta do filtro para a frequência de componentes u e v .

Na Figura 2.1 é apresentado um exemplo que ilustra as componentes real e imaginária da resposta ao impulso de um filtro de Gabor no domínio espacial, podendo ser observado o efeito de modulação senoidal do envelope gaussiano. Já na Figura 2.2, é apresentada a resposta em frequência do mesmo filtro, caracterizando uma resposta do tipo passa-banda. Foram utilizados os parâmetros $\alpha^2=50$ e $\beta^2=50$, implicando comportamento semelhante da gaussiana nas direções x e y . Os parâmetros de frequência são $u_0=0,05$ ciclos/pixel e $v_0=0$ ciclos/pixel, implicando uma oscilação com período de 20 pixels na direção x . A respectiva frequência de oscilação é claramente observada na resposta em frequência, correspondendo às coordenadas do valor de pico da função.

Os filtros de Gabor bidimensionais podem formar uma base de *wavelets* com os requisitos de ortogonalidade e suporte compacto relaxados, utilizando parametrização apropriada para dilatação, rotação e translação. Através da referida parametrização, um conjunto de filtros de Gabor pode ser projetado para representar informações da textura de uma imagem em todo o domínio frequência, detectando, individualmente, padrões senoidais em direções diversas (DAUGMAN, 1993).

A filtragem de uma imagem com o objetivo de detectar a presença de uma determinada frequência é realizada através da convolução da mesma com a resposta ao impulso do filtro de Gabor sintonizado na respectiva frequência. Tendo em vista que a imagem obtida após a filtragem tem valores de intensidade complexos, é usual considerar-se o módulo da mesma para avaliação da presença da respectiva frequência na imagem analisada.

Seletividades elevadas implicam grandes dimensões espaciais para o filtro de Gabor, conduzindo a elevada complexidade computacional na realização da filtragem. Desta

forma, deve-se chegar a um compromisso prático entre seletividade, para discriminar adequadamente componentes em frequência distintas, e complexidade computacional. A limitação das dimensões do filtro pode ser realizada em função da característica do envelope gaussiano, desprezando-se os termos da resposta ao impulso que possuem valores muito reduzidos. Em (ANGELO, 2000) é sugerida a limitação da resposta ao impulso do filtro de Gabor ao intervalo entre -3α e $+3\alpha$ para a dimensão x e entre -3β e $+3\beta$ para a dimensão y .

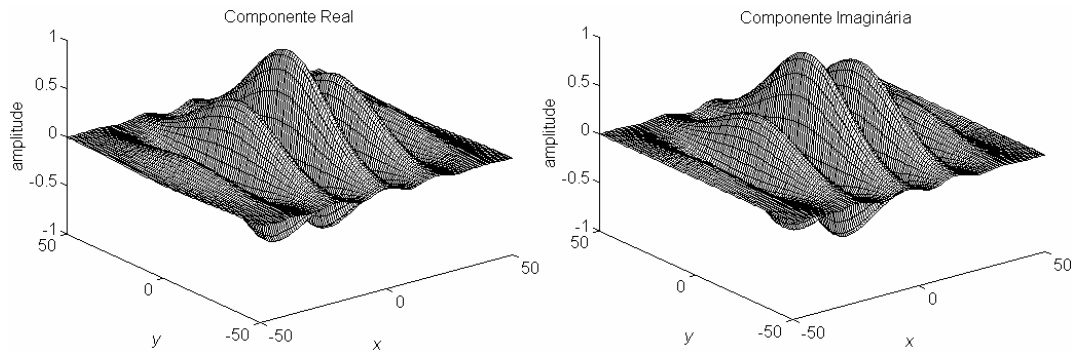


Figura 2.1: Componentes real e imaginária da resposta ao impulso de um filtro de Gabor

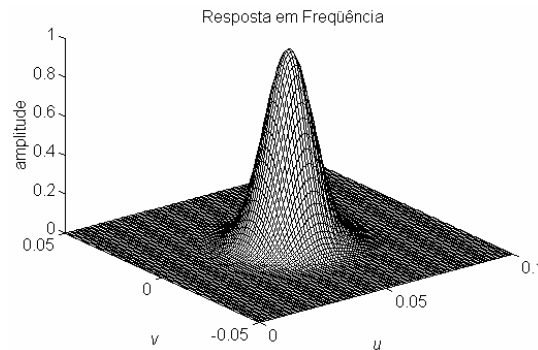


Figura 2.2: Resposta em frequência do filtro de Gabor representado pela resposta ao impulso da Figura 2.1

A Figura 2.3 apresenta um exemplo de utilização de um filtro de Gabor para detectar texturas com frequência espacial de $u_0=0,1$ períodos/pixel e $v_0=0,1$ períodos/pixel em uma imagem formada por regiões com variações senoidais de intensidade com frequências distintas. As imagens apresentadas na Figura 2.3(b) e na Figura 2.3(c) representam o módulo da imagem da Figura 2.3(a) filtrada utilizando, respectivamente, $\alpha^2=50$ e $\beta^2=50$ no primeiro caso e $\alpha^2=400$ e $\beta^2=400$ no segundo caso. O exemplo permite verificar que a seletividade em frequência do filtro de Gabor está relacionada com as dimensões da janela espacial utilizada.

A Figura 2.4 apresenta um exemplo de aplicação de um filtro de Gabor para detectar uma componente de frequência específica em uma imagem com texturas formadas pela adição de componentes com frequências e orientações diversas. No caso, a imagem original, representada na Figura 2.4(b), é formada pela soma de diversas imagens com texturas senoidais (Figura 2.4(c), (d), (e) e (f)). O módulo da imagem obtida com a aplicação de um filtro de Gabor utilizando $\alpha^2=400$, $\beta^2=400$, $u_0=0$ e $v_0=0,1$

períodos/pixel e dimensão 121x121 pixels é apresentado na Figura 2.4(a), observando-se a correta detecção da componente desejada.

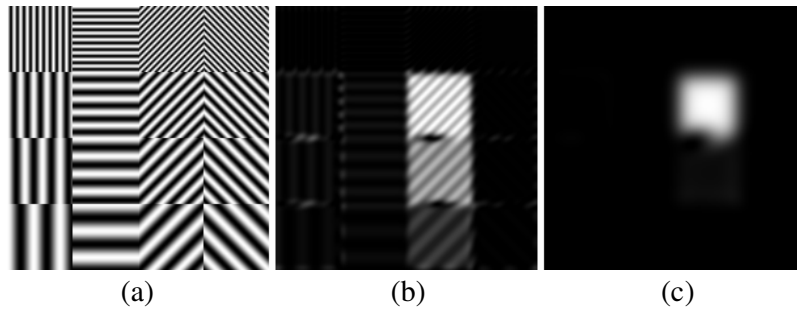


Figura 2.3: Utilização de um filtro de Gabor para detectar texturas com frequência espacial de $u_0=0,1$ e $v_0=0,1$ períodos/pixel. (a) Imagem original. (b) Módulo da imagem filtrada com $\alpha^2=50$ e $\beta^2=50$. (c) Módulo da imagem filtrada com $\alpha^2=400$ e $\beta^2=400$

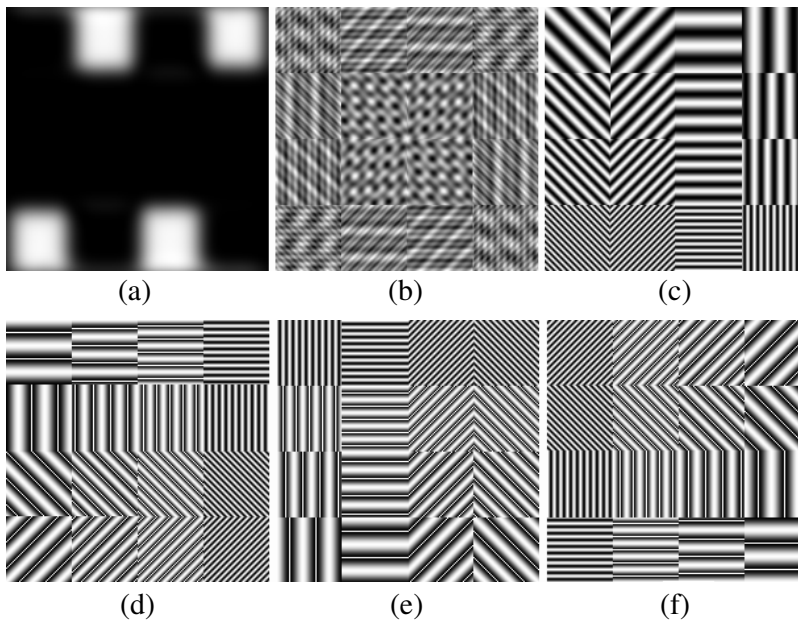


Figura 2.4: Utilização de um filtro de Gabor para detectar componentes específicas em textura complexas. (a) Módulo da imagem filtrada com $\alpha^2=400$, $\beta^2=400$, $u_0=0$ e $v_0=0,1$ períodos/pixel. (b) Soma das imagens (c), (d), (e) e (f)

Um conjunto de filtros de Gabor pode, desta forma, ser utilizado para caracterizar regiões de imagens como parte de procedimentos de segmentação, gerando vetores de características associados a cada pixel. A escolha adequada dos parâmetros para caracterizar determinadas texturas pode implicar uma grande quantidade de filtros e também filtros com grande número de coeficientes, resultando em elevada complexidade computacional e restringindo a aplicação prática do procedimento.

2.1.4 Segmentação Baseada em Detecção de Bordas

Os algoritmos utilizados para segmentação de imagens podem ser baseados na similaridade ou na descontinuidade das características vinculadas aos pixels da imagem.

Os algoritmos baseados em descontinuidades utilizam, via de regra, procedimentos de detecção de bordas. Uma borda é o limite entre duas regiões com propriedades relativamente distintas de níveis de cinza (GONZALES; WOODS, 2000) ou, de uma forma mais geral, de características diversas dos pixels.

Os operadores utilizados para detecção de bordas podem ser classificados em operadores gradiente, que medem o gradiente em duas direções ortogonais, e operadores bússola (*compass operators*), que medem o gradiente em direções selecionadas (JAIN, 1989). No processamento digital de imagens, as máscaras utilizadas para detecção de bordas determinam aproximações em termos de diferenças finitas dos gradientes ortogonais ou do gradiente direcional.

Os operadores gradiente podem ser representados por um par de máscaras, utilizadas para estimar o gradiente da imagem em duas direções ortogonais, $g_1(m,n)$ e $g_2(m,n)$. Por simplicidade, a magnitude do gradiente é normalmente determinada pela forma apresentada em (2.6). Um pixel localizado nas coordenadas (m,n) pertencente à uma borda quando $g(m,n)$ for superior a um limiar preestabelecido, na prática cerca de 90 a 95% do valor máximo do gradiente para toda a imagem (JAIN, 1989; DAS; ANAND, 1995).

$$g(m,n) = |g_1(m,n)| + |g_2(m,n)| \quad (2.6)$$

Alguns operadores do tipo gradiente normalmente encontrados são os de Prewitt, Sobel, Roberts e o Isotrópico, apresentando os dois primeiros melhores resultados (JAIN, 1989). Já os operadores bússola, tais como o operador de Kirsh, utilizam máscaras para estimação do gradiente em várias direções, sendo o gradiente na posição (m,n) , definido como o máximo entre os módulos dos gradientes nas diversas direções (2.7) (JAIN, 1989).

$$g(m,n) = \max_k (|g_k(m,n)|) \quad (2.7)$$

Os métodos mencionados para estimação do gradiente não funcionam bem quando as bordas não se apresentam de forma bem definida. Neste caso, o uso de operadores como o Laplaciano é mais adequado. Por outro lado, a maior sensibilidade ao ruído associada à produção de bordas duplas e à inabilidade de detectar a direção da borda tornam o operador Laplaciano pouco satisfatório para detectar bordas (JAIN, 1989).

Uma das maiores desvantagens de operadores como Prewitt e Sobel é o fato de que suas máscaras são fixas, independente do nível de ruído da imagem, resultando em baixa performance nestas condições (DAS; ANAND, 1995). Uma alternativa bastante efetiva para projetar máscaras de detecção de bordas, a qual leva em conta a presença de ruído, é o método do Gradiente Estocástico. Neste método, são realizadas estimativas com base em um filtro FIR, obtendo-se um desempenho na detecção de bordas maior do que os métodos anteriormente mencionados (JAIN, 1989; DAS; ANAND, 1995).

Em (SHIN; GOLDFOF; BOWYER, 2001) é realizada uma comparação entre oito diferentes procedimentos para detecção de bordas (Anisotrópico, Begholm, Canny, Heitger, Rothwell, Sarkar, Sobel e SUSAN), tendo-se concluído que o detector de Canny apresentou melhor performance, de forma geral, para a aplicação em questão. Na mesma publicação é também afirmado que não existe um método formal genericamente aceito para avaliar resultados de detecção de bordas e que nenhum método de detecção tem sido genericamente adotado pela comunidade de pesquisadores. Por outro lado, contraditoriamente, afirma-se que o detector de Canny é largamente considerado como

um método padrão para detecção de bordas, o que também é afirmado em (HEATH et al., 1997).

Já em (SHARIFI; FATHY; MARYAM, 2002), é apresentado um estudo classificando e comparando sete algoritmos para detecção de bordas: ISEF, Canny, Marr-Hildreth, Sobel, Kirsh, Lapla1 e Lapla2, os quais são os mais comumente utilizados. Conclui-se que, sob condições de ruído, os métodos ISEF e Canny apresentaram melhor performance e afirma-se que nenhum dos métodos propostos é totalmente satisfatório em aplicações reais. Nesta mesma referência, os algoritmos são classificados em cinco classes:

- Detectores de borda por gradiente
- Detectores de cruzamento por zero
- Laplaciano de Gaussiana (LoG)
- Detectores de borda gaussianos
- Detectores de bordas coloridos

Também classificados como operadores clássicos, os detectores de borda por gradiente apresentam simplicidade e sensibilidade ao ruído (ex.: Sobel, Prewitt e Kirsh). Já os detectores de cruzamento por zero apresentam sensibilidade ao ruído e podem duplicar algumas bordas (ex.: Laplaciano). O detector LoG apresenta funcionamento inadequado nas quinas e em regiões homogêneas, enquanto que os detectores de bordas gaussianos possibilitam boa detecção em condições ruidosas e possuem computação complexa (ex.: Canny e ISEF). Os detectores de bordas coloridos incluem métodos de fusão, métodos de gradiente multidimensional e métodos vetoriais, sendo mais eficientes para o reconhecimento de objetos, além de possuírem computação complexa (SHARIFI; FATHY; MARYAM, 2002).

Em (VALVERDE, et al., 2001), é proposto um critério para avaliação de desempenho de procedimentos de detecção de bordas em imagens ruidosas. Foram avaliados 14 métodos: Canny, Diferença, Diferença Separada, Frei-Chen, Prewitt, Roberts, Sobel, Laplaciano Vizinhança 4, Laplaciano Vizinhança 8, Laplaciano Vizinhança 8 Separável, Base, Kirsh, Qiu e Schalkoff. A melhor performance média, utilizando a medida proposta, foi alcançada pelo método de Canny.

Em (HEATH et al., 1997) é apresentada uma metodologia baseada em análise visual para medir a performance relativa de algoritmos de detecção de bordas. São relacionados 21 algoritmos publicados em jornais conceituados na área. Destes, foram avaliados os métodos de Canny, Nalwa-Binford, Inverson-Zucker, Bergholm e Rothwell, por serem considerados representativos do estado da arte. Os resultados novamente indicam que o algoritmo de Canny obtém a melhor performance, desde que seus parâmetros sejam convenientemente adaptados para cada imagem.

2.1.5 Influência do Ruído na Segmentação de Imagens

Um dos problemas que dificultam significativamente a segmentação de imagens, o qual é frequentemente encontrado em aplicações práticas, é a degradação do sinal a ser segmentado por ruído ou distorções de naturezas diversas. De maneira geral, resultados obtidos com diferentes procedimentos na segmentação de imagens corrompidas por ruído comprovam ser necessária a utilização de alguma estratégia de filtragem sobre a imagem original. Tal filtragem tem como objetivo enfatizar a homogeneidade de características dentro das regiões de interesse (segmentos), ao mesmo tempo em que são acentuadas ou preservadas as diferenças de características entre regiões diversas.

Para atenuar o problema causado pelo ruído em aplicações de segmentação de imagens, são amplamente utilizadas técnicas de suavização em uma etapa anterior ou incorporados ao procedimento de segmentação, incluindo métodos iterativos de filtragem não linear. De uma forma geral, métodos iterativos de suavização de imagens estão sujeitos ao denominado “problema da terminação” (CHEN; WANG; LIU,), o qual se refere à obtenção de uma imagem totalmente homogênea, e portanto inútil, após um determinado número de aplicações do método.

A escolha de uma técnica de filtragem como apoio a procedimentos de segmentação de imagens deve ater-se àquelas que simultaneamente reduzam o ruído ou distorção e preservem características importantes da imagem, tais como as bordas dos objetos, por exemplo. Um exemplo de tais técnicas é o filtro de Lee (JU; MOLONEY, 1997), o qual tem a habilidade de atenuar ruído multiplicativo do tipo *speckle*, mantendo as bordas da imagem após várias iterações. O filtro FPS (*Feature Preservation Smoothing*) (CHEN; WANG; LIU, 2000), é outro exemplo de filtro que homogeneiza a imagem preservando as bordas, o qual apresenta também imunidade ao denominado “problema da terminação”. Já na segmentação de imagens baseada nas bordas dos objetos, a utilização de procedimentos de detecção de bordas robustos faz-se normalmente necessária.

2.1.6 Avaliação da Qualidade da Segmentação

Um aspecto importante em aplicações de segmentação de imagens diz respeito à avaliação da qualidade dos resultados obtidos. Enquanto o desenvolvimento de procedimentos para segmentação de imagens tem atraído significativa atenção, o mesmo não tem ocorrido com métodos de avaliação de qualidade (ZHANG, 1996). Alguns poucos procedimentos têm sido estabelecidos para este propósito, incluindo o cálculo da diferença pixel a pixel entre a imagem ou um objeto segmentado e a mesma imagem ou objeto segmentado manualmente por um especialista.

Em (PICCOLI, 1999), as dispersões *intra-cluster* e *inter-cluster* são utilizadas como medidas quantitativas da qualidade das segmentações obtidas, utilizando, contudo, avaliações visuais subjetivas de um especialista humano (um médico no caso) como principal indicador.

Uma classificação para os métodos de avaliação de algoritmos de segmentação é apresentada em (ZHANG, 1996), envolvendo em três grupos: os métodos analíticos, os de qualidade (*goodness*) empírica e os de discrepância empírica. Os métodos analíticos trabalham com os princípios e propriedades de cada algoritmo de segmentação. Já os métodos empíricos julgam indiretamente os algoritmos de segmentação aplicando imagens de teste e medindo a qualidade dos resultados. Nos métodos de qualidade empírica são avaliados parâmetros tais como a uniformidade *intra-região*, o contraste *inter-regiões* ou a forma das regiões, enquanto que na discrepância empírica são avaliados o número de pixels incorretamente segmentados, a posição dos pixels incorretamente segmentados, a discrepância baseada no número de objetos da imagem e a discrepância baseada em características dos objetos segmentados. Ainda em (ZHANG, 1996), cinco métodos de avaliação de segmentação, considerados de aplicação geral, foram avaliados: qualidade baseada em uniformidade de níveis de cinza, qualidade baseada em contraste, discrepância baseada na probabilidade de erro, discrepância baseada em distância normalizada e discrepância baseada em diferenças absolutas de áreas, tendo o último apresentado melhores resultados. Os métodos empíricos baseados em discrepância são considerados mais adequados que os analíticos para avaliação de performance em segmentação de imagens (ZHANG, 1996).

Em (YASNOFF; MUI; BACUS, 1977), duas medidas de erro, o percentual de área incorretamente classificada e uma distância de erro entre pixels foram definidas e

avaliadas em termos de sua correlação com a observação humana. Concluiu-se que, para comparação da segmentação da mesma imagem por vários algoritmos e de diferentes imagens pelo mesmo algoritmo, as medidas de erro propostas são coerentes com a percepção humana.

Uma metodologia para avaliar algoritmos de segmentação de imagens médicas que utiliza como referência as fronteiras de regiões delimitadas por vários especialistas é proposta em (CHALANA; KIM, 1996). Foram consideradas as variabilidades interespecialista e intra-especialista na estimativa de curvas médias para representar as fronteiras dos objetos corretamente segmentados (referência denominada *pseudo ground truth*¹). Para medir as distâncias entre curvas foram utilizadas a distância média entre pontos e a distância de Hausdorff. Esta última, considera a máxima distância encontrada entre pontos das curvas que representam as fronteiras das regiões segmentadas e pode não representar adequadamente os resultados de avaliações subjetivas.

Em (HUANG; DOM, 1995) são propostos dois conjuntos de medidas para avaliação quantitativa de resultados de segmentação de imagens. O primeiro é projetado para situações onde o *ground truth* é disponível e o segundo para quando o *ground truth* não é disponível. No primeiro caso, é utilizada a discrepância entre parâmetros tais como os coeficientes de uma função polinomial aproximando as curvas ou coeficientes de Fourier descrevendo as formas das mesmas. Também podem ser utilizadas distâncias ponderadas ou não entre as curvas. Outra possibilidade é utilizar a distância normalizada de Hamming para avaliar a precisão em termos das localizações e posições das áreas segmentadas. Quando o *ground truth* não é disponível, utiliza-se como critério geral a maximização da homogeneidade dentro de regiões segmentadas e da heterogeneidade entre diferentes regiões. Resultados experimentais demonstraram que as medidas obtidas com base no conhecimento do *ground truth* concordam com a avaliação visual subjetiva.

Outras medidas objetivas podem ser aplicadas, tais como a correlação e a uniformidade por exemplo, porém o julgamento de um observador humano é tido como o melhor meio de avaliar a qualidade de um processo de segmentação de imagens (PAL; PAL, 1993).

2.2 Redes Neurais em Segmentação de Imagens

A justificativa para a grande popularidade dos métodos baseados na utilização de redes neurais artificiais em processamento de imagens está basicamente vinculada a algumas características interessantes apresentadas por algumas topologias destas redes, porém não necessariamente presentes em todas elas:

- Capacidade de generalização, podendo-se obter resultados adequados para entradas não consideradas no projeto da rede;
- Existência de procedimentos de projeto (treinamento) baseados na apresentação de entradas e verificação de resultados, os quais minimizam a necessidade de criação ou utilização de modelos matemáticos elaborados;
- Robustez em relação a falhas operacionais nos elementos processadores (neurônios) da rede se a mesma for projetada com suficiente redundância;

¹ Em cartografia e sensoriamento remoto, o termo *ground truth* é traduzido como “verdade terrestre” ou “realidade do terreno” (WEBSTER’S, 2004) e representa uma referência para calibração de um instrumento ou de um processo. Tendo em vista a especificidade destas traduções, o termo original em inglês é mantido ao longo deste trabalho.

- Alto grau de paralelismo, implicando a possibilidade de implementações mais rápidas quando utilizado um *hardware* adequado;
- Bons resultados práticos, mesmo com a utilização de imagens distorcidas pela introdução de ruído ou por algum outro fenômeno não desejado.

2.2.1 Técnicas de Segmentação de Imagens Baseadas em Redes Neurais

A bibliografia aponta a existência de diversas aplicações de redes neurais em segmentação de imagens, principalmente direcionadas a imagens monocromáticas. Tais aplicações valem-se, em sua grande maioria, de métodos que podem ser considerados clássicos no campo das redes neurais: os mapas de Kohonen (SOFM - *Self-Organizing Feature Map*) (HAYKIN, 1999), as redes de Hopfield (HAYKIN, 1999) e as redes MLP (*MultiLayer Perceptron*) associadas ao algoritmo de treinamento *Backpropagation* (HAYKIN, 1999; PANDYA; MACY, 1996). Tais algoritmos têm demonstrado bom desempenho mesmo em imagens ruidosas e, é claro, apresentam diversas vantagens inerentes às respectivas topologias de redes neurais.

As redes MLP são provavelmente a topologia mais popular de redes neurais artificiais, encontrando-se diversas aplicações práticas das mesmas. Em (BLANZ; GISH, 1990), por exemplo, é utilizada uma rede do tipo MLP para segmentação de imagens, na qual o número de neurônios na camada de entrada depende do número de características (feições) de cada pixel e o número de neurônios na camada de saída é igual ao número de classes pré-definidas. Já em (BABAGUCHI et al., 1990), uma rede treinada com MLP é utilizada para segmentar imagens utilizando limiarização. O histograma da imagem é apresentado à entrada da rede, a qual fornece como saída o limiar (*threshold*) a ser aplicado na segmentação da imagem. O mesmo tipo de rede é utilizada em (DAHMER, 1998), onde a segmentação é realizada com base em atributos de textura da imagem. Em (STEMM, 1994), uma rede MLP é utilizada para segmentar imagens em um sistema de visão para deslocamento de um robô autônomo.

A utilização de redes MLP apresenta como desvantagem a necessidade de conhecimento antecipado, na fase de treinamento, de um número significativo de imagens adequadamente segmentadas por algum outro processo reconhecidamente eficiente ou por um especialista humano (aprendizado supervisionado).

Outro algoritmo robusto para extração de objetos de imagens contaminadas por ruído, apresentado em (GHOSH; PAL; PAL, 1991-b; GHOSH; PAL; PAL, 1992), utiliza redes de Hopfield. A função energia da rede é construída de maneira que em cada estado estável da mesma é obtido um segmento da imagem. Em (POLI; VALLI, 1995) são apresentados resultados referentes à utilização de redes de Hopfield na segmentação de imagens em 2 e 3 dimensões com aplicações em imagens médicas obtidas através de raios-x e tomografia computadorizada.

A utilização dos mapas auto-organizáveis propostos por Kohonen para segmentar imagens ecocardiográficas fetais é apresentada em (PICCOLI, 1999), tendo sido obtidos resultados considerados bons por médicos especialistas. Em (STOCKER et al., 1996) é apresentada uma comparação entre mapas auto-organizáveis e redes MLP na segmentação de imagens de ressonância magnética, sendo que o procedimento baseado em mapas auto-organizáveis apresentou melhores resultados. Já em (AHMED; FARAG, 1997), são comparados os resultados da aplicação de mapas de Kohonen e de redes de Hopfield na segmentação de imagens de ressonância magnética, concluindo-se também que as redes de Kohonen apresentaram melhores resultados.

Em (McCAULEY; THANE; WHITTAKER, 1994) é utilizada uma rede lógica adaptativa (ALN - *Adaptive Logic Network*) para identificar regiões de gordura em

imagens de ultra-sonografia. As ALNs são redes neurais construídas utilizando árvores binárias cujos nós representam portas lógicas.

Uma topologia de rede neural que detecta bordas em potencial com diferentes orientações é apresentada em (CORTES; HERTZ, 1989). As imagens podem ser então segmentadas com base nas bordas dos objetos identificadas.

Em (CHEN; TSAO; LIN, 1991), é proposta uma rede neural tipo CSNN (*Constraint Satisfaction Neural Network*), a qual relaciona um conjunto de objetos, um conjunto de rótulos, uma coleção de relações de restrição e uma restrição topológica descrevendo as relações de vizinhança entre vários objetos. O referido método foi aplicado com sucesso em segmentação de imagens de tomografia computadorizada e de ressonância magnética. A robustez do algoritmo não foi investigada para imagens excessivamente ruidosas e as redes obtidas apresentam número elevado de neurônios, mesmo para imagens de tamanho moderado.

Em algumas referências, tais como em (GHOSH; PAL; PAL, 1991-a), podem ser encontradas sugestões para a realização em *hardware* da respectiva rede neural proposta. Já em (RATHA; JAIN, 1995), são apresentadas considerações sobre a implementação de uma rede MLP para segmentação de imagens em FPGA (*Field Programmable Gate Array*), a qual pode ser utilizada como processador auxiliar em estações de trabalho SUN. Outra implementação aplicada à segmentação de imagens utilizando FPGAs é apresentada em (PÉREZ-URIBE; SANCHEZ, 1996), onde é utilizada uma rede neural que altera seu tamanho dinamicamente segundo a necessidade identificada na fase de treinamento.

Com base na bibliografia disponível, pode-se concluir que a utilização de técnicas baseadas em redes neurais na segmentação de imagens apresenta-se como uma alternativa atraente, com resultados práticos bastante satisfatórios, mesmo em imagens de baixa qualidade, distorcidas ou corrompidas por ruído. O estudo de técnicas de implementação destas redes em *hardware* dedicado, visando maior velocidade de processamento, também se apresenta como um campo de pesquisa interessante.

2.2.2 Redes Neurais Oscilatórias

Recentemente, topologias alternativas de redes neurais artificiais inspiradas no suposto funcionamento do córtex visual humano, as denominadas redes neurais oscilatórias, têm sido aplicadas em procedimentos de segmentação de imagens com resultados favoráveis. Nesta área de estudo, é particularmente relevante o trabalho desenvolvido por DeLiang Wang, o qual, segundo a bibliografia disponível, tem conduzido a quase totalidade das pesquisas com este tipo de rede em segmentação de imagens. O trabalho desenvolvido por Wang e seus colaboradores está centrado na proposição e aplicação de um novo modelo de rede neural artificial denominada LEGION (*Locally Excitatory Globally Inhibitory Oscillator Network*) (WANG; TERMAN, 1996).

A rede LEGION é construída com base em um arranjo de osciladores de relaxação localmente conectados e com um mecanismo de inibição global. Este mecanismo de conexões permite que osciladores vizinhos submetidos a entradas semelhantes oscilem em sincronismo de fase e em anti-sincronismo² com outros grupos de osciladores da rede. Desta forma, cada grupo de osciladores correspondente a uma região coerente do sinal de entrada da rede é ativado em um intervalo de tempo diferente dos demais

² O termo anti-sincronismo é empregado ao longo deste trabalho para indicar neurônios osciladores ou grupos de neurônios osciladores que estão ativos em intervalos de tempo distintos.

grupos, propiciando o efeito de uma segmentação temporal dos diversos objetos ou regiões homogêneas do sinal de entrada. Tal característica facilita a posterior identificação e quantificação de cada um dos segmentos obtidos na saída da rede.

Diversas aplicações práticas têm sido desenvolvidas para a rede LEGION, comprovando o funcionamento adequado da mesma. Por exemplo, em (LIU; CHEN; WANG, 2001) é encontrado um procedimento envolvendo redes MLP associadas a redes LEGION, o qual é utilizado para extração de regiões hidrográficas em imagens de sensoriamento remoto. Já em (SHAREEF; WANG; YAGEL, 1997), pode ser encontrado um estudo sobre a utilização de redes LEGION aplicadas à segmentação de imagens médicas de tomografia computadorizada e de ressonância magnética.

Um exemplo distinto de aplicação da rede LEGION em segmentação de sinais é encontrado em (WANG; BROWN, 1999). Neste caso, uma rede LEGION é utilizada para segmentar sinais de áudio (unidimensionais), mais especificamente para separar a voz de um locutor dos sinais interferentes oriundos do ambiente, utilizando atributos relativos à característica espectral do sinal a ser segmentado (voz do locutor), os quais são adequadamente transformados para posterior utilização da rede.

Os diversos resultados encontrados na bibliografia demonstram que a rede LEGION apresenta-se como a proposta mais consolidada de rede neural oscilatória para segmentação de imagens. Como atrativo prático, sua topologia maciçamente paralela possibilita implementações rápidas com a utilização de *hardware* específico. Sua peculiar capacidade para separar temporalmente os objetos das imagens, facilitando a posterior identificação dos mesmos, também se apresenta como forte atrativo. Por outro lado, limitações ainda apresentadas pelo modelo, tais como sua complexidade computacional para implementação digital e o elevado número de parâmetros, com ajuste conjunto pouco intuitivo, fazem da rede LEGION um assunto relevante para desenvolvimento de pesquisas, tanto no que se refere à sua estrutura, quanto a aplicações práticas da mesma.

2.3 Chips de Visão com Arquiteturas Maciçamente Paralelas

A crescente demanda por sistemas de visão artificial que implementam algoritmos complexos de forma suficientemente rápida tem justificado o desenvolvimento de circuitos integrados denominados *chips* de visão (KOMURO; ISHII; ISHIKAWA, 2001) ou retinas de silício (BARBARO et al., 2002). Nestes *chips*, os fotodetectores (do tipo CCD – *Charge Coupled Device*, por exemplo) são integrados em conjunto com uma rede maciçamente paralela de elementos de processamento (EPs) para execução de operações específicas sobre os pixels da imagem captada (Figura 2.5). O termo maciçamente paralela, neste caso, refere-se ao fato de que cada fotodetector possui um EP correspondente na rede de processamento.

As técnicas de processamento digital de imagens são normalmente consideradas fortes candidatas à aplicação de processamento paralelo, tendo em vista a grande quantidade de dados e a elevada complexidade de muitos algoritmos envolvidos. As operações denominadas de baixo nível, onde o processamento pixel a pixel de uma imagem gera uma outra imagem tipicamente de mesmas dimensões são particularmente adequadas a implementações em arquiteturas maciçamente paralelas (DOWNTON; CROOKES, 1998). Tais implementações enquadram-se na categoria SIMD (*Single Instruction stream Multiple Data stream*) (DOWNTON; CROOKES, 1998). As arquiteturas do tipo SIMD operam sobre vetores de dados, sendo todos os EPs sincronizados para executarem uma mesma tarefa de forma concorrente sobre as respectivas entradas.

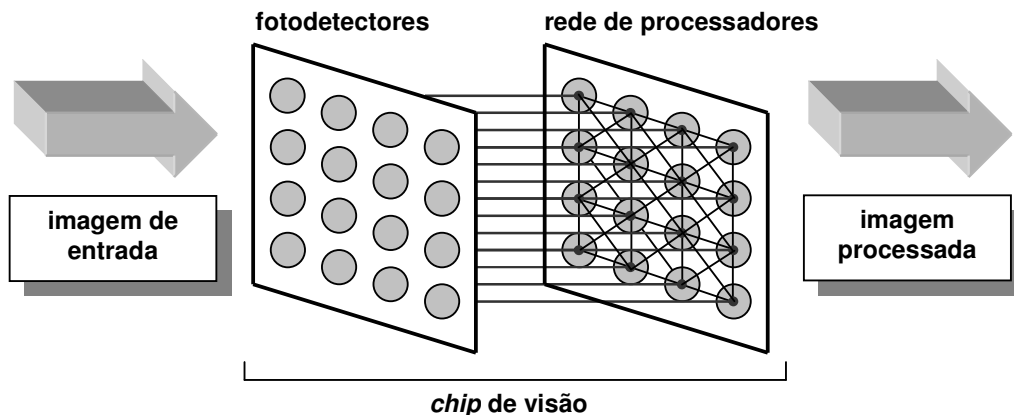


Figura 2.5: Integração dos sensores e rede de EPs em um *chip* de visão

Tendo em vista a grande quantidade de EPs, os chips de visão com arquiteturas maciçamente paralelas para processamento de imagens são normalmente implementadas para aplicações específicas (BARBARO et al., 2002; BOUBEKEUR; SAUCIER, 1990; FARROHA; DESHMUKH, 1995), evitando o uso de processadores de uso geral, mais complexos. Entretanto, em (KOMURO; ISHII; ISHIKAWA, 2002) pode ser encontrada uma proposta para implementação de um *chip* de visão programável, com 64x64 EPs, utilizando tecnologia digital. A família AR LSI, proposta pela Mitsubishi, apresenta retinas artificiais de até 352x288 pixels em cores, com menor custo e menor consumo em comparação com sensores CCD tradicionais (KYUMA et al., 1999). Nesta mesma referência, são propostas aplicações como detecção de bordas e reconhecimento de impressões digitais, tendo-se obtido, para uma versão de 32x32 pixels, uma capacidade de processamento de até 500 fps (quadros por segundo). É mencionado o desenvolvimento de uma versão de 640x480 pixels com capacidade para processar 100 fps (KYUMA et al., 1999). Uma retina artificial programável é apresentada em (PAILLET; MERCIER; BERNARD, 1999), tendo sido implementado um protótipo com 128x128 pixels, com expectativa de obtenção de uma versão de 512x512 pixels. Já em (BARBARO et al., 2002), é descrita a implementação de um *chip* analógico com 100x100 EPs com a finalidade de extrair o gradiente da imagem captada.

De forma geral, implementações analógicas de chips de visão requerem circuitos mais simples do que implementações digitais, apresentando, por outro lado, maior dificuldade quanto à reprogramação da função executada (KOMURO; ISHII; ISHIKAWA, 2001). Assim, quando é requerida flexibilidade para alterar a função a ser executada, justificam-se as implementações digitais.

2.4 Conclusão

As técnicas de segmentação de imagens exercem um papel fundamental dentro de sistemas de visão artificial, tendo em vista que os diversos objetos simultaneamente presentes em uma cena devem ser normalmente isolados para que possam ser submetidos a processos de identificação e quantificação.

Diferentes técnicas para segmentação de imagens podem ser encontradas na bibliografia, não existindo um método que possa ser considerado adequado para todos os tipos de aplicações e nem tampouco todos os métodos funcionam bem para qualquer tipo de imagem. Desta forma, a determinação da técnica para segmentação de imagens é normalmente um problema dependente da aplicação.

Um aspecto bastante importante em segmentação de imagens diz respeito à avaliação dos resultados do processo. Poucos procedimentos têm sido estabelecidos para este propósito, não existindo uma teoria consolidada sobre o assunto. A opinião de um observador humano é considerado ainda o melhor meio de avaliar a qualidade dos resultados de um procedimento de segmentação de imagens.

As técnicas de segmentação de imagens baseadas em redes neurais têm apresentado bons resultados mesmo com imagens corrompidas por ruído ou distorções de naturezas diversas, podendo apresentar vantagens como o paralelismo, o qual pode traduzir-se em maior rapidez de operação quando utilizado um *hardware* adequado. Como desvantagens, alguns métodos apresentam necessidade de treinamento, implicando o conhecimento antecipado de um número elevado de amostras corretamente segmentadas e elevado tempo de processamento.

Dentre os procedimentos para segmentação de imagens utilizando redes neurais encontram-se métodos baseados em técnicas clássicas, incluindo mapas de Kohonen, redes de Hopfield e redes MLP associadas ao algoritmo *Backpropagation*. Outras topologias de redes neurais menos populares também podem ser encontradas na bibliografia relativa à segmentação de imagens. Algumas referências apresentam propostas para implementação das respectivas redes em *hardware* específico, com o objetivo principal de explorar o ganho de velocidade devido ao paralelismo apresentado por algumas topologias empregadas.

A utilização de redes de osciladores para simular a capacidade de segmentação temporal atribuída ao cérebro humano tem se apresentado como uma alternativa mais recente e com resultados satisfatórios. O pesquisador DeLiang L. Wang e seus colaboradores têm desenvolvido trabalhos onde é proposta uma rede denominada LEGION, construída com base em osciladores de relaxação localmente conectados e com um mecanismo de inibição global. Os resultados encontrados na bibliografia demonstram que a rede LEGION tem apresentado bom desempenho quando utilizada em procedimentos de segmentação de imagens. Como aspectos positivos, a rede separa temporalmente os objetos das imagens e apresenta topologia maciçamente paralela, possibilitando implementações mais rápidas com *hardware* adequado. Como aspecto negativo relevante em termos de aplicação prática, a rede LEGION apresenta elevada complexidade computacional para implementação em *hardware* digital, além de muitos parâmetros, com procedimento de ajuste pouco prático.

Recentemente, a concepção dos chamados *chips* de visão têm atraído uma quantidade crescente de pesquisadores. Nestes *chips*, a integração dos sensores de imagens com uma rede maciçamente paralela de processadores possibilita a realização de operações específicas sobre os pixels da imagem de forma bastante rápida. Devido à sua arquitetura, os *chips* de visão são uma aplicação oportuna para a rede LEGION e para outros modelos de redes maciçamente paralelas para segmentação de imagens.

Com base no estudo realizado, conclui-se que o desenvolvimento de aplicações e a proposição de novos modelos de redes neurais artificiais com inspiração biológica no mecanismo de segmentação executado pelo cérebro humano, apresenta-se como um campo fértil para pesquisas. Da mesma forma, o desenvolvimento de arquiteturas de *hardware* para implementação de tais redes, explorando sua natureza maciçamente paralela, é aspecto também relevante. Finalmente, a implementação de *chips* de visão apresenta-se como aplicação promissora em termos da possibilidade da separação dos objetos de uma imagem de maneira mais rápida do que hoje é possível com a utilização de arquiteturas de processamento tradicionais.

3 A REDE NEURAL OSCILATÓRIA LEGION

Neste capítulo é apresentado um estudo sobre a estrutura, os princípios de funcionamento, vantagens, limitações e aplicações práticas da rede LEGION (*Locally Excitatory Globally Inhibitory Oscillator Network*). O desenvolvimento e a apresentação deste estudo fez-se necessária por tratar-se de assunto relativamente recente, não sedimentado e com bibliografia bastante restrita, além do que foi utilizado como base para o desenvolvimento do tema de tese.

Para compreensão do funcionamento da rede LEGION, é apresentado seu elemento de processamento, o neurônio oscilador de Terman-Wang, o qual tem comportamento modelado através de um par de equações diferenciais. Já no estudo da rede LEGION propriamente dita, é apresentado e analisado o funcionamento de redes de neurônios osciladores. O efeito de sincronismo local entre neurônios osciladores e de anti-sincronismo entre grupos de osciladores são também abordados. Um algoritmo para implementação eficiente da rede LEGION em máquinas seriais está também incluído no texto, além de um procedimento para filtragem de imagens, denominado FPS (*Feature-Preserving Smoothing*), originalmente desenvolvido para ajuste dos pesos da rede. Finalmente, são apresentados resultados obtidos em aplicações práticas originais utilizando a rede LEGION em segmentação de imagens, as quais foram implementadas ao longo do desenvolvimento da presente tese.

3.1 Neurônio Oscilador de Terman-Wang

No final da década de 80, foram descobertas oscilações de aproximadamente 40 Hz no córtex visual e em outras áreas do cérebro humano. Foi verificado que estas oscilações apresentam forte correlação com a coerência do estímulo visual, ocorrendo sincronismo de fase entre neurônios fisicamente próximos que recebem estímulos semelhantes, o que pode caracterizar uma região homogênea da imagem visualizada. Por outro lado, neurônios fisicamente próximos que recebem estímulos diferentes ou neurônios fisicamente distantes apresentam anti-sincronismo de fase (WANG, 1999). Tal propriedade é denominada correlação oscilatória e é uma forma especial de correlação temporal, onde cada objeto ou região homogênea de um estímulo recebido, no caso uma imagem, é representada por um grupo de osciladores sincronizados entre si e anti-sincronizados dos demais grupos de osciladores (WANG, 1999).

Desde a descoberta das oscilações neurais coerentes, diversas topologias de redes e estruturas de neurônios osciladores artificiais têm sido estudadas com o objetivo de criar modelos para o fenômeno. Verificou-se que redes de osciladores senoidais ou harmônicos, as quais necessitam de conexões entre todos os neurônios da rede para atingir sincronismo, não conduzem a resultados satisfatórios. Por outro lado, os osciladores de relaxação provaram ser uma alternativa adequada para a aplicação em questão (WANG, 1999).

O oscilador de Terman-Wang é uma proposta mais recente, a qual, por motivos matemáticos e computacionais, relacionados com complexidade e flexibilidade, é adotado como elemento de processamento básico para composição de redes neurais oscilatórias com aplicação em segmentação de sinais de naturezas diversas. Outro aspecto importante com relação ao modelo de oscilador proposto por Terman e Wang é o de que o mesmo pode ser implementado utilizando circuitos eletrônicos analógicos, o que se torna relevante no momento em que deseja-se obter implementações em *hardware* para redes neurais formadas por tais osciladores (WANG, 1999).

3.1.1 Modelo do Oscilador de Terman-Wang

O comportamento do oscilador de Terman-Wang pode ser descrito pelas equações diferenciais (3.1) e (3.2) (WANG, 1999). Sob o ponto de vista biológico, a variável de estado $x(t)$ pode ser entendida como o potencial da membrana da célula nervosa, ou seja, a grandeza física que representa a saída no neurônio.

$$\frac{dx(t)}{dt} = 3x(t) - x^3(t) + 2 - y(t) + I \quad (3.1)$$

$$\frac{dy(t)}{dt} = \varepsilon(\alpha(1 + \tanh(x(t)/\beta)) - y(t)) \quad (3.2)$$

onde:

$x(t)$ e $y(t)$ são variáveis de estado que representam o sistema, sendo $x(t)$ a saída;
 α , β e ε são parâmetros do modelo e
 I é uma entrada externa.

As equações (3.3) e (3.4) representam as denominadas *nullclines*³ do sistema de equações diferenciais apresentado por (3.1) e (3.2), as quais são obtidas quando ambas as derivadas são igualadas a zero (WANG, 1999).

$$y_1(t) = 3x(t) - x^3(t) + 2 + I \quad (3.3)$$

$$y_2(t) = \alpha(1 + \tanh(x(t)/\beta)) \quad (3.4)$$

A Figura 3.1 apresenta o plano de fase do oscilador de Terman-Wang com as respectivas *nullclines* considerando os parâmetros $I=1$, $\beta=0,2$ e $\alpha=3$. As *nullclines* funcionam como atratores para a fase ou estado do sistema ($[x(t) \ y(t)]^T$), existindo, no caso, um ponto de estabilidade (A), onde não ocorre oscilação, para o qual ambas as derivadas são simultaneamente nulas. De uma forma geral, todos os pontos de interseção de *nullclines* implicam respostas não oscilatórias. Para que o sistema sustente a oscilação, é portanto necessário que não existam cruzamentos de *nullclines* na trajetória de oscilação.

A Figura 3.2 apresenta a trajetória de fase do oscilador de Terman-Wang considerando os parâmetros $I=1$, $\beta=0,2$, $\alpha=3$. Os pontos indicados A, B e C correspondem a exemplos de condições iniciais para as equações diferenciais

³ As *nullclines* são as funções obtidas quando as derivadas de cada uma das variáveis de estado em relação ao tempo são igualadas a zero em um sistema de equações diferenciais. O termo em língua inglesa foi mantido tendo em vista não ter sido encontrado o termo equivalente em língua portuguesa nas fontes consultadas.

representadas por (3.1) e (3.2). Quando os pontos B ou C são utilizados como condições iniciais, o sistema adquire uma resposta oscilatória cuja trajetória de fase é representada pela linha mais espessa, com o sentido indicado. Os segmentos horizontais da trajetória correspondem a variações rápidas em $x(t)$ e variações muito lentas em $y(t)$. Já os segmentos quase verticais da trajetória de fase correspondem a variações muito lentas em $x(t)$ e variações rápidas em $y(t)$. A condição inicial representada por A ($x(0)=0$ e $y(0)=3$) implica uma situação de equilíbrio, com ambas as derivadas nulas, para a qual não ocorre oscilação, devendo portanto ser evitada. Uma maneira de evitar que o sistema assuma um estado estável indesejado é a adição de um ruído de pequena amplitude (ou de pequena variância, no caso de ruído gaussiano) na entrada externa do oscilador (WANG, 1999).

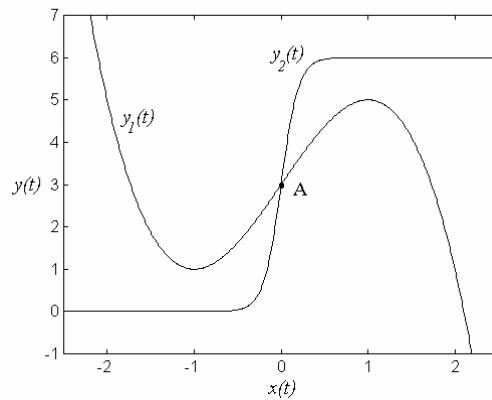


Figura 3.1: Plano de fase do oscilador de Terman-Wang e *nullclines* obtidas para $I=1$, $\beta=0,2$ e $\alpha=3$

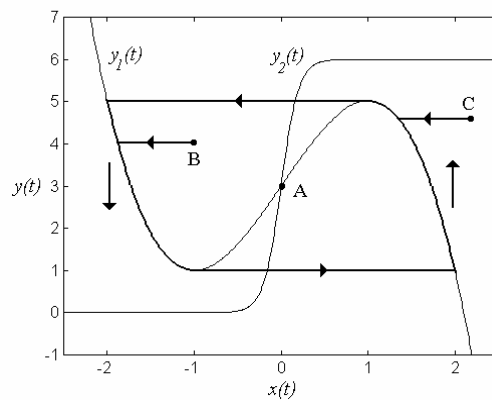


Figura 3.2: Trajetória de fase do oscilador de Terman-Wang para $I=1$, $\beta=0,2$, $\alpha=3$. Exemplos de condições iniciais: A, B e C

A Figura 3.3 representa a saída $x(t)$ do oscilador de Terman-Wang em função do tempo (com parâmetros $I=1$, $\beta=0,2$, $\alpha=3$, $\varepsilon=0,01$, $x(0)=0$ e $y(0)=0$), caracterizando o comportamento do mesmo como o de um oscilador de relaxação. Os intervalos de tempo nos quais o valor da amplitude da saída é alto ou positivo são chamados de fases ativas, enquanto que os intervalos de tempo nos quais o valor da amplitude da saída é baixo ou negativo são chamados de fases silenciosas (WANG, 1999).

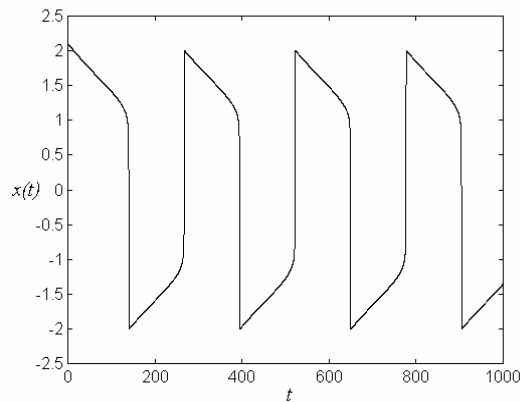


Figura 3.3: Saída do oscilador de Terman-Wang em função do tempo para $I=1$, $\beta=0,2$, $\alpha=3$, $\varepsilon=0,01$, $x(0)=0$ e $y(0)=0$

3.1.2 Parâmetros do Oscilador de Terman-Wang

Para melhor compreensão do comportamento do oscilador de Terman-Wang é necessária a avaliação da influência individual de cada parâmetro (I , α , β e ε).

O parâmetro I pode ser entendido como uma variável de entrada externa do oscilador. O valor de I implica deslocamento vertical da *nullcline* representada por (3.3). Se o referido deslocamento for suficientemente grande, conforme apresentado na Figura 3.4, ocorrerão pontos de interseção das *nullclines* dentro da trajetória de oscilação (A e B), implicando respostas não oscilatórias. No caso de $I \leq 0$, diz-se que o oscilador é excitável ou não estimulado (Figura 3.4(a)) (WANG, 1999).

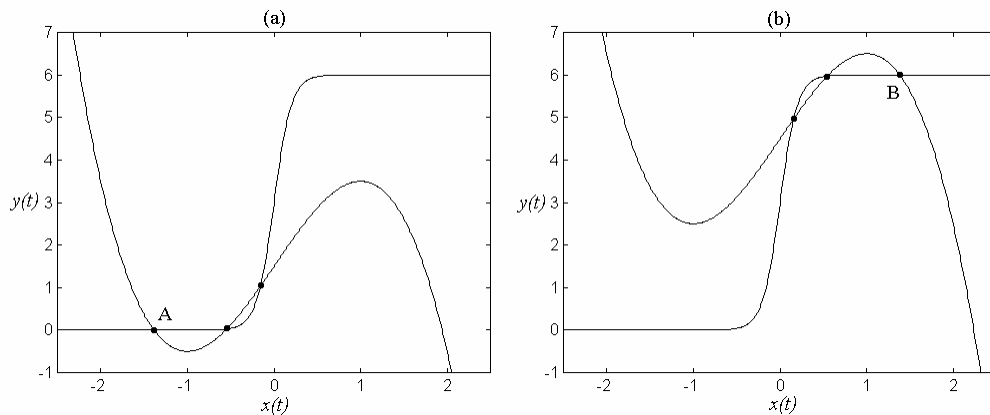


Figura 3.4: Efeito de I na condição de oscilação oscilador de Terman-Wang com $I=-1$ (a), $I=3$ (b), $\beta=0,2$ e $\alpha=3$

Para valores da variável de entrada I que resultam em comportamento oscilatório, pode ser observada uma correlação entre sua amplitude e o valor da razão entre os tempos de permanência nas fases ativa e silenciosa, conforme exemplificado através da Figura 3.5. Para que não ocorram interseções das *nullclines* na trajetória de oscilação, o valor da excitação externa recebida pelo oscilador deve estar portanto restrita à faixa $0 < I < (2\alpha - 4)$, situação na qual o oscilador é referido como habilitado (WANG, 1999).

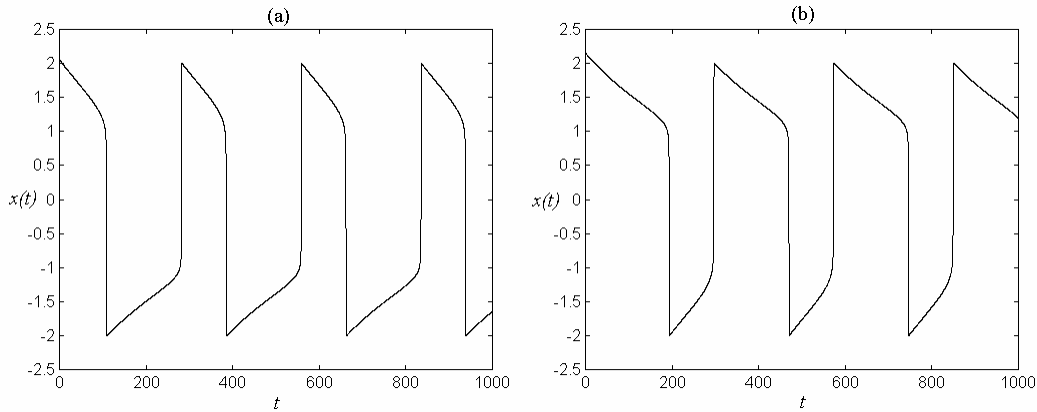


Figura 3.5: Efeito da variação de I na saída do oscilador de Terman-Wang. $I=0,5$ (a), $I=1,5$ (b), $\beta=0,2$, $\alpha=3$, $\varepsilon=0,01$, $x(0)=0$ e $y(0)=0$

O parâmetro α é responsável pelo tempo relativo de permanência nas fases ativa e silenciosa, controlando a amplitude da sigmóide representada por (3.4). Quanto maior o valor de α , menor o tempo de permanência na fase ativa. Valores reduzidos de α , os quais implicam um ponto de interseção das *nullclines* na trajetória de oscilação, resultam em respostas não oscilatórias. O parâmetro β controla a declividade da sigmóide representada e deve ter um valor adequado para que ocorra um único ponto de interseção entre as *nullclines*, o qual não deve estar na trajetória de oscilação. Já o parâmetro ε controla a taxa de variação de $y(t)$, determinando os tempos de permanência nas fases ativa e silenciosa e influenciando a frequência de oscilação do sistema. Normalmente, $\varepsilon \in (0,1)$ e $\varepsilon \ll 1$ (CHEN; WANG; LIU, 2000). Uma avaliação mais detalhada sobre o assunto pode ser encontrada em (FERNANDES, 2001-a).

Todos os resultados apresentados foram obtidos com auxílio de um aplicativo específico desenvolvido durante o trabalho de tese, o qual é apresentado no Apêndice 1.

3.2 Estrutura da Rede

A rede LEGION foi concebida associando-se a propriedade de sincronismo local entre osciladores acoplados a um mecanismo de inibição global para obter anti-sincronismo entre diversos grupos de osciladores. Desta forma, a rede obtida conta com duas estruturas de conexões, uma excitatória e outra inibitória (Figura 3.6).

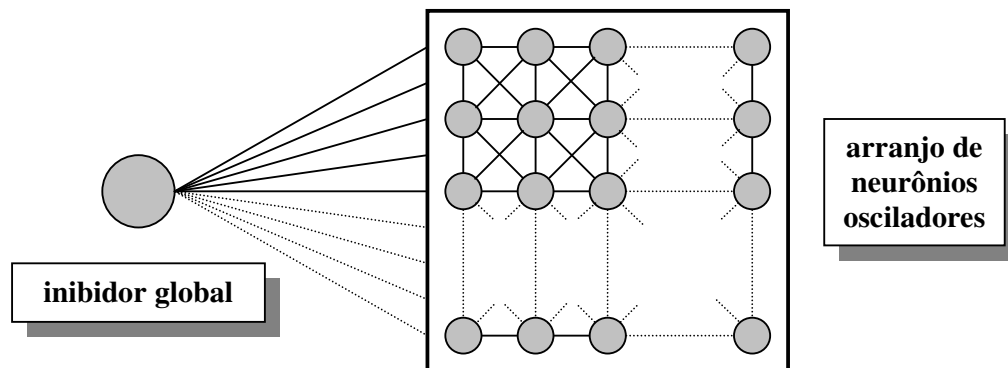


Figura 3.6: Exemplo de topologia de uma rede LEGION bidimensional

3.2.1 Neurônio Oscilador Básico

A rede LEGION é baseada em um conjunto de neurônios osciladores de Terman-Wang com termos adicionais em suas entradas, conforme representado pelas equações (3.5) e (3.6) (TERMAN; WANG, 1995).

$$\frac{dx_i(t)}{dt} = 3x_i(t) - x_i^3(t) + 2 - y_i(t) + I_i + S_i(t) + \rho_i \quad (3.5)$$

$$\frac{dy_i(t)}{dt} = \varepsilon(\alpha(1 + \tanh(x_i(t)/\beta)) - y_i(t)) \quad (3.6)$$

onde:

- $x_i(t)$ e $y_i(t)$ são as variáveis de estado que representam cada oscilador da rede, sendo $x_i(t)$ as saídas dos mesmos;
- $S_i(t)$ é a excitação lateral recebida dos osciladores vizinhos na rede;
- α , β e ε são parâmetros dos osciladores;
- I_i são as entradas externas e
- ρ_i é um ruído gaussiano com variância ρ e média $-\rho$.

A função do ruído gaussiano de pequena variância adicionado à entrada de cada oscilador é a de evitar que as condições iniciais da rede impliquem estados de estabilidade não desejados e também evitar o possível sincronismo entre diferentes blocos de osciladores (WANG, 1999).

A excitação lateral é definida como o acoplamento recebido dos demais osciladores da rede, sendo representada pela equação (3.7).

$$S_i(t) = \sum_{k \in N_i(1)} W_{ik} S_\infty(x_k(t), \theta_x) - W_z S_\infty(z(t), \theta_z) \quad (3.7)$$

com

$$\frac{dz(t)}{dt} = \phi(\sigma_\infty(t) - z(t)) \quad (3.8)$$

$$\sigma_\infty(t) = \begin{cases} 1 & , x_i(t) \geq \theta_z \\ 0 & , x_i(t) < \theta_z \end{cases} \quad (3.9)$$

$$S_\infty(x, \theta) = \frac{1}{1 + e^{-\kappa(x-\theta)}} \quad (3.10)$$

onde:

- W_{ik} são os pesos referentes às conexões entre os osciladores k e i ;
- W_z é o peso referente ao inibidor global;
- θ_x é um limiar acima do qual um oscilador pode ser afetado por seus vizinhos;
- θ_z é um limiar acima do qual um oscilador pode ser afetado pelo inibidor global;
- $N_i(R)$ é a vizinhança com distância R pixels do oscilador i ;
- κ é um parâmetro que ajusta a inclinação da respectiva sigmóide;
- $z(t)$ é a saída do inibidor global da rede e
- ϕ é um parâmetro que ajusta a taxa de variação da saída do inibidor global.

A função sigmóide representada por (3.10) pode ser substituída por uma função do tipo degrau (Heaviside), implicando $S(x, \theta)=0$ para $x < \theta$ e $S(x, \theta)=1$ para $x \geq \theta$, o que diminui a complexidade computacional em aplicações práticas.

Somente os osciladores pertencentes a uma vizinhança próxima possuem acoplamento, conforme pode ser observado no exemplo de topologia de rede LEGION bidimensional apresentada na Figura 3.6. A vizinhança de um oscilador, $N_i(R)$, é definida como o conjunto de osciladores que se encontram a uma distância inferior a $R+1$ em relação ao oscilador i . O inibidor global, por outro lado, é conectado a todos os osciladores da rede.

3.2.2 Sincronismo Entre Osciladores Acoplados

A capacidade de obter sincronismo de fase entre osciladores vizinhos submetidos a entradas externas semelhantes é um dos princípios básicos para o correto funcionamento da rede LEGION.

A Figura 3.7 exemplifica o caso de dois neurônios osciladores de Terman-Wang, denominados N_1 e N_2 , acoplados entre si. Os dois osciladores encontram-se inicialmente nos estados $N_1(0)$ e $N_2(0)$, respectivamente, ambos na fase silenciosa e desacoplados entre si. Com o passar do tempo, o neurônio N_1 atinge o estado de transição ($N_1(1)$) antes do neurônio N_2 , alterando sua saída rapidamente para a fase ativa. Nesta situação, o neurônio N_2 é estimulado pela saída do neurônio N_1 , fazendo com que a *nullcline* cúbica do mesmo seja deslocada para cima, implicando a transição rápida de N_2 para a fase ativa e o deslocamento de *nullcline* de N_1 de forma semelhante. Pode ser observado que no processo de transição entre as fases silenciosas e ativas e vice-versa, a distância entre os estados dos neurônios é reduzida gradativamente, o que indica a tendência ao sincronismo de fase entre as saídas dos mesmos. Tal efeito foi denominado “modulação rápida de limiar” (*fast threshold modulation*) (WANG, 1999).

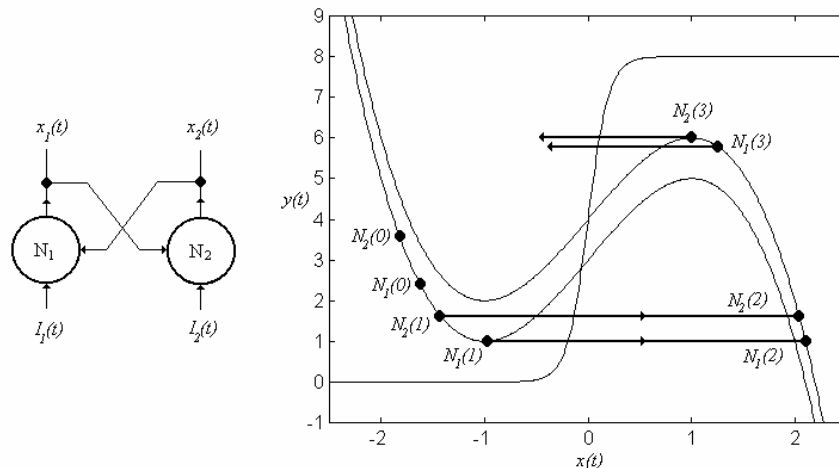


Figura 3.7: Mecanismo de sincronismo entre dois osciladores acoplados

Em um estudo realizado, Somers e Kopell provaram um teorema que afirma que um par de osciladores de relaxação acoplados pode ter um domínio de atração para o sincronismo com evolução geométrica ou exponencial (SOMERS; KOPELL, 1993). Estendendo tal constatação, Terman e Wang provaram um teorema afirmando que, para uma rede de osciladores de relaxação localmente acoplados de dimensões arbitrárias, existe um domínio de atração para o qual toda a rede tende ao sincronismo com uma taxa exponencial (WANG, 1993).

Assim, devido ao modelo matemático adotado, o oscilador de Terman-Wang apresenta-se como uma alternativa adequada para a construção de redes de osciladores que possuam a capacidade de obter sincronismo local entre osciladores com entradas externas semelhantes.

3.2.3 Inibição Global

A Figura 3.8 ilustra o funcionamento do mecanismo de inibição global da rede LEGION. São considerados dois neurônios osciladores, N_1 e N_2 , desacoplados e conectados ao inibidor global. Os dois neurônios encontram-se inicialmente nos estados $N_1(0)$ e $N_2(0)$, na fase silenciosa. Com o passar do tempo, N_1 atinge o estado de transição $N_1(1)$ antes de N_2 , alterando sua saída rapidamente para a fase ativa. Nesta situação, a entrada de N_2 é inibida pelo inibidor global, o qual é ativado por N_1 , fazendo com que as *nullclines* cúbicas de ambos sejam deslocadas para baixo, implicando o surgimento de pontos de estabilidade nas trajetórias de oscilação. Desta forma, enquanto N_1 estiver ativado, N_2 será impedido de pular para a fase ativa devido ao ponto de estabilidade criado em sua trajetória de oscilação. Ao retornar para a fase silenciosa, N_1 desabilita o inibidor global, fazendo com que as cúbicas dos neurônios sejam deslocadas para cima, eliminando os pontos de estabilidade e habilitando os mesmos a pularem para a fase ativa. No processo de transição entre as fases ativa e silenciosa, N_2 assume uma posição mais próxima do estado de transição em relação à N_1 , o que candidata o mesmo preferencialmente a pular para fase ativa, inibindo então N_1 .

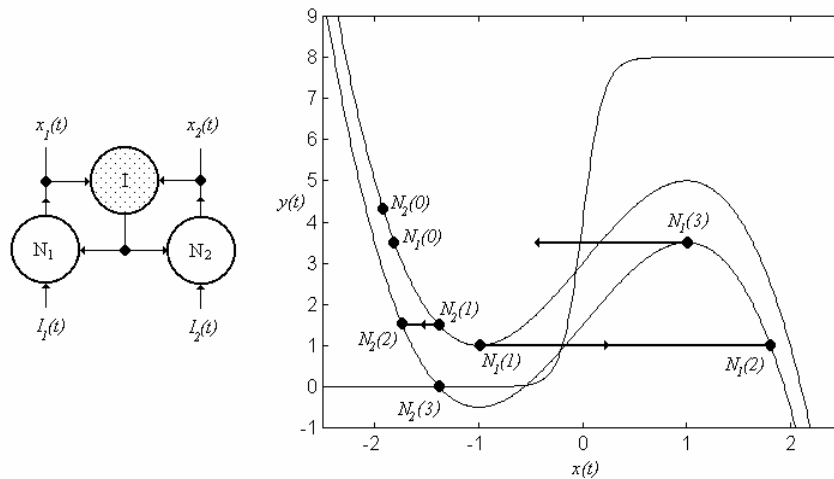


Figura 3.8: Mecanismo de inibição entre dois osciladores não acoplados

O processo descrito garante que se os neurônios osciladores N_1 e N_2 partirem de estados iniciais distintos, nunca estarão simultaneamente ativos. Se, por outro lado, N_1 e N_2 estiverem acoplados entre si e a excitação relativa ao acoplamento entre eles for mais intensa do que a inibição global, o neurônio N_1 , ao pular para a fase ativa, recrutará N_2 também para a fase ativa. Desta forma, cada grupo de neurônios osciladores acoplados tende a pular para a fase ativa em tempos distintos, caracterizando uma discriminação temporal de diferentes regiões do sinal de entrada da rede de osciladores.

3.2.4 Uma Rede LEGION Unidimensional

A Figura 3.9 apresenta uma rede LEGION unidimensional composta de 4 osciladores, implementada para verificar as propriedades anteriormente mencionadas.

Cada oscilador é somente acoplado com seus dois osciladores vizinhos em uma topologia tipo anel. Desta forma, é esperado que osciladores vizinhos, quando submetidos ao mesmo nível de sinal de entrada (I_i), devam oscilar com sincronismo de fase entre si e em anti-sincronismo com os demais osciladores da rede excitados por entradas com níveis diferentes.

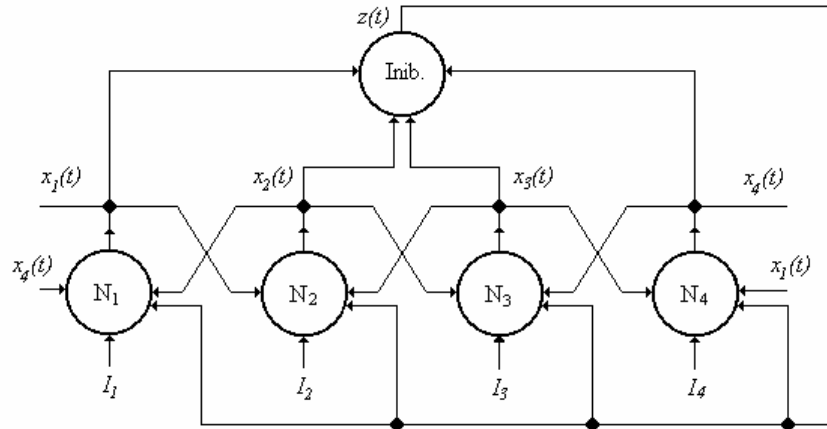


Figura 3.9: Rede LEGION unidimensional com 4 osciladores e conexão em anel

A Figura 3.10 apresenta as saídas dos osciladores da rede da Figura 3.9 em função do tempo, com parâmetros $\varepsilon=0,02$, $\alpha=3$, $\beta=0,1$, $W_{ik}=0,5$, $W_z=1,6$, $\theta_x=-0,5$, $\theta_z=0,1$, $\Phi=3$, $I_1=0,3$, $I_2=0,4$, $I_3=0,5$ e $I_4=0,5$. Pode ser observado que as saídas dos osciladores 3 e 4 atingem sincronismo rapidamente, indicando semelhança entre as respectivas entradas, ao contrário das saídas dos osciladores 1 e 2. O sinal de saída do inibidor global, demonstra que a rede é inibida (nível alto) sempre que algum dos osciladores estiver ativo. Já a Figura 3.11, apresenta os resultados obtidos com padrões de entrada $I_1=0,3$, $I_2=0,3$, $I_3=0,5$ e $I_4=0,5$, implicando oscilações em anti-sincronismo entre os osciladores com entradas distintas e oscilações em fase entre os osciladores com entradas semelhantes. Na Figura 3.12, a posição das entradas foi modificada ($I_1=0,3$, $I_2=0,5$, $I_3=0,3$ e $I_4=0,5$). Neste caso, os osciladores não atingem sincronismo pois não existe semelhança entre as entradas de neurônios vizinhos.

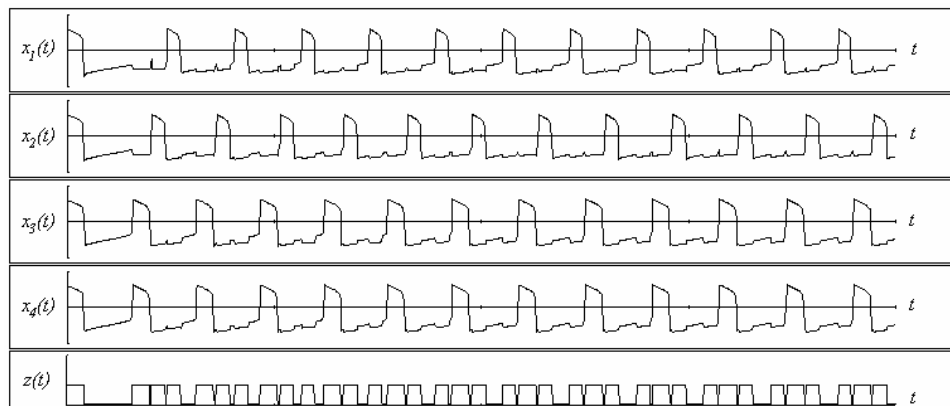


Figura 3.10: Saídas dos osciladores e do inibidor global da rede exemplo com 3 padrões de entrada distintos ($I_1=0,3$, $I_2=0,4$, $I_3=0,5$ e $I_4=0,5$)

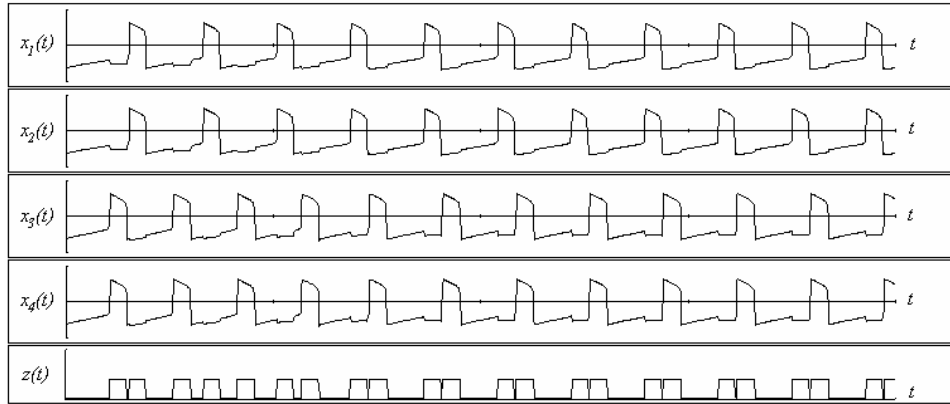


Figura 3.11: Saídas dos osciladores e do inibidor global da rede exemplo com 2 padrões de entrada distintos ($I_1=0,3$, $I_2=0,3$, $I_3=0,5$ e $I_4=0,5$)

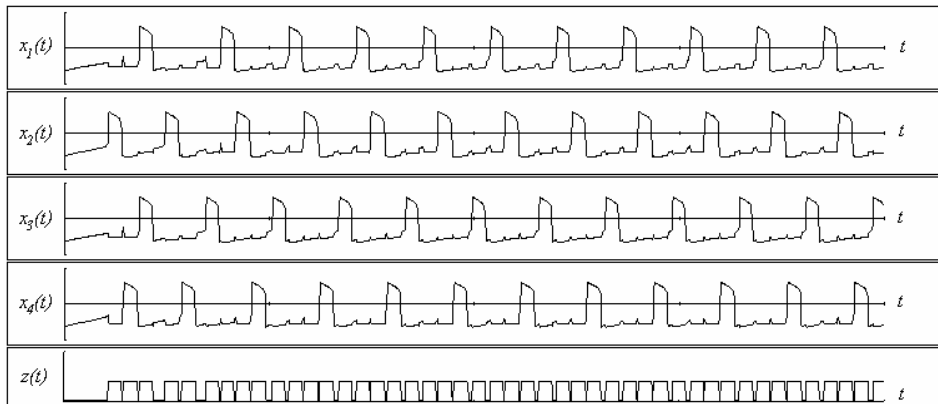


Figura 3.12: Saídas dos osciladores e do inibidor global da rede exemplo com 2 padrões de entrada distintos ($I_1=0,3$, $I_2=0,5$, $I_3=0,3$ e $I_4=0,5$)

Para obter os resultados apresentados, os conjuntos de equações diferenciais dos osciladores da rede foram resolvidos através do método de Runge-Kutta (KREYSZIG, 1982), implicando complexidade computacional relativamente elevada, ainda que a rede possua dimensões reduzidas. Em todas as simulações foi adicionado às entradas um ruído uniforme com média $-0,1I_i$ e amplitude máxima igual a $0,2I_i$ com o objetivo de evitar a possível ocorrência de estados estáveis nos osciladores e o sincronismo entre diferentes grupos de osciladores. Mais detalhes sobre o programa implementado para obter os resultados apresentados podem ser obtidos no Apêndice 1.

Um aspecto observado nas simulações da rede utilizada no exemplo apresentado, foi a dificuldade, para certos valores de entradas e parâmetros, de ser obtido o sincronismo entre osciladores adjacentes e o anti-sincronismo entre osciladores não adjacentes. Tal constatação indica a necessidade de um criterioso ajuste dos parâmetros da rede em função das entradas, podendo estabelecer-se como uma limitação de ordem prática.

3.2.5 Normalização dos Pesos Dinâmicos

Na concepção da rede LEGION, foi considerada a existência de dois tipos de pesos referentes às conexões entre os neurônios osciladores, denominados pesos dinâmicos (W_{ik}) e pesos permanentes (T_{ik}). Os pesos permanentes representam as conexões físicas

estruturais da rede, sendo estáticos ou de variação muito lenta, ao contrário dos pesos dinâmicos, os quais variam rapidamente em função da entrada da rede. Os pesos permanentes podem ser constantes ou ponderados em função da distância com uma distribuição gaussiana, por exemplo (WANG; TERMAN, 1997). Nenhum estudo sobre a modificação dos pesos permanentes foi encontrado na bibliografia, tratando-se de um aspecto a ser pesquisado.

Com o objetivo de melhorar a capacidade de sincronismo entre osciladores, foi proposto um mecanismo de normalização dos pesos dinâmicos, representado pelas equações (3.11) e (3.12) (WANG; TERMAN, 1997).

$$\frac{du_i}{dt} = \eta(1 - u_i)I_i - \nu u_i \quad (3.11)$$

$$\frac{dW_{ik}}{dt} = W_T T_{ik} u_i u_k - W_{ik} \sum_{j \in N(i)} T_{ij} u_i u_j \quad (3.12)$$

onde:

- u_i determina se um oscilador está estimulado ou não (se $I_i > 0$ então $u_i \rightarrow 1$),
- η determina a taxa de variação de u_i , sendo da ordem de 1,
- ν determina a taxa de decréscimo de u_i quando o estímulo é retirado ($I_i = 0$),
- W_{ik} é o peso dinâmicos entre os osciladores i e k e
- W_T é a soma dos pesos dinâmicos de um oscilador.

Os pesos W_{ik} são inicialmente iguados a zero para toda a rede. Se um oscilador não recebe estímulo externo, seus pesos dinâmicos são fixados em zero. Recebendo estímulo externo, um oscilador tem seus pesos dinâmicos em regime permanente representados por (3.13) e (3.14) (WANG; TERMAN, 1997), onde pode ser verificado que o somatório dos pesos dinâmicos de um oscilador deve ser igual a W_T . Se as entradas da rede forem mantidas constantes por um tempo suficientemente grande, pode ser observado que $u_i \rightarrow 1$ para entradas não nulas e $u_i \rightarrow 0$ para entradas nulas. Assim, os pesos dinâmicos referentes a um oscilador com entrada nula serão nulos.

$$W_{ik} = \frac{W_T T_{ik} u_i u_k}{\sum_{j \in N(i)} T_{ij} u_i u_j} \quad (3.13)$$

$$\sum_{k \in N(i)} W_{ik} = W_T \quad (3.14)$$

De acordo com (WANG; TERMAN, 1997), a normalização dos pesos não é uma condição necessária para o correto funcionamento da rede LEGION, melhorando, entretanto, o sincronismo entre osciladores vizinhos submetidos à excitações semelhantes. Os pesos W_{ik} podem ser determinados no início do processo com base nas entradas externas da rede, o que não implica complexidade computacional elevada vinculada a tal procedimento.

3.2.6 Osciladores Líderes

Em (WANG; TERMAN, 1997) é apresentada uma modificação na excitação do oscilador de Terman-Wang, a qual é representada pela equação (3.15). A idéia básica é

a de que um conjunto de osciladores com entradas semelhantes deve possuir ao menos um oscilador, denominador líder, o qual deve receber grande excitação lateral de sua vizinhança. Por outro lado, osciladores isolados, pertencentes a fragmentos ruidosos, não podem ser caracterizados como líderes. Desta forma, um oscilador com grande potencial lateral pode liderar a ativação de um bloco de osciladores (“legião”) correspondente a um padrão de entrada homogêneo.

$$\frac{dx_i(t)}{dt} = 3x_i(t) - x_i^3(t) + 2 - y_i(t) + I_i H(p_i + e^{-\gamma t} - \theta) + S_i(t) + \rho_i \quad (3.15)$$

$$\frac{dp_i}{dt} = \lambda(1 - p_i) H\left(\sum_{k \in N_i(R)} T_{ik} H(x_k - \theta_x) - \theta_p\right) - \mu p_i \quad (3.16)$$

com

$$H(v) = \begin{cases} 1 & , v \geq 0 \\ 0 & , v < 0 \end{cases} \quad (3.17)$$

onde:

- γ é um fator de decaimento positivo de mesma ordem que ε ;
- p_i é o potencial lateral que determina se um oscilador é líder ou não;
- λ é uma constante de tempo ($\lambda > 0$);
- μ é um fator de decaimento;
- T_{ik} são os pesos das conexões permanentes e
- θ , θ_x e θ_p são limiares com valores entre 0 e 1.

Da equação diferencial (3.16), deduz-se que se a excitação que um oscilador recebe de sua vizinhança excede um limiar θ_p , o potencial lateral p_i se aproxima de $\lambda/(\lambda + \mu)$ e o oscilador i é considerado um líder se $\lambda/(\lambda + \mu) > \theta$. Por outro lado, se o oscilador não recebe excitação lateral suficiente, p_i tende a zero com o passar do tempo. Para que a excitação lateral seja elevada, deve haver sincronismo entre os osciladores da vizinhança, sendo necessário que todos eles atinjam um determinado limiar θ_x ao mesmo tempo. O termo $e^{-\gamma t}$ habilita como líderes os osciladores estimulados durante um determinado período de tempo, com o objetivo de iniciar o processo (WANG; TERMAN, 1997).

Para verificar o comportamento da rede com a modificação proposta, foram realizadas simulações utilizando uma configuração bidimensional da rede LEGION com 20x20 osciladores, implementada de acordo com as equações (3.15), (3.6), (3.7) (utilizando a função degrau), (3.8), (3.11) e (3.12). A Figura 3.13 apresenta os resultados obtidos com a referida rede na segmentação de uma imagem binária sintética de tamanho 20x20 pixels (Figura 3.13(a)). As amplitudes dos pixels pretos foram fixadas em 0,2 e a dos pixels brancos em 0 e mantidas constantes, sendo possível determinar os pesos das conexões dinâmicas no início do processo. O estado inicial de cada oscilador é aleatório, conforme apresentado na Figura 3.13(b). Foi também acrescentado às entradas um ruído gaussiano com média igual a -0,02 e variância igual a 0,02, equivalentes a 10% do nível máximo estipulado para a excitação externa. Após um determinado tempo, as saídas dos osciladores passam a fornecer seqüencialmente cada uma das letras, isoladamente, conforme apresentado na Figura 3.13 (c), (d), (e) e (f). A separação de cada uma das letras temporalmente é uma característica bastante interessante e que facilita a posterior aplicação de algum procedimento de identificação de caracteres.

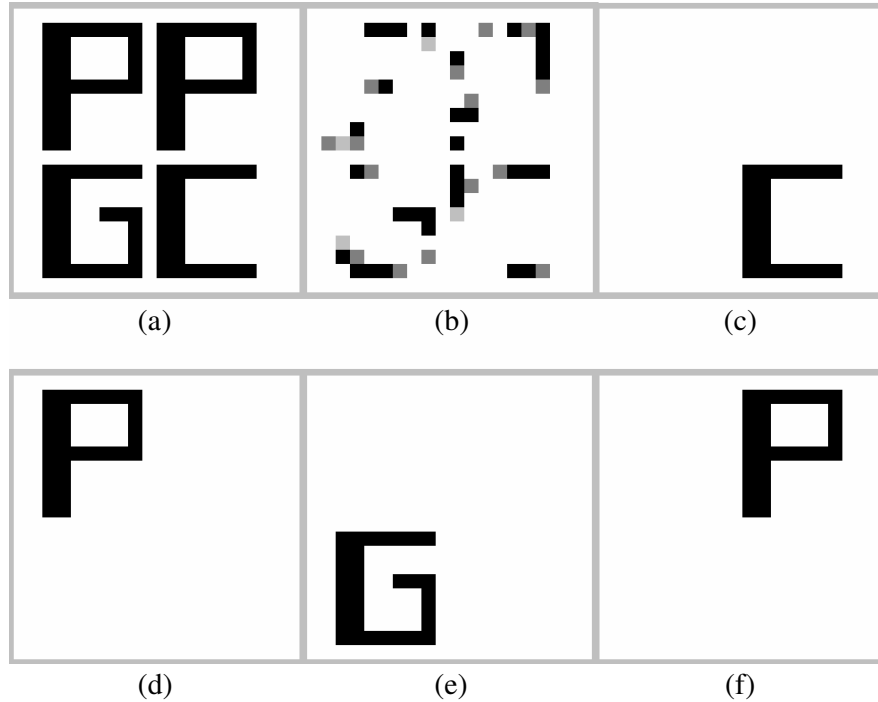


Figura 3.13: Segmentação com rede LEGION. Imagem de entrada (a). Saídas em tempos distintos (b), (c), (d), (e) e (f)

Os parâmetros da rede de osciladores utilizados para obter os resultados da Figura 3.13 foram $\varepsilon=0,02$, $\alpha=6$, $\beta=0,1$, $\theta_z=0,1$, $\theta_x=-1,1$, $\eta=1,0$, $\nu=0,01$, $\theta_p=5$, $\theta=0,9$, $\phi=3$, $\lambda=0,1$, $\mu=0,01$, $\gamma=0,005$, $W_z=1,5$, $W_T=6$ e $T_{ij}=2$. Tanto para as conexões fixas quanto para as conexões dinâmicas foi considerada uma relação de vizinhança que considera que cada oscilador da rede pode receber excitação lateral dos 4 pixels vizinhos posicionados segundo as direções vertical e horizontal.

O resultado apresentado, assim como outros resultados de simulações utilizando a rede LEGION em segmentação de imagens demonstraram que a mesma se apresenta como uma ferramenta eficiente para este tipo de aplicação. Como aspectos positivos, destacam-se a propriedade de segmentação temporal apresentada pela rede e sua topologia maciçamente paralela. Como aspectos negativos, foi constatada a elevada complexidade computacional para implementação em *hardware* digital, visto o grande número de equações diferenciais a serem solucionadas, e ainda uma certa dificuldade de sincronismo entre os neurônios osciladores para determinadas combinações de imagens de entrada e de parâmetros da rede.

3.3 Algoritmo para Implementação da Rede

A utilização da versão contínua da rede LEGION, baseada em equações diferenciais, implica elevada complexidade computacional na segmentação de imagens reais compostas por grande número de pixels. A aplicação prática de tal topologia é adequada com a utilização de um *hardware* analógico que explore o paralelismo da mesma.

Com o objetivo de reduzir o esforço computacional na utilização de computadores seriais, Wang desenvolveu uma versão algorítmica simplificada da rede LEGION, a qual é baseada nos mesmos princípios operacionais da versão contínua apresentada (WANG; TERMAN, 1997).

No algoritmo desenvolvido, adequado para segmentação de imagens, somente a variável de estado $x(t)$ é considerada, pois representa a saída efetiva do oscilador. Foi considerada a relação de vizinhança $N_i(1)$ que considera os 8 pixels vizinhos para determinação de excitação lateral de cada oscilador, o que, entretanto, pode ser facilmente modificado segundo a conveniência. O algoritmo pode ser dividido em três grandes etapas, as quais são apresentadas a seguir:

1) Fase de inicialização

- Fazer o inibidor global $z(t)$ igual a zero:

$$z(0) = 0$$

- Determinar os pesos dinâmicos W_{ik} :

$$W_{ik} = \frac{I_M}{1 + |I_i - I_k|}, k \in N(i) \quad (3.18)$$

onde

I_M é a máxima amplitude encontrada entre os pixels da imagem.

- Encontrar os neurônios osciladores líderes p_i :

$$p_i = H\left(\sum_{k \in N(i)} W_{ik} - \theta_p\right) \quad (3.19)$$

- Posicionar as saídas de todos os osciladores aleatoriamente na fase silenciosa:

$$-2 < x_j(0) < -1$$

2) Determinação do primeiro oscilador a passar para a fase ativa

- Considerando que todos os osciladores (x_k) estão na fase silenciosa, escolher o líder ($p_j=1$) que estiver mais próximo do ponto de transição para a fase ativa (estado $x=-1$) e transportá-lo para a fase ativa (estado $x=1$). Habilitar também o inibidor global ($z(t)>0$) e evoluir os estados dos demais osciladores na fase silenciosa:

$$x_j(t+1) = 1 \text{ e } z(t+1) = 1 \text{ para } x_j(t) \geq x_k(t), \forall k \quad (3.20)$$

$$x_k(t+1) = x_k(t) - 1 - x_j(t), k \neq j \quad (3.21)$$

3) Dinâmica da rede

- Manter os osciladores que estão na fase ativa neste estado se o inibidor global foi incrementado. Retornar os osciladores que estão na fase ativa para a fase silenciosa (estado $x=-2$) se o inibidor global tiver sido decrementado ou mantido constante. Para cada oscilador que retorna à fase silenciosa decrementar o

inibidor global. Quando não existirem mais neurônios osciladores na fase ativa ($z(t)=0$), retornar à etapa 2.

$$x_i(t+1) = x_i(t) \text{ se } z(t) > z(t-1) \text{ e } x_i(t) = 1 \quad (3.22)$$

$$x_i(t+1) = -2 \text{ e } z(t) = z(t-1) - 1 \text{ se } z(t) \leq z(t-1) \text{ e } x_i(t) = 1 \quad (3.23)$$

- Determinar a excitação lateral S_i dos osciladores que estão na fase silenciosa segundo um dos critérios representados por (3.24), (3.25) e (3.26). Aqueles que possuírem potencial suficiente passarão para a fase ativa e o inibidor global será incrementado.

$$S_i(t) = \sum_{k \in N(i)} W_{ik} H(x_k(t) + 1) - W_z H(z(t) - 0,5) \quad (3.24)$$

ou

$$S_i(t) = \text{Max}_{k \in N(i)} (W_{ik} H(x_k(t) + 1)) - W_z H(z(t) - 0,5) \quad (3.25)$$

ou

$$S_i(t) = \frac{\sum_{k \in N(i)} W_{ik} H(x_k(t) + 1)}{\ln(1 + \sum_{k \in N(i)} H(x_k(t) + 1))} - W_z H(z(t) - 0,5) \quad (3.26)$$

$$x_i(t+1) = 1 \text{ e } z(t) = z(t-1) + 1 \text{ se } S_i(t) > 0 \text{ e } x_i(t) < 1 \quad (3.27)$$

$$x_i(t+1) = x_i(t) \text{ e } z(t) = z(t-1) \text{ se } S_i(t) \leq 0 \text{ e } x_i(t) < 1 \quad (3.28)$$

- Executar a dinâmica da rede até que seja alcançada a condição de desvio para a etapa 2 (nenhum oscilador na fase ativa).

A equação (3.18) indica que, quanto maior a semelhança entre dois pixels adjacentes, mais forte é a excitação lateral entre os respectivos osciladores, o que é intuitivo em se tratando de um agrupamento por semelhança. Deve ser ressaltado, entretanto, que a equação (3.18) considera a semelhança do atributo intensidade do sinal de entrada. Outros atributos poderiam ser considerados no processo de segmentação, tal como a textura, por exemplo, implicando a modificação adequada da referida equação.

Outro aspecto relevante relacionado com a equação (3.18) diz respeito à relação não linear entre os pesos dinâmicos e as diferenças de intensidades entre entradas adjacentes. A Figura 3.14 apresenta a curva que relaciona os pesos dinâmicos com as diferenças entre entradas adjacentes para o caso onde o sinal de entrada é uma imagem com intensidades entre 0 e 255 (8 bits), podendo ser verificada a relação não linear entre as mesmas. Neste caso, os pesos determinados ponderam com muito maior relevância as pequenas diferenças entre entradas adjacentes. Para o caso de considerar-se uma imagem normalizada, com valores de intensidade entre 0 e 1, um efeito não linear semelhante pode ser obtido através da utilização da equação (3.29) para determinação dos pesos da rede LEGION.

$$W_{ik} = \frac{1}{1 + 255|I_i - I_k|}, k \in N_i(1) \quad (3.29)$$

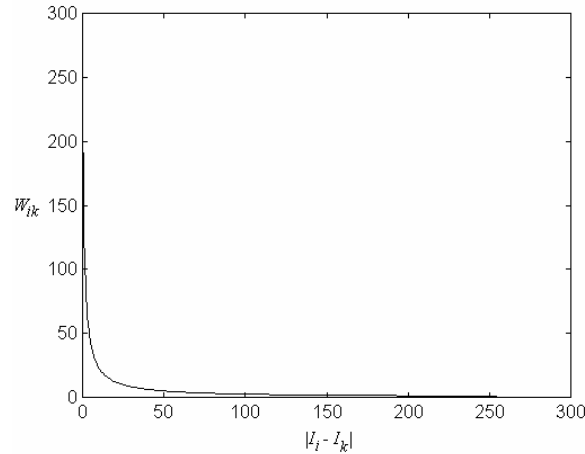


Figura 3.14: Relação não linear entre os pesos e as diferenças entre entradas adjacentes para uma imagem de 8 bits

O algoritmo para a rede LEGION apresentado prevê que os líderes e os pesos das conexões dinâmicas devam ser determinados na fase de inicialização, não necessitando atualização posterior e, portanto, não implicando elevada complexidade computacional relativa no processo. Deve ser observado também que o número de líderes não corresponde ao número de segmentos discriminados, já que cada segmento pode possuir vários líderes.

Um procedimento alternativo para determinação dos líderes é apresentado em (CHEN; WANG; LIU, 2000), o qual é representado pelas equações (3.30) e (3.31) que são as condições a serem cumpridas por um oscilador para que o mesmo seja considerado líder. Tais condições elegem os pixels localizados no centro de regiões homogêneas como líderes com base, respectivamente, na média (3.32) e na variância (3.33) das intensidades dos pixels calculados em uma vizinhança de raio predeterminado. As médias e as variâncias podem ser normalizados para o intervalo [0;1] através de (3.34) e (3.35).

$$|\mu_{ij}(R) - \mu_{ij}(1)| \leq T_\mu \quad (3.30)$$

$$|\sigma_{ij}^2(R) - \sigma_{ij}^2(1)| \leq T_\sigma \quad (3.31)$$

$$\mu_{ij}(R) = \frac{\sum_{(k,l) \in N_{ij}(R) \cup (i,j)} I_{kl}}{|N_{ij}(R)| + 1} \quad (3.32)$$

$$\sigma_{ij}^2(R) = \frac{\sum_{(k,l) \in N_{ij}(R) \cup (i,j)} (I_{kl} - \mu_{ij}(R))^2}{|N_{ij}(R)| + 1} \quad (3.33)$$

$$\tilde{\mu}_{ij}(R) = \frac{\mu_{ij}(R) - \mu_{\min}(R)}{\mu_{\max}(R) - \mu_{\min}(R)} \quad (3.34)$$

$$\tilde{\sigma}_{ij}^2(R) = \frac{\sigma_{ij}^2(R) - \sigma_{\min}^2(R)}{\sigma_{\max}^2(R) - \sigma_{\min}^2(R)} \quad (3.35)$$

onde:

$\mu_{ij}(R)$ é a média das intensidades dos pixels calculada em uma vizinhança de raio R do pixel ij ;

$\sigma_{ij}^2(R)$ é a variância das intensidades dos pixels calculada em uma vizinhança de raio R do pixel ij ;

$\mu_{\max}(R)$ é a média máxima das intensidades dos pixels calculada em uma vizinhança de raio R do pixel ij ;

$\mu_{\min}(R)$ é a média mínima das intensidades dos pixels calculada em uma vizinhança de raio R do pixel ij ;

$\sigma_{\max}^2(R)$ é a variância máxima das intensidades dos pixels calculada em uma vizinhança de raio R do pixel ij ;

$\sigma_{\min}^2(R)$ é a variância mínima das intensidades dos pixels calculada em uma vizinhança de raio R do pixel ij ;

$\tilde{\mu}_{ij}(R)$ é a média normalizada das intensidades dos pixels calculada em uma vizinhança de raio R do pixel ij ;

$\tilde{\sigma}_{ij}^2(R)$ é a variância normalizada das intensidades dos pixels calculada em uma vizinhança de raio R do pixel ij ;

$|N_{ij}(R)|$ é o número de pixels pertencentes a vizinhança de raio R do pixel ij ;

T_{μ} é um limiar relacionado com a média em torno de um líder e

T_{σ} é um limiar relacionado com a variância em torno de um líder.

Existem dois parâmetros ajustáveis no algoritmo da rede LEGION apresentado: W_z e θ_p . O primeiro controla a intensidade do inibidor global, sendo que, para valores mais elevados, são obtidas mais regiões segmentadas e de menor tamanho. O segundo parâmetro determina a quantidade de líderes gerados, sendo que, para valores elevados, são obtidos menos líderes e menos regiões segmentadas.

Em aplicações práticas de segmentação de imagens, o algoritmo pode ser interrompido após todos os líderes terem completado pelo menos um ciclo, passando da fase silenciosa para a fase ativa e retornando à fase silenciosa (WANG; TERMAN, 1997). Para obter então os segmentos encontrados, basta verificar os grupos de osciladores com saídas idênticas, quando todos os osciladores estão na fase silenciosa, correspondendo cada grupo a um segmento da imagem.

As equações (3.24), (3.25) e (3.26) representam três critérios diferentes para determinação da excitação lateral recebida por cada um dos osciladores da rede. O primeiro, denominado critério da soma ou do somatório, pondera as semelhanças entre um pixel e sua vizinhança. O segundo critério, denominado critério do máximo ou da maximização, enfatiza a relação entre um pixel e seu vizinho mais semelhante localizado na fase ativa. Em (WANG; TERMAN, 1997), os dois primeiros critérios foram comparados na segmentação de imagens de sensoriamento remoto, concluindo-se que o critério da soma resulta em regiões mais homogêneas em comparação com o critério do máximo, podendo entretanto ignorar pequenos detalhes da imagem no processo de segmentação. Já o terceiro critério é apresentado em (CHEN; WANG; LIU, 2000) como uma melhor alternativa aos dois anteriores, tendo sido avaliados resultados

com a utilização de imagens sintéticas, imagens de sensoriamento remoto e imagens médicas de ressonância magnética.

A Figura 3.15 apresenta um exemplo de segmentação realizado com a implementação da versão algorítmica da rede LEGION. A imagem original, com dimensões 100x310 pixels, observada na posição superior esquerda, possui 5 letras de mesmo tom de cinza isoladas umas das outras e sobrepostas a um fundo com diferente tom de cinza. A rede com parâmetros $\theta_p=4$ e $W_z=0,5$ determinou, a partir da intensidade de cada pixel da imagem original, a existência de 7 segmentos, os quais são representados pelas demais imagens. O primeiro segmento, apresentado na posição superior direita, corresponde ao fundo, tendo sido a parte interior da letra "R" desconsiderada por não estar conectada ao mesmo. Esta pequena região aparece, entretanto, como um segmento isolado, o qual pode ser observado na imagem da posição direita inferior. Os segmentos obtidos são apresentados seqüencialmente na saída da rede, observando-se que todas as letras foram corretamente segmentadas e isoladas no tempo, facilitando a aplicação posterior de um procedimento de reconhecimento de caracteres. Em todas as imagens referentes aos segmentos obtidos, a cor preta representa osciladores na fase silenciosa e a cor branca representa os osciladores na fase ativa. Para a imagem de entrada apresentada foram obtidos os mesmos resultados com a utilização dos três procedimentos para a determinação da excitação lateral apresentados. O aplicativo desenvolvido para obtenção destes resultados é apresentado no Apêndice 1.

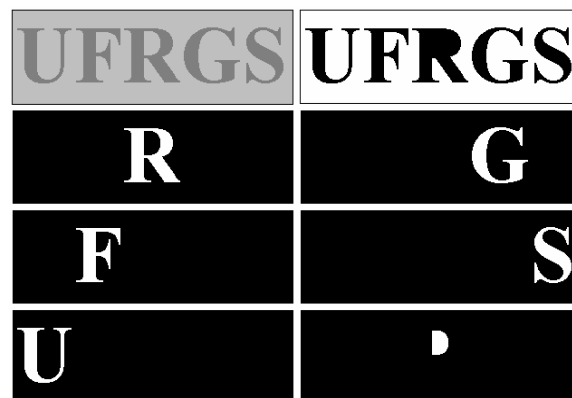


Figura 3.15: Segmentação obtida com a utilização da versão algorítmica da rede LEGION com parâmetros $\theta_p=4$ e $W_z=0,5$

Os resultados obtidos demonstram que a rede LEGION, em sua versão algorítmica, apresenta-se como uma ferramenta eficiente para aplicações de segmentação de imagens. O aumento significativo na velocidade de segmentação das imagens torna esta versão da rede LEGION mais adequada para implementações que utilizam computadores seriais em comparação à versão contínua da rede, mais adequada à implementações paralelas.

3.4 Outros Aspectos Relacionados à Rede LEGION

A seguir são abordados outros aspectos relevantes relacionados com as versões apresentadas da rede LEGION. Tais aspectos envolvem o número máximo de segmentos que podem ser discriminados pela rede, o denominado efeito de percepção dúplex e a influência do ruído nos resultados obtidos. Com relação ao último aspecto, é

apresentado um algoritmo para filtragem não linear de imagens, o qual foi originalmente desenvolvido para ser utilizado como um mecanismo para o ajuste dos pesos da rede LEGION.

3.4.1 Capacidade de Segmentação

Terman e Wang provaram que existe um domínio de parâmetros e de condições iniciais para o qual a rede LEGION atinge sincronismo entre blocos de osciladores acoplados e anti-sincronismo entre diferentes blocos de osciladores em um número de ciclos menor ou igual ao número de regiões a serem segmentadas presentes no sinal de excitação externo da rede (WANG, 1999). O ciclo em questão equivale a passagem de todos os osciladores correspondentes aos segmentos a serem discriminados da fase silenciosa para a fase ativa e da fase ativa para a fase silenciosa.

Por outro lado, foi constatado que, para um determinado conjunto de parâmetros, a rede LEGION, em sua versão contínua, pode discriminar um número limitado de segmentos, o qual depende da razão entre os tempos de permanência nas fases ativa e silenciosa dos osciladores. Este limite é denominado capacidade de segmentação da rede e situa-se na faixa de 5 a 10 segmentos na prática (WANG, 1999). Esta característica, apesar de apresentar-se como uma restrição à aplicação prática da rede em segmentação de imagens, tem correlação biológica pois também é observada em seres humanos, os quais apresentam uma limitação quantitativa para discriminação simultânea de diversos objetos (WANG; TERMAN, 1997). A versão algorítmica da rede LEGION, por sua vez, não apresenta a referida limitação quanto ao número de segmentos simultaneamente discriminados, característica bastante positiva para aplicações práticas da mesma.

3.4.2 Percepção Dúplex

O fenômeno denominado percepção dúplex ocorre nos seres humanos quando alguma região daquilo que é percebido faz parte da composição de elementos distintos (WANG; TERMAN, 1996). Tal fenômeno pode ocorrer na visão quando uma determinada região da imagem pode ser entendida como pertencente simultaneamente a objetos distintos. Na percepção de fala, por exemplo, uma determinada região do espectro de frequências pode, simultaneamente, fazer parte da composição de fonemas adjacentes e diferentes entre si.

A Figura 3.16 apresenta, como exemplo, o resultado da segmentação de uma imagem com a rede LEGION, onde pode ser verificado o fenômeno da percepção dúplex. A imagem original (Figura 3.16(a)), com 7×7 pixels e 8 bits, contém uma região de intensidade 150 conectada simultaneamente a duas regiões com intensidade 90 e separadas entre si. Utilizando os parâmetros $\theta_p=2,5$ e $W_z=0,04$, foram obtidos três segmentos, podendo ser observado que a região de intensidade 150 foi incluída simultaneamente em dois segmentos distintos. Com o aumento de W_z entretanto, a imagem foi segmentada em 4 regiões, onde a área de intensidade 150 surge como um segmento isolado.

A utilização do critério do máximo na determinação da excitação lateral dos neurônios osciladores da rede LEGION impossibilita a ocorrência do fenômeno da percepção dúplex, sendo portanto sua utilização mais indicada do que a do critério da soma, sob tal ponto de vista (WANG; TERMAN, 1996). Com relação ao critério representado por (3.26), o qual é apontado por (CHEN; WANG; LIU, 2000) como mais adequado sob o ponto de vista da qualidade da segmentação, apresenta também possibilidade de percepção dúplex, o que pode também limitar sua aplicação prática.

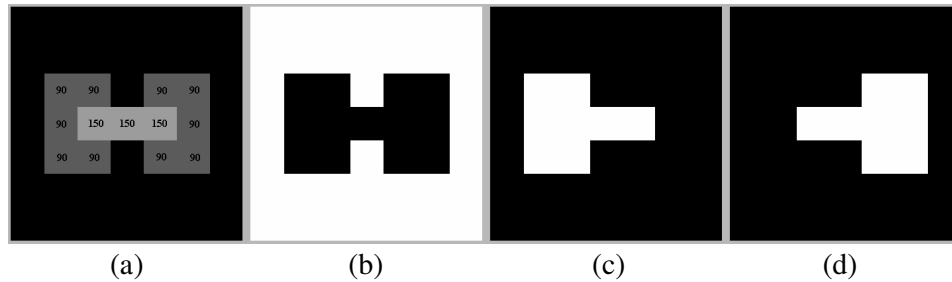


Figura 3.16: Fenômeno da percepção dúplex no resultado da segmentação obtido com a rede LEGION

Apesar de ser um fenômeno também observado em seres humanos, a ocorrência de percepção dúplex na segmentação de imagens não está em conformidade com a condição apresentada pela equação (2.1), na qual um pixel da imagem não pode pertencer simultaneamente a dois segmentos distintos. Dependendo da aplicação, os parâmetros da rede LEGION e o procedimento de determinação da excitação lateral deverão ser cuidadosamente escolhidos para que o fenômeno da percepção dúplex não seja observado.

3.4.3 O Algoritmo FPS

Um dos problemas que dificultam a segmentação de imagens, o qual é freqüentemente encontrado em aplicações práticas é a degradação do sinal de entrada por ruído ou distorções de naturezas diversas. Para atenuar o problema causado pelo ruído em aplicações de segmentação de imagens, são normalmente utilizadas técnicas de suavização em uma etapa anterior ao procedimento de segmentação, incluindo métodos iterativos de filtragem não linear. Estas técnicas de suavização também podem ser agregadas aos próprios métodos de segmentação, não sendo explicitamente observadas.

Os resultados apresentados em (FERNANDES, 2001-a) e obtidos com a utilização da rede LEGION na segmentação de imagens corrompidas por ruído comprovam também ser necessária a utilização, em alguns casos, de alguma estratégia de filtragem sobre imagem original, com o objetivo de enfatizar as regiões homogêneas da mesma.

Em (CHEN; WANG; LIU, 2000; CHEN; WANG, 1999) uma rede LEGION é utilizada para segmentar imagens de sensoriamento remoto e de ressonância magnética, tendo sido proposto um algoritmo de suavização denominado FPS (*Feature-Preserving Smoothing*), o qual reduz o ruído preservando detalhes presentes na imagem original e não está sujeito ao denominado “problema da terminação”. O “problema da terminação” pode ser observado, por exemplo, quando uma imagem é submetida recursivamente a uma filtragem do tipo passa-baixas, resultando, após um grande número de iterações, em uma imagem completamente homogênea e inútil sob o ponto de vista da segmentação (CHEN; WANG; LIU, 2000). Também, a definição a priori do número ideal de iterações do filtro para um procedimento genérico é uma tarefa não trivial.

O algoritmo FPS pode ser utilizado em conjunto com a rede LEGION, promovendo uma atualização gradativa dos pesos cujo efeito é equivalente ao de uma filtragem que atenua o ruído e simultaneamente preserva detalhes da imagem (CHEN; WANG, 1999). O algoritmo pode também ser utilizado em uma etapa anterior à rede, atuando diretamente sobre os pixels da imagem como um filtro não linear. As equações (3.36) e (3.37) apresentam o referido procedimento de suavização para aplicação diretamente sobre a imagem (CHEN; WANG, 1999).

$$I_{ij}^{t+1} = I_{ij}^t + \frac{\sum_{(m,n) \in N_{ij}(1)} (I_{mn}^t - I_{ij}^t) e^{-(k_1 \Phi(\tilde{\sigma}_{mn}^2, \theta_\sigma) + k_2 D_{mn}^t)}}{\sum_{(m,n) \in N_{ij}(1)} e^{-(k_1 \Phi(\tilde{\sigma}_{mn}^2, \theta_\sigma) + k_2 D_{mn}^t)}} e^{-k_1 \Phi(\tilde{\sigma}_{ij}^2, \theta_\sigma)} \quad (3.36)$$

$$D_{ij}^t = \frac{|I_{i-1,j}^t - I_{i+1,j}^t| + |I_{i,j-1}^t - I_{i,j+1}^t| + |I_{i-1,j-1}^t - I_{i+1,j+1}^t| + |I_{i-1,j+1}^t - I_{i+1,j-1}^t|}{4} \quad (3.37)$$

$$\Phi(\tilde{\sigma}_{ij}^2, \theta_\sigma) = \begin{cases} \tilde{\sigma}_{ij}^2 & , \tilde{\sigma}_{ij}^2 \geq \theta_\sigma \\ 0 & , \tilde{\sigma}_{ij}^2 < \theta_\sigma \end{cases} \quad (3.38)$$

onde:

I_{ij}^t é a intensidade do pixel ij na iteração t ;

D_{ij}^t é a uma medida das discontinuidades locais em torno de um pixel;

$\tilde{\sigma}_{ij}^2(R)$ é a variância normalizada das intensidades dos pixels calculada em uma vizinhança de raio R do pixel ij ;

θ_σ é um limiar para a influência da variância e

k_1 e k_2 são parâmetros de ponderação.

A Figura 3.17(a) apresenta uma imagem em tons de cinza (8 bits/pixel) contendo regiões com fronteiras bem definidas, a qual é corrompida por ruído gaussiano com desvio padrão 15,97 ($\sigma^2=255$). A simples aplicação de um filtro passa-baixas do tipo de média (GONZALES; WOODS, 2000) na tentativa de reduzir o ruído resulta na degradação de detalhes da imagem, tais como as bordas que limitam as diversas regiões.

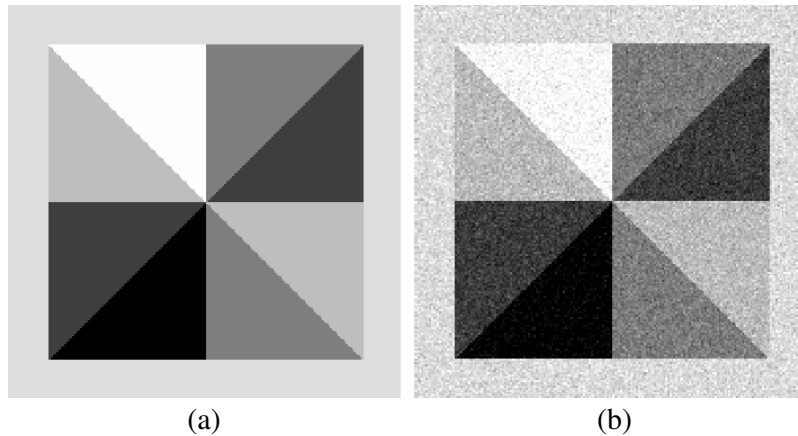


Figura 3.17: (a) Imagem original e (b) imagem corrompida por ruído

A Figura 3.18 apresenta resultados da aplicação do algoritmo FPS na imagem da Figura 3.17(b) com $R=2$, $\theta_\sigma=0,5$, $k_1=50$ e $k_2=50$. O ruído sofre atenuação significativa já desde a primeira iteração, não sendo observada degradação significativa das bordas da imagem original. Após um determinado número de iterações, a imagem resultante do processo de filtragem não apresenta mais modificações, comprovando a inexistência do “problema da terminação” (Figura 3.18(c) e Figura 3.18(d)).

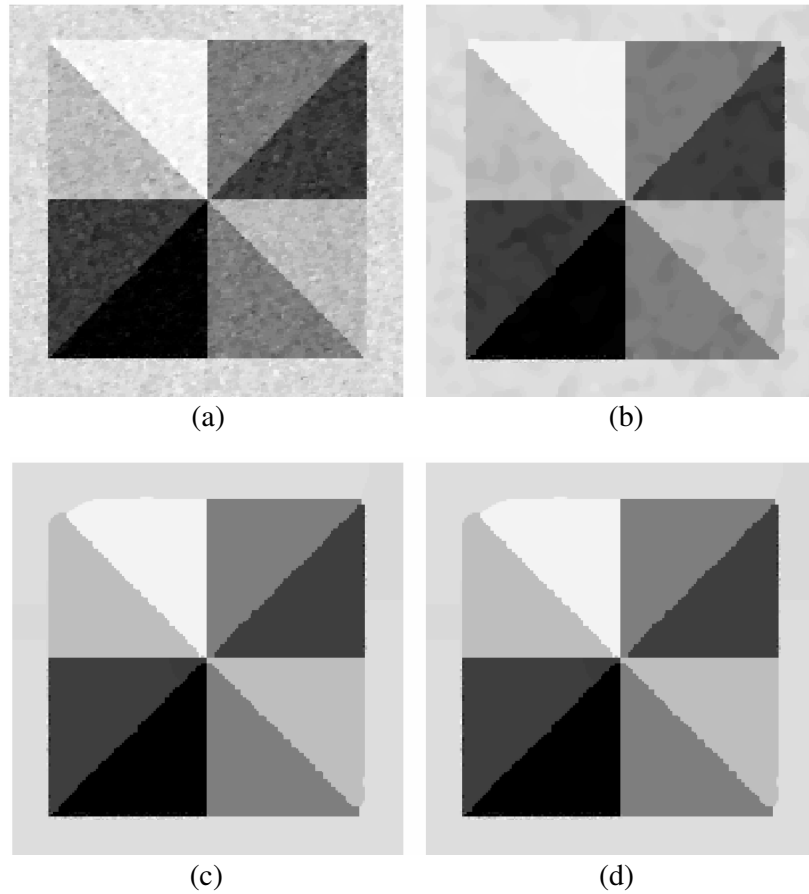


Figura 3.18: Imagens filtradas com o algoritmo FPS: (a) 1 iteração, (b) 10 iterações, (c) 1000 iterações e (d) 10000 iterações

A Figura 3.19 apresenta a imagem obtida após a realização de 100 iterações do algoritmo FPS utilizando os parâmetros $R=2$, $\theta_o=0,5$, $k_1=50$ e $k_2=50$ sobre a imagem da Figura 3.17(a), não degradada por ruído. Pode ser observado que pouca distorção ocorreu na imagem resultante, comprovando a característica do algoritmo FPS de preservar detalhes da imagem original.



Figura 3.19: Imagem original filtrada com 100 iterações do algoritmo FPS

3.5 Aplicações Práticas Desenvolvidas

Com o objetivo de melhor avaliar os resultados obtidos com a versão algorítmica da rede LEGION na segmentação de imagens naturais, foram desenvolvidas aplicações originais em problemas práticos de segmentação de imagens. Tais problemas de segmentação de imagens dizem respeito à identificação automática de caracteres em placas de automóveis, à contagem e medição de objetos (bolhas de gás hélio) em imagens de microscopia eletrônica e à segmentação e medição de estruturas cardíacas em imagens de ecocardiografia fetal.

3.5.1 Segmentação de Caracteres de Placas de Automóveis

Procedimentos de identificação automática de placas de automóveis podem ser baseados em um reconhecedor de caracteres (OCR), o qual, para funcionamento adequado, deve receber em sua entrada os caracteres devidamente isolados. A tarefa de segmentação e separação temporal dos caracteres da placa pode ser realizada com a utilização de uma rede LEGION.

A Figura 3.20 apresenta uma imagem de uma placa de automóvel, com dimensões de 224x608 pixels e resolução de 8 bits em escala de níveis de cinza, a qual foi obtida com uma câmera fotográfica digital. Para identificação posterior do número da placa, foi utilizada uma rede LEGION com o objetivo de segmentar os 4 algarismos e as 3 letras principais que identificam a mesma. A Figura 3.21 apresenta os 12 segmentos obtidos pela rede LEGION com 224x608 neurônios e parâmetros $\theta_p=5$ e $W_z=0,1$ e utilização do critério da máxima excitação lateral. O segmento apresentado no canto superior esquerdo representa o fundo relativamente homogêneo. Todos os 7 caracteres foram adequadamente segmentados, facilitando a atuação posterior de um identificador de caracteres. Os demais segmentos obtidos correspondem ao lacre da placa (cor branca na imagem original) e às regiões interiores dos algarismos 0 e 8, os quais não são conectados com o restante do fundo. Pode ser também observado que os caracteres referentes à identificação do estado e da cidade não foram segmentados pois correspondem a regiões muito pequenas e ruidosas, não tendo gerado osciladores líderes para promover sua ativação. Em todos os segmentos apresentados na Figura 3.21, a cor branca corresponde ao osciladores ativos, enquanto que a cor preta corresponde aos osciladores com saídas na fase silenciosa.



Figura 3.20: Imagem de uma placa de automóvel obtida com uma câmera fotográfica digital

Os resultados obtidos indicam que a rede LEGION pode ser empregada com sucesso neste tipo de aplicação, sendo, para tanto, necessário o ajuste criterioso de seus parâmetros, a escolha do procedimento para determinação da excitação lateral e, em alguns casos, a aplicação de algum procedimento para reduzir distorções ou ruídos excessivos na imagem a ser segmentada.

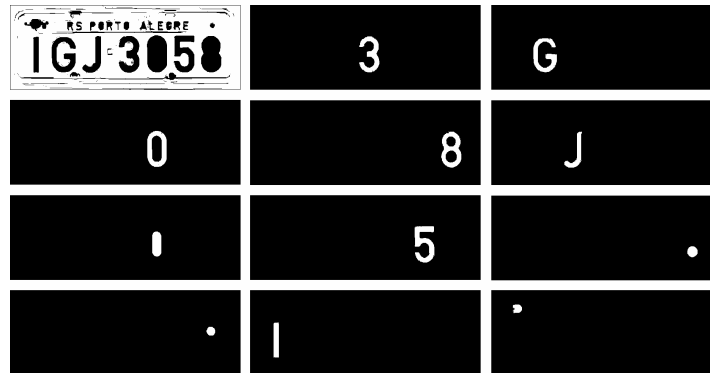


Figura 3.21: Segmentos obtidos com a aplicação da imagem da Figura 3.20 a uma rede LEGION

3.5.2 Segmentação de Imagens de Microscopia Eletrônica

O Instituto de Física da UFRGS vem realizando experimentos que implicam a análise de imagens de microscopia eletrônica relativas à implantação de hélio em amostras de silício com o objetivo de reduzir defeitos na estrutura cristalina do semiconductor. Devido às suas propriedades físicas, o hélio tende a acumular-se na forma de pequenas bolhas concentradas em uma determinada profundidade da amostra de silício. Para avaliação do processo, cujos resultados dependem de fatores como a temperatura, por exemplo, é necessário determinar o volume de gás presente em uma determinada região, o que pode ser estimado através da quantidade e do diâmetro das bolhas presentes em imagens de microscopia eletrônica (da SILVA et al., 2001). Este procedimento vem sendo realizado para diversas imagens, tratando-se de uma tarefa exaustiva para realização manual. A seguir são apresentados alguns resultados da aplicação da rede LEGION na proposta de implementação de um sistema semi-automático para realização das medidas necessárias (FERNANDES; NAVAU; FICHTNER, 2001-c).

A Figura 3.22 apresenta uma imagem com dimensões de 1482x2060 pixels e resolução de 8 bits em escala de níveis de cinza, a qual foi obtida com um microscópio eletrônico de transmissão (TEM), representando uma amostra de silício na qual foram implantados íons de hélio.

Para segmentar as bolhas de hélio e separar temporalmente as mesmas, facilitando a implementação de um processo semi-automático de contagem e medida do diâmetro de cada uma, optou-se pela utilização de uma rede LEGION (versão algorítmica) com as mesmas dimensões da imagem (1482x2060 osciladores). A Figura 3.23 apresenta uma região extraída da imagem da Figura 3.22, a qual é utilizada para exemplificar os resultados do procedimento de segmentação.

A Figura 3.24 apresenta o resultado da filtragem da imagem da Figura 3.23 com o algoritmo FPS (CHEN; WANG, 1999), podendo ser observada a preservação dos contornos das bolhas de hélio.

Os segmentos relativos às 9 bolhas segmentadas pela rede LEGION, utilizando $\theta_p=4$ e $W_z=0,4$ podem ser visualizados na Figura 3.25. O fundo preto representa os osciladores na região silenciosa (inativos), enquanto que as regiões brancas representam os osciladores com saída na fase ativa. Um segmento adicional, apresentado na Figura 3.26, referente ao fundo de silício, também é fornecido pela rede, o qual, devido à área muito maior do que os segmentos relevantes, pode ser facilmente detectado e descartado em um procedimento automatizado.

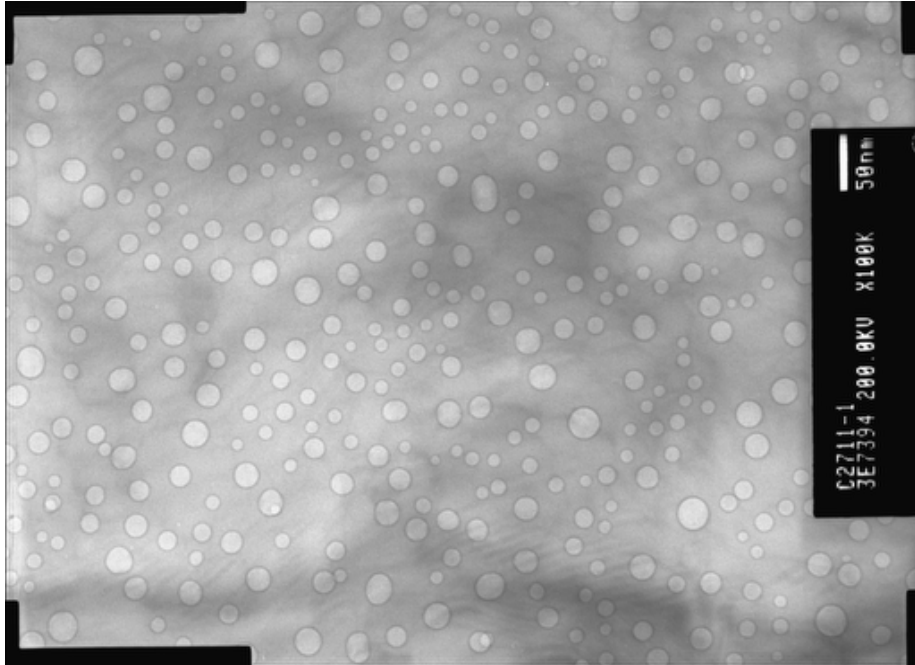


Figura 3.22: Imagem captada de um TEM representando uma amostra de silício com bolhas de hélio

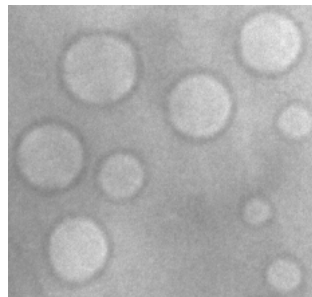


Figura 3.23: Região extraída da imagem da Figura 3.22

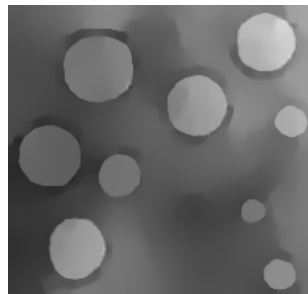


Figura 3.24: Imagem da Figura 3.23 filtrada com o algoritmo FPS

Os pixels pertencentes a regiões não homogêneas da imagem podem não ser incorporados a segmentos extraídos pela rede. A Figura 3.27 apresenta os pixels que não foram segmentados (representados em preto), que, em sua maioria, correspondem às regiões próximas às bordas das bolhas.

As áreas relativas de cada uma das bolhas podem ser rapidamente obtidas a partir dos segmentos apresentados na Figura 3.25, como a razão entre o número de osciladores com saída na fase ativa e o número total de osciladores da rede, o qual corresponde ao número de pixels da imagem. Sendo conhecida a relação pixel/nm da imagem obtida do TEM (referência presente na Figura 3.22), o volume de hélio na região representada pela imagem pode ser então estimado. Deve ser observado que a determinação visual da área de cada bolha (ou do diâmetro médio) pode ser imprecisa nos casos nos quais as bolhas não são aproximadamente circulares. Tal limitação não é observada no método proposto, já que a área relativa de cada segmento obtido pode ser determinada com maior exatidão, independentemente da forma do mesmo.

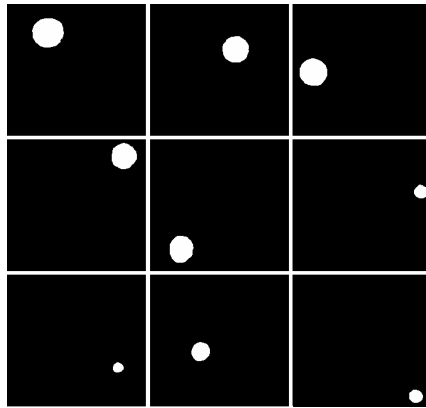


Figura 3.25: Resultado da segmentação da imagem da Figura 3.24 com uma rede LEGION

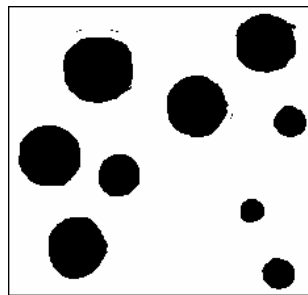


Figura 3.26: Segmento relativo ao fundo de silício homogêneo

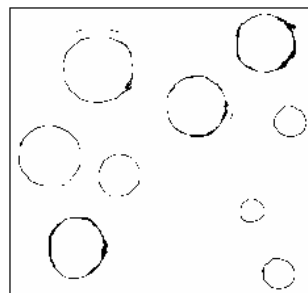


Figura 3.27: Pixels não segmentados pela rede (representados em preto)

Bolhas incompletas, presentes nos extremos das imagens, bolhas com contornos degradados e bolhas sobrepostas conduzem a medidas erradas. Para contornar tal limitação, foi adotado um procedimento interativo original, onde são marcados pontos sobre as bolhas a serem segmentadas. Tais pontos são utilizados para determinar os osciladores líderes da rede LEGION. Desta forma, somente as regiões homogêneas em torno dos líderes indicados serão segmentadas pela rede. As bolhas não marcadas, em número reduzido, podem ser posteriormente medidas manualmente ou através de um procedimento específico.

Como conclusão da aplicação desenvolvida foi verificado que um aspecto relevante diz respeito, assim como na versão contínua da rede LEGION, à determinação dos parâmetros, θ_p e W_z , os quais devem ser interativamente ajustados para a correta operação da rede. A qualidade da segmentação fornecida pela rede também está intimamente relacionada com a qualidade da imagem a ser segmentada, conduzindo a necessidade de aplicação de um procedimento anterior de suavização, tal como o algoritmo de filtragem FPS apresentado.

3.5.3 Segmentação de Imagens de Ecocardiografia Fetal

No caso de imagens de ecocardiografia fetal, a segmentação pode auxiliar na detecção e na diagnose de problemas cardíacos congênitos, enfatizando estruturas relevantes e conduzindo a avaliações mais confiáveis em exames pré-natais. Medidas características tais como a espessura do septo interventricular e volumes ou áreas dos ventrículos, podem ser utilizadas para identificar algumas anomalias, sendo, portanto, importante sua obtenção de maneira rápida e precisa.

Em (FERNANDES; SIQUEIRA; NAVAU, 2001-b) é apresentado um procedimento para extração e medida da região do ventrículo esquerdo (VE) em imagens de ecocardiografia fetal, baseado em um SOM pós-processado com a utilização do algoritmo *K-Means* (PANDYA; MACY, 1996) e em uma rede LEGION com indicação dos osciladores líderes pelo usuário. O procedimento SOM/*K-Means* realiza a classificação dos pixels como tecido ou cavidade (SIQUEIRA et al., 2000), enquanto que a rede LEGION extrai a região de interesse para realização de medidas. Devido à existência de pequenas conexões entre estruturas distintas em algumas imagens, é realizada uma operação de abertura morfológica (GONZALES; WOODS, 2000) na imagem submetida à rede LEGION, minimizando satisfatoriamente o problema.

A Figura 3.28(a) apresenta uma imagem onde foi acrescentada uma marca sobre o VE com o objetivo de extrair somente aquele segmento. A Figura 3.28(b) apresenta o resultado da segmentação utilizando o procedimento SOM/*K-Means*. Já a Figura 3.28(c) apresenta a saída da rede LEGION, podendo ser observado que a extração da região do VE não foi realizada com sucesso devido à uma pequena conexão existente entre o mesmo e o átrio esquerdo, resultado da baixa qualidade da imagem.

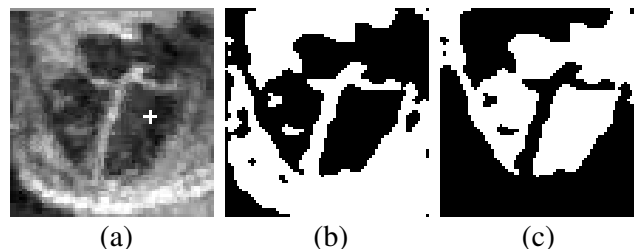


Figura 3.28: Segmentação incorreta do VE. (a) Imagem original com indicação da região a ser segmentada. (b) Segmentação SOM/*K-Means*. (c) Saída da rede LEGION

Para eliminar as pequenas conexões entre estruturas foi utilizado uma operação de abertura morfológica sobre a imagem obtida do procedimento *SOM/K-Means*, utilizando um elemento estruturante unitário com dimensões 3x3 pixels. O resultado da aplicação da operação proposta é apresentado na Figura 3.29, podendo ser observada a correta separação da região do VE.

A imagem obtida na Figura 3.29(c) possibilita a estimativa imediata da área do VE, sendo a mesma proporcional à razão entre o total de osciladores ativos (área branca) e o número total de osciladores da rede LEGION (número de pixels da imagem). No caso apresentado, a área do VE é equivalente a 11,49% ou 612 pixels², para uma rede 74x72 osciladores. De posse da escala de referência (mm/pixel) fornecida pelo equipamento de ultra-sonografia pode então ser determinada a área real do VE em mm².

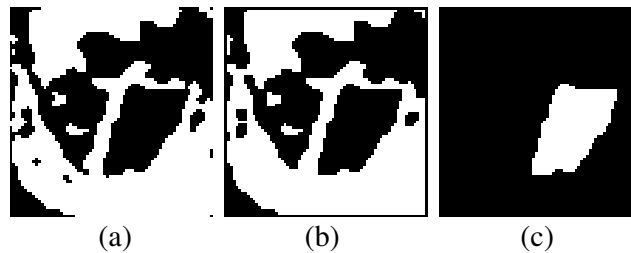


Figura 3.29: Segmentação correta do VE. (a) Imagem obtida com a segmentação *SOM/K-Means*. (b) Resultado da operação de abertura morfológica. (c) Saída da rede LEGION

Como conclusão do trabalho descrito verificou-se que os segmentos obtidos na saída da rede LEGION preservaram adequadamente a forma e o tamanho do VE, possibilitando a rápida estimativa da área do mesmo. Os resultados foram considerados satisfatórios por médicos especialistas. Outro aspecto relevante diz respeito à indicação manual dos líderes na rede LEGION, induzindo a mesma a extrair somente os segmentos indicados pelo usuário, a qual é uma solução original não encontrada na bibliografia e com relevância em aplicações práticas da rede.

3.6 Conclusão

Como aspectos positivos relacionados com a rede LEGION, versão contínua, ressaltam-se sua peculiar habilidade de separar os segmentos da imagem temporalmente e sua topologia maciçamente paralela, a que faz da mesma uma candidata natural à implementação de *chips* de visão. Por outro lado, a complexidade computacional para implementação em *hardware* digital, a capacidade de segmentação limitada e o elevado número de parâmetros, com procedimento de ajuste pouco prático, apresentam-se como fatores restritivos para aplicações práticas.

O oscilador de Terman-Wang, utilizado como elemento de processamento na rede LEGION, é baseado em um par de equações diferenciais, o que implica elevada complexidade computacional para o emprego em máquinas não paralelas. Um algoritmo para implementação da rede LEGION em computadores seriais foi proposto por Wang, tendo apresentado um ganho significativo de velocidade em relação à simulação da versão contínua da rede e ainda suplantando a limitação daquela versão quanto à capacidade de segmentação no que se refere ao número de objetos simultaneamente discriminados. Neste caso, entretanto, perde-se o atrativo da topologia maciçamente paralela apresentado pela versão contínua da rede.

Um dos pontos chave para a correta operação da rede LEGION, tanto em sua versão contínua quanto em sua versão algorítmica simplificada é a determinação dos pesos das conexões entre osciladores vizinhos. Os referidos pesos podem ser determinados no início do processo, em um passo único, com base em determinados atributos do sinal aplicado à entrada da rede, o que implica relativamente baixa complexidade computacional para desenvolver tal tarefa.

Algumas formas do algoritmo proposto para implementação da rede LEGION apresentam o fenômeno da percepção dúplex, o qual ocorre quando alguma região da imagem é percebida como sendo pertencente a mais de um objeto ou segmento diferente. Este fenômeno, apesar de ocorrer com observadores humanos, é contraditório com o princípio básico da segmentação de imagens, o qual estabelece que uma determinada região da imagem deve pertencer a um único objeto segmentado. Uma forma alternativa para o algoritmo da rede LEGION é apresentada, a qual não apresenta a possibilidade de percepção dúplex.

Resultados práticos demonstram que a degradação de imagens com a introdução de ruído é um fator limitante quanto ao funcionamento adequado dos algoritmos de segmentação, inclusive para a rede LEGION. Com objetivo de reduzir tais efeitos negativos, o algoritmo FPS, desenvolvido por Chen e Wang, apresenta como característica principal a propriedade de reduzir drasticamente o ruído da imagem sem degradar detalhes significativos da mesma, tais como as bordas dos objetos, apresentando também imunidade ao denominado “problema da terminação”. O algoritmo FPS pode ser utilizado isoladamente como um eficiente filtro não linear ou ainda como um procedimento de adaptação dos pesos da rede LEGION.

O trabalho desenvolvido por DeLiang Wang inclui aplicações da rede LEGION para segmentação de sinais de naturezas diversas, incluindo imagens médicas, imagens de sensoriamento remoto e sinais de fala, tendo sido relatados resultados bastante favoráveis. O emprego da rede LEGION em aplicações práticas implementadas durante o desenvolvimento deste trabalho também apresentou resultados animadores. Foram implementadas aplicações relativas à identificação automática de caracteres em placas de automóveis, contagem e medição de bolhas de hélio em imagens de microscopia eletrônica e segmentação e medição de estruturas cardíacas em imagens de ecocardiografia fetal.

Os resultados obtidos indicam que a rede LEGION pode ser empregada com sucesso em segmentação de imagens, sendo, para tanto, necessário o ajuste criterioso de seus parâmetros, os quais devem ser interativamente determinados para a correta operação da rede. A escolha cuidadosa do procedimento para determinação da excitação lateral também pode ser necessária em alguns casos. A qualidade da segmentação fornecida pela rede também está intimamente relacionada com a qualidade da imagem a ser segmentada, conduzindo, em alguns casos, à necessidade de aplicação de um procedimento anterior de suavização, tal como a filtragem FPS apresentada.

Outro aspecto importante diz respeito à implementação de um procedimento para indicação manual dos líderes na versão algorítmica da rede LEGION, induzindo a mesma a extrair somente os segmentos indicados pelo usuário. Este procedimento também é uma solução original proposta no presente trabalho e possui relevância em aplicações práticas onde não há interesse em segmentar todos os objetos presentes na imagem de entrada da rede.

4 ONNIS-GI: PROPOSTA DE UMA REDE NEURAL OSCILATÓRIA PARA SEGMENTAÇÃO DE IMAGENS

Este capítulo apresenta detalhadamente a proposta de uma nova rede neural, em duas versões, denominadas ONNIS (*Oscillatory Neural Network for Image Segmentation*) e ONNIS-GI (*Oscillatory Neural Network for Image Segmentation with Global Inhibition*) incluindo a concepção dos respectivos elementos de processamento e as topologias de conexões de cada uma delas. A segmentação da imagem de entrada com base em atributos diversos, tal como textura por exemplo, é abordada com o objetivo de ressaltar a flexibilidade apresentada pelos novos modelos propostos. A determinação dos pesos da rede com base em procedimentos de determinação de bordas também é abordada, tendo em vista a possibilidade de obter menor sensibilidade ao ruído na segmentação de imagens. Finalmente, também é apresentada uma versão algorítmica da rede ONNIS-GI, adequada à implementação da mesma em máquinas seriais.

4.1 Motivação

Os resultados encontrados na bibliografia, bem como as simulações e as aplicações práticas desenvolvidas, indicam que a rede LEGION apresenta-se como uma ferramenta de concepção recente, particularmente adequada para segmentação de imagens. Ressalta-se a sua peculiar capacidade de separar temporalmente os segmentos encontrados, o que facilita a posterior aplicação de procedimentos para identificação e quantificação dos mesmos. A versão contínua desta rede apresenta o forte atrativo de possuir uma topologia maciçamente paralela, o que possibilita implementações em *hardware* mais rápidas, assim como a construção de *chips* de visão analógicos com capacidade de segmentação de imagens. Por outro lado, sua estrutura baseada em equações diferenciais apresenta-se computacionalmente complexa para utilização de *hardware* digital, paralelo ou não. A necessidade de adição de um sinal aleatório às suas entradas, o elevado número de parâmetros e seu processo de ajuste conjunto pouco prático, assim como a capacidade de segmentação limitada, apresentam-se como fatores fortemente limitantes para aplicações práticas.

As características peculiares e atraentes da rede LEGION, bem como os bons resultados obtidos na segmentação de imagens, motivaram sua utilização como base para a concepção de uma nova rede neural com estrutura mais adequada para implementação de *chips* de visão utilizando circuitos digitais.

4.2 As Redes ONNIS e ONNIS-GI

No desenvolvimento da rede proposta, inicialmente em sua versão denominada ONNIS, incluindo a estrutura de seu elemento de processamento (EP), o qual simula um neurônio oscilador utilizando lógica digital, foram estabelecidas as seguintes premissas:

- A rede deve ser capaz de segmentar e separar no tempo todas as regiões homogêneas da imagem de entrada, segundo um critério de semelhança pré-estabelecido;
- A rede deve possuir arquitetura maciçamente paralela (um EP por pixel da imagem de entrada);
- As estruturas de conexões entre EPs devem priorizar a utilização de conexões locais entre os mesmos;
- A estrutura interna do EP deve ser a mais simples possível e facilmente realizável com dispositivos digitais;
- A rede deve ter o menor número possível de parâmetros a serem ajustados pelo usuário e
- A rede deve apresentar flexibilidade para implementação de procedimentos de segmentação baseados em atributos diversos dos pixels da imagem de entrada.

A capacidade das redes ONNIS e ONNIS-GI para segmentar e separar no tempo todas as regiões da imagem de entrada com características homogêneas (segmentos) implica que a mesma opere de forma similar à rede LEGION, porém, ao contrário desta, sem possuir qualquer limitação quanto ao número de segmentos a serem discriminados (capacidade de segmentação ilimitada).

A restrição imposta para implementação da rede utilizando arquitetura bidimensional maciçamente paralela vincula-se à idéia de utilizar a mesma para a realização de *chips* de visão para segmentação de imagens. Esta premissa implica que a rede proposta possua um número de EPs igual ao número de pixels da imagem a ser segmentada.

A prioridade na utilização de conexões locais tem como objetivo minimizar as rotas de conexão entre os EPs na implementação de um *chip* de visão. Rotas de conexão muito longas, volumosas e em grande quantidade geralmente oneram a implementação de um *chip*, digital ou não, podendo inclusive inviabilizar sua realização prática.

A simplicidade na concepção da estrutura interna dos EPs, tendo em mente uma posterior implementação em *hardware* digital dedicado, implica a utilização de operações e estruturas lógicas simples e facilmente implementáveis em dispositivos digitais. Deve ser ressaltado que, sendo a rede maciçamente paralela, um grande número de EPs (equivalente ao número de pixels da imagem de entrada) deverá ser implementado em um mesmo *chip*, enfatizando mais ainda o aspecto da simplicidade da estrutura interna dos mesmos.

A utilização de um número mínimo de parâmetros ajustáveis pelo usuário, e de fácil ajuste, deve conferir à rede maior autonomia e simplicidade para utilização em aplicações práticas.

Finalmente, tendo em vista que um procedimento de segmentação pode ser implementado com base em diferentes atributos ou características dos pixels da imagem, como intensidade, cor ou textura, por exemplo, a rede deve ser concebida de forma a apresentar flexibilidade para ser facilmente adaptada às diferentes aplicações.

4.2.1 Elemento de Processamento da Rede ONNIS

A filosofia que fundamentou a concepção do EP original da rede ONNIS consiste na criação de uma estrutura digital simples que imite o funcionamento característico de um neurônio oscilador da rede LEGION. Para atingir tal objetivo, foi associado inicialmente um contador binário a cada EP da rede, fazendo com que EPs vizinhos

submetidos a entradas semelhantes, as quais são atributos dos pixels da imagem de entrada, tenham seus contadores sincronizados em um mesmo estado, diferente dos estados dos demais grupos de EPs (Figura 4.1). O número de estados dos contadores, o qual é idêntico para todos EPs da rede, deve ser suficientemente grande para acomodar todos os diversos grupos de EPs correspondentes aos segmentos encontrados no sinal de entrada. Para uma aplicação geral, o número de estados dos contadores deve ser igual ao número total de EPs, equivalente, no caso de uma rede bidimensional, ao número de pixels da imagem a ser segmentada. As saídas dos EPs devem ser mantidas inativas até que os respectivos contadores atinjam um estado predefinido (N_n-1), ativando as saídas e habilitando o surgimento do respectivo segmento na saída da rede. No próximo pulso de *clock*, todos os contadores serão incrementados, implicando o *reset* dos EPs que estão na fase ativa, conduzindo seus contadores do estado N_n-1 para o estado 0, e passando o próximo grupo de EPs, correspondente a um outro segmento do sinal de entrada, para a fase ativa.

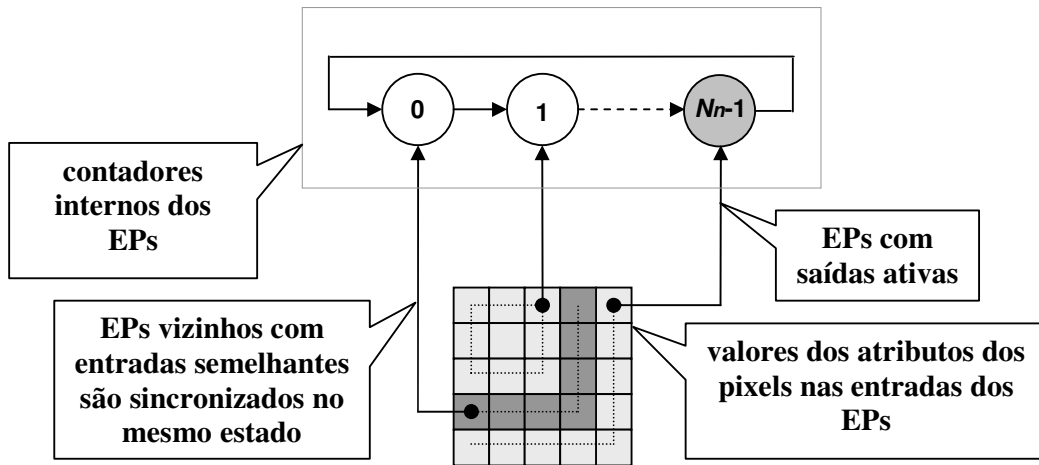


Figura 4.1. Filosofia utilizada na concepção do EP original da rede ONNIS

Com base nas premissas estabelecidas e na filosofia de operação representada pela Figura 4.1 foi concebida a estrutura representada pela Figura 4.2. Com o objetivo de possibilitar uma melhor compreensão do funcionamento do modelo do EP inicialmente proposto, as definições das constantes e variáveis utilizadas são relacionadas abaixo:

- N_x : número de colunas em um arranjo de EPs bidimensional;
- N_y : número de linhas em um arranjo de EPs bidimensional;
- N_n : número total de EPs na rede (em uma rede bidimensional $N_n=N_xN_y$);
- N_s : parâmetro que limita o número máximo de segmentos que podem ser discriminados pela rede;
- L_w : limiar para determinação da semelhança entre as entradas de EPs vizinhos;
- $v_c(i,j,t)$: estado do contador interno do EP situado na linha i e coluna j no tempo (iteração) t ;
- $v_o(i,j,t)$: saída do comparador do EP;
- $v_e(i,j,t)$: saída excitatória do EP;
- $v_{io}(i,j,t)$: saída inibitória do EP;

- $v_{ii}(i,j,t)$: entrada inibitória do EP;
- $v_l(i,j,t)$: sinal gerado internamente para indicação de EP líder;
- $v_{li}(i,j,t)$: sinal para indicação de EP líder;
- $I_a(i,j,t)$: sinal de entrada do EP (no caso de uma imagem, pode representar um conjunto de atributos ou um único atributo tal como a intensidade do respectivo pixel, por exemplo) e
- $w(i,j,k,l,t)$: peso referente à comparação da entrada do EP situado na linha i e coluna j com a entrada do EP situado na linha k e coluna l , no tempo t .

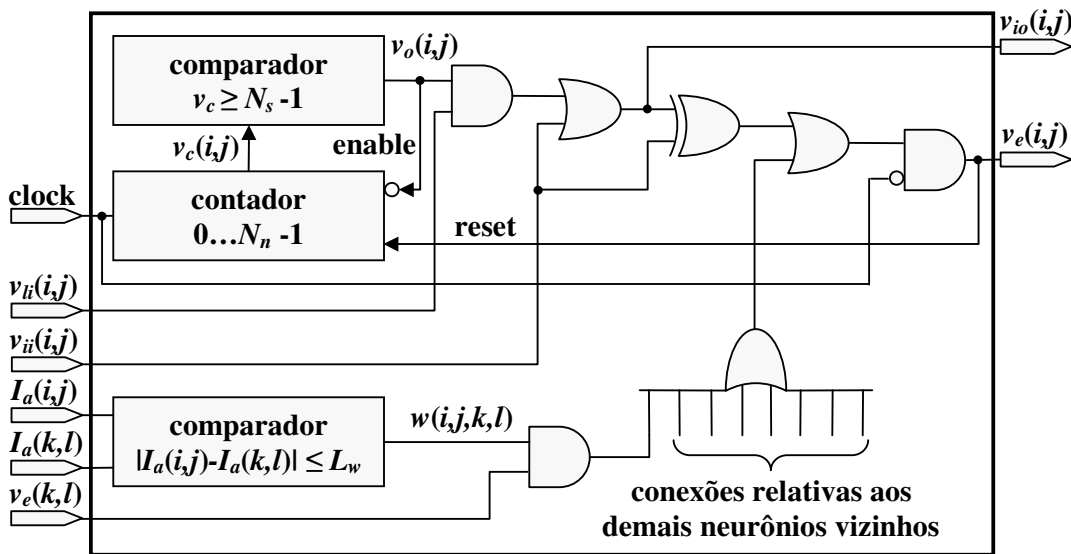


Figura 4.2: Estrutura interna do EP da rede ONNIS

Para uma aplicação genérica da rede ONNIS em segmentação de imagens, um contador com número de estados igual ao número de EPs da rede (N_n) é utilizado para definir o estado interno do EP em função do tempo ($v_c(i,j,t)$). O referido contador tem seu estado retornado a zero (*reset* síncrono) sempre que o EP estiver em sua fase ativa ($v_e(i,j,t)=1$), caso contrário o contador evolui sua contagem até atingir o estado N_n-1 , permanecendo aí até receber um *reset*. A utilização de N_n estados no contador possibilita a discriminação máxima de N_n segmentos (um estado para cada segmento), caso extremo no qual não existe semelhança entre qualquer par de entradas de EPs vizinhos. Como, de uma forma geral, o número de segmentos é bem menor do que o número de pixels de uma imagem, deverão surgir vários estados dos contadores sem qualquer segmento associado, implicando imagens nulas na saída da rede. Para minimizar tal problema, pode ser determinado um número máximo de segmentos à priori ($N_s \leq N_n$), limitando o número de estados dos contadores dos EPs. Como desvantagem desta técnica, pode não ocorrer a discriminação de alguns segmentos quando o número de segmentos presentes na imagem for superior a N_s (N_s subestimado). A equação (4.1) apresenta o comportamento da saída do contador interno do EP em função do tempo. A saída do contador interno do EP é enviada a um comparador cuja saída $v_o(i,j,t)$ é ativada somente quando sua entrada for igual ou superior ao número de segmentos N_s a ser discriminado pela rede menos um (4.2). A saída do comparador é utilizada para indicar internamente que o EP está em condições de passar para a fase ativa e também para desabilitar a contagem quando ($v_c(i,j,t) \geq N_s-1$).

$$v_c(i, j, t+1) = \begin{cases} v_c(i, j, t) + 1 & \text{se } v_c(i, j, t) < N_n - 1 \text{ e } v_e(i, j, t) = 0 \\ & \text{e } v_o(i, j, t) = 0 \\ v_c(i, j, t) & \text{se } v_c(i, j, t) = N_n - 1 \text{ e } v_e(i, j, t) = 0 \\ 0 & \text{se } v_e(i, j, t) = 1 \end{cases} \quad (4.1)$$

$$v_o(i, j, t) = \begin{cases} 1 & \text{se } v_c(i, j, t) \geq N_s - 1 \\ 0 & \text{se } v_c(i, j, t) < N_s - 1 \end{cases}, N_s \leq N_n \quad (4.2)$$

A equação (4.3) representa a saída inibitória do EP da rede ONNIS, $v_{io}(i, j, t)$, a qual é obtida através de uma operação lógica binária (booleana) entre a saída de seu comparador interno, $v_o(i, j, t)$, o sinal externo para indicação de líder, $v_{li}(i, j, t)$ e o sinal inibitório recebido de seu EP precedente, $v_{ii}(i, j, t)$. Assim, a saída inibitória do EP será ativada sempre que seu estado interno for igual ou superior a $N_s - 1$ e que o mesmo seja habilitado como um líder ($v_{li}(i, j, t) = 1$) ou ainda que o mesmo tenha recebido uma inibição de seu EP precedente. A posterior utilização de uma operação lógica *xor* entre a saída inibitória do EP e sua entrada inibitória implica a possibilidade do mesmo passar para a fase ativa ($v_e(i, j, t) = 1$) com base em seu estado interno ou recebendo excitação lateral de seus vizinhos.

$$v_{io}(i, j, t) = (v_o(i, j, t) \text{ and } v_{li}(i, j, t)) \text{ or } v_{ii}(i, j, t) \quad (4.3)$$

O sinal externo para indicação de EP líder, $v_{li}(i, j, t)$, é utilizado para qualificar inicialmente um EP a passar para a fase ativa. Somente os EPs com $v_{li}(i, j, t) = 1$ serão potencialmente habilitados a passar para a fase ativa. Os EPs com $v_{li}(i, j, t) = 0$ poderão ser ativados somente através de um EP vizinho com entrada semelhante e que esteja com sua saída na fase ativa.

A saída excitatória do EP, representada pela equação (4.4), será ativada em função do seu estado interno, da condição de líder e da inibição recebida de seu EP precedente ou se o mesmo receber um sinal excitatório de um EP vizinho cuja entrada seja semelhante à sua ($v_e(k, l, t) = 1$), implicando o respectivo peso $w(i, j, k, l, t) = 1$. Um EP pode receber sinais excitatórios de EPs diversos, pertencentes a uma região de vizinhança predefinida. A equação (4.4) considera a conexão do EP com um único EP vizinho. No caso de utilizar-se uma vizinhança maior, como vizinhança 8, por exemplo, as saídas excitatórias dos EPs vizinhos combinadas com os respectivos pesos (operação lógica *and*) deverão ser combinados através de uma operação lógica *or*.

$$v_e(i, j, t) = \overline{clock} \text{ and } ((v_{ii}(i, j, t) \text{ xor } v_{io}(i, j, t)) \text{ or } (w(i, j, k, l, t) \text{ and } v_e(k, l, t))) \quad (4.4)$$

O sinal de *clock* global utilizado, além de prover uma referência necessária para o sincronismo das transições de todos os contadores dos EPs da rede, é utilizado para “resetar” a saída da mesma, evitando que os EPs permaneçam com suas saídas indefinidamente ativadas. Assim, conforme pode ser observado através da equação (4.4), a saída excitatória de um EP da rede somente poderá passar para a sua fase ativa durante o tempo no qual o sinal de *clock* estiver em nível lógico baixo. Como consequência deste fato, os diversos segmentos da imagem de entrada somente aparecerão na saída da rede nestes períodos de tempo.

Os denominados pesos da rede são determinados através da comparação das intensidades dos atributos de entrada $I_a(i,j,t)$ do EP com as respectivas intensidades dos atributos de entrada $I_a(k,l,t)$ de cada EP vizinho. Nos casos onde a diferença destas entradas situar-se, em módulo, abaixo de um limiar prefixado (L_w) o respectivo peso será unitário, indicando semelhança entre os pixels da imagem. Se os atributos vinculados aos pixels da imagem de entrada $I_a(i,j,t)$ forem escalares, a determinação dos pesos pode ser realizada através de operações de implementação simples, como representado através da equação (4.5).

$$w(i, j, k, l, t) = \begin{cases} 0 & \text{se } |I_a(i, j, t) - I_a(k, l, t)| > L_w \\ 1 & \text{se } |I_a(i, j, t) - I_a(k, l, t)| \leq L_w \end{cases} \quad (4.5)$$

4.2.2 Estruturas de Conexões Excitatórias e Inibitórias

O modelo de EP proposto pode ser utilizado para formar redes de uma ou mais dimensões com capacidade de segmentar os sinais de entrada e separar no tempo os respectivos segmentos encontrados. Para segmentação de imagens, é proposta uma topologia bidimensional com as mesmas dimensões da imagem a ser segmentada, implicando correspondência direta entre os pixels da imagem e os EPs da rede.

A rede ONNIS apresenta duas estruturas de conexões locais entre EPs, uma delas denominada estrutura de conexões excitatórias e outra denominada estrutura de conexões inibitórias. Sob o ponto de vista de uma rede neural, estas estruturas realizam as conexões excitatórias e inibitórias entre os diversos neurônios.

A Figura 4.3 apresenta um exemplo da estrutura das conexões excitatórias proposta para uma pequena rede de 4x4 EPs, correspondendo a uma imagem de dimensões 4x4 pixels. Considerando a restrição de serem utilizadas somente conexões locais, da mesma forma que na rede LEGION foi adotada uma relação de vizinhança onde um EP é conectado simultaneamente aos 8 EPs mais próximos. Desta forma, a saída excitatória de um EP estará incondicionalmente em sua fase ativa quando pelo menos um de seus EPs vizinhos mais próximos, submetido a uma entrada semelhante ($w(i,j,k,l)=1$), estiver na fase ativa, conforme representado pela equação (4.6). Com relação aos EPs situados nas fronteiras da rede, os pesos relacionados com as conexões inexistentes podem ser considerados nulos, assim como suas respectivas entradas excitatórias.

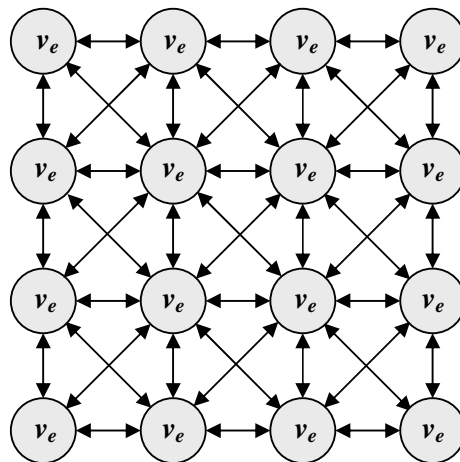


Figura 4.3: Exemplo da estrutura de conexões excitatórias

$$\begin{aligned}
 v_e(i, j, t) = & \overline{clock} \text{ and } ((v_{ii}(i, j, t) \text{ xor } v_{io}(i, j, t)) \text{ or} \\
 & (w(i-1, j-1, t) \text{ and } v_e(i-1, j-1, t)) \text{ or} \\
 & (w(i-1, j, t) \text{ and } v_e(i-1, j, t)) \text{ or} \\
 & (w(i-1, j+1, t) \text{ and } v_e(i-1, j+1, t)) \text{ or} \\
 & (w(i, j+1, t) \text{ and } v_e(i, j+1, t)) \text{ or} \\
 & (w(i+1, j+1, t) \text{ and } v_e(i+1, j+1, t)) \text{ or} \\
 & (w(i+1, j, t) \text{ and } v_e(i+1, j, t)) \text{ or} \\
 & (w(i+1, j-1, t) \text{ and } v_e(i+1, j-1, t)) \text{ or} \\
 & (w(i, j-1, t) \text{ and } v_e(i, j-1, t)))
 \end{aligned} \tag{4.6}$$

Na concepção da estrutura de conexões excitatórias apresentada, diferentes relações de vizinhança poderiam ser adotadas. A utilização da vizinhança 8 prende-se ao fato de que a mesma implica somente conexões locais entre EPs. A utilização de vizinhanças maiores pode conduzir a um melhor resultado na segmentação de imagens, ao custo de uma maior complexidade da estrutura do EP, de um maior número de conexões e de conexões mais longas, implicando maior custo para implementação da rede em *hardware* dedicado.

A Figura 4.4 apresenta um exemplo da estrutura das conexões inibitórias adotada, também para uma rede de 4x4 EPs. Esta estrutura estabelece uma prioridade entre os EPs da rede de tal forma que, no caso de vários EPs, pertencentes a segmentos diferentes, capacitarem-se a passar para a fase ativa, somente um deles, o de maior prioridade o fará. Em outras palavras, o EP com maior prioridade inibe os restantes, mantendo-os inativos. Por outro lado, as conexões excitatórias farão com que o único EP habilitado pela estrutura de conexões inibitórias habilite também todos os EPs que compõem o respectivo segmento através das respectivas conexões excitatórias com peso unitário. Diferentes estruturas de conexões inibitórias poderiam ser utilizadas tais como uma estrutura helicoidal iniciando em um extremo ou no centro da rede, por exemplo. Em termos operacionais, a diferença na utilização de outras estruturas de conexões inibitórias está vinculada basicamente à ordem com a qual os segmentos serão apresentados na saída da rede.

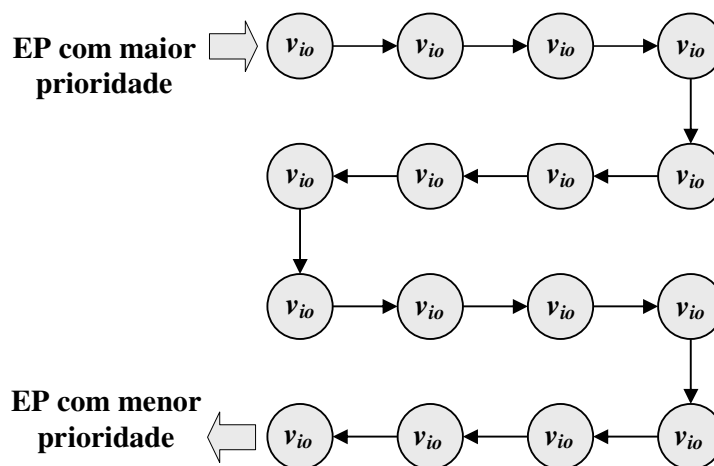


Figura 4.4: Exemplo da estrutura de conexões inibitórias

Na topologia de conexões inibitórias apresentada, o EP com prioridade mais alta recebe sempre inibição nula em sua entrada. Já o EP com menor prioridade fornece em sua saída inibitória um sinal binário que pode ser utilizado como um indicador de presença de um segmento na saída da rede. Em outras palavras, a saída inibitória do EP com menor prioridade será unitária somente quando pelo menos um EP estiver com sua saída na fase ativa. Em uma implementação prática, este sinal pode ser utilizado para desprezar imagens nulas, sem qualquer segmento na saída da rede, ou para transferir a imagem de saída para outro dispositivo ou etapa de processamento posterior.

4.2.3 EPs Líderes

A determinação e utilização dos denominados neurônios osciladores líderes na rede LEGION implica a não discriminação de segmentos muito ruidosos pela rede, melhorando a qualidade da segmentação. De forma semelhante, a idéia pode ser aplicada na rede ONNIS com o intuito de melhorar a segmentação.

Na geração externa do sinal utilizado para qualificar um EP como líder, $v_{li}(i,j,t)$, o mesmo pode ser estabelecido por algum critério tal como a posição correspondente na imagem, por exemplo. Assim, poderia ser estabelecido que somente EPs correspondentes à região central da imagem, por exemplo, ou a uma região indicada pelo usuário poderiam passar à fase ativa, implicando somente o aparecimento daquele segmento na saída da rede. Na geração interna do sinal pode ser utilizado o critério de que um EP líder deve possuir todos os pesos referentes às conexões excitatórias unitários, o que corresponde a um pixel no centro de uma região homogênea. Tal condição pode ser facilmente realizada através de uma operação *and* entre os pesos excitatórios determinados.

A Figura 4.5 apresenta a estrutura adotada para determinação dos EPs líderes na rede proposta. O sinal $v_{li}(i,j,t)$ é obtido através de uma operação lógica *and* entre todos os pesos vinculados à vizinhança de um determinado EP (vizinhança 8, no caso) e um sinal de habilitação externo, $v_l(i,j,t)$, conforme representado pela equação (4.7). O sinal de habilitação externo pode ser definido de forma interativa pelo usuário da rede, possibilitando que a mesma discrimine somente determinados segmentos de interesse.

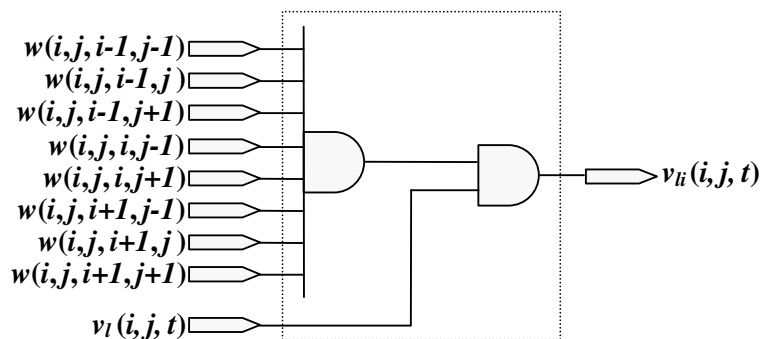


Figura 4.5: Estrutura utilizada para determinação dos EPs líderes

$$v_{li}(i, j, t) = v_l(i, j, t) \text{ and } w(i-1, j-1, t) \text{ and } w(i-1, j, t) \text{ and} \\ w(i-1, j+1, t) \text{ and } w(i, j+1, t) \text{ and} \\ w(i+1, j+1, t) \text{ and } w(i+1, j, t) \text{ and} \\ w(i+1, j-1, t) \text{ and } w(i, j-1, t) \quad (4.7)$$

4.2.4 Inicialização Global

Utilizando a estrutura do EP inicialmente proposta, com as respectivas conexões excitatórias e inibitórias, a rede apresenta ciclicamente N_s imagens em sua saída. Sendo o número de segmentos da imagem desconhecido a priori, uma estimativa do parâmetro N_s deve ser realizada de forma interativa pelo usuário. Sendo N_s menor do que o número de segmentos presentes na imagem, alguns segmentos não serão discriminados pela rede. Uma alternativa para esta limitação é estimar N_s suficientemente grande, o que, por outro lado, poderá resultar em um grande número de imagens sem qualquer segmento na saída da rede. Tendo em vista a possível inconveniência que este aspecto pode ter em aplicações práticas, é proposta uma modificação na rede inserindo-se um procedimento de inicialização global automático, o qual dispensa a estimativa do número de segmentos a priori por parte do usuário.

O denominado procedimento de inicialização global tem como objetivo produzir um *reset* simultâneo nos contadores de todos os EPs da rede na presença de uma imagem sem qualquer segmento na saída da mesma. Tal objetivo pode ser alcançado utilizando-se o sinal da saída inibitória do EP com menor prioridade da rede, $v_{io}(i_m, j_m, t)$, o qual é distribuído através de uma conexão global entre este EP e todos os demais EPs da rede, conforme observado na Figura 4.6. Os índices i_m e j_m indicam a posição do EP de menor prioridade na rede (linha e coluna respectivamente).

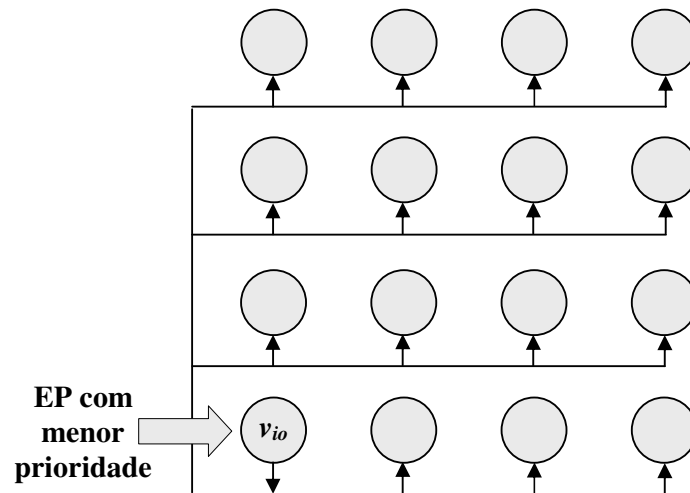


Figura 4.6: Exemplo da estrutura de conexões de inicialização global

Quando não está presente nenhum segmento na saída da rede, a saída inibitória do EP de menor prioridade é nula, determinando que todos os contadores internos dos EPs sejam inicializados no estado N_n-1 . Nesta situação, os EPs referentes ao segmento de maior prioridade serão ativados. Considerando $N_s=N_n$, no próximo ciclo de *clock* o segmento de maior prioridade será desativado e o próximo grupo de EPs será ativado. Assim, obedecendo a prioridade estabelecida pela estrutura de conexões inibitórias, todos os segmentos da imagem de entrada serão seqüencialmente apresentados na saída da rede até obter-se uma imagem nula, a qual reiniciará o ciclo. A versão da rede proposta que utiliza o mecanismo de inicialização global foi denominada ONNIS-GI.

4.2.5 Elemento de Processamento da Rede ONNIS-GI

Com a adoção do procedimento de inicialização global proposto podem ser utilizados somente dois estados para os contadores internos dos EPs. Tal fato permite a

substituição dos contadores por simples flip-flops, simplificando significativamente a estrutura interna do EP. Com relação a este aspecto, deve ser ressaltado que os contadores dos EPs devem ter, em uma implementação genérica, um número de estados igual ao número de pixels da imagem de entrada. Com a eliminação dos contadores e dos respectivos comparadores associados, a complexidade da estrutura do EP torna-se independente do número de pixels da imagem a ser segmentada, um aspecto também bastante positivo em termos práticos.

A Figura 4.7 apresenta uma proposta para a estrutura interna do EP convenientemente modificado para operar com o mecanismo de inicialização global apresentado. O contador de N_n estados e o comparador original foram substituídos por um simples flip-flop JK, o qual pode ser entendido como um contador de um bit, possibilitando somente dois estados internos para o EP.

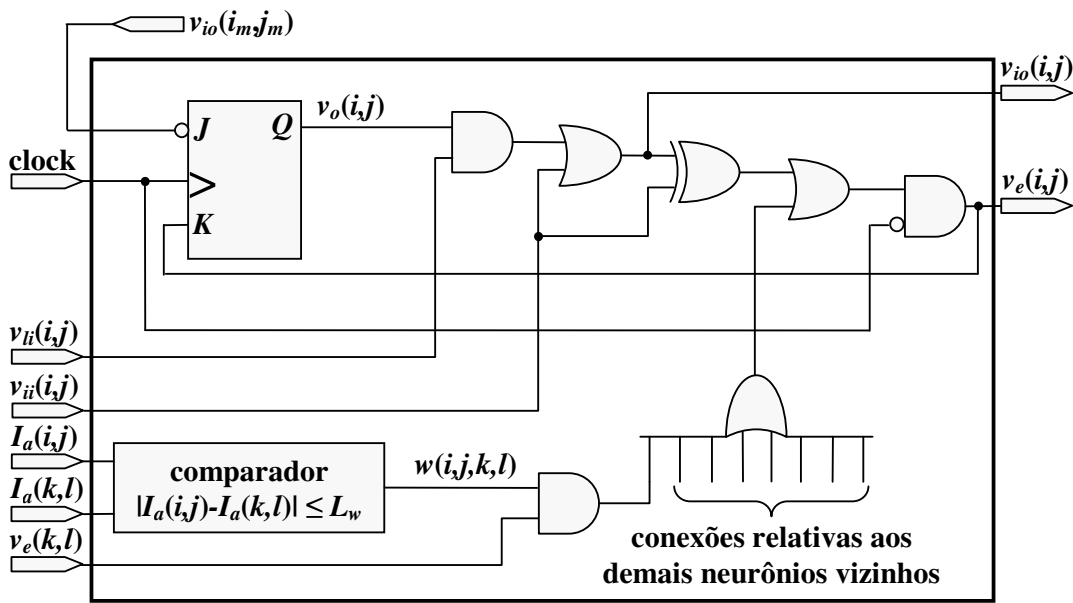


Figura 4.7: Estrutura interna do EP da rede ONNIS-GI

A modificação introduzida no EP da rede proposta com a utilização do procedimento de inicialização global implica a substituição das equações (4.1) e (4.2) pela equação (4.8).

$$v_o(i, j, t) = \begin{cases} 1 & \text{se } v_{io}(i_m, j_m, t-1) = 0 \text{ e } v_e(i, j, t-1) = 0 \\ 0 & \text{se } v_{io}(i_m, j_m, t-1) = 1 \text{ e } v_e(i, j, t-1) = 1 \\ v_o(i, j, t-1) & \text{se } v_{io}(i, j, t-1) = 1 \text{ e } v_e(i, j, t-1) = 0 \end{cases} \quad (4.8)$$

Em contrapartida à necessidade de utilização de uma conexão global entre o EP de menor prioridade e os demais EPs da rede, a utilização da estrutura de inicialização global proposta associada ao EP simplificado apresenta vantagens relevantes do ponto de vista prático. As diversas imagens sem segmentos são reduzidas a uma única imagem, a qual pode também ser utilizada para sincronizar a transferência das imagens de saída da rede para um estágio posterior de processamento. A substituição dos contadores e dos comparadores internos dos EPs por meros flip-flops JK possibilita uma simplificação significativa do *hardware*. Na implementação de um algoritmo baseado

no funcionamento da rede, a quantidade de memória necessária fica também reduzida, já que o estado interno de cada EP necessita de somente um bit para seu armazenamento. Finalmente, o número de segmentos da imagem de entrada não precisa ser estimado a priori, sendo agora implicitamente determinado pela rede.

Em virtude das vantagens apresentadas em relação à rede ONNIS, a rede ONNIS-GI foi adotada para obtenção dos resultados apresentados a partir deste ponto. Assim, a rede resultante utilizada apresenta um único parâmetro, L_w , a ser ajustado pelo usuário, já que não se faz mais necessária a utilização do parâmetro N_s , o qual deveria ser estimado a priori.

4.2.6 Algoritmo para Implementação da Rede ONNIS-GI

A arquitetura da rede proposta, em suas duas versões, implica que as mesmas sejam adequadas à implementação em *hardware* maciçamente paralelo, aproveitando assim a velocidade inerente a tais topologias. Entretanto, a implementação da proposta em máquinas seriais, mesmo podendo apresentar redução de velocidade significativa em relação à implementação paralela, pode ser útil se utilizada simplesmente como um algoritmo de segmentação de imagens. Assim, de maneira análoga ao realizado por Wang para a rede LEGION, foi desenvolvida uma versão algorítmica da rede ONNIS-GI para computadores seriais, a qual é apresentada a seguir:

1) Fase de inicialização

- Inicializar todos os flip-flops com suas saídas ativas:

$$v_o(i, j, 0) = 1 \quad (4.9)$$

- Fixar a entrada inibitória do EP de maior prioridade da rede:

$$v_{ii}(1, 1, t) = 0 \quad (4.10)$$

- Determinar os pesos da rede:

$$w(i, j, k, l, t) = \begin{cases} 0 & \text{se } |I_a(i, j, 0) - I_a(k, l, 0)| > L_w \\ 1 & \text{se } |I_a(i, j, 0) - I_a(k, l, 0)| \leq L_w \end{cases} \quad (4.11)$$

- Encontrar os EPs líderes:

$$\begin{aligned} v_{li}(i, j, t) = & w(i-1, j-1, 0) \text{ and } w(i-1, j, 0) \text{ and} \\ & w(i-1, j+1, 0) \text{ and } w(i, j+1, 0) \text{ and} \\ & w(i+1, j+1, 0) \text{ and } w(i+1, j, 0) \text{ and} \\ & w(i+1, j-1, 0) \text{ and } w(i, j-1, 0) \end{aligned} \quad (4.12)$$

2) Dinâmica da rede

- Determinar as saídas inibitórias dos EPs:

$$v_{io}(i, j, t) = (v_o(i, j, t) \text{ and } v_{li}(i, j, t)) \text{ or } v_{ii}(i, j, t) \quad (4.13)$$

- Determinar as saídas excitatórias dos EPs:

$$\begin{aligned}
 v_e(i, j, t) = & \overline{clock} \text{ and } ((v_{ii}(i, j, t) \text{ xor } v_{io}(i, j, t)) \text{ or} \\
 & (w(i-1, j-1, t) \text{ and } v_e(i-1, j-1, t)) \text{ or} \\
 & (w(i-1, j, t) \text{ and } v_e(i-1, j, t)) \text{ or} \\
 & (w(i-1, j+1, t) \text{ and } v_e(i-1, j+1, t)) \text{ or} \\
 & (w(i, j+1, t) \text{ and } v_e(i, j+1, t)) \text{ or} \\
 & (w(i+1, j+1, t) \text{ and } v_e(i+1, j+1, t)) \text{ or} \\
 & (w(i+1, j, t) \text{ and } v_e(i+1, j, t)) \text{ or} \\
 & (w(i+1, j-1, t) \text{ and } v_e(i+1, j-1, t)) \text{ or} \\
 & (w(i, j-1, t) \text{ and } v_e(i, j-1, t)))
 \end{aligned} \tag{4.14}$$

- Determinar as novas saídas dos flip-flops dos EPs:

$$v_o(i, j, t+1) = \begin{cases} 1 & \text{se } v_{io}(i_m, j_m, t) = 0 \text{ e } v_e(i, j, t) = 0 \\ 0 & \text{se } v_{io}(i_m, j_m, t) = 1 \text{ e } v_e(i, j, t) = 1 \\ v_o(i, j, t) & \text{se } v_{io}(i_m, j_m, t) = 1 \text{ e } v_e(i, j, t) = 0 \end{cases} \tag{4.15}$$

- Executar a dinâmica da rede até que todos os EPs tenham suas saídas excitatórias nulas. O número de segmentos encontrados corresponde ao número de iterações realizadas.

A equação (4.12) utilizada para determinação dos EPs líderes poderia ser substituída pela equação (4.7), a qual considera um sinal de habilitação externo, $v_l(i, j, t)$, o qual pode ser definido de forma interativa pelo usuário da rede, possibilitando que a mesma discrimine somente determinados segmentos de interesse.

O algoritmo desenvolvido para simular o comportamento da rede ONNIS-GI foi utilizado como base para o aplicativo apresentado na seção 5.2 e utilizado para obter os demais resultados apresentados ao longo deste trabalho.

4.3 Segmentação com Base em Atributos Diversos

A determinação dos pesos da rede ONNIS-GI com base na equação (4.5) implica que a mesma segmente a imagem com base na semelhança entre a intensidade das entradas de EPs vizinhos. De uma forma geral, qualquer atributo ou conjunto de atributos do sinal de entrada pode ser utilizado para determinação dos pesos da rede proposta, implicando somente uma modificação adequada na realização da respectiva etapa da estrutura. Diferentes vetores de atributos podem ser utilizados para diferenciar os diversos segmentos de uma imagem. A determinação adequada de tais atributos, por outro lado, é um problema de reconhecimento de padrões, o qual está intimamente relacionado com as características da imagem e o tipo de segmentação desejada.

4.3.1 Segmentação de Imagens em Níveis de Cinza

Imagens em níveis de cinza (*grayscale*) são normalmente representadas com a utilização de 8 bits/pixel, implicando a possibilidade de 256 níveis de cinza para cada pixel, sendo o nível 0 correspondente à cor preta e o nível 255 correspondente à cor

branca. A determinação dos pesos da rede para segmentar uma imagem em níveis de cinza requer apenas a comparação da diferença das intensidades de pixels vizinhos com um limiar estabelecido (L_w). Tal procedimento pode ser também adotado sempre que um pixel for representado por um único atributo na forma de um escalar.

A Figura 4.8 apresenta uma imagem em níveis de cinza com 135x75 pixels (a) e os quatro segmentos obtidos com a utilização da rede proposta (b). O segmento mais à esquerda corresponde ao fundo, os dois seguintes às letras B e R, enquanto que o segmento mais à direita equivale à parte interna da letra B. Os pixels brancos representam os EPs que têm suas saídas ativas. Para determinação dos pesos da rede foi utilizando limiar $L_w=0$.

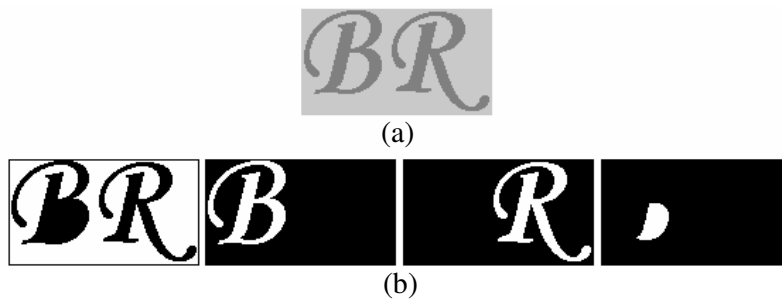


Figura 4.8: Segmentação de imagem em níveis de cinza utilizando a rede ONNIS-GI

4.3.2 Segmentação Utilizando Atributos de Cor

Quando as entradas da rede ONNIS-GI, $I_a(i,j,t)$, representam vetores de atributos relacionados com os pixels da imagem de entrada, os pesos podem ser determinados através da utilização de uma medida de distância vetorial, tal como a distância Euclidiana, por exemplo.

Para exemplificar a capacidade da rede em segmentar imagens coloridas utilizou-se uma versão em cores da imagem da Figura 4.8 (a). Neste caso, cada pixel da imagem foi representado por um vetor de três elementos correspondentes à intensidade de cada uma das componentes de cor (vermelho, verde e azul). Aos dois caracteres presentes na imagem foi atribuída a cor (125, 125, 0), enquanto que ao fundo foi atribuída a cor (0, 0, 250). A determinação dos pesos da rede proposta com base na intensidade de cada pixel, calculada como a média aritmética das intensidades de cada componente de cor, não permitiu a discriminação dos segmentos da imagem. Desta forma, a rede forneceu um único segmento em sua saída, considerando que o fundo e as letras pertencem a uma região homogênea. Já com a utilização da distância Euclidiana calculada no espaço das cores para determinar os pesos foi possível à discriminação correta dos segmentos, de forma idêntica ao resultado apresentado na Figura 4.8 (b).

4.3.3 Segmentação Utilizando Atributos de Textura

A utilização de vetores de atributos apropriados para determinação dos pesos da rede ONNIS-GI pode conduzir à segmentação da imagem de entrada com base nas texturas das diversas regiões da mesma.

A Figura 4.9 apresenta os resultados obtidos com a utilização da rede ONNIS-GI na segmentação de uma imagem composta por regiões com texturas diferentes. A imagem a ser segmentada (a), com dimensões de 40x40 pixels, é composta por um retângulo de cor branca posicionado na parte interna de uma região reticulada formada por pixels pretos e brancos, a qual é contornada por uma borda preta. Com a segmentação baseada na intensidade de cada pixel, foram obtidos os dois segmentos apresentados na Figura

4.9 (b). O primeiro segmento representa a borda preta juntamente com os quadrados pretos do reticulado, pois apresentam mesma intensidade e são conectados. O segundo segmento apresenta o retângulo branco e os quadrados brancos do reticulado, os quais também possuem mesma intensidade e são conectados.

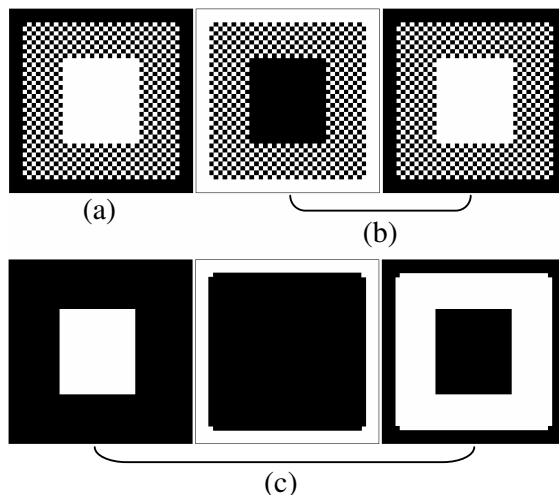


Figura 4.9: Utilização da rede ONNIS-GI para segmentação de uma imagem considerando um atributo de textura

Para conseguir discriminar a região do reticulado foi utilizado um atributo de textura espacial que considera todo pixel preto contornado por no mínimo 3 pixels brancos e todo pixel branco contornado por no máximo 5 pixels brancos como pertencentes ao reticulado. Estes pixels receberam valor um enquanto os demais, não pertencentes ao reticulado, receberam valor zero gerando uma nova imagem referente ao atributo de textura da imagem original. Os resultados da segmentação utilizando esta nova imagem são representados pelas imagens da Figura 4.9 (c), relacionadas, respectivamente, com o retângulo central, a borda homogênea e a região do reticulado. Todos os segmentos são representadas em branco (EPs com saídas ativas).

Deve ser observado que a simples aplicação do critério de pertinência de um pixel à região reticulada poderia ser utilizado diretamente para segmentar corretamente esta região. Por outro lado, não seria possível discriminar temporalmente a região homogênea central da região homogênea da borda, as quais representam objetos diferentes, o que é realizado de forma adequada pela rede ONNIS-GI.

4.3.4 Segmentação Utilizando Filtros de Gabor

Um conjunto adequado de filtros de Gabor pode ser utilizado para caracterizar regiões de imagens com texturas diferentes, gerando vetores de atributos associados a cada pixel. A determinação dos pesos da rede ONNIS-GI pode ser então realizada com base em uma medida de distância adequada entre os referidos vetores. Assim, vetores semelhantes devem implicar peso unitário, enquanto vetores não semelhantes devem implicar peso nulo.

A Figura 4.10 apresenta uma imagem com 200x200 pixels composta por quatro regiões com três tipos de texturas diferentes. A região superior esquerda possui uma textura formada por uma oscilação senoidal no nível de cinza dos pixels, entre 0 e 255, com período igual a 10 pixels na horizontal e zero (nível de cinza contínuo) na vertical. A região inferior direita apresenta oscilação semelhante, porém com zero pixels por

período na horizontal e 10 pixels por período na direção vertical. Já as regiões superior direita e inferior esquerda têm textura idênticas, formadas pela média das intensidades relativas às regiões superior esquerda e inferior direita.

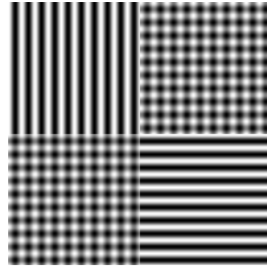


Figura 4.10: Imagem utilizada para exemplificar a segmentação utilizando filtros de Gabor e a rede ONNIS-GI

A Figura 4.11 apresenta o resultado obtido com a aplicação de filtros de Gabor sobre a imagem da Figura 4.10. A Figura 4.11 (a) foi obtida com os parâmetros $\alpha^2=200$, $\beta^2=200$, $u_0=0,1$ e $v_0=0$, correspondendo a intensidade do pixel ao módulo da amplitude da componente de oscilação horizontal presente na imagem naquele ponto. Já a Figura 4.11 (b) foi obtida com os parâmetros $\alpha^2=200$, $\beta^2=200$, $u_0=0$ e $v_0=0,1$, correspondendo à detecção da componente de oscilação vertical. Pode ser observado que nas regiões onde a textura corresponde à média das duas componentes de oscilação, a amplitude de cada uma delas é menor do que nas regiões onde cada uma aparece isoladamente, como era esperado.

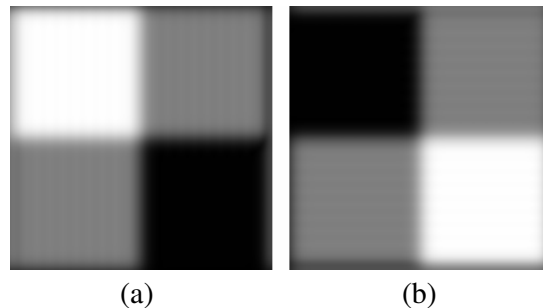


Figura 4.11: Filtragem da imagem da Figura 4.10 utilizando filtros de Gabor

Na utilização da rede ONNIS-GI para separar as regiões com diferentes texturas, os pesos são determinados com base nas imagens obtidas com a utilização dos filtros de Gabor. A Figura 4.12 apresenta a saída da rede através das imagens (a), (b) e (c), e os segmentos correspondentes da imagem de entrada através das imagens (e), (f) e (g). As imagens (e), (f) e (g) foram obtidas utilizando-se uma operação lógica *and* entre a saída de cada EP da rede e o valor da intensidade do respectivo pixel da imagem de entrada original. Pode ser verificado que a rede separou corretamente as regiões com texturas distintas da imagem. Devido à utilização de filtros de Gabor com tamanho constante, a amplitude de saída dos mesmos decresce próximo às regiões de fronteira entre diferentes texturas. Tal característica implica que os pixels pertencentes a estas regiões de fronteira não sejam incluídos pela rede nos segmentos detectados. Já na Figura 4.12 (a) pode ser observado que a rede não separou as duas regiões de mesma textura pois existe uma conexão física, ainda que mínima, entre as mesmas.

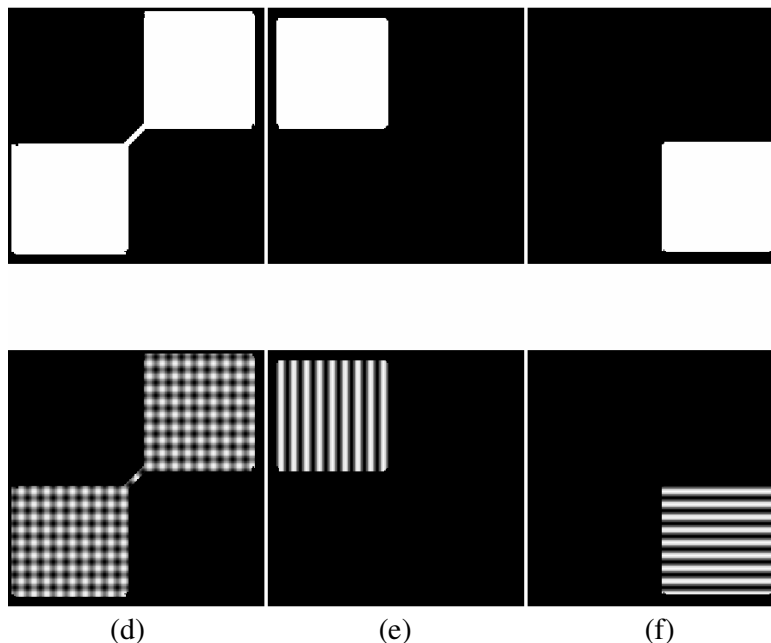


Figura 4.12: Segmentação da imagem da Figura 4.10 utilizando a rede ONNIS-GI com determinação de pesos baseada em filtros de Gabor

Apesar dos resultados preliminares apresentarem-se estimulantes, a utilização de filtros de Gabor para determinação dos pesos da rede ONNIS-GI apresenta algumas limitações. Uma delas diz respeito à elevada complexidade, o que deve certamente comprometer a aplicação destes filtros na solução de problemas de segmentação de imagens em tempo real. Por outro lado, a determinação a priori do número de filtros e de seus parâmetros para uma aplicação genérica apresenta-se como um procedimento ainda não convenientemente formalizado, conduzindo à necessidade de um elevado nível de interatividade entre o usuário e o processo. O assunto, entretanto, é promissor para o desenvolvimento de pesquisas futuras.

4.4 Determinação dos Pesos Através de Detecção de Bordas

A determinação dos pesos da rede proposta com base na diferença entre atributos de pixels adjacentes pode ser comparado a um procedimento de determinação de bordas. Se os atributos dos pixels adjacentes forem suficientemente diferentes, os mesmos pertencerão a segmentos diferentes e o peso assim determinado (com valor zero) corresponde a uma borda infinitesimal posicionada entre os dois pixels. Desta forma, o estudo e aplicação de algoritmos robustos de detecção de bordas que apresentem possibilidade de implementação em arquitetura maciçamente paralela pode trazer melhoria na qualidade de segmentação da rede proposta na presença de ruído na imagem de entrada.

4.4.1 Os Pesos da Rede como Bordas da Imagem

A Figura 4.13 apresenta uma imagem (a) e os respectivos pesos da rede ONNIS-GI determinados utilizando como base a semelhança entre as intensidades dos pixels adjacentes considerando $L_w=0$. Nas imagens que representam os 8 conjuntos de pesos (Figura 4.13 (b) até Figura 4.13 (i)), branco corresponde aos pesos unitários, enquanto preto corresponde aos pesos nulos. Tais imagens correspondem, respectivamente à

comparação da intensidade do pixel com índice (i,j) com seus 8 vizinhos de índices $(i-1, j-1)$, $(i-1, j)$, $(i-1, j+1)$, $(i, j+1)$, $(i+1, j+1)$, $(i+1, j)$, $(i+1, j-1)$, $(i, j-1)$. Pode ser claramente observada a semelhança entre os pesos nulos e as bordas dos segmentos da imagem original. As imagens da Figura 4.13 (c) e (g) podem ser relacionadas com as componentes horizontais das bordas, enquanto que as imagens (e) e (i) podem ser relacionadas com as componentes verticais. Já as imagens (b), (f), (d) e (h) podem ser relacionadas com componentes diagonais das bordas nas direções de 45° e -45° . Desta forma, os pesos das redes ONNIS e ONNIS-GI podem ser entendidos como bordas infinitesimais entre os segmentos, podendo ser calculados através de procedimentos de determinação de bordas.

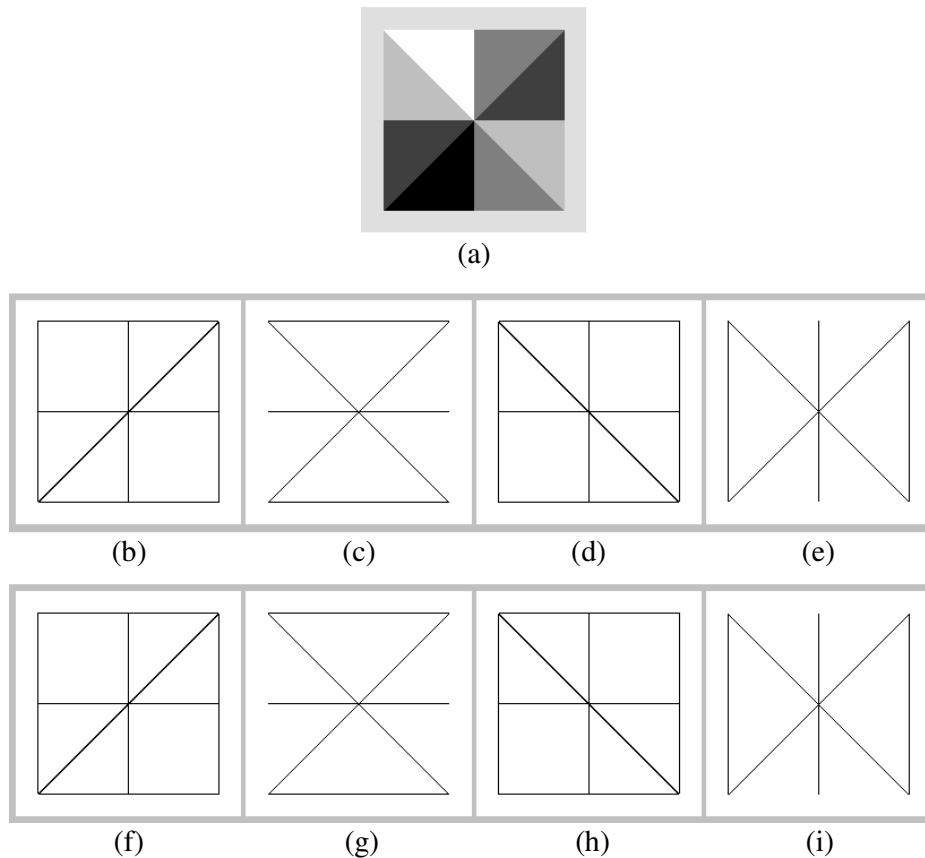


Figura 4.13: Pesos determinados com base na semelhança entre as intensidades de pixels adjacentes

A presença de ruído na imagem de entrada da rede pode conduzir à determinação imprecisa dos pesos e à segmentação incorreta das regiões da imagem. Para exemplificar este fato, a Figura 4.14 apresenta os pesos referentes à imagem da Figura 4.13 (a) com a adição de ruído gaussiano de média nula e variância igual a 25, e valores de L_w respectivamente iguais a 0, 5, 10 e 25. Para que as bordas da imagem possam ser claramente detectadas, L_w deve ser aumentado, o que, por outro lado, pode conduzir à supressão de alguns pontos pertencentes às bordas (Figura 4.14 (d)). Nesta situação, a rede pode proceder a segmentação incorreta da imagem. Outro aspecto negativo vinculado ao aumento do limiar L_w é a possibilidade de aglutinação de regiões adjacentes com pequenas diferenças entre os atributos dos pixels considerados para a segmentação da imagem.

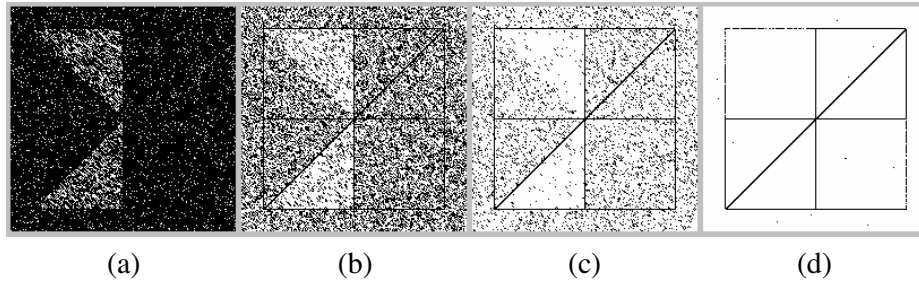


Figura 4.14: Influência do ruído na determinação dos pesos para diversos valores do parâmetro L_w

Tendo em vista o exposto, conclui-se que os pesos da rede ONNIS-GI podem ser determinados a partir de procedimentos de determinação de bordas. Também são relevantes a avaliação da influência do ruído na determinação dos pesos e o estabelecimento de procedimentos que tornem a determinação dos pesos mais robusta.

4.4.2 Um Procedimento Robusto para Determinação dos Pesos

Para melhorar o desempenho da rede ONNIS-GI na presença de ruído devem ser utilizados procedimentos que permitam a determinação robusta dos pesos nestas condições. A abordagem deste problema pode ser encarada sob dois diferentes pontos de vista, sendo o primeiro deles a utilização de filtragem para suavização da imagem de entrada, tornando a mesma mais homogênea, sem degradar significativamente as bordas das diversas regiões. Neste caso, podem ser adotados procedimentos mais simples para determinação dos pesos da rede, ficando a maior complexidade computacional vinculada ao procedimento de filtragem. O segundo ponto de vista diz respeito à utilização de procedimentos de detecção de bordas robustos ao ruído para determinação direta dos pesos da rede proposta.

Entre os procedimentos clássicos para detecção de bordas, o operador de Sobel destaca-se pela sua simplicidade computacional. A utilização de operações locais entre os pixels é também um fator atraente para implementação em arquiteturas maciçamente paralelas. Considerando-se que este operador agrega intrinsecamente uma filtragem do tipo passa-baixas, é obtida uma maior robustez em relação à presença de ruído na imagem a ser processada na determinação das bordas.

O operador de Sobel foi concebido para detectar as componentes das bordas nas direções vertical e horizontal através da convolução da imagem com as máscaras m_h e m_v definidas respectivamente em (4.16) e (4.17) (JAIN, 1989). Sob o ponto de vista de filtragem para suavização, as máscaras de Sobel podem ser compreendidas como a convolução da imagem individualmente com máscaras de suavização vertical e horizontal e posteriormente com máscaras que implementam as derivadas de primeira ordem também nas direções vertical e horizontal. As referidas máscaras de suavização e de diferenciação combinadas através da operação de convolução também podem ser observadas em (4.16) e (4.17).

$$m_v = \begin{bmatrix} -1 & 0 & +1 \\ -2 & 0 & +2 \\ -1 & 0 & +1 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 \\ 2 & 2 & 0 \\ 1 & 1 & 0 \end{bmatrix} * [-1 \ 1 \ 0] \quad (4.16)$$

$$m_h = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ +1 & +2 & +1 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 1 \\ 1 & 2 & 1 \\ 0 & 0 & 0 \end{bmatrix} * \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} \quad (4.17)$$

onde:

m_v é a máscara para obtenção das componentes verticais das bordas;
 m_h é a máscara para obtenção das componentes horizontais das bordas e
o símbolo * representa o operador convolução.

A utilização do operador de Sobel para determinação dos pesos da rede proposta apresenta dois inconvenientes. O primeiro deles diz respeito à determinação de bordas com espessura de um pixel, enquanto que os pesos da rede proposta são equivalentes a bordas infinitesimais entre dois pixels da imagem. O segundo inconveniente está relacionado com a determinação dos pesos nas direções diagonais. Neste caso, a utilização das componentes horizontal e vertical das bordas para determinação das bordas diagonais pode conduzir a distorções significativas, comprometendo os resultados da segmentação.

Inspirado no funcionamento do operador de Sobel, é proposto um procedimento para determinação robusta dos pesos da rede. Este procedimento consiste em filtrar inicialmente a imagem com a máscara m_o apresentada na equação (4.18). Através da mesma equação pode ser observado que a filtragem da imagem com a utilização da máscara m_o é equivalente a filtrações sucessivas nas direções horizontal e vertical, resultando em um efeito de suavização onidirecional. A imagem obtida através deste procedimento de filtragem é então utilizada para a determinação dos pesos da rede proposta utilizando a técnica anteriormente descrita. Resultados numéricos sobre o desempenho da rede proposta utilizando este procedimento para determinação dos pesos são apresentados no Capítulo 5.

$$m_o = \begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix} / 16 = \begin{bmatrix} 1 & 2 & 1 \end{bmatrix} / 4 * \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} / 4 \quad (4.18)$$

onde:

m_o é a máscara para a suavização onidirecional da imagem.

O procedimento proposto baseado no operador de Sobel apresenta baixa complexidade computacional e pode ser implementado com a utilização unicamente de operações entre EPs vizinhos, uma vantagem sob o ponto de vista de arquiteturas maciçamente paralelas. Outros procedimentos também poderiam ser utilizados como base para obter os pesos da rede. Uma das possibilidades seria a otimização dos coeficientes da máscara m_o em função do nível de ruído presente na imagem utilizando o método do gradiente estocástico (JAIN, 1989), por exemplo. Um procedimento baseado no método de Canny, reconhecidamente um dos que apresentam melhores resultados na obtenção de bordas na presença de ruído, apesar de computacionalmente complexo, é uma outra possibilidade. Finalmente, é importante ressaltar que o benefício decorrente da utilização de diferentes procedimentos para determinação dos pesos da rede proposta deve ser avaliado em termos do compromisso entre a qualidade da segmentação obtida e da complexidade computacional resultante. O aspecto relativo à utilização prioritária de operações entre EPs vizinhos é um outro ponto que pode

inviabilizar a utilização prática da algum procedimento a ser adotado no caso de uma implementação maciçamente paralela.

4.5 Conclusão

Neste capítulo foi apresentada a proposta de uma rede neural adequada à segmentação de imagens. A rede é basicamente um arranjo bidimensional de EPs com as mesmas dimensões da imagem de entrada e associado a estruturas de conexões locais excitatórias e inibitórias.

Foram apresentadas duas versões da rede proposta. A primeira delas, denominada ONNIS, foi inspirada diretamente no comportamento da rede LEGION. A segunda versão da rede proposta, denominada ONNIS-GI, utiliza um mecanismo de inibição global e apresenta uma estrutura adicional de conexões globais, sendo mais adequada a aplicações práticas do que a rede ONNIS.

Com base nos estudos e resultados obtidos, pode-se concluir que a rede proposta, em suas duas versões, apresenta características operacionais semelhantes à rede LEGION. A implementação em arquitetura maciçamente paralela é adequada à realização de *chips* de visão para segmentação de imagens, tendo também a capacidade de separar os objetos segmentados temporalmente, facilitando a posterior identificação. Quando comparada especificamente com a rede LEGION em sua versão contínua, as redes ONNIS e ONNIS-GI, apresentam as seguintes vantagens:

- Pode ser implementada mais facilmente utilizando tecnologia digital, ao contrário da proposta de Wang, a qual utiliza funções matemáticas que implicam maior complexidade computacional;
- Possui poucos parâmetros, de ajuste intuitivo;
- Não apresenta o efeito de percepção dúplex;
- Não possui capacidade de segmentação limitada;
- Apresenta flexibilidade para implementação de procedimentos de segmentação mais sofisticados e
- Não necessita da adição de um valor aleatório às entradas para diminuir a possibilidade de sincronismo entre segmentos distintos e evitar estados estáveis.

Na implementação da versão contínua da rede LEGION, foram adotadas operações matemáticas que implicam elevada carga computacional ou complexidade de *hardware* para sua implementação prática com circuitos digitais. As redes ONNIS e ONNIS-GI, por outro lado, são implementadas com operações lógicas simples, facilitando sua realização com dispositivos lógicos programáveis.

O neurônio oscilador de Terman-Wang apresenta uma série de parâmetros com ajuste pouco intuitivo. Da mesma forma, os parâmetros da rede LEGION não apresentam relação intuitiva com o procedimento de segmentação, tornado seu ajuste complicado e altamente dependente do sinal a ser segmentado. A rede ONNIS possui dois parâmetros enquanto que a rede ONNIS-GI possui um único parâmetro, cujo ajuste, apesar de dependente do sinal a ser segmentado, tem significado intuitivo.

O efeito indesejado da percepção dúplex não é verificado nas redes propostas. Tal característica implica que qualquer pixel da imagem de entrada seja obrigatoriamente vinculado a somente um segmento desta imagem.

A capacidade de segmentação limitada, uma grande desvantagem de ordem prática apresentada pela rede LEGION também não é verificada nas redes ONNIS e ONNIS-

GI, tornando as mesmas mais adequadas a aplicações onde o número de segmentos é elevado e desconhecido à priori.

A utilização de dispositivos digitais também torna as redes propostas mais flexíveis quanto ao aspecto da reprogramação. Tal flexibilidade pode ser explorada para implementação de procedimentos de segmentação mais sofisticados, utilizando atributos diversos, como a segmentação baseada em texturas, por exemplo.

Na rede LEGION, para facilitar o anti-sincronismo entre osciladores pertencentes a segmentos distintos e que se encontram na mesma fase é adicionado um ruído gaussiano de pequena amplitude às entradas da rede. Tal procedimento também evita o surgimento de estados estáveis indesejados tornando, porém, mais complexa a implementação da rede e dificultando o sincronismo de neurônios vizinhos com entradas semelhantes e baixa intensidade. O mecanismo de conexões excitatórias e inibitórias adotado nas redes ONNIS e ONNIS-GI torna desnecessária a adição de ruído às entradas da rede. Uma alternativa para evitar a adição de ruído às entradas da rede LEGION é inicializar os estados dos neurônios aleatoriamente em posições diferentes. Este procedimento, o qual é adotado na versão algorítmica da rede LEGION, além de não garantir o anti-sincronismo, pois é aleatório, implica maior complexidade na implementação da rede. Outra desvantagem desta solução é a necessidade de um dispositivo de *reset* para a rede sempre que algum parâmetro ou as entradas forem modificadas. As redes ONNIS e ONNIS-GI apresentam insensibilidade às condições iniciais dos neurônios, o que dispensa mecanismos adicionais de *reset* e ajuste do estado inicial.

O estudo da aplicação de procedimentos robustos para determinação de bordas para determinação dos pesos das redes propostas pode conduzir a um melhor desempenho das mesmas quanto à qualidade da segmentação em presença de ruído. Dentro deste contexto, é apresentada uma proposta para determinação dos pesos com base no operador de Sobel.

De forma geral, as vantagens apresentadas pela rede ONNIS-GI em relação à rede LEGION, e sua menor complexidade em relação à rede ONNIS, tornam a mesma mais adequada para implementações de *chips* de visão em tecnologia digital com capacidade de segmentar imagens.

Finalmente, ressalta-se também que, em relação a alguns modelos de redes neurais artificiais, as redes propostas apresentam como aspecto positivo a não necessidade de treinamento para determinação de seus pesos, os quais são calculados diretamente a partir do sinal de entrada da mesma.

5 RESULTADOS OBTIDOS COM A REDE ONNIS-GI

Para verificar o desempenho da rede ONNIS-GI foram realizados dois tipos de implementação. A primeira diz respeito à simulação do comportamento da rede através de um algoritmo implementado em computadores PC utilizando ambiente Delphi. Tal algoritmo simula o comportamento da versão da rede que utiliza o mecanismo de inicialização global. A segunda implementação refere-se à utilização do programa Max+plus II da Altera (ALTERA, 2004) para simulação da rede, em suas duas versões, e verificação da viabilidade de realização da mesma em dispositivos FPGA.

O presente capítulo apresenta resultados obtidos com a simulação da rede ONNIS-GI em microcomputadores PC, através de sua versão algorítmica implementada. Primeiramente, são descritos o procedimento para avaliação da qualidade dos resultados e o ambiente de simulação utilizado. Após, são apresentados os resultados obtidos na segmentação de imagens artificiais, naturais e em presença de ruído e variação de iluminação.

5.1 Procedimento de Avaliação da Qualidade da Segmentação

Para comparar diferentes procedimentos de determinação dos pesos da rede ou a utilização de diferentes valores de parâmetros para um mesmo procedimento foram adotados os seguintes critérios:

- Verificação do número de segmentos detectados pela rede;
- Percentual de pixels pertencentes à região segmentada e classificados como não pertencentes (falsa rejeição), denominado e_1 ;
- Percentual de pixels não pertencentes à região segmentada e classificados como pertencentes (falsa detecção), denominado e_2 , e
- Percentual total de pixels incorretamente classificados, denominado e_t .

A determinação dos percentuais de erro só foi realizada naquelas situações onde a rede conseguiu discriminar o número correto de segmentos da imagem. Entretanto, mesmo em número correto, os segmentos encontrados podem diferir em muito do resultado correto. Por outro lado, tal fato pode ser facilmente identificado através dos valores dos erros que, nesta situação, tornam-se bastante elevados podendo inclusive ultrapassar 100% dos pixels do segmento.

Na avaliação da qualidade da segmentação foram utilizadas duas imagens em 256 níveis de cinza geradas artificialmente, as quais foram corretamente segmentada por processo manual para obtenção das respectivas referências (*ground truth*). Uma destas imagens é apresentada na Figura 5.1, juntamente com os valores dos pixels relativos a cada uma das regiões da mesma. Na concepção desta imagem, a qual tem dimensões de 256x256 pixels, foram utilizadas bordas verticais, horizontais, diagonais e circulares. Os

valores dos pixels das regiões foram escolhidos com o objetivo de obter vários níveis de diferença entre as intensidades de regiões adjacentes.

Através da imagem apresentada na Figura 5.1 também pode ser verificada a diferença de sensibilidade do olho humano para diferentes intensidades. Os segmentos com intensidades 8 e 16 parecem visualmente compor um único segmento, enquanto que os segmentos com intensidade 240 e 248 podem ser mais facilmente discriminados, ainda que possuam a mesma diferença de intensidade dos segmentos mais escuros. A obtenção, com a rede proposta, de resultados semelhantes àqueles percebidos através visão humana pode ser alcançada através de alguma transformação adequada na imagem de entrada. Tendo em vista que, em algumas aplicações, a discriminação linear de diferenças de intensidades, independente do nível das mesmas, pode ser encarada como uma vantagem em relação ao mecanismo de visão humana, não é adotada qualquer transformação da imagem para segmentação utilizando a rede proposta. É importante também ressaltar que esta propriedade não linear do olho humano refere-se ao atributo intensidade, não podendo ser generalizada para outros atributos dos pixels da imagem.

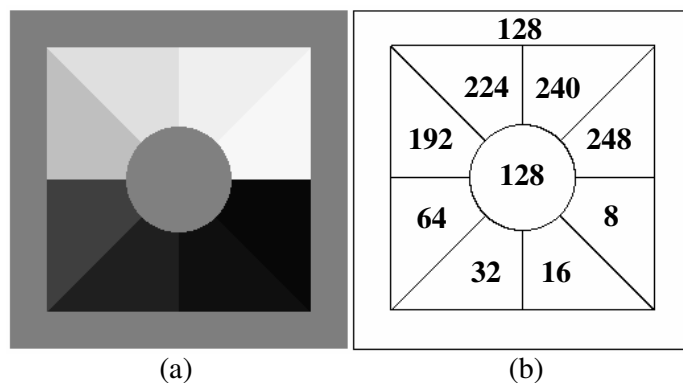


Figura 5.1: Imagem de teste (a). Bordas e valores dos pixels das regiões (b)

A Figura 5.2 apresenta os 10 segmentos obtidos com a versão algorítmica da rede ONNIS-GI com determinação dos pesos com base na diferença entre as intensidades de pixels adjacentes ($L_w=0$) a partir da imagem de teste da Figura 5.1. Os segmentos foram apresentados seqüencialmente na saída da rede, sendo a forma e o tamanho dos mesmos exatamente igual àquela obtida com a segmentação manual da imagem. Tal resultado já demonstra a correta operação da rede proposta.

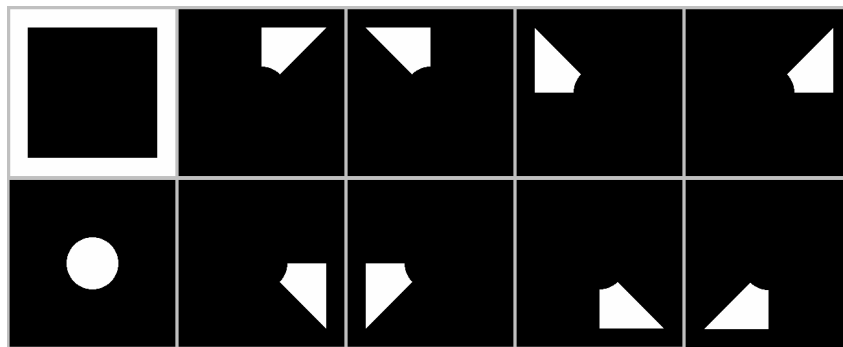


Figura 5.2: Segmentos da imagem de teste

Por outro lado, para avaliar a robustez da rede ONNIS-GI em relação à presença de ruído nas imagens segmentadas, foi utilizada uma segunda imagem de referência (Figura 5.3), à qual foi adicionado ruído gaussiano, com média nula e diferentes variâncias. Esta imagem, com dimensões de 256x256 pixels, é composta por três segmentos, os quais dizem respeito a um objeto octogonal possuindo um furo circular central disposto sobre um fundo homogêneo. O objeto octogonal possui intensidade igual a 150 e área correspondente a 30,13% (19744 pixels) da área total da imagem. A intensidade do fundo é igual a 100.



Figura 5.3: Imagem utilizada para avaliação da sensibilidade ao ruído

A influência do ruído sobre o procedimento de segmentação proposto está vinculada à relação entre a amplitude do mesmo e as amplitudes das bordas da imagem. Desta forma, definiu-se uma medida de relação sinal/ruído, representada pela equação (5.1), a qual relaciona a amplitude da borda a ser detectada e valor eficaz do ruído sobreposto. Para um ruído gaussiano de média nula, o valor eficaz do mesmo, v_r , corresponde ao seu desvio padrão. O valor de v_b corresponde à amplitude da borda considerada.

$$RSR_b = 20 \log_{10} \left(\frac{v_b}{v_r} \right) \quad (5.1)$$

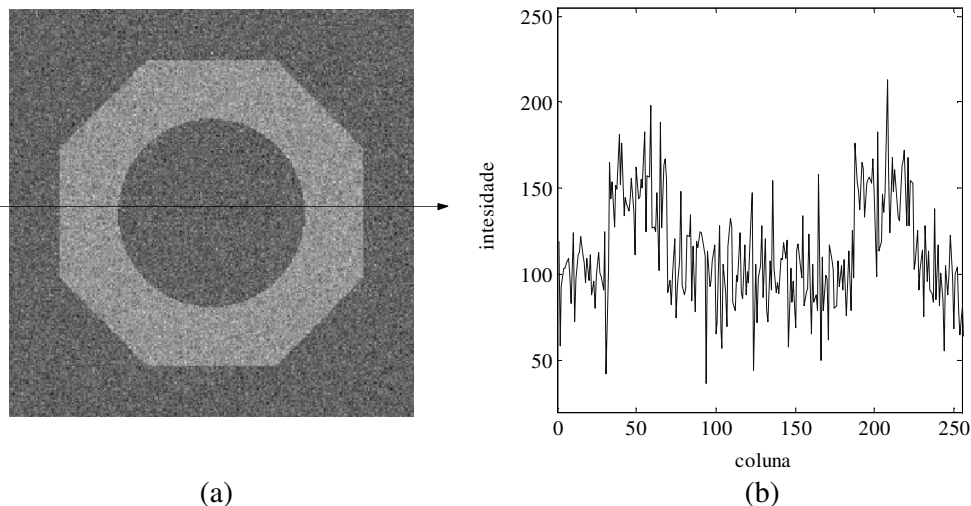


Figura 5.4: Imagem de teste corrompida por ruído gaussiano e $RSR_b=7,96$ dB

A Figura 5.4 (a) apresenta a imagem da Figura 5.3 com a adição de um ruído gaussiano com valor eficaz igual a 20. Tendo em vista que a amplitude das bordas presentes na imagem é igual a 50, a relação sinal/ruído computada com a utilização da equação (5.1) é de 7,96 dB. É apresentado também o perfil das intensidades correspondente à linha central da imagem (Figura 5.4 (b)), onde pode ser observada a dificuldade imposta pelo ruído à correta detecção das bordas.

5.2 Ambiente de Simulação

Para comprovar o correto funcionamento e o desempenho da rede ONNIS-GI, foi desenvolvido um aplicativo em ambiente Delphi baseado no algoritmo apresentado na seção 4.2.6. A rede foi implementada para realizar a segmentação da imagem de entrada segundo a semelhança de intensidade entre pixels vizinhos e utiliza o procedimento de inicialização global proposto, minimizando o número de imagens nulas na saída.

A Figura 5.5 apresenta a interface gráfica do aplicativo desenvolvido, onde podem ser visualizadas a imagem original, à esquerda, e uma imagem que apresenta um a um os segmentos obtidos em função do tempo, à direita. Os pixels desta imagem correspondem às saídas excitatórias dos EPs da rede, sendo que a cor branca representa os EPs ativos, enquanto que a cor preta representa os EPs inativos. As imagens utilizadas na entrada da rede são do tipo BMP, coloridas ou monocromáticas. As imagens coloridas são internamente substituídas por uma imagem em níveis de cinza correspondente à intensidade.

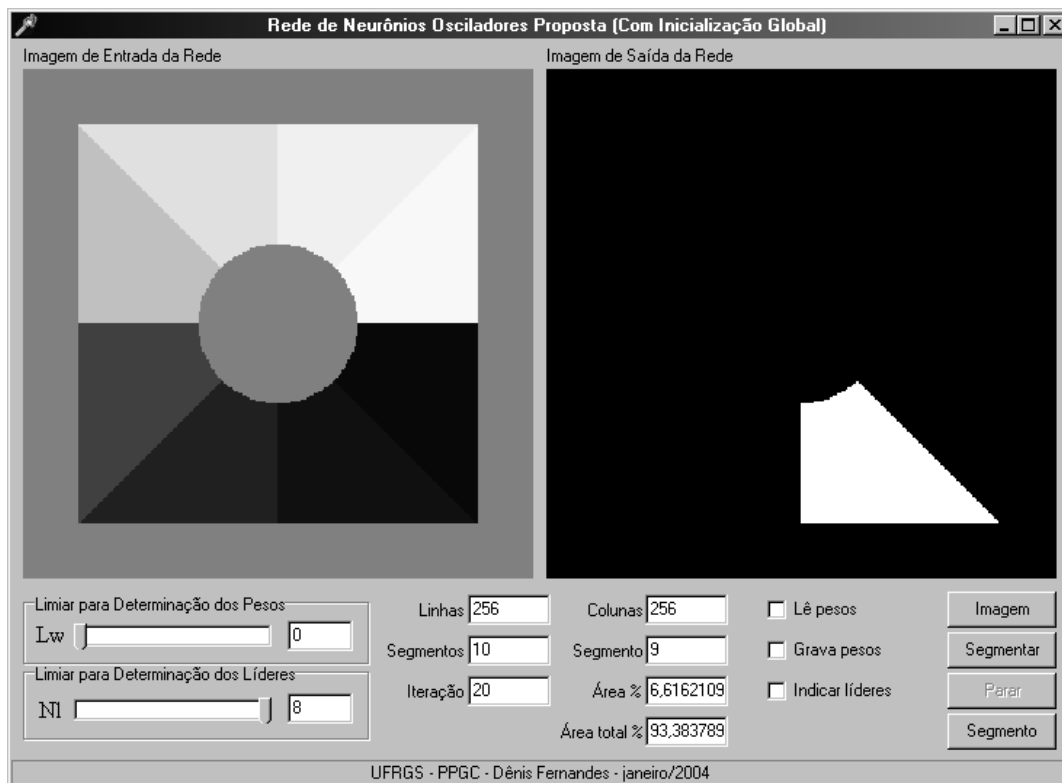


Figura 5.5: Tela do aplicativo desenvolvido para avaliação do comportamento da rede ONNIS-GI

Além dos segmentos da imagem de entrada, o aplicativo desenvolvido também fornece o número total de segmentos discriminados pela rede proposta, o número de ordem do segmento a partir da inicialização global, a área relativa do segmento em relação ao tamanho da imagem e a área total segmentada. São apresentados automaticamente o número de linhas e o número de colunas da imagem de entrada, correspondendo às dimensões da rede em questão. O tamanho de rede de EPs é automaticamente ajustado ao tamanho da imagem de entrada, já que a cada pixel da imagem deve corresponder um EP da rede.

É possível ajustar o parâmetro L_w com o objetivo de obter o melhor resultado possível na segmentação da imagem. Este ajuste pode ser realizado continuamente, sem parar o processamento da rede já que o resultado fornecido pela mesma depende somente do valor deste parâmetro e da imagem de entrada. O parâmetro N_l , apesar de ter sido considerado fixo na concepção da rede ($N_l=8$), pode ser ajustado para verificar sua influência no resultado da segmentação. Este parâmetro representa o número de EPs vizinhos semelhantes considerados para determinação dos EPs líderes ($0 \leq N_l \leq 8$).

Para determinação automática do número de segmentos da imagem de entrada, o algoritmo conta o número de iterações entre duas inicializações globais consecutivas. Em uma implementação em *hardware*, bastaria simplesmente contar o número de pulsos de *clock* entre pulsos consecutivos de inicialização global.

Como a segmentação é realizada somente com base na intensidade associada a cada pixel, as imagens coloridas são automaticamente convertidas para imagens de intensidade em tons de cinza (*grayscale*) representadas em 8 bits. A intensidade é obtida através da média aritmética das componentes R, G e B de cada pixel da imagem. Para implementar a segmentação com base em atributos diversos, é possível carregar um arquivo de pesos em formato específico, gerado externamente. Quando necessário, os arquivos de pesos para diferentes atributos da imagem de entrada foram gerados através de uma rotina implementada em Matlab. O aplicativo desenvolvido permite também salvar o arquivo de pesos utilizado pela rede.

Com a finalidade de segmentar somente determinadas regiões da imagem de entrada, foi implementado um procedimento para indicação de EPs líderes por parte do usuário. Assim, basta “clique” sobre as regiões desejadas da imagem de entrada para que somente os EPs correspondentes sejam habilitados como líderes. As posições marcadas na imagem são utilizadas para gerar os sinais $v_{li}(i,j)$ dos EPs líderes. Desta forma, somente os segmentos que possuem líderes indicados aparecerão na saída da rede.

É importante ressaltar que, tendo sido a rede ONNIS-GI concebida em arquitetura maciçamente paralela visando a implementação em *hardware* dedicado, sua simulação em máquinas seriais não apresenta o grande atrativo relativo à velocidade de processamento. Pode também ser claramente verificado o aumento drástico no tempo de processamento em função do tamanho da imagem, obtendo-se, contudo, os resultados desejados. Tratando-se também de um aplicativo destinado somente à avaliação do funcionamento da rede proposta, a questão relativa à otimização do software com o objetivo de minimizar o tempo de processamento foi desconsiderada.

5.3 Resultados Obtidos na Segmentação de Imagens Artificiais

A Figura 5.6 apresenta um exemplo de segmentação realizado com a implementação algorítmica da rede ONNIS-GI. A imagem original, com dimensões 100x310 pixels, observada na posição superior esquerda, possui 5 letras de mesmo tom de cinza isoladas umas das outras e sobrepostas a um fundo com diferente tom de cinza. A rede proposta determinou, a partir da intensidade de cada pixel da imagem original, a existência de 7

segmentos, os quais são representados pelas demais imagens. Os segmentos obtidos são apresentados seqüencialmente na saída da rede, observando-se que todas as letras foram corretamente segmentadas e isoladas no tempo, facilitando a aplicação posterior de um procedimento de reconhecimento de caracteres. Em todas as imagens referentes aos segmentos obtidos a cor preta representa EPs na fase silenciosa e a cor branca representa os EPs na fase ativa.

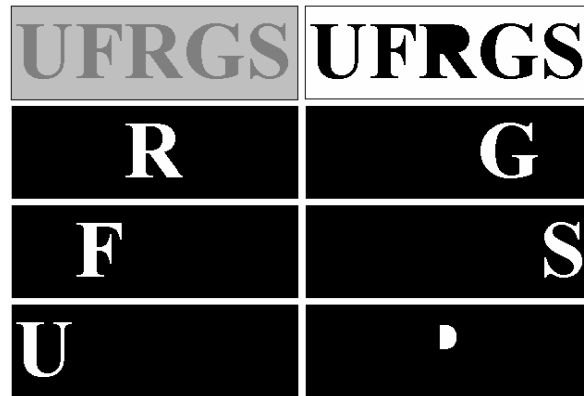


Figura 5.6: Segmentação de uma imagem obtida com a rede ONNIS-GI

5.4 Resultados Obtidos na Segmentação de Imagens Naturais

A Figura 5.7 apresenta o resultado obtido com a rede ONNIS-GI na segmentação da imagem da Figura 3.24, podendo ser observados os segmentos relativos às 9 bolhas de hélio e um segmento adicional, referente ao fundo de silício. As bolhas foram corretamente segmentadas pela rede, preservando tamanho e forma das mesmas, resultado bastante semelhante àquele encontrado com a utilização da rede LEGION (Figura 3.25 e Figura 3.26).

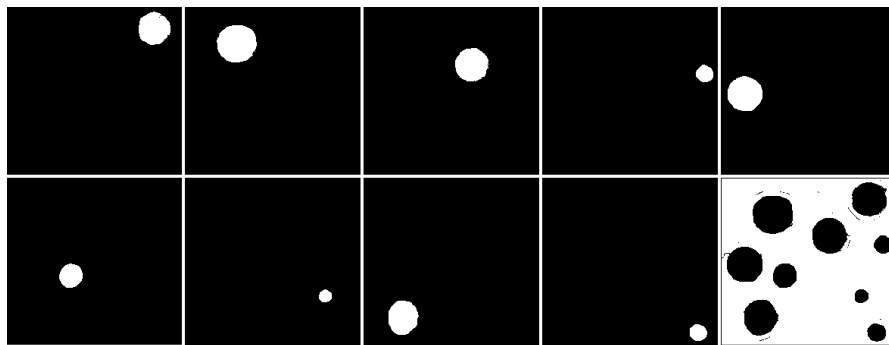


Figura 5.7: Segmentos obtidos com a aplicação da rede ONNIS-GI para segmentação da imagem da Figura 3.24

Já a Figura 5.8 apresenta os pixels (representados em preto) não segmentados pela rede. Tais pixels, por não pertencerem a regiões homogêneas da imagem de entrada, particularmente as regiões próximas às bordas das bolhas, não foram agregados a nenhum dos segmentos obtidos na saída da rede. Este resultado também se aproxima muito daquele obtido com a utilização da rede LEGION na segmentação da respectiva imagem (Figura 3.27).

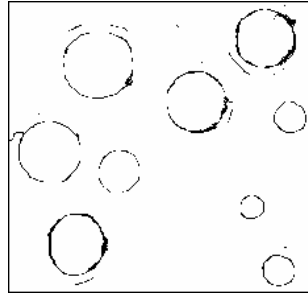


Figura 5.8: Pixels não segmentados pela rede

O resultado apresentado, entre outros, indica que a rede ONNIS-GI tem capacidade semelhante àquela apresentada pela rede LEGION no que diz respeito a aplicações em segmentação de imagens reais.

5.5 Desempenho da Rede ONNIS-GI na Presença de Ruído

O desempenho da rede ONNIS-GI na presença de ruído foi avaliado através de dois procedimentos, cujos resultados são apresentados a seguir. O primeiro procedimento considera a determinação dos pesos diretamente a partir da imagem de entrada utilizando a equação (4.5). O segundo procedimento utiliza uma pré filtragem da imagem com a máscara de coeficientes proposta em 4.4.2.

5.5.1 Resultados Obtidos sem Filtragem Preliminar da Imagem

A Tabela 5.1 apresenta os resultados obtidos com a utilização da rede ONNIS-GI para segmentação de imagens na presença de ruído sem qualquer filtragem. Os resultados foram obtidos utilizando-se a imagem de teste apresentada na Figura 5.3, tendo sido adicionado ruído gaussiano com valores eficazes (v_r) 0, 1, 5, 10 e 25 e parâmetro L_w entre 50 e 0.

A relação sinal/ruído RSR_b apresentada na Tabela 5.1 foi calculada utilizando a equação (5.1) situando-se ente os valores infinito (para $v_r=0$) e 6,02 dB (para $v_r=25$). Tendo em vista a característica aleatória do ruído, foram realizadas 100 simulações para cada valor eficaz de ruído e para cada valor de L_w . A coluna denominada $Oc.$ indica o número de vezes nas quais a rede conseguiu discriminar 3 segmentos a partir da imagem de teste. Os casos em que o número de segmentos foi diferente de 3 não foram utilizados no cálculo dos erros apresentados. Os erros percentuais e_1 , e_2 e e_r , conforme definidos na seção 5.1, foram calculados somente para o segmento correspondente à região octogonal presente na imagem, utilizada como referência.

A primeira linha de dados da tabela indica que para uma relação sinal/ruído infinita e $L_w=50$, não foi possível discriminar os segmentos desejados ($Oc.=0$), não tendo sido determinados os erros. Já da segunda linha de dados da tabela conclui-se que, também para uma relação sinal/ruído infinita, consegue-se discriminar o segmento desejado com erro nulo em 100% das ocorrências para $0 \leq L_w \leq 49$, situação esta exemplificada através da Figura 5.9. A imagem original é representada em (a), enquanto que os 3 segmentos obtidos na saída da rede são representados através de (b), (c) e (d). Já com uma relação sinal/ruído de 13,98 dB, a rede consegue discriminar os 3 segmentos em 95% dos casos, com $L_w=7$ e erro total máximo de 13,7%. Já com relação sinal/ruído igual a 6,02 dB, mesmo tendo sido encontrados 3 segmentos, em muitos casos sua áreas não correspondiam aos segmentos desejados, implicando erros elevados, em alguns casos maiores do que 100%.

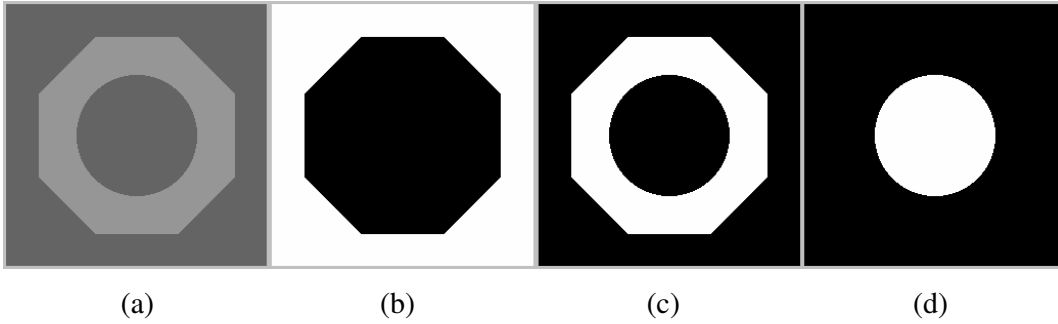


Figura 5.9: Segmentos obtidos da imagem de teste com $RSR_b = \infty$ e $L_w = 0$

A Figura 5.10 apresenta um exemplo dos resultados obtidos com relação sinal/ruído de 13,98 dB e $L_w = 7$, no qual foi possível separar adequadamente as três regiões da imagem de entrada da rede. Já no exemplo representado pela Figura 5.11, obtido com $RSR_b = 6,02$ dB e $L_w = 13$, foram encontrados também três segmentos, porém com grande discrepância de forma e área quando comparados aos segmentos corretos. Tanto na Figura 5.10 quanto na Figura 5.11, a imagem original, corrompida pelo ruído gaussiano encontra-se representada em (a), enquanto que os 3 segmentos obtidos na saída da rede são representados através das imagens (b), (c) e (d).

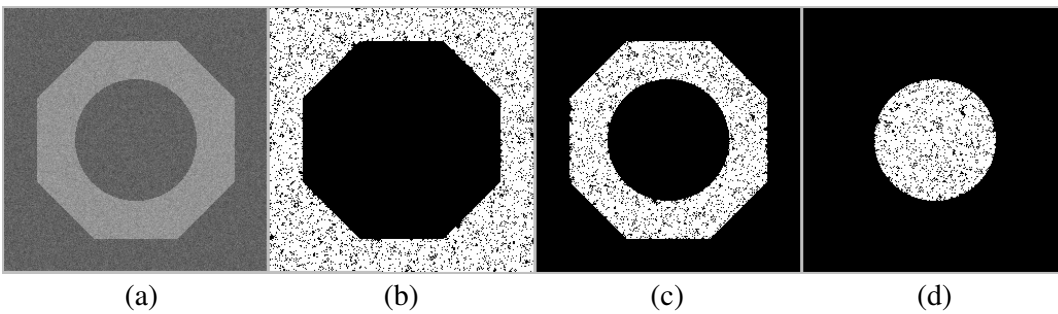


Figura 5.10: Segmentos obtidos da imagem de teste com $RSR_b = 13,98$ dB e $L_w = 7$

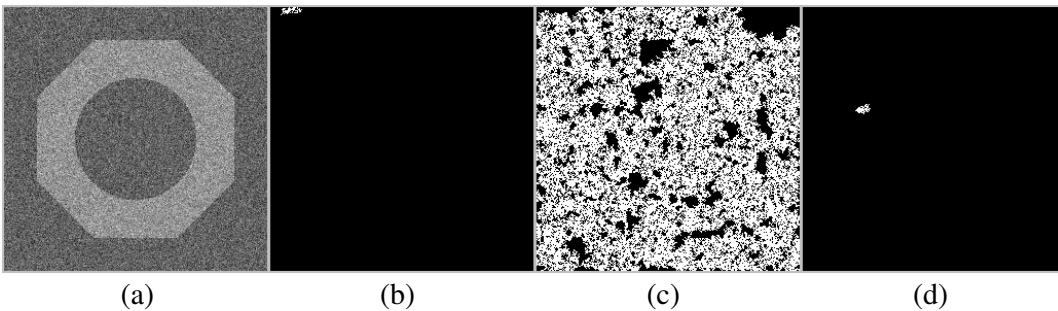


Figura 5.11: Segmentos obtidos da imagem de teste com $RSR_b = 6,02$ dB e $L_w = 13$

5.5.2 Resultados com o Procedimento Robusto Proposto

A Tabela 5.2 apresenta os resultados obtidos com a utilização do procedimento robusto proposto para determinação dos pesos da rede. De forma análoga ao item anterior, os resultados foram obtidos utilizando-se a imagem de teste apresentada na Figura 5.3, com ruído gaussiano adicionado e variando-se do parâmetro L_w , com 100 simulações para cada valor eficaz de ruído e para cada valor de L_w .

Tabela 5.2: Erro obtido com a utilização do procedimento robusto proposto para determinação dos pesos da rede

v_r	RSR_b	L_w	Oc.	$e_{1\ min\ \%}$	$e_{1\ med\ \%}$	$e_{1\ max\ \%}$	$e_{2\ min\ \%}$	$e_{2\ med\ \%}$	$e_{2\ max\ \%}$	$e_{1\ min\ \%}$	$e_{1\ med\ \%}$	$e_{1\ max\ \%}$
0	-	50...12	0	-	-	-	-	-	-	-	-	-
0	-	11...9	100	0	0	0	0	0	0	0	0	0
0	-	8...6	100	2,92E+00	2,92E+00	2,92E+00	0	0	0	2,92E+00	2,92E+00	2,92E+00
0	-	5...3	100	4,46E+00	4,46E+00	4,46E+00	0	0	0	4,46E+00	4,46E+00	4,46E+00
0	-	2...0	100	6,28E+00	6,28E+00	6,28E+00	0	0	0	6,28E+00	6,28E+00	6,28E+00
1	33,98	50...11	0	-	-	-	-	-	-	-	-	-
1	33,98	10	100	0	0	0	0	0	0	0	0	0
1	33,98	9	100	0	0	5,00E-03	0	0	0	0	0	5,00E-03
1	33,98	8	100	9,10E-02	8,06E-01	2,89E+00	0	0	0	9,10E-02	8,39E-01	2,89E+00
1	33,98	7	100	1,62E-01	2,89E+00	2,92E+00	0	0	0	1,62E-01	2,87E+00	2,92E+00
1	33,98	6	100	2,92E+00	2,94E+00	3,00E+00	0	0	0	2,92E+00	2,94E+00	3,00E+00
1	33,98	5	100	3,91E+00	4,13E+00	4,32E+00	0	0	0	3,91E+00	4,14E+00	4,32E+00
1	33,98	4	100	4,45E+00	4,46E+00	4,46E+00	0	0	0	4,45E+00	4,46E+00	4,46E+00
1	33,98	3	100	4,46E+00	4,46E+00	4,46E+00	0	0	0	4,46E+00	4,46E+00	4,46E+00
1	33,98	2	100	4,53E+00	4,60E+00	4,66E+00	0	0	0	4,53E+00	4,60E+00	4,66E+00
1	33,98	1	100	5,71E+00	6,16E+00	6,28E+00	0	0	0	5,71E+00	6,16E+00	6,28E+00
1	33,98	0	100	6,28E+00	6,28E+00	6,28E+00	0	0	0	6,28E+00	6,28E+00	6,28E+00
5	20	50...9	0	-	-	-	-	-	-	-	-	-
5	20	8	38	4,60E-02	8,10E-02	1,37E-01	0	1,20E-02	9,10E-02	4,60E-02	9,40E-02	1,82E-01
5	20	7	91	1,22E-01	2,40E-01	1,16E+00	0	5,00E-03	1,27E-01	1,22E-01	2,37E-01	1,16E+00
5	20	6	99	3,70E-01	1,48E+00	3,36E+00	0	0	4,60E-02	3,70E-01	1,49E+00	3,36E+00
5	20	5	100	1,12E+00	3,26E+00	3,87E+00	0	0	0	1,12E+00	3,23E+00	3,87E+00
5	20	4	100	2,82E+00	4,05E+00	4,32E+00	0	0	0	2,82E+00	4,08E+00	4,32E+00
5	20	3	100	4,03E+00	4,37E+00	4,47E+00	0	0	0	4,03E+00	4,37E+00	4,47E+00
5	20	2	100	4,46E+00	4,56E+00	4,63E+00	0	0	0	4,46E+00	4,55E+00	4,63E+00
5	20	1	100	4,97E+00	5,20E+00	5,44E+00	0	0	0	4,97E+00	5,20E+00	5,44E+00
5	20	0	6	1,72E+01	1,89E+01	1,98E+01	0	0	0	1,72E+01	1,89E+01	1,98E+01
10	13,98	50...8	0	-	-	-	-	-	-	-	-	-
10	13,98	7	13	2,43E-01	3,29E-01	5,42E-01	0	2,00E-02	6,60E-02	2,58E-01	3,31E-01	5,67E-01
10	13,98	6	76	4,81E-01	8,59E-01	1,29E+00	0	1,00E-02	9,10E-02	4,91E-01	8,48E-01	1,30E+00
10	13,98	5	99	1,33E+00	1,85E+00	2,42E+00	0	2,00E-03	4,60E-02	1,33E+00	1,83E+00	2,43E+00
10	13,98	4	100	2,54E+00	3,13E+00	3,75E+00	0	0	1,50E-02	2,54E+00	3,11E+00	3,75E+00
10	13,98	3	100	3,81E+00	4,18E+00	4,60E+00	0	0	0	3,81E+00	4,19E+00	4,60E+00
10	13,98	2	94	5,48E+00	6,06E+00	6,59E+00	0	0	0	5,48E+00	6,07E+00	6,59E+00
10	13,98	1	12	1,58E+01	1,74E+01	1,90E+01	0	0	0	1,58E+01	1,76E+01	1,90E+01
10	13,98	0	10	9,99E+01	1,00E+02	1,00E+02	0	2,00E-01	4,20E-01	9,99E+01	1,00E+02	1,00E+02
25	6,02	50...7	0	-	-	-	-	-	-	-	-	-
25	6,02	6	4	9,99E+01	1,00E+02	1,00E+02	2,00E-02	8,00E-02	1,47E-01	1,00E+02	1,00E+02	1,00E+02
25	6,02	5	28	6,46E+00	6,01E+01	1,00E+02	6,60E-02	3,56E+01	1,62E+02	6,96E+00	9,36E+01	1,69E+02
25	6,02	4	12	1,41E+01	5,03E+01	1,00E+02	0	1,61E+01	4,82E+01	1,45E+01	7,85E+01	1,48E+02
25	6,02	3...2	0	-	-	-	-	-	-	-	-	-
25	6,02	1	4	1,00E+02	1,00E+02	1,00E+02	5,60E-02	1,62E-01	2,53E-01	1,00E+02	1,00E+02	1,00E+02
25	6,02	0	0	-	-	-	-	-	-	-	-	-

Com uma relação sinal/ruído de 13,98 dB e parâmetro $L_w=4$, por exemplo, foi possível separar adequadamente as três regiões presentes na imagem de entrada da rede em 100% dos casos simulados com erro total máximo de 3,75%. Este resultado é superior aquele apresentado sem a utilização do procedimento robusto (3 segmentos em 95% dos casos, com $L_w=7$ e erro total máximo de 13,7%). A Figura 5.12 apresenta um exemplo dos resultados obtidos com a utilização do procedimento robusto proposto. Comparando este resultado com aquele apresentado na Figura 5.10 pode ser observada uma maior homogeneidade nos segmentos obtidos, resultado mais próximo da segmentação ideal.

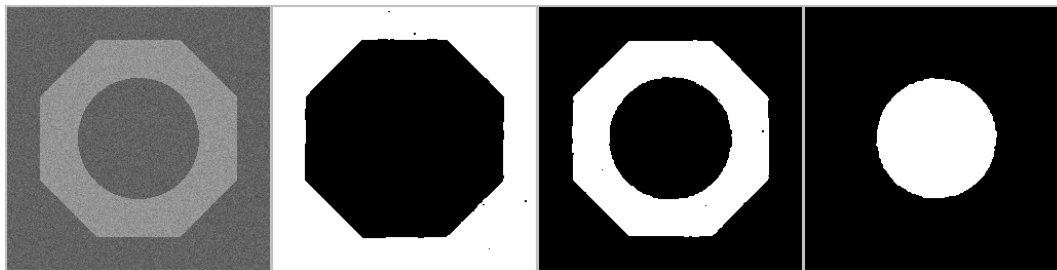


Figura 5.12: Exemplo de segmentação utilizando o procedimento robusto proposto para determinação dos pesos ($RSR_b=13,98$ dB e $L_w=4$)

Para imagens com baixa relação sinal/ruído, a utilização do procedimento robusto proposto tende a fornecer geralmente melhores resultados do que a não utilização do mesmo. Por outro lado, para relações sinal/ruído elevadas a situação se inverte devido ao efeito de suavização das bordas, ainda que reduzido, introduzido pelo procedimento de filtragem proposto. Tal fato conduz à expectativa de que o ajuste dos coeficientes da máscara do filtro de acordo com a intensidade do ruído presente na imagem possa conduzir a resultados superiores, ao custo, é claro, de um aumento na complexidade computacional.

5.6 Efeito da Variação de Iluminação na Imagem de Entrada

Assim como a presença de ruído, a variação espacial da iluminação sobre uma imagem é um fator que dificulta sua segmentação adequada. Variações de iluminação espacialmente lentas podem não ser tão problemáticas. Por outro lado, variações rápidas de iluminação, caracterizando áreas de sombra, podem levar o procedimento de segmentação a discriminar equivocadamente tais regiões como sendo objetos da imagem. A seguir são apresentados resultados obtidos com a aplicação da rede ONNIS-GI na segmentação de imagens onde a componente de iluminação sofre variação espacial ao longo da imagem.

Um modelo simplificado para a descrição de uma imagem representa a mesma como o produto ponto a ponto de duas componentes, uma denominada refletância e a outra denominada iluminação. Tais componentes dizem respeito respectivamente às características de reflexão dos objetos presentes na imagem e à quantidade de luz incidente sobre os mesmos (GONZALES; WOODS, 2000). Para variar a iluminação em uma imagem basta portanto multiplicar os valores de seus pixels por uma função com as mesmas dimensões da imagem e com valores ente 0 (iluminação mínima) e 1 (iluminação máxima).

A Figura 5.13 apresenta um exemplo onde a imagem da teste da Figura 5.1 foi submetida à uma componente de iluminação com variação linear no sentido vertical. No caso, a iluminação decresce de seu valor máximo (100%) na parte superior da imagem até o valor mínimo de 20% na parte inferior. Com a utilização de $L_w=1$, e sem qualquer filtragem preliminar, a rede proposta consegue discriminar todos os 10 segmentos da imagem, conforme apresentados na Figura 5.2. Neste mesmo exemplo pode ser observado que a rede proposta obteve maior poder de discriminação nas regiões de menor luminosidade (parte inferior da imagem) em relação à análise visual humana, a qual perde eficiência em cenas escuras.

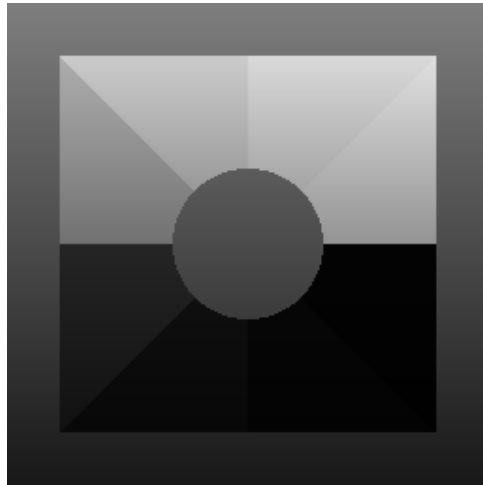


Figura 5.13: Imagem com variação vertical linear de iluminação

A Figura 5.14 apresenta uma situação na qual a variação de iluminação ao longo da imagem ocorre de maneira abrupta. Neste caso, a metade esquerda da imagem apresentada tem sua iluminação reduzida a 20% de seu valor original. A Figura 5.15 apresenta os 11 segmentos obtidos utilizando a rede proposta também com $L_w=1$ e sem qualquer filtragem preliminar. Pode ser observado que a rede aglutinou segmentos distintos em um único segmento (terceira imagem superior a partir da esquerda) e também separou segmentos originalmente homogêneos correspondentes ao círculo central e à borda.

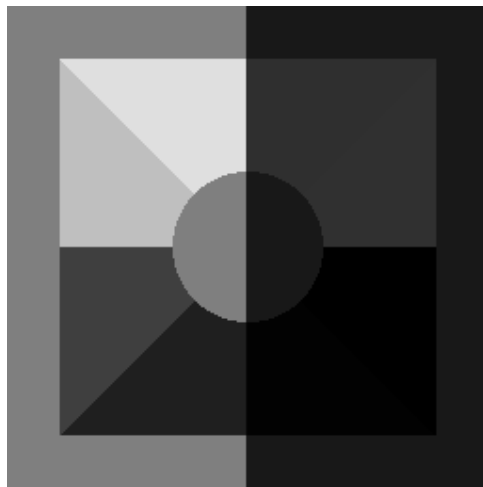


Figura 5.14: Imagem com variação abrupta de iluminação

Com base nos resultados apresentados, conclui-se que a segmentação utilizando a rede proposta apresenta sensibilidade a mudanças de iluminação sobre a imagem, fato igualmente comum a outros procedimentos de segmentação. Tal sensibilidade, a qual depende do conteúdo da imagem, principalmente da definição e da intensidade das bordas, pode ser minimizada através do ajuste do parâmetro L_w de forma criteriosa.

Nos casos onde a variação de iluminação prejudicar significativamente o resultado da segmentação, algum procedimento específico deve ser adotado. Uma alternativa

seria a utilização de técnicas de filtragem homomórfica para minimizar o efeito da variação da iluminação, realçando detalhes em regiões mais escuras da imagem (JAIN, 1989). Uma outra possibilidade seria a adaptação de L_w ao conteúdo da imagem, com valores menores naquelas regiões onde o nível de iluminação for reduzido. O efeito destas propostas carece, no entanto, de estudos mais aprofundados para determinar suas reais vantagens e limitações para resolver o problema em questão.

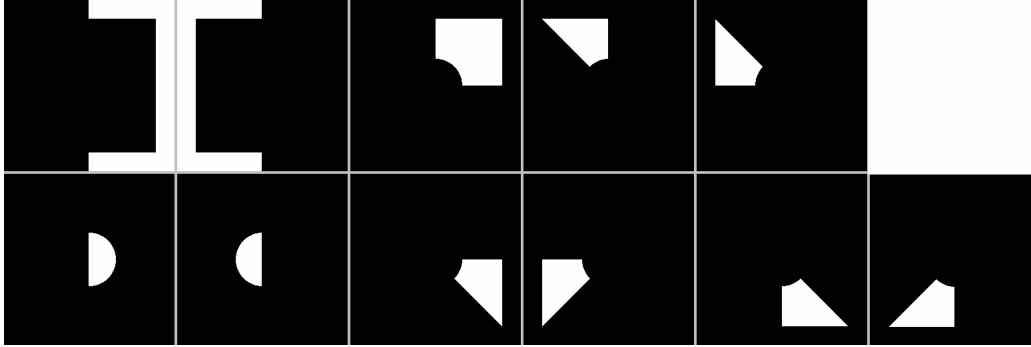


Figura 5.15: Segmentos obtidos a partir da imagem da Figura 5.14

5.7 Comparação com um Procedimento de Segmentação Clássico

Na rede proposta, em qualquer de suas duas versões, a ativação do EP de maior prioridade conduz a uma ativação seqüencial de EPs correspondentes ao mesmo segmento. Tal comportamento é similar ao que ocorre no clássico método de segmentação denominado crescimento de regiões por agregação de pixels (GONZALES; WOODS, 2000). Esta comparação permite estabelecer uma equivalência de desempenho quanto à qualidade da segmentação entre a rede proposta e o referido método de segmentação de imagens.

O procedimento de segmentação por crescimento de regiões consiste no agrupamento de pixels ou sub-regiões em regiões maiores. Na agregação de pixels utiliza-se como base um conjunto de pixels denominados sementes. A partir das sementes, são formados segmentos através da anexação de pixels que possuam propriedades similares. Um dos critérios utilizados para promover a agregação de pixels é o da diferença absoluta entre as intensidades de pixels adjacentes, o que é um procedimento equivale àquele utilizado para determinação dos pesos da rede proposta com base na equação (4.5).

A Figura 5.16 apresenta um exemplo de segmentação utilizando o método de crescimento de regiões por agregação de pixels. Neste caso, foi utilizada a mesma imagem apresentada na Figura 5.10(a) como entrada do processo. A imagem da Figura 5.16(a) apresenta a indicação do pixel utilizado como semente para o início do procedimento de segmentação. Para manter uma equivalência com o resultado apresentado na Figura 5.10, onde foi utilizado $L_w=7$, os pixels cujas intensidades diferem em no máximo 7 em relação a algum de seus vizinhos pertencentes à região da semente foram seqüencialmente aglutinados ao segmento. A Figura 5.16(b) indica o resultado após 10 iterações, enquanto que a Figura 5.16(c) apresenta o resultado após a iteração 100. Já a Figura 5.16(d) apresenta o resultado final da segmentação, após 225 iterações, o qual é idêntico ao resultado apresentado na Figura 5.10(c), obtido com a rede ONNIS-GI.

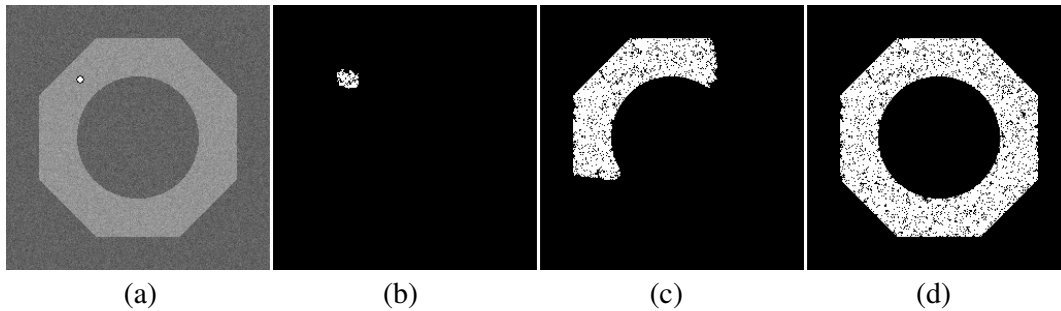


Figura 5.16: Segmento obtido através do método de crescimento de regiões por agregação de pixels

5.8 Conclusão

No presente capítulo foram relatados os resultados práticos obtidos com a utilização da rede ONNIS-GI em segmentação de imagens. A partir de um procedimento para avaliação da qualidade da segmentação e de um aplicativo desenvolvido para simular o comportamento da rede foram realizados vários testes avaliando inclusive o efeito do ruído e da iluminação não homogênea sobre os resultados obtidos.

O procedimento de avaliação da qualidade de segmentação baseia-se no número ou percentual de pixels corretamente classificados como pertencentes a um determinado segmento da imagem de entrada. Os resultados numéricos obtidos com a medida de qualidade estabelecida, revelam que a mesma apresentou-se adequada e compatível com a qualidade visualmente percebida.

Já um aplicativo desenvolvido em ambiente Delphi implementa a rede ONNIS-GI para uma aplicação genérica. Apesar de não ter sido otimizado em relação à velocidade de processamento, e sim em função da flexibilidade para fins de avaliação da qualidade dos resultados, o aplicativo apresentou-se muito mais rápido do que um aplicativo semelhante desenvolvido para simular a rede LEGION.

Tanto os resultados obtidos na segmentação de imagens artificiais quanto aqueles obtidos na segmentação de imagens naturais demonstraram que a rede proposta tem habilidade de segmentação similar à da rede LEGION, com diversas vantagens de ordem prática.

Os resultados obtidos com a adição de ruído gaussiano à imagem de entrada demonstraram que a rede possui intrinsecamente uma certa robustez ao ruído. Neste caso, a qualidade dos resultados depende muito da relação entre o nível do ruído e a amplitude das bordas presentes na imagem. Verificou-se também que o ajuste cuidadoso do parâmetro L_w é necessário nesta situação.

A utilização do procedimento robusto proposto para determinação dos pesos de rede apresentou, em alguns casos, melhoria significativa dos resultados para imagens com baixa relação sinal/ruído. Já nos casos onde a relação sinal/ruído é elevada, o procedimento robusto apresentou resultados inferiores, indicando que a adaptação dos pesos da máscara de filtragem em função da relação sinal/ruído pode ser uma alternativa mais frutífera.

Um outro efeito que pode comprometer a qualidade da segmentação é a variação da iluminação ao longo da imagem de entrada da rede. Verificou-se que variações espacialmente lentas de iluminação podem ser desconsideradas pela rede e que, por outro lado, variações abruptas de iluminação podem comprometer significativamente os resultados da segmentação. Soluções específicas, como a utilização de filtragem homomórfica por exemplo, poderiam ser desenvolvidas para amenizar este problema.

A comparação com o clássico método de crescimento de regiões por agregação de pixels comprova que a rede proposta apresenta resultados idênticos quando o critério para agregação de pixels é semelhante ao critério utilizado para cálculo dos pesos da rede. Desta forma, pode ser afirmado que a rede proposta, em suas duas versões, apresenta qualidade de segmentação equivalente ao procedimento de segmentação por agregação de pixels.

Finalmente, com base nos resultados obtidos, conclui-se que a rede ONNIS-GI apresenta a capacidade esperada de segmentar imagens. Por outro lado, mesmo não sendo uma exclusividade da mesma, a sensibilidade ao ruído e à variação de iluminação pode ser um empecilho de ordem prática, podendo ser necessário o estudo de algum procedimento para sua minimização em determinadas aplicações.

6 IMPLEMENTAÇÃO EM FPGA

O presente capítulo tem como finalidade apresentar os resultados obtidos na simulação de pequenos protótipos das redes ONNIS e ONNIS-GI para implementação em FPGA com a utilização do programa Max+plus II da Altera. O objetivo deste estudo é o de obter argumentos que comprovem a viabilidade de implementação prática das redes utilizando dispositivos digitais.

Tendo em vista a capacidade dos dispositivos FPGA disponíveis e a dificuldade prática para implementar uma interface de dados maciçamente paralela para um número elevado de EPs, foram implementadas redes de pequenas dimensões. Este fato, contudo, não invalida a verificação do correto funcionamento da rede, além do que os resultados obtidos podem ser generalizados para redes de quaisquer dimensões.

Ao longo deste capítulo são apresentados detalhes sobre as duas implementações realizadas. A primeira implementação é relativa à rede sem inicialização global e a segunda relativa à rede com inicialização global, a qual resulta em uma menor complexidade do *hardware*.

6.1 Rede ONNIS

A Figura 6.1 apresenta uma rede de 3x5 EPs ($N_n=15$) sem inicialização global simulada e fisicamente implementada em um *chip* FPGA. A rede possui como entrada um *clock*, comum a todos os EPs, e, como saída, os sinais excitatórios dos 15 EPs (S_1, S_2, \dots, S_{15}). Para simplificação do *hardware* os valores de N_s, L_w e as entradas referentes às intensidades dos pixels foram considerados parâmetros da rede na etapa de compilação do Max+plus II, o que não constitui qualquer restrição quanto à correta operação da rede. Em uma aplicação real, entretanto, tais entradas devem ser externamente acessíveis para possibilitar a modificação dos parâmetros e dos valores dos pixels da imagem de entrada por parte do usuário.

A estrutura interna do EP da rede da Figura 6.1 é apresentada na Figura 6.2. Com o objetivo de simplificar a implementação, cada EP da rede é responsável pela determinação de somente 4 do total de 8 pesos relativos às suas conexões excitatórias. Desta forma, um EP combina as excitações recebidas de 4 de seus EPs vizinhos com os respectivos pesos calculados internamente e envia sua excitação já devidamente combinada com os respectivos pesos a estes mesmos EPs vizinhos. Este artifício também torna necessário o recebimento das intensidades de somente 4 EPs vizinhos para cálculo dos respectivos pesos, diminuindo o número total de conexões da rede para manutenção da relação de vizinhança de 8 pixels. A Figura 6.3, a qual apresenta a lógica interna do EP, ilustra este aspecto. A estrutura utilizada para determinação dos pesos da rede, baseada no valor absoluto da diferença de intensidade de pixels adjacentes, é apresentada na Figura 6.4.

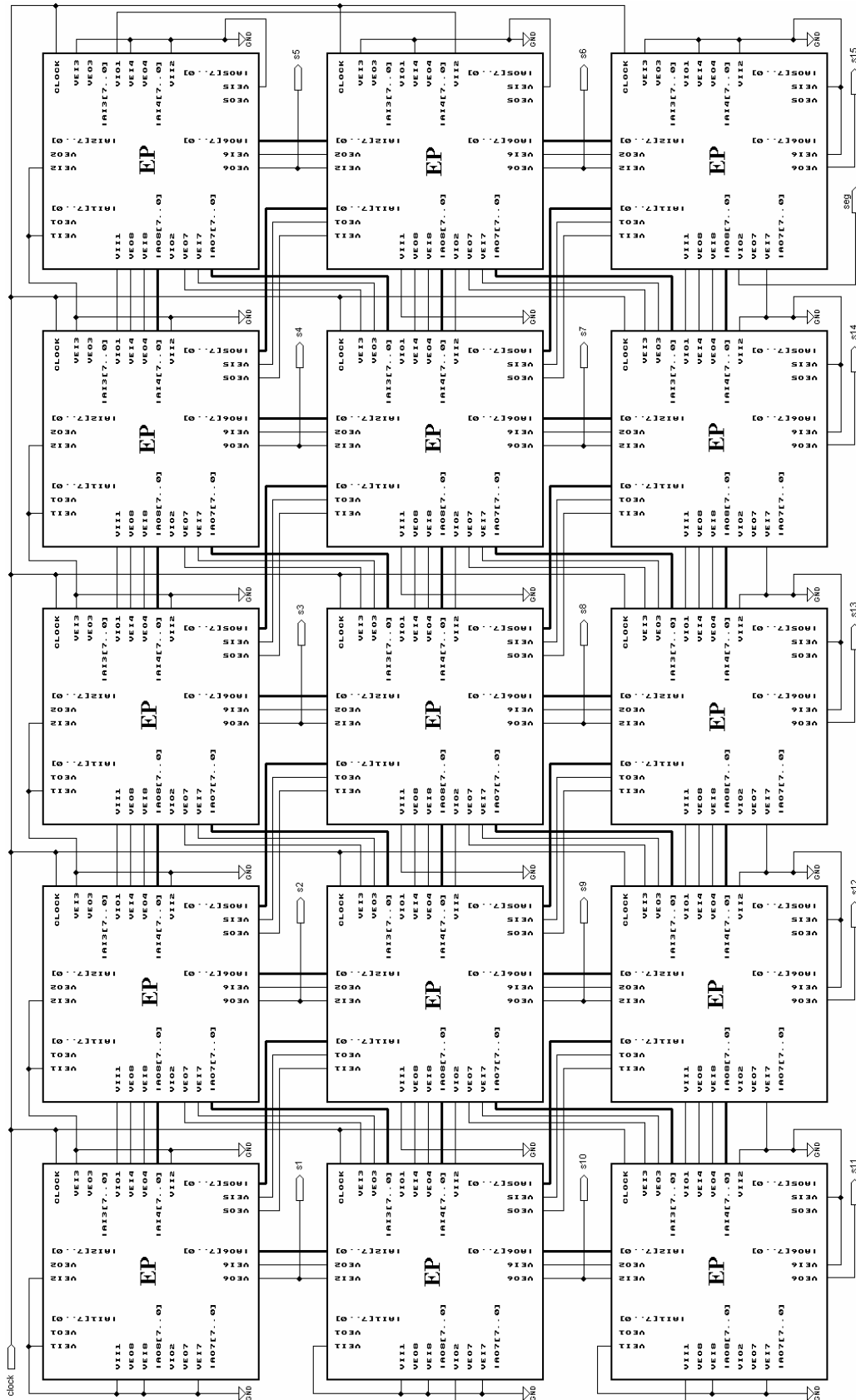


Figura 6.1: Implementação da rede ONNIS

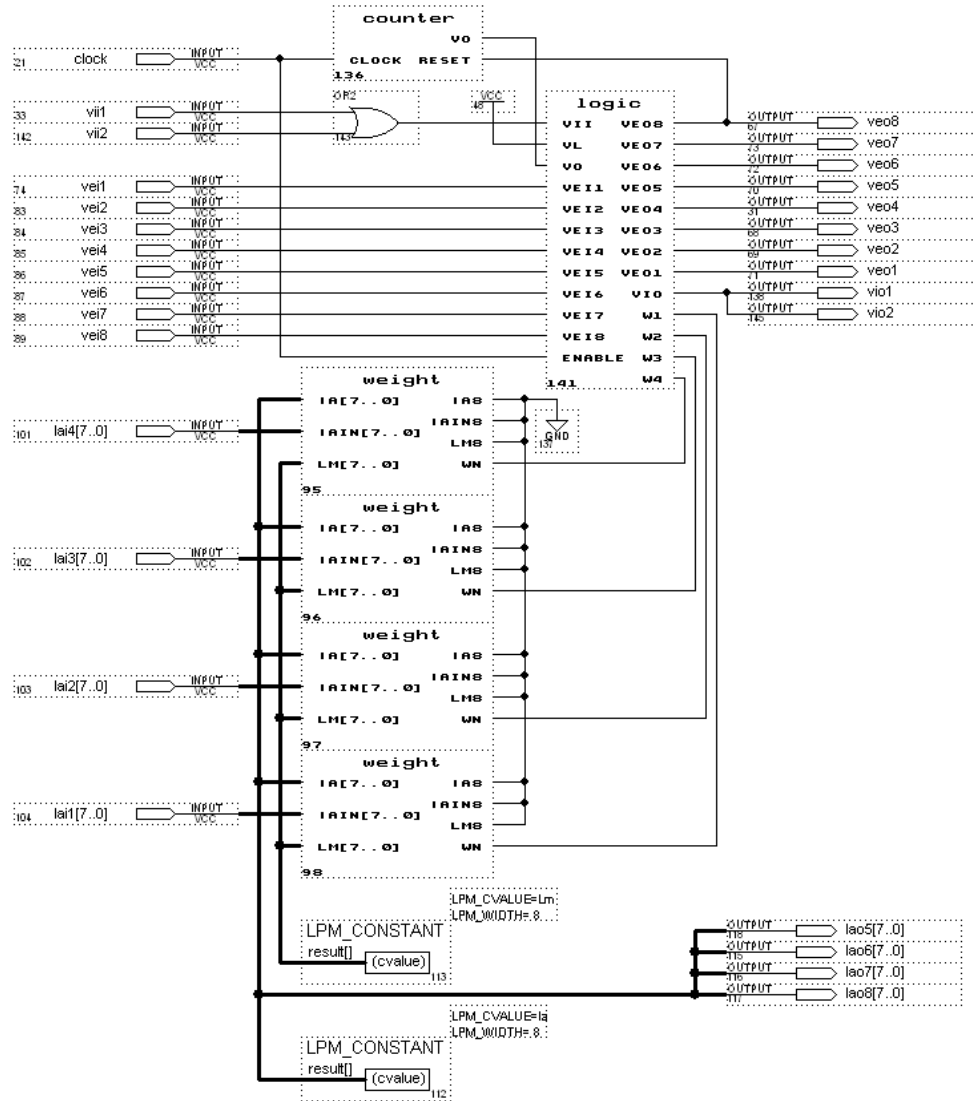


Figura 6.2: Estrutura interna do EP da rede ONNIS

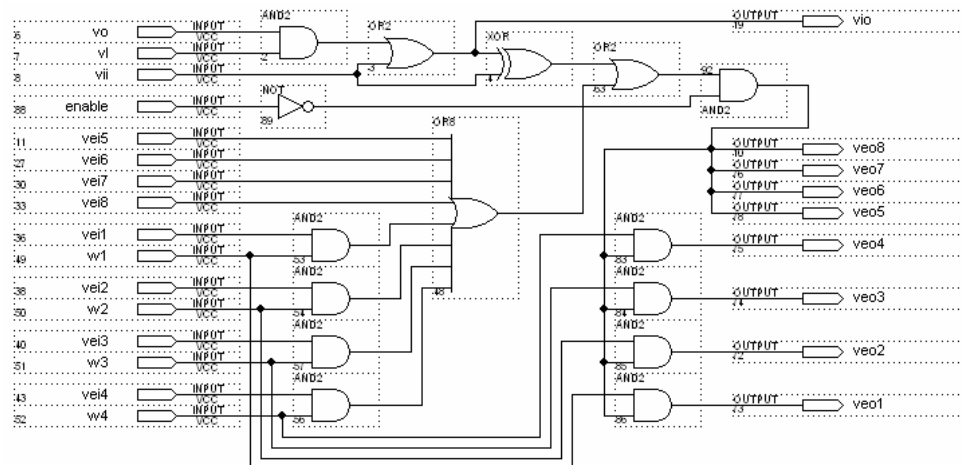


Figura 6.3: Lógica interna do EP da rede ONNIS

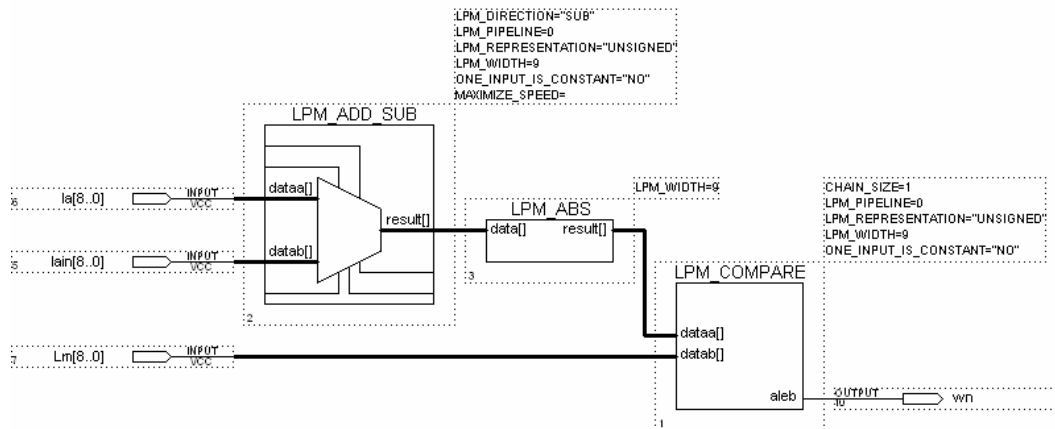


Figura 6.4: Etapa para determinação do peso excitatório do EP

Na implementação dos EPs foram utilizados contadores de 4 bits, ilustrados através da Figura 6.5, implicando contagem de 0 até 15 (16 estados). Para redes de maiores dimensões um maior número de estados será necessário, vinculando assim a complexidade do EP ao tamanho da rede (N_n).

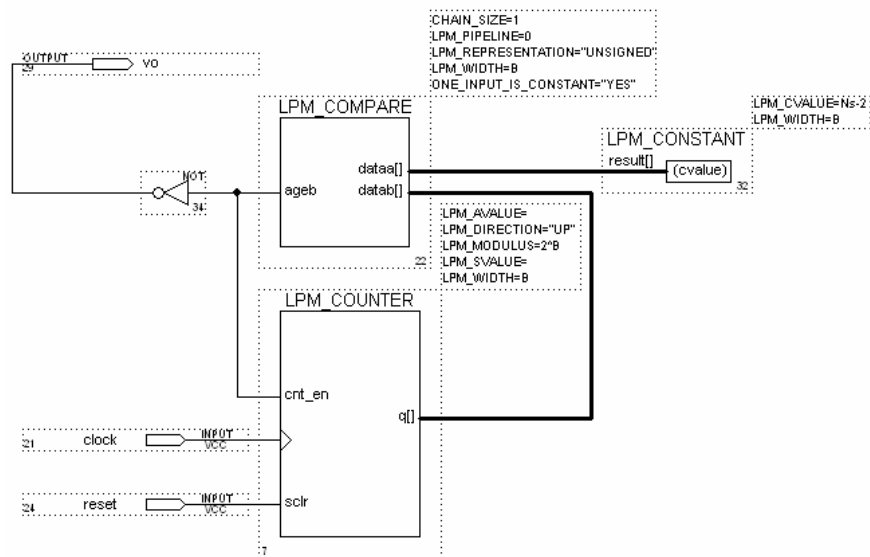


Figura 6.5: Contador interno do EP da rede ONNIS

Com a utilização das características de um dispositivo FPGA Altera EPF10K10LC84-3 na simulação obteve-se uma necessidade de 124 células lógicas para implementação da rede, equivalente à cerca de 21% da capacidade do *chip*. Com relação a este aspecto, deve ser lembrado que a rede implementada foi bastante simplificada devido ao fato de considerarem-se suas entradas como parâmetros de compilação.

Como exemplo, a Figura 6.6 representa uma pequena imagem de 3x5 pixels em níveis de cinza com os respectivos valores de intensidade de cada pixel (8 bits/pixel), a qual foi utilizada como entrada da rede. Considerando-se $L_w=20$, por exemplo, devem ser formados os grupos de pixels com níveis de cinza semelhantes apresentados na mesma figura. Tais grupos de pixels devem ser apresentados isoladamente e em tempos diferentes na saída da rede para caracterizar o correto funcionamento da mesma.

0	1	120	249	250
2	122	121	252	251
3	4	5	253	6

Figura 6.6: Imagem de entrada da rede simulada no Max+plus II e implementada em FPGA

A Figura 6.7 apresenta as formas de onda (níveis lógicos 0 e 1) obtidas através da simulação, onde pode ser verificado o agrupamento dos EPs com entradas semelhantes e a separação temporal de cada um destes grupos. O parâmetro N_s deve ter um valor mínimo igual ao número de segmentos contidos na imagem. A utilização de valores superiores acarreta somente o aparecimento de períodos sem qualquer segmento na saída da rede. O resultado apresentado foi obtido com $N_s=6$. Como todos os EPs partiram do estado zero, o primeiro segmento é apresentado na saída no sexto pulso de *clock*. A partir daí, os quatro segmentos são apresentados seqüencialmente na saída, existindo dois períodos de *clock* (N_s menos 4 segmentos) sem qualquer saída ativa entre cada seqüência. A saída *seg* refere-se à saída inibitória do EP com menor prioridade e pode ser utilizada como um indicador de presença de um segmento ativo na saída da rede. Este sinal, juntamente com o *clock*, pode ser utilizado para transferir dados referentes aos segmentos encontrados da saída da rede para algum dispositivo externo.

A implementação da rede apresentada em um *chip* FPGA comprovou os resultados previstos através do simulador. Para verificar o estado das saídas dos EPs foram conectados LEDs a cada uma delas, tendo sido os mesmos acionados de forma análoga às formas de onda apresentadas na Figura 6.7.

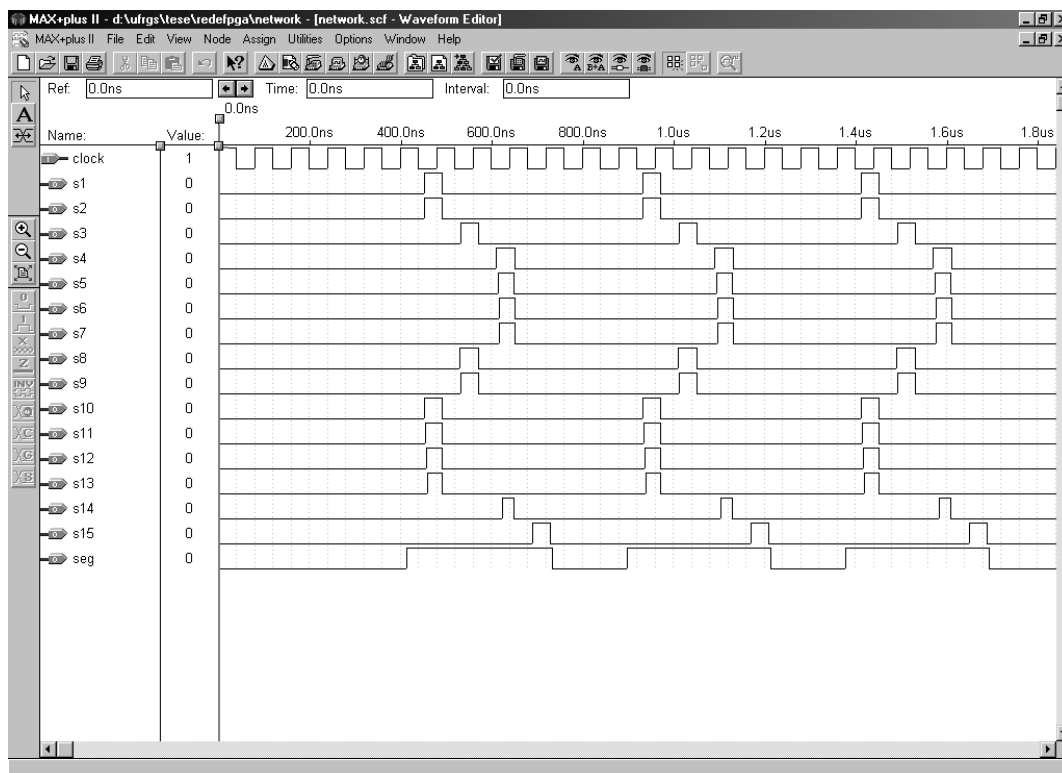


Figura 6.7: Resultado da simulação da rede ONNIS no Max+plus II

6.2 Rede ONNIS-GI

O procedimento de inicialização global implica basicamente a modificação do contador vinculado ao EP da rede. Ao invés de serem utilizados contadores como os da Figura 6.5, foram utilizados flip-flops JK, conforme ilustrado através da Figura 6.8. A estrutura interna do EP da rede com inicialização global é apresentada na Figura 6.9.

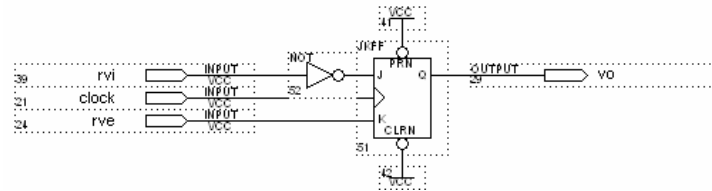


Figura 6.8: Contador interno do EP da rede ONNIS-GI

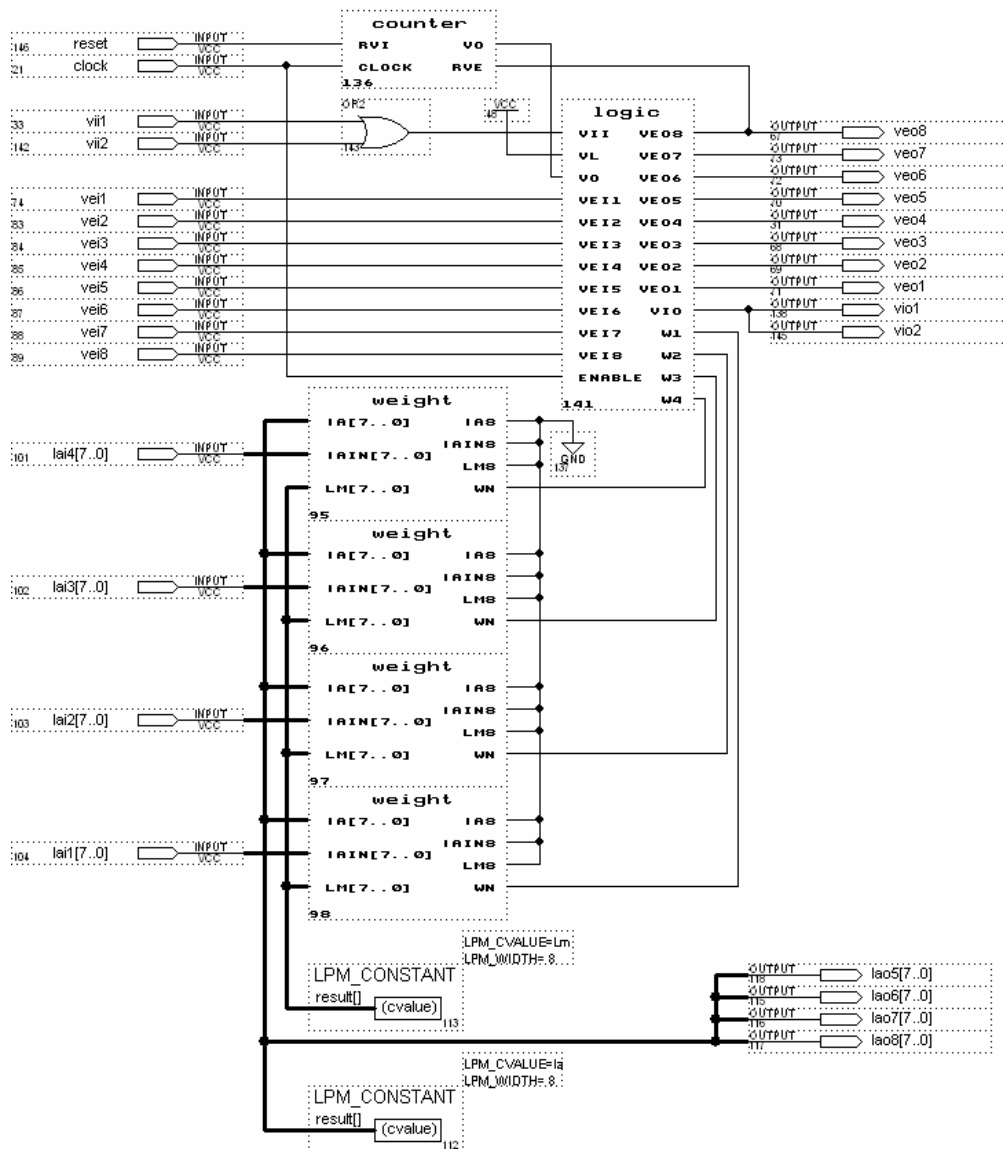


Figura 6.9: Estrutura interna do EP da rede ONNIS-GI

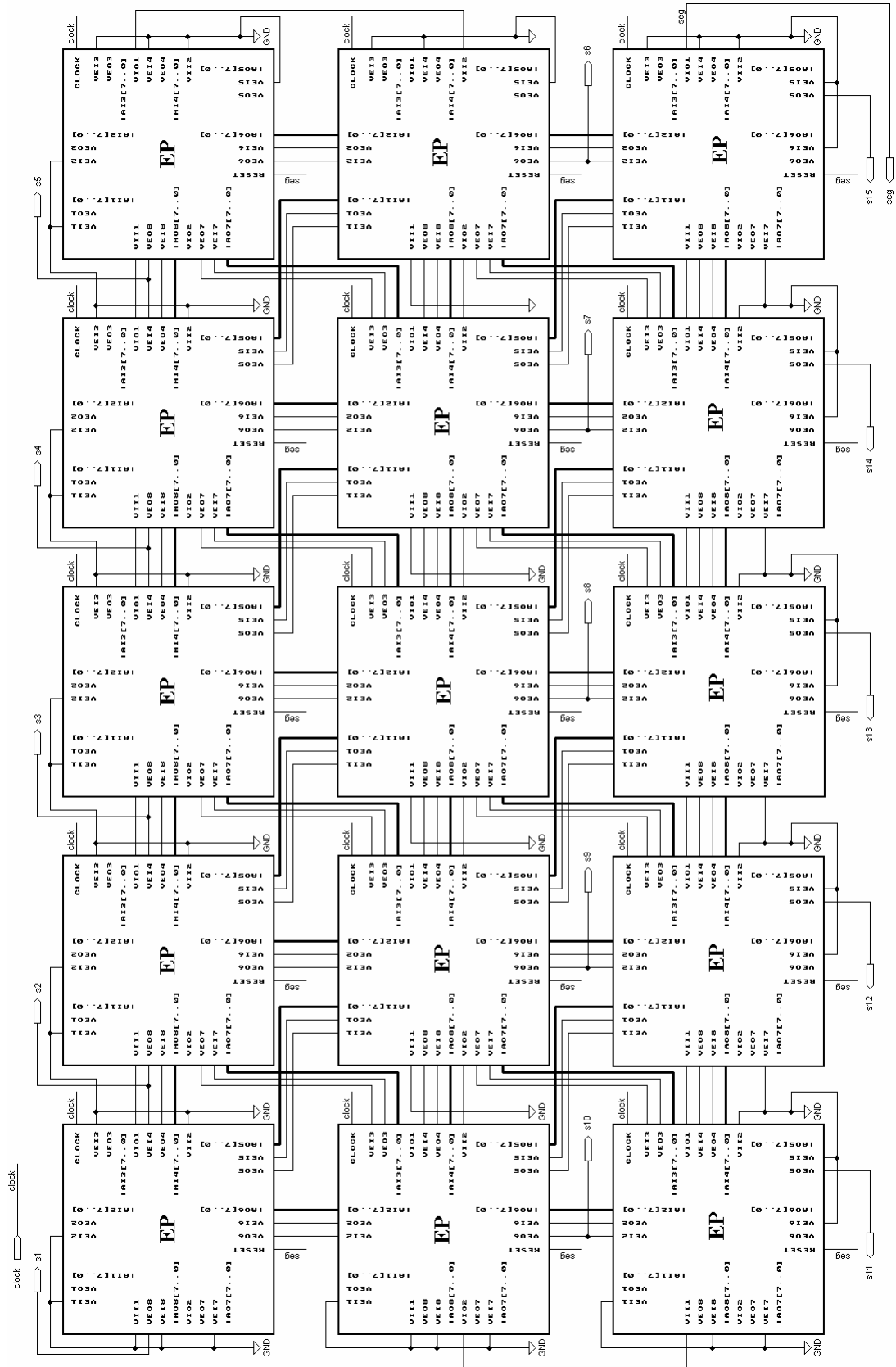


Figura 6.10: Implementação da rede da rede ONNIS-GI

A lógica combinacional interna do EP da rede com inicialização global permanece como na Figura 6.3, enquanto que a estrutura utilizada para determinação dos pesos é a mesma da Figura 6.4. A Figura 6.10 apresenta a rede de 3x5 EPs ($N_n=15$) implementada utilizando o procedimento de inicialização global.

Da mesma forma que a rede apresentada na seção 6.1, a rede possui como entrada um *clock*, comum a todos os EPs, e, como saída, os sinais excitatórios dos 15 neurônios (S_1, S_2, \dots, S_{15}). Para simplificação do *hardware*, o valor de L_w e as entradas referentes às intensidades dos pixels foram considerados parâmetros da rede na etapa de compilação, o que também não constitui qualquer alteração quanto à operação da rede. Por outro lado, o parâmetro N_s não se faz mais necessário, decorrência da utilização do procedimento de inicialização global.

A Figura 6.11 apresenta as formas de onda obtidas através da simulação utilizando-se como entrada a imagem da Figura 6.6. A inibição global implica o surgimento de somente uma imagem em branco na saída da rede a cada ciclo, uma delas observada cerca de 450 ns após o início do processo. Com todos os EPs partindo do estado zero, o primeiro segmento é apresentado na saída já no segundo pulso de *clock*. A partir daí, os quatro segmentos são apresentados sequencialmente na saída, existindo um período de *clock* sem qualquer saída ativa entre cada seqüência. Já sem a utilização de inibição global, com parâmetro $N_s = 6$, são obtidas duas imagens nulas na saída da rede, como pode ser observado cerca de 800 ns após o início do processo (Figura 6.7).

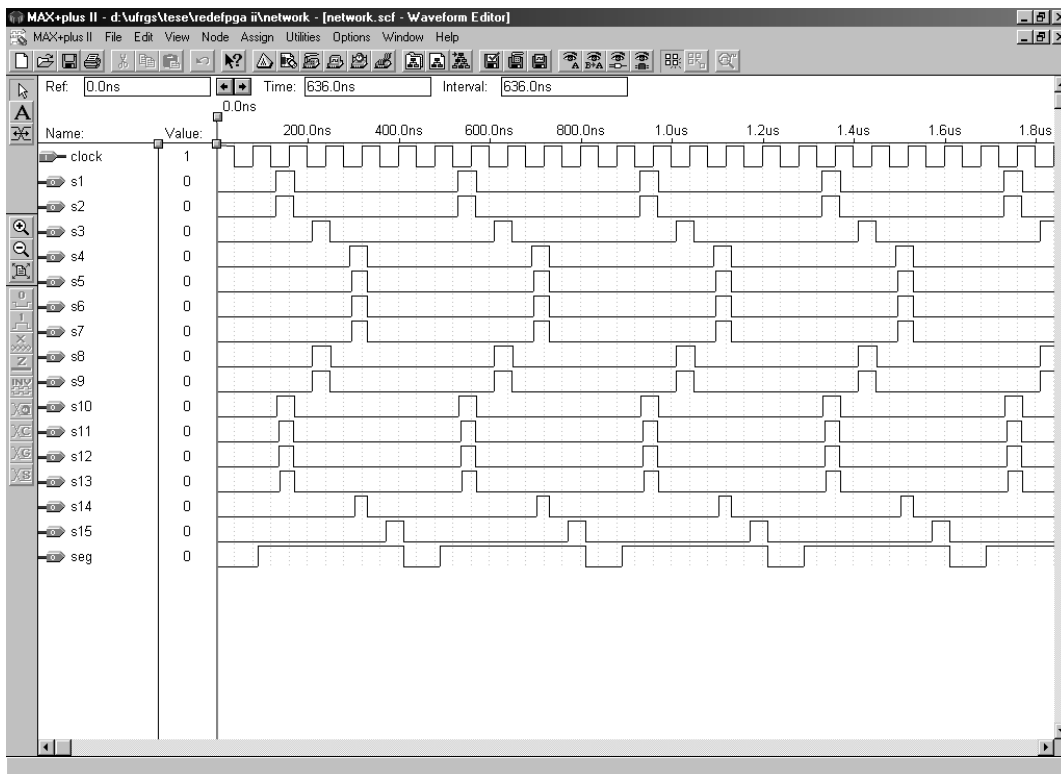


Figura 6.11: Resultado da simulação da rede ONNIS-GI

A inibição global não implica que a complexidade do EP seja uma função do tamanho da rede (N_n), reduzindo portanto a complexidade do *hardware* para implementação da mesma. Com a utilização das características de um dispositivo FPGA Altera EPF10K10LC84-3 na simulação, obteve-se uma necessidade de 48 células

lógicas para implementação da rede, equivalente à cerca de 8% da capacidade do *chip*. Comparando-se este resultado com aquele obtido sem a utilização de inibição global pode-se constatar uma redução significativa da complexidade do *hardware*. Devem também ser ressaltado que esta redução de complexidade será proporcionalmente tanto maior, quanto maior for o número de EPs da rede.

6.3 Estimativa da Velocidade de Processamento

A velocidade de processamento das arquiteturas das duas versões de rede propostas é função do período do *clock* utilizado, o qual deve ser suficientemente grande para possibilitar a ativação integral de todos os EPs do maior segmento presente na imagem de entrada. Considerando que o número de portas lógicas entre a entrada excitatória e a saída excitatória do EP é maior do que o número de portas lógicas entre a entrada inibitória e a saída inibitória, o atraso na ativação do EP ao receber uma entrada excitatória é o que determina o período mínimo do *clock* (ver Figura 4.2 e Figura 4.7). Desta forma, quanto maior a cadeia de EPs a serem sequencialmente ativados, o que deve ocorrer durante o semiciclo baixo do *clock*, menor a frequência máxima do *clock* e menor a velocidade de processamento da rede.

Para uma aplicação geral, deve ser determinado o segmento de maior comprimento que pode ocorrer na imagem de entrada. Com base neste segmento, pode ser então estabelecida a máxima frequência do *clock* da rede. A Figura 6.12 apresenta um exemplo típico de um segmento de comprimento máximo no qual a ativação dos EPs é realizada de forma sequencial, sendo que o EP de menor prioridade sofrerá o maior atraso em relação à ativação do EP de maior prioridade.

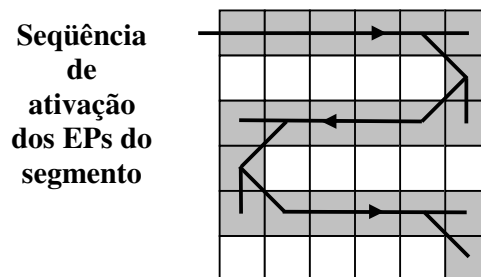


Figura 6.12: Exemplo típico de segmento com máximo atraso para ativação completa

Generalizando a situação apresentada na Figura 6.12 para uma rede de dimensões arbitrárias, com $N_x \geq N_y$, a equação (6.1) determina o tempo necessário (t_s) para que a rede proposta apresente um segmento de comprimento máximo em sua saída a partir borda de descida do *clock*. O tempo em questão é função basicamente do atraso para ativação de um EP ao receber uma entrada excitatória ($4t_d$) e do segmento de maior comprimento que pode ocorrer em uma imagem de dimensões $N_x N_y$ pixels. Também é considerado o tempo que o EP de maior prioridade tem sua saída excitatória ativada após a borda de descida do *clock* ($2t_d$). Uma parcela de tempo adicional (t_m) considera o tempo mínimo necessário para manutenção do segmento completo ativado na saída da rede, o que se faz necessário para possibilitar a transferência adequada da imagem de saída para um estágio posterior de processamento. Assim, com base no atraso determinado pela equação (6.1), um *clock* de onda quadrada (*duty cycle* igual a 50%) terá frequência máxima igual a $1/(2t_s)$. Tendo em vista o menor atraso na ativação das saídas inibitórias dos EPs, a possibilidade de utilização de um sinal de *clock* com *duty*

cycle menor do que 50% pode possibilitar maiores frequências de oscilação e maior velocidade de processamento da rede.

$$t_s = \text{int}\left(\frac{N_y + 1}{2}\right)(N_x - 1)4t_d + 2t_d + t_m, \quad N_x \geq N_y \quad (6.1)$$

onde:

- N_x é o número de colunas da rede ou da imagem;
- N_y é o número de linhas da rede ou da imagem;
- t_s é o atraso para ativação completa do segmento (em segundos);
- t_m é o tempo para manutenção da saída ativa (em segundos) e
- t_d é o atraso máximo de uma porta lógica em segundos (em segundos).

A velocidade de processamento da arquitetura proposta é função, além do número total de EPs (N_n) e dos atrasos lógicos (t_d), também do número de regiões a serem segmentadas (N_s). Assim, considerando N_s segmentos presentes na imagem de entrada, a rede poderia segmentar até $1/(2t_s(N_s+1))$ fps (*frames per second*), onde um *frame* corresponde a uma imagem de entrada ou N_s imagens (segmentos) de saída. Por exemplo, considerando um atraso de 1ns por porta lógica (t_d) e t_m da mesma ordem, uma rede com 32x32 EPs ($N_x=32$ e $N_y=32$), e 10 segmentos na imagem de entrada ($N_s=10$), poderia segmentar até 22876 fps (22876 imagens de entrada ou 228760 segmentos de saída por segundo). Ainda que carecendo de maiores detalhes para uma comparação adequada com outras propostas, em (KYUMA et al., 1999) é mencionada uma retina artificial para reconhecimento de imagens de 32x32 pixels com capacidade de processar 500 fps.

6.4 Conclusão

Os resultados das simulações realizadas com o Max+plus II e a implementação em um *chip* FPGA demonstraram que as redes ONNIS e ONNIS-GI operam segundo a forma desejada, agrupando corretamente as regiões cujos pixels têm características semelhantes e apresentam seqüencialmente de forma isolada cada grupo homogêneo em sua saída.

Tendo em vista sua maior complexidade, a rede ONNIS necessita de maior quantidade de recursos para implementação em *hardware* do que a rede ONNIS-GI. Esta diferença de complexidade é tanto maior quanto maior for o número de EPs na rede, pois a estrutura do EP sem inicialização global tem complexidade dependente do tamanho da rede, ao contrário do EP com inicialização global.

Na rede ONNIS foi observada a presença de várias imagens em branco em cada ciclo de segmentação, em número dependente do valor de N_s . Já na rede ONNIS-GI, onde N_s não é utilizado, é observada somente uma imagem nula na saída durante cada ciclo de segmentação. Esta imagem pode ser utilizada como um indicador externo de que todos os segmentos já foram extraídos pela rede. Neste caso, também o número de segmentos encontrados pela rede pode ser facilmente determinado através da contagem do número de pulsos de *clock* entre cada duas imagens em branco consecutivas, desde que a imagem de entrada e L_w sejam mantidos constantes no intervalo.

Uma avaliação da velocidade de processamento da rede, a qual é função basicamente do atraso imposto pela lógica combinacional utilizada nos EPs, do número de EPs e dos número de segmentos na imagem de entrada, indica que podem potencialmente ser obtidos resultados bem superiores aos encontrados na bibliografia.

Esta análise, contudo, deve ser tratada com alguma restrição, pois as referências consultadas não fornecem dados suficientes para uma comparação mais apurada.

Em todas as simulações realizadas foi possível confirmar a facilidade de sincronismo entre EPs e a necessidade de poucas iterações (pulsos de *clock*) e em número facilmente previsível para obter o resultado desejado. A implementação de um pequeno protótipo em um chip FPGA comprovou na prática os resultados obtidos nas simulações.

7 CONCLUSÃO E DESDOBRAMENTO DO TRABALHO

No presente capítulo é apresentada uma conclusão geral a respeito do trabalho realizado. Tendo em vista os resultados positivos obtidos, também é realizada uma avaliação sobre o desdobramento da pesquisa em trabalhos futuros. Finalmente, uma previsão das publicações que podem ainda ser obtidas a partir do presente trabalho é também apresentada.

7.1 Conclusão Geral

Como resultado do presente estudo foi desenvolvida uma rede neural para segmentação de imagens, em duas versões, as quais podem ser implementadas em arquitetura maciçamente paralela utilizando tecnologia digital, sendo particularmente adequadas à implementação de *chips* de visão.

O desenvolvimento das pesquisas relativas ao tema de tese implicou inicialmente o estudo de procedimentos de segmentação de imagens, de aplicações de redes neurais nesta área, com ênfase nas redes neurais oscilatórias devido à sua natureza maciçamente paralela, e de *chips* de visão empregando também arquiteturas maciçamente paralelas.

Durante o estudo da aplicação de redes neurais em segmentação de imagens foi identificada uma nova abordagem, as denominadas redes neurais oscilatórias, as quais têm apresentado resultados promissores e são relativamente ainda pouco pesquisadas. Dentro deste contexto, a denominada rede LEGION é a proposta mais consolidada, tendo sido encontrados vários estudos sobre sua aplicação em segmentação de imagens e implementação em *hardware* dedicado.

Apesar dos resultados estimulantes obtidos, a rede LEGION apresenta algumas limitações de ordem prática relativas ao modelo complexo para implementação em *hardware* digital, à sua capacidade de segmentação limitada e ao elevado número de parâmetros e seu ajuste pouco prático.

Foram desenvolvidas algumas aplicações práticas utilizando a rede LEGION, em sua versão algorítmica, para segmentação de imagens. Os resultados obtidos foram bastante interessantes, o que conduziu à utilização desta rede como modelo para o desenvolvimento de uma nova rede neural para segmentação de imagens.

O desenvolvimento de *chips* de visão ou retinas de silício apresenta-se também como uma área de pesquisa nova e ainda com poucos resultados práticos. Nestes *chips*, o sensor de imagem propriamente dito é integrado em conjunto com uma rede maciçamente paralela de processadores para possibilitar um processamento muito rápido da imagem captada. Foram encontradas propostas para implementação de *chips* de visão utilizando tecnologia analógica ou digital. A utilização de tecnologia digital proporciona maior flexibilidade para reprogramação da função executada pelo *chip*, o que conduziu a presente pesquisa na direção de obter uma rede facilmente implementável com dispositivos digitais.

Com base no princípio de funcionamento da rede LEGION, a qual é inspirada no comportamento dos neurônios do córtex visual humano, foi então desenvolvida uma nova rede neural em duas versões. As redes desenvolvidas, denominada ONNIS e ONNIS-GI, apresentam capacidade de segmentar a imagem de entrada e de separar os segmentos no tempo e possuem arquitetura maciçamente paralela com estrutura facilmente implementável utilizando dispositivos digitais. Possuem poucos parâmetros e de ajuste intuitivo, apresentando também facilidade para implementação de procedimentos de segmentação de imagens mais elaborados como, por exemplo, a utilização de atributos de textura.

A rede ONNIS-GI, baseada na utilização de inibição global, apresentou maiores vantagens de ordem prática ao custo de uma conexão global adicional entre os EPs da rede.

Resultados de simulações da rede proposta em computadores PC utilizando um aplicativo desenvolvido em ambiente Delphi indicam que a mesma alcança seus objetivos iniciais em segmentação de imagens.

Para a avaliação da qualidade da segmentação obtida em relação à presença de ruído na imagem de entrada da rede foi estabelecido um critério de avaliação que considera o percentual de pixels corretamente classificados para um determinado segmento. Através da realização de uma série de simulações concluiu-se que a rede proposta apresenta uma certa imunidade ao ruído a qual se dá através do criterioso ajuste do parâmetro L_w . A utilização de um procedimento de filtragem inspirado no detector de bordas de Sobel conduziu a melhores resultados na presença de níveis mais altos de ruído na imagem de entrada. Foi também mostrado que a rede proposta produz resultados, em termos da qualidade de segmentação, equivalentes àqueles obtidos com o clássico método Crescimento de Regiões utilizando agregação de pixels e o critério da diferença absoluta entre pixels vizinhos.

A simulação utilizando o programa Max+plus II da Altera e a implementação de protótipo em FPGA ratificaram os resultados obtidos nas simulações anteriores. Foi verificado que a rede ONNIS-GI utiliza menos recursos de *hardware* o que, dependendo do tamanho da rede, pode implicar uma diferença de complexidade bastante significativa.

Uma avaliação preliminar leva a conclusão de que a nova rede pode apresentar velocidade de processamento bastante elevada. Uma comparação mais detalhada com outras propostas não foi possível por carência de detalhes nas referências consultadas.

Finalmente, com base nos resultados obtidos nas simulações e na implementação em FPGA conclui-se que a rede ONNIS e principalmente a rede ONNIS-GI apresentam resultados estimulantes em segmentação de imagens e diversas vantagens de ordem prática sobre a rede LEGION. Assim, verifica-se que o desenvolvimento de estudos visando o aprimoramento da rede proposta e de aplicações práticas para a mesma constitui-se em um campo de trabalho promissor.

7.2 Desdobramento da Pesquisa

O resultados obtidos com as simulações da rede proposta criaram algumas expectativas quanto a estudos abordando aplicações práticas da mesma, minimização de sua sensibilidade ao ruído, segmentação baseada em textura e implementação física de *chips* de visão.

A versão algorítmica da rede ONNIS-GI pode ser utilizada como parte de um sistema mais sofisticado de visão artificial, segmentando objetos de imagens para sua posterior identificação ou quantificação. Desta forma, a utilização da rede em aplicações

práticas de segmentação de imagens é um dos campos que se apresenta com potencial para o desenvolvimento de pesquisas adicionais.

O estudo de procedimentos para minimização da sensibilidade da rede ao ruído presente na imagem de entrada é outro ponto que se apresenta como objeto de estudos adicionais.

A segmentação baseada em textura é um campo de pesquisa atraente. A utilização de filtros de Gabor para caracterizar as diversas texturas presentes em uma imagem tem alguns aspectos práticos ainda não formalmente resolvidos. O exemplo apresentado neste trabalho indica que a utilização de filtros de Gabor para a determinação dos pesos da rede proposta pode conduzir a um procedimento de segmentação baseado em textura bastante interessante e portanto merecendo atenção adicional.

Finalmente, sendo uma das motivações iniciais para o desenvolvimento do tema de tese, a implementação física de *chips* de visão utilizando a rede proposta reveste-se de um atrativo especial. A idéia aqui é a de implementar fisicamente um *chip*, integrando sensores do tipo CCD, por exemplo, com a rede proposta. Desta forma, o sensor fornecerá em sua saída diretamente os segmentos da imagem captada com velocidade bastante elevada.

7.3 Artigos a Serem Elaborados a Partir da Tese

O Anexo 1 apresenta os cinco artigos publicados durante o desenvolvimento do presente trabalho. Estes artigos incluem aplicações desenvolvidas para a rede LEGION e a concepção da rede proposta, não esgotando porém a possibilidade de novas publicações baseadas em aspectos não abordados nos artigos já publicados.

Dentre os possíveis temas para publicações baseadas no trabalho de tese desenvolvido, pode ser citada a concepção e utilização do procedimento de inicialização global e a avaliação das vantagens obtidas em relação à rede inicialmente concebida.

A utilização da rede ONNIS-GI em conjunto com filtros de Gabor pode também servir como base para uma publicação interessante, abordando o tema de segmentação baseada em textura.

Outra possibilidade de publicação refere-se à avaliação objetiva do desempenho da rede proposta na presença de ruído.

Finalmente, uma publicação direcionada à implementação da rede ONNIS-GI em FPGA, mesmo implicando protótipos de pequenas dimensões com a tecnologia atualmente disponível, deve apresentar interesse de ordem prática.

REFERÊNCIAS

AHMED, M. N.; FARAG, A. A. Two-Stage Neural Network for Volume Segmentation of Medical Images. In: INTERNATIONAL CONFERENCE ON NEURAL NETWORKS, 1997. **Proceedings...** Houston: IEEE, 1997. v.3, p.1373-1378.

ALTERA. **MAX+PLUS II Development Tools**. Disponível em: <<http://www.altera.com/literature/lit-mp2.jsp>>. Acesso em: mar. 2004.

ANGELO, N. P. **Aplicação de Filtros de Gabor no Processo de Classificação de Imagens Digitais com Base em Atributos de Textura**. 2000. 137 f. Dissertação (Mestrado) - Centro Estadual de Pesquisa em Sensoriamento Remoto, UFRGS, Porto Alegre.

BABAGUCHI, N. et al. Connectionist model binarization. In: INTERNATIONAL CONFERENCE ON PATTERN RECOGNITION, 10., 1990. **Proceedings...** Atlantic City: IEEE, 1990. v.2, p.51-56.

BARBARO, M. et al. 100x100 Pixel Silicon Retina for Gradient Extraction with Steering Filter Capabilities and Temporal Output Coding. **IEEE Journal of Solid-State Circuits**, Pisacataway, NJ, v.37, n.2, p.160-172, Feb. 2002.

BLANZ, W. E.; GISH, S. L. A connectionist classifier architecture applied to image segmentation. In: INTERNATIONAL CONFERENCE ON PATTERN RECOGNITION, 10., 1990. **Proceedings...** Atlantic City: IEEE, 1990. v.2, p.272-277.

BOUBEKEUR, A.; SAUCIER, G. A Silicon Compiler for Massively Parallel Image Processing ASICs. In: SYMPOSIUM ON THE FRONTIERS OF MASSIVELY PARALLEL COMPUTATION, 3., 1990. **Proceedings...** College Park: IEEE, 1990. p.519-524.

CAMPBELL, N. W.; THOMAS, B. T.; TROSCIANKO, T. Segmentation of Natural Images Using Self-Organizing Feature Maps. In: BRITISH MACHINE VISION CONFERENCE, 1996. **Proceedings...** Edinburgh: The British Machine Vision Association, 1996. p.223-232.

CAMPBELL, N. W.; THOMAS, B. T.; TROSCIANKO, T. Neural Networks for the Segmentation of Outdoor Images. In: INTERNATIONAL CONFERENCE ON ENGINEERING APPLICATION OF NEURAL NETWORKS, 1996. **Proceedings...** London: System Engineering Association, 1996. p.343-346.

CAMPBELL, N. W.; THOMAS, B. T.; TROSCIANKO, T. Automatic Segmentation and Classification of Outdoor Images Using Neural Networks. **International Journal of Neural Systems**, [S. l.], v.8, n.1, p.137-144, Feb. 1997.

CHALANA, V.; KIM, Y. A Methodology for Evaluation of Image Segmentation Algorithms on Medical Images. In: SPIE CONFERENCE ON MEDICAL IMAGING, 1996. **Proceedings...** Newport Beach: SPIE, 1996. v. 2710.

CHEN, C. T.; TSAO, E. C.; LIN, W. C. Medical image segmentation by a constraint satisfaction neural network. **IEEE Transactions on Nuclear Sciences**, Piscataway, v.38, n.2, 678-686, Apr.1991.

CHEN, K.; WANG, D.; LIU, X. Weight Adaptation and Oscillatory Correlation for Image Segmentation. **IEEE Transactions on Neural Networks**, Piscataway, NJ, v.11, n.5, p.1106-1123, Sept. 2000.

CHEN, K.; WANG, D. Image Segmentation Based on a Dynamically Coupled Neural Oscillator Network. In: INTERNATIONAL JOINT CONFERENCE ON NEURAL NETWORKS, 1999. **Proceedings...** Washington, DC: IEEE, 1999. v.4, p.2653-2658.

CORTES, C.; HERTZ, J. A. A network system for image segmentation. In: INTERNATIONAL JOINT CONFERENCE ON NEURAL NETWORKS, 1989. **Proceedings...** Washington, DC: IEEE, 1989. v.1, p.121-125.

DAHMER, A. **Segmentação de Imagens Ecocardiográficas Utilizando Redes Neurais e Medidas de Textura**. 1998. 98 f. Dissertação (Mestrado em Ciência da Computação) – Instituto de Informática, UFRGS, Porto Alegre.

DAS, M.; ANAND, J. Robust Edge Detection in Noisy Images Using an Adaptive Stochastic Gradient Technique. In: INTERNATIONAL CONFERENCE ON IMAGE PROCESSING, 1995. **Proceedings...** Washington, DC: IEEE, 1995. v.2, p.149-152.

DAUGMAN, J. G. High Confidence Visual Recognition of Persons by a Test of Statistical Independence. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, Piscataway, v.15, n.11, p.1148-1161, Nov. 1993.

DOWNTON, A.; CROOKES, D. Parallel Architectures for Image Processing. **IEEE Electronics & Communication Engineering Journal**, Piscataway, v.10, n.3, p.139-151, 1998.

FARROHA, B. S.; DESHMUKH, R. G. A Novel Approach to Design a Massively Parallel Application Specific Architecture for Image Recognition Systems. In: SOUTHEASTCON VISUALIZE THE FUTURE, 1995. **Proceedings...** Raleigh: IEEE, 1995. p. 293-299.

FERNANDES, D. **Segmentação de Imagens Baseada em Redes Neurais Oscilatórias**. 2001. 65 f. Trabalho Individual (Doutorado em Ciência da Computação) - Instituto de Informática, UFRGS, Porto Alegre.

FERNANDES, D.; SIQUEIRA, M.; NAVAU, P. O. A. Segmentation of Fetal Echocardiographic Images Using Self-Organizing Maps and Oscillatory Neural

Networks. In: IBEROAMERICAN SYMPOSIUM ON PATTERN RECOGNITION, 6., 2001. **Proceedings...** Florianópolis: IAPR, 2001.

FERNANDES, D.; NAVAU, P. O. A.; FICHTNER, P. F. P. Segmentation of TEM Images Using Oscillatory Neural Networks. In: BRAZILIAN SYMPOSIUM ON COMPUTER GRAPHICS AND IMAGE PROCESSING, 14., 2001. **Proceedings...** Florianópolis: IEEE, 2001. p.289-296.

FERNANDES, D.; STEDILE, J. P.; NAVAU, P. O. A. Architecture of Oscillatory Neural Network for Image Segmentation. In: SYMPOSIUM ON COMPUTER ARCHITECTURE AND HIGH PERFORMANCE COMPUTING, 14., 2002. **Proceedings...** Vitória: IEEE, 2002. p.29-36.

FERNANDES, D.; NAVAU, P. O. A. Rede Neural Maciçamente Paralela para Segmentação de Imagens. **Cadernos de Informática UFRGS**, v.3, n.1, p.151-156, jun. 2003.

FERNANDES, D.; NAVAU, P. O. A. **An Oscillatory Neural Network for Image Segmentation**. Berlin: Springer, 2003. p.667-674. (Lecture Notes in Computer Science - Progress in Pattern Recognition, Speech and Image Analysis, v.2905)

GHOSH, A.; PAL, N. R.; PAL, S. K. Image segmentation using a neural network. **Biologic Cybernetics**, [S. l.], v.66, p.151-158, 1991.

GHOSH, A.; PAL, N. R.; PAL, S. K. Object extraction using a self-organizing neural network. **Intelligent Robotics**. New Delhi: McGraw-Hill, 1991.

GHOSH, A.; PAL, N. R.; PAL, S. K. Object background classification using Hopfield type neural network. **International Journal of Pattern Recognition and Artificial Intelligence**, River Edge, v.6, n.5, p.989-1008, 1992.

GONZALES, R. C.; WOODS, R. E. **Processamento de Imagens Digitais**. São Paulo: Edgard Blücher, 2000. 509p.

HARALICK, R. M. Statistical and structural approaches to texture. In: YOUNG, Tzai Y.; FU, King-Sun. **Handbook of Pattern Recognition and Image Processing**. San Diego: Academic Press, 1986. p.247-79.

HAYKIN, S. **Redes Neurais - Princípios e prática**. 2.ed. São Paulo: Bookman, 1999. 900p.

HEATH, M. D. et al. A Robust visual Method for Assessing the Relative Performance of Edge-Detection Algorithms. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, Pisacataway, NJ, v.19, n.12, p.1338-1359, Dec.1997.

HUANG, Q.; DOM, B. Quantitative Methods of Evaluating Image Segmentation. In: INTERNATIONAL CONFERENCE ON IMAGE PROCESSING, 1995. **Proceedings...** Washington, DC: IEEE, 1995. v.3. p.53-56.

JAIN, A. K. **Fundamentals of Digital Image Processing**. Englewood Cliffs: Prentice-Hall, 1989. 569p.

JU, C.; MOLONEY, C. An Edge-Enhanced Modified Lee Filter for the Smoothing of SAR Image Speckle Noise. In: IEEE WORKSHOP ON NONLINEAR SIGNAL AND IMAGE PROCESSING, 1997. **Proceedings...** Michigan: IEEE, 1997.

KOMURO, T.; ISHII, I.; ISHIKAWA, M. Digital Vision Chips and High-Speed Vision Systems. In: SYMPOSIUM ON VLSI CIRCUITS DIGEST OF TECHNICAL PAPERS, 2001. **Proceedings...** Kyoto: IEEE, 2001. p.1-4.

KOMURO, T.; ISHII, I.; ISHIKAWA, M. A New Architecture of Programmable Digital Vision Chip. In: SYMPOSIUM ON VLSI CIRCUITS DIGEST OF TECHNICAL PAPERS, 2002. **Proceedings...** Honolulu: IEEE, 2002. p.266-269.

KREYSZIG, E. **Matemática Superior**. Rio de Janeiro: Livros Técnicos e Científicos, 1982. v.1, p.168-174.

KYUMA, K. et al. Concept, Development, Mass Production, and Applications of Artificial Retina Chips. In: INTERNATIONAL CONFERENCE ON INTELLIGENT PROCESSING AND MANUFACTURING OF MATERIALS, 2., 1999. **Proceedings...** Honolulu: IEEE, 1999. p.1297-1303.

LIU, X.; CHEN, K.; WANG, D. L. Extraction of hydrographic regions from remote sensing images using an oscillator network with weight adaptation. **IEEE Transactions on GeoScience and Remote Sensing**, Pisacataway, NJ, v.39, n.1, p.207-211, Jan. 2001.

McCAULEY, J. D.; THANE, B. R.; WHITTAKER, A. D. Fat Estimation in Beef Ultrasound Images Using Texture and Adaptive Logic Networks. **Transactions of the American Society of Agricultural Engineers**, Saint Joseph, MI, v.37, n.3, p.997-1002, 1994.

MOHAMED, N. A.; AHMED, M. N.; FARAG, A. Modified fuzzy c-mean in medical image segmentation. In: IEEE INTERNATIONAL CONFERENCE ON ACOUSTICS, SPEECH, AND SIGNAL PROCESSING, 1999. **Proceedings...** Phoenix: IEEE, 1999. v.6, p.3429-3432.

PAILLET, F.; MERCIER, D.; BERNARD, T. M. Second Generation Programmable Artificial Retina. In: ANNUAL ASIC/SOC CONFERENCE, 20., 1999. **Proceedings...** Washington, DC: IEEE, 1999. p.304-309.

PAL, N. R.; PAL, S. K. A Review on Image Segmentation Techniques. **Pattern Recognition**, Oxford, v.26, n.9, p.1277-1294, 1993.

PANDYA, A. S.; MACY, Robert B. **Pattern Recognition with Neural Networks in C++**. Boca Raton: CRC Press, 1996.

PÉREZ-URIBE, A.; SANCHEZ, E. **FPGA Implementation of an Adaptable-Size Neural Network**. Berlim: Springer, 1996. p.383-388. (Lecture Notes in Computer Science, v.1112).

PICCOLI, L. **Segmentação e Classificação de Imagens Ecocardiográficas Utilizando Redes Neurais**. 1999. 90 f. Dissertação (Mestrado em Ciência da Computação) - Instituto de Informática, UFRGS, Porto Alegre.

- POLI, R.; VALLI, G. **Optimum Segmentation of Medical Images with Hopfield Neural Networks**. 1995. 11 f. Technical Report CSRP-95-12 - School of Computer Science, The University of Birmingham, Birmingham.
- RATHA, N. K.; JAIN, A. K. High Performance Custom Computing for Image Segmentation. In: HIGH PERFORMANCE COMPUTING CONFERENCE, 1995. **Proceedings...** New Delhi: [s. n.], 1995. p.67-72.
- RUSS, J. C. **The Image Processing Handbook**. 2nd ed. Boca Raton: CRC Press, 1995. 674p.
- SHAREEF, N.; WANG, D. L.; YAGEL, R. Segmentation of Medical Images Using LEGION. **IEEE Transactions on Medical Image**, Pisacataway, NJ, v.18, n.1, p.74-91, Jan. 1999.
- SHARIFI, M.; FATHY, M.; MARYAM, T. M. A Classified and Comparative Study of Edge Detection Algorithms. In: INTERNATIONAL CONFERENCE ON INFORMATION TECHNOLOGY, 2002. **Proceedings...** Las Vegas: IEEE, 2002. p.117-120.
- SHIN, M. C.; GOLDFOF, D. B.; BOWYER, K. W.; NIKIFOROU, S. Comparison of Edge Detection Algorithms Using a Structure from Motion Task. **IEEE Transactions on Systems, Man, and Cybernetics**, Pisacataway, NJ, v.31, n.4, p.589-601, Aug. 2001.
- Da SILVA, D. L. et al. The effects of implantation temperature on He bubble formation in silicon. **Nuclear Instruments and Methods in Physics Research B**, Amsterdam, 175-77, p.335-339, 2001.
- SIQUEIRA, M. L. et al. Echocardiographic Image Sequence Segmentation Using Self-Organizing Maps. In: IEEE WORKSHOP ON NEURAL NETWORKS FOR SIGNAL PROCESSING, 2000. **Proceedings...** Sydney: IEEE, 2000. p.594-603.
- SOMERS, D.; KOPELL, N. Rapid synchrony through fast threshold modulation. **Biologic Cybernetics**, [S. l.], v.68, p.393-407, 1993.
- STEMM, M. **Using Artificial Neural Networks and Image Segmentation to Assist in Mobile Robot Navigation**. 1994. Bachelors Thesis - School of Computer Science, Carnegie Mellon University, Pittsburgh, PA.
- STOCKER, A. et al. Stability Study of Some Neural Networks Applied to Tissue Characterization of Brain Magnetic Resonance Images. In: INTERNATIONAL CONFERENCE ON PATTERN RECOGNITION, 13., 1996. **Proceedings...** Vienna: IEEE, 1996. v.4, p.472-476.
- TERMAN, D.; WANG, D. Global competition and local cooperation in a network of neural oscillators. **Physica D**, Amsterdam, n.81, p.148-176, 1995.
- THURROTT, P. et al. **Delphi 3 SuperBible**. USA: Waite Group Press, 1997. 1312 p.
- VALVERDE, F. L. et al. An Evaluation Criterion for Edge Detection Techniques in Noisy Images. In: INTERNATIONAL CONFERENCE ON IMAGE PROCESSING, 2001. **Proceedings...** Thessaloniki: IEEE, 2001. v.1, p.7-10.

WANG, D. Modeling global synchrony in the visual cortex by locally coupled neural oscillators. In: ANNUAL CONFERENCE ON COGNITIVE SCIENCE SOCIETY, 15.,1993. **Proceedings...** [S. l.: s. n.], 1993. p.1058-1063.

WANG, D.; TERMAN, D. **Image Segmentation Based on Oscillatory Correlation**. 1996. Technical Report 19 - The Ohio State University Center for Cognitive Science, The Ohio State University, Columbus, USA.

WANG, D.; TERMAN, D. Image Segmentation Based on Oscillatory Correlation. **Neural Computation**, Cambridge, MA, v.9, p.805-836, May 1997.

WANG, D. Relaxation Oscillators and Networks. In: WEBSTER, J. G. (Ed.). **Wiley Encyclopedia of Electrical and Electronics Engineering**. New York: John Willey, 1999. v.18, p.396-405.

WANG, D.; BROWN, G. J. Separation of Speech from Interfering Sounds Based on Oscillatory Correlation. **IEEE Transactions on Neural Networks**, Pisacataway, NJ, v.10, n.3, p.684-697, Mar.1999.

WEBSTER'S. **Webster's Online Dictionary - The Rosetta Edition**. Disponível em: <<http://www.websters-online-dictionary.org/>>. Acesso em: abr. 2004.

YASNOFF, W. A.; MUI, J. K.; BACUS, J. W. Error Measures for Scene Segmentation. **Pattern Recognition**, Oxford, v.9, p.217-231, 1977.

ZHANG, Y. J. A Survey on Evaluation Methods for Image Segmentation. **Pattern Recognition**, Oxford, v.29, n.8, p.1335-1346, 1996.

GLOSSÁRIO

- Amostragem** - Seleção de valores de uma função contínua, de uma ou mais dimensões, para valores discretos de seu domínio, geralmente regularmente espaçados.
- Anti-Sincronismo** - O termo anti-sincronismo é empregado ao longo deste documento para indicar osciladores ou grupos de osciladores que estão ativos em intervalos de tempo distintos.
- Ground Truth** - Representa um conjunto de dados de referência para calibração de um instrumento ou de um processo.
- Nullclines** - Funções obtidas quando as derivadas de cada uma das variáveis de estado em relação ao tempo são igualadas a zero em um sistema de equações diferenciais. O termo em língua inglesa foi mantido tendo em vista não ter sido encontrado o termo equivalente em língua portuguesa nas fontes consultadas.
- Quantização** - Representação de uma grandeza contínua através de um conjunto de valores discretos em quantidade determinada pelo número de bits disponível.

APÊNDICE A APLICATIVOS DESENVOLVIDOS

Com o objetivo de avaliar o funcionamento e os resultados obtidos com a rede LEGION e com o algoritmo FPS na segmentação de imagens, foram desenvolvidos diversos programas aplicativos, todos em ambiente Delphi (THURROTT et al., 1997), os quais são apresentados neste apêndice.

A Figura APA.1 apresenta a interface gráfica do aplicativo para visualização do funcionamento do neurônio oscilador de Terman-Wang. São apresentados gráficos com a resposta temporal do oscilador e o plano de fase com as *nullclines* associadas. É possível ajustar o nível da entrada do oscilador, os parâmetros α , β e ε , além das condições iniciais $x(0)$ e $y(0)$. Ao ser iniciada a simulação, um ponto, inicialmente posicionado no estado inicial $x(0)$, movimenta-se sobre o gráfico da saída do oscilador em função do tempo, enquanto outro ponto correlacionado se desloca no plano de fase indicando o estado atual do oscilador. A solução do conjunto de equações diferenciais foi realizada através do método de Runge-Kutta (KREYSZIG, 1982), solução adotada em todos os demais aplicativos que utilizam o modelo contínuo da rede LEGION.

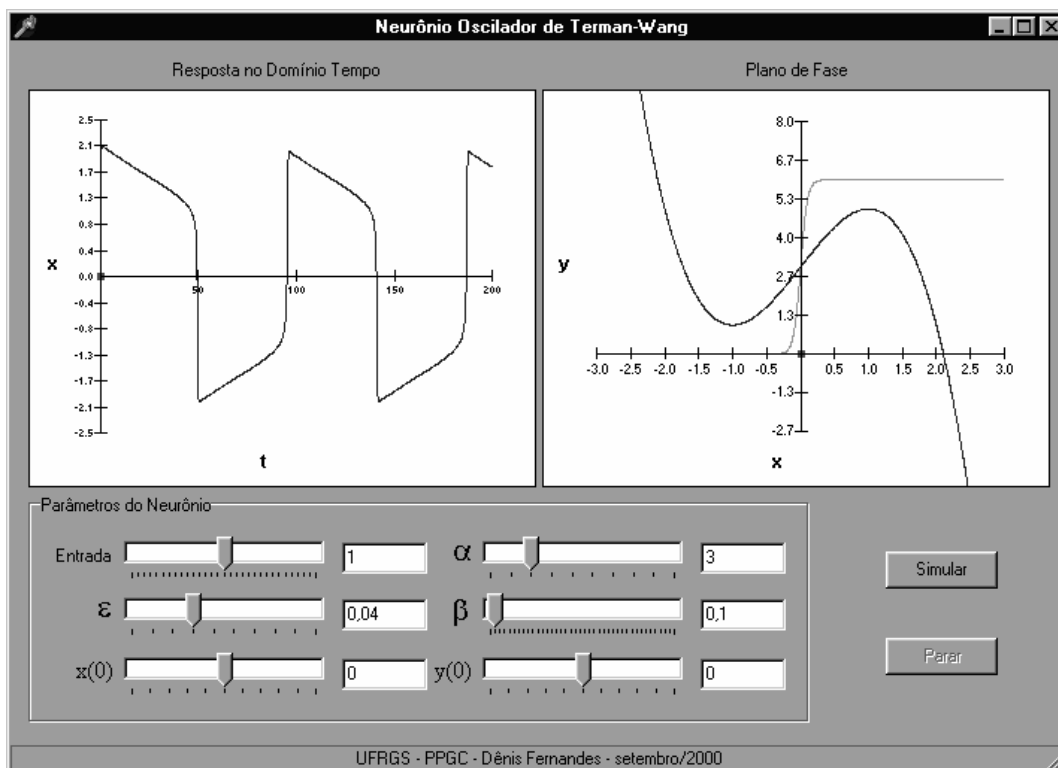


Figura APA.1: Aplicativo para simulação do neurônio oscilador de Terman-Wang

A Figura APA.2 apresenta o aplicativo desenvolvido para simulação de uma rede composta por quatro osciladores de Terman-Wang conectados em uma configuração do tipo anel. O referido aplicativo foi utilizado para obtenção dos resultados apresentados na seção 3.2.4. São apresentados gráficos referentes às saídas de cada um dos osciladores e do inibidor global em função do tempo, onde pode ser observado o sincronismo temporal ou não das saídas dos diversos osciladores. As entradas externas de cada um dos osciladores podem ser independentemente ajustadas permitindo assim a criação de segmentos homogêneos na entrada da rede. Os parâmetros inerentes dos osciladores α , β e ε e os parâmetros vinculados à rede, Φ , W_{ik} , W_z , θ_x , e θ_z , podem ser ajustados para possibilitar a visualização do efeito de cada um no desempenho do sistema. Neste aplicativo não foi utilizado o procedimento de normalização dos pesos dinâmicos, com o objetivo de comprovar a não obrigatoriedade de utilização do mesmo apontada pela bibliografia consultada.

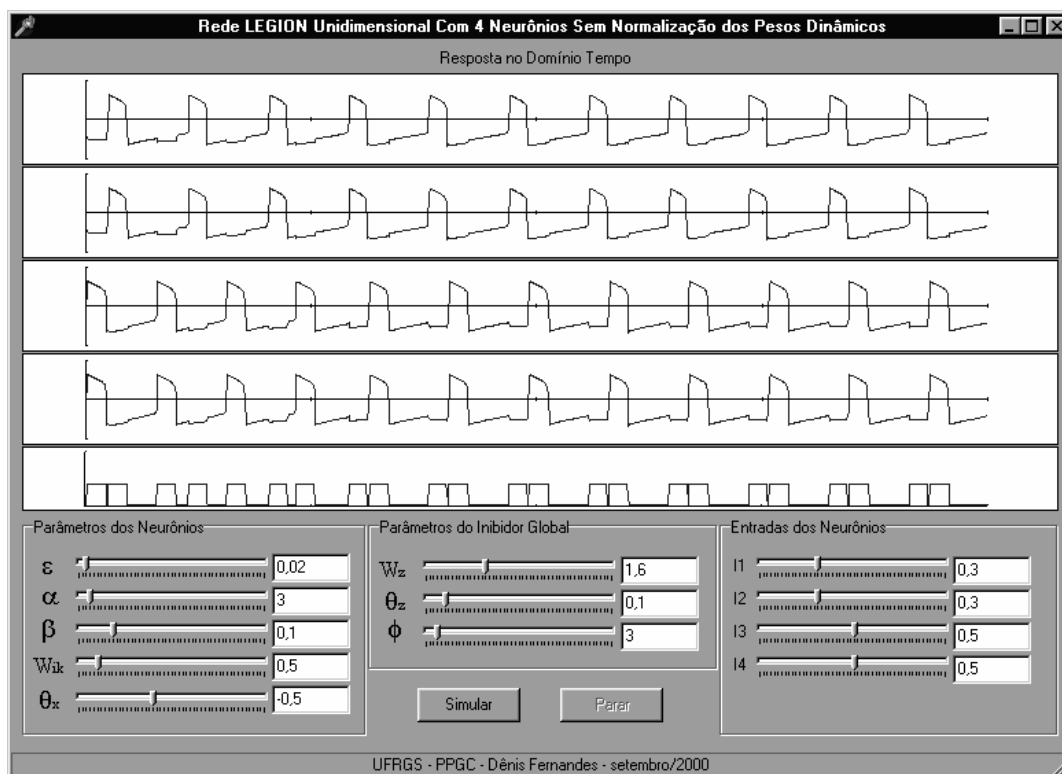


Figura APA.2: Aplicativo para simulação de uma rede de quatro neurônios osciladores de Terman-Wang em anel

A Figura APA.3 apresenta a tela do aplicativo desenvolvido para analisar o comportamento de uma rede LEGION bidimensional contínua sem o procedimento de normalização dos pesos dinâmicos. É utilizada uma rede de 20x20 osciladores com conexões excitatórias entre um determinado oscilador e seus vizinhos esquerdo, direito, superior e inferior (vizinhança 4). Os parâmetros inerentes aos osciladores, α , β e ε , e os parâmetros vinculados à rede, Φ , W_{ik} , W_z , θ_x , e θ_z , podem ser ajustados para possibilitar a visualização do efeito de cada um no resultado obtido. A imagem apresentada à esquerda, com dimensões de 20x20 pixels, é a imagem binária a ser segmentada, apresentando dois quadrados brancos (nível 1) sobre fundo preto (nível 0), sendo cada pixel da imagem a entrada de um oscilador da rede LEGION. A imagem da direita

apresenta a saída dos osciladores da rede LEGION, sendo que a cor preta corresponde ao menor nível da fase silenciosa (-2), a cor branca corresponde ao maior nível da fase ativa (+2) e os diversos níveis de cinza correspondem aos valores intermediários das saídas dos osciladores da rede. A opção "binária" possibilita a apresentação das saídas da rede negativas como preto (fase silenciosa) e as saídas maiores ou iguais a zero como brancas (fase ativa). O número da iteração apresentado corresponde ao número de passos para os quais foram calculadas as saídas de cada um dos osciladores utilizando o método de Runge-Kutta. Especificamente, a situação apresentada na Figura APA.3 demonstra que, naquela iteração, a rede LEGION conseguiu segmentar satisfatoriamente o quadrado menor, apresentando-o isoladamente em sua saída. Naquele instante, o oscilador correspondente ao pixel inferior direito do quadrado encontra-se com saída ligeiramente inferior em comparação com os demais osciladores do segmento.

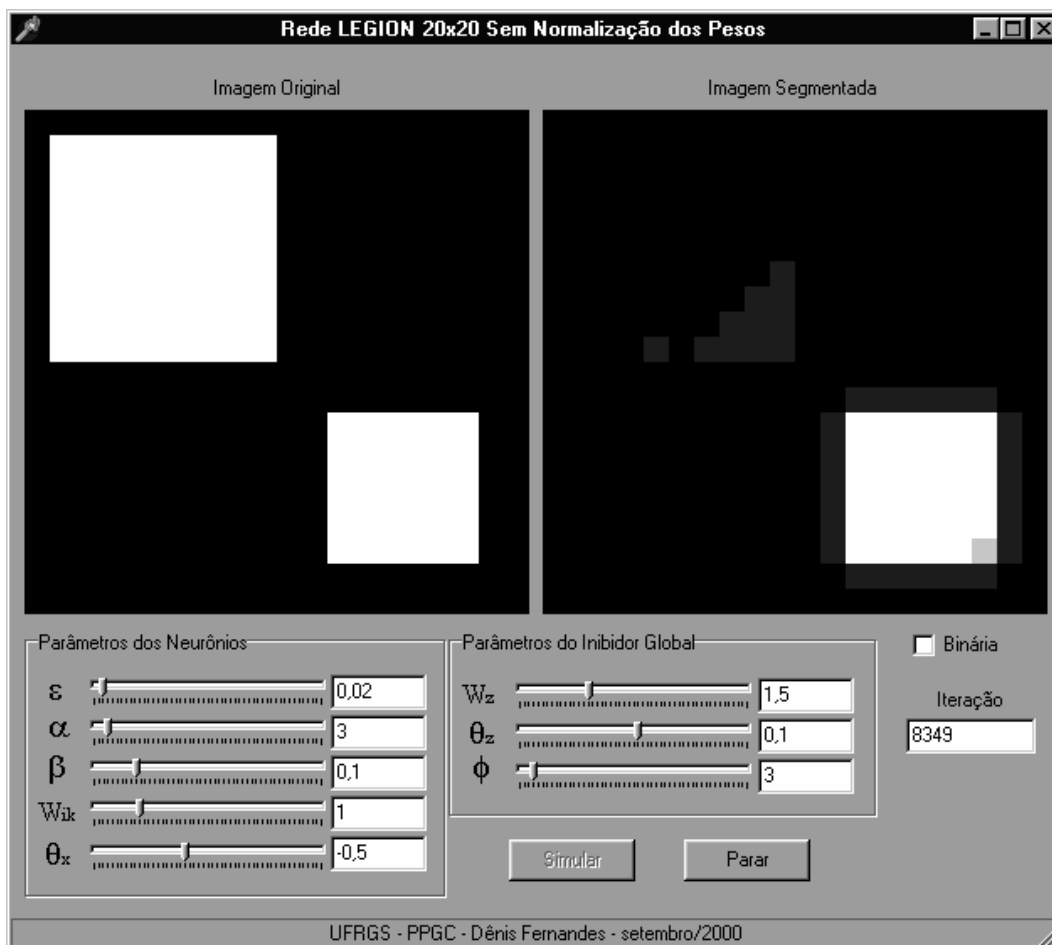


Figura APA.3: Aplicativo para simulação de uma rede LEGION com 20x20 osciladores, sem normalização dos pesos dinâmicos

A Figura APA.4 apresenta o aplicativo desenvolvido para simulação da rede LEGION contínua utilizando o procedimento de normalização dos pesos dinâmicos. Trata-se basicamente do mesmo programa apresentado na Figura APA.3 com a inclusão do procedimento de normalização dos pesos dinâmicos proposto por Terman e Wang, tendo sido o parâmetro W_{ik} substituído por W_T , o qual corresponde ao somatório dos

pesos das conexões entre um oscilador e seus vizinhos. Na situação em questão, foi apresentada uma imagem binária à rede composta por 4 caracteres brancos sobre um fundo preto. No instante de tempo apresentado (iteração), a rede LEGION segmentou perfeitamente uma das letras "P" apresentando a mesma isoladamente, tendo sido as demais letras também satisfatoriamente segmentadas e apresentadas seqüencialmente no tempo. Foi possível observar que o procedimento de normalização dos pesos dinâmicos implica maior facilidade da rede em segmentar imagens mais complexas.

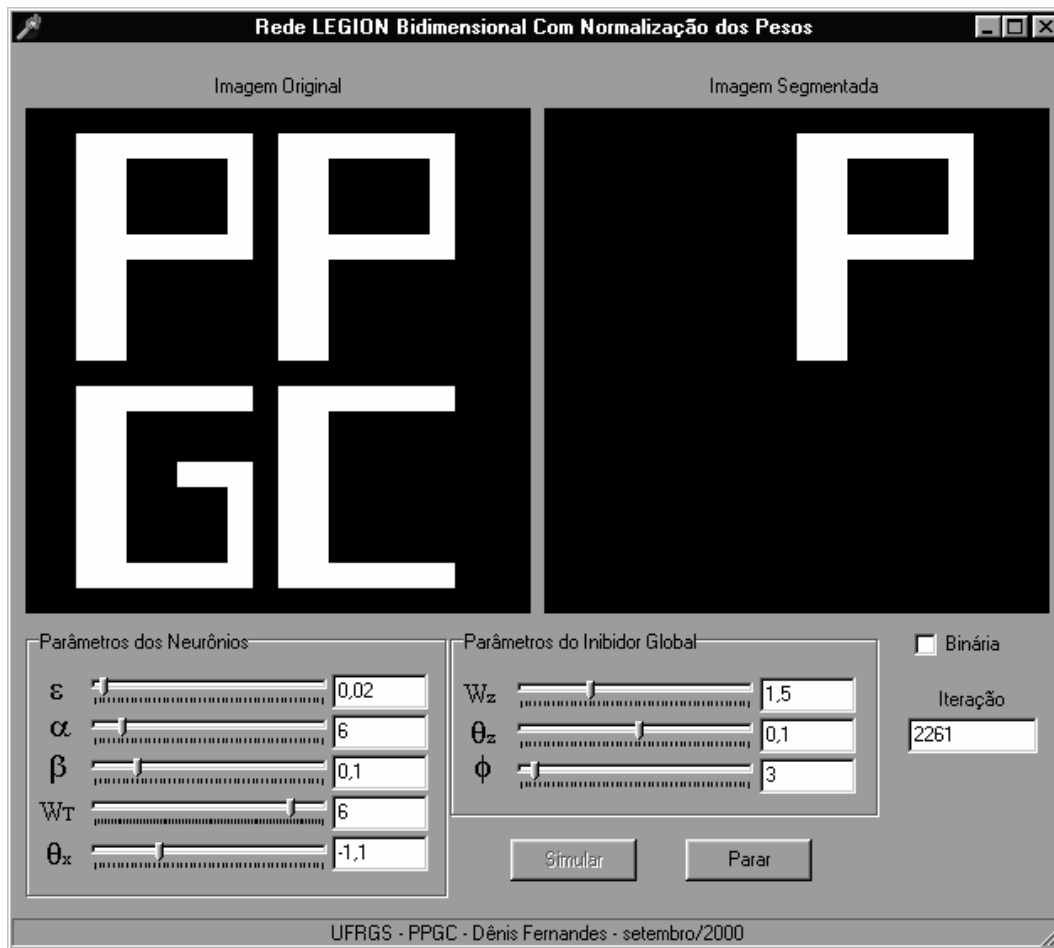


Figura APA.4: Aplicativo para simulação de uma rede LEGION com 20x20 osciladores e normalização dos pesos dinâmicos

A Figura APA.5 apresenta o aplicativo desenvolvido para simulação da rede LEGION em sua versão algorítmica. É possível ajustar os parâmetros θ_p e W_z e escolher os mecanismo para determinação da excitação lateral entre “soma”, “máximo” ou “logarítmico”. A imagem a ser segmentada pode ser livremente escolhida, com a restrição de ser aceito somente o formato BMP com tamanho limitado a 600x600 pixels. A determinação dos pesos dinâmicos é realizada com base na intensidade de cada pixel da imagem a ser segmentada. Desta forma, o atributo considerado para segmentação é a intensidade do pixel, independente de sua cor. Os segmentos encontrados podem ser apresentados em branco sobre fundo preto, em preto sobre fundo branco ou em sua cor natural sobre fundo preto ou branco com o objetivo de facilitar a visualização dos mesmos. Também são fornecidas opções para montagem da imagem com os segmentos

discriminados, facilitando a identificação das regiões da imagem que não foram segmentadas. Para cada segmento discriminado, é fornecida sua área relativa em relação à área total da imagem. Ao discriminar todos os segmentos possíveis, o procedimento é parado, sendo então apresentado o número de segmentos encontrados, os quais podem então ser visualizados um a um. É possível também indicar pontos sobre a imagem de entrada, os quais são utilizados para determinação dos osciladores líderes da rede. Desta forma, somente os segmentos vinculados aos líderes indicados pelo usuário serão discriminados pela rede. No caso apresentado, uma imagem em tons de cinza (8 bits) com 221x232 pixels foi apresentada à entrada da rede LEGION, tendo sido indicado um ponto sobre a imagem de entrada. O respectivo segmento homogêneo da imagem original é apresentado em branco sobre fundo preto na imagem da esquerda, demonstrando sua correta discriminação.

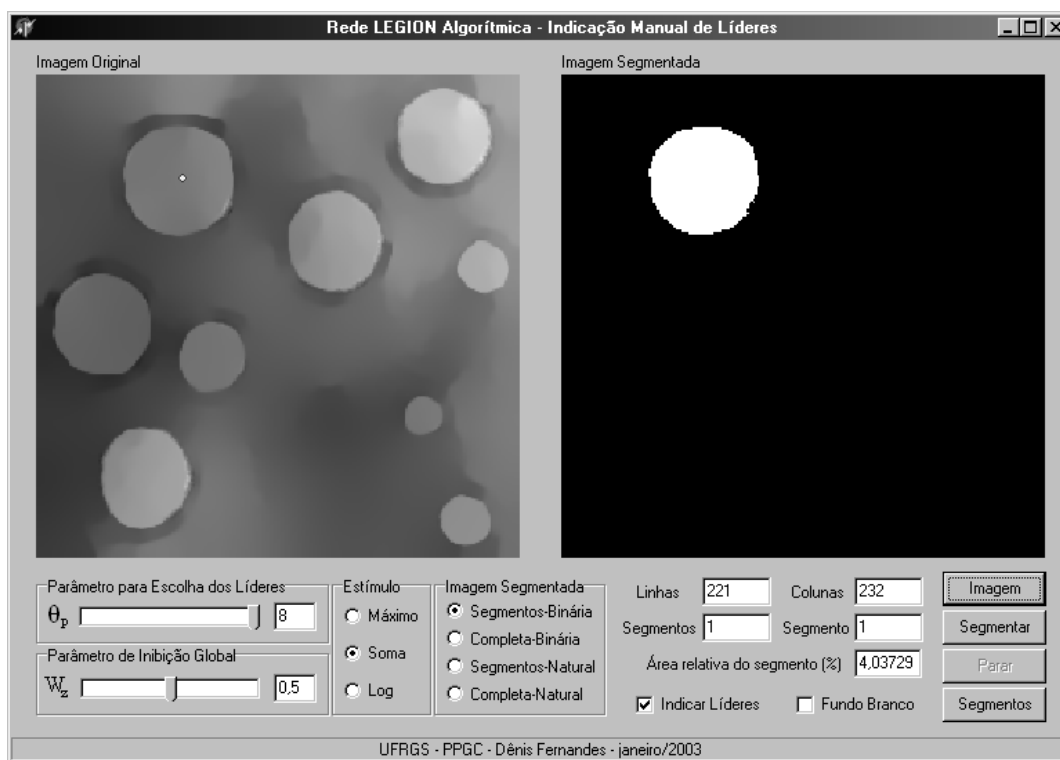


Figura APA.5: Aplicativo desenvolvido para simulação da rede LEGION em sua versão algorítmica

A Figura APA.6 apresenta o aplicativo desenvolvido para avaliar os resultados do algoritmo FPS na filtragem de imagens. São apresentadas simultaneamente na tela a imagem original e a imagem filtrada, sendo possível a variação dos parâmetros R , θ_σ , k_1 e k_2 . Também é possível o ajuste do número de iterações para filtragem da imagem de entrada. O caso apresentado na Figura APA.6 ilustra a filtragem de uma imagem natural de baixa qualidade, captada através de uma câmera digital, com dimensões de 160x160 pixels e 6 bits por amostra. A imagem filtrada foi obtida após a aplicação de 20 iterações (2×10^1) do algoritmo FPS com parâmetros $R=1$, $\theta_\sigma=10$, $k_1=30$ e $k_2=30$. Apesar do aspecto bastante artificial adquirido, a imagem filtrada apresenta regiões nitidamente mais homogêneas, discriminando melhor as bordas dos grandes objetos, sendo portanto mais adequada à aplicação de um posterior processo de segmentação.



Figura APA.6: Aplicativo desenvolvido para simulação do algoritmo FPS

A Figura APA.7 apresenta a interface gráfica do aplicativo desenvolvido para obtenção dos resultados que utilizam o filtro de Gabor. Neste aplicativo é possível selecionar a imagem de entrada (somente do tipo .BMP) e especificar os parâmetros α^2 , β^2 , u_0 e v_0 do filtro de Gabor. Adicionalmente, um limiar pode ser também selecionado com o objetivo de indicar os valores do módulo da imagem filtrada que se encontram acima do percentual especificado em relação ao valor máximo do módulo. No exemplo apresentado, são detectadas as regiões da imagem de entrada que possuem texturas com a presença de componentes senoidais de frequência espacial igual a zero ciclos por pixel (nível de cinza constante) na direção horizontal 0,2 ciclos por pixel na direção vertical. O resultado apresentado demonstra que o filtro de Gabor com os parâmetros especificados tem capacidade de discriminar adequadamente as regiões de ocorrência da componente de frequência especificada.

Todos os aplicativos apresentados foram desenvolvidos priorizando-se a flexibilidade na modificação de seus parâmetros e o aspecto visual da apresentação dos resultados, o que significou, em alguns casos, comprometimento da velocidade de processamento. Desta forma, não foi possível estimar com maior precisão a complexidade computacional para implementação da rede LEGION, tendo sido observado que, para imagens maiores, da ordem de 100x100 pixels ou mais, o tempo de processamento em computadores PC com processadores do tipo Pentium IV, mesmo utilizando a versão algorítmica da rede, pode se apresentar como um fator limitante adicional para aplicações práticas em tempo real. Tal fato é explicado pelo elevado número de osciladores da rede (um para cada pixel da imagem), o que, por outro lado, apresenta-se como uma característica interessante para implementação de um *hardware* paralelo analógico específico.

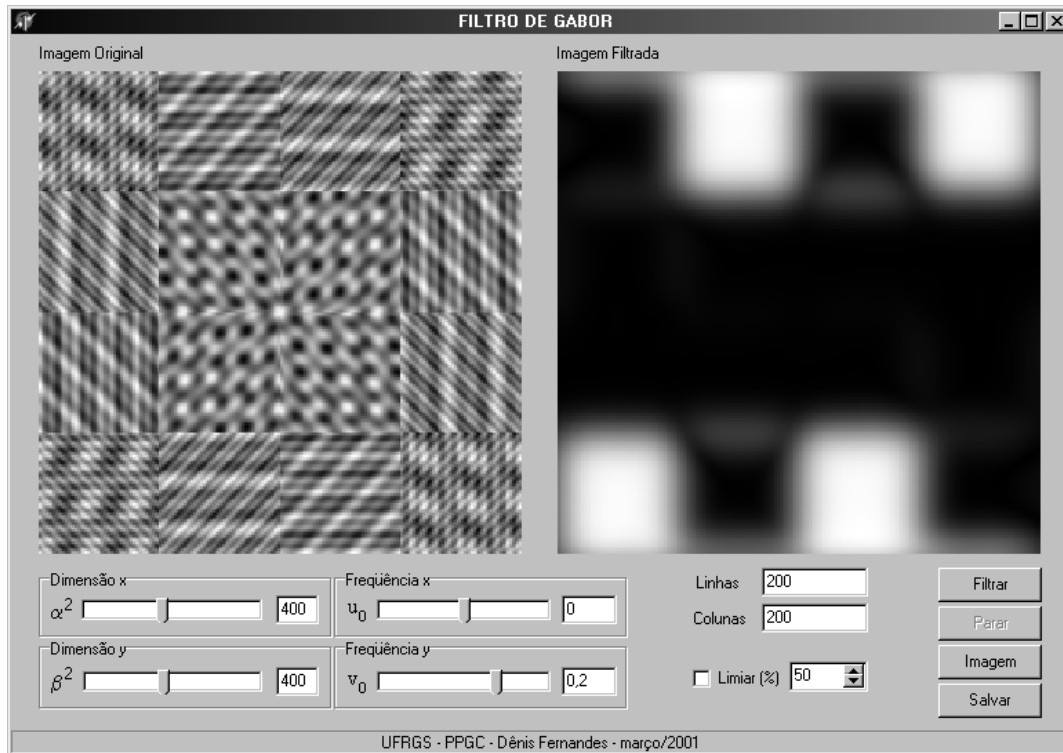


Figura APA.7: Aplicativo desenvolvido para o filtro de Gabor

ANEXO A ARTIGOS PUBLICADOS

Aseguir são apresentados os cinco artigos publicados durante o desenvolvimento da tese⁴. Procurou-se preservar o formato original de publicação de cada artigo, porém, as restrições impostas pelas margens do presente documento implicaram uma pequena diferença de escala entre os documentos originais e as versões aqui apresentadas.

O artigo *Segmentation of Fetal Echocardiographic Images Using Self-Organizing Maps and Oscillatory Neural Networks* (FERNANDES; SIQUEIRA; NAVAU, 2001-b) foi apresentado no *VI Iberoamerican Symposium on Pattern Recognition*. O tema abordado diz respeito à utilização da rede LEGION em um procedimento para extração e medida da área do ventrículo esquerdo em imagens de ecocardiografia fetal.

O segundo artigo, *Segmentation of TEM Images Using Oscillatory Neural Networks* (FERNANDES; NAVAU; FICHTNER, 2001-c), foi apresentado no *XIV Brazilian Symposium on Computer Graphics and Image Processing*. O artigo apresenta a utilização da rede LEGION, juntamente com o filtro FPS, em um sistema para contagem e medição da área de bolhas de hélio em amostras de silício utilizando imagens obtidas de microscópios de transmissão eletrônica.

Já o terceiro artigo, *Architecture of Oscillatory Neural Network for Image Segmentation* (FERNANDES; STEDILE; NAVAU, 2002), foi apresentado no *14th Symposium on Computer Architecture and High Performance Computing*. O artigo apresenta a nova arquitetura de rede proposta sem inibição global e alguns resultados preliminares em segmentação de imagens. Também estão incluídos resultados referentes à implementação de um pequeno protótipo em FPGA.

O quarto artigo, **Rede Neural Maciçamente Paralela para Segmentação de Imagens** (FERNANDES; NAVAU, 2003-a), foi apresentado I Workshop do Grupo de Processamento Paralelo e Distribuído e publicado nos Cadernos de Informática – Processamento Paralelo e Distribuído na Informática/UFRGS. Trata-se de uma versão mais resumida e em português do terceiro artigo, com a inclusão de observações e referências à aplicação da rede proposta em *chips* de visão. Além disso, foi dada menor ênfase à rede LEGION.

Finalmente, o artigo *An Oscillatory Neural Network for Image Segmentation* (FERNANDES; NAVAU, 2003-b) foi publicado na *Lecture Notes in Computer Science – Progress in Pattern Recognition, Speech and Image Analysis*, referente aos anais do *CIARP'2003 - 8th Iberoamerican Congress on Pattern Recognition*, realizado em Cuba. O artigo apresenta a aplicação da rede proposta sem inibição global (ONNIS) na segmentação de bolhas de hélio, de maneira análoga àquela realizada com a rede LEGION no artigo 2.

⁴ Dois artigos sobre a rede ONNIS-GI, um apresentado no WSPPD'2004 (UFRGS, Porto Alegre, julho/2004) e outro aceito para apresentação no ISSPIT'2004 (IEEE, Roma, dezembro/2004) não foram incluídos nesta relação pois não haviam sido publicados ou apresentados até a entrega do presente documento para avaliação pela banca.

Segmentation of Fetal Echocardiographic Images Using Self-Organizing Maps and Oscillatory Neural Networks

Dênis Fernandes¹, Mozart Lemos de Siqueira², and Philippe Olivier Alexandre Navaux²

¹PUCRS - Faculdade de Engenharia - Pontifícia Universidade Católica do Rio Grande do Sul, Av. Ipiranga, 6681, 90619-900 Porto Alegre, RS, Brasil
denis@ee.pucrs.br

²UFRGS - Instituto de Informática - Universidade Federal do Rio Grande do Sul, Av. Bento Gonçalves, 9500, 91501-970 Porto Alegre, RS, Brasil
{mozart, navaux}@inf.ufrgs.br

Abstract

This article presents a new approach for segmentation of fetal echocardiography images and the application of the same one in a semi-automatic system of analysis. The proposed method uses self-organized maps to approach the probability density functions of the image patterns and an oscillatory neural network to isolate the desired structures, facilitating the extraction of measurements. The clusters obtained with the application of the self-organized maps are post-processed using the K-Means algorithm with the objective to obtain 2 groups of neurons with similar weights, yielding the segmentation of the original images in 2 classes. The images segmented on the basis of the post-processed maps are then presented to a LEGION oscillatory neural network for extraction of the interest region. The LEGION network generates a binary image, which is used for measurement of the cardiac structure in focus. The results of the procedure had been classified as satisfactory by medical specialists.

1. Introduction

Innumerable applications imply the extraction of several features from images so that the same ones can be properly understood by an automatic or semi-automatic system of analysis. Images conveniently post-processed can be segmented through some adequate technique, with the objective to detect regions with similar features (segments). The posterior application of a classifier relates a label with each one of the identified regions, according to a previous base of knowledge. With the results of the classification, descriptions or interpretations of the images, related with the context of the application, can be elaborated or measures of the present structures can be obtained.

Several techniques of image segmentation based on artificial neural networks have been developed, particularly using MLP (MultiLayer Perceptron)

networks associated to the Backpropagation algorithm [6], Hopfield networks [6] or Kohonen maps (SOM - Self-Organizing Maps or SOFM - Self-Organizing Feature Maps) [6], with typical examples founded respectively in [1,4,10]. Such algorithms have presented good results, even in the presence of noise or distortions over the image to be segmented. Also, they present some of the inherent advantages of neural networks including the robustness in relation to the occurrence of faults and the parallelism that can be expressed as fast operation when using an adequate hardware implementation. As disadvantages, some of these methods present the necessity of training, what can be a problem when a long training time or a large number of samples previously segmented are necessary. The increased computational complexity for implementation in real time using non parallel machines can be also a restrictive factor in some practical applications.

Recently, alternative topologies of artificial neural networks, called oscillatory neural networks, have been applied in procedures of image segmentation with favorable results. The LEGION (Locally Excitatory Globally Inhibitory Oscillator Network) network, developed by DeLiang Wang and David Terman [15,16], from The Ohio State University, is an example of oscillatory network with applications in image segmentation. For example, in [8] it is found a procedure involving MLP networks associated to oscillatory networks used for extraction of hydrographic regions in images of remote sensing. In [12], can be found a study on the use of networks of oscillatory neurons applied to the segmentation of medical images of computerized tomography (CT) and of magnetic resonance (MRI).

In the specific case of fetal echocardiographic images, the segmentation can assist in the detection and diagnose of congenital cardiac problems, emphasizing important structures and leading to more trustworthy evaluations in prenatal examinations. Moreover, characteristic measures, such as the thickness of the interventricular septum and volumes of the ventricles,

Artigo publicado nos anais do VI Iberoamerican Symposium on Pattern Recognition

used to identify some anomalies, can be obtained with more accuracy.

The information present in fetal ultrasonic images is normally degraded by the characteristic noise of the process (speckle noise) [10], which mainly corrupts the edges of the structures present in the image. The variability of the images obtained, whose quality depends on the position of the transducer, position of the fetus and composition of the tissue between the transducer and the fetus, can be also an important factor to be considered in the implementation of automatic or semi-automatic procedures of segmentation.

With the objective to reduce the referred problems, some specific techniques of segmentation have been proposed [1,11,13,14]. In [14], it is proposed a MLP network whose inputs are the average value and the variance of the pixel intensities and parameters of texture obtained from the gray-level co-occurrence matrix (contrast, entropy and homogeneity), which are used to detect the edges of the left ventricle. The same parameters are used in [1] to classify, using a MLP network, the whole image in 2 main classes: cavity and tissue.

In this work, a mixed solution, based on a SOM post-processed by the K-Means algorithm [9] and on a LEGION network [15,16] is presented, which is used for the extraction of structures in images of fetal echocardiography. After the phase of unsupervised clustering executed by the SOM, the K-Means algorithm is applied with the objective to agglutinate the neurons with similar weights in two classes, which will be used to classify the homogeneous regions present in the images [10]. The subsequent use of the LEGION network, with a small modification, makes possible the extraction of the desired structures and the presentation of the same ones in individual binary images, facilitating the automatic measurement.

The following section presents one brief description of the fetal ultrasound images used to evaluate the proposed method. Section 3 presents a discussion about the selection of the image features used in the proposed process. In section 4, the procedure of clustering through the use of the SOM and K-Means algorithm is presented. Section 5 presents the LEGION network and a proposal of modification in its algorithm in order to obtain only the desired structures from its output. In section 6, we present the results obtained from the application of the proposed procedure for the measurement of the area of the left ventricle in images of fetal echocardiography. Finally, section 7 presents the final conclusions on the developed study.

2. Images of Fetal Echocardiography

In the south of Brazil, around 8 to 10 in each 1000 newborn babies present cardiac problems, which can be more adequately treated when previously detected [2]. Inside this context, the fetal echocardiography represents

an important tool, fast and safe, for prenatal detection of congenital defects in the structure of the heart.

The images of the fetal heart are obtained by the use of a transducer, freely maneuverable, which gets data of the heart through the reflection of ultrasound in the internal structures, generating a video sequence that can be observed by a specialist. Based on these sequences, different anomalies can be diagnosed. Some illnesses are characterized by the shift or the abnormal form and size of certain structures of the heart. For example, the abnormality of Ebstein consists of the shift of the tricuspid valve, reducing the size of the right ventricle. Another common problem is related with the hypertrophy of the septum [17].

There are different viewing standards of the fetal heart, each one adjusted to observe some specific structure or function. The four chamber view is considered the most complete, being possible to observe several structures simultaneously and to detect almost two thirds of the known cardiac malfunctions [7]. Fig. 1 presents an example of an image obtained in four chambers view, in which some important structures are indicated, including the left ventricle (LV), the right ventricle (RV), the left atrium (LA), the right atrium (RA), the tricuspid valve (TV), the mitral valve (MV), the moderator band (MB), the interventricular septum (IS), and the descending aorta (DA).

Such images, which were used for evaluation of the proposed procedure can be obtained in digital form, in the DICOM (Digital Imaging and Communications in Medicine) standard format, directly from the echocardiographic equipment.

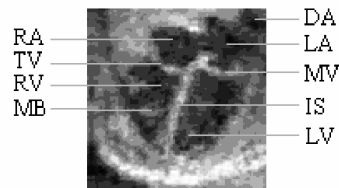


Figure 1. Heart of a fetus in four chambers view with indication of relevant structures.

3. Selection of Features

The selection of the adequate features plays an important role in the field of image segmentation. With the techniques of feature selection it is possible to diminish the dimensionality of the problem with the objective to reduce its computational complexity, preserving the level of separability. The feature selection must be based on the separability of classes. Features that make possible greater separability of classes are more significant and must be chosen for the implementation of the image segmentation process.

There are different methods to evaluate the separability of the several classes. One common

technique is the use of the Bhattacharyya distance [3]. This distance is capable to measure the overlapping area of two probability density functions ($p(x/w_i)$), related to the classes w_1 and w_2 . How much lesser the area, greater is the Bhattacharyya distance and better is the separability. The equation (1) describes the Bhattacharyya distance (B) between the classes w_1 and w_2 , represented by their respective probability density functions.

$$B = -\ln \int_{-\infty}^{\infty} \sqrt{p(x/w_1)p(x/w_2)} dx \quad (1)$$

The features that are used to classify and to segment echocardiographic images, includes mean value of the gray levels, variance, entropy, contrast and homogeneity [1,10,13,14]. The three last ones are extracted from the local gray level co-occurrence matrix. The separability between the classes must be evaluated for each combination of features. Table 1 presents the values of the Bhattacharyya distance determined for several combinations of features for the classes of tissue and cavity, obtained from typical fetal echocardiographic images in four chamber view [10,13].

Features	Bhattacharyya Distance
mean	7,673
variance	0,290
contrast	0,270
entropy	0,211
homogeneity	0,189
mean and variance	12,507
mean, contrast and entropy	13,396

Table 1. Bhattacharyya distance between classes for different features [13].

From the presented values, it can be verified that the mean value of the pixels intensities is the most significant feature for classes separability, followed by the variance of the pixels intensities. The simultaneous use of these two features results in a significant increase of the Bhattacharyya distance. On the other hand, the simultaneous use of the mean, contrasted and entropy, presents a small profit in relation to the use of mean and variance. In this way, considering the additional computational effort to determinate contrast and entropy, we opted to use the mean and the variance of the pixels intensities to segment the echocardiographic images.

Figure 2 presents an image of a fetal heart in four chambers view and the respective mean and variance of pixels intensities using a sliding window of 3x3 pixels. As much the mean, the variance had been scaled for full representation with 8 bits of resolution (levels between 0 and 255). It can be seen that the mean emphasizes the homogeneity of the regions while the variance emphasizes the limits between adjacent homogeneous regions.

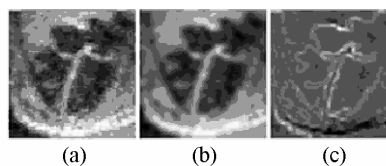


Figure 2. (a) Original image; (b) Mean; (c) Variance.

4. Segmentation Using Self-Organizing Maps and the K-Means Algorithm

The self-organizing map proposed by Kohonen is a unsupervised approach to extract information of unknown data and organize them into groups with similar features.

In the adopted SOM, the units of the input layer, one for each feature of the image, are connected to all linear units of a two-dimensional competitive layer. The winner neuron is determined by the minimum Euclidean distance of its weights w_i in relation to the input vector x_j representing the features mean and variance for each pixel. The learning rule for updating the weights of the neurons in the competitive layer with the presentation of a vector in the input is expressed by the equation (2). In this equation, a weight factor $\Phi_{r_i, r_i^*}(k)$ is present, which is related to the neighborhood around the winner neuron (3). This factor implies a weighted update of the neighboring neurons weights, which decrease as a function of the time. The update of all neurons also suffers an attenuation during the learning process, so that the computed weights do not be strongly affected with a new presentation of the input vectors. $\rho(k)$ is the learning rate for the k -th presentation of the input data, $\sigma(k)$ is a parameter relative to the extension of the neighborhood region, r_i are the coordinates of the neuron to be updated, and r_i^* are the coordinates of the winner neuron [6].

$$w_i(t+1) = w_i(t) + \rho(k) \Phi_{r_i, r_i^*}(k) (x_j(t) - w_i(t)) \quad (2)$$

$$\Phi_{r_i, r_i^*}(k) = e^{-\frac{\|r_i - r_i^*\|^2}{2\sigma^2(k)}} \quad (3)$$

In the proposed application, random samples of the pixels features of the images under analysis has been selected, which were used as entries for the SOM training.

With the objective to separate the neurons of the trained SOM in two classes (tissue and cavity), which will be used to segment the image, the use of the K-Means algorithm is proposed [13]. This procedure classifies the weights of the neurons according to minimum Euclidean distance and determines new average values for each cluster, minimizing the total distance in relation to the mean value inside of one cluster. For the measures obtained from a structure like the LV, the use of 2 classes is adequate, however, less clear structures can have its detection degraded.

To execute the procedure of image segmentation, the average of the intensities and the variance are determined for each pixel, both using a window of 3×3 pixels. The vector determined for these features is compared with the weights of the centers of the classes determined by the K-Means algorithm, choosing the winner class according to the minimum Euclidean distance. Fig. 3 presents the result of the segmentation of a fetal echocardiographic image in classes of tissue (white regions) and cavities (black regions), where a satisfactory emphasis of the physical limits of the structures can be observed.

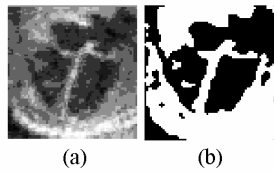


Figure 3. (a) Original image; (b) 2 classes segmented image.

5. Extraction of Structures Using the LEGION Network

The Terman-Wang oscillator is a relatively recent proposal, which, for mathematical and computational reasons, has been used for composition of oscillatory artificial neural networks with application in image segmentation. The behavior of the continuous Terman-Wang oscillator can be described by the pair of differential equations (4) e (5). Under the biological point of view, the state variable $x(t)$ can be understood as the potential of the membrane of the nervous cell, or either, the physical variable that represents the output of the neuron. α , β and ε are parameters of the model, I is an external input, which is related to features of image pixels, $S_i(t)$ is the lateral excitation received from the neighbour oscillators and $y(t)$ is an internal state variable [16].

$$\frac{dx(t)}{dt} = 3x(t) - x^3(t) + 2 - y(t) + I + S_i(t) \quad (4)$$

$$\frac{dy(t)}{dt} = \varepsilon(\alpha(1 + \tanh(x(t)/\beta)) - y(t)) \quad (5)$$

$$S_i(t) = \sum_{k \in N_i(1)} W_{ik} S_{\infty}(x_k(t), \theta_x) - W_z S_{\infty}(z(t), \theta_z) \quad (6)$$

$$\frac{dz(t)}{dt} = \phi(H(x_i(t) - \theta_x) - z(t)) \quad (7)$$

$$S_{\infty}(x, \theta) = \frac{1}{1 + e^{-\kappa(x-\theta)}} \quad (8)$$

The LEGION network consists basically of an array of Terman-Wang oscillators with a local coupling to make possible phase synchronism between neighbor oscillators with similar external inputs and a mechanism of global inhibition ($z(t)$) to get asynchronism between different groups of oscillators physically distant, as it can

be observed in the topology of the network presented in the fig. 4. W_{ik} are the weights related to the connections between neighboring oscillators k and i , which represents the similarity between the respective external inputs. W_z is the weight related to global inhibitor, θ_x and θ_z are thresholds, $N_i(R)$ is the neighborhood of radius R of the oscillator i , $H(x-\theta)$ is the function of Heaviside, and κ and ϕ are additional parameters [15].

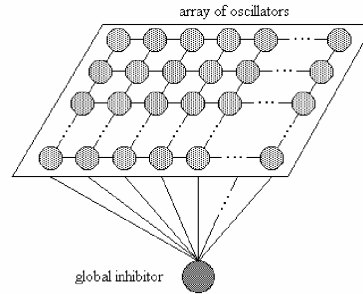


Figure 4. Topology of a two-dimensional LEGION network.

In [15] is also presented a modification in the input of the Terman-Wang oscillator considering the basic idea that a set of oscillators with similar inputs must possess at least one oscillator, denominated leader, which must receive great lateral excitation from its neighborhood. On the other hand, isolated oscillators, pertaining to noisy fragments, cannot be characterized as leaders. In this way, an oscillator with great lateral potential can lead the activation of a group ("legion") of oscillators corresponding to a homogeneous pattern of input (segment). It should be observed that the number of leaders does not correspond to the number of discriminated segments, since each homogeneous region can possess several leaders, although only one leader is really necessary to activate the block.

When presenting an image to the input of a LEGION network, both with same dimensions, the several segments are obtained from the output, individually and each one in a different interval of time, according to the features used for determination of the network weights. Thus, it can be seen that the LEGION network provides a spatial and temporal segmentation of the input signal. Fig. 5 presents an image and the respective time sequence of segments obtained from the output of a LEGION network, using the intensity of the pixels as feature for segmentation. The obtained segments are presented in white over black background.



Figure 5. Original image (left) and the 3 segments obtained from the output of a LEGION network.

Artigo publicado nos anais do VI Iberoamerican Symposium on Pattern Recognition

With the objective of reducing the computational effort when using serial computers, Terman and Wang had developed a specific algorithm for the implementation of the LEGION network in such machines [15], which was adopted in the present work. This form, the images segmented with the use of the SOM and the K-Means algorithm are presented to a LEGION network with same dimensions of the images, which supplies in its output, sequentially, each one of the segments with homogeneous intensities.

To obtain only the segments of interest in the output of the LEGION network, an modification in the procedure of determination of the leader oscillators was implemented, which consists of the manual indication of them. Thus, upon desiring the discrimination of exclusively one segment, it is marked one or more points over the respective region on the original image. The related points will correspond to the leader oscillators, which will be the only ones qualified to promote the activation of groups of oscillators locally coupled. Fig. 6 presents an example where a point over the letter "A" was indicated over the original image, which leads to the identification of only the region of the referring segment in the output of the LEGION network.

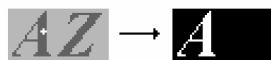


Figure 6. Original image with indication of the region to be segmented and the respective segment obtained from the LEGION network output.

6. Experimental Results

For evaluation of the performance of the presented procedure, several images of fetal hearts in four chambers view had been selected, obtained in intervals of time for which the mitral valve was closed, making possible the individual segmentation of the left ventricle for its respective area estimation.

Fig. 7(a) presents an image where it was made a mark on the LV with the objective of inducing the LEGION network to extract only the region related to the respective segment. Fig. 7(b) presents the result of the segmentation using the SOM/K-Means procedure. Fig. 7(c) presents the output of the LEGION network, being able to be observed that the extraction of the LV region was not carried through successfully due to a small connection with the left atrium, resulting of the low quality of the image in the region in question.

To eliminate the small connections between structures, results mainly from the poor quality of the images, an operation of morphologic binary opening [5] was applied over the image obtained from the SOM/K-Means procedure, using a unitary structuring element with dimensions of 3x3 pixels. The result of the application of the proposal operation is presented in fig.

8, being able to be observed the correct splitting of the region of the LV. However, in some images, like the case showed in fig. 9, where the frontier between the LV and the LA is strongly corrupted by noise, even the morphologic operation was not able to separate the region of the LV. In this case, it is necessary to choose another image of the sequence to apply the process.

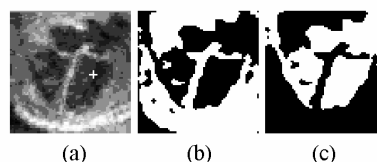


Figure 7. Incorrect segmentation of the LV. (a) Original image with indication of the region to be segmented. (b) SOM/K-Means Segmentation. (c) LEGION output.

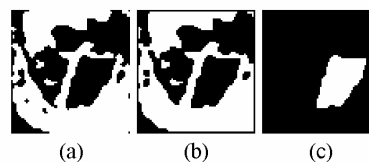


Figure 8. Correct segmentation of the LV using morphologic binary opening. (a) Image obtained with the SOM/K-Means segmentation. (b) Result of the operation of morphologic binary opening. (c) Output of the LEGION network.

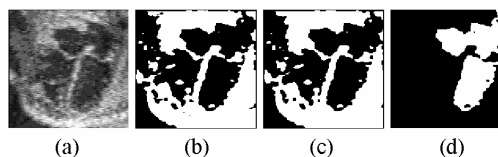


Figure 9. Incorrect segmentation of the LV. (a) Original image. (b) SOM/K-Means result. (c) Binary opening result. (d) LEGION result.

The image observed in fig. 8(c) makes possible the immediate estimation of the area of the LV, being the same one proportional to the ratio between the number of active oscillators (white area) and the total number of oscillators of network LEGION (equals to the number of pixels of the image). In the presented case, the area of the LV is equivalent to 11.49% of the image area or 612 pixels², for a network of 74x72 oscillators. Based on a reference scale (mm/pixel) supplied by the equipment of ultrasound, the real area of the LV, in mm², can be determined. Fig. 10 shows other examples where the LV region was segmented with success using the proposed procedure. The influence of the holes observed inside the segmented LVs in the area estimation is a topic for future investigation.

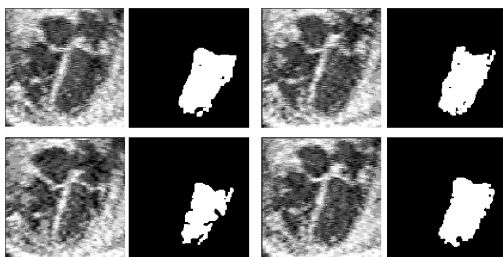


Figure 10. Examples of correct segmentation of the LV.

7. Conclusion

A new procedure for extraction and measurement of regions in fetal echocardiography images was presented, which is based on a post-processed SOM through the K-Means algorithm and on a LEGION network with leaders oscillators indicated by the user. The SOM/K-Means procedure carries on the classification of the images in classes of tissue and cavity, while the LEGION network extracts the region of interest for purpose of measurement. Due to the possible existence of small connections between distinct structures in some images, an operation of morphologic binary opening is carried on the image submitted to the LEGION network, which minimized the problem satisfactorily.

The manual indication of the leaders in the LEGION network, inducing the same one to extract only the indicated segments, is an original approach with relevance in practical applications of the network, avoiding a posteriori selection of segments and increasing the velocity of the process.

For performance evaluation of the considered procedure, several simulations were carried on where the corresponding regions to the LV in four chambers view fetal heart images had been marked. The respective segments, obtained from the output of the LEGION network, had adequately preserved the form and the size of the LV, making possible to estimate its area with more accuracy. The results had been considered satisfactory by medical specialists.

As future perspectives, we intend to correlate the area of the LV with the respective volume, providing a direct estimation of the LV volume, which is a more useful measure. Also, it is necessary to evaluate the error introduced by the holes observed inside the segmented LV in their area or volume estimation.

8. References

[1] DAHMER, A. *Segmentação de Imagens Ecocardiográficas Utilizando Redes Neurais e Medidas de Textura*. Porto Alegre: UFRGS-PGCC, 1998. 98p. Master Thesis.

[2] FIRPO, C. *Estudo evolutivo do comportamento do septum primum durante a vida intra-uterina e sua*

relação com a presença de extra-sístoles atriais no feto. Porto Alegre: UFRGS, 1995. Master Thesis.

[3] FUKUNAGA, K. *Introduction to Statistical Pattern Recognition*. 2 ed. New York: Academic Press, 1990. 591p.

[4] GHOSH, A.; PAL, N. R.; PAL, S. K. Object background classification using Hopfield type neural network. *Int. Journal on Pattern Recognition and Artificial Intelligence*, Bern, v.6, n.5, p.989-1008, 1992.

[5] GONZALES, R. C.; WOODS, R. E. *Processamento Digital de Imagens*. São Paulo: Edgard Blücher, 2000. 509 p.

[6] HAYKIN, S. *Neural Networks - A Comprehensive Foundation*. New York: Macmillian College Publishing Company, 1994. 696p.

[7] HUHTA, J. C.; ROTONDO, K. M. Fetal Echocardiography. In *Seminars in Roentgenology*, 1991. *Proceedings...*, 1991, v01.26, p.5-11.

[8] LIU, X.; CHEN, K.; WANG, D. L. Extraction of hydrographic regions from remote sensing images using an oscillator network with weight adaptation. *IEEE Transactions on GeoScience and Remote Sensing*. New York, v.39, n.1, p.207-211, jan. 2001.

[9] PANDYA, A. S.; MACY, R. B. *Pattern Recognition with Neural Networks in C++*. Boca Raton: CRC Press, 1996. 410p.

[10] PICCOLI, L. *Segmentação e Classificação de Imagens Ecocardiográficas Utilizando Redes Neurais*. Porto Alegre: UFRGS-PGCC, 1999. 90p. Master Thesis.

[11] POLI, R.; VALLI, G. *Optimum Segmentation of Medical Images with Hopfield Neural Networks*. 1995. The University of Birmingham, Scholl of Computer Science. (Technical Report CSRP-95-12).

[12] SHAREEF, N.; WANG, D. L.; YAGEL, R. Segmentation of Medical Images Using LEGION. *IEEE Transactions on Medical Imaging*. New York, v.18, n.1, p.74-91, jan. 1999.

[13] SIQUEIRA, M. L.; GASPERIN, C.V.; ZIELINSKY, P.; SCHARCANSKI, J.; NAVAU, P. O. A. Echocardiographic Image Sequence Segmentation Using Self-Organizing Maps. In *IEEE Workshop on Neural Networks for Signal Processing - NNSP*, 2000. *Proceedings...* Sydney, IEEE, 2000, p. 594-603.

[14] SÜSSNER, M.; BUDIL, T.; PORENTA, G. Segmentation and Edge-Detection of Echocardiograms Using Artificial Neural Networks. In *International Conference EANN*, 1995. *Proceedings...*, 1995.

[15] WANG, D. L.; TERMAN, D. Image Segmentation Based on Oscillatory Correlation. *Neural Computation*, Cambridge, v.9, p.805-836, 1997.

[16] WANG, D. L. Relaxation Oscillators and Networks. New York: *Wiley Encyclopedia of Electrical and Electronics Engineering*. John Wiley & Sons, 1999. v.18, p.396-405.

[17] ZIELINSKY, P. *Malformações Cardíacas Fetais: Diagnóstico e Conduta*. Arquivo Brasileiro de Cardiologia, dec. 1998.

Segmentation of TEM Images Using Oscillatory Neural Networks

DÊNIS FERNANDES¹

PHILIPPE OLIVIER ALEXANDRE NAVAU²

PAULO FERNANDO PAPALEO FICHTNER³

¹PUCRS - Faculdade de Engenharia - Pontifícia Universidade Católica do Rio Grande do Sul, Av. Ipiranga, 6681, 90619-900 Porto Alegre, RS, Brasil

denis@cc.pucrs.br

²UFRGS - Instituto de Informática - Universidade Federal do Rio Grande do Sul, Av. Bento Gonçalves, 9500, 91501-970 Porto Alegre, RS, Brasil

navaux@inf.ufrgs.br

³UFRGS - Instituto de Física - Universidade Federal do Rio Grande do Sul, Av. Bento Gonçalves, 9500, 91501-970 Porto Alegre, RS, Brasil

fichtner@if.ufrgs.br

Abstract. Oscillatory neural networks are a relatively recent approach for the problem of image segmentation. Inside of this context, the oscillator neuron of Terman-Wang is presented, which one is used as base element of an oscillatory network called LEGION (Locally Excitatory Globally Inhibitory Oscillator Network). The continuous version of the LEGION network, based on a set of differential equations, presents high computational complexity and has limited capacity of segmentation, what restricts its practical application, being adequate for implementation in parallel hardware topologies. To reduce the computational complexity in serial computers, an algorithm proposed by Terman and Wang is presented, which implies significant gain of speed in comparison to the continuous version and, in contrast, capacity to discriminate a unlimited number of segments. An interactive version of this algorithm was proposed and the results obtained in segmentation of transmission electron microscopy (TEM) images were evaluated, with the objective of measuring helium bubbles in silicon samples. As final result we found that the LEGION network presents itself as a singular alternative to solve problems of image segmentation, which provides simultaneously both spatial and temporal discrimination of segments.

1 Introduction

Innumerable applications imply the extraction of features from an image in order that the same can be understood by a system of artificial vision. The pre-processing implies the restoration, the improvement or simply the adequate representation of the image data, improving some features, as the contours of objects, for example, or the transformation of the image for some other domain more appropriated to the application. The conveniently pre-processed image is then segmented through the use of some appropriated technique, with the objective to isolate regions with similar features (segments). The posterior application of a classifier relates a label with each one of the regions or segments identified according to a previous base of knowledge. With the results of the classification procedure, a description or an interpretation of the image, appropriated to the context of the application, can finally be elaborated.

The procedures used in image segmentation can be classified in 4 great categories [7]: a) the classic methods based on amplitude thresholding, detection of edges or growing/contraction of regions; b) statistical methods,

such as the classifier of maximum likelihood; c) methods based on fuzzy logic, as the algorithm fuzzy c-means and d) methods based on artificial neural networks with varied topologies.

Several techniques of image segmentation based on artificial neural networks have been developed, particularly using MLP networks (MultiLayer Perceptron) associated to the Backpropagation algorithm [5,8], Hopfield networks [5] or Kohonen maps (SOFM - Self-Organizing Feature Map) [5], with typical examples found respectively in [3,4,9]. Such algorithms have presented good results even in the presence of noise or distortions in the image to be segmented, bringing with itself the inherent advantages of the neural networks related to the robustness in relation to the occurrence of faults and to the parallelism that can be expressed as velocity of operation when used the adequate hardware. As disadvantages, some of these methods present the training necessity, what can be problematic due to the long necessary time, the number of available samples previously segmented and also very high complexity, in some cases, for implementation using non parallel machines.

Artigo publicado nos anais do XIV Brazilian Symposium on Computer Graphics and Image Processing

Recently, alternative topologies of artificial neural networks, called oscillatory neural networks or simply oscillatory networks, have been applied in procedures of image segmentation with favorable results. In this area, is particularly interesting the work developed by DeLiang Wang and David Terman [14], who have lead most of the practical research with this type of network. For example, in [6] it is found a procedure involving MLP networks associated to the oscillatory networks, used for extraction of hydrographic regions in images of remote sensing. In [10], can be found a study about the use of networks of oscillator neurons applied to the segmentation of medical computerized tomographic images (CT) and magnetic resonance (MRI). A different application can be found in [17], where an oscillatory network is used to separate the speech of a speaker from interfering signals.

The study of such topologies of neural networks, which have direct biological inspiration on the mechanism of segmentation executed by the human brain, and applications of the same ones, is presented as a fertile field, as well as the development of architectures of hardware for practical implementation of the networks, exploring the parallel nature of them.

In this work, a topology of oscillatory neural network called LEGION is presented (Locally Excitatory Globally Inhibitory Oscillator Network), in both its continuous and discrete versions, based on the called Terman-Wang oscillator neuron. Also are presented some results relative to the use of a modified LEGION network for segmentation of images of silicon samples obtained by transmission electron microscopy (TEM).

2 The Terman-Wang Oscillator Neuron

The Terman-Wang oscillator is a relatively recent proposal, which, due to mathematical and computational reasons, has been used for composition of oscillatory artificial neural networks with application in segmentation of signals. The behavior of the Terman-Wang oscillator can be described through the pair of non-linear differential equations (1) and (2) [16]. The state variables $x(t)$ and $y(t)$ represent the system. Under the biological point of view, $x(t)$ can be understood as the potential of the nervous cell membrane, or the physical variable that represents the output in the neuron. α , β and ε are parameters of the model and I is an external input.

$$\frac{dx(t)}{dt} = 3x(t) - x^3(t) + 2 - y(t) + I \quad (1)$$

$$\frac{dy(t)}{dt} = \varepsilon(\alpha(1 + \tanh(x(t)/\beta)) - y(t)) \quad (2)$$

Fig. 1 presents the phase trajectory of the Terman-Wang oscillator with parameters $I=1$, $\beta=0,2$ and $\alpha=3$.

The functions $y_1(t)$ and $y_2(t)$ are obtained by making both the derivatives represented by (1) e (2) equal to zero. These functions are called *nullclines* of the system. The points, B and C correspond to examples of initial conditions for the presented differential equations. When the points B or C are used as initial conditions, the system acquires oscillatory response whose trajectory of phase is represented by the bold line, with the indicated direction. The initial condition represented by $x(0)=0$ and $y(0)=3$ implies an equilibrium situation for which oscillation does not occur, and must be avoided. Normally, a random noise of small amplitude can be added to the input in order to avoid such equilibrium state.

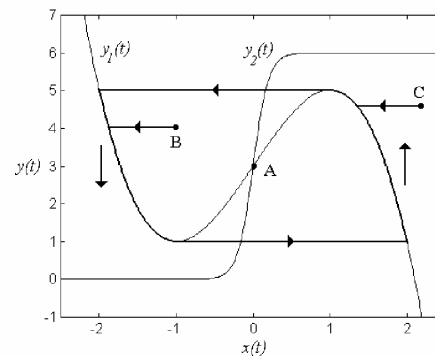


Figure 1 Phase trajectory of the Terman-Wang oscillator for $I=1$, $\beta=0,2$, $\alpha=3$.

Fig. 2 represents the output of the Terman-Wang oscillator in function of the time, characterizing the behavior of a relaxation oscillator. The intervals of time where the value of the output ($x(t)$) is high or positive are called active phases, while the intervals of time where the value of the output is low or negative are called silent phases [16].

In the case of $I \leq 0$, one says that the oscillator is excitable or it is not stimulated and there is no oscillation in the system [16]. For values of I that result in oscillatory behavior, a correlation can be observed between the value of the input and the value of the ratio of the times of permanence in the active and silent phases. In order to such oscillation occurs, the value of the external input received by the oscillator must be restricted to the band $0 < I < (2\alpha - 4)$, situation which the oscillator is said stimulated [16].

The Terman-Wang oscillator mathematically has a similar behavior to other models, being simpler than the Morris-Lecar oscillator and presenting bigger flexibility than the oscillators of van der Pol and of FitzHugh-Nagumo [16]. Such properties make the Terman-Wang

Artigo publicado nos anais do XIV Brazilian Symposium on Computer Graphics and Image Processing

oscillator a better alternative to implement networks of oscillators with several practical applications, including image segmentation.

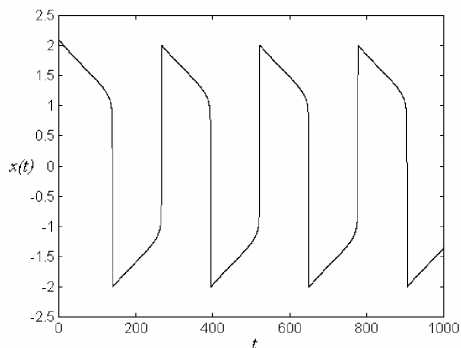


Figure 2 Output of the Terman-Wang oscillator in function of the time for $I=1$, $\beta=0.2$, $\alpha=3$, $\varepsilon=0.01$, $x(0)=0$ and $y(0)=0$.

3 Continuous LEGION Network

In the end of 80's decade, oscillations of approximately 40 Hz had been discovered in the visual cortex and in other areas of the human brain. It was verified that such neural oscillations present strong correlation with the coherence of the visual stimulus, occurring synchronism of phase between physically near neurons that receive similar external stimuli, what can characterize a homogeneous region of the perceived image. On the other hand, physically near neurons that receive different external stimuli or physically distant neurons that receive similar external stimuli do not present the cited synchronism of phase [16].

Since the discovery of the coherent neural oscillations, diverse types of networks composed by oscillator neurons have been studied with the objective to create artificial models for the phenomenon. Extending the study of Somers and Kopell [12], Terman and Wang had proved that in a network of relaxation oscillators with local coupling and arbitrary dimensions, an attraction domain exists for which whole the network tends to the synchronism with an exponential rate [13]. Using such property of local synchronism between coupled oscillators and adding a mechanism of global inhibition to get desynchronism among several groups of oscillators, Terman and Wang obtained a network they called LEGION. Moreover, a gaussian signal of small variance is added to the input of each oscillator to prevent that the initial conditions of the network imply not desired states of stability and also to prevent the possible synchronism

between distant oscillators with similar inputs [16]. The lateral excitation of an oscillator of the LEGION network (S_i), also added to the input, is defined as the coupling received from the others oscillators of the network and from the global inhibitor, being represented by the equation (3). Only the oscillators that belong to a near neighborhood possess coupling, as it can be observed in the example of topology presented in fig. 3. The global inhibitor ($z(t)$), on the other hand, is connected to all the oscillators of the network. W_{ik} are weights related to the connections between oscillators k and i , W_z is the weight of the global inhibitor, θ_x and θ_z are thresholds, $N_i(R)$ is the neighborhood of radius R of the oscillator i , $H(x-\theta)$ is the function of Heaviside, k is a parameter that controls the inclination of the function defined by (5) and ϕ adjusts the rate of variation of the global inhibitor output.

$$S_i(t) = \sum_{k \in N_i(R)} W_{ik} S_{\infty}(x_k(t), \theta_x) - W_z S_{\infty}(z(t), \theta_z) \quad (3)$$

$$\frac{dz(t)}{dt} = \phi(H(x_i(t) - \theta_z) - z(t)) \quad (4)$$

$$S_{\infty}(x, \theta) = \frac{1}{1 + e^{-k(x-\theta)}} \quad (5)$$

A mechanism used for normalization of the weights is proposed in [15]. The normalization of the weights is not a necessary condition for the correct functioning of the LEGION network, but it improves the quality of the synchronism between neighbor oscillators that are submitted to similar external stimuli. The W_{ik} weights can be determined in the beginning of the process representing the similarity between the external inputs of neighboring oscillators [15].

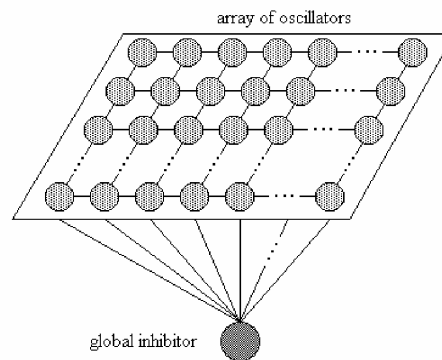


Figure 3 Example of a topology of a 2-dimensional LEGION network.

The fig. 4 presents the behavior of a LEGION network with 4 neurons hardwired in a 1-dimensional ring

Artigo publicado nos anais do XIV Brazilian Symposium on Computer Graphics and Image Processing

architecture, where it can be observed the synchronism between neighbor oscillators with similar external inputs and desynchronism between oscillators with distinct external inputs. The temporal discrimination of groups of physically near neurons with similar inputs can be observed in the output of the network.

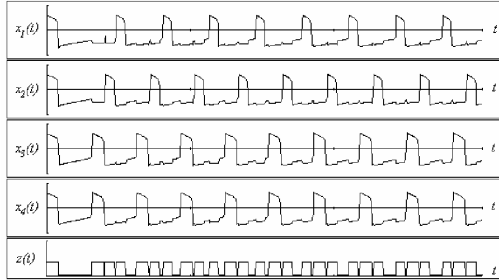


Figure 4 Outputs of the oscillators and of the global inhibitor of a LEGION network with external inputs $I_1=0,3$, $I_2=0,4$, $I_3=0,5$ and $I_4=0,5$.

For a set of parameters, a continuous LEGION network can discriminate a limited number of segments, which depends on the ratio between the permanence times in the active and silent phases. This limit is called capacity of segmentation of the network and places in the band of 5 to 10 segments [16]. This property, although presenting itself as a restriction to practical applications of the network in segmentation, has biological correlation because it is also observed in human beings, which present a quantitative limitation for discrimination of several objects at the same time [15].

In [15] is also presented a modification in the excitation of the Terman-Wang oscillator, considering the basic idea that a set of oscillators with similar inputs must possess at least one oscillator, called leader oscillator, which must receive great lateral excitation from its neighborhood. On the other hand, isolated oscillators that belong to noisy regions, cannot be characterized as leaders. So, an oscillator with great lateral potential can lead the activation of a group of oscillators that correspond to a homogeneous region of the external input.

The results found in the bibliography demonstrate that the continuous LEGION network presents itself as a potentially efficient tool for image segmentation. As positive aspect, can be pointed the property of simultaneous spatial and temporal segmentation presented by the network. As negative aspect, can be pointed the high computational complexity, due to the great number of differential equations to be solved. On the other hand, the use of the model based on differential equations is

presented as an alternative for realization of a full parallel network using analog circuits, associating the high speed presented for such circuits and the inherent high level of parallelism of the LEGION network.

4 Discrete LEGION Network

With the objective to reduce the computational effort in the use of serial computers, Terman and Wang had developed an algorithmic discrete version of the LEGION network, with basically the same properties of the continuous version [15]. The limitation related to the number of segments to be distinguish presented by the continuous network is not verified in this proposal, adding a positive feature to the same one. In the developed algorithm, the state variable $x(t)$ is the only one considered, since the same represents the output of the oscillator. A neighborhood relation was used considering the 8 neighbor pixels for determination of the lateral excitation of each oscillator in a 2-dimensional network used for applications of image segmentation, what can be easily modified. The algorithm can be divided in three great stages, which are presented to follow:

1) Phase of initialization

- Make the global inhibitor $z(t)$ equal to zero:

$$z(0) = 0$$

- Calculate the weights of the connections between neighbor oscillators (W_{ik}) considering the input image, where I_M the maximum intensity found among the pixels of the same one:

$$W_{ik} = \frac{I_M}{1 + |I_i - I_k|}, k \in N(i) \quad (6)$$

- Find the leaders p_i using the function of Heaviside with a threshold θ_p :

$$p_i = H \left(\sum_{k \in N(i)} W_{ik} - \theta_p \right) \quad (7)$$

- Set the outputs of all the oscillators in random positions in the silent phase:

$$-2 < x_i(0) < -1$$

2) Determination of the first oscillator to pass to the active phase

- Considering that all the oscillators (x_k) are in the silent phase, choose the leader oscillator ($p_j=1$) which will be the oscillator with its state next to the point of transition for the active phase (state $x=-1$) and carry it to the active phase (state $x=1$), increment the global inhibitor ($z(t)=1$) and recalculate the states of the others oscillators in the silent phase:

Artigo publicado nos anais do XIV Brazilian Symposium on Computer Graphics and Image Processing

$$[x_j(t+1)=1 \text{ and } z(t+1)=1]$$

$$\text{for } [x_j(t) \geq x_k(t) \text{ , } \forall k] \quad (8)$$

$$\text{and } x_k(t+1) = x_k(t) - 1 - x_j(t) \text{ , } k \neq j \quad (9)$$

3) Dynamics of the network

- Keep the oscillators that are in the active phase if the global inhibitor was incremented. Return the oscillators that are in the active phase to the beginning of the silent phase (defined as the state $x=-2$) if the global was decreased or kept constant. For each oscillator that returns to the silent phase decrement the global inhibitor. When no more oscillator neurons exist in the active phase ($z(t)=0$), return to the stage 2.

$$[x_i(t+1) = x_i(t)]$$

$$\text{if } [z(t) > z(t-1) \text{ and } x_i(t) = 1] \quad (10)$$

$$\text{and } [x_i(t+1) = -2 \text{ and } z(t) = z(t-1) - 1]$$

$$\text{if } [z(t) \leq z(t-1) \text{ and } x_i(t) = 1] \quad (11)$$

- Find the lateral excitation of the oscillators that are in the silent phase according to the criterion established for (12) or other alternatives presented in [1]. Those oscillator neurons that possess enough potential will pass to the active phase and the global inhibitor will be incremented.

$$S_i(t) = \sum_{k \in N(t)} W_{ik} H(x_k(t)+1) - W_z H(z(t)-0.5) \quad (12)$$

$$[x_i(t+1) = 1 \text{ and } z(t) = z(t-1) + 1]$$

$$\text{if } [S_i(t) > 0 \text{ and } x_i(t) < 1] \quad (13)$$

$$\text{and } [x_i(t+1) = x_i(t) \text{ and } z(t) = z(t-1)]$$

$$\text{if } [S_i(t) \leq 0 \text{ and } x_i(t) < 1] \quad (14)$$

- Execute the dynamics of the network until reach the condition of no oscillators in the active phase and then go to the stage 2.

The algorithm presented for the LEGION network considers that the leaders and the weights of the connections must be determined in the initialization phase, not needing posterior update. It must be observed that the number of leaders does not correspond to the number of discriminated segments, since each segment can generate several leaders.

The results presented in the bibliography demonstrate that the LEGION network, in its algorithmic version are efficient for segmentation of signals of diverse natures and, in particular, for segmentation of images. The significant increase in the speed makes the discrete version of LEGION network more suitable for implementations that use serial computers.

5 Sensitivity to Noise of the LEGION Network

Preliminary evaluations had demonstrated that the result of the segmentation using LEGION network in its two versions is sensible to the noise present in the image signal, due basically to the use of the intensity similarity criterion between adjacent pixels.

In [1,2] is proposed an algorithm for smoothing called FPS (Feature-Preserving Smoothing), which reduces the noise while preserving relevant details found in the original image. The related algorithm can be used with the algorithms presented for the LEGION network, promoting a gradual update of the weights, whose effect is equivalent to the one of a filtering that attenuates the noise and simultaneously preserves details of the original image [2]. The FPS algorithm can also be directly used in a previous stage of the LEGION network, acting on the pixels of the image as a non-linear filter. The equation (15) presents the cited procedure of smoothing for application over images [2]. I_{ij}^t is the intensity of pixel ij at iteration t , D_{ij}^t is a measure of the local discontinuities around the pixel, σ_{ij}^t is the normalized variance of the intensities of the pixels calculated in a neighborhood of radius R of the pixel ij , θ_σ is a threshold for the influence of the variance and k_1 and k_2 are weight parameters.

$$I_{ij}^{t+1} = I_{ij}^t + \frac{\sum_{(m,n) \in N_{ij}(1)} (I_{mn}^t - I_{ij}^t) e^{-(k_1 \Phi(\sigma_{mn}^2, \theta_\sigma) + k_2 D_{mn}^t)}}{\sum_{(m,n) \in N_{ij}(1)} e^{-(k_1 \Phi(\sigma_{mn}^2, \theta_\sigma) + k_2 D_{mn}^t)}} e^{-k_1 \Phi(\sigma_{ij}^2, \theta_\sigma)} \quad (15)$$

$$D_{ij}^t = \frac{1}{4} \left(|I_{i-1,j}^t - I_{i+1,j}^t| + |I_{i,j-1}^t - I_{i,j+1}^t| + |I_{i-1,j-1}^t - I_{i+1,j+1}^t| + |I_{i-1,j+1}^t - I_{i+1,j-1}^t| \right) \quad (16)$$

$$\Phi(\sigma_{ij}^2, \theta_\sigma) = \begin{cases} \sigma_{ij}^2 & , \sigma_{ij}^2 \geq \theta_\sigma \\ 0 & , \sigma_{ij}^2 < \theta_\sigma \end{cases} \quad (17)$$

6 Segmentation of TEM Images

The fig. 5 presents an image with dimensions of 1482x2060 pixels and resolution of 8 bits in gray level scale, which was obtained through transmission electron microscopy (TEM), representing a silicon sample in which had been implanted ions of helium with the objective to reduce defects in the crystalline structure of the semiconductor. Due to its physical properties, the helium tends to accumulate itself in small bubbles concentrated in one determined depth of the silicon sample. For evaluation of the process, whose result depend on factors as the temperature of implantation, it is necessary to determine the amount of helium bubbles in

Artigo publicado nos anais do XIV Brazilian Symposium on Computer Graphics and Image Processing

one determined region, as well as the volume of the present gas, which can be estimated through the diameter or the area of the bubbles [11]. This procedure must be carried through for several images, being a complex task for manual execution by a human being.

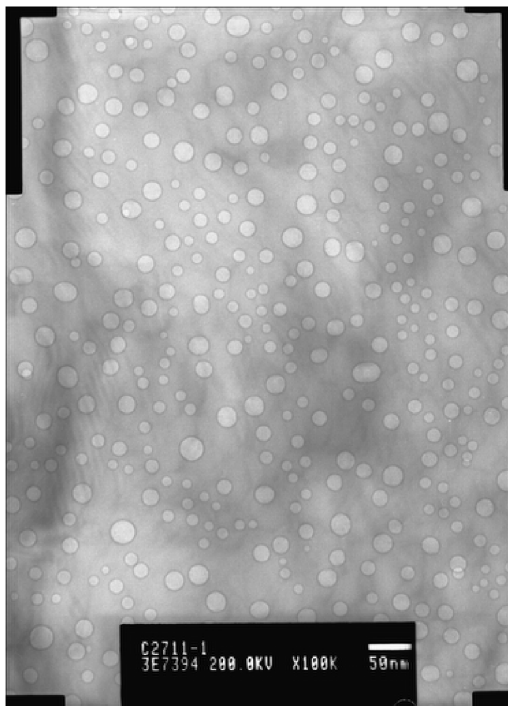


Figure 5 Image of TEM representing a silicon sample with helium bubbles.

To segment the helium bubbles and simultaneously separate the same ones in time, facilitating the implementation of an automatic process of counting and measuring its areas or diameters, it was opted to use a discrete LEGION network with the same dimensions of the image (1482x2060 oscillator neurons). The fig. 6 presents an region extracted from the image of the fig. 5, which is used to verify the qualitative results of the proposed segmentation procedure.

Fig. 7 presents the result of filtering the image of fig. 6 using the FPS algorithm, being able to be observed the preservation of the contours of the helium bubbles. The used parameters had been $R=3$, $\theta_o=90$, $k_1=20$, $k_2=5$ and 10 iterations.

Fig. 8 presents the segments related to the 9 bubbles segmented for the LEGION network using the image of

the fig. 7 as input. The parameters used in the network had been $\theta_p=4$ and $W_s=0,4$. The black color represents the oscillators in the silent region (inactive), while that the white regions represent the oscillators with output in the active phase. An additional segment, presented in fig. 9, presents the silicon background and also is supplied by the network, which, due to its very big area, can easily be detected and discarded in the automatic procedure of measure. Pixels pertaining to not homogeneous regions are not incorporated to the extracted segments got from the LEGION network. Fig. 10 presents pixels that had not been segmented by the network (represented in black color) relative to the image of fig. 7.

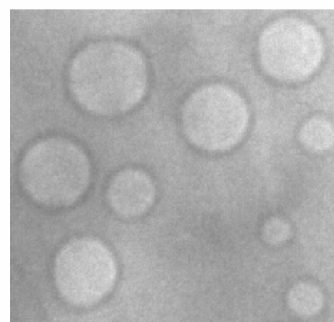


Figure 6 Region extracted from the image of fig. 5.

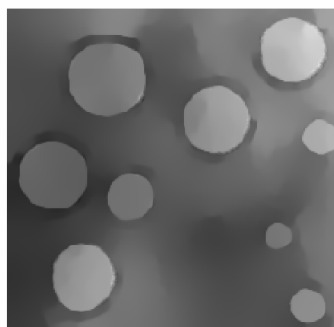


Figure 7 Image of fig. 6 filtered with the FPS algorithm.

The relative areas of each one of the bubbles can quickly be obtained from the segments presented in fig. 7, as the ratio between the number of oscillators with output in the active phase and the total number of oscillators of the network, which corresponds to the number of pixels of the image. Being known the relation in pixel/nm of the image obtained of a TEM (showed near de right bottom corner in fig. 5), the volume of helium in the region

Artigo publicado nos anais do XIV Brazilian Symposium on Computer Graphics and Image Processing

represented by the image can then be estimated. It must be observed that the visual determination of the area of each bubble (or of the average diameter) can be inexact in the cases in which the bubbles are not perfectly circular. Such limitation is not observed in the considered method, since the relative area of each segment can be determined with exactness, independently of the form of the same. On the other hand, incomplete bubbles, located at the extremities of the images, bubbles with very degraded contours and bubbles superimposed lead to wrong measures. To avoid such limitation, an interactive procedure was adopted, where points on the regions to be segmented are marked (bubbles). Such points are used to determine the leader oscillators of the LEGION network. In this way, the homogeneous regions around the indicated leaders will be the only regions segmented by the network. The bubbles not marked, in number relatively reduced, can be measured later through a manual procedure.

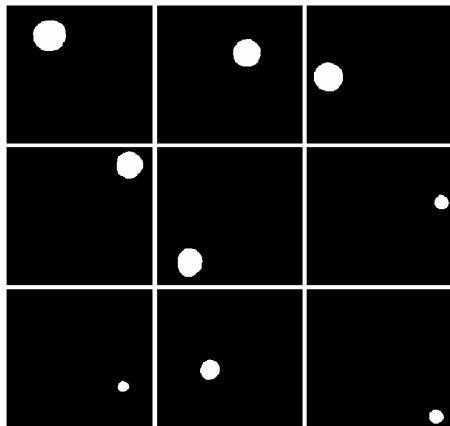


Figure 8 Result of the segmentation of the image of fig. 7 with a LEGION network.

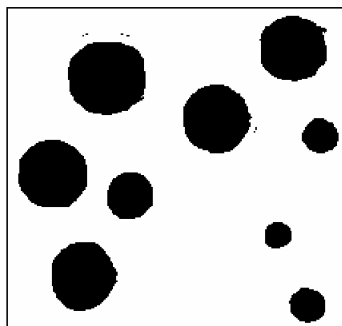


Figure 9 Segment related to the homogeneous silicon background.

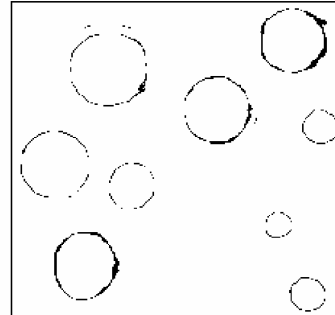


Figure 10 Pixels not segmented by the network (represented in black color).

7 Conclusion

The use of oscillator networks to simulate the capacity of temporal segmentation that is given credit to possess the human brain is presented as a recent alternative and with satisfactory results. The researchers DeLiang L. Wang and David Terman have developed works where a called LEGION network (Locally Excitatory Globally Inhibitory Oscillator Network) is proposed, which is constructed based on relaxation oscillators locally connected and with a mechanism of global inhibition. This mechanism of connections allows that neighboring oscillators submitted to similar external inputs oscillate in synchronism of phase and desynchronism with other groups of oscillators of the network. Of this form, each group of oscillators corresponding to coherent regions of the input signal of the network is activated in a different time interval, propitiating both spatial and temporal segmentation of the input signal simultaneously.

The Terman-Wang oscillator, used as the basic element of processing of the LEGION network, is based on a pair of differential equations, what implies high computational complexity for application of the network in non-parallel machines. An algorithm for implementation of the LEGION network in serial computers was proposed by Terman and Wang, having presented a significant gain of speed in relation to the continuous version of the network and still supplanting the limitation of the continuous version related to the capacity of segmentation for simultaneous discrimination.

One of the key points for the correct operation of the LEGION network, as much in its continuous version how much in its discrete version it is the determination of the weights of the connections between neighboring oscillators. According to proposal of Terman and Wang, the related weights can be determined in the beginning of the process, in a single step, based on some attributes of

Artigo publicado nos anais do XIV Brazilian Symposium on Computer Graphics and Image Processing

the signal applied to the input of the network, what implies relatively low computational complexity to develop such task. Another relevant aspect says respect to the determination of the parameters of the network, θ_p and W_z , which must interactively be determined for the correct operation of the network. The quality of the segmentation supplied by the network is also strongly related with the quality of the image to be segmented, leading to the necessity of application of a previous procedure of smoothing, such as the FPS filtering presented, which has the property of drastically attenuate the noise in the image, preserving relevant characteristic, as the borders between distinct regions.

The work developed by Terman and Wang includes applications of LEGION network for segmentation of signals of diverse natures, including medical images, images of remote sensing and signals of speech, having been found favorable results. In this work, the application of the LEGION network in the segmentation of TEM images to estimate the volume of helium in silicon samples, qualitatively shows that this network presents itself as a singular tool for applications that requires simultaneously both spatial and temporal image segmentation. The procedure of interactive indication of leaders, proposed in this work, forces the LEGION network to segment only the respective regions, avoiding undesirable segments to appear as outputs of the network.

Bibliography

- [1] CHEN, Ke; WANG, DeLiang; LIU, Xiuwen. "Weight Adaptation and Oscillatory Correlation for Image Segmentation". *Technical Report OSU-CISRC-8/98-TR37*. 1998. Available in <http://www.cis.ohio-state.edu/~dwang/papers/> (17/8/2000).
- [2] CHEN, Ke; WANG, DeLiang. "Image Segmentation Based on a Dynamically Coupled Neural Oscillator Network". *Proceedings of IJCNN*. 1999.
- [3] DAHMER, Alessandra. *Segmentação de Imagens Ecocardiográficas Utilizando Redes Neurais e Medidas de Textura*. Master Thesis. Porto Alegre: UFRGS-PGCC, 1998. 98p.
- [4] GHOSH, A.; PAL, N. R.; PAL, S. K. "Object background classification using Hopfield type neural network". *Int. J. Pattern Recognition Artificial Intelligence*, v.6, n.5, p.989-1008, 1992.
- [5] HAYKIN, S. *Neural Networks - A Comprehensive Foundation*. New York: Macmillian College Publishing Company, 1994. 696p.
- [6] LIU, Xiuwen; CHEN, Ke; WANG, DeLiang L. "Extraction of hydrographic regions from remote sensing images using an oscillator network with weight adaptation". *IEEE Transactions on GeoScience and Remote Sensing*. 2000. Available in <http://www.cis.ohio-state.edu/~dwang/papers/LCW00.ps.gz> (17/8/2000).
- [7] MOHAMED, N. A.; AHMED, M. N.; FARAG, A. "Modified Fuzzy C-Mean in Medical Image Segmentation". 1998. Available in <http://www.cvip.uofl.edu/Papers/fuzzy3.ps.gz> (16/8/2000).
- [8] PANDYA, Abhijit S.; MACY, Robert B. *Pattern Recognition with Neural Networks in C++*. Boca Raton: CRC Press, 1996. 410p.
- [9] PICCOLI, Luciano. *Segmentação e Classificação de Imagens Ecocardiográficas Utilizando Redes Neurais*. Master Thesis. Porto Alegre: UFRGS-PGCC, 1999. 90p.
- [10] SHAREEF, Naeem; WANG, DeLiang L.; YAGEL, Roni. *Segmentation of Medical Images Using LEGION*. 1997. Available in <http://www.cis.ohio-state.edu/~shareef/OSUTR26/journal-1.html> (17/8/2000).
- [11] da SILVA, D. L.; FICHTNER, P. F. P.; PEEVA, A.; BEHAR, M.; KOEGLER, H.; SKORUPA, W. "The effects of implantation temperature on He bubble formation in silicon". *Nuc. Inst. and Methods*. B175-77. P.336-340, 2001.
- [12] SOMERS, D.; KOPELL, N. "Rapid synchrony through fast threshold modulation". *Biologic Cybernetics*. v.68, p.393-407, 1993.
- [13] WANG, DeLiang. "Modeling global synchrony in the visual cortex by locally coupled neural oscillators". *Proceedings of 15th Annual Conference on Cognitive Science Society*, 1058-1063, 1993.
- [14] WANG, DeLiang; TERMAN, David. "Image Segmentation Based on Oscillatory Correlation". *Technical Report 19*, The Ohio State University Center for Cognitive Science, 1996.
- [15] WANG, DeLiang; TERMAN, David. "Image Segmentation Based on Oscillatory Correlation". *Neural Computation*, v.9, p.805-836, 1997.
- [16] WANG, DeLiang. "Relaxation Oscillators and Networks". *Wiley Encyclopedia of Electrical and Electronics Engineering*, v.18, p.396-405, 1999.
- [17] WANG, DeLiang; BROWN, Guy J. "Separation of Speech from Interfering Sounds Based on Oscillatory Correlation". *IEEE Transactions on Neural Networks*, v.10, n.3, p.684-697, 1999.

Artigo publicado nos anais do 14th Symposium on Computer Architecture and High Performance Computing

Architecture of Oscillatory Neural Network for Image Segmentation

Dênis Fernandes^{1*}, Jeferson Polidoro Stedile^{2**} and Philippe Olivier Alexandre Navaux²

¹PUCRS - Pontifícia Universidade Católica do Rio Grande do Sul - Faculdade de Engenharia, Porto Alegre, RS, Brazil
denis@ee.pucrs.br

²UFRGS - Universidade Federal do Rio Grande do Sul - Instituto de Informática, Porto Alegre, RS, Brazil
jstedile@terra.com.br, navaux@inf.ufrgs.br

Abstract

Oscillatory neural networks are a recent approach for applications in image segmentation. In this context, the LEGION (Locally Excitatory Globally Inhibitory Oscillator Network) is the most consistent proposal. As positive aspects, the network has got parallel architecture and capacity to separate the segments in time. On the other hand, the structure based on differential equations presents high computational complexity and limited capacity of segmentation, which restricts practical applications. In this paper, a proposal of parallel architecture for implementation of an oscillatory neural network suitable for image segmentation is presented. The proposed network keeps the positive features of the LEGION network, offering lower complexity for implementation in digital hardware and capacity of segmentation not limited, as well as few parameters, with intuitive setting. Preliminary results confirm the successful operation of the proposed network in applications of image segmentation.

1. Introduction

Several techniques of image segmentation based on artificial neural networks have been developed, particularly using MLP (MultiLayer Perceptron) networks associated to the Backpropagation algorithm, Hopfield networks or Kohonen maps (SOFM - Self-Organizing Feature Map), with typical examples found in [5][6][10][13]. Such algorithms have presented good results even with the existence of noise or distortions in the image to be segmented, having several advantages

inherent to neural network topologies. One of the drawbacks of some of these methods is the necessity of training, which can be problematic due to the long time needed and the high number of available samples previously segmented. Another negative aspect is the high computational complexity presented by some of these architectures.

Recently, alternative topologies of artificial neural networks, called oscillatory neural networks, have been applied in procedures of image segmentation with favorable results [2][3][4][11][16]. The study of such topologies of neural networks, which have biological inspiration on the mechanism of segmentation supposedly executed by the human brain, as well as their applications, are fertile ground for work. The necessity of efficiency in applications of image processing highly justifies the development of new hardware architectures for practical implementation of such networks, exploring their parallel nature [2][4].

The LEGION network (Locally Excitatory Globally Inhibitory Oscillator Network) [17] is the most consistent proposal of oscillatory neural network with practical applications in image segmentation found in the bibliography. Applications of the LEGION network include, for example, a procedure associated to MLP networks used for extraction of hydrographic regions in remote sensing images [12]. In [14], a study on the use of the LEGION network applied to the segmentation of medical images of computerized tomography and magnetic resonance can be found. In [7], a LEGION network is used to segment images of electronic microscopy in a practical procedure for counting and measuring helium bubbles implanted in silicon. In [8], a LEGION network is associated to Kohonen maps to isolate the region of the left ventricle in a procedure to analyze fetal echocardiographic images. Other applications of the LEGION network can also be found in the bibliography, such as, for example, the separation of a speaker's talk from interfering noise [19].

* Professor at Electrical Engineering Department of PUCRS and Ph.D. student at PGCC of UFRGS.

** Student of Electrical Engineering at PUCRS and CNPq scholarship researcher at UFRGS.

Artigo publicado nos anais do 14th Symposium on Computer Architecture and High Performance Computing

As a positive characteristic of the LEGION network, its highly parallel architecture is very attractive, which can be expressed as high operation speed in implementations using adequate hardware. Another interesting aspect is related to the capacity of the network to segregate in time the various segments of the input image, which can facilitate the later identification and quantification of these regions. As negative aspects, its high computational complexity for implementation in digital hardware, due to its structure based on sets of differential equations, and the limited capacity for simultaneous discrimination of different segments can be pointed out. The high amount of parameters and their little intuitive setting are also limiting factors.

The present work aims at presenting a parallel architecture for implementation of an oscillatory neural network in digital hardware, which is suitable for applications in image segmentation. The proposed network offers similar principles of operation to the ones of the LEGION network, possessing parallel architecture and capacity to discriminate in time the segments found. The lower computational complexity, due to the predominant use of binary logical operations, and the non-limitation of the number of segments to be discriminated are presented as the main advantages of the proposed network in relation to the LEGION network. The use of few parameters, with intuitive setting, is also presented as a positive aspect, as well as the reduced number of iterations and the easy prevision of the time necessary to get the intended result.

Results found in practical implementations are presented and prove the successful operation of the proposed architecture according to the expectations.

2. The LEGION Network

In the late 80's, oscillations of approximately 40 Hz were discovered in the visual cortex and other areas of the human brain. It was verified that such neural oscillations have a strong correlation with the coherence of the visual stimulus, and synchronism of phase occurs between neurons which are physically near and that receive similar stimulus, which can characterize a homogeneous region of the perceived image. On the other hand, physically near neurons that receive different stimulus or physically distant neurons that receive equal stimulus do not present such synchronism of phase [18]. Using such property of local synchronism between oscillators and adding a mechanism of global inhibition to get anti-synchronism among groups of oscillators, the LEGION network was conceived.

The fig. 1 presents a 2-dimensional architecture of LEGION network adjusted to applications of image segmentation. The network is basically formed by a set of Terman-Wang oscillators [18]. Only neighboring

oscillators are connected, as presented in fig. 1. A global inhibitor, on the other hand, is connected to all the oscillators of the network.

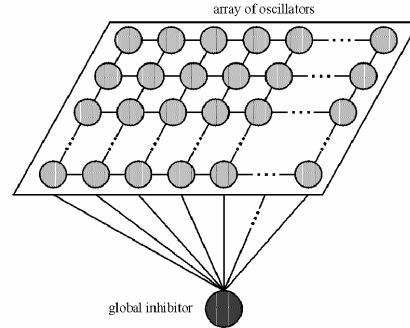


Figure 1. Example of LEGION network with 2-dimensional topology.

The behavior of the Terman-Wang oscillator can be described through the differential equations (1) and (2) [18]. From the biological point of view, $x(t)$ can be understood as the potential of the nervous cell membrane, or either, the physical variable that represents the output in the neuron. α , β and ε are parameters of the model and I is an external input. A gaussian signal of small variance is added to the input of each oscillator to avoid that the initial conditions of the network imply unwanted states of stability and also to prevent the possible synchronism between distant groups of oscillators with similar inputs [18]. The lateral excitation of an oscillator of the LEGION network, $S_i(t)$, is also added to the input, and it is defined as the coupling received from the other oscillators of the network, being represented by the equation (3). W_{ik} represents the weights related to the connections between the neighboring oscillators k and i , being correlated with the similarity between the external inputs of such oscillators. W_z is the weight related to the global inhibitor, θ_x e θ_z are thresholds, $N_i(R)$ means the neighborhood of radius R around the oscillator i , $H(x)$ is the Heaviside function, κ is a parameter of the sigmoid defined by (5) and ϕ adjusts the variation rate of the global inhibitor output $z(t)$ (4) [17].

$$\frac{dx(t)}{dt} = 3x(t) - x^3(t) + 2 - y(t) + I + S_i(t) + \rho \quad (1)$$

$$\frac{dy(t)}{dt} = \varepsilon(\alpha(1 + \tanh(x(t)/\beta)) - y(t)) \quad (2)$$

$$S_i(t) = \sum_{k \in N_i(1)} W_{ik} S_{\infty}(x_k(t), \theta_x) - W_z S_{\infty}(z(t), \theta_z) \quad (3)$$

$$\frac{dz(t)}{dt} = \phi(H(x_i(t) - \theta_z) - z(t)) \quad (4)$$

$$S_{\infty}(x, \theta) = \frac{1}{1 + e^{-\kappa(x-\theta)}} \quad (5)$$

A mechanism of weight normalization is presented in [17]. It is shown that the normalization of the weights is not a necessary condition for the correct functioning of the network, however it improves the synchronism between neighboring oscillators submitted to similar inputs.

For a determined set of parameters, a LEGION network can discriminate a limited number of segments, which depends on the ratio between the times of permanence of the neurons in the phases active and silent. This limit is called capacity of segmentation of the LEGION network and it is placed in the band of 5 to 10 segments [18]. Although such feature presents itself as a restriction to the application of the network in image segmentation, it has biological correlation, because it is also observed in human beings, which present a quantitative limitation for simultaneous discrimination of several objects [17].

In [17], a modification in the excitation of the Terman-Wang oscillator is also presented, considering the idea that a set of oscillators with similar inputs must possess at least one leader oscillator, which must receive great lateral excitation from its neighborhood. On the other hand, isolated oscillators, belonging to noisy fragments, cannot be characterized as leaders. This way, an oscillator with great lateral potential can lead the activation of a group of oscillators corresponding to a homogeneous region of the input.

Several simulations demonstrate that the LEGION network is capable to produce satisfactory results in image segmentation applications [9]. On the other hand, the necessity of a careful setting of the network parameters for each different image to be segmented was evidenced. Such procedure is interactive and not much practical due to the high number of parameters and their complex joint influence on the results obtained. In several cases, some difficulty in getting synchronism between oscillators was observed as well as the necessity of a great and not easily estimable number of iterations to reach the desired result.

One concludes through several tests and results found in the bibliography that the LEGION network is a solution suitable for image segmentation applications. As positive aspects, its property of time separation of the segments and the parallel topology, which can bring on fast implementations using adequate hardware architecture, can be pointed out. As negative aspects, the network presents high computational complexity for implementation in digital hardware, given the great number of differential equations to be solved, and also limited capacity of segmentation and high number of parameters with little intuitive setting procedures.

3. Proposal of the Network Architecture

In face of the practical limitations presented by the LEGION network, a new architecture of oscillatory neural network was conceived, suitable for image segmentation applications and also to implementation in digital hardware with parallel topology.

Like the LEGION network, in case of image segmentation applications, the proposed network is implemented in a 2-dimensional topology with the same dimensions of the image (equal number of pixels and oscillator neurons). Each pixel of the input image will have therefore a corresponding oscillator in the network.

The 2-dimensional proposed network presents two structures of connections, making it different, in this aspect, from the LEGION network. Fig. 2 presents an example of the structure of the excitatory connections. A neighborhood relation was adopted where an oscillator neuron has its excitatory output $v_e(i, j)$ connected simultaneously to the 8 nearest neurons. This way, the excitatory output of the oscillator neuron will be always active when at least one of its nearest neighbors, submitted to a similar input, is in the active phase. For the neurons situated in the borders of the network, the weights related with the nonexistent connections are null, as well as the respective excitatory inputs.

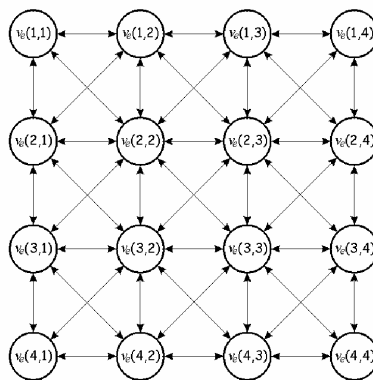


Figure 2. Example of the structure of excitatory connections proposed.

Fig. 3 presents an example of the structure of inhibitory connections adopted. Each neuron has its inhibitory output $v_i(i, j)$ connected only to one neighboring neuron. This structure establishes a priority among the oscillator neurons of the network in a way that, if there are several neurons enabled to pass to the active phase, pertaining to different segments, only one of them, the one with the highest priority, will do. In other words, the neuron with the highest priority inhibits the remainders, keeping them in the silent phase. On the other hand, the excitatory

Artigo publicado nos anais do 14th Symposium on Computer Architecture and High Performance Computing

connections will cause the only neuron qualified by the structure of inhibitory connections to qualify all the other neurons belonging to the respective segment. Different structures of inhibitory connections could be used such as a helicoidal structure initiating at one extremity or at the center of the network, for example. In the topology of inhibitory connections presented, the neuron with the highest priority always receives null inhibition in its input. The inhibitory output of the neuron with the lowest priority in the network will be active only when at least one neuron of the network has its output in the active phase. In a practical implementation, this signal can be used to detect null images, without any segment, present in the output of the network.

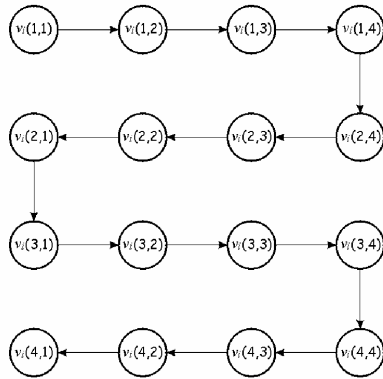


Figure 3. Example of the structure of inhibitory connections proposed.

4. Proposal of the Network Neuron Structure

During the development of the structure of the basic processing element (oscillator neuron) of the proposed network, simplicity was prioritized, bearing in mind a later implementation in a dedicated digital hardware. So, the use of operations and logical structures with easy implementation in digital programmable devices such as the FPGAs (Field Programmable Gate Arrays) was adopted.

Fig. 4 presents the structure of the oscillator neuron of the proposed network adjusted to a 2-dimensional topology. In order to make better understanding of the functioning of the neuron possible, the definitions of the constants and variables are listed below:

- N_n : total number of oscillator neurons in the network;
- N_s : adjustable parameter that sets the maximum number of segments to be separated by the network;
- L_m : adjustable parameter that sets a threshold to similarity determination between the inputs of neighboring neurons;

- $v_c(i,j,t)$: state of the internal counter belonging to the neuron placed at line i and column j in time (iteration) t ;
- $v_o(i,j,t)$: output signal of the internal comparator belonging to the neuron placed at line i and column j in time t ;
- $v_e(i,j,t)$: excitatory output of the neuron placed at line i and column j in time t ;
- $v_i(i,j,t)$: inhibitory output of the neuron placed at line i and column j in time t ;
- $v_i(i',j',t)$: inhibitory output of the precedent neuron related to the neuron placed at line i and column j in time t ;
- $v_l(i,j,t)$: leader indication signal related to the neuron placed at line i and column j in time t ;
- $I_o(i,j,t)$: input signal of the neuron placed at line i and column j in time t (in case of an image, it can represent a set of features or only one feature such as the intensity of the related pixel, for example);
- $w(i,j,k,l,t)$: weight related to comparison between the input of the neuron placed at line i and column j and the input of the neuron placed at line k and column l , in time t .

A binary counter with a total number of states higher or equal to the number of neurons of the network (N_n) is used to define the internal state of the neuron as a function of the time, $v_c(i,j,t)$. The referred counter always has its state returned to zero (synchronous reset) if the neuron is in its active phase ($v_e(i,j,t)=1$), otherwise its state is incremented until reaching the N_s-1 state, remaining so until receiving one reset, as represented by the equation (6). The use of N_n states in the counter makes the discrimination of a maximum of N_n segments possible (one state for each segment), an extreme case in which there is no similarity between any pair of inputs of neighboring neurons.

$$v_c(i,j,t) = \begin{cases} v_c(i,j,t-1)+1 & \text{if } \begin{cases} v_c(i,j,t-1) < N_s - 1 \\ v_e(i,j,t-1) = 0 \end{cases} \\ 0 & \text{if } v_e(i,j,t-1) = 1 \\ v_c(i,j,t-1) & \text{if } \begin{cases} v_c(i,j,t-1) \geq N_s - 1 \\ v_e(i,j,t-1) = 0 \end{cases} \end{cases} \quad (6)$$

The output of the internal counter of the neuron is sent to a comparator whose output $v_o(i,j,t)$ is equal to one when its input is equal or higher than the number of segments to be discriminated by the network minus one, being, otherwise, null (7). Such procedure qualifies the network to separate the N_s first segments according to the sequence established by the structure of inhibitory connections of the network. In the extreme case in which $N_s=N_n$, the network can separate up to N_n segments.

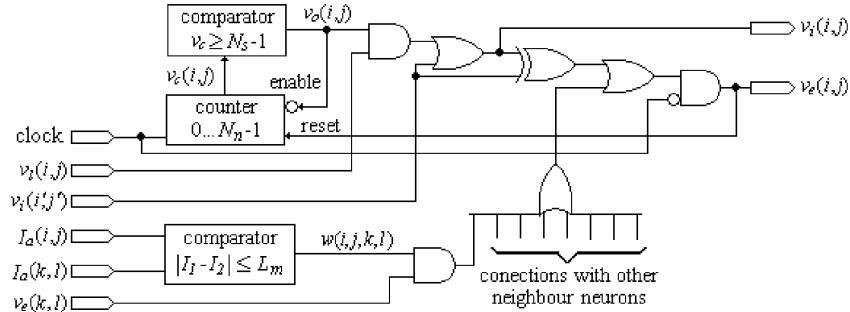


Figure 4. Neuron internal structure of the proposed network.

$$v_o(i, j, t) = \begin{cases} 1 & \text{if } v_c(i, j, t) \geq N_s - 1 \\ 0 & \text{if } v_c(i, j, t) < N_s - 1 \end{cases}, N_s \leq N_n \quad (7)$$

The equation (8) represents the inhibitory output of the oscillator neuron $v_i(i, j, t)$, which is obtained through a binary logical operation among the output of its internal comparator $v_o(i, j, t)$, the signal of qualification $v_l(i, j, t)$ and the inhibitory output received from its preceding oscillator $v_i(i', j', t)$. Thus, the inhibitory output of the neuron will be high only when its internal state is equal or higher than $N_s - 1$ and it is qualified as a leader ($v_l(i, j, t) = 1$) or that it has received an inhibition from its preceding neuron.

$$v_i(i, j, t) = (v_o(i, j, t) \text{ and } v_l(i, j, t)) \text{ or } v_i(i', j', t) \quad (8)$$

The qualification signal $v_l(i, j, t)$ can be generated internally or externally and only the oscillators with $v_l(i, j, t) = 1$ will be initially qualified to pass to the active phase (leader oscillators). Oscillators with $v_l(i, j, t) = 0$ can only be activated through a neighbor with similar input that is in the active phase. In the external generation of the signal, the leader oscillators can be established by some criteria such as the corresponding position in the image, for example. Thus, it could be established that corresponding oscillators to the central region of the image would be the only ones qualified to pass to the active phase, causing the appearance of only that segment at the output of the network. In the internal generation of the signal, one can use, for example, the criteria that a leader oscillator must possess all the weights related to the unitary excitatory connections, which corresponds to a pixel in the center of a homogeneous region. Such condition can easily be carried out by using a Boolean operation and among the excitatory weights of the respective neuron.

The excitatory output of the oscillator neuron $v_e(i, j, t)$, represented by equation (9), is unitary when the neuron is qualified, considering its internal state, the condition of leader and the inhibition received from its preceding

neuron, or if it receives an excitatory signal ($v_e(k, l, t) = 1$) from a neighboring neuron whose input is similar to its own, implying the respective weight $w(i, j, k, l, t) = 1$. A neuron can receive excitatory signals from several neurons pertaining to a predefined region of neighborhood. The clock signal is used to define the initial states of all the excitatory outputs of the network. Neurons that have their outputs in high logical level are said to be in the active phase ($v_e(i, j, t) = 1$), and otherwise, they are said to be in the silent phase ($v_e(i, j, t) = 0$).

$$v_e(i, j, t) = \overline{\text{clock}} \text{ and } ((v_l(i, j, t) \text{ xor } v_i(i', j', t)) \text{ or } (w(i-1, j-1, t) \text{ and } v_e(i-1, j-1, t)) \text{ or } (w(i-1, j, t) \text{ and } v_e(i-1, j, t)) \text{ or } (w(i-1, j+1, t) \text{ and } v_e(i-1, j+1, t)) \text{ or } (w(i, j+1, t) \text{ and } v_e(i-1, j+1, t)) \text{ or } (w(i+1, j+1, t) \text{ and } v_e(i+1, j+1, t)) \text{ or } (w(i+1, j, t) \text{ and } v_e(i+1, j, t)) \text{ or } (w(i+1, j-1, t) \text{ and } v_e(i+1, j-1, t)) \text{ or } (w(i, j-1, t) \text{ and } v_e(i, j-1, t))) \quad (9)$$

The weights of the network are determined through the comparison of the input attribute intensities $I_a(i, j, t)$ of the neuron with the respective inputs of the neighboring neurons $I_a(k, l, t)$. In the cases where the differences between such inputs are, in module, below a predetermined threshold L_m , the respective weights will be unitary. If the inputs $I_a(i, j, t)$ are scalar, the determination of the weights can be carried out by means of operations of simple implementation, as exemplified in (10).

$$w(i, j, k, l, t) = \begin{cases} 0 & \text{if } |I_a(i, j, t) - I_a(k, l, t)| > L_m \\ 1 & \text{if } |I_a(i, j, t) - I_a(k, l, t)| \leq L_m \end{cases} \quad (10)$$

The version of the oscillator neuron presented implies that the proposed network groups the input patterns based

Artigo publicado nos anais do 14th Symposium on Computer Architecture and High Performance Computing

on the similarity between the level or the intensity of the inputs of neighboring oscillators. However, different attributes of the input signal of the network can be used to carry through the segmentation, such as the average or the variance in the pixel neighborhood. On the other hand, $I_o(i,j,t)$ can also represent vectors of attributes related to the pixels of the image. In this case, the weights can be determined through the use of a measure of vectorial distance, such as the Euclidean distance, implying more complex structures for its implementation. In color images, for example, the Euclidean distance calculated in the space of the colors could be used to determine the weights, which would result in segmentation by color similarity. Generally, any attribute or set of attributes of the input signal can adequately be used to determine the weights of the network.

5. Implementation of the Network and Tests

To verify the functioning of the proposed network, two types of implementation were carried through. The first one is related to the simulation of the behavior of the network through an algorithm implemented in a PC computer using Delphi language. The second one is the use of the Max+plusII program from Altera [1] for simulation of the network and verification of the viability of its implementation in FPGA devices.

Fig. 5 presents an example of image segmentation carried out through the implementation of an algorithm that simulates the behavior of the proposed network in a PC computer. The original image, with 100x310 pixels, placed at the top left position, possesses 5 characters in the same gray level, isolated from each other, on a background with a different gray level. The proposed network found, from the intensity of each pixel of the original image (range from 0 to 255), the existence of 7 segments, which are shown in the other images. The first segment, presented at the top right position, corresponds to the background, having been the interior part of the character "R" not presented because it is not physically connected. This region appears as an isolated segment, which can be seen in the image at the bottom right position. The segments are sequentially presented in the output of the network. The characters were correctly segmented and isolated in time, facilitating the application of a procedure of character recognition. In the images related to the segments, the black tone represents oscillators in the silent phase while the white tone represents oscillators in the active phase.

Fig. 6 presents a network with 3x5 neurons ($N_n=15$) which was simulated with the use of Max+plusII. The network has one input, a clock signal, common to all the neurons, and, as outputs, the excitatory signals of the 15 neurons (S_1, S_2, \dots, S_{15}). The values of N_s , L_m and the inputs of the pixel intensities were considered as

parameters of the network in the stage of compilation. In a real application, however, such inputs must be externally accessible. In order to simplify the structure, each neuron is responsible for the determination of only 4 weights related to the excitatory connections. Thus, a neuron combines the excitations received from 4 of its neighboring neurons with the respective weights internally calculated and sends the excitation already combined with the respective weight to these neighboring neurons. Such artifice also makes it necessary to receive the intensities of only 4 neighboring neurons for calculation of the respective weights, diminishing the total number of connections of the network.

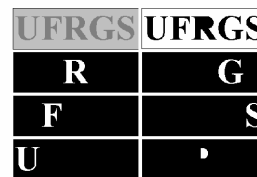


Figure 5. Segmentation using an algorithm simulating the proposed network.

The results of the simulations carried through with the Max+plusII demonstrate that the proposed network operates satisfactorily, grouping the similar clusters of pixels correctly and presenting sequentially each group in its output. To exemplify, fig. 7 represents an image with the respective values of intensity of each pixel (8 bits/pixel), which is used as the input of the network. Considering, for example $L_m=20$, the represented gray levels must form similar groups of oscillators. Fig. 8 presents the waveforms (logical levels 0 and 1) obtained through the simulation, where the grouping of the neurons with similar inputs and the time separation of each one of these groups can be verified. N_s must have a minimum value equal to the number of segments contained in the image. The use of higher values causes the appearance of time intervals without any segment in the output of the network. The result presented was obtained with $N_s=5$. As all the oscillators had started at state 0, the first segment is presented in the output after the fifth pulse of clock. From there, the four segments are sequentially presented in the output, with a clock period without any active output between each sequence. The output seg is the inhibitory output of the neuron with the lowest priority in the network and it can be used as an indicator of the existence of an active segment in the output of the network. This signal and the clock, can also be used to transfer the output values of the network to some external device.

In all the simulations it was possible to confirm the easiness of synchronism between oscillators and the necessity of a low and easily predictable number of iterations (pulses of clock) to get the desired result.

Artigo publicado nos anais do 14th Symposium on Computer Architecture and High Performance Computing

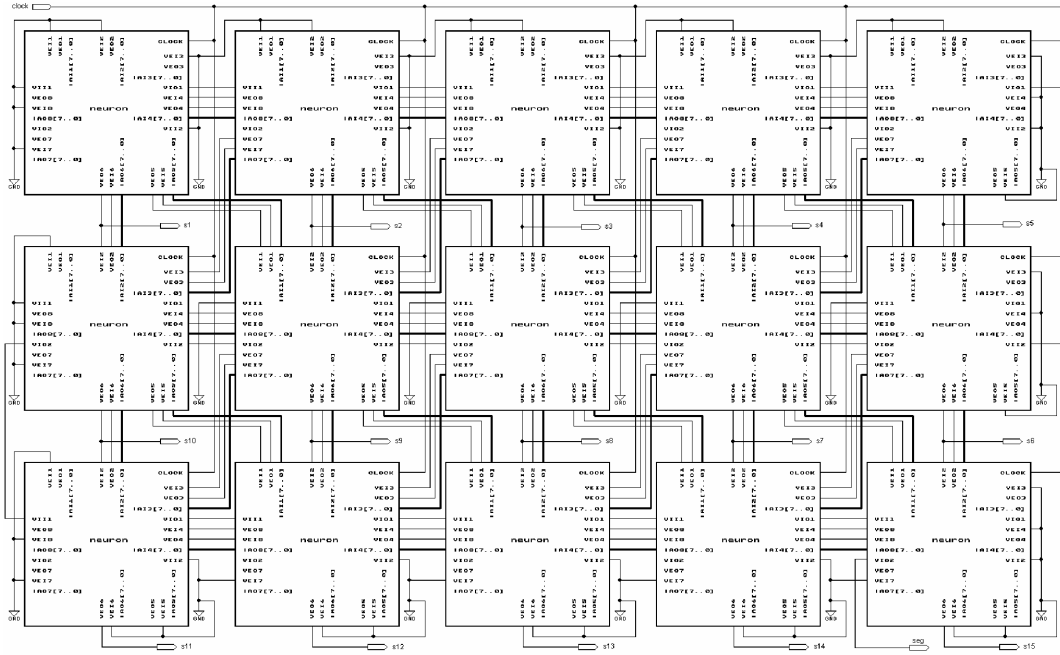


Figure 6. Implementation of the proposed network using the Max+plusII.

0	1	120	249	250
2	122	121	252	251
3	4	5	253	6

Figure 7. Input values used in the implementation with the Max+plusII.

6. Conclusion

The use of oscillator networks to simulate the capacity of image segmentation of the human brain is a recent proposal with satisfactory results. In the context of image segmentation, the LEGION network is the most consolidated model. Its parallel architecture and the capacity to separate the segments in time are strongly attractive. On the other hand, the structure based on differential equations implies high complexity for implementation using digital machines, parallel or not. Another disadvantage of this network concerns the limitation of the segmentation capacity of the number of objects simultaneously discriminated. The high number of parameters and their little intuitive setting procedures also

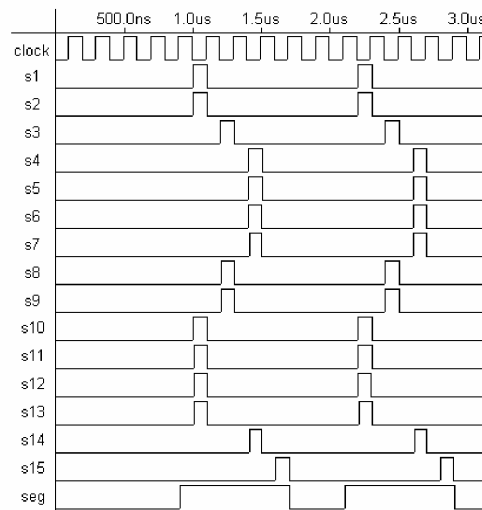


Figure 8. Result of the implementation using Max+plusII.

Artigo publicado nos anais do 14th Symposium on Computer Architecture and High Performance Computing

The architecture of the oscillatory neural network presented in this paper has the same advantages of the LEGION network related to the highly parallel topology and the capacity of segmenting and separating in time the segments of the input images. On the other hand, the proposed network does not present limitation related to the number of segments simultaneously discriminated and, because it uses simple binary logical operations, it offers lower computational complexity and the possibility of implementation in programmable logical devices. The use of only two parameters, L_m and N_s , of intuitive meaning and setting, also presents itself as an advantage of the proposed network. Other aspects that are not less positive concern the easiness of synchronism between oscillators and the necessity of a lower and easily predictable number of iterations in order to get the desired result, as opposed to the LEGION network. The addition of a random signal to the inputs of the network is not necessary, which also is a factor of complexity reduction of the proposed network.

The possibility of weights determination of the proposed network on the basis of diverse attributes of the input image can lead to more sophisticated procedures of segmentation, such as texture segmentation, for example. It is also pointed out that, being the weights of the network obtained directly from the attributes of the image to be segmented, the necessity of training of the network does not exist, in contrast to other architectures of neural networks.

The results obtained with the implementation of the proposed network in PC computers, through an algorithm that simulates its behavior, prove the consistency of the proposal and its several advantages in comparison with the LEGION network. The implementation of a network of small dimensions using the Max+plusII program from Altera also confirms the correct functioning of the proposal and its advantages.

On the basis of the results obtained, it is concluded that the new network architecture presented is an attractive alternative for applications involving image segmentation. Its several advantages imply greater viability and flexibility for practical implementation.

7. Bibliographical References

- [1] ALTERA, *MAX+PLUS II Software Overview*, <http://www.altera.com/products/software/maxplus2/mp2-index.html> (06/20/2002).
- [2] ANDO, H., MIYAKE, M., MORIE, T., NAGATA, M., IWATA, A. "A Nonlinear Oscillator Network Circuit for Image Segmentation with Double-threshold Phase Detection", *JCANN 99*, IEE, Edinburgh, UK, 1999, p. 655-660.
- [3] BUHMANN, J., von der MALSBERG, C. "Sensory Segmentation by Neural Oscillators", *JCNN 91*, IEEE, Seattle, EUA, 1991, pp. II603-II607.
- [4] COSP, J., MADRENAS, J. "A Neural Network for Scene Segmentation Based on Compact Astable Oscillators", *ICANN 99*, IEE, Edinburgh, UK, 1999, p. 690-695.
- [5] DAHMER, A. *Segmentação de Imagens Ecocardiográficas Utilizando Redes Neurais e Medidas de Textura*, Master Thesis, UFRGS-PPGC, Porto Alegre, 1998.
- [6] DAHMER, A., PICCOLI, L., SCHARCANSKI, J., NAVAU, P. O. A. "Fetal Echocardiographic Image Segmentation Using Neural Networks", *IPA'99 Image Processing and Applications Conference*, IEE, Manchester, UK, 1999.
- [7] FERNANDES, D., NAVAU, P. O. A., FICHTNER, P. F. P. "Segmentation of TEM Images Using Oscillatory Neural Networks", *SIBGRAPI 2001*, IEEE, Florianópolis, 2001, p.289-296.
- [8] FERNANDES, D., SIQUEIRA, M. L., NAVAU, P. O. A. "Segmentation of Fetal Echocardiographic Images Using Self-Organizing Maps and Oscillatory Neural Networks", *SIARP 2001*, IARP, Florianópolis, 2001, p.55-60.
- [9] FERNANDES, D. *Segmentação de Imagens Baseada em Redes Neurais Oscilatórias*, TI976, UFRGS-PPGC, Porto Alegre, 2001.
- [10] GHOSH, A., PAL, N. R., PAL, S. K. "Object background classification using Hopfield type neural network". *Int. J. Pattern Recognition Artificial Intelligence*, 1992, v.6, n.5, p.989-1008.
- [11] KUROKAWA, H., MORI, S. "A Local Connected Neural Oscillator Network for Sequential Character Segmentation", *ICNN 97*, IEEE, Texas, EUA, 1997, p.838-843.
- [12] LIU, X., CHEN, K., WANG, D. L. "Extraction of hydrographic regions from remote sensing images using an oscillator network with weight adaptation". *Transactions on GeoScience and Remote Sensing*, IEEE, 2001, 39(1), 207-211.
- [13] PICCOLI, L. *Segmentação e Classificação de Imagens Ecocardiográficas Utilizando Redes Neurais*. Master Thesis. UFRGS-PPGC, Porto Alegre, 1999.
- [14] SHAREEF, N., WANG, D. L., YAGEL, R. "Segmentation of Medical Images Using LEGION", *Transactions on Medical Imaging*, IEEE, 1999, v.18, n.1.
- [15] SIQUEIRA, M. L. *Obtenção de Medidas Cardíacas Fetais Através de Imagens Ecocardiográficas Segmentadas*, Master Thesis, UFRGS-PPGC, Porto Alegre, 2002.
- [16] TERMAN, D., WANG, D. L. "Global competition and local cooperation in a network of neural oscillators", *Physica D*, 1995, v.81(1-2), pp.148-176.
- [17] WANG, D., TERMAN, D. "Image Segmentation Based on Oscillatory Correlation". *Neural Computation*, MIT, 1997, v.9, p.805-836.
- [18] WANG, D. "Relaxation Oscillators and Networks". *Wiley Encyclopedia of Electrical and Electronics Engineering*, Wiley & Sons, 1999, v.18, p.396-405.
- [19] WANG, D., BROWN, G. J. "Separation of Speech from Interfering Sounds Based on Oscillatory Correlation". *Transactions on Neural Networks*, IEEE, 1999, v.10, n.3, p.684-697.

Rede Neural Maciçamente Paralela para Segmentação de Imagens

Dênis Fernandes¹
Philippe Olivier Alexandre Navaux²

Resumo

Redes neurais oscilatórias são uma abordagem recente para aplicações em segmentação de imagens. Como aspectos positivos, ressaltam-se a arquitetura maciçamente paralela e a capacidade de separar temporalmente os segmentos. Tais aspectos tornam estas redes adequadas à implementação de *chips* de visão, integrando-se uma rede de processadores com o sensor de imagens. Limitações práticas são encontradas nas estruturas propostas, tais como o uso de equações diferenciais, implicando elevada complexidade para implementação em *hardware* digital, e capacidade de segmentação limitada. Neste artigo, é apresentada uma proposta de rede neural oscilatória para segmentação de imagens com arquitetura maciçamente paralela. A rede é adequada à implementação em *hardware* digital, possuindo capacidade de segmentação não limitada e poucos parâmetros, com ajuste intuitivo. Simulações e a implementação de um pequeno protótipo em FPGA comprovam o correto funcionamento da rede proposta em aplicações de segmentação de imagens.

1 Introdução

A demanda por sistemas de visão artificial que implementam algoritmos complexos tem justificado o desenvolvimento de *chips* de visão ou retinas de silício. Nestes *chips*, os fotodetectores são integrados em conjunto com uma rede maciçamente paralela de elementos de processamento (EPs) para execução de operações específicas sobre a imagem (Figura 1).

Técnicas de processamento de imagens são fortes candidatas à aplicação de processamento paralelo. Operações onde o processamento pixel a pixel de uma imagem gera outra imagem de mesmas dimensões são adequadas a implementações em arquiteturas maciçamente paralelas, enquadrando-se na categoria SIMD (*Single Instruction stream Multiple Data stream*).

Arquiteturas maciçamente paralelas para processamento de imagens são normalmente implementadas para aplicações específicas (FARROHA, B. S.; DESHMUKH, R. G., 1995), evi-

¹denis@ee.pucrs.br

²navaux@inf.ufrgs.br

Rede Neural Maciçamente Paralela para Segmentação de Imagens

tando o uso de processadores de uso geral, mais complexos. Entretanto, em (KOMURO, T.; ISHII, I.; ISHIKAWA, M., 2002) é encontrada uma proposta de implementação de um *chip* de visão programável, com 64x64 EPs, utilizando tecnologia digital. A família AR LSI, da Mitsubishi, apresenta retinas artificiais de até 352x288 pixels em cores. Encontra-se em desenvolvimento uma versão de 640x480 pixels com capacidade de 100 fps (KYUMA, K. et al., 1999). Uma retina artificial programável é apresentada em (PAILLET, F.; MERCIER, D.; BERNARD, T. M., 1999), tendo sido implementado um protótipo com 128x128 pixels, com expectativa de obtenção de uma versão de 512x512 pixels. Já em (BARBARO, M. et al., 2002), é descrito um *chip* com 100x100 EPs e processamento analógico, com a finalidade de extrair o gradiente da imagem. Implementações analógicas requerem circuitos mais simples do que as digitais, apresentando maior dificuldade quanto à reprogramação.

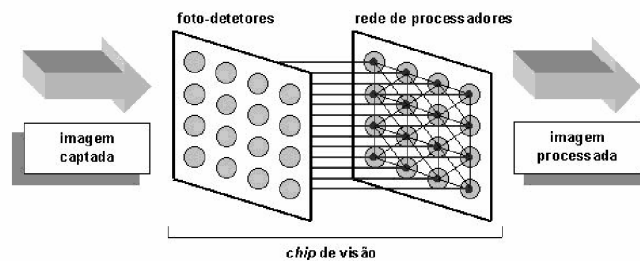


Figura 1: Integração dos sensores e da rede EPs em um *chip* de visão ou retina artificial.

Recentemente, redes neurais oscilatórias têm sido aplicadas em segmentação de imagens com resultados favoráveis (BUHMANN, J. VON DER MALSBURGH, C., 1991; COSP, J.; MADRENAS, J., 1999; KUROKAWA, H.; MORI, S., 1997; TERMAN, D.; WANG, D. L., 1995). A rede LEGION (*Locally Excitatory Globally Inhibitory Oscillator Network*) é a proposta mais consistente de rede neural oscilatória para segmentação de imagens (FERNANDES, D.; NAVAU, P.; FICHTNER, P., 2001; FERNANDES, D.; SIQUEIRA, M.; NAVAU, P., 2001). Esta rede pode separar temporalmente os segmentos, facilitando a identificação e quantificação. Sua natureza maciçamente paralela também torna a mesma adequada à implementação de *chips* de visão para segmentação de imagens. Já a elevada complexidade para implementação em *hardware* digital, resultado da estrutura baseada em equações diferenciais, a capacidade limitada para discriminação de segmentos e a quantidade de parâmetros com ajuste pouco intuitivo são aspectos limitantes.

O presente trabalho propõe uma rede neural oscilatória para segmentação de imagens com possibilidade de implementação em um *chip* de visão digital. A rede apresenta arquitetura maciçamente paralela e discrimina temporalmente os segmentos. A baixa complexidade e a não limitação quanto ao número de segmentos são atrativos da proposta. A utilização de poucos parâmetros, com ajuste intuitivo, o reduzido número de iterações e a fácil previsibilidade do tempo de processamento são também vantagens. Resultados obtidos, incluindo um

pequeno protótipo em FPGA, comprovam o correto funcionamento da proposta.

2 Arquitetura Proposta para a Rede Neural Oscilatória Maciçamente Paralela

Na rede proposta, cada pixel da imagem de entrada possui um EP correspondente. São utilizadas duas estruturas de conexões: conexões excitadoras e conexões inibitórias. A Figura 2 apresenta um exemplo da estrutura de conexões excitadoras. A saída excitadora de um EP será ativada quando pelo menos um de seus 8 vizinhos, submetido a uma entrada semelhante, estiver ativo. A Figura 3 apresenta um exemplo da estrutura de conexões inibitórias. Cada EP tem sua saída inibitória conectada a um único vizinho. No caso de existirem vários EPs vinculados a segmentos diferentes capacitados a passar para a fase ativa, somente o de maior prioridade o fará. As conexões excitadoras farão com que o único EP habilitado pelas conexões inibitórias habilite todos os demais EPs do respectivo segmento.

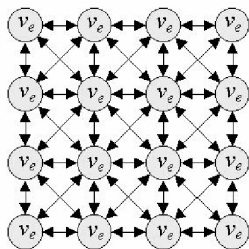


Figura 2: Conexões excitadoras.

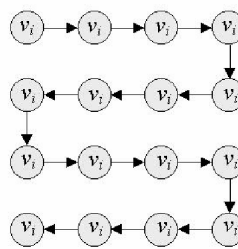


Figura 3: Conexões inibitórias.

3 Estrutura do Elemento de Processamento da Rede

O desenvolvimento do EP da rede proposta foi realizado visando a implementação em *hardware* digital. A idéia básica consiste em associar um contador binário a cada EP, fazendo com que EPs vizinhos com entradas semelhantes sejam sincronizados em um mesmo estado, diferente dos estados dos demais grupos de EPs. A Figura 4 apresenta a estrutura interna de um EP da rede. As definições das constantes e variáveis utilizadas são relacionadas abaixo:

- N_n : número total de EPs na rede;
- N_s : número máximo de segmentos a serem discriminados pela rede;
- L_m : limiar para determinação da semelhança entre as entradas de EPs vizinhos;
- $w(i,j,k,l,t)$: peso referente à comparação da entrada do EP situado na linha i e coluna j com a entrada do EP situado na linha k e coluna l , no tempo t ;
- $I_a(i,j,t)$: sinal de entrada do EP situado na linha i e coluna j no tempo (iteração) t , representando atributos vinculados ao pixel;

Rede Neural Maciçamente Paralela para Segmentação de Imagens

- $v_c(i, j, t)$: estado do contador do EP situado na linha i e coluna j no tempo t ;
- $v_o(i, j, t)$: sinal de saída do comparador do EP situado na linha i e coluna j no tempo t ;
- $v_e(i, j, t)$: sinal da saída excitadora do EP situado na linha i e coluna j no tempo t ;
- $v_i(i, j, t)$: sinal da saída inibitória do EP situado na linha i e coluna j no tempo t ;
- $v_i(i', j', t)$: sinal da entrada inibitória do EP situado na linha i e coluna j no tempo t e
- $v_l(i, j, t)$: sinal para indicação de líder do EP situado na linha i e coluna j no tempo t .

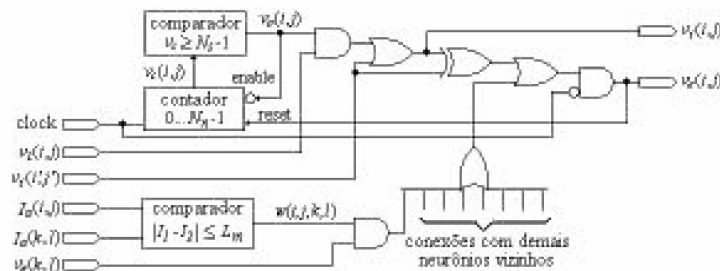


Figura 4: Estrutura interna do EP da rede proposta.

O sinal $v_l(i, j, t)$ pode ser gerado internamente ou externamente e somente EPs com $v_l(i, j, t)=1$ serão inicialmente habilitados a passar para a fase ativa (EPs líderes). EPs com $v_l(i, j, t)=0$ poderão ser ativados somente através de um EP vizinho com entrada semelhante que esteja na fase ativa. Na geração externa do sinal, os EPs líderes podem ser estabelecidos por um critério como a posição na imagem, por exemplo. Na geração interna do sinal pode ser utilizado o critério de que um EP líder deve possuir todos os pesos referentes às conexões excitadoras unitárias, o que corresponde a um pixel no centro de uma região homogênea.

Os pesos da rede são determinados através da comparação das intensidades dos atributos de entrada do EP com as entradas dos EPs vizinhos. Sendo a diferença das entradas menor do que um limiar (L_m), os respectivos pesos serão unitários. Se $I_a(i, j, t)$ é um vetor de atributos relacionado com os pixels da imagem, os pesos podem ser determinados através da utilização de uma medida de distância vetorial, implicando estruturas mais elaboradas.

4 Implementação da Rede Proposta e Testes

Foram realizados dois tipos de implementação. A primeira diz respeito à simulação através de um algoritmo em computadores PC. A segunda refere-se à utilização do programa Max+plusII da Altera para simulação e verificação da viabilidade de realização da rede em FPGAs.

A Figura 5 apresenta um exemplo de segmentação realizado com a implementação do algoritmo que simula a rede proposta. A imagem original, com 100x310 pixels, encontra-se

na posição superior esquerda. A rede encontrou, a partir da intensidade de cada pixel, 7 segmentos, os quais são representados pelas demais imagens. O primeiro segmento, na posição superior direita, corresponde ao fundo, tendo sido a parte interior da letra "R" desconsiderada por não estar conectada ao mesmo. Os segmentos são apresentados seqüencialmente na saída da rede, facilitando a aplicação de um procedimento de reconhecimento de caracteres. Nas imagens referentes aos segmentos, branco representa os EPs na fase ativa.

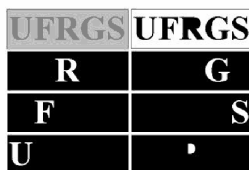


Figura 5: Segmentação utilizando um algoritmo simulando a operação da rede proposta.

Uma rede de 3x5 EPs ($N_n = 15$) foi simulada com a utilização do programa Max+plusII e implementada em um *chip* FPGA. Os resultados das simulações e a implementação demonstraram que a rede proposta funciona da forma desejada, agrupando corretamente pixels semelhantes e apresentando seqüencialmente cada grupo em sua saída.

Considerando um atraso de 1ns por porta lógica, estima-se que uma rede com N_n EPs possa segmentar N_s regiões em $4N_n(2N_s - 1)$ ns. Assim, uma rede de 32x32 EPs, com $N_s = 10$, poderia segmentar até 12850 fps. Carecendo de maiores detalhes, em (KYUMA, K. et al., 1999) é mencionada uma retina artificial para reconhecimento de imagens de 32x32 pixels com capacidade de processar 500 fps.

5 Conclusão

A arquitetura de rede neural oscilatória proposta possui topologia maciçamente paralela e separa temporalmente os segmentos da imagem. A rede não apresenta limitação quanto ao número de segmentos e é adequada para implementação de *chips* de visão digitais. A utilização de somente dois parâmetros, de ajuste intuitivo, também é uma vantagem atraente. Outros aspectos relevantes são a facilidade de sincronismo e a necessidade de poucas iterações e em número facilmente previsível para obter o resultado. A não necessidade de inicialização e de adição de um sinal aleatório na entradas, como em outros modelos, também se apresentam como fatores de redução da complexidade.

A possibilidade de determinação dos pesos com base em atributos diversos da imagem pode conduzir a procedimentos de segmentação mais sofisticados. Ressalta-se também que, sendo os pesos da rede obtidos diretamente dos atributos da imagem a ser segmentada, não existe necessidade de treinamento da rede, ao contrário de outras redes neurais.

Com base nos resultados obtidos, conclui-se que a rede proposta apresenta-se como uma

Referências Bibliográficas

alternativa atrativa para segmentação de imagens em *hardware* digital maciçamente paralelo. Suas diversas vantagens implicam viabilidade para implementação prática de *chips* de visão em tecnologia digital, possibilitando elevada velocidade no processamento de imagens.

Referências Bibliográficas

BARBARO, M. et al. *A 100x100 Pixel Silicon Retina for Gradient Extraction With Steering Filter Capabilities and Temporal Output Coding*. In: *IEEE Journal of Solid-State Circuits*, 2002, v.37, n.2, p.160-172.

BUHMANN, J., VON DER MALSBERG, C. *Sensory Segmentation by Neural Oscillators*. In: *IJCNN 91*, IEEE, Seattle, EUA, 1991, p. II603-II607.

COSP, J., MADRENAS, J. *A Neural Network for Scene Segmentation Based on Compact Astable Oscillators*. In: *ICANN 99*, IEE, Edinburgh, UK, 1999, p. 690-695.

FARROHA, B. S., DESHMUKH, R. G. *A Novel Approach to Design a Massively Parallel Application Specific Architecture for Image Recognition Systems*. In: *Southeastcon '95 Visualize the Future*, IEEE, Raleigh, EUA, 1995, p. 293-299.

FERNANDES, D., NAVAUX, P., FICHTNER, P. *Segmentation of TEM Images Using Oscillatory Neural Networks*. In: *SIBGRAPI 2001*, IEEE, Florianópolis, 2001, p.289-296.

FERNANDES, D., SIQUEIRA, M., NAVAUX, P. *Segmentation of Fetal Echocardiographic Images Using Self-Organizing Maps and Oscillatory Neural Networks*. In: *SIARP 2001*, IARP, Florianópolis, 2001, p.55-60.

KOMURO, T.; ISHII, I.; ISHIKAWA, M. *A New Architecture of Programmable Digital Vision Chip*. In: *Symposium on VLSI Circuits Digest of Technical Papers*, IEEE, 2002, p.266-269.

KUROKAWA, H.; MORI, S. *A Local Connected Neural Oscillator Network for Sequential Character Segmentation*. In: *ICNN 97*, IEEE, Texas, EUA, 1997, p.838-843.

KYUMA, K. et al. *Concept, Development, Mass Production, and Applications of Artificial Retina Chips*. In: *2nd International Conference on Intelligent Processing and Manufacturing of Materials*, IEEE, Honolulu, EUA, 1999, p.1297-1303.

PAILLET, F.; MERCIER, D.; BERNARD, T. M. *Second Generation Programmable Artificial Retina*. In: *20th Annual ASIC/SOC Conference*, IEEE, Washington, EUA, 1999, p.304-309.

TERMAN, D.; WANG, D. L. *Global competition and local cooperation in a network of neural oscillators*. In: *Physica D*, 1995, v.81(1-2), pp.148-176.

An Oscillatory Neural Network for Image Segmentation

Dênis Fernandes¹ and Philippe Olivier Alexandre Navaux²

¹PUCRS - Faculdade de Engenharia - Pontifícia Universidade Católica do Rio Grande do Sul, Av. Ipiranga, 6681, 90619-900 Porto Alegre, RS, Brazil
denis@ee.pucrs.br

²UFRGS - Instituto de Informática - Universidade Federal do Rio Grande do Sul, Av. Bento Gonçalves, 9500, 91501-970 Porto Alegre, RS, Brazil
navaux@inf.ufrgs.br

Abstract. Oscillatory neural networks are a recent approach for applications in image segmentation. Two positive aspects of such networks are its massively parallel topology and the capacity to separate the segments in time. On the other hand, limitations that restrict the practical application are found in the proposed oscillatory networks, such as the use of differential equations, implying high complexity for implementation in digital hardware, and limited capacity of segmentation. In the present paper, an oscillatory neural network suitable for image segmentation in digital vision chips is presented. This network offers several advantages, including unlimited capacity of segmentation. Preliminary results confirm the successful operation of the proposal in image segmentation and its good potential for real time video segmentation.

1 Introduction

The increasing demand for artificial vision systems, which implement complex algorithms with high speed, justifies the development of vision chips [10] or silicon retinas [3]. In these chips, the photo detectors corresponding to the pixels of the image are jointly integrated with a massively parallel network of processing elements (PEs) for execution of specific operations over the input image. Analog implementation of vision chips requires simpler circuits than the digital one, presenting, on the other hand, lower flexibility to reprogram the executed function [10]. So, when flexibility is required, the digital implementation is more attractive.

Recently, alternative topologies of artificial neural networks, the oscillatory neural networks, which are inspired on the mechanism of segmentation executed by the human brain, have been applied in image segmentation with satisfactory results [4][6][11][14]. The study of these networks is a fertile field of work, as well as the development of dedicated architectures for efficient implementation [2][6].

The LEGION network (Locally Excitatory Globally Inhibitory Oscillator Network) [15] is the most consistent proposal of oscillatory neural network for image segmentation found in the bibliography. Its applications include segmentation of remote sensing images [12], medical images [8][13], and electronic microscope images [7]. An interesting aspect of this network is the capability to separate the segments of the image in time, facilitating later identification and quantification. Its massively parallel

nature is adequate to the implementation of vision chips for image segmentation. The negative aspects include high computational complexity for implementation in digital hardware as a consequence of the use of differential equations, limited segmentation capacity, and also the high amount of parameters and their little intuitive setting.

This paper introduces a new model of oscillatory neural network inspired on LEGION network, suitable to applications of image segmentation and implementation in vision chips with digital technology. The network presents massively parallel topology and it is able to separate the segments in time. The low complexity and the non-limitation regarding the number of segments to be discriminated are advantages of the network. The use of few parameters with intuitive setting is also a positive aspect, as well as the reduced number of iterations and the easy predictability of the time to reach the results. The digital structure also presents flexibility for easy implementation of more sophisticated segmentation procedures, using different attributes of the image. Results found in practical implementations are presented, proving that the operation of the proposed network was performed according to the expectations, and showing its potential efficiency to real time video segmentation.

2 The Proposed Oscillatory Neural Network

In the late 80's, oscillations of approximately 40 Hz were discovered in the visual cortex of the human brain. Such oscillations have strong correlation with the visual stimulus and synchronism of phase occurs between physically near neurons that receive similar stimulus, which can characterize a homogeneous region of the perceived image. Physically near neurons that receive different stimulus or physically distant neurons do not present such synchronism of phase [4][16].

A new oscillatory neural network was conceived [9] using the property of local synchronism between neurons and adding a mechanism of global inhibition implemented with local connections to get anti-synchronism among different groups of neurons. The proposed network is suitable to application in image segmentation and implementation in digital hardware with massively parallel topology.

2.1 Structures of Connections

The proposed network is implemented in a two-dimensional topology with the same size of the image to be segmented. Two structures of connections among neurons, called excitatory connections and inhibitory connections, are used.

Figure 1(a) presents an example of the excitatory connection structure. A neuron has its excitatory output v_e simultaneously connected to the 8 nearest neurons. So, the excitatory output of a neuron will be activated when at least one of its nearest neighbors, with similar input, is active.

Figure 1(b) presents an example of the inhibitory connections structure adopted. Each neuron has its inhibitory output v_{io} connected to only one neighboring neuron. This structure establishes a priority order and neurons with higher priority inhibit the remainders. The excitatory connections will cause the only neuron qualified by the inhibitory connections to excite the other neurons belonging to same segment. The

inhibitory output of the neuron with the lowest priority will be active only when at least one neuron is active. In a practical implementation, this signal can be used to detect images without any segment present at the output of the network.

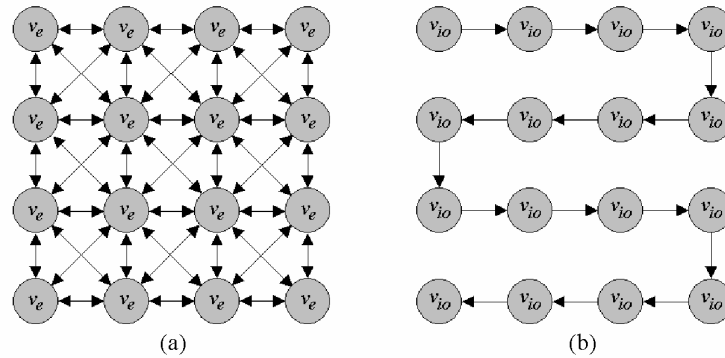


Fig. 1. Examples of structures of excitatory connections (a) and inhibitory connections (b)

2.2 The Network Neuron

The basic idea for the conception of the proposed neuron consists of associating a binary counter to each of them in a way that neighboring neurons with similar inputs are synchronized in the same state, different from the states of the other groups (Figure 2). The neurons stay inactive, until its counters reach a predefined state.

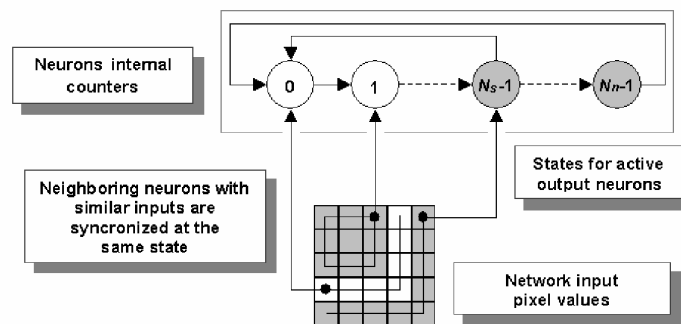


Fig. 2. Basic idea behind the conception of the proposed neuron

Figure 3 presents the internal logic structure of the proposed neuron. The definitions of the constants and variables are listed below:

- N_n : total number of neurons in the network;
- N_s : maximum number of segments allowed to be separated by the network;
- L_w : threshold for weight determination;

- $v_c(i,j,t)$: state of the internal binary counter belonging to the neuron placed at line i and column j in time (iteration) t ;
- $v_o(i,j,t)$: output signal of the neuron internal comparator;
- $v_e(i,j,t)$: excitatory output of the neuron;
- $v_{io}(i,j,t)$: inhibitory output of the neuron;
- $v_{ii}(i,j,t)$: inhibitory input of the neuron;
- $v_l(i,j,t)$: leader indication signal;
- $I_a(i,j,t)$: input signal of the neuron representing a set of features of the related pixel;
- $w(i,j,k,l,t)$: weight related to the comparison between the input of the neuron placed at line i and column j and the input of the neuron placed at line k and column l .

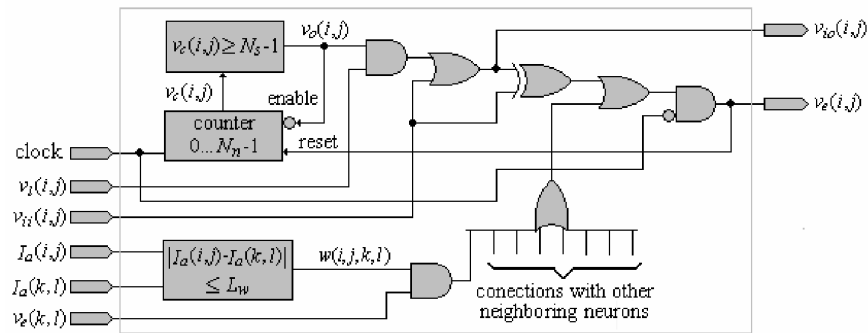


Fig. 3. Neuron internal structure of the proposed network

A binary counter with N_n states defines the neuron state as a function of the time. The counter receives a synchronous reset if the neuron is active ($v_e(i,j,t)=1$), otherwise its state is increased until N_s-1 , remaining there until it is reset. The use of N_n states makes the discrimination of a maximum of N_n segments possible, even though, in this case, there is no similarity between any inputs of neighboring neurons. The proposed neuron structure qualifies the network to separate the N_s first segments according to the sequence established by the structure of inhibitory connections.

The leader signal can be generated internally or externally and only neurons with $v_l(i,j,t)=1$ will be initially qualified to pass to the active phase (leaders). Neurons with $v_l(i,j,t)=0$ can be activated through an active neighbor with similar input. In the external generation, some criteria such as the position in the image (for example) can be established. It could be established that neurons in the central region of the image would be the only ones qualified to pass to the active phase, causing the appearance of only that segment at the output of the network. In the internal generation of the signal, one can use, for example, the criteria that a leader must have all the excitatory weights active, which correspond to a pixel in the center of a homogeneous region.

The weights of the network are determined through the comparison of each neuron input attribute intensities with the respective inputs of its neighboring neurons (1). In the cases where the differences between such inputs are, in module, below a threshold L_w , the respective weights will be unitary. Different attributes of the image can be used to carry through the segmentation. On the other hand, $I_a(i,j,t)$ can represent vec-

tors of attributes related to the pixels. In this case, the weights can be determined using a measure of vectorial distance, implying more complex structures. In color images, the Euclidean distance could be used to determine the weights, which would result in segmentation by color similarity.

$$w(i, j, k, l, t) = \begin{cases} 0 & \text{if } |I_a(i, j, t) - I_a(k, l, t)| > L_w \\ 1 & \text{if } |I_a(i, j, t) - I_a(k, l, t)| \leq L_w \end{cases} \quad (1)$$

3 Implementation and Results

To verify the functioning of the proposed network, two types of implementation were carried out. The first one is related to the simulation of the behavior of the network through an algorithm implemented in a PC computer. The second one is the use of the Max+plusII program from Altera [1] for simulation of the network and verification of the viability of its implementation in digital devices [9].

3.1 Segmentation of an Artificial Image

Figure 4 presents an example carried out through the algorithm that simulates the proposed network. The input image, with 100x310 pixels, is placed at the top left position. The additional 7 images are the non-null segments got from the output of a network with 100x310 neurons. In these 7 images, each pixel is a neuron output and the white represents the active neurons. The weights were calculated on the basis of the neighboring pixel intensity differences with $L_w=0$. The neurons with all the excitatory weights unitary were considered leaders. The first segment is the background, having the interior parts of the characters “P” and “A” not presented, as they are not physically connected. All the characters were correctly isolated in time, facilitating the application of a character recognition procedure.



Fig. 4. Segmentation using the algorithm that simulates the proposed network operation

3.2 A Practical Application

Figure 5 presents a gray level image with 1482x2060 pixels and 8 bits, which was obtained through transmission electron microscopy (TEM). This image represents a silicon sample where ions of helium were implanted with the objective to reduce defects in the crystalline structure. The helium accumulates itself in bubbles concentrated in some areas of the silicon. For evaluation of the process, it is necessary to

determine the gas volume in the sample. Such volume can be found through the counting of bubbles and estimation of their areas. This procedure must be carried through for several images, being a complex task for human manual execution [7].

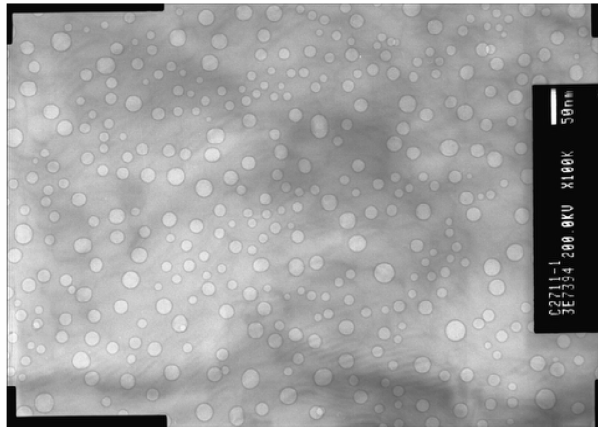


Fig. 5. Image of TEM representing a silicon sample with helium bubbles

The proposed network was used to segment the helium bubbles and simultaneously separate them in time, facilitating the implementation of an automatic process of counting and area measuring. Figure 6(a) presents a region extracted from Figure 5, which is used to verify the qualitative results of the proposed network.

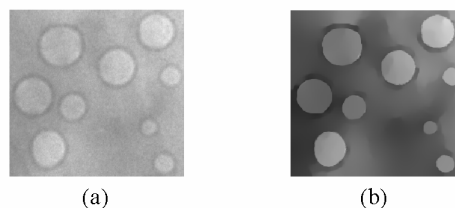


Fig. 6. Region extracted from Figure 5 (a) and filtered with the FPS algorithm (b)

As other segmentation procedures, the proposed network is sensitive to noise in the input image. Such problem makes the use of a smoothing filter necessary. Figure 6(b) presents the result of filtering the Figure 6(a) image using the FPS (Feature Preserving Smoothing) algorithm [5]. The noise reduction and the preservation of the contours of the helium bubbles can be easily observed.

Figure 7 presents the non-null segments got from the output of a network with 1482x2060 neurons. Besides the 9 bubbles, a segment representing the silicon background was also supplied (bottom right position), which can be easily detected and discarded in the automatic procedure of measuring. The network weights were calculated on the basis of the neighboring pixel intensity differences with $L_w=3$.

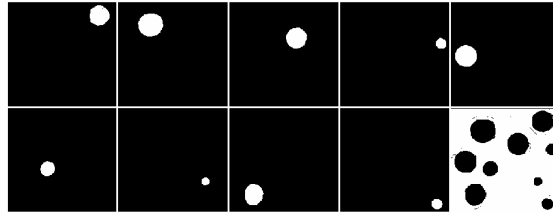


Fig. 7. Segments of Figure 6(b) image obtained with the proposed neural network

The bubble areas can be obtained computing the ratio between the number of active neurons and the total number of neurons. Knowing the relation in pixel/mm (Figure 5), the helium volume can be estimated. Visual area determination can be inexact if the bubbles are not perfectly circular. Such limitation is not observed in the proposed method. Incomplete, degraded and superimposed bubbles can lead to wrong measures. For these situations, the manual process can be used.

4 Conclusion

The use of oscillator networks to simulate the capacity of image segmentation of the human brain is a recent proposal with good results. In this context, the LEGION network is the most consolidated model. Its massively parallel topology and the capacity to separate the segments in time are highly attractive. On the other hand, the structure based on differential equations has high complexity for digital machines. Another disadvantage is the limitation to segment a high number of objects simultaneously. The high number of parameters and their little intuitive setting are also drawbacks.

The oscillatory neural network presented by this paper has massively parallel topology and capacity to separate the homogeneous regions of the input image in time. The network does not present limitation related to the number of segments, has lower complexity, is suitable for implementation in digital vision chips, and uses few parameters, with easy setting. Other positive aspects are the easiness of synchronism and the necessity of a low and easily predictable number of iterations to get the result. The addition of a random signal to the inputs, like in the LEGION, is not necessary, which is also a factor of complexity reduction. Since the weights are obtained directly from the attributes of the input image, there is no need of training, in contrast with other neural networks. Finally, the weight determination on the basis of diverse attributes of the input image, can lead to more sophisticated segmentation procedures.

The good results obtained with the proposed network through an algorithm that simulates its behavior prove the consistency and its several positive characteristics. The simulation of a small size network and its implementation in an FPGA chip confirms the correct functioning of the proposal. Also, preliminary statistical analysis have showed that the proposed network can segment images with additive Gaussian noise and 20dB of signal-to-noise ratio with less than 0.02% of misclassified pixels.

As there are no references to implementations of LEGION network using massively parallel digital hardware, the evaluation of the complexity gain of the proposed

network is not possible. On the other hand, computational simulations show that the proposed network is much faster than the LEGION network version based on differential equations. Specific studies for complexity reduction are under development in order to implement practical networks using commercial FPGA devices.

Based on the results, it is concluded that the proposed oscillatory neural network is attractive for applications of image segmentation with implementation in massively parallel digital hardware. Its several advantages imply viability for implementation of vision chips using digital technology, allowing high speed in image segmentation.

References

1. ALTERA, MAX+PLUS II Software Overview, <http://www.altera.com/products/software/maxplus2/mp2-index.html> (20/06/2002)
2. ANDO, H., MIYAKE, M., MORIE, T., NAGATA, M., IWATA, A. A Nonlinear Oscillator Network Circuit for Image Segmentation with Double-threshold Phase Detection, ICANN 99, IEE, Edinburgh, UK, 1999, p. 655-660
3. BARBARO, M., BURGI, P., MORTARA, A., NUSSBAUM, P., HEITGER, F. A 100x100 Pixel Silicon Retina for Gradient Extraction With Steering Filter Capabilities and Temporal Output Coding, IEEE Journal of Solid-State Circuits, 2002, v.37, n.2, p.160-172
4. BUHMANN, J., von der MALSBURG, C. Sensory Segmentation by Neural Oscillators, IJCNN 91, IEEE, Seattle, USA, 1991, p. II603-II607
5. CHEN, Ke, WANG, DeLiang. Image Segmentation Based on a Dynamically Coupled Neural Oscillator Network, IJCNN 99, IEEE, Washington D. C., USA, 1999
6. COSP, J., MADRENAS, J. A Neural Network for Scene Segmentation Based on Compact Astable Oscillators, ICANN 99, IEE, Edinburgh, UK, 1999, p. 690-695
7. FERNANDES, D., NAVAU, P. O. A., FICHTNER, P. F. P. Segmentation of TEM Images Using Oscillatory Neural Networks, SIBGRAPI 2001, IEEE, Florianópolis, Brazil, 2001, p.289-296
8. FERNANDES, D., SIQUEIRA, M. L., NAVAU, P. O. A. Segmentation of Fetal Echocardiographic Images Using Self-Organizing Maps and Oscillatory Neural Networks, SIARP 2001, IARP, Florianópolis, Brazil, 2001, p.55-60
9. FERNANDES, D., STEDILE, J. P., NAVAU, P. O. A. Architecture of Oscillatory Neural Network for Image Segmentation. SBAC-PAD 2002, IEEE, Vitória, Brazil, 2002
10. KOMURO, T., ISHII, I., ISHIKAWA, M. Digital Vision Chips and High-Speed Vision Systems, Symposium on VLSI Circuits Digest of Technical Papers, IEEE, Kyoto, Japan, 2001, p.1-4
11. KUROKAWA, H., MORI, S. A Local Connected Neural Oscillator Network for Sequential Character Segmentation, ICNN 97, IEEE, Texas, USA, 1997, p.838-843
12. LIU, X., CHEN, K., WANG, D. L. Extraction of hydrographic regions from remote sensing images using an oscillator network with weight adaptation, Transactions on GeoScience and Remote Sensing, IEEE, 2001, 39(1), 207-211
13. SHAREEF, N., WANG, D. L., YAGEL, R. Segmentation of Medical Images Using LEGION, Transactions on Medical Imaging, IEEE, 1999, v.18, n.1
14. TERMAN, D., WANG, D. L. Global competition and local cooperation in a network of neural oscillators, Physica D, 1995, v.81(1-2), pp.148-176
15. WANG, D., TERMAN, D. Image Segmentation Based on Oscillatory Correlation, Neural Computation, MIT, 1997, v.9, p.805-836
16. WANG, D. Relaxation Oscillators and Networks. Wiley Encyclopedia of Electrical and Electronics Engineering, Wiley & Sons, 1999, v.18, p.396-405