

MINISTÉRIO DA CIÊNCIA E TECNOLOGIA
INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS

**DETECÇÃO DE MOVIMENTO DE OBJETOS
NÃO RÍGIDOS:
UMA ABORDAGEM BASEADA EM REDES NEURAIS**

Lúbia Vinhas

Dissertação final do curso de Mestrado em Computação Aplicada, orientada pelo
Dr. Paulo Ouverá Simoni, em cumprimento total dos requisitos exigidos para a
obtenção do título de Mestre.

INPE
São José dos Campos
Junho de 1994


Aprovada pela Banca Examinadora
em cumprimento a requisito exigido
para a obtenção do Título de Mestre
em Computação Aplicada

Dra. Maria Suelena Santiago Barros




Presidente

Dr. Paulo Ouverá Simoni



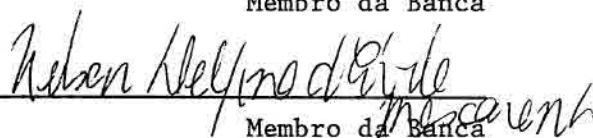
Orientador

Dr. Atair Rios Neto



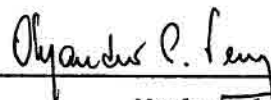
Membro da Banca

Dr. Nelson Delfino d'Ávila Mascarenhas



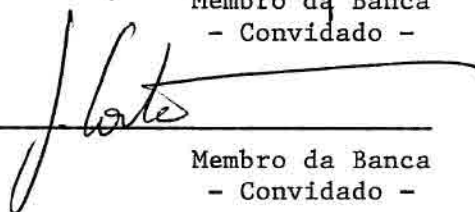
Membro da Banca

Dr. Alejandro Cesar Frery Orgambide



Membro da Banca
- Convidado -

Dr. Luciano da Fontoura Costa



Membro da Banca
- Convidado -

Candidata: Lúbia Vinhas

São José dos Campos, 11 de abril de 1994

AGRADECIMENTOS

Agradeço ao Dr. Paulo Simoni pela orientação e ajuda, e ao Dr. Alejandro Frery por suas valiosas opiniões.

Ao Machine Perception Laboratory, da University of California at Los Angeles, pelo permissão de uso do software *SFINX*.

Gostaria de deixar registrada a ajuda “não técnica” mas de fundamental importância de algumas pessoas. Os colegas da pós (principalmente da CAP e da sala 37) pelos momentos de amizade e companheirismo. À super Marisa pela presença sutil, porém constante e decisiva. Às minhas irmãs Mara e Heluiza, que sempre foram minha base afetiva em S. José. Ao amigo Alejandro por sua paciência. Finalmente, ao Rei, pelo incentivo, apoio e amor.

RESUMO

Nesse trabalho estuda-se o problema da detecção de movimento do objetos através da análise de sequências de imagens digitais. O estudo é aplicável, em particular, ao movimento de objetos não rígidos. Uma parte do problema consiste na detecção local de movimento em cada ponto de uma imagem. Para isso, propõe-se uma extensão de um modelo biológico de detecção para sinais unidimensionais, o Detector de Reichardt, para sinais bidimensionais. Na formalização do modelo proposto empregam-se conceitos da área de de processamento de imagens, enquanto que o modelo original é descrito segundo a teoria de sistemas. O detector proposto é sensível à direção. A seguir trata-se da compatibilização de medidas locais de modo a obter uma solução global mais consistente para a imagem inteira. Propõe-se, para isto, um modelo onde a compatibilização é tratada como um problema combinatório que pode ser resolvido com o uso de uma rede neural do tipo Hopfield.

ABSTRACT

In this work a study of the problem of detecting object motion is done for input given by image sequences, in particular for the case of non-rigid object motion. An initial part of the problem is the local detection of motion in a single image point. For this purpose an extension of the biological detection model for unidimensional signals, the Reichardt Detector, is proposed for bidimensional signals. The proposed detector is sensitive to direction. The problem of improving the compatibility of the local measurements is then treated as a search for global solution for the whole image. A model is proposed in which the search for compatibility is treated as a combinatory problem solved by means of a Hopfield neural network.

SUMÁRIO

	Pág.
LISTA DE FIGURAS	v
CAPÍTULO 1 INTRODUÇÃO	1
CAPÍTULO 2 DETECÇÃO DE MOVIMENTO	5
2.1 O processamento visual natural	5
2.2 Detectores de Movimento	8
2.2.1 Detectores baseados em correspondência de atributos	9
2.2.2 Detectores baseados em gradiente	10
2.2.3 Detectores biológicos	11
2.3 Modelos de Detecção	19
CAPÍTULO 3 REDES NEURAIIS	21
3.1 Componentes fisiológicos	22
3.2 Conceitos e Definições em Redes Neurais Artificiais	23
3.3 Aprendizagem	26
3.4 Paradigmas de Redes Neurais	27
3.5 Implementações	28
CAPÍTULO 4 UMA SOLUÇÃO PARA O PROBLEMA DA DETECÇÃO DE MOVIMENTO BASEADA EM REDES NEURAIIS	31
4.1 O problema tratado	31
4.2 Formulação do problema	33

4.3	Detectores locais	35
4.4	Compatibilização	38
4.4.1	Modelo $x - y$	42
4.4.2	Rede de Hopfield	45
CAPÍTULO 5 TESTES DO MODELO		51
5.1	Testes do modelo	51
5.1.1	Imagens sintéticas com objetos rígidos	51
5.1.2	Imagens reais	58
5.1.3	Compatibilização por Rede de Hopfield	67
5.1.4	Informação computacional	69
CAPÍTULO 6 CONCLUSÕES E EXTENSÕES		71
REFERÊNCIAS BIBLIOGRÁFICAS		75

LISTA DE FIGURAS

	Pág.	
2.1	Representação esquemática do olho humano.	6
2.2	Subáreas do córtex visual.	7
2.3	O detector de Reichardt	13
2.4	Exemplo da aplicação de um Detector de Reichardt a um par de imagens sintéticas.	14
2.5	Dois sinais periódicos.	15
3.1	Alguns tipos de neurônios.	22
3.2	Representação diagramática de um neurônio artificial.	25
4.1	As quatro direções propostas para a detecção de movimento	34
4.2	Compatibilizador de direções.	35
4.3	Exemplo de aplicação da Regra do Maior Valor	38
4.4	Exemplo de aplicação da Regra da Padronização	39
4.5	Exemplo de configuração de uma rede de Hopfield para o problema da compatibilização.	46
4.6	Exemplo de suavidade espacial.	49
4.7	Exemplo de falta de suavidade espacial.	49
5.1	Imagem $I_{r_0}(s, 1)$ com objeto rígido.	52
5.2	Imagem $I_{r_0}(s, 2)$ com objeto rígido e movimento arbitrário para a esquerda.	52
5.3	Imagem de direção X_0 da sequência I_{r_0}	53
5.4	Imagem de direção X_1 da sequência I_{r_0}	53
5.5	Imagem de direção X_2 da sequência I_{r_0}	54

5.6	Imagem de direção X_3 da sequência I_{r0}	54
5.7	Resultado da detecção com compatibilização por Regra do Máximo na sequência L_{r0}	55
5.8	Resultado da detecção com compatibilização por Rede de Hopfield na sequência L_{r0}	55
5.9	Imagem $I_{r1}(s, 1)$ com objeto rígido.	56
5.10	Imagem $I_{r1}(s, 2)$ com objeto rígido e movimento arbitrário de 45 graus para cima.	57
5.11	Resultado da detecção com compatibilização por Regra do Máximo na sequência L_{r1}	57
5.12	Resultado da detecção com compatibilização por Rede de Hopfield na sequência L_{r1}	58
5.13	(a) A imagem base, (b) a imagem base deslocada no sentido arbitrário para cima	59
5.14	(a) A imagem base, (b) a imagem base deslocada no sentido arbitrário para direita	60
5.15	(a) A imagem base, (b) a imagem real seguinte, com os deslocamentos naturais.	60
5.16	Detecção por regra do máximo, sem campo receptivo, na sequência L_1 , com deslocamentos arbitrários para cima.	61
5.17	Detecção por regra do máximo, sem campo receptivo, na sequência L_2 , com deslocamentos arbitrários para direita.	62
5.18	Detecção por regra do máximo, sem campos receptivos, na sequência L_3 , com deslocamentos naturais.	63
5.19	Detecção por regra do máximo, com campo receptivo, na sequência L_1 , com deslocamentos arbitrários para cima.	64
5.20	Detecção por regra do máximo, com campo receptivo, na sequência L_2 , com deslocamentos arbitrários para direita.	65
5.21	Detecção por regra do máximo, com campo receptivo, na sequência L_3 , com deslocamentos naturais.	66

5.22	Compatibilização por rede de Hopfield, na sequência L_1 , com deslocamentos arbitrários para direita.	67
5.23	Compatibilização por rede de Hopfield, na sequência L_2 , com deslocamentos arbitrários na direção vertical, no sentido para cima.	68
5.24	Compatibilização por rede de Hopfield, na sequência L_3 , com deslocamentos naturais	68

CAPÍTULO 1

INTRODUÇÃO

A disciplina Visão Por Computador preocupa-se com a análise de imagens no sentido de conseguir resultados similares àqueles conseguidos pelo sistema visual de um observador humano. Uma das linhas de pesquisa nessa área é a relação entre a estrutura neural do sistema visual humano e os fenômenos de percepção visual, através de observações sobre a experiência visual. Procura-se descobrir propriedades dos mecanismos neurais que possam ser úteis na compreensão do processamento visual. É de interesse que essas descobertas possam contribuir na simulação de fenômenos perceptíveis ao olho humano através de redes neurais artificiais (44).

O fenômeno perceptivo de interesse nesse trabalho é a percepção visual de movimento. É tarefa do sistema visual compreender o ambiente tridimensional a sua volta a partir das imagens bidimensionais das cenas que são projetadas na retina. O movimento de objetos e seres no ambiente é um importante substrato para essa compreensão. A partir daí é possível a recuperação da estrutura tridimensional dos objetos, sua identificação e segmentação, além de contribuir para que o observador possa se movimentar de forma segura entre os objetos que estão a sua volta. A detecção de movimento é um processo executado pelo sistema visual que examina um estímulo físico, as intensidades radiométricas da cena que são projetadas na retina, e controla julgamentos, percepções e ações relacionadas com o movimento dos objetos.

Vários pesquisadores nas áreas de psicologia, fisiologia e computação têm se preocupado em entender essa capacidade e transportá-la para sistemas de visão por computador (2).

Nesse trabalho será abordado o problema da detecção de movimento em uma sequência de imagens. O estudo compreende, mais especificamente,

a detecção em imagens cujos objetos que possam estar se movimentando são objetos não rígidos, sujeitos a deformações. Trata-se do fenômeno conhecido como **movimento aparente**. Apesar de cada imagem representar um instante de tempo, ou seja não existe o estímulo real de movimento, o sistema visual humano é capaz de fornecer ao observador a impressão convincente de movimento. É essa característica que se pretende obter em um sistema de detecção automática de movimento.

Para esse estudo foram consideradas imagens sintéticas de objetos rígidos e partes de imagens reais de satélites meteorológicos onde aparecem nuvens, as quais podem se movimentar em várias direções ao mesmo tempo e podem se deformar. Essas imagens servem para se estudar a detecção de movimento de objetos não rígidos.

Nesse trabalho pretende-se estudar o problema da detecção de movimento propondo uma abordagem baseada em redes neurais. Esse paradigma de computação inspira-se nos modelos funcionais de neurônios e redes de neurônios. Essa abordagem justifica-se visto que o problema está diretamente correlacionado com uma tarefa perceptiva executada pelo sistema de processamento visual na retina e no córtex visual.

A organização do trabalho é a seguinte: no Capítulo 2 faz-se uma revisão dos conceitos relevantes ao problema da percepção de movimento, abordando alguns modelos propostos na literatura para essa tarefa e os problemas que esse modelos podem apresentar. Propõe-se uma extensão generalizada para o caso de imagens, e uma formalização computacional do modelo de detecção de movimento local do tipo detector de Reichardt.

No Capítulo 3 faz-se uma revisão sobre o paradigma de redes neurais.

No Capítulo 4 mostra-se uma formalização e uma discussão a respeito desse problema e propõe-se um modelo de detecção de movimento onde a detecção local de movimento é feita por detectores de Reichardt sensíveis a direção (31) (citado em (40)) e pela versão elaborada desse detector (41). Propõe-se que as

respostas de detectores sensíveis a diferentes direções sejam combinadas para a obtenção de uma solução que represente o deslocamento dos objetos na imagem. Essa combinação é feita através de uma regra estritamente pontual, que leva em conta apenas os valores dos detectores no ponto. Outra abordagem proposta neste trabalho é tratar a compatibilização como um problema do tipo combinatório, com solução dada por uma rede do tipo Hopfield. Essa solução leva em conta não só os valores dos detectores nos pontos, mas também em uma vizinhança considerada.

No Capítulo 5 mostram-se alguns resultados da aplicação desse modelo a imagens sintéticas com corpos rígidos e deslocamentos arbitrários e a partes de imagens meteorológicas reais com deslocamentos arbitrários e deslocamentos naturais.

No Capítulo 6 apresentam-se conclusões e algumas possíveis extensões para esse trabalho.

CAPÍTULO 2

DETECÇÃO DE MOVIMENTO

2.1 O processamento visual natural

O processamento visual natural (nos animais e no homem) é estudado através de experimentos fisiológicos e de comportamento (44). Estruturas neurofisiológicas altamente complexas e organizadas estão presentes em todas as etapas do processamento visual. O funcionamento harmônico e as informações que elas fornecem fazem com que seja possível o relacionamento do indivíduo com o mundo a sua volta por meio da informação visual.

De uma forma simplificada o processamento visual num estágio primário consiste da captação, do processamento e da análise de informação na forma de radiação eletromagnética de um certo espectro proveniente do mundo exterior. A luz penetra o olho humano através de uma membrana transparente, a **córnea**. Atrás da córnea existe um músculo opaco em cujo centro há uma abertura chamada pupila. A pupila controla a quantidade de luz que é focalizada pela ação de um músculo retráctil, o cristalino, sobre a parte fotossensível do olho, a **retina**. A retina é a parte do olho responsável pela conversão da luz em impulsos elétricos (ver Figura 2.1). A retina é composta por um grande número de elementos individuais, sensíveis à luz, chamados de **receptores**. Os sinais dos receptores são transmitidos ao cérebro através do nervo óptico (22).

A informação visual, já convertida em impulsos elétricos, é processada prioritariamente na camada externa do cérebro, o **córtex**. Existem áreas especializadas em diferentes funções no córtex, por exemplo o córtex visual, onde acontece a maioria do processamento visual.

Análises detalhadas do córtex visual têm demonstrado que existem subáreas voltadas para o processamento de informações visuais específicas, como

por exemplo, informação relativa à cor, à profundidade, ao movimento, etc (46). As áreas específicas para a detecção de movimento são conhecidas nos primatas como a área V1, onde existem células com respostas sensíveis a movimento, e a área MT (*Middle Temporal*), onde ocorre a integração dessas respostas (23,30) (ver Figura 2.2).

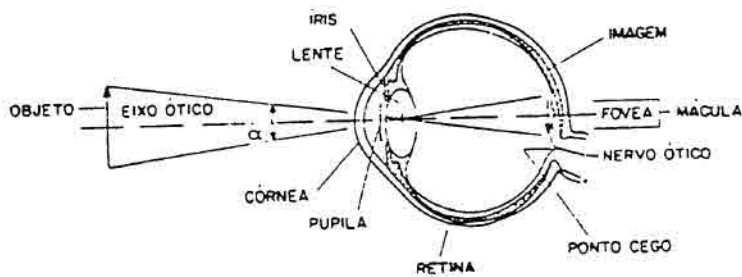


Fig. 2.1 - Representação esquemática do olho humano.

FONTE: (25), p. 1.17.

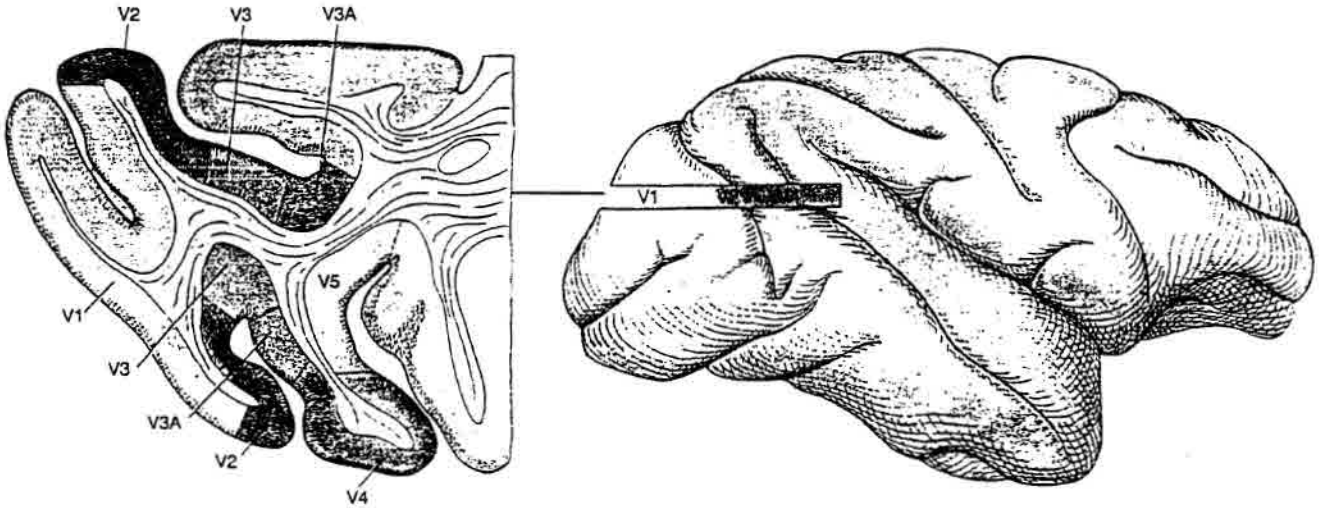


Fig. 2.2 - Subáreas do córtex visual.

FONTE: (12), p. 44.

Do ponto de vista fenomenológico algumas propriedades de interesse na percepção de movimento, segundo (42), são as seguintes:

- o estímulo para detecção de movimento é a variação contínua, no tempo, das intensidades eletromagnéticas sobre os fotorreceptores da retina. Por outro lado, o movimento também é percebido, mesmo não havendo estrita continuidade no tempo, quando um objeto é apresentado em posições diferentes, desde que sejam suficientemente próximas no espaço e a intervalos suficientemente pequenos. Esse fenômeno é chamado de **movimento aparente**. Um exemplo de percepção de movimento aparente acontece no caso da percepção de movimento nos filmes;
- observadores humanos são capazes de perceber diferentes movimentos em diferentes partes do campo visual, tais como quando dois objetos movimentam-se relativamente um ao outro;

- breves exposições ao estímulo são suficientes para a percepção de movimento, em geral exposições da ordem de 200 a 400 ms são suficientes pra que um observador seja capaz de detectar o movimento de um sinal do tipo senoidal; esse valor depende do contraste do estímulo;
- em geral a discriminação quantitativa de movimento é uma tarefa complexa, isto é, determinou-se que o ser humano tem um capacidade reduzida de inferir velocidades exatas a partir da informação visual.

É importante lembrar que a retina é formada por um número finito de fotorreceptores. Para cada receptor, a informação de movimento é conseguida com base na variação de intensidade eletromagnética nele incidente. Porém, a percepção de movimento não provém unicamente dessa informação pontual, mas de uma integração dessa informação de forma global que é feita no cérebro.

Nesse trabalho, serão estudadas algumas técnicas para a detecção automática de movimento. O termo **detector** referir-se-á à unidade elementar de determinação de movimento (em analogia com os fotorreceptores), e o termo **modelo** a um sistema que é capaz de integrar respostas de vários detectores.

Um elemento importante para a detecção de movimento em sequências de imagens é a capacidade de determinar os diferentes elementos constitutivos na imagem através das diferenças relativas de intensidade radiométrica, isto é, do contraste na imagem.

2.2 Detectores de Movimento

Os detectores já propostos na literatura podem ser classificados em três abordagens gerais: os baseados em correspondência de atributos, os baseados em gradiente espaço-temporal e os biológicos. A seguir serão apresentados os aspectos básicos dessas três abordagens.

2.2.1 Detectores baseados em correspondência de atributos

Dada uma sequência de imagens sucessivas no tempo, com pelo menos duas imagens, o primeiro passo da abordagem por correspondência de atributos é a detecção de atributos locais destacados em cada imagem. Podem ser, por exemplo, os atributos tipicamente utilizados nas áreas de visão por computador e processamento de imagens, como bordas, arcos, linhas, ou segmentos. Mais genericamente, pode-se detectar atributos arbitrários, intimamente ligados à natureza das imagens sendo observadas e que estão simultaneamente bem localizados e com alto contraste, sem considerar suas formas ou identidades específicas.

Tendo obtido os atributos, é feita uma correspondência entre eles para se obter pares, onde cada elemento do par pertence a uma de duas imagens subsequentes, ou seja, o que se pretende é localizar, para um ponto específico de uma imagem, seu correspondente na imagem seguinte. Uma discussão detalhada do problema da correspondência pode ser encontrada em (39).

A direção do movimento é estimada usando a posição relativa entre os atributos correlacionados, e a velocidade é estimada usando a distância entre os atributos e o intervalo de tempo entre as imagens.

Essa abordagem apresenta problemas tais como:

- devido a fatores como a oclusão, por exemplo, que pode fazer com que os atributos apareçam e desapareçam, nem sempre é fácil estabelecer a correspondência entre os atributos. E disso depende a precisão e a confiabilidade dos resultados da detecção de movimento. Principalmente em situações reais onde não se pode ter controle sobre as condições do ambiente e pode existir ruído nas imagens utilizadas;
- o processo de obtenção de correspondência de atributos costuma ter um custo computacional elevado (2);
- a escolha dos atributos a serem considerados não é simples, e depende em geral do problema em questão. É desejável que os pares de atributos sejam

obtidos em número e densidade espacial tais que permitam uma detecção adequada dos movimentos relativos ocorridos na cena. Por um lado, poucos atributos não conduzem a uma detecção satisfatória, mas, por outro, muitos atributos requerirão uma carga computacional muito elevada para o estabelecimento das correspondências.

Do ponto de vista neurofisiológico não existem evidências de que esse processo ocorra no sistema visual natural.

Esses problemas fazem com que a aplicação dos detectores baseados em correspondência de atributos seja restrita a problemas onde possam ser feitas simplificações. A mais comum delas é supor que as cenas contem apenas objetos rígidos onde o movimento pode ser determinado através de um número conhecido de parâmetros, pois, todo movimento rígido pode ser modelado através de uma rotação do objeto sobre um eixo passando pela origem de um sistema de coordenadas, seguido de uma translação nesse sistema de coordenadas. Não há restrição quanto a sua aplicação ao caso de movimento aparente.

Exemplos de trabalhos que utilizam detectores baseados nessa abordagem podem ser encontrados em (34,36,37,48).

2.2.2 Detectores baseados em gradiente

O movimento dos objetos na cena se traduz em variações da intensidade em cada ponto ao longo do espaço e no decorrer do tempo. Os detectores baseados em gradiente utilizam-se desse fato. Mais formalmente seja $I : \mathbb{R}^2 \times \mathbb{R}^+ \rightarrow \mathbb{R}$ a função que representa a intensidade de radiação da imagem nas duas coordenadas espaciais e num determinado instante de tempo, isto é, a intensidade de radiação no ponto (x, y) no instante t é dada por $I(x, y, t)$. Na abordagem baseada em gradiente empregam-se as variações temporal e espacial das intensidades da imagem, dado que a intensidade em cada ponto pode variar no tempo como decorrência do movimento dos objetos. Essas variações são usadas para se estimar o movimento em cada ponto da imagem. Pode-se formar a expansão da

função $I(x, y, t)$ em série de Taylor de acordo com a equação (2.1):

$$I(x + dx, y + dy, t + dt) = I(x, y, t) + \frac{\partial I(x, y, t)}{\partial x} dx + \frac{\partial I(x, y, t)}{\partial y} dy + \frac{\partial I(x, y, t)}{\partial t} dt + O \quad (2.1)$$

onde O denota os termos que dependem das derivadas de maior ordem da série. Pode-se desconsiderar O para pequenas variações de posição e de tempo (dx, dy, dt) . Além disso, supõe-se que no intervalo de tempo $(t, t + dt)$ a função intensidade definida sobre um ponto na imagem, sob translação de uma pequena distância (dx, dy) não se altera, ou seja, que vale:

$$I(x + dx, y + dy, t + dt) = I(x, y, t) \quad \forall x, \forall y, \forall t. \quad (2.2)$$

Combinando-se as equações (2.1) e (2.2) acima, e considerando o vetor velocidade como sendo $\mathbf{v} = (u, v)$, onde $u = \frac{dx}{dt}$ e $v = \frac{dy}{dt}$ são as componentes do vetor velocidade no ponto (x, y) , obtêm-se:

$$\frac{\partial I(x, y, t)}{\partial x} u + \frac{\partial I(x, y, t)}{\partial y} v + \frac{\partial I(x, y, t)}{\partial t} = 0.$$

Na equação (2.1) as derivadas parciais $\frac{\partial I(x, y, t)}{\partial x}$, $\frac{\partial I(x, y, t)}{\partial y}$ e $\frac{\partial I(x, y, t)}{\partial t}$ podem ser determinadas diretamente da imagem. Essa equação não é suficiente para determinar as duas componentes do vetor velocidade, assim é necessário introduzir alguma outra restrição para possibilitar a solução do sistema.

Várias versões de “equações de restrição” têm sido propostas. Horn e Schunk propõem uma restrição onde supõem objetos opacos e de tamanho finito, e levam em conta a suavidade da velocidade ao longo de pontos vizinhos, que devem ter velocidades similares (19). Nagel propõe uma restrição mais complexa baseada explicitamente nas propriedades geométricas da cena tri-dimensional (28). Uma discussão sobre a validade dessas restrições baseada em resultados experimentais pode ser encontrada em (45).

2.2.3 Detectores biológicos

Os detectores biológicos têm sido propostos com base em pesquisas sobre a visão nos animais. O trabalho pioneiro é o detector proposto por

Reichardt (32) (citado em (40)), a partir de experiências com o sistema visual de insetos. Na literatura consultada (40,41) esses detectores e seu comportamento são estudados através da Teoria de Sistemas. A seguir será apresentada uma discussão análoga, porém usando uma abordagem segundo o paradigma computacional e voltada para o processamento de imagens.

A idéia original do detector proposto por Reichardt consiste no uso de um filtro casado. A entrada desse filtro consiste de dois sinais, e a saída é máxima quando esses sinais coincidem. No caso do detector de Reichardt uma entrada é o deslocamento no espaço e no tempo da outra entrada, e a saída está relacionada ao deslocamento espaço-temporal.

Uma possível generalização para o caso idealizado de imagens bidimensionais contínuas no espaço e no tempo é apresentada a seguir. A detecção de movimento no ponto $(x, y) \in \mathbb{R}^2$, entre os instantes $0 \leq t_1 < t_2 < +\infty$ na direção associada à reta que passa pelo ponto no ângulo $0 \leq \theta < \pi$, é obtida da seguinte maneira: seja $\Delta > 0$ o número pontos de deslocamento de interesse dessa detecção; os deslocamentos nas direções horizontal e vertical são dados, respectivamente, por $\Delta_x = \Delta \cdot \text{sen}(\theta)$ e $\Delta_y = \Delta \cdot \text{cos}(\theta)$. Define-se:

$$\begin{aligned} v_1(x, y, \Delta, \theta, t_1, t_2) &= I(x, y, t_1) \cdot I(x + \Delta_x, y + \Delta_y, t_2) \\ v_2(x, y, \Delta, \theta, t_1, t_2) &= I(x, y, t_2) \cdot I(x + \Delta_x, y + \Delta_y, t_1), \end{aligned} \quad (2.3)$$

e o valor associado à direção θ , no ponto (x, y) nos instantes t_1, t_2 e ao deslocamento Δ , é dado por:

$$D(x, y, \Delta, \theta, t_1, t_2) = v_1 - v_2. \quad (2.4)$$

Esse procedimento é ilustrado na Figura 2.3, onde T representa a separação temporal entre as imagens, isto é $T = t_2 - t_1$.

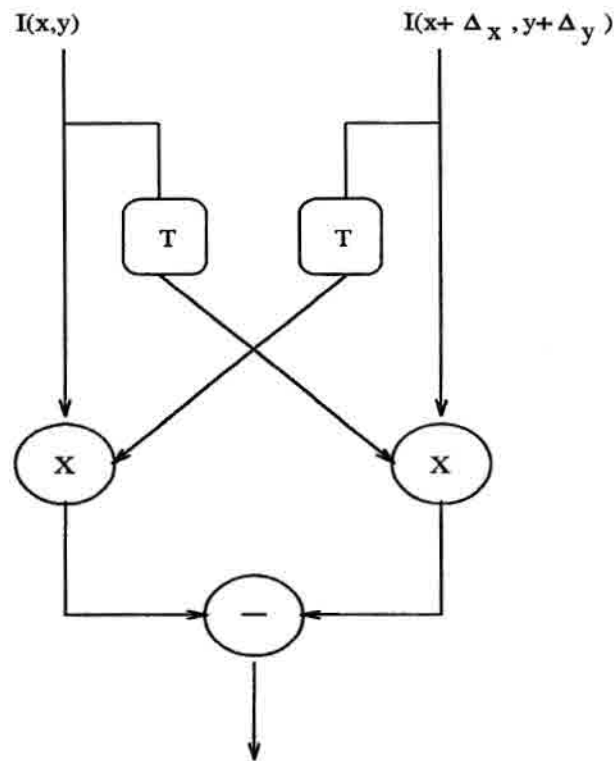


Fig. 2.3 - O detector de Reichardt.

FONTE: Adaptado de (40), p. 298.

Mostra-se a seguir um exemplo da aplicação desse procedimento para um par de imagens sintéticas (ver Figura 2.4). Nos instantes t_1 e t_2 as imagens são dadas por:

$$I(x, y, t_1) = \begin{cases} 10 & \text{para } (x_1, y_1) \text{ e } (x_2, y_2) \\ 0 & \text{caso contrário} \end{cases}$$

$$I(x, y, t_2) = \begin{cases} 10 & \text{para } (x_0, y_0) \text{ e } (x_3, y_3) \\ 0 & \text{caso contrário.} \end{cases} \quad (2.5)$$

Toma-se $\theta = \pi/4$ e $\Delta = 1$. Para esse caso, os valores de $D(x_0, y_0, \theta, \Delta, t_1, t_2)$ e

$D(x_2, y_2, \theta, \Delta, t_1, t_2)$ são calculados por:

$$D(x_0, y_0, \theta, \Delta, t_1, t_2) = 0 \cdot 0 - 10 \cdot 10 = -100,$$

$$D(x_2, y_2, \theta, \Delta, t_1, t_2) = 10 \cdot 10 - 0 \cdot 0 = 100.$$

Assim temos que, em cada ponto da imagem inicial, a saída do detector quando positiva indica que esse ponto sofre um deslocamento Δ no sentido θ . Quando a saída do detector for negativa indica que o ponto $(x + \Delta_x, y + \Delta_y)$ sofreu um deslocamento Δ no sentido $\theta + \pi$.

O valor $v_1(x, y, \Delta, \theta, t_1, t_2)$, pode ser visto como a saída de um subdetector responsável pela detecção de movimento no sentido θ , e $v_2(x, y, \Delta, \theta, t_1, t_2)$ como a saída de um subdetector responsável pela detecção de movimento no sentido $\theta + \pi$.

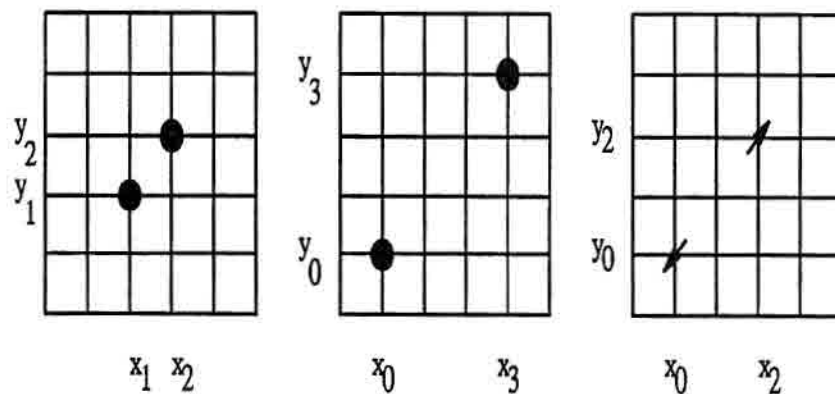


Fig. 2.4 - Exemplo da aplicação de um Detector de Reichardt a um par de imagens sintéticas.

x	0	1	2	3	4	
$t = 1$	1	2	3	4	1	$\cdots I(x, 1)$
$t = 2$	4	1	2	3	4	$\cdots I(x, 2)$

Fig. 2.5 - Dois sinais periódicos.

Em (41) é apresentado um formalismo rigoroso para o estudo desse detector. Os autores fizeram uma série de hipóteses a respeito dos sinais de entrada, por exemplo, considerando-se os sinais como unidimensionais e periódicos. A partir dessas suposições puderam ser obtidos alguns resultados quanto ao comportamento desse tipo de detector. Esses resultados permitem verificar que esse detector apresenta problemas de desempenho, entre outros para o caso de sinais periódicos, dependendo da frequência do sinal de entrada.

Para ilustrar um desses problemas considere-se o seguinte exemplo típico de processamento de sinais unidimensionais. Seja $I : \mathbb{N} \times \mathbb{R}^+ \rightarrow \mathbb{R}$ dada por:

$$\begin{aligned} I(x, t_1) &= (x \bmod 4) \\ I(x, t_2) &= I(x + 1, t_1), \end{aligned}$$

e que são mostrados na Figura 2.5. Evidentemente esses sinais são periódicos de período 4. Dados esses dois sinais, e supondo desconhecido o deslocamento relativo, deseja-se observar o resultado apresentado pelo detector de Reichardt. Como nesse caso os sinais são unidimensionais, $\theta = 0$ e $\Delta = \Delta_x$. Suponha que deseja-se detectar o movimento no ponto $x = 0$ para o deslocamento $\Delta_x = 1$ entre os instantes t_1 e t_2 . Assim sendo, tem-se:

$$D(x, y, \theta, \Delta, t_1, t_2) = D(x, \Delta_x, t_1, t_2) = -7.$$

Desta maneira verifica-se que para esse tipo de sinal o detector forneceu um resultado errado, um valor de saída negativo, isto é, sinalizou um deslocamento no sentido $\theta + \pi$ (da direita para a esquerda), sendo que na realidade o deslocamento é no sentido contrário.

Os autores (40) analisaram sob quais condições o desempenho do detector é pobre e apontam algumas modificações visando superar esses problemas. Essas modificações são em duas linhas:

- 1) não usar como entrada apenas o valor do sinal no ponto, mas considerar, de alguma maneira, os valores em uma certa vizinhança do ponto de interesse;
- 2) não usar a saída como um valor quantitativo na detecção de movimento, mas considerá-lo como mera evidência das **chances** de ter acontecido movimento nesse sentido.

O termo “chances” acima não é empregado no sentido matemático da palavra; isto é, se $\text{IPr}(A)$ é a probabilidade do evento A , então $\text{IPr}(A)/(1 - \text{IPr}(A))$ é definida como a **chance relativa** do evento A . No contexto desse trabalho, a palavra “chance” é empregada informalmente para expressar, em forma numérica, evidência probabilística.

No que diz respeito ao primeiro ponto, uma das formas possíveis de se levar em conta a vizinhança do ponto de interesse (**contexto**) é a introdução de uma função local de ponderação, chamada de **campo receptivo**. Essa modificação torna o detector fisiologicamente mais plausível quando comparado com as respostas de células sensíveis a movimento encontradas na área V1 de mamíferos (30,40). O problema da ponderação é parcialmente abordado em (40), através do uso da Teoria de Sistemas para sinais determinísticos simples. No Capítulo 4 o mesmo problema será tratado para o caso da detecção de movimento em sequências de imagens.

No que diz respeito ao segundo ponto acima, é importante frisar que podem existir várias propostas diferentes para fazer com que as saídas dos detectores sejam indicadores das chances de haver acontecido movimento em cada sentido. Os autores de (40) apontam algumas hipóteses, expressas de maneira generalizada através das propriedades **P1 - P3** mais a frente, que poderiam ser úteis para o problema de transformar números em chances.

A seguir é apresentado um formalismo que resulta útil para incorporar essas hipóteses e que, no Capítulo 4, é empregado para (usando o paradigma probabilístico) fornecer uma solução interessante e computacionalmente

viável. Cabe ressaltar que esse formalismo é uma generalização dos resultados vistos em (40). Os autores trabalham com o caso unidimensional onde existe uma única direção à qual são aplicados vários detectores. A generalização é para o caso de detecção de movimento em imagens, isto é, o problema é bidimensional e deve-se trabalhar com $2K$ sentidos correspondentes às K direções de interesse; a cada direção é aplicado um detector.

Reescrevendo as hipóteses **P1-P3** propostas em (40) e generalizando para o caso de interesse, o problema pode ser formulado da seguinte maneira: deseja-se detectar o movimento em K direções a partir de uma sequência a tempo contínuo de imagens digitais discretas. Denota-se essa sequência como uma função da forma $L : S \times \mathbb{R}^+ \rightarrow \{0, \dots, 255\}$, onde $S = \{0, \dots, M\} \times \{0, \dots, N\}$ é o domínio das imagens, e $t \in \mathbb{R}^+$ representa o tempo e $\{0, \dots, 255\}$ é o conjunto dos tons de cinza presentes na imagem.

Em cada $s \in S$ são aplicados K detectores, cada um respondendo na direção dada por k em dois sentidos especificados pelos ângulos $k\pi/K$ e $(k\pi/K) + \pi$, para todo $0 \leq k < K$. Cada detector será denotado D_k ; eles são da forma $D_k(s) : L \rightarrow \mathbb{R}$ e tipicamente sua saída depende do ponto s ao qual ele é aplicado e de uma vizinhança espacial $v(s)$ desse ponto. Isto é, se L e L' são duas sequências diferentes de imagens tais que $L(u) = L'(u)$ para todo $u \in \{s\} \cup v(s)$, onde $v(s) \subset S$ é a vizinhança espacial de s , então $D_k(L)(s) = D_k(L')(s)$. Toda vez que os detectores forem aplicados a uma sequência de imagens L , por simplicidade de notação esse argumento será omitido, isto é, a função $D_k(L)(s)$ será escrita $D_k(s)$. Dada uma sequência L , se a saída $D_k(s) > 0$ ($D_k(s) < 0$ respectivamente), então considera-se que existe evidência de movimento na coordenada $s \in S$ no sentido associado ao ângulo $k\pi/K$ ($k\pi/K + \pi$, respectivamente); caso $D_k(s) = 0$ não existe evidência de movimento na direção associada ao ângulo $k\pi/K$ no ponto $s \in S$. O resultado de aplicar o detector D_k à sequência L será denotado $X_k = (X_k(s))_{s \in S} = (D_k(L)(s))_{s \in S}$, para todo $0 \leq k < K$ e para todo $s \in S$. Note-se que por definição, $X_k(s) \in \mathbb{R}$.

Para transformar números em chances relativas os autores de (40) propõem o uso de uma função $V : \mathbb{R}^k \times \{0, \dots, 2K - 1\} \rightarrow [0, 1]$ da forma

$$V_\ell(X_0, \dots, X_{K-1})(L, s)_{0 \leq \ell < K}$$

com as seguintes propriedades:

P1 $(V_\ell(L, s))_{0 \leq \ell < 2K}$ é a probabilidade de a sequência L apresentar movimento, na coordenada $s \in S$, no sentido dado pelo ângulo $\theta = \ell\pi/K$ para todo $\ell \in \{0, \dots, 2K - 1\}$, sempre que exista pelo menos um k tal que $D_k(s) \neq 0$. Nesse caso, é possível impor a seguinte propriedade:

$$\sum_{0 \leq \ell < 2K} V_\ell(X_0, \dots, X_{K-1})(L, s) = 1;$$

P2 se L e L' são duas sequências tais que $D_k(L)(s) = -D_k(L')(s)$ então vale que $V_\ell(X_0, \dots, X_{K-1})(L, s) = V_{\ell+K}(X_0, \dots, X_{K-1})(L', s)$. Essa propriedade, chamada pelos autores de **antissimetria**, pode ser interpretada dizendo que a forma de atribuir chances a sentidos opostos, em função das saídas dos detectores na mesma direção, independe do sentido, isto é nenhum dos dois sentidos é favorecido;

P3 as funções V_ℓ são não decrescentes em todos os seus argumentos X_k . Por exemplo são não-decrescentes em X_0 : se $x < y$, então $V_\ell(x, \dots, X_{K-1})(\cdot, \cdot) \leq V_\ell(y, \dots, X_{K-1})(\cdot, \cdot)$. Analogamente, é não decrescente em X_1, \dots, X_{K-1} .

Os autores não especificam qual função V usar.

Suponha-se agora dados dois sinais para os quais deseja-se detectar em cada ponto, para um certo deslocamento, o movimento em vários sentidos. Suponha-se ainda que a função V esteja especificada. Uma vez aplicados os detectores é desejável que a cada ponto seja associado um único sentido de deslocamento baseado nos valores de saída dos detectores sensíveis a direção em cada ponto, tal que se obtenha uma certa consistência global na solução do problema. Por consistência global entende-se, nesse trabalho, que a cada ponto será associado uma direção de deslocamento e entre pontos vizinhos, esses deslocamentos serão o mais similares possíveis (esta noção será definida mais precisamente no Capítulo 4.

A proposta de Van Santen e Sperling (41), chamada de ERD (**Elaborated Reichardt Detector**: detector elaborado de Reichardt) consiste em alterar o modelo original com a introdução dos campos receptivos. Além disso, em cada ponto são aplicados vários detectores sensíveis à direções diferentes, e as respostas desses detectores (suas variações no tempo, para o caso contínuo, e entre pares de imagens diferentes, no caso de movimento aparente) devem ser combinadas por uma função chamada de **voting rule**. Exemplos de algoritmos para o caso de movimento contínuo são: escolher em cada ponto o deslocamento associado ao detector cuja saída seja máxima durante um intervalo de tempo observado. Outro algoritmo seria somar as saídas dos detectores de acordo com uma função de ponderação, definida *a priori*, e executar um processo de decisão de acordo com esse valor. Os autores argumentam que para muitos estímulos (sinais de entrada), a escolha dessa voting rule depende, em geral, do tipo de movimento que se pretende observar, e da forma do sinal de entrada.

Outras propostas de detectores biológicos podem ser encontradas em (1) e (42). Para esses detectores vale dizer que podem ser equivalentes, em termos das operações que executam e de resultados, ao ERD, como pode ser visto em (41). As diferenças entre eles podem ser apontadas apenas com base em experimentos psicológicos e fisiológicos que examinam as respostas da detecção de movimento em detalhes, apesar de terem sido propostos partindo de filosofias diferentes. A escolha entre um ou outro deve ser feita com base, por exemplo, na conveniência matemática das operações que executam.

2.3 Modelos de Detecção

Como já foi dito acima, um modelo de detecção deve incluir alguma forma de compatibilizar saídas de diversos detectores, em cada ponto, em uma solução globalmente consistente.

Esse problema foi abordado em (30), onde os autores propõem que a tarefa de compatibilizar as evidências de movimento de cada detector seletivo a uma direção sejam simplificadas, ou seja que a compatibilização ocorra apenas em

um subconjunto de detectores. As tarefas de escolher o subconjunto e de compatibilizar os detectores são realizadas por um mecanismo ativo, onde os dois estágios ocorrem em paralelo, e sua implementação é feita através de redes de neurônios interconectados. Essas redes são treinadas com um conjunto de imagens sintéticas, cujo movimento é conhecido. Nesse trabalho apresenta-se duas propostas de compatibilização das saídas de detectores sensíveis a diferentes direções. A primeira é chamada de Regra do Máximo, e faz a atribuição de um valor de deslocamento em um ponto com base nos valores de saída dos detectores nesse ponto. A segunda, através de uma rede neural do tipo Hopfield, atribui a cada ponto um valor de deslocamento de acordo com os valores de saída dos detectores sensíveis a direção no ponto e em alguma vizinhança. Esses modelos serão apresentados no Capítulo 4.

No Capítulo seguinte faz-se uma revisão sobre redes neurais e sua utilização.

CAPÍTULO 3

REDES NEURAIS

Existem tarefas como: solução de problemas científicos e matemáticos; criação, manipulação e manutenção de bancos de dados; comunicação eletrônica; processamento de textos, gráficos entre outras, que são perfeitamente tratadas pela computação convencional. Entretanto, existem aplicações que gostaria-se de automatizar, em computadores sequenciais, o que no entanto não é viável, devido a grande complexidade associada com sua programação.

Uma tecnologia diferente do processamento convencional do tipo Máquina de Turing, que inspira-se na pesquisa nos princípios funcionais dos neurônios e de suas interligações, é a **Neurocomputação**, ou **Conexionismo**. Os sistemas neurais artificiais, ou simplesmente **redes neurais**, são os modelos de processamento que têm sido usados para tarefas complexas em classificação, reconhecimento de padrões, interpolação de funções, otimização combinatória e tarefas de previsão, por exemplo, de séries temporais.

Esses problemas têm em comum algumas características como: em geral suas forma de solução são altamente paralelas, os seres humanos sabem como fazê-las de forma rápida e eficiente, podem ser coletados grandes conjuntos de pares de exemplos de tarefas e sua solução, e também o fato de que cada tarefa envolve a associação de objetos entre conjuntos.

A neurocomputação preocupa-se em capturar os princípios funcionais dos neurônios, redes de neurônios e do sistema nervoso que levam a soluções de problemas complexos e aplicá-los aos sistemas computacionais. Pode-se pensar que em retribuição, os modelos de redes neurais artificiais servem para validar um possível desempenho de modelos estudados experimentalmente na neurofisiologia (35).

3.1 Componentes fisiológicos

O cérebro humano é muito complexo e por isso pode-se dizer que ele é ainda pouco entendido. Não existem respostas satisfatórias para questões fundamentais como “o que é a mente?” ou “como pensamos?”. Não se conhece como o cérebro representa informação a nível mais alto. Apesar disso, estudos levados nas décadas passadas tem trazido alguma base explicativa para o funcionamento de certas áreas do cérebro e do sistema nervoso associado a ele.

O elemento básico do sistema nervoso é o neurônio. As componentes do neurônio são: o corpo central da célula, e ligado a ele, os dendrites e um axônio. Existem muitos tipos de neurônios, com configurações e funcionamentos diferentes, que se diferenciam por sua forma (12). A Figura 3.1 mostra algumas formas de neurônios.

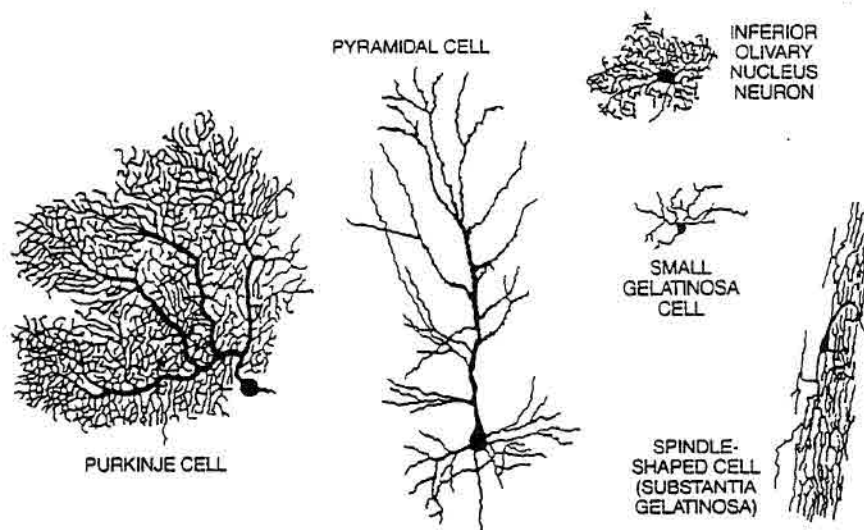


Fig. 3.1 - Alguns tipos de neurônios.

FONTE: Adaptada de (12), p. 29.

Os neurônios estão agrupados em redes neurais. Assim o axônio de um neurônio liga-se aos dendrites de outros neurônios em contatos chamados de

sinapses.

Estima-se que o cérebro tem da ordem de dez a quinhentos bilhões de neurônios, arranjados em cerca de mil módulos principais, tendo cada módulo cerca de 500 redes neurais. Cada rede neural, tem por sua vez, da ordem de cem mil neurônios. O axon de cada neurônio pode conectar-se com cerca de cem (ou até vários milhares) de outros neurônios, dependendo de sua localização e tipo (6).

O estado interno de um neurônio é caracterizado em parte por uma diferença de potencial elétrico através da membrana que o separa do plasma intracelular. Essa diferença de potencial é chamada de **potencial de geração**. As sinapses com diversos outros neurônios são capazes de alterar o seu potencial de geração através da liberação de substâncias químicas chamadas de neurotransmissores, e quando este atinge um determinado limiar um impulso é produzido e propagado nos ramos do axônio até as sinapses com os dendrites de outros neurônios.

As atividades sinápticas podem ser excitatórias ou inibitórias dependendo do tipo de interação local entre o neurotransmissor e a membrana do neurônio. Uma atividade é excitatória se ela aumenta a probabilidade da produção de um impulso como atividade pós-sináptica do receptor, e inibitória caso ela diminua essa probabilidade. As mudanças induzidas por cada entrada do neurônio são combinadas para produzir uma mudança no seu potencial de geração.

3.2 Conceitos e Definições em Redes Neurais Artificiais

Uma rede neural artificial, que a partir daqui será tratada apenas como **Rede Neural**, será definida como uma estrutura do tipo paralela distribuída para processamento de informação que consiste de elementos de processamento, nós ou neurônios, interconectados via ligações unidirecionais chamadas de conexões. A alta conectividade das redes neurais fazem com que sejam tolerantes a falhas de alguns dos elementos de processamento sem implicar em perdas muito significativas quanto ao desempenho da rede.

Quanto ao seu funcionamento pode-se ressaltar os seguintes aspectos:

- sabe-se que são adaptativas: elas podem tomar dados e aprender a partir deles, inferindo as soluções a partir dos dados. Essa capacidade difere das técnicas convencionais de software, porque não depende do conhecimento *a priori*, por parte de programadores, isto é, de regras explícitas que descrevam numericamente como obter essas soluções;
- podem generalizar: elas podem processar corretamente dados apenas similares àqueles usados inicialmente no seu treinamento (o treinamento das redes será discutido mais adiante). Com isso podem tratar dados imperfeitos ou incompletos, o que é de grande utilidade em aplicações práticas visto que os dados do mundo real são, por natureza, ruidosos, ou seja, distorcidos, afetados pelo ambiente;
- são não-lineares, podendo capturar relações complexas entre as entradas que são apresentadas.

Cada nó pode possuir uma memória local e executar processamento local de informação. Possui uma única saída conectada a vários outros nós, onde cada conexão transporta o mesmo sinal, de qualquer tipo matemático desejado, que é a saída do processamento local. O processamento de informação, executado dentro de cada nó, pode ser linear ou não-linear, e é definido arbitrariamente com a restrição de que seja local, isto é, dependa exclusivamente dos sinais atuais chegando ao nó via conexões, e de valores armazenados na memória local. As intensidade das conexões entre os nós (as sinapses), são representadas pela associação de pesos a cada uma delas. O sinal do peso significa excitação quando positivo e inibição quando negativo.

Considerando-se uma rede neural com K nós pode-se defini-la matematicamente da seguinte forma: cada nó será denotado genericamente por m , isto é, $m \in \{1, \dots, K\}$. O nó m tem $E(m)$ entradas e, portanto, a rede possui $E = \sum_{m=1}^K E(m)$ entradas. O valor da entrada j do nó m denota-se por x_{jm} , com

$1 \leq j \leq E(m)$ e $1 \leq m \leq K$. À entrada x_{jm} associa-se um peso real denotado ω_{jm} . A saída do nó m convencionalmente é dada por:

$$y_m = \mathcal{F} \left(\sum_{j=1}^{E(m)} x_{jm} \omega_{jm} \right), \quad (3.1)$$

onde \mathcal{F} é a função de transferência do nó m . Uma representação de um nó artificial pode ser visto na Figura 3.2.

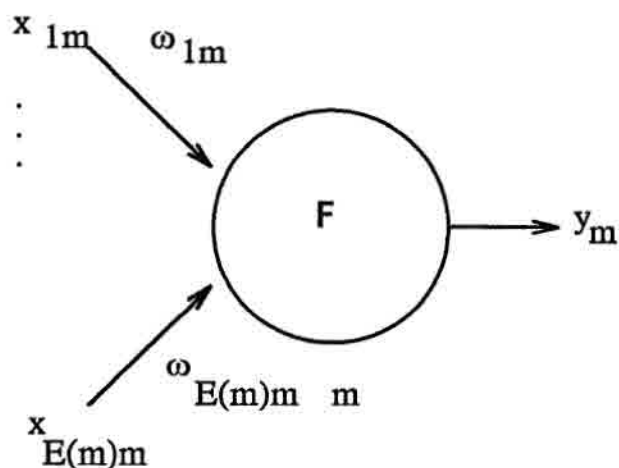


Fig. 3.2 - Representação diagramática de um neurônio artificial.

A topologia da rede é dada pela associação dos dados em certas entradas $x_{..}$ e pela coleta de resultados em certas saídas $y_{..}$. Para a compreensão da topologia, os nós podem ser arranjados em $T + 1 \in \mathbb{N}$ camadas onde:

- a entrada dos nós da camada $t = 0$, são os dados
- a saída dos nós da camada $t = T$, são os resultados
- a saída dos nós da camada $t \leq T - 1$ servem de entrada para os nós camada $t + 1$.

A escolha de uma topologia em particular, ou seja, a escolha do número de nós da rede e sua organização em camadas, bem como da função de transferência, dependem de considerações a respeito da aplicação (volume dos dados, tempo disponível para efetuar o treinamento, hardware ou simulador a ser empregado, precisão dos resultados desejados, etc.). É comum que todo nó pertencente a uma camada tenha a mesma função de transferência.

As funções de transferência podem operar continuamente ou por eventos. Se a operação é por eventos, deve haver uma entrada chamada **ativação** que permite à função nó operar. Nós do tipo contínuo operam continuamente.

Em (11) propõe-se uma formalização rigorosa bem como uma uniformização de nomenclatura e notação, com o intuito de permitir uma base teórica sólida, do ponto de vista matemático, para o paradigma de neurocomputação. Propõe-se, também, uma classificação hierárquica dos diferentes tipos de redes. Uma discussão sobre as capacidades e limitações das redes neurais, através de um tratamento matemático rigoroso pode ser vista em (4).

3.3 Aprendizagem

Por **aprendizagem** entende-se o mecanismo pelo qual os nós de alguma forma mudam os seus valores em resposta ao ambiente (6). É conhecido que na maioria das redes neurais adaptáveis, a aprendizagem é realizada através da modificação dos pesos associados às conexões entre os nós.

Para efeito de formalização matemática, supõe-se que todos pesos são números reais. Para cada nó m tem-se $E(m)$ entradas e a cada entrada um peso associado, ou seja cada, nó possui associado a ele um vetor peso $\omega_m = (\omega_{1m}, \dots, \omega_{E(m)m})^T$, e para uma rede com N nós (indexados de 1 a N) tem-se um vetor peso ω para a rede toda, onde $\omega = (\omega_1^T, \dots, \omega_N^T)$.

Para redes neurais que usam o treinamento por modificação de pesos, a aprendizagem realiza-se através do desenvolvimento de regras que melhor

guiem o vetor peso para uma posição que dê à rede o desempenho desejado.

Uma abordagem comumente encontrada é a de mudar ω para uma posição que possibilite à rede neural minimizar (ou maximizar, respectivamente) uma função custo ou perda (função desempenho, respectivamente), tal como o erro médio quadrático (medida de qualidade, respectivamente).

As categorias de treinamento utilizadas em neurocomputação são:

- *supervisionada*: a rede é alimentada com uma sequência suficientemente representativa de exemplos do tipo (x_i, y_i) de pares de entradas e saídas corretas. Para cada entrada x_i uma saída y_i é fornecida à rede de modo a informar precisamente qual a saída que a rede deve fornecer;
- *reforço*: similar a aprendizagem supervisionada, mas em vez de ser dada a saída correta a rede recebe somente uma graduação que indica a qualidade da resposta fornecida;
- *auto-organizável*: a rede se modifica em respostas às entradas x sem recorrer a nenhum valor de saída y ou a alguma graduação de desempenho.

3.4 Paradigmas de Redes Neurais

Pelo menos cinquenta tipos de redes neurais tem sido utilizadas em pesquisa ou em desenvolvimento. No entanto, treze são as mais comumente estudadas. As grandes diferenças entre elas estão na regra de aprendizagem incorporadas, nas suas funções de transferência, na topologia das conexões e nos pesos associados às suas conexões. A seguir comenta-se os aspectos básicos de alguns desses tipos de redes. Para uma discussão detalhada sobre esses tipos de redes, suas aplicações e implementações, ver (13,35).

- *Perceptron*: o mais antigo elemento de processamento que é chamado de neurônio, foi proposto em 1958. Esse modelo de processamento deu origem a várias topologias de redes neurais artificiais. Pode ser utilizado em uma

forma simplificada, com apenas um neurônio, ou com diversos neurônios arranjados em camadas. É tipicamente utilizado em aplicações de reconhecimento de padrões lineares para duas classes.

- *Mapas Auto-organizáveis*: referem-se a redes de neurônios que aprendem através de um processo de auto-organização com a apresentação de somente, um padrão de entrada. Usam o princípio de que entradas similares são mapeadas a neurônios próximos. Foi estudada pioneiramente por Kohonen e aplicada com sucesso em tarefas de reconhecimento de fonemas e sua transformação em caracteres.
- *Back-propagation (Retropropagação do erro)*: a rede neural mais conhecida e utilizada devido a sua simplicidade de compreensão e implementação. Alguns exemplos de aplicações podem ser encontrados com sucesso nas áreas de processamento de imagens e de som, aplicações financeiras e de reconhecimento de padrões.
- *Memória Associativa Bidirecional*: usadas para fazer associações análogas às realizadas pelo cérebro, tendo com entradas padrões incompletos. Usadas em aplicações em processamento de imagens, controle e alocação de recursos.
- *Hopfield*: tem o mesmo princípio de funcionamento das memórias associativas. Uma extensão de suas aplicações é na área de otimização combinatória. Essa aplicação será discutida em detalhes no próximo capítulo.

3.5 Implementações

As primeiras implementações de redes neurais foram quase exclusivamente simulações em computadores do tipo máquina de Turing, e até hoje é o tipo mais comum de implementação. A maioria dos modelos de redes neurais, geralmente, é suportada por simulações em software (26).

Um outro método para implementar as redes neurais, mais eficiente que as simulações, é a sua emulação dentro de arquiteturas paralelas. A estru-

tura dos elementos de processamento e suas interconexões estão em correspondência com os processadores e suas ligações (35).

Os neurocomputadores representam uma nova classe de computadores, otimizados para processarem redes neurais. Sua arquitetura é diferente das máquinas paralelas. Apesar de utilizarem um alto grau de paralelismo, a programação sequencial é ainda usada. De fato, os mecanismos de processamento de informação são projetados para implementar um sistema de equações diferenciais associado às redes neurais (35).

Aplicações em tempo real geralmente necessitam de maior velocidade do que aquela conseguida com neurocomputadores que simulam um grande número de processadores. Pesquisa-se uma forma de construir redes neurais diretamente em circuitos integrados. Ao invés de simular paralelismo, esse está inerente no projeto do hardware.

Existem ainda implementações ópticas, ou seja que, ao invés de trabalharem com sinais elétricos, como é o caso dos semicondutores de silício, envolvem o uso de componentes ópticos. Estas implementações estão restritas a laboratórios com equipamentos sofisticados, porém, são tidas como promissoras (35) devido a sua velocidade de operação.

CAPÍTULO 4

UMA SOLUÇÃO PARA O PROBLEMA DA DETECÇÃO DE MOVIMENTO BASEADA EM REDES NEURAIIS

4.1 O problema tratado

Nesse trabalho estuda-se um modelo de detecção de movimento de nuvens em imagens de satélites meteorológicos. Nessas imagens, as nuvens movimentam-se em várias direções, ao mesmo tempo em que podem sofrer deformações; por isso não é possível identificar atributos determinísticos, como bordas por exemplo. Além disso, a intensidade nos pontos que formam as nuvens não é constante. Assim sendo, as nuvens não podem ser vistas como objetos rígidos e é com esse enfoque, o da detecção de movimento de objetos não rígidos, que pretende-se desenvolver esse estudo.

Além do interesse em movimento de corpos não rígidos, pode-se ressaltar o interesse específico em movimento de nuvens como é demonstrado em (27). O autor fala sobre um sistema de combate a tempestades de granizo, que opera sobre imagens de radar, onde aparecem nuvens que estão se deslocando. A detecção do movimento das nuvens nas imagens obtidas pelo radar é feita manualmente. Os usuários reconhecem que o sistema teria uma melhora substancial no seu funcionamento se a tarefa de identificação dos campos de deslocamento das nuvens pudesse ser feita de forma automática.

Deseja-se estudar esse problema com um enfoque voltado para o paradigma de redes neurais, descrito no Capítulo anterior. As redes neurais são implementações de modelos de processamento de informação inspirados em estudos sobre o funcionamento de certas áreas cerebrais. Estes estudos têm permitido chegar a modelos bastante correlacionados com aqueles estudados experimentalmente na

neurociência, ainda que os modelos funcionais de neurônios e redes neurais propostos na neurocomputação sejam simplificações dos elementos fisiológicos naturais. Pode-se dizer que esses elementos e funções ainda não são totalmente compreendidos pela neurofisiologia, porém, muitos fenômenos perceptivos têm sido simulados via redes neurais artificiais com sucesso (44).

Os autores de (29) mostram que as respostas neurais em muitas regiões do cérebro estão organizadas em padrões espaciais característicos que refletem, de alguma forma, aspectos quanto a seu funcionamento. Mas ainda, não se conhecem totalmente os princípios gerais que relacionam as estruturas às suas funcionalidades. A proposição de modelos de processamento altamente paralelos que imitam as estruturas neurológicas pode gerar avanços na compreensão de como o cérebro executa determinadas tarefas; um exemplo desse tipo de modelagem pode ser visto em (43), onde os autores propõem um modelo que simula o processamento visual primário de um mamífero (um gato adulto) baseado em dados anatômicos e fisiológicos.

Para o caso da detecção de movimento, em particular, deseja-se aproveitar o alto paralelismo das redes neurais, visto que as operações executadas pelos processadores de cada coordenada para a qual está sendo detectado movimento são as mesmas e podem ser executadas em paralelo. Existem simuladores capazes de fornecer uma implementação eficiente quando a função executada é a mesma (26).

Um exemplo de rede neural baseada em elementos fisiológicos e anatômicos do sistema visual para percepção de movimento pode ser visto em (47). Os autores constroem uma rede neural artificial, associando cada nó da rede a uma célula do córtex visual, com conexões locais em uma vizinhança, tentando reproduzir uma organização, ainda sendo estudada em fisiologia, de neurônios do córtex visual. Para cada coordenada das imagens da sequência existe um conjunto de nós seletivos a um tipo de deslocamento. Os autores propõem um algoritmo de avaliação do estado de saída desses nós, que representará o deslocamento associado a esse ponto. Na abordagem proposta nesse trabalho, a informação de vizinhança é levada em consideração quando se usa a informação de suavidade espacial (descrita mais adiante),

porém sem reproduzir estruturas anatômicas do processamento visual natural.

A seguir será descrita de uma maneira mais formal e genérica o problema da detecção de movimento.

4.2 Formulação do problema

Considere a situação de se ter N imagens tomadas de uma mesma região em sequência crescente no tempo, e que possa existir movimento e deformação entre alguns dos objetos imageados. Deseja-se detectar automaticamente em cada ponto uma direção que represente o movimento desse ponto ao longo da sequência de N imagens.

Com base no que foi dito no Capítulo 2 nesse trabalho, o problema foi estudado pelo uso de detectores seletivos à direção. Isto é, dada uma sequência de imagens serão formadas K imagens que representam as saídas de K detectores de movimento.

Essas saídas representam os possíveis deslocamentos em $2K$ sentidos diferentes de $\pi/(2K)$ (veja Figura 4.1) e esses campos são compatibilizados em um único campo de deslocamento global onde procura-se uma consistência global.

A notação empregada é a seguinte: seja S o suporte (ou domínio (5)) das imagens, que sempre será considerado um retângulo finito $\mathbb{N} \times \mathbb{N}$, onde \mathbb{N} é o conjunto dos naturais. Em particular toma-se $S = \{0, \dots, M-1\} \times \{0, \dots, N-1\}$. A cardinalidade de um conjunto A é denotada $\#A$. Uma imagem de intensidades radiométricas é dada pela coleção de $\#S$ pixels, onde o valor do nível de cinza de cada pixel é denotado por $I(s)$, isto é, $I: S \rightarrow \Lambda$, onde $\Lambda \subset \mathbb{N}$. No caso considerado nesse trabalho, $\Lambda = \{0, \dots, 255\}$, porém existem outros formatos possíveis.

Desta maneira, a sequência de imagens de intensidades de entrada será denotada por $L: S \times \mathbb{N} \rightarrow \Lambda$, com $L = (I(s, t))_{s \in S, t \in \{1, 2, \dots\}}$. Esta sequência constitui-se nos dados de entrada do problema, e o conjunto de todas as possíveis sequências de imagens será denotado \mathcal{L} .

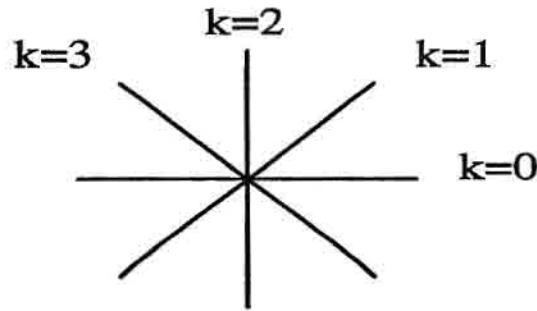


Fig. 4.1 - As quatro direções propostas para a detecção de movimento

A detecção para a direção dada por $k \in \{0, \dots, K - 1\}$ na coordenada $s \in S$ baseada na sequência $L \in \mathcal{L}$ de imagens é representada por:

$$X_k(s) = D_k(L, s),$$

onde a função $D_k: \mathcal{L} \times S \rightarrow \mathbb{R}$ é chamada de detector da direção k . O resultado $X_k = (X_k(s))_{s \in S, 0 \leq k < K}$ é chamado de **imagem de direção** k e está associado ao movimento na direção k em seus dois sentidos possíveis, ou seja, nos sentidos dados pelos ângulos $k\pi/K$ e $(k\pi/K) + \pi$.

Suponha-se agora fixados os detectores D_k e formadas as imagens de direção X_k para a sequência L . A fase seguinte é a proposição de um **modelo** de detecção de movimento, ou seja, a proposição de uma regra que compatibilize as k imagens de direção em um único campo vetorial $\widehat{X} = F(X_0, \dots, X_{K-1})$ que seja representativo do movimento observado nos objetos imageados na sequência L . Os domínio e contradomínio de F serão especificados mais adiante. Ver Figura 4.2.

A fim de simplificar a apresentação dos próximos itens, de agora em diante será considerado que valem as seguintes hipóteses:

- a sequência L é formada por 2 imagens digitais discretas, isto é, $L = (I(s, 1), I(s, 2))_{s \in S}$;
- o número de detectores está fixo em $K = 4$;

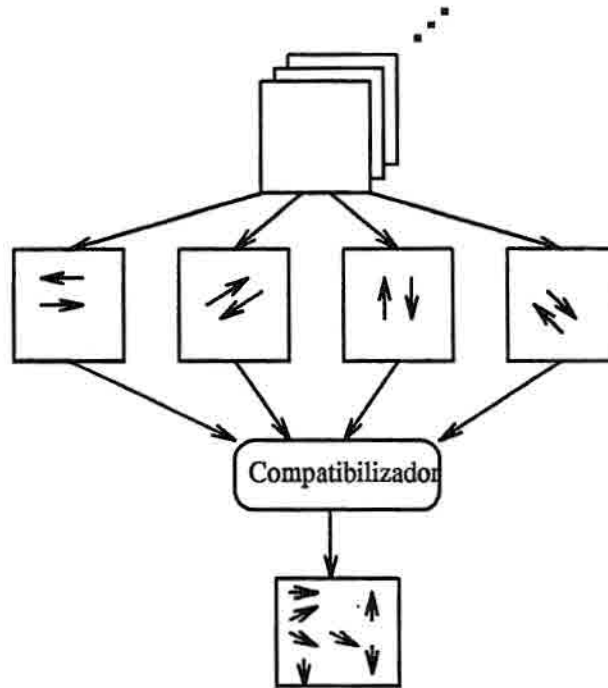


Fig. 4.2 - Compatibilizador de direções.

4.3 Detectores locais

Pretende-se estudar o comportamento do detector biológico do tipo detector de Reichardt (40) e a versão proposta por Van Santem e Sperling (o Detector Elaborado de Reichardt —ERD—) (41). Empregaram-se esses detectores em sua forma generalizada tal como foi descrita no Capítulo 2. Aplicam-se quatro detectores, seletivos aos sentidos já mencionados, em cada ponto de uma imagem da sequência.

Essa organização pode ser vista como o processamento de uma rede neural artificial onde cada nó representa uma célula capaz de processar informação de movimento. Cada nó executa a mesma função, tem o mesmo padrão de conectividade com os outros nós e os mesmos pesos associados fixos *a priori*. Esse tipo de rede neural será chamada, aproveitando a nomenclatura usada em (26), de **rede implícita**. Do ponto de vista neurofisiológico, as redes implícitas podem ser vistas como modelos as redes neurais naturais responsáveis pelos fenômenos per-

ceptivos, e cuja arquitetura foi desenvolvida pela natureza durante a evolução das espécies.

Os campos receptivos adotados para o ERD, ou seja o contexto espacial, foram a média das intensidades numa vizinhança $v(s)$ para todo ponto $s \in S$. Esse campo receptivo foi usado por ser de fácil implementação e ser capaz de capturar uma informação de contexto espacial em cada ponto. Nota-se que, como no caso das imagens de nuvens não existem objetos rígidos, esse campo receptivo pode ser bem representativo do contexto em um ponto, supondo-se que as deformações são capturadas dentro dessa vizinhança. O uso desses campos receptivos pretende amenizar o efeito nocivo que a presença de ruído pode acarretar na detecção de movimento.

Assim, seja L a sequência de imagens de entrada. Após a aplicação dos quatro detectores obtém-se as imagens de direção X_0, X_1, X_2, X_3 com $X_k = (X_k(s))_{s \in S, 0 \leq k < 4}$.

Como foi mencionado no Capítulo 3, é necessário converter a saídas dos detectores em números associados às chances de ter-se observado movimento em cada um dos sentidos possíveis.

A seguir serão propostas algumas funções que permitem realizar essa conversão. Dado que a discussão a seguir é válida para cada $s \in S$, esse índice será omitido por simplicidade. Seja $V: \mathbb{R}^4 \times \{0, \dots, 3\} \rightarrow [0, 1]$ onde, por definição, $V(\cdot, \ell) = V_\ell(\cdot) = (V_\ell(s))_{s \in S, 0 \leq \ell < 3}$ e além disso, V satisfaz as propriedades **P1-P3**, propostas na Seção 2.2.3. Para tanto, por simplicidade e sem perda de generalidade, suponha-se que exista pelo menos um $0 \leq k < 4$ tal que $X_k(s) \neq 0$, e considere-se as seguintes funções:

- Regra do Maior Valor: essa regra atribui probabilidade unitária à direção cujo detector dá máxima saída, em valor absoluto, e ao sentido correspondente ao sinal dessa saída. Mais formalmente, V_ℓ para $0 \leq \ell < 8$ é dada por:

1) para $0 \leq \ell < 4$

$$V_\ell = \begin{cases} 1 & \text{se } |X_\ell| > |X_m| \forall \ell \neq m \in \{0, \dots, 3\} \text{ e } X_\ell > 0 \\ 0 & \text{caso contrário,} \end{cases} \quad (4.1)$$

2) para $4 \leq \ell < 7$

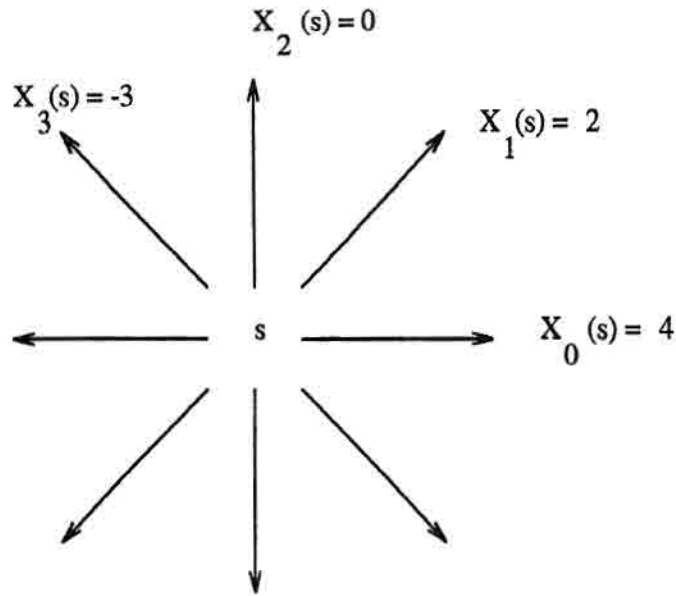
$$V_\ell = \begin{cases} 1 & \text{se } |X_{\ell \bmod 4}| > |X_{m \bmod 4}| \forall \ell \neq m \in \{4, \dots, 7\} \text{ e } X_\ell < 0 \\ 0 & \text{caso contrário.} \end{cases} \quad (4.2)$$

Por exemplo na Figura 4.3 estão mostrados alguns valores de X_0, X_1, X_2, X_3 e os valores de V_ℓ correspondentes à aplicação da Regra do Maior Valor.

- Regra da Padronização: esta regra atribui probabilidade nula às direções associadas aos detectores de saída nula, e probabilidades proporcionais aos valores de saídas dos detectores das outras direções, levando em conta o sinal da saída. Formalmente, V_ℓ é dada por:

$$V_\ell = \begin{cases} Z^{-1} X_\ell & \text{se } X_\ell > 0, \forall 0 \leq \ell < 4 \\ Z^{-1} |X_{\ell \bmod 4}| & \text{se } X_{\ell \bmod 4} < 0, \forall 4 \leq \ell < 7 \\ 0 & \text{caso contrário,} \end{cases} \quad (4.3)$$

e $Z = \sum_{\ell=0}^3 |X_\ell|$. Por exemplo, na Figura 4.4 abaixo, estão mostrados alguns valores de X_0, X_1, X_2, X_3 e os valores de V_ℓ correspondentes a aplicação da Regra de Padronização.



$$V_{(0)} = 1, \quad V_{(1)} = V_{(2)} = V_{(3)} = V_{(4)} = V_{(5)} = V_{(7)} = 0$$

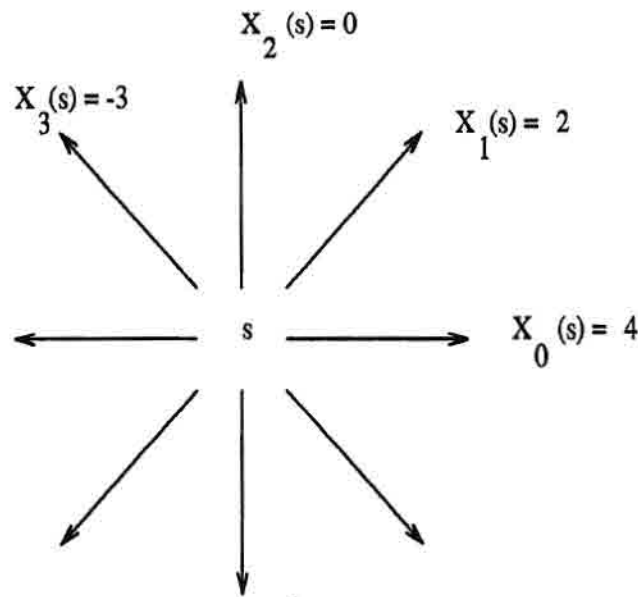
Fig. 4.3 - Exemplo de aplicação da Regra do Maior Valor

Considerando-se os valores de probabilidade relativas de se obter movimento em cada sentido, na próxima Seção discute-se o problema de compatibilizar esses valores em uma solução global.

4.4 Compatibilização

Como já é previsto no ERD (41), propõe-se nesse trabalho estudar uma função de compatibilização para as saídas dos detectores espalhados pelo campo visual, ou seja, pelos pontos da imagem.

Para tanto, deve-se procurar uma função F que compatibilize as saídas X_ℓ em um único campo de deslocamento $\widehat{X} = ([\widehat{X}(s)]_{s \in S})$, isto é, $\widehat{X} = F(X_0, \dots, X_{K-1})$. Dependendo da forma da função F será a complexidade de se achar a solução \widehat{X} . Por analogia com o problema da segmentação Bayesiana de imagens (14) essa função pode ser:



$$V_{(0)} = 4/9, V_{(1)} = 2/9, \quad V_{(2)} = V_{(3)} = V_{(4)} = V_{(5)} = 0, V_{(7)} = 3/9$$

Fig. 4.4 - Exemplo de aplicação da Regra da Padronização

- **estritamente pontual:** onde a direção final encontrada em cada ponto depende das saídas dos detectores em cada direção apenas nesse ponto. A função F pode ser decomposta em $\#S$ funções, sendo que cada uma depende apenas dos valores na coordenada correspondente, isto é, $F = (F_0, \dots, F_{\#S-1})$ e, no ponto s , $F_s(X_0, \dots, X_3) = F_s(X_0(s), \dots, X_3(s))$.
- **sensíveis à vizinhança:** são aquelas onde não é possível apresentar uma decomposição da função F como no caso anterior. Isto é, na coordenada s , a solução $\widehat{X}(s)$, depende de alguma maneira de X_0, X_1, X_2, X_3 em alguma vizinhança do ponto s . Esse tipo de solução é desejável porque permite incorporar a noção de suavidade espacial. A desvantagem é que, em geral, podem levar a problemas de alta complexidade do tipo problemas combinatorios (14,31,38).

A seguir serão apresentados exemplos desses dois tipos de função compatibilizadora. Como função estritamente pontual, considere a função F onde

a direção escolhida em cada ponto da imagem é dada pelo detector que tenha apresentado maior valor de saída naquele ponto. Ou seja, $X(s) = F_s(X_0, X_1, X_2, X_3)$ para todo $s \in S$ e:

$$F_s(X_0(s), X_1(s), X_2(s), X_3(s)) = \begin{cases} k & \text{se } |X_k(s)| > |X_\ell(s)| \forall \ell \neq k \text{ e } X_k(s) > 0 \\ k + 4 & \text{se } |X_k(s)| > |X_\ell(s)| \forall \ell \neq k \text{ e } X_k(s) < 0. \end{cases} \quad (4.4)$$

Esta função leva em conta o fato de que, nos detectores utilizados, o sinal da saída é um indicador de evidência de movimento em um dos sentidos associados a essa direção.

Serão vistos agora dois exemplos simples de funções sensíveis à vizinhança. Primeiro considere-se a função F onde o valor de direção em todos os pontos da imagem é dada pelo detector que tenha o maior valor de saída entre todos os pontos. Ou seja, $X(s) = F(X_0, X_1, X_2, X_3)$ para todo $s \in S$ e:

$$F(X_0, X_1, X_2, X_3) = \begin{cases} k & \text{se } |X_k(s)| \geq |X_\ell(t)| \forall \ell \neq k, \forall t \neq s \text{ e } X_k(s) > 0 \\ k + 4 & \text{se } |X_k(s)| \geq |X_\ell(t)| \forall \ell \neq k, \forall t \neq s \text{ e } X_k(s) < 0 \end{cases} \quad (4.5)$$

Agora considere a função F onde o valor de direção em todos os pontos da imagem é dada pelo detector que tenha o maior número de pontos onde seu valor de saída seja maior que os dos outros detectores. Ou seja, $X(s) = F(X_0, X_1, X_2, X_3) = \text{máx}_n^{-1}(C_n)$ para todo $s \in S$, para todo $0 \leq n \leq 7$ e onde C_n é dado por :

$$C_n = \begin{cases} \#\{s \in S : |X_k(s)| \geq |X_\ell(s)| \forall \ell \neq k \text{ e } X_k(s) > 0\} & \text{para } k \leq 3 \\ \#\{s \in S : |X_k(s)| \geq |X_\ell(s)| \forall \ell \neq k \text{ e } X_k(s) < 0\} & \text{para } k > 3. \end{cases} \quad (4.6)$$

A solução para esses dois tipos de funções (estritamente pontual e sensíveis à vizinhança) são pouco interessantes porque ambas atribuem a cada posição a mesma direção de movimento. Mesmo no caso de uma sequência de imagens onde exista um único objeto movimentando-se em uma direção, a solução estaria errada para os pontos que estão fora do objeto, ou seja, o fundo. Em um caso mais geral, como o tratado nesse trabalho, onde cada ponto da imagem está sujeito a deslocamentos em diferentes direções, essas soluções são claramente inadequadas.

Propõe-se, então, estudar uma solução capaz de atribuir sentidos de deslocamento diferentes a pontos diferentes na imagem. Isto é desejável para distinguir pontos onde não há movimento de pontos onde há; por outro lado, deseja-se reter parte da suavidade obtida pelas soluções sensíveis a vizinhança para minimizar o efeito de ruídos na solução \widehat{X} . Assim sendo deseja-se uma função F que retenha as boas características das funções globais e locais.

Sejam $x, y \in \mathbb{R}^2$ e seja $\langle x, y \rangle$ o produto escalar de x e y . Sejam V_0, \dots, V_7 as probabilidades associadas aos oito sentidos encontradas de acordo com as saídas dos quatro detectores aplicados. Esses resultados serão usados para construir Y_0, \dots, Y_7 , onde $Y_\ell = (Y_\ell(s))_{s \in S, 0 \leq \ell \leq 7}$ e $Y_\ell(s) \in \{(x, y) \in \mathbb{R}^2: x^2 + y^2 \leq 1\}$, da seguinte forma:

$$Y_\ell = V_\ell \cdot \vec{\ell},$$

onde cada vetor unitário no sentido ℓ , denotado por $\vec{\ell}$, é definido por

$$\vec{\ell} = (\cos \ell\pi/4, \text{sen} \ell\pi/4)$$

para todo $0 \leq \ell \leq 7$, e $Y_\ell(s) = (Y_{\ell 1}(s), Y_{\ell 2}(s))$.

Existe um número finito de soluções dada a natureza discreta do problema em questão. Cada solução será denotada $\widehat{X} \in \Xi = \{\vec{0}, \dots, \vec{7}\}^S$.

Como pode ser visto em (21,38) muitos problemas relacionados aos processos que recuperam as propriedades físicas tri-dimensionais de objetos imageados, a partir de suas imagens bidimensionais, podem ser formulados em termos de minimização de uma função custo. Exemplos desses problemas são: detecção de bordas, interpolação de superfícies ou detecção de movimento. E esse será o enfoque dado a esse trabalho. Ou seja, será conveniente associar a cada solução possível um número real que é chamado **energia**. Essa energia será diretamente relacionada à qualidade da solução global sendo que quanto menor a energia “melhor” será considerada a solução. A formulação do problema em termos de minimização de energia traz vantagens como, entre outras, a de ser uma forma elegante de se representar restrições ou o conhecimento *a priori* do mundo. A seguir, apresentam-se duas formas de especificar a energia na obtenção de soluções globais.

4.4.1 Modelo $x - y$

Nos últimos anos têm-se empregado modelos provindos da mecânica estatística em processamento de imagens. Por exemplo, na suavização (7,8) e filtragem de imagens (38).

Um dos modelos mais populares é o modelo de Pott-Strauss. Este modelo calcula a energia de uma configuração somando, sobre todos os pares de vizinhos, o valor “1” se os vizinhos do par tem o mesmo valor, e “0” caso contrário. Essa função dos valores de coordenadas vizinhos é chamada de potencial. E quanto mais potenciais positivos houver menor será a energia da configuração. Dessa maneira o modelo de Pott-Strauss favorece configurações onde pares de vizinhos tenham o mesmo valor, fazendo com que essas configurações tenham baixo valor de energia (14).

Como já foi mencionado no Capítulo 2, o problema da compatibilização consiste em procurar configurações dos detectores associados a cada ponto, mais suaves que as fornecidas apenas pelos detectores locais. O uso do modelo de Pott-Strauss não seria completamente satisfatório, pois não leva em conta a orientação espacial, que é uma característica essencial do problema em questão.

Contudo, existe um outro modelo da mecânica estatística que incorpora as noções de dependência espacial, comum ao modelo de Pott-Strauss e de valores vetoriais em cada coordenada, requerida para o problema da detecção. Esse modelo é chamado **modelo $x - y$** . O modelo $x - y$ sem campo externo é definido, na notação de (24), como:

$$E = -1/2 \sum_{i,j} J_{i,j} (s_i \cdot s_j) \quad (4.7)$$

onde s_j é um vetor em \mathbb{R}^2 , “ \cdot ” denota o produto escalar entre vetores, e $J_{i,j}$ é definido por:

$$J_{i,j} = \begin{cases} J & \text{se } i \text{ e } j \text{ são vizinhos} \\ 0 & \text{caso contrário,} \end{cases}$$

com $J \in \mathbb{R}$.

Em analogia ao que se faz com o modelo Pott-Strauss sem campo externo para adaptá-lo ao problema da segmentação (14), à equação (4.7) é acrescentado um campo externo. A energia de uma solução, $E : \underbrace{\mathbb{R}^S \times \dots \times \mathbb{R}^S}_8 \times \Xi \rightarrow \mathbb{R}$, leva em conta os dados de entrada (Y_0, \dots, Y_7) e pode-se escreve-la como:

$$E_\lambda(\widehat{X}; Y_0, \dots, Y_7) = -\lambda E_1(\widehat{X}) - (1 - \lambda) E_2(\widehat{X}; Y_0, \dots, Y_7)$$

com $\lambda \in [0, 1]$, onde:

- E_1 representa a suavidade espacial, isto é, $E_1(\widehat{X})$ será mínimo quando $\widehat{X}(s) = \vec{\ell}$ para todo $s \in S$ e qualquer $\vec{\ell}$. Esse termo, muitas vezes chamado de **energia livre da configuração \widehat{X}** , serve para amenizar o efeito nocivo que as alterações dos objetos podem produzir na solução. A especificação de E_1 deve capaz de favorecer soluções que apresentem um certo grau de alinhamento em relação a uma direção qualquer.
- E_2 representa a adequação da solução aos dados de entrada, isto é, quanto mais E é dependente dos Y_i , $0 \leq i < 8$, dados de entrada (e portanto mais independente da suavidade espacial) menor será E_2 .

O parâmetro $\lambda \in [0, 1]$ controla o peso relativo dos termos E_1 e E_2 no cálculo da energia total E_λ da solução. Assim sendo, para $\lambda = 0$, $E_\lambda(\widehat{X}; Y_0, \dots, Y_7) = -E_2(\widehat{X}; Y_0, \dots, Y_7)$, ou seja, a solução leva em conta apenas os dados de entrada, desconsiderando por completo a suavidade espacial. Por outro lado, para $\lambda = 1$, $E_\lambda(\widehat{X}; Y_0, \dots, Y_7) = -E_1(\widehat{X})$, ou seja, a solução leva em conta apenas a suavidade espacial, desconsiderando por completo os dados de entrada.

A escolha do parâmetro λ influi diretamente na solução. Para os casos extremos mencionados acima ($\lambda = 0, \lambda = 1$) é imediato obter configurações \widehat{X} que minimizam a energia E_λ . Porém, para $\lambda \in (0, 1)$, não se conhecem técnicas diretas para obtenção da solução \widehat{X} que minimize E_λ , pois o problema de achar tal solução é um problema combinatório da classe dos NP-completos (31), (ver (7,8,15,16) para problemas análogos em processamento de imagens).

Será apresentada agora uma proposta de especificação de E_λ de

acordo com as características mostradas acima. A energia livre E_1 é dada por:

$$E_1(\widehat{X}) = \sum_{\langle\langle s,t \rangle\rangle} \langle \widehat{X}(s), \widehat{X}(t) \rangle \quad \text{onde } \langle\langle s,t \rangle\rangle = \{(s,t) \in S \times S : t \in v(s)\}, \quad (4.8)$$

e $v(s)$ determina uma vizinhança de interesse para a coordenada s , $\widehat{X}(s) \in \{\vec{0}, \dots, \vec{7}\}$.

O termo E_2 , a energia de interação, é dada por:

$$E_2(\widehat{X}; Y_0, \dots, Y_7) = \sum_{s \in S} |Y_{\widehat{X}'(s)}(s)|, \quad (4.9)$$

onde $\widehat{X}'(s) \in \{0, \dots, 7\}$ denota o sentido de $\widehat{X}(s)$.

A seguir apresenta-se um algoritmo capaz de gerar sucessivas configurações válidas com energia não crescente e passível de ser implementada numa máquina sequencial. Esse algoritmo é semelhante ao algoritmo ICM apresentado em (14), para a segmentação de imagens de radar de abertura sintética.

- 1) Fixar $\lambda_0 = 0$ e obter \widehat{X}_0 tal que $E_{\lambda_0}(\widehat{X}_0; Y_0, \dots, Y_7)$ seja mínima. Atribuir $t = 1$; fixar $\lambda \in (0, 1]$
- 2) atualizar para toda coordenada (em uma ordem a ser especificada para cada t) o seu valor de forma que \widehat{X}_t seja substituído por:

$$\widehat{X}_{t+1}(s) = \max_{\vec{\ell} \in \{\vec{0}, \dots, \vec{7}\}}^{-1} \left[\lambda \sum_{u \in v(s)} \langle \vec{\ell}, \widehat{X}_t(u) \rangle + (1 - \lambda) \langle \vec{\ell}, Y_t(s) \rangle \right]; \quad (4.10)$$

- 3) enquanto algum critério de parada não for satisfeito voltar ao passo 2 com $t = t + 1$;
- 4) retornar \widehat{X}_t .

Embora esse algoritmo não tenha sido usado diretamente no presente trabalho, dado que o enfoque proposto é o de usar o paradigma da neurocomputação, serviu como base para a especificação de uma rede do tipo Hopfield que se pretende empregar com o mesmo propósito. A especificação dessa rede será vista a seguir.

4.4.2 Rede de Hopfield

Uma rede de Hopfield, que será analisada aqui sob o ponto de vista meramente digital, é composta de $N \geq 2$ nós. Cada nó tem um valor de saída $\eta_i \in [0, 1]$. Os nós da rede estão totalmente interconectados, isto é, a cada par de nós (i, j) está associado um valor $T_{i,j}$ que representa o peso da conexão entre o nó i e o nó j . Para todo $t \geq 1$, com $t \in \mathbb{N}$, o estado do nó i será dado por:

$$\eta_i(t) = f\left(\sum_{0 \leq j \leq N} T_{i,j} \cdot \eta_j(t-1)\right),$$

onde f é, neste trabalho, da classe das funções do tipo sigmóides dadas por:

$$f(x) = 1/(1 + e^{-\alpha x}),$$

em que $0 < f(x) < 1$ para todo $x \in \mathbb{R}$, e α é uma constante real positiva que controla a “inclinação” da função. Quando α cresce a função f aproxima-se da função do tipo “degrau unitário”.

Os autores de (18) mostram que as saídas dos nós da rede para o caso de conexões simétricas $T_{i,j} = T_{j,i}$, sempre converge assintoticamente para estados estáveis. Ou seja, estados onde não há mais alteração no valor das saídas. Além disso, quando f aproxima-se assintoticamente da função “degrau unitário”, os estados estáveis da rede são mínimos locais da seguinte função:

$$E = -1/2 \sum_{i=1}^N \sum_{j=1}^N T_{i,j} \eta_i \eta_j - \sum_{i=1}^N \eta_i I_i, \quad (4.11)$$

onde I_i é uma entrada externa do nó i . Uma discussão sobre as propriedades de convergência do modelo de rede de Hopfield pode ser vista em (20).

Redes neurais com essa configuração básica podem ser usadas para calcular soluções de problemas específicos em otimização combinatória, através da escolha adequada dos valores das conexões $T_{i,j}$ e das entradas externas $I_{i,j}$, os quais representarão, apropriadamente, a função a ser minimizada. Em (18) é apresentada uma configuração de rede de Hopfield para a solução do problema do caixeiro viajante.

Considere-se agora o problema da compatibilização das saídas locais dos quatro detectores de movimento em cada coordenada $s \in S$. Visto que esse é um problema combinatório, podem ser obtidas soluções aproximadas através da minimização de energia por uma rede de Hopfield.

A rede terá N nós, onde $N = \#S \cdot 8$, arranjados em forma de matriz e referenciados como (s, ℓ) com $s \in S$ e $0 \leq \ell \leq 7$. A função f será dada por:

$$f(x) = 1/2(1 + \tanh(x/u_0)), \quad (4.12)$$

onde o parâmetro u_0 define a aproximação da função f á função do tipo “degrau unitário”.

O valor de saída de cada nó (s, ℓ) , no instante t , é dado por $\eta_{s,\ell}(t)$ para $s \in S$ e $0 \leq \ell \leq 7$. Os nós (s, \cdot) são associados a cada sentido possível de deslocamento na coordenada $s \in S$. Os nós (\cdot, ℓ) são associados a cada coordenada que pode apresentar deslocamento no sentido $0 \leq \ell \leq 7$.

Assim, considere-se o exemplo onde $S = \{0, 1, 2\}$ e $0 \leq \ell \leq 3$. A rede teria então $N = 12$ nós. Dispostos na forma de matriz, uma possível configuração dos estados dos nós da rede pode ser vista na Figura 4.5. Assim os elementos da matriz podem ser compreendidos como: na coordenada $s = 0$ o vetor deslocamento atribuído pelo modelo é $\vec{1}$, na coordenada $s = 1$ o vetor $\vec{0}$ e, na coordenada $s = 2$ o vetor $\vec{3}$.

$s \setminus \ell$	0	1	2	3
0	0	1	0	0
1	1	0	0	0
2	0	0	0	1

Fig. 4.5 - Exemplo de configuração de uma rede de Hopfield para o problema da compatibilização.

Denota-se por $\mathcal{R} = \{0, \dots, \#S - 1\} \times \{0, \dots, 7\}$ o conjunto dos índices dos N nós da rede. Para que fique completamente especificada se faz necessário atribuir um valor $T_{s,\ell,s',\ell'}$ para cada $((s, \ell), (s', \ell')) \in \mathcal{R} \times \mathcal{R}$.

Seja $\eta(0) = (\eta_{s,\ell}(0))_{s \in S, 0 \leq \ell \leq 7}$ a configuração inicial. Para todo $t \geq 1$ com $t \in \mathbb{N}$, o estado do nó (s, ℓ) será atualizado em função do estado dos outros nós e dos pesos associados às conexões entre eles. Isto é:

$$\eta_{s,\ell}(t+1) = f \left(\sum_{(s',\ell') \in \mathcal{R}} T_{s,\ell,s',\ell'} \cdot \eta_{s',\ell'}(t) \right), \quad (4.13)$$

ou seja, a cada instante $t > 0$ são efetuadas as seguintes operações:

- 1) é escolhido o nó (s, ℓ) a ser visitado. Esta escolha pode ser feita de forma determinística ou, tipicamente, aleatória;
- 2) é calculado o valor de $\sum_{(s',\ell') \in \mathcal{R}} T_{s,\ell,s',\ell'} \cdot \eta_{s',\ell'}(t)$;
- 3) é calculado o valor de $f(\cdot)$;
- 4) o valor $f(\cdot)$ é atribuído a saída do nó (s, ℓ) ;
- 5) enquanto algum critério de parada não for satisfeito, atribui-se $t = t + 1$ e volta-se ao passo 1;
- 6) retorna-se $\eta_{s,\ell}(t)$.

Para que os valores das saídas dos N neurônios da rede descrita acima forneçam uma solução para o problema da compatibilização, essa deve estar associada a uma função de energia do tipo dado pela equação (4.11), onde o estado de mais baixa energia corresponda à melhor solução global. Essa função energia deve favorecer configurações de estados de saída dos nós onde valem as seguintes restrições:

- 1) a configuração de estados seja válida, isto é, como no exemplo mostrado na Figura 4.5, a cada coordenada deseja-se atribuir um único vetor deslocamento;
- 2) deve-se levar em conta os dados locais, isto é, favorecer configurações que em cada coordenada apresentem o vetor deslocamento com maior chance relativa;

- 3) a cada coordenada deve ser atribuído algum sentido;
- 4) os vetores de deslocamento na imagem possuem suavidade espacial em uma vizinhança considerada.

O termo **suavidade espacial** quer dizer, nesse trabalho, que na vizinhança considerada, supõe-se que os deslocamentos entre os pontos vizinhos sejam o mais similares possíveis. Como exemplo considere as Figuras 4.6 e 4.7, na primeira considera-se que existe mais suavidade espacial entre todos os pontos. Esse conceito foi incorporado pelo uso da operação produto escalar entre dois vetores deslocamento unitários. Na Tabela 4.1 mostra-se como essa operação retorna valores cuja magnitude está diretamente ligada à similaridade entre dois vetores deslocamento. Note-se que o produto escalar é invariante em rotação, portanto, estão mostrados apenas os cinco casos que, rotacionados, fornecem todas as combinações possíveis.

TABELA 4.1 -
PRODUTO ESCALAR ENTRE DOIS VETORES UNITÁRIOS.

	→	↗	↑	↖	←
→	1	$\frac{\sqrt{2}}{2}$	0	$-\frac{\sqrt{2}}{2}$	1
↗	$\frac{\sqrt{2}}{2}$	1	$\frac{\sqrt{2}}{2}$	0	$-\frac{\sqrt{2}}{2}$
↑	0	$\frac{\sqrt{2}}{2}$	1	$\frac{\sqrt{2}}{2}$	0
↖	$-\frac{\sqrt{2}}{2}$	0	$\frac{\sqrt{2}}{2}$	1	$\frac{\sqrt{2}}{2}$
←	1	$-\frac{\sqrt{2}}{2}$	0	$\frac{\sqrt{2}}{2}$	1

Levando-se em conta essas restrições propõe-se a seguinte função energia:

$$\begin{aligned}
 E(\widehat{X}; Y_0, \dots, Y_7) = & A \sum_{s \in S} \sum_{\ell=0}^7 \sum_{\ell' \neq \ell} \eta_{s,\ell} \cdot \eta_{s,\ell'} - \\
 & B \sum_{s \in S} \sum_{s' \in v(s)} \sum_{\ell=0}^7 \sum_{\ell'=0}^7 \langle \vec{\ell} \cdot \eta_{s,\ell}, \vec{\ell}' \cdot \eta_{s',\ell'} \rangle - C \sum_{s \in S} \sum_{\ell=0}^7 |Y_\ell(s)| \cdot \eta_{s,\ell} + \\
 & D \left(\sum_{s \in S} \sum_{\ell=0}^7 \eta_{s,\ell} - \#S \right)^2, \tag{4.14}
 \end{aligned}$$

onde $Y_\ell(s)$ e $\langle \cdot, \cdot \rangle$ são como definidos na Seção 4.4, além disso

$$|Y_\ell(s)| = \sqrt{Y_{\ell 1}(s)^2 + Y_{\ell 2}(s)^2}.$$

O primeiro termo de E refere-se à restrição 1, ou seja, será tanto maior quanto mais indicações de deslocamentos existirem para uma única coordenada. O segundo termo está associado a restrição 4, isto é, será tanto menor quanto mais suavidade espacial houver em uma dada vizinhança $v(s)$. O terceiro termo está associado à restrição 3, isto é, será tanto menor quanto maior for a probabilidade associada ao sentido do vetor deslocamento em cada coordenada. O quarto termo está associado à restrição de que em cada coordenada deve haver pelo menos um vetor deslocamento associado. As constantes A, B, C e D definem o peso que será dado a cada uma dessas restrições.

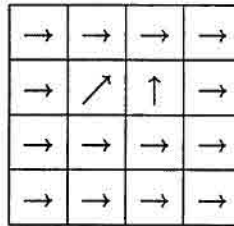


Fig. 4.6 - Exemplo de suavidade espacial.

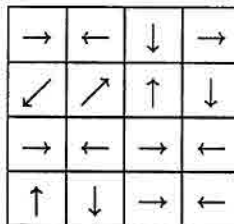


Fig. 4.7 - Exemplo de falta de suavidade espacial.

Denotar-se-á por $\mathbb{I}_A(x)$ a função característica de um conjunto A , isto é

$$\mathbb{I}_A(x) = \begin{cases} 1 & \text{se } x \in A \\ 0 & \text{caso contrário.} \end{cases}$$

Os valores de $T_{s,\ell,s',\ell'}$ são especificados, com base nessa função de energia, da seguinte maneira:

$$T_{s,\ell,s',\ell'} = -A \cdot \mathbb{I}_{\{s\}}(s') + B \cdot \mathbb{I}_{v(s)}(s') \langle \vec{\ell}, \vec{\ell}' \rangle + C \cdot \mathbb{I}_{\{s\}}(s') \mathbb{I}_{\{\ell\}}(\ell') |y_\ell(s)| + D \quad (4.15)$$

Em (17) prova-se que se um problema pode ser caracterizado através de uma função de energia, e essa função de energia dá origem a conexões simétricas, então basta achar um algoritmo que produza sucessivas configurações com energias não crescentes. Considerando-se que a dinâmica de uma rede de Hopfield convenientemente especificada obedece essa exigência, basta modelar o problema em questão em termos de uma função de energia que implique em conexões simétricas para que os seus estados estáveis sejam mínimos locais da função de energia envolvida, conforme especificado para a modelagem proposta para o problema em questão.

Para a especificação de $T_{s,\ell,s',\ell'}$ dada na equação (4.15) vale que $T_{s,\ell,s',\ell'} = T_{s',\ell',s,\ell}$ e garante-se, portanto, que a evolução das configurações de estados das saídas dos nós da rede, segundo o algoritmo mostrado na Seção 4.4.2, converge para um estado estável e de mínimo local da função energia dada pela equação (4.14). Ou seja, uma possível e boa solução para o problema.

Um estado estável é considerado uma solução caso obedeça a restrição de que a cada ponto está associado um único vetor deslocamento.

O parâmetro u_0 foi considerado fixo e igual a 0.01.

No Capítulo seguinte apresentam-se alguns testes do modelo proposto.

CAPÍTULO 5

TESTES DO MODELO

Nesse capítulo serão mostradas alguns resultados referentes ao modelo de detecção proposto. Serão apresentados inicialmente testes com imagens representando objetos rígidos e com movimento conhecido, e, posteriormente, testes com imagens reais e com movimentos naturais.

5.1 Testes do modelo

5.1.1 Imagens sintéticas com objetos rígidos

Um estudo foi feito com duas sequências de imagens sintéticas, L_{r0} e L_{r1} , que contém objetos rígidos.

A sequência $L_{r0} = (I_{r0}(s, 1), I_{r0}(s, 2))$, com $s \in S = \{0, \dots, 9\} \times \{0, \dots, 9\}$ representa o movimento de um objeto rígido, que se movimenta entre os instantes 1 e 2 um pixel para a esquerda. As imagens I_{r0} e I_{r1} estão mostradas, forma de matriz de intensidades radiométricas, nas Figuras 5.1 e 5.2. A imagem $I_{r0}(s, 2)$ é o resultado de se aplicar um deslocamento arbitrário a todos os pontos da imagem $I_{r0}(s, 1)$ na direção horizontal sentido para a esquerda.

As imagens de direção para a sequência L_{r0} , para os detectores X_0, X_1, X_2, X_3 , podem ser vistas nas Figuras 5.3, 5.4, 5.5, 5.6, respectivamente.

O resultado da detecção por detector de Reichardt, sem campo receptivo, e compatibilização por Regra do Máximo, na sequência L_{r0} , é mostrado na Figura 5.7. Nas figuras a seguir, os pontos onde nenhum dos detectores seletivos à direção sinalizou movimento serão representados por espaços em branco.

O resultado da detecção por detector de Reichardt e compatibilização por Rede de Hopfield, na sequência L_{r0} , é mostrado na Figura 5.8.

$$I_{r0}(s, 1) =$$

0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0
0	0	0	220	220	220	220	0	0	0
0	0	0	220	220	220	220	0	0	0
0	0	0	220	220	220	220	0	0	0
0	0	0	220	220	220	220	0	0	0
0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0

Fig. 5.1 - Imagem $I_{r0}(s, 1)$ com objeto rígido.
$$I_{r0}(s, 2) =$$

0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0
0	0	220	220	220	220	0	0	0	0
0	0	220	220	220	220	0	0	0	0
0	0	220	220	220	220	0	0	0	0
0	0	220	220	220	220	0	0	0	0
0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0

Fig. 5.2 - Imagem $I_{r0}(s, 2)$ com objeto rígido e movimento arbitrário para a esquerda.

$$X_0 =$$

0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	-1.0	0.0	0.0	-1.0	0.0	0.0	0.0
0.0	0.0	0.0	-0.5	0.0	0.0	-0.5	0.0	0.0	0.0
0.0	0.0	0.0	-0.5	0.0	0.0	-0.5	0.0	0.0	0.0
0.0	0.0	0.0	-0.5	0.0	0.0	-0.5	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

Fig. 5.3 - Imagem de direção X_0 da sequência I_{r_0} .
$$X_1 =$$

0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	-0.5	0.0	0.0	-0.5	0.0	0.0	0.0
0.0	0.0	0.0	-0.5	0.0	0.0	-0.5	0.0	0.0	0.0
0.0	0.0	0.0	-0.5	0.0	0.0	-0.5	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

Fig. 5.4 - Imagem de direção X_1 da sequência I_{r_0} .

$$X_2 =$$

0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

Fig. 5.5 - Imagem de direção X_2 da sequência I_{r0} .
$$X_3 =$$

0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	1.0	0.0	1.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	1.0	0.0	1.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	1.0	0.0	1.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

Fig. 5.6 - Imagem de direção X_3 da sequência I_{r0} .

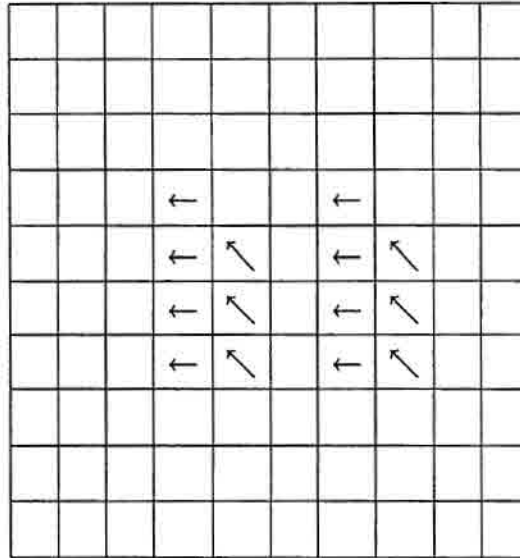


Fig. 5.7 - Resultado da detecção com compatibilização por Regra do Máximo na sequência L_{r0} .

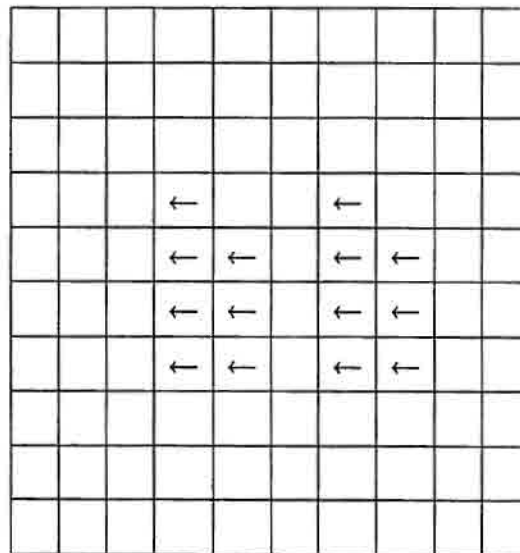


Fig. 5.8 - Resultado da detecção com compatibilização por Rede de Hopfield na sequência L_{r0} .

A sequência $L_{r1} = (I_{r1}(s, 1), I_{r1}(s, 2))$, com $s \in S = \{0, \dots, 9\} \times \{0, \dots, 9\}$, representa o movimento de um objeto rígido, que se movimenta entre os instantes 1 e 2 um pixel na direção quarenta e cinco graus, sentido para cima. Essas imagens estão mostradas, na forma de matriz de intensidades radiométricas, nas Figuras 5.9 e 5.10. A imagem $I_{r1}(s, 2)$ é o resultado de se aplicar um deslocamento arbitrário em todos os pontos na direção de quarenta e cinco graus, no sentido para cima, nos pontos da imagem $I_{r1}(s, 1)$.

$$I_{r1}(s, 1) =$$

0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0
0	0	0	220	250	220	220	0	0	0
0	0	0	220	250	220	220	0	0	0
0	0	0	220	250	220	220	0	0	0
0	0	0	220	250	220	220	0	0	0
0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0

Fig. 5.9 - Imagem $I_{r1}(s, 1)$ com objeto rígido.

O resultado da detecção por detector de Reichardt e compatibilização por Regra do Máximo, na sequência L_{r1} é mostrado na Figura 5.11.

O resultado da detecção por detector de Reichardt e compatibilização por Rede de Hopfield, na sequência L_{r1} é mostrado na Figura 5.12.

$$I_{r1}(s, 2) =$$

0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0
0	0	0	0	220	250	220	220	0	0
0	0	0	0	220	250	220	220	0	0
0	0	0	0	220	250	220	220	0	0
0	0	0	0	220	250	220	220	0	0
0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0

Fig. 5.10 - Imagem $I_{r1}(s, 2)$ com objeto rígido e movimento arbitrário de 45 graus para cima.

			→	↗	↑	↖			
			→	↗	↗	→	→		
			→	↗	↗	→	→		
			↗	↗	↑	↖			

Fig. 5.11 - Resultado da detecção com compatibilização por Regra do Máximo na sequência L_{r1} .

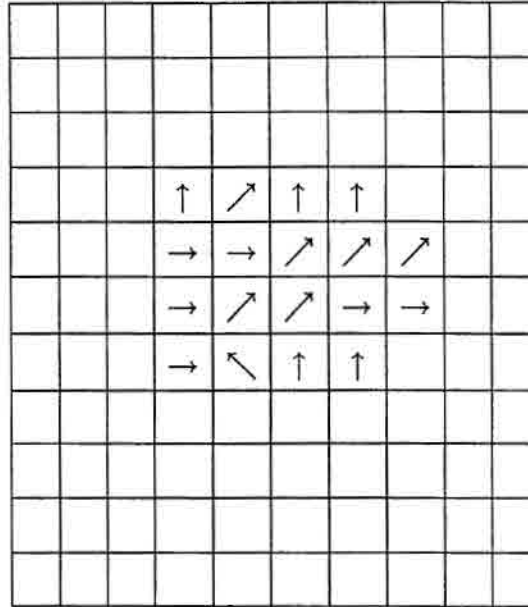


Fig. 5.12 - Resultado da detecção com compatibilização por Rede de Hopfield na sequência L_{r1} .

5.1.2 Imagens reais

Um estudo foi feito com três sequências de imagens L_1 , L_2 e L_3 , formadas, cada uma, por duas imagens, em que testes foram efetuados com e sem campo receptivos. A primeira das duas imagens é a mesma para as três sequências e será chamada de imagem base e denotada $I_{base}(s, 1)$. A imagem $I_{base}(s, 1)$ é uma parte de uma imagem meteorológica real tomada a partir do satélite METEOSAT-3, com intervalo de tempo de meia hora entre cada imagem. Essa parte corresponde aos pontos que pertencem ao domínio dado pelo conjunto $S' = \{0, \dots, 19\} \times \{0, \dots, 19\}$, ou seja, imagens com quatrocentos pontos.

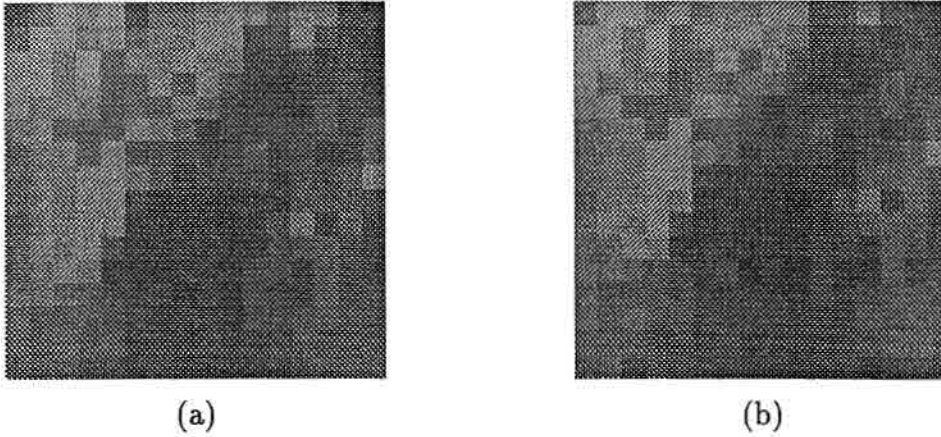
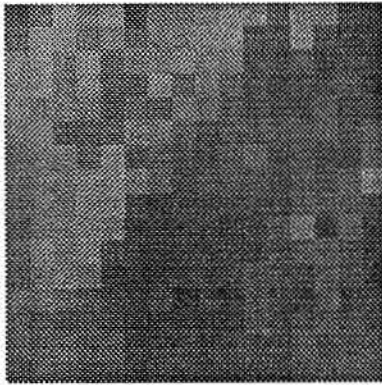
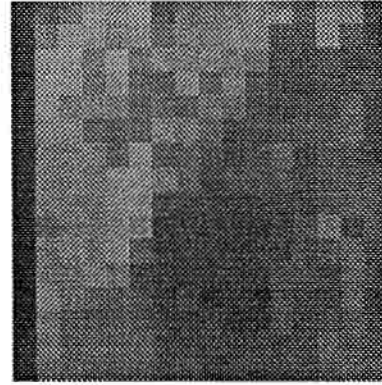


Fig. 5.13 - (a) A imagem base, (b) a imagem base deslocada no sentido arbitrário para cima

A sequência L_1 é formada pela imagem base e a imagem construída impondo-se deslocamento arbitrários em cada ponto na direção vertical, no sentido para cima. A sequência L_2 é formada pela imagem base e a imagem construída impondo-se deslocamento arbitrários em cada ponto na direção horizontal, no sentido para direita. A sequência L_3 é formada pela imagem base e a imagem real subsequente no tempo tomada pelo satélite meteorológico sobre a mesma região. Essas sequências estão mostradas respectivamente nas Figuras 5.13, 5.14 e 5.15, onde os pontos estão aumentados para efeito de visualização.

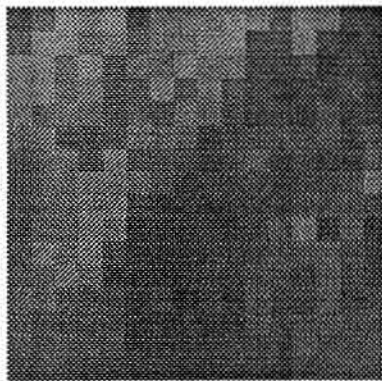


(a)

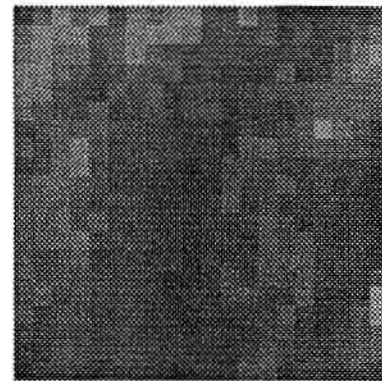


(b)

Fig. 5.14 - (a) A imagem base, (b) a imagem base deslocada no sentido arbitrário para direita.



(a)



(b)

Fig. 5.15 - (a) A imagem base, (b) a imagem real seguinte, com os deslocamentos naturais.

Na Figura 5.16 mostra-se o resultado de se detectar movimento na sequência L_1 , com a entrada dos detectores locais sendo dada apenas pelos valores de intensidades radiométricas nos pontos. Isto é, não foram introduzidos campos receptivos. As Figuras 5.17 e 5.18 referem-se às sequências L_2 e L_3 , respectivamente.

→	←	←	→	→	←	→	←	←	←	→	→	←	→	←	←	→	←	→	
←	↑	↘	←	↗	↘	↘	←	↘	↗	↓	←	↑	↓	↘	↘	→	↓	→	↓
↑	↘	↗	↑	↓	↑	←	↑	←	↘	↘	↗	↓	↘	↘	→	→	↘	←	↑
↑	→	←	↗	→	←	↑	↓	↘	↓	↘	↗	→	←	↑	→	↘	←	↑	↘
↗	←	↑	↘	↘	↗	↓	↑	↘	↘	↘	←	↑	→	↘	←	↑	↘	↓	↘
↗	→	↓	↘	←	↑	↑	↘	↓	↘	←	↑	←	↘	↗	→	↗	↓	↘	↘
↓	↘	→	←	↗	↓	↘	→	↘	↗	↑	↘	↗	↓	↘	←	←	↑	↘	↓
↗	↑	↗	↑	←	↑	↘	↓	↑	↘	→	↘	↘	↘	←	↑	↘	↓	↘	↘
↑	↓	↘	←	↗	↘	↘	↘	↘	↗	→	↘	←	↑	↗	↗	↓	↘	→	↓
↓	↘	←	←	↘	↘	↘	←	→	↓	↗	↗	↑	←	↑	↘	↗	↗	↓	↘
↑	↘	←	↘	↓	↘	←	→	↗	→	↗	↗	↘	↓	↘	↘	↗	↘	↓	↘
↗	←	↑	↘	↘	↘	↗	↗	←	↗	↑	↘	↓	↘	→	↓	↗	←	↗	↑
←	↘	↘	↓	↘	←	↑	←	↑	↓	↗	→	←	→	↓	↘	←	↘	↓	↘
↗	→	↘	↘	↗	↗	↓	↗	↓	←	↑	→	↗	↗	↑	↗	↑	↓	→	↘
→	←	↑	↘	↑	↘	↗	↑	↘	↗	↓	↘	↑	↓	↘	←	↗	↑	↗	↘
→	↗	↓	↘	↘	↘	↑	→	←	→	←	←	↗	↑	↘	↘	↘	↓	↘	↘
←	↑	←	↑	↘	↗	↓	↗	↑	←	→	↗	↗	↗	→	↗	↓	↘	↗	↑
↑	↓	↘	↘	→	↘	↗	←	↗	→	↓	↗	→	↗	←	↑	↗	↗	←	↑
→	↘	↓	↘	↘	↘	↘	↘	→	↘	↑	←	→	↘	←	→	←	↑	↗	↓
↓	↗	↑	↘	↘	↘	↑	↓	↘	↘	↓	←	←	↗	↑	↗	↑	↘	↑	↘

Fig. 5.16 - Detecção por regra do máximo, sem campo receptivo, na sequência L_1 , com deslocamentos arbitrários para cima.

São apresentados a seguir os resultados da introdução de um campo receptivo na entrada dos detectores.

Na Figura 5.19 mostra-se o resultado de se detectar movimento na sequência L_1 , com a entrada dos detectores locais sendo dada pelo campo receptivo da média aritmética dos valores de intensidades radiométricas, numa vizinhança 3×3 . As Figuras 5.20 e 5.21 referem-se às sequências L_2 e L_3 respectivamente.

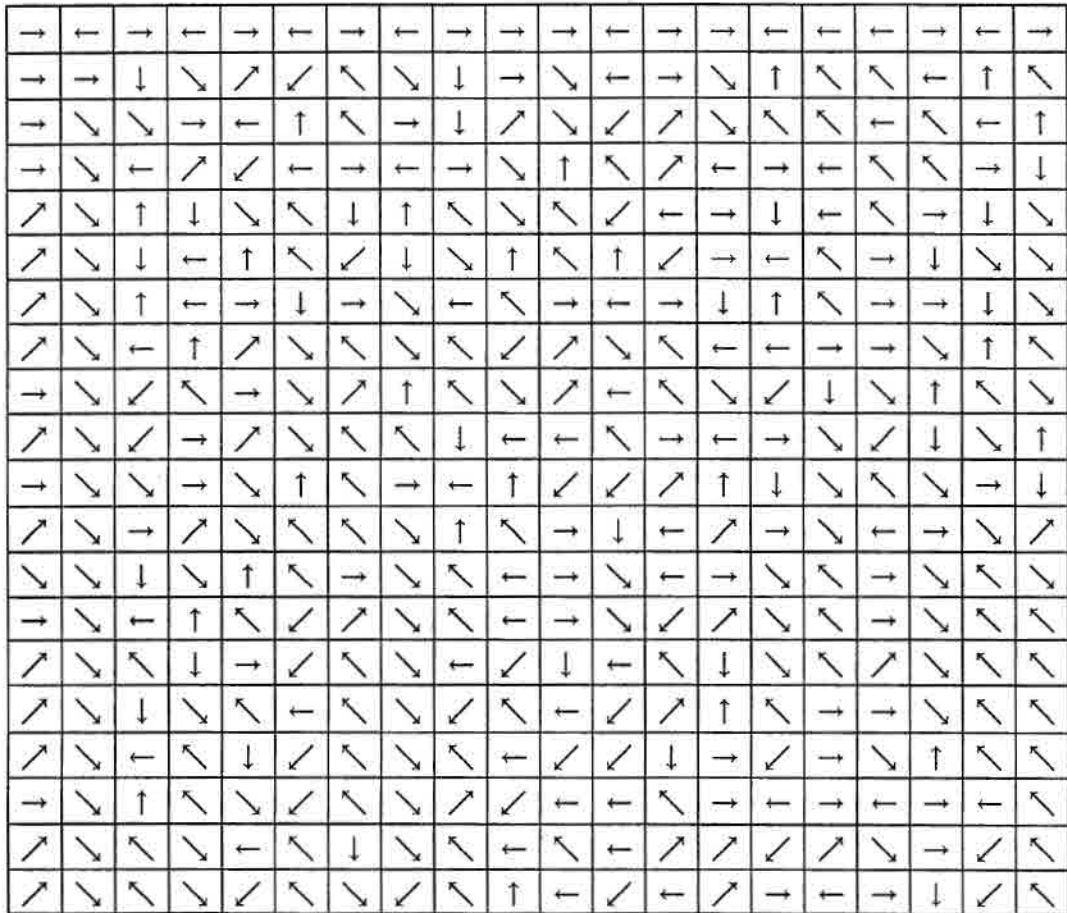


Fig. 5.17 - Detecção por regra do máximo, sem campo receptivo, na sequência L_2 , com deslocamentos arbitrários para direita.

→	←	←	→	→	←	→	→	←	←	→	→	←	→	→	←	←	→	→	
↓	↘	→	←	↑	→	↓	↘	↑	↙	↓	←	↑	↓	↘	↙	↙	↓	↘	↘
→	↓	←	↑	↘	↗	←	↑	↘	↘	↘	↙	↓	↘	↗	↑	↘	↘	←	↑
↓	↘	←	↘	↗	←	↑	↙	→	→	↘	←	↘	→	←	↙	↗	↗	↑	↘
↓	←	↘	↘	↘	↙	↓	↑	→	↘	↘	↙	↓	↗	↑	↙	←	↑	↘	↓
←	↘	↓	↘	←	↑	→	↙	↓	↘	←	↘	↗	↑	↗	↑	↙	↙	↓	↘
↑	↘	→	←	↑	↓	→	↘	↘	↗	↑	↘	↘	↓	↘	↓	←	↘	↗	↑
←	→	↗	←	↘	↑	→	↘	↑	↘	→	↘	→	←	←	↗	↑	↘	↑	↑
↑	→	↘	←	↘	↓	↘	↘	↙	→	→	↘	↗	↑	↘	↙	↓	↘	→	↘
↓	↘	←	↘	↘	↘	↑	→	↓	↗	←	↘	↙	↓	→	↙	↘	↘	↘	↘
↑	↘	←	↘	↓	↘	↘	↘	←	↑	↙	↙	↗	↑	↘	↗	↑	↘	↘	↘
←	↑	↙	→	↘	↘	↙	→	↙	↗	↑	↘	↙	→	↘	←	↘	←	↗	↘
↙	→	↙	↓	↘	↘	↘	↓	↙	↙	→	↘	↙	→	↘	←	↘	↘	↓	↘
↑	→	↘	←	↘	←	↘	↘	↓	←	→	←	→	↓	↘	←	↘	↘	↘	↘
↓	←	↗	↑	↘	↑	↗	↑	↘	↙	↓	←	→	↘	↘	←	↘	↘	→	↘
→	↙	→	↘	↓	↘	↗	↑	↘	→	↗	↙	↓	↑	↘	↙	↘	↘	→	↘
←	↑	↘	↑	↘	↙	↓	←	↗	→	↗	↗	↙	→	↓	↙	↓	↘	↗	↑
↑	↓	↘	↑	→	↘	↙	↘	↓	↗	←	↗	↑	↗	←	↘	↗	↗	↗	↑
↗	↘	↓	↘	↘	↑	↘	↘	→	↗	↑	↙	↘	↗	↙	↘	↑	→	↘	↓
↓	↘	↗	↑	↙	↘	↑	↘	→	→	↙	→	↓	←	↑	←	↘	↗	↗	↑

Fig. 5.18 - Detecção por regra do máximo, sem campos receptivos, na sequência L_3 , com deslocamentos naturais.

	→	←	←	→	←	→	←	→	←	←	←	→	→	→	→	←	→		
	↓	→	↘	↘	↑	←	↑	↓	→	↘	↘	↓	↘	↘	→	↘	←	↑	→
	←	↘	→	↘	↓	↘	←	↑	→	↘	↘	↘	↘	→	↘	←	↑	↘	
	↘	→	↓	↘	←	↘	↘	→	↘	↘	↘	↑	↘	→	↘	←	↘	↓	
	↘	↓	↘	←	↘	↘	↓	↘	↘	↘	↑	↘	→	↘	↘	←	↘	↘	
	↘	↘	←	↘	↘	↘	→	↘	↑	↘	→	↘	↘	←	↑	↘	↘	↘	
	↘	↘	←	↘	↘	↘	↘	→	→	↘	↘	←	↘	↘	↘	↘	↘	↑	
	↓	←	↘	↘	↘	↘	↘	↘	↘	↓	↘	↘	←	↘	↘	↘	↓	↘	↘
	←	←	↘	↘	↓	↘	↘	→	↘	←	↑	↘	↘	↘	↘	↘	↓	↘	
	←	↘	↘	→	↘	↘	↘	↘	↘	↑	↘	↘	↘	↘	↓	↘	↓	↘	↘
	↑	↘	→	↘	↘	↘	↘	↘	↘	↑	↘	↓	↘	↘	→	↘	↘	↓	↘
	↘	↓	↘	↘	↘	↘	↘	↘	←	→	↘	↓	↘	↘	↘	←	↘	↓	
	↓	↘	↑	↘	↘	→	↘	←	↘	→	↘	↘	→	↘	↘	↑	↘	↑	
	↘	↑	↓	↘	↘	↘	↑	↘	↘	←	↘	↑	↘	↑	↘	↘	→	↘	
	↘	→	↘	↘	↘	↘	↘	←	↘	↘	↘	↓	↘	↘	→	↓	↘	↘	
	→	↘	↑	↘	→	↘	↘	↑	↘	↘	↘	↘	↓	↘	↓	↘	↘	↑	
	↓	↘	↓	↘	↘	↘	↘	→	→	↘	→	→	↘	↘	↘	↑	←	↘	
	↘	↘	↘	↘	↑	↘	→	↓	↘	↘	↘	↘	↑	↘	↘	↑	↘	↘	

Fig. 5.19 - Detecção por regra do máximo, com campo receptivo, na sequência L_1 , com deslocamentos arbitrários para cima.

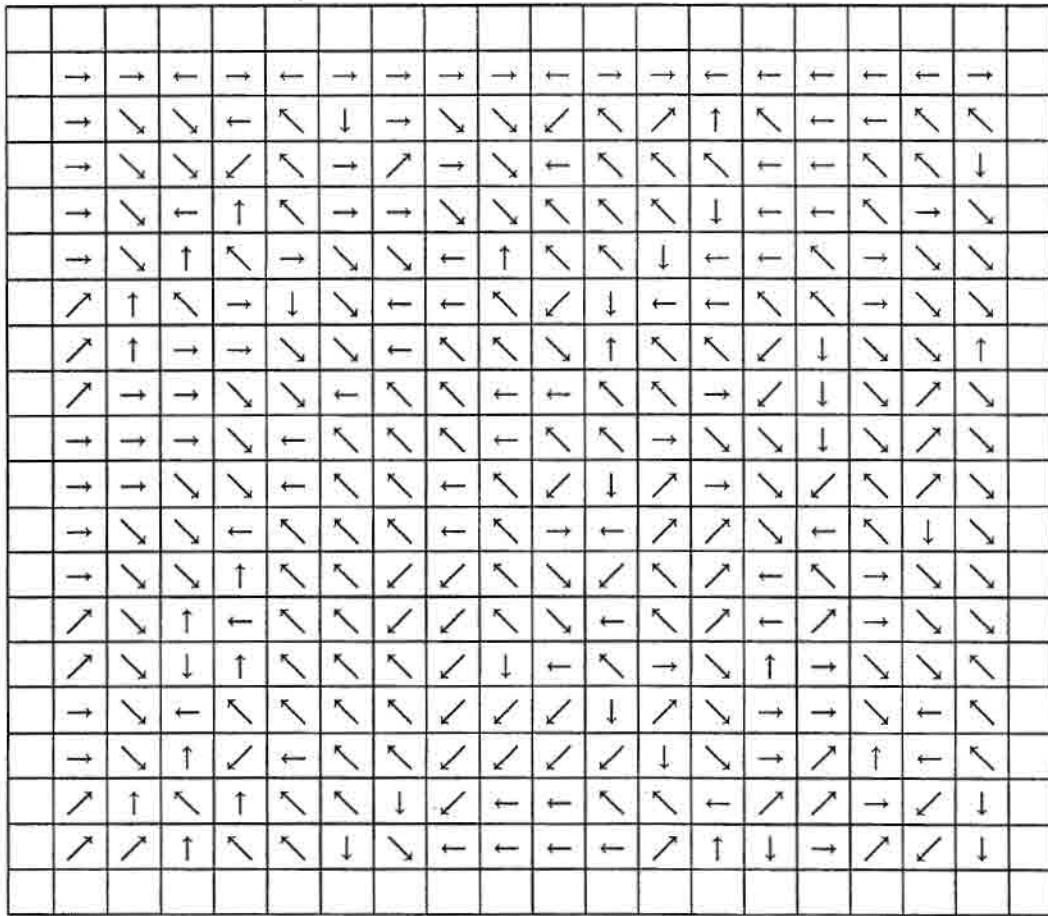


Fig. 5.20 - Detecção por regra do máximo, com campo receptivo, na sequência L_2 , com deslocamentos arbitrários para direita.

	→	←	←	→	→	←	←	→	→	←	→	→	→	←	←	→	→		
	←	←	↖	↖	↗	←	↖	↖	→	↘	↓	↘	↘	←	←	↗	↗	↑	
	←	↖	↖	↖	↗	↑	↖	→	↘	↘	↘	→	↘	↙	←	←	↑	↖	
	←	↖	↓	←	←	↖	→	→	↘	↘	↗	↗	↗	↑	←	↖	↖	↖	
	→	↓	↙	←	↖	→	↘	↘	↘	↗	↑	↗	↗	↗	↙	↙	↓	↖	
	→	↘	←	↖	↖	→	↘	↘	↑	↖	→	↘	↘	←	↑	↖	↖	↖	
	→	↘	←	↖	→	↘	↘	↘	→	→	↘	↘	←	↖	↖	↖	↓	↘	
	↓	←	↖	↖	↓	↘	↘	↙	↓	↘	↗	↑	↖	↙	↓	↘	↘	↓	
	←	←	↖	→	↘	↘	↘	→	↘	←	↖	↖	↖	→	↙	↙	↓	↘	
	←	↖	→	↘	↘	↘	↘	↗	↑	←	↖	↖	→	↗	↘	↙	↖	↘	
	↙	↓	↘	↘	↘	↓	↙	←	↖	↖	↙	↓	→	↘	↘	←	→	↘	
	→	↓	↘	↘	↖	↖	↙	↙	↖	←	→	↓	↘	↘	←	↖	→	↘	
	↗	↘	↑	↖	↖	↖	←	↙	↖	↙	↓	↘	↘	←	↖	↖	→	↘	
	↗	↑	↖	↖	↖	↗	↑	↖	→	↙	↓	↘	←	←	↖	↖	→	↘	
	↗	↑	←	↙	↖	↓	←	→	→	↘	↙	↙	↙	↙	↙	→	→	↘	
	↑	↗	↑	↖	↙	↙	↙	→	↗	↗	↑	↖	←	←	↖	→	↗	↑	
	↗	↘	↓	↘	↘	↙	↖	↗	↗	←	↑	↖	←	↖	↖	↗	↗	↑	
	↓	↘	↘	↗	↑	↖	↖	↗	↘	↙	↙	←	←	←	↖	↗	↗	↘	

Fig. 5.21 - Detecção por regra do máximo, com campo receptivo, na sequência L_3 , com deslocamentos naturais.

5.1.3 Compatibilização por Rede de Hopfield

A seguir serão apresentados os resultados da compatibilização Rede de Hopfield. Para os valores iniciais, atribuiu-se a cada nó $\eta_{s,\ell}$ o valor de $Y_\ell(s)$, ou seja, foram usados como valores iniciais de saída as chances relativas encontradas pela detecção local.

A rede foi colocada em funcionamento conforme o procedimento descrito na Seção 4.4.2 e para os três casos testados, a rede convergiu para uma solução viável, ou seja, um estado onde a cada ponto da imagem foi associado um único valor de deslocamento.

Na Figura 5.22 mostra-se o resultado da compatibilização dos deslocamentos locais detectados na sequência L_1 , nas coordenadas $s \in \{0, \dots, 9\} \times \{0, \dots, 9\}$, com a entrada dos detectores locais sendo dada pelo campo receptivo da média aritmética numa vizinhança 3×3 . As Figuras 5.23 e 5.24 referem-se aos resultados para as sequências L_1 e L_2 , respectivamente.

	→	→	→	→	→	←	←	↑	
	↑	↗	→	↗	↗	→	←	↗	
	→	→	→	↗	↗	↖	↖	↗	
	→	↑	↑	↗	→	→	↖	↗	
	↑	↑	↑	↗	→	→	→	↗	
	↑	→	↗	→	↖	→	↗	↗	
	↘	↘	→	→	→	↗	↗	↘	
	↘	↘	→	→	↑	←	↖		

Fig. 5.22 - Compatibilização por rede de Hopfield, na sequência L_1 , com deslocamentos arbitrários para direita.

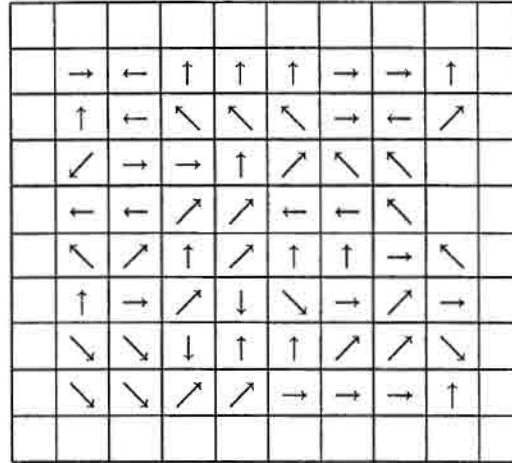


Fig. 5.23 - Compatibilização por rede de Hopfield, na sequência L_2 , com deslocamentos arbitrários na direção vertical, no sentido para cima.

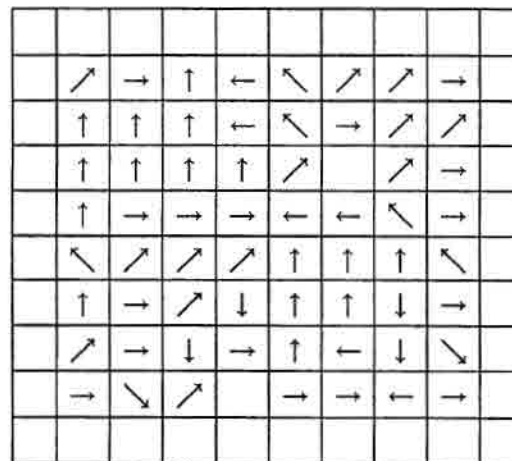


Fig. 5.24 - Compatibilização por rede de Hopfield, na sequência L_3 , com deslocamentos naturais

5.1.4 Informação computacional

Esses testes foram feitos usando estações de trabalho SUN e sistema operacional UNIX. Usou-se o simulador de redes neurais SFINX (26), desenvolvido pelo Machine Perception Laboratory, University of California at Los Angeles (UCLA).

Esse simulador, além de oferecer uma interface amigável para o tratamento e visualização de imagens, permite simular vários paradigmas de redes neurais entre eles redes do tipo Hopfield e redes implícitas (ou seja, redes com padrões de conexão fixos e onde todos os nós executam as mesmas funções).

O processamento dos detectores sensíveis à direção foi implementado através redes implícitas com funções nós escritas em linguagem C.

CAPÍTULO 6

CONCLUSÕES E EXTENSÕES

O detector biológico escolhido teve um desempenho satisfatório no sentido de que consegue capturar bem deslocamentos, lembrando que as saídas dos detectores locais representam evidências de ter ocorrido um deslocamento em uma dada direção. Para os casos testados onde conhecia-se o deslocamento correto as respostas foram coerentes, no sentido de que localmente o detector indicou chances não nulas nas direções onde poderia ter ocorrido movimento (ver Figuras 5.3, 5.4, 5.5, 5.6). Isso pode ser notado nos experimentos descritos na Seção 5.1.1 com imagens de objetos rígidos, onde os detectores locais não indicaram deslocamento na região correspondente ao fundo.

Porém, para o caso de imagens reais e com objetos não rígidos, no caso as imagens de nuvens, não existia uma separação nítida entre o objeto e o fundo e na maioria dos pontos o nível de cinza poderia ser indício de pequenas nuvens ou partes, o que causou respostas não nulas nas saídas, para a maioria dos pontos, de todos os quatro detectores.

A compatibilização por Regra do Máximo apresentou resultados pouco satisfatórios, pois não foi possível obter uma uniformidade de deslocamentos corretos, mesmo no caso de deslocamentos bem definidos em imagens com objetos rígidos. Isso acontece, devido ao fato de que, na presença de evidências iguais em direções diferentes, a Regra do Máximo faz uma escolha arbitrária de uma dessas direções. Esse é o efeito da discretização dos possíveis movimentos em um número finito de direções.

Esse problema é bem resolvido, no caso de imagens de objetos rígidos, com a compatibilização por Rede de Hopfield, com a atribuição de um valor maior para o parâmetro que controla a restrição de suavidade espacial. Os valores dos parâmetros para os testes com imagens de objetos rígidos foram $A = 500$,

$B = 700, C = 500, D = 300.$

Como sugerido em (18,48) a escolha desses parâmetros foi feita a partir de testes de forma a conseguir soluções que fossem pelo menos possíveis, ou seja, soluções onde a cada coordenada só haveria um deslocamento associado. Pequenas variações nos valores dos parâmetros A, B, C e D não causaram mudanças no comportamento da convergência para alguma configuração.

Para o caso das imagens reais, a compatibilização por Regra do Máximo mostrou-se insatisfatória mesmo para o caso de movimentos arbitrários sem deformações. Para os deslocamentos naturais novamente essa regra de compatibilização apresentou resultados pouco satisfatórios quando comparado com a sensação de deslocamento captado por um observador humano tem quando observa a sequência testada.

A compatibilização por Rede de Hopfield apresentou para os casos tratados uma pequena melhora nos casos onde o deslocamento era conhecido. No caso dos deslocamentos naturais notou-se uma pequena melhora, novamente quando compara-se os deslocamentos apontados por esse modelo com a sensação de movimento captada por um observador. Isso se deve novamente à restrição de suavidade espacial, imposta na equação dos pesos. Para os casos de imagens reais os valores de parâmetros usados foram $A = 500, B = 500, C = 500$ e $D = 300.$

O tamanho das imagens testadas para a compatibilização não é suficiente para que se possa tirar conclusões mais fortes a respeito do desempenho desse modelo. Os testes com imagens maiores não foram feitos devido ao problema computacional que se tem para a implementação da rede de Hopfield. Para uma imagem de tamanho $N \times N$ e L direções, é necessário construir uma rede com $N^2 \cdot L$ nós e $N^4 \cdot L^2.$ Esse foi o principal fator limitante para a observação de novos resultados.

Esse trabalho contribui para o estudo do comportamento dos detectores biológicos, sua formulação para o caso de imagens digitais, discretas e no "pior caso", ou seja, imagens ruidosas com objetos não rígidos. Desenvolveu-se

uma formulação matemática coerente. Apesar das limitações quanto a equipamento e software, considera-se este um ponto de partida para a aplicação desse tipo de detectores a imagens e problemas reais.

O enfoque combinatório dado ao problema da compatibilização mostrou-se promissor para o caso de imagens de objetos rígidos e pode ser estendida a outros problemas em visão computacional (21). A proposta de solução por rede de Hopfield a problemas análogos, do tipo combinatórios, é de grande interesse devido a bons resultados obtidos, como os mostrados em (18,47,48).

A extensão natural desse trabalho seria a proposta de uma forma computacional eficiente de aplicar a compatibilização por rede de Hopfield á imagens reais de tamanho maior, que seriam mais representativas do problema tratado. Essa implementação deve levar em conta também o tempo necessário para se obter boas soluções.

Uma proposta neste sentido, seria a utilização de um processamento piramidal, que tem exibido bons resultados em problemas de visão computacional e processamento de imagens (33).

Outro ponto que deve ser considerado é a proposta de uma forma de avaliar resultados obtidos, ou as saídas fornecidas pelo modelo, de forma a ter-se uma medida quantitativa e qualitativa do desempenho do modelo. Uma forma seria usar os deslocamentos apontados no modelo para a obtenção de previsões de comportamento de uma sequência e comparar resultados com a sequência real. Além disso, os deslocamentos obtidos poderiam ser utilizados para a segmentação de objetos (nuvens) e sua possível categorização.

Outra possível extensão seria usar, como sugerido em (3), redes neurais do tipo *backpropagation* para aprender a evolução de padrões espaço-temporais, do tipo evolução de campos de deslocamento de nuvens em sequências de imagens meteorológicas, e também para a categorização desses campos de deslocamento.

REFERÊNCIAS BIBLIOGRÁFICAS

- (1) Adelson, E.H.; Bergen, J.R. Spatiotemporal energy models for the perception of motion. *Journal of the Optical Society of America A*, 2(2):284–299, Feb. 1985.
- (2) Aggarwal, J.K.; Nandhakumar, N. On the computational of motion from sequences of images : a review. *Proceedings of the IEEE*, 76(8):917–935, Aug. 1988. Special Issue on Computer Vision.
- (3) Alonso, E. M. B.; Rodrigues, V.; Amorim, M. C. A. Generating future states in satellite imagery by neural networks.
- (4) Amari, S. Mathematical foundations of neurocomputing. *Proceedings of the IEEE*, 78(9):1443–1464, Sep. 1990. Special Issue on Neural Networks, 1.
- (5) Banon, G. *Processamento digital de imagens*. Notas de aula dadas no INPE. São José dos Campos, 2. período, 1991. Manuscrito.
- (6) Beale, R.; Jackson, T. *Neural computing: an introduction*. Philadelphia, PA, IOP Publishing Ltda, 1990.
- (7) Besag, J. On the statistical analysis of dirty pictures (with discussion). *Journal of the Royal Statistical Society B*, 48(3):259–302, 1986.
- (8) Besag, J. Towards Bayesian image analysis. *Journal of Applied Statistics*, 16(3):395–407, 1989.
- (9) Edwards, R.G.; Sokal, A.D. Generalization of the Fortuin-Kasteleyn-Swendsen-Wang representation and Monte Carlo algorithms. *Physical Review D*, 38(6):2009–2012, Sep. 1988.
- (10) Edwards, R.G.; Sokal, A.D. Dynamical critical behavior of Wolff's collective-mode Monte Carlo algorithm for the two-dimensional $\mathcal{O}(n)$ nonlinear σ model. *Physical Review D*, 40(4):1374–1377, Aug. 1989.

- (11) Fiesler, E. Neural network classification and formalization. In: Fulcher, J. ed. *Neural network standards*. North-Holland, Amsterdam, 1994, s.p. (Computer Standards & Interfaces, 16.).
- (12) Fischback, G.D. Mind and brain. *Scientific American*, 267(3):24-23, Sep. 1992. Special Issue Mind and Brain.
- (13) Freeman, J.A.; Skapura, D.M. *Neural networks algorithms, applications, and programming techniques*. Reading, MA, Addison-Wesley, 1991.
- (14) Frery, A.C. *Algumas ferramentas estatísticas para síntese, processamento e análise de imagens de radar de abertura sintética*. (Tese de Doutorado) INPE, São José dos Campos, 1993.
- (15) Geman, D. Random fields and inverse problems in imaging. In: Hennequin, P.L., ed. *École d'été de probabilités de Saint-Flour XVIII-1988*. Berlin, Springer-Verlag, 1988. (Lecture Notes in Mathematics, 1427).
- (16) Geman, S.; Geman, D. Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-6(6):721-741, Nov. 1984.
- (17) Hopfield, J.J. Neurons with graded response have collective computational properties like those of two-state neurons. *Proceedings of National Academy of Sciences USA*, 81:3088-3092, May 1984.
- (18) Hopfield, J.J.; Tank, D.W. "Neural" computation of decisions on optimization problems. *Biological Cybernetics*, 54:145-152, 1985.
- (19) Horn, B.K.P.; Schunck, B.G. Determining optical flow. *Artificial Intelligence*, 17:185-203, 1981.
- (20) Jehoshua, B. On the convergence properties of the Hopfield model. *Proceedings of IEEE*, 78(10):1579-1585, Oct. 1990. Special Issue on Neural Networks, 2.

- (21) Koch, C.; Marroquin, J.; Yuille, A. Analog "neuronal" networks in early vision. Boston, MIT, 1985. (MIT-AI Memo 751, 1985).
- (22) Lindsay, P.H.; Norman, D.A. *Human information processing: an introduction to psychology*. New York, NY, Academic Press, 1977.
- (23) Livingstone, M.; Hubel, D. Segregation of form, color, movement and depth: anatomy, physiology and perception. *Science*, 240(4853):740-750, May 1988.
- (24) Ma, S. *Statistical Mechanics* Philadelphia, PA, World Scientific Publishing Co. Pte. Ltd., 1985.
- (25) Mascarenhas, N.D.A.; Velasco, F.R.D. *Processamento digital de imagens*. Buenos Aires, Editorial Kapelusz, 1989. Escola Brasileira-Argentina de Informática, 4.
- (26) Mesrobian, E.; Stiber, M.; Skrzypek, J. *SFINX: structure and functions in neural connections*. Los Angeles, CA, 1990. (MPL-TR 89-9).
- (27) Molina, M.E. El sistema de lucha antigranizo. *Diario Los Andes*, Mendoza, Argentina, 23 fev. 1993. c.1, p.3, c.1.
- (28) Nagel, H. On a constraint equation for the estimation of displacements rates in image sequences. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-11(1):13-30, Jan. 1989.
- (29) Nelson, M.E.; Bower, J.M. Brain maps and parallel computers. *TINS*, 13(10):403-408, 1990.
- (30) Nowlan, S. J.; Sejnowski, T. J. Filter selection model for generating visual motion signals, In: Giles, C.L.; Hanson, S.J.; Cowan, J.D. *Advances in Neural Information Processing Systems*. San Mateo, CA, Morgan Kaufman, 369-376, 1993.
- (31) Papadimitriou, C.H.; Steiglitz, K. *Combinatorial optimization algorithms and complexity*. Englewood Cliffs, NJ, Prentice-Hall, 1982.

- (32) Reichardt, W. Autokorrelationsauswertung als Funktionsprinzip des zentralnervensystems. *Z. Naturforsch*, 12b:447-457, 1957.
- (33) Rosenfeld, A. *Multiresolution image processing and analysis*. New York, Springer-Verlag, 1984.
- (34) Sabata, B.; Aggarwal, J.K. Estimation of motion from a pair of range images: a review. *CVGIP: Image Understanding*, 54(3):309-325, Nov. 1991.
- (35) Simpson, P.K. *Artificial neural systems. Foundations, paradigms, applications and implementations*. San Diego, CA, Pergamon Press, 1989.
- (36) Singer, M.H. Significant feature detection and matching in image pairs. In: 1987 International Joint Conference on Artificial Inteligence (IJCAI'87). Milan, Aug 1987. *Proceedings*. p. 829-831.
- (37) Spetsakis, M.E.; Aloimonos, J.Y. Structure from motion using line correspondences. *International Journal of Computer Vision*, 4(3):171-225, Apr. 1990.
- (38) Torreão, J.R.A. *Métodos estocásticos em computação visual*. Porto Alegre, Instituto de Informática da UFRGS, 1992. 148 p. Notas de curso da Escola de Computação, 8.
- (39) Ullman, S. *The interpretation of visual motion*. Cambridge, MIT Press, 1979.
- (40) Van Santen, J.P.H.; Sperling, G. Temporal covariance model of human motion perception. *Journal of the Optical Society of America A*, 1(5):451-473, May 1984.
- (41) Van Santen, J.P.H.; Sperling, G. Elaborated Reichardt detectors. *Journal of the Optical Society of America A*, 2(2):300-321, Feb. 1985.

- (42) Watson, A.B.; Ahumada, A.J. Model of human visual-motion sensing. *Journal of the Optical Society of America A*, 2(2):322-342, Feb. 1985.
- (43) Wehmeier, U. Modeling the mammalian visual system. In: Koch, C.; Segev, I. (ed.) *Methods in neuronal modeling. From synapses to networks*. USA, MIT Press, 1989.
- (44) Whestheimer, G. Relating neural mechanisms to visual perception. Historical and philisofical considerations. In: Spillman, L.; Werner, J.S. ed. *Perception: the neurophysiological foundations*. San Diego, CA, Academic Press, Harcourt Brace Jovanovich Publishers, 1990.
- (45) Willick, D.; Yang, Y.H. Experimental evaluation of motion constraints equations. *CVGIP: Image Understanding*, 54(2):206-214, Sep. 1991.
- (46) Zeki, S. The visual image in mind and brain. *Scientific American*, 267(3)42-51, Sep. 1992. Special Issue Mind and Brain.
- (47) Zhou, Y.T.; Chellappa, R. A network for motion perception. In: 1991 International Joint Conference on Neural Networks (IJCNN). Singapore, Nov. 18-21, 1991. *Proceedings* p. 875-884.
- (48) Zhu, P.Y.; Kasvand, T.; Krzyzak, A. Motion estimation based on point correspondence using neural network. In: 1991 International Joint Conference on Neural Networks (IJCNN). Singapore, Nov. 18-21, 1991. *Proceedings* p. 869-874.