UNIVERSIDADE FEDERAL FLUMINENSE

SILVIA MARA DA COSTA CAMPOS VICTER

FUNÇÕES DE GABOR SINTONIZADAS: APLICAÇÃO À ANÁLISE E À SÍNTESE DE SINAIS

NITERÓI 2012

UNIVERSIDADE FEDERAL FLUMINENSE

SILVIA MARA DA COSTA CAMPOS VICTER

FUNÇÕES DE GABOR SINTONIZADAS: APLICAÇÃO À ANÁLISE E À SÍNTESE DE SINAIS

Tese de Doutorado apresentada ao Programa de Pós-Graduação em Computação da Universidade Federal Fluminense como requisito parcial para a obtenção do Grau de Doutor em Computação. Área de concentração: Computação Visual e Interfaces.

Orientador: Prof. Dr. JOSÉ RICARDO DE ALMEIDA TORREÃO

> NITERÓI 2012

SILVIA MARA DA COSTA CAMPOS VICTER

FUNÇÕES DE GABOR SINTONIZADAS: APLICAÇÃO À ANÁLISE E À SÍNTESE DE SINAIS

Tese de Doutorado apresentada ao Programa de Pós-Graduação em Computação da Universidade Federal Fluminense como requisito parcial para a obtenção do Grau de Doutor em Computação. Área de concentração: Computação Visual e Interfaces.

Aprovada em 22 de março de 2012.

BANCA EXAMINADORA

Prof. Dr. JOSÉ RICARDO DE ALMEIDA TORREÃO - Orientador, UFF

Prof. Dr. ANSELMO ANTUNES MONTENEGRO, UFF

 ${\rm Prof}^a.$ ${\rm Dr}^a.$ AURA CONCI, UFF

Prof. Dr. EDUARDO A.B. DA SILVA, UFRJ

Prof. Dr. MARIO F. MONTENEGRO CAMPOS, UFMG

Niterói 2012

Dedico aos meus queridos filhos Vinícios e Felipe e ao meu marido Marcelo.

Agradecimentos

Ao meu orientador, José Ricardo Torreão, os maiores e mais sinceros agradecimentos. Agradeço-lhe a confiança demonstrada e as suas preciosas críticas e sugestões; pela paciência, colaboração, e conhecimentos repassados durante todo o desenvolvimento do trabalho.

Ao professor João Luíz Fernandes, por sua disponibilidade e boa vontade em colaborar durante todas as etapas do trabalho.

Aos meus pais, pela educação, amor, dedicação e exemplo de vida.

Ao meu marido, por todo o seu amor, carinho, admiração, e compreensão durante o tempo de desenvolvimento deste trabalho.

Aos meus filhos, por serem tão maravilhosos.

A toda a família e amigos que acreditaram em mim e torceram pela minha vitória.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior, CAPES, pela bolsa concedida durante os anos dedicados à pesquisa.

A Deus, que me deu luz, inspiração e, principalmente, muita força de vontade para superar todas as dificuldades e a seguir sempre adiante.

Resumo

As funções de Gabor sintonizadas foram originalmente propostas num contexto de visão computacional, como modelo para o desfocamento de imagens. Demonstrou-se então que um sinal espacial pode ser expresso como a superposição de funções de Gabor (senóides moduladas por Gaussianas), cujos parâmetros de largura e de fase são determinados pela transformada de Fourier do sinal. Aqui nós estendemos este modelo para a representação de sinais no domínio da frequência, empregando funções de Gabor cujos parâmetros são obtidos a partir da transformada de Fourier inversa, e demonstramos que ambos os tipos de funções codificadoras podem ser usados para definir transformadas de Gabor sintonizadas. Nestas, cada sinal - espacial ou na frequência - é analisado por núcleos de Gabor associados a sua própria representação, de forma que, num certo sentido, o próprio sinal é a sua função analisadora, o que confere às transformadas sintonizadas propriedades semelhantes às da transformada de Wigner. Como exemplo de aplicação, a primeira forma da transformada sintonizada foi empregada com sucesso na caracterização da epilepsia a partir de eletro-encefalogramas. Numa outra linha de investigação, nós verificamos que a versão bidimensional do modelo de representação sintonizada pode ser usada para descrever os campos receptivos de células do córtex visual, que experimentos recentes demonstram ser dependentes do estímulo. Uma extensão da mesma abordagem, baseada em funções de codificação rotacionalmente simétricas, também foi proposta e utilizada para descrever o processo de branqueamento de sinais que se supõe ser realizado pelas células da retina e do núcleo geniculado lateral. Os nossos modelos sintonizados de representação neuronal foram então empregados na definição de um algoritmo de casamento estereoscópico que emula o processamento efetuado pelos estágios cortical e pré-cortical do cérebro.

Palavras-chave: Funções de Gabor. Análise de Sinais. Modelagem Neuronal.

Abstract

The work here reported deals with signal-tuned Gabor functions, and their application to signal analysis and synthesis. Such functions have been originally proposed in a computer vision framework, as a model for defocusing. It has been shown that any spatial signal can be expressed as a superposition of Gabor functions (Gaussian-modulated sinusoids), whose phase and width are determined by the signal's Fourier transform. We have extended such model for the representation of signals in the frequency domain, by using Gabor functions whose parameters are obtained from the inverse Fourier transform, and we have also shown that both coding functions can be used for the definition of signaltuned Gabor transforms. Such transforms analyze each input signal - in the spatial or frequency domains – by means of Gabor kernels associated with the signal's representation – thus, in a certain sense, the analyzed signal is its own analyzing function, what leads to similar properties as those of the Wigner transform. As a practical application, the first form of the signal tuned transform has been successfully used for the characterization of epilepsy from EEG signals. In a different line of investigation, we have shown that the 2-D extension of the space-domain coding model can be used for describing the receptive fields of neurons of the visual cortex, which have been proven, by recent experiments, to depend on the presented stimulus. An extension of the same approach, now based on rotationally symmetric coding functions, has also been introduced for modeling the whitening transformation supposedly performed by the cells of the retina and of the lateral geniculate nucleus. Finally, our signal-tuned neuronal representation models have been employed in a stereoscopic estimation algorithm which emulates the processing performed by the cortical and pre-cortical stages of the brain.

Keywords: Gabor Functions. Signal Analysis. Neuronal Modeling.

Lista de Figuras

2.1	Em (a), o sinal $I_1(x)$. Em (b), o seu espectrograma $ T(\omega, x) $. Em (c), a magnitude da sua transformada de Wigner, $ W(\omega, x) $	12
2.2	Em (a), sinal $I_2(x) = \delta(x - 200)$. Em (b), o seu espectrograma $ T(\omega, x) $.	13
2.3	Em (a), sinal $I_3(x) = \sum_k \delta(x - kL)$. Em (b), o seu espectrograma $ T(\omega, x) $.	14
2.4	Espectrograma $ T(\omega, x) $ do sinal $I_4(x) = e^{i(\pi/50)x}$.	15
2.5	Espectrograma $ T^{(2)}(\omega, x) $ do sinal $I_1(x)$	20
2.6	Espectrograma $ T^{(2)}(\omega, x) $ do sinal $I_2(x)$	20
2.7	Espectrograma $ T^{(2)}(\omega, x) $ do sinal $I_4(x)$	20
2.8	Em (a), sinal $I_5(x)$ e o mapa de magnitude de suas funções de representação. Em (b), à esquerda, é mostrado o espectrograma $ T(\omega, x) ^{1/4}$ e, à direita, o espectrograma $ G(\omega, x) ^{1/4}$.	28
2.9	Em (a), sinal $I_6(x)$. Em (b), o seu espectrograma $ T(\omega, x) ^{1/4}$. Em (c), o seu espectrograma $ G(\omega, x) ^{1/4}$ obtido com a transformada de Gabor	29
2.10	Em (a), sinal $I_7(x)$. Em (b), são apresentadas duas vistas do gráfico tridi- mensional do seu espectrograma $ T(\omega, x) $. Em (c), é apresentado o espec- trograma $ G(\omega, x) $, obtido com a transformada de Gabor	31
2.11	Transformada wavelet (Equação (2.61)) para o sinal $I_7(x)$. Parâmetros: (a) $\beta = 0,2$. (b) $\beta = 0,6$. (c) $\beta = 1,2$	32
2.12	Em (a), sinal $I_8(x)$. Em (b), à esquerda, é mostrada a magnitude da transformada de Fourier, e, à direita, o espectrograma $ T(\omega, x) $. Em (c), à esquerda, é mostrado o espectro de Fourier suavizado, e, à direita, o espectrograma $ T(\omega, x) $ obtido a partir deste. Em (d), é apresentada a magnitude da Transformada S. Em (e), é apresentado o espectrograma $ G(\omega, x) $, obtido a partir da transformada de Gabor	34
		01

2.13	Em (a), sinal $I_9(x)$. Em (b), o espectrograma $ T(\omega, x) $ obtido a partir do espectro de Fourier suavizado. Em (c), o espectrograma $ G(\omega, x) $, obtido a partir da transformada de Gabor.	35
2.14	Em (a), à esquerda, sinal $I_{10}(x)$ e, à direita, a magnitude da transformada de Fourier. Em (b), à esquerda, o espectrograma $ \hat{T}(\omega, x) $ e, à direita, o espectrograma $ T(\omega, x) $. Em (c), o espectrograma $ G(\omega, x) $ obtido a partir da transformada de Gabor	36
2.15	Em (a), sinal This is a Test e, em (b), o seu espectrograma $ T(\omega, x) $. Em (c), o espectrograma $ G(\omega, x) $, obtido a partir da transformada de Gabor.	37
2.16	Em (a), sinal <i>Glockenspiel</i> e, em (b), o seu espectrograma $ T(\omega, x) $. Em (c), o espectrograma $ G(\omega, x) $, obtido a partir da transformada de Gabor.	38
3.1	Em (a), é mostrado um exemplo de sinal de EEG para um paciente normal. Em (b), são mostrados exemplos de <i>paroxismos epileptiformes interictais</i> , e, em (c), exemplos de <i>paroxismos epileptiformes ictais</i> (ver texto) [25]	40
3.2	Bandas de frequência usadas na análise de sinais de EEG. Figura extraída de [28]	42
3.3	Esquema de localização dos eletrodos de superfície de acordo com o sistema internacional 10-20, como visto a partir da esquerda (a), e acima da cabeça (b). A nomenclatura das posições dos eletrodos deriva da sua localização anatômica, a saber: F - lobo Frontal, T - lobo Temporal, C - lobo Central, P - lobo Parietal, O - lobo Occipital. Z se refere ao eletrodo localizado na linha medial. Figura extraída de [30]	42
3.4	Amostra de sinais de EEG obtidos via eletrodos de superfície em várias áreas do córtex cerebral. Figura extraída de [27].	43
3.5	Esquema de implantação de eletrodos intra-craniais, para avaliação pré- cirúrgica de pacientes epilépticos. Em (a), eletrodos de profundidade im- plantados simetricamente nas formações do hipocampo. Em (b) e (c), faixas de eletrodos implantados nas regiões do neocórtex (áreas mais evoluí- das do córtex) [13]	44
3.6	Exemplos de segmentos de EEG de cada uma das cinco classes. De cima para baixo, classes A a E . Nos eixos verticais, as amplitudes dos sinais	
	estão em μV [13]	48

3.7	O mesmo sinal da Fig. 3.6e, apresentado em sua duração total (23,6 se- gundos)	49
3.8	Intensidade relativa de banda (<i>RIR</i>), para cada banda de frequência, obtida a partir da transformada de Gabor sintonizada, (a) para o segmento de sinal apresentado na Fig. 3.6e, e (b), para o mesmo sinal, mas na duração total de 23,6 segundos (Fig. 3.7).	50
3.9	Espectrograma $ T(\omega, x) $ dos sinais da Figura 3.6, referentes às classes de A a D (de cima para baixo)	52
3.10	(a) Segmento de um sinal de EEG de um paciente com diagnóstico de epilepsia. Amplitude em μV no eixo vertical. (b) Espectrograma da transformada sintonizada $ T(\omega, x) $ (c) Espectrograma da transformada de Gabor $ G(\omega, x) $, utilizando janela de 0,25 segundo. (d) Espectrograma da transformada $S.$	53
3.11	(a) Segmento de um sinal de EEG de um paciente com diagnóstico de epilepsia. Amplitude em μV no eixo vertical. (b) Espectrograma da transformada sintonizada $ T(\omega, x) $ (c) Espectrograma da transformada de Gabor $ G(\omega, x) $, utilizando janela de 0,25 segundo. (d) Espectrograma da transformada S	55
4.1	Estrutura de um neurônio [8]	57
4.2	Caminho Visual [8]	57
4.3	Tipos de campos receptivos <i>centro-periferia</i> . À esquerda, um campo centro- on e à direita, um campo centro-off. O símbolo "+"denota regiões exci- tatórias, enquanto o "-"denota regiões inibitórias [8]	59
4.4	Exemplos de funções de representação da Equação (4.1). As frequências consideradas, (ω_x, ω_y) , de cima para baixo e da esquerda para a direita, são: $(0,0), (0,1), (1,0), (0,2), (1,1) \in (1,2).$	60
4.5	Em cima: imagens originais, de tamanho 128×192 . Embaixo: represen- tação obtida a partir do nosso modelo, computada usando janelas 3×3 , para as frequências $(0,0)$, $(0,1)$, $(1,0)$, $(0,2)$, $(1,1)$ e $(1,2)$	61
4.6	Magnitude das funções de representação da Equação (4.6), obtidas de um fragmento 3×3 da imagem natural na Figura 4.7a. As frequências representadas (ω_x, ω_y) , de cima para baixo, são: (0,1), (0,2), e (3,1)	65

4.7	Em (a), à esquerda, a imagem original e à direita a imagem branqueada. b) Os espectros log-log correspondentes (ver o texto). Foram utilizadas janelas 5×5 , com $\kappa = 0.05$. Essas imagens, e também as das Figuras 4.8 a 4.10, são 192×192	66
4.8	Em (a), à esquerda, imagem original e à direita a imagem branqueada. b) Os espectros log-log correspondentes (ver o texto). Foram utilizadas janelas 3×3 , com $\kappa = 0.05$	67
4.9	Em (a), à esquerda, imagem original e à direita, a imagem branqueada. b) Os espectros log-log correspondentes (ver o texto). Foram utilizadas janelas 5×5 , com $\kappa = 0.05$	68
4.10	Em (a), à esquerda, imagem original e à direita, a imagem branqueada. b) Os espectros log-log correspondentes (ver o texto). Foram utilizadas janelas 3×3 , com $\kappa = 1$	69
5.1	Geometria da visão estereoscópica [8]	71
5.2	Vista superior do Caminho Visual [52]	76
5.3	Estereograma de pontos aleatórios denso. Em (a), são mostradas as im- agens originais, em (b), as imagens branqueadas, e, em (c), os mapas de disparidades real e estimado	81
5.4	Estereograma de disparidade contínua. Em (a), são mostradas as imagens originais, em (b), as imagens branqueadas, e, em (c), os mapas de dispari- dades real e estimado.	82
5.5	Estereograma <i>Sawtooth</i> . Em (a), são mostradas as imagens originais, em (b), as imagens branqueadas, e, em (c), os mapas de disparidades real e estimado.	83
5.6	Estereograma <i>Motherboard</i> . Em (a), são mostradas as imagens originais, em (b), as imagens branqueadas, e, em (c), o mapa de disparidades	85
5.7	Estereograma <i>Pentagon</i> . Em (a), são mostradas as imagens originais, em (b), as imagens branqueadas, em (c), o nosso mapa de disparidades, e, em (d), o mapa de disparidades obtido com as abordagens de Green anteriores:	
	à esquerda, aquela de [62]; à direita, aquela de [11]	86

5.8	Estereograma $Tree$. Em (a), são mostradas as imagens originais, em (b),	
	as imagens branqueadas, em (c), o nosso mapa de disparidades, e, em (d),	
	o mapa de disparidades obtido com as abordagens de Green anteriores: à	
	esquerda, aquela de [62]; à direita, aquela de [11]	87
5.9	Estereograma $Shrub$. Em (a), são mostradas as imagens originais, em (b),	
	as imagens branqueadas, em (c), o nosso mapa de disparidades, e, em (d),	
	o mapa de disparidades obtido com as abordagens de Green anteriores: à	
	esquerda, aquela de [62]; à direita, aquela de [11]. \ldots	88
5.10	Estereograma Meter. Em (a), são mostradas as imagens originais, em (b),	
	as imagens branqueadas, em (c), o nosso mapa de disparidades e, em (d),	
	o mapa de disparidades obtido com as abordagens de Green anteriores: à	
	esquerda, aquela de [62]; à direita, aquela de [11]. \ldots	89

Lista de Tabelas

3.1	Estados mentais associados às bandas de frequência	41
3.2	Descrição das cinco classes de sinais de EEG consideradas [13]	46

Sumário

1	Intr	odução		1
	1.1	Prime	ra abordagem de representação	3
		1.1.1	Extensão 2-D	5
	1.2	Segun	la abordagem de representação	6
		1.2.1	Extensão 2-D	7
2	Apli	icação à	Análise de Sinais	8
	2.1	Introd	ıção	8
	2.2	A Tra	nsformada de Gabor Sintonizada	9
		2.2.1	Primeira Forma da Transformada: $T(\omega,x)$	9
		2.2.2	Propriedades analíticas da $T(\omega, x)$:	6
			2.2.2.1 Momentos da frequência:	6
			2.2.2.2 Translação Espacial:	6
			2.2.2.3 Translação na Frequência:	6
			2.2.2.4 Primeiro momento condicional do espaço:	6
			2.2.2.5 Momento cruzado do espaço e da frequência:	7
		2.2.3	Segunda Forma da Transformada: $T^{(2)}(\omega, x)$	9
		2.2.4	Propriedades analíticas da $T^{(2)}(\omega, x)$:	1
			2.2.4.1 Transformada da delta de Dirac:	1
			2.2.4.2 Transformada do trem de impulsos:	1
			2.2.4.3 Transformada da exponencial complexa:	2
			2.2.4.4 Densidade de Energia:	2

		2.2.4.5	Correlação em frequência:	22
		2.2.4.6	Momentos da posição:	22
		2.2.4.7	Translação Espacial:	23
		2.2.4.8	Translação na Frequência:	23
		2.2.4.9	Primeiro momento condicional da frequência:	23
		2.2.4.10	Momento cruzado da posição e da frequência:	23
	2.3	Estudo Experim	nental da $T(\omega, x)$	27
3	Apli	icação à Análise d	le Sinais de EEG para a Caracterização de Epilepsias	39
	3.1	Introdução		39
	3.2	Coleta de sinais	de EEG e caracterização dos sinais de epilepsia $\ .\ .\ .$.	41
	3.3	Análise Tempo-	Frequência de sinais de EEG	44
	3.4	Experimentos .		46
4	Apli	icação à Modelag	em Neuronal	56
4	Apl i 4.1	icação à Modelag Neurofisiologia e	e m Neuronal da Visão: O Caminho Visual	56 56
4	Apl i 4.1	i cação à Modelag Neurofisiologia 4.1.1 Campos	em Neuronal da Visão: O Caminho Visual	56 56 58
4	Apl i 4.1 4.2	icação à Modelag Neurofisiologia 4.1.1 Campos Modelagem das	em Neuronal da Visão: O Caminho Visual	56 56 58 59
4	Apl 4.1 4.2 4.3	icação à Modelag Neurofisiologia d 4.1.1 Campos Modelagem das Modelagem das	em Neuronal da Visão: O Caminho Visual	56 56 58 59 61
4	 Apli 4.1 4.2 4.3 Ester 	icação à Modelag Neurofisiologia o 4.1.1 Campos Modelagem das Modelagem das ereoscopia	em Neuronal da Visão: O Caminho Visual	 56 56 58 59 61 70
4	 Apli 4.1 4.2 4.3 Este 5.1 	icação à Modelag Neurofisiologia o 4.1.1 Campos Modelagem das Modelagem das ereoscopia Introdução	em Neuronal da Visão: O Caminho Visual	 56 56 58 59 61 70 70
4	 Apli 4.1 4.2 4.3 Ester 5.1 5.2 	icação à Modelag Neurofisiologia o 4.1.1 Campos Modelagem das Modelagem das ereoscopia Introdução Estimação de D	em Neuronal da Visão: O Caminho Visual	 56 58 59 61 70 70 73
4	 Apli 4.1 4.2 4.3 Ester 5.1 5.2 5.3 	icação à Modelag Neurofisiologia o 4.1.1 Campos Modelagem das Modelagem das ereoscopia Introdução Estimação de D Experimentos .	em Neuronal da Visão: O Caminho Visual	 56 58 59 61 70 73 80
4 5	Apli 4.1 4.2 4.3 Este 5.1 5.2 5.3 Con	icação à Modelag Neurofisiologia d 4.1.1 Campos Modelagem das Modelagem das reoscopia Introdução Estimação de D Experimentos . clusões e Trabalh	em Neuronal da Visão: O Caminho Visual	 56 58 59 61 70 70 73 80 90
4 5 6 Aj	Apli 4.1 4.2 4.3 Este 5.1 5.2 5.3 Con	icação à Modelag Neurofisiologia d 4.1.1 Campos Modelagem das Modelagem das reoscopia Introdução Estimação de D Experimentos . clusões e Trabalh	em Neuronal da Visão: O Caminho Visual	 56 58 59 61 70 70 73 80 90 94

96

Apêndice C	101
Apêndice D	105
Apêndice E	107
Referências	112

Capítulo 1

Introdução

O trabalho aqui apresentado tem como base uma proposta de representação de sinais inicialmente sugerida em [1], e que foi por nós estendida e aplicada a problemas em duas áreas distintas, a análise espaço-frequência (ou tempo-frequência) de sinais, e a modelagem neuronal.

A versão original da representação considerada surgiu no contexto de uma pesquisa em visão computacional, como modelo para o processo de desfocamento de imagens. Nela, um sinal espacial qualquer é expresso em termos de funções de representação bem localizadas na posição e na frequência, escolhidas sob a forma de funções de Gabor espaciais (senóides moduladas por Gaussianas) cujos parâmetros de largura e de fase são obtidos a partir da transformada de Fourier do sinal dado. Verificamos que este modelo pode ser também estendido para a representação de sinais no domínio da frequência: neste caso, as funções de codificação são funções de Gabor na frequência, com parâmetros dados pela magnitude e pela fase da transformada inversa do sinal codificado.

Ambas as versões das funções codificadoras — no espaço e na frequência — podem ser usadas como base para transformadas de Gabor, definindo o que chamamos de *transformadas sintonizadas de Gabor*, cuja 'sintonia' advém do fato de que cada sinal dado é analisado pelas funções de Gabor associadas à sua própria representação. De certa forma, o próprio sinal é a sua função analisadora, o que confere às transformadas sintonizadas propriedades semelhantes às da transformada de Wigner [2], considerada ótima sob vários aspectos [3].

Além da sua aplicação à análise de sinais, verificamos que o modelo de representação introduzido em [1], devidamente estendido para duas dimensões, pode ser empregado para descrever a codificação de imagens pelos campos receptivos das chamadas *células simples* do córtex visual, tradicionalmente modelados como funções de Gabor [4]. Resultados neurofisiológicos recentes têm indicado que, ao contrário do que anteriormente se supunha, o campo receptivo de um neurônio não é um arranjo espacial fixo, mas varia com o estímulo apresentado [5, 6, 7]. O fato de que os parâmetros das funções de representação no nosso modelo dependem do sinal codificado, o credencia como um potencial candidato para a descrição dos *campos receptivos dependentes do sinal*. Aqui exploramos esta possibilidade, propondo também uma nova versão do modelo 2D de representação sintonizada, baseada agora não nas funções de Gabor, mas em funções de codificação *rotacionalmente simétricas*, por nós empregadas para a modelagem dos campos receptivos de células da *retina* e do *núcleo geniculado lateral*, que apresentam tal simetria [8]. Seguindo a interpretação de [9, 10], nós assumimos que estes dois estágios pré-corticais do *caminho visual* – o trajeto seguido pela informação visual ao longo do cérebro – têm a função de descorrelacionar os sinais transmitidos ao córtex visual primário, de forma que as funções sintonizadas rotacionalmente simétricas codificam versões *branqueadas* das imagens de entrada.

Como aplicação dos nossos modelos de campos receptivos dependentes do estímulo, propusemos uma *abordagem de inspiração biológica* para a estimação de disparidades estereoscópicas, com base no método da função de Green, introduzido em [11]. Nessa abordagem, os campos receptivos circularmente simétricos são usados para branquear as imagens de entrada, refletindo a transformação efetuada pelos estágios pré-corticais do cérebro. A operação dos neurônios *binoculares* do córtex visual (neurônios que respondem à estimulação de ambos os olhos) é então modelada em termos dos campos receptivos sintonizados de Gabor. O algoritmo proposto mostra-se consistente com o modelo de Volterra, tradicionalmente empregado para descrever as respostas das células neuronais [12], e incorpora também o dado neurofisiológico da segregação dos hemicampos visuais no *quiasma óptico* — espécie de encruzilhada dos eixos ópticos provenientes dos dois olhos [8]. Com o seu emprego, conseguimos obter estimativas comparáveis, e mesmo superiores, às fornecidas pela abordagem original de [11].

Esta tese encontra-se organizada como segue. No restante deste capítulo, introduzimos as duas versões da representação de Gabor sintonizada que fundamenta o presente estudo, sendo a segunda delas uma das nossas contribuições originais. Nos capítulos seguintes, são apresentadas extensões e aplicações da abordagem sintonizada, tanto à análise quanto à síntese de sinais. Especificamente, o Capítulo 2 introduz as transformadas de Gabor sintonizadas nos domínios do espaço e da frequência, salientando as suas propriedades, e a aplicação à análise de sinais espaciais não-estacionários. O Capítulo 3, por sua vez, se detém num estudo mais detalhado do emprego da primeira forma da transformada sintonizada para a caracterização de epilepsias, a partir de sinais de EEG (eletro-encefalografia) [13]. No Capítulo 4, passamos a enfocar a abordagem sintonizada sob o aspecto da síntese de sinais, utilizando-a como ferramenta para a modelagem neural. Mais especificamente, consideramos as funções sintonizadas de Gabor como modelo para a codificação de sinais pelas células simples do córtex visual, e introduzimos funções sintonizadas rotacionalmente simétricas que desempenham um papel semelhante no caso das células da retina e do núcleo geniculado lateral — neste caso codificando versões branqueadas dos sinais de entrada. Finalmente, no Capítulo 5, fazendo uso dos resultados precedentes, nós propomos um algoritmo de estimação de disparidades estereoscópicas que emula o processamento neural nos estágios cortical e pré-cortical do caminho visual. O Capítulo 6 fecha o documento com a apresentação das nossas conclusões e propostas para trabalhos futuros.

No que se segue, e no restante desta tese, nós empregaremos as seguintes definições para a transformada de Fourier e a sua inversa:

Sendo I(x) um sinal unidimensional de quadrado integrável, definido num domínio ilimitado, a sua transformada de Fourier é dada por

$$\tilde{I}(\omega) = |\tilde{I}(\omega)|e^{i\varphi_{\tilde{I}}(\omega)} \equiv \mathcal{F}[I(x)](\omega) = \int_{-\infty}^{\infty} e^{-i\omega x} I(x) dx$$
(1.1)

e a sua transformada inversa é obtida como

$$I(x) = |I(x)|e^{i\varphi_I(x)} \equiv \mathcal{F}^{-1}[\tilde{I}(\omega)](x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{i\omega x} \tilde{I}(\omega) d\omega$$
(1.2)

Para sinais bidimensionais, valem as extensões usuais das fórmulas acima.

1.1 Primeira abordagem de representação

Conforme demonstrado em [1], qualquer sinal unidimensional, I(x), de quadrado integrável, pode ser representado como

$$I(x) = \int_{-\infty}^{\infty} e^{i\omega x} * \frac{1}{(2\pi)^{3/2}} e^{i[\omega x + \varphi_{\tilde{I}}(\omega)]} e^{-\frac{x^2}{2\sigma^2(\omega)}} d\omega$$
(1.3)

onde o asterisco denota uma convolução espacial. Na Equação (1.3), $\varphi_{\tilde{I}}(\omega)$ é a fase da transformada de Fourier do sinal, como já definido, e $\sigma(\omega)$ é o módulo da mesma transformada:

Como se demonstra no Apêndice A, a Equação (1.3) é equivalente à transformada de Fourier inversa. Assim como esta, ela expressa o sinal em termos de uma expansão em um conjunto de *funções de base* exponenciais complexas, mas os 'coeficientes' dessa expansão não são funções apenas da frequência ω , e sim núcleos espaciais de Gabor – portanto, funções relativamente bem localizadas, tanto da posição como da frequência – cujos parâmetros se relacionam à transformada de Fourier do sinal. Explicitamente, I(x)é expresso em termos de *funções codificadoras*, ou *de representação*, da forma

$$\psi_{\omega}(x) = \frac{1}{(2\pi)^{3/2}} e^{i[\omega x + \varphi_{\tilde{I}}(\omega)]} e^{-\frac{x^2}{2\sigma^2(\omega)}}$$
(1.5)

com os parâmetros $\varphi_{\tilde{I}} \in \sigma$ obtidos como acima.

Efetuando-se a convolução na Equação (1.3), ela pode ser reescrita como

$$I(x) = \int_{-\infty}^{\infty} e^{i\omega x} \left[\psi_{\omega}(x) \cdot e^{i\omega x} \right] d\omega$$
(1.6)

onde a operação entre colchetes é um produto interno. Comparando-a à forma padrão para uma expansão no conjunto de base de Fourier,

$$I(x) = \int_{-\infty}^{\infty} e^{i\omega x} \left[I(x) \cdot e^{i\omega x} \right] d\omega$$
(1.7)

verificamos que, no que se refere à frequência ω , a função $\psi_{\omega}(x)$ é equivalente ao sinal I(x), correspondendo a uma representação oscilatória, espacialmente bem localizada, do conteúdo do sinal naquela frequência. Isto pode ser tornado mais explícito se tomamos a transformada de Fourier de $\psi_{\omega}(x)$. Levando em conta que ω é um parâmetro fixo, nós calculamos a transformada em uma frequência genérica, ω' :

$$\mathcal{F}[\psi_{\omega}(x)](\omega') = \frac{1}{2\pi} \tilde{I}(\omega) e^{-\frac{|\tilde{I}(\omega)|^2}{2}(\omega'-\omega)^2}$$
(1.8)

e daí obtemos

$$\mathcal{F}[\psi_{\omega}(x)](\omega'=\omega) = \frac{1}{2\pi}\tilde{I}(\omega)$$
(1.9)

o que mostra que, na frequência ω , o conteúdo da função analisadora $\psi_{\omega}(x)$ é o mesmo do sinal. Assim, o conjunto { $\psi_{\omega}(x)$, $\forall \omega$ } fornece uma representação espaço-frequência completa de I(x), o que pode ser visto reescrevendo-se a Equação (1.3) como

$$I(x) = \frac{1}{(2\pi)^{3/2}} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{i[\omega x + \varphi_{\tilde{I}}(\omega)]} e^{-\frac{(x-\xi)^2}{2\sigma^2(\omega)}} d\omega d\xi$$
(1.10)

É interessante comparar este resultado com uma expansão de Gabor tradicional [14],

$$I(x) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} A(\omega,\xi) e^{i\omega x} e^{-\frac{(x-\xi)^2}{2\beta^2}} d\omega d\xi$$
(1.11)

que corresponde a uma soma ponderada de funções de Gabor cuja largura é escolhida *a priori* e mantida fixa para todas as frequências (o parâmetro β , acima, não depende de ω). Em contraste, a representação sintonizada utiliza funções de Gabor cuja largura e fase são determinadas pela transformada de Fourier do sinal representado. Ademais, em (1.10), todas as funções de codificação têm peso unitário. Se uma determinada frequência não está presente no sinal, é a própria função de codificação associada que se anula, já que para aquela frequência específica, $\sigma(\omega) = 0$, por definição.

A demonstração das identidades (1.3), (1.4), (1.6) e (1.10) é fornecida no Apêndice A.

1.1.1 Extensão 2-D

A representação da Equação (1.3) pode ser estendida para sinais bidimensionais, conforme demonstrado em [1].

Sendo I(x, y) um sinal bidimensional de quadrado integrável, vale a relação

$$I(x,y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{i[\omega_x x + \omega_y y]} * e^{i[\omega_x x + \omega_y y + \varphi_{\tilde{I}}(\omega_x, \omega_y)]} e^{-\frac{x^2 + y^2}{2\sigma^2(\omega_x, \omega_y)}} d\omega_x d\omega_y$$
(1.12)

onde $\varphi_{\tilde{I}}(\omega_x, \omega_y)$ é a fase da transformada de Fourier do sinal, e onde $\sigma(\omega_x, \omega_y)$ é obtido, a partir do módulo desta transformada, como

$$\sigma(\omega_x, \omega_y) = \frac{1}{(2\pi)^{3/2}} \sqrt{|\tilde{I}(\omega_x, \omega_y)|}$$
(1.13)

O sinal I(x, y) é então representado em termos de funções de representação bidimensionais dadas por:

$$\psi_{\omega_x,\omega_y}(x,y) = e^{i[\omega_x x + \omega_y y + \varphi_{\tilde{I}}(\omega_x,\omega_y)]} e^{-\frac{x^2 + y^2}{2\sigma^2(\omega_x,\omega_y)}}$$
(1.14)

com os parâmetros $\varphi_{\tilde{I}}$
e σ obtidos como acima.

Propriedades semelhantes às da representação unidimensional valem no caso 2-D, como pode ser facilmente verificado.

1.2 Segunda abordagem de representação

Uma representação semelhante à da Equação (1.3), mas agora no domínio da frequência, pode ser obtida para a *transformada de Fourier* de qualquer sinal de quadrado integrável, sob a forma:

$$\tilde{I}(\omega) = \int_{-\infty}^{\infty} e^{-i\omega x} * \frac{1}{(2\pi)^2} e^{-i[\omega x - \varphi_I(x)]} e^{-\frac{\omega^2}{2\Sigma^2(x)}} dx$$
(1.15)

onde o asterisco agora denota uma convolução em frequência, $\varphi_I(x)$ é a fase do sinal I(x)- ou seja, da trasformada de Fourier inversa de $\tilde{I}(\omega)$ -, e $\Sigma(x)$ se relaciona ao módulo deste sinal, como

$$\Sigma(x) = (2\pi)^{3/2} |I(x)| \tag{1.16}$$

Explicitando-se a convolução na Equação (1.15), ela também pode ser reescrita como uma expansão de Gabor com coeficientes unitários:

$$\tilde{I}(\omega) = \frac{1}{(2\pi)^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-i[\omega x - \varphi_I(x)]} e^{-\frac{(\omega - \Omega)^2}{2\Sigma^2(x)}} dx d\Omega$$
(1.17)

A Equação (1.15) expressa a transformada de Fourier do sinal I(x) em termos de uma expansão em funções de base exponenciais complexas, mas, em contraste com a transformada de Fourier tradicional, nessa expressão os 'coeficientes' não dependem apenas da posição, mas são funções de Gabor na frequência, da forma

$$\psi_x(\omega) = \frac{1}{(2\pi)^2} e^{-i[\omega x - \varphi_I(x)]} e^{-\frac{\omega^2}{2\Sigma^2(x)}}$$
(1.18)

As funções codificadoras $\psi_x(\omega)$ têm propriedades semelhantes às das $\psi_\omega(x)$, da primeira representação. Por exemplo, a Equação (1.15) pode ser reescrita como

$$\tilde{I}(\omega) = \int_{-\infty}^{\infty} e^{-i\omega x} \left[\psi_x(\omega) \cdot e^{-i\omega x} \right] dx$$
(1.19)

o que mostra que a função de representação $\psi_x(\omega)$ é localmente equivalente, na posição x, ao sinal I(x). Isto também pode ser comprovado tomando-se a transformada de Fourier inversa de $\psi_x(\omega)$ (novamente, levando em conta que x é aqui um parâmetro fixo, deve-se calcular a transformada em uma posição genérica, x'). Obtemos

$$\mathcal{F}^{-1}[\psi_x(\omega)](x') = \frac{1}{2\pi} I(x) e^{-4\pi^3 |I(x)|^2 (x'-x)^2}$$
(1.20)

e, portanto,

$$\mathcal{F}^{-1}[\psi_x(\omega)](x'=x) = \frac{1}{2\pi}I(x)$$
(1.21)

As demonstrações das identidades em (1.15), (1.16)
e (1.19) aparecem no Apêndice A.

1.2.1 Extensão 2-D

Uma extensão bidimensional para a representação da Equação (1.15) também pode ser facilmente obtida sob a forma:

$$\tilde{I}(\omega_x,\omega_y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-i[\omega_x x + \omega_y y]} * \frac{1}{(2\pi)^2} e^{-i[\omega_x x + \omega_y y - \varphi_I(x,y)]} e^{-\frac{\omega_x^2 + \omega_y^2}{2\Sigma^2(x,y)}} dxdy$$
(1.22)

 $\operatorname{com} \Sigma(x, y)$ agora dado por

$$\Sigma(x,y) = (2\pi)^{3/2} \sqrt{|I(x,y)|}$$
(1.23)

onde $I(x,y) \equiv |I(x,y)|e^{i\varphi_I(x,y)}$ é a transformada de Fourier inversa de $\tilde{I}(\omega_x, \omega_y)$.

O sinal $\tilde{I}(\omega_x, \omega_y)$ é então representado em termos de funções de representação bidimensionais dadas por:

$$\psi_{x,y}(\omega_x,\omega_y) = \frac{1}{(2\pi)^2} e^{-i[\omega_x x + \omega_y y - \varphi_I(x,y)]} e^{-\frac{\omega_x^2 + \omega_y^2}{2\Sigma^2(x,y)}}$$
(1.24)

com os parâmetros φ_I e Σ obtidos como acima.

Capítulo 2

Aplicação à Análise de Sinais

2.1 Introdução

A transformada de Fourier informa o conteúdo em frequência de um dado sinal – ou seja, o peso de cada frequência em sua composição –, mas este é obtido globalmente, pela soma das contribuições coletadas ao longo de toda a extensão do sinal. Deste modo, qualquer informação local (sobre *onde* cada componente de frequência teria maior ou menor importância) se perde na transformada. Dada uma função da posição, I(x), a transformada de Fourier retorna uma função apenas da frequência, $\tilde{I}(\omega)$.

As transformadas espaço-frequência surgiram como uma alternativa à transformada de Fourier, introduzindo uma certa medida de localidade na análise do sinal [3, 15, 16]. A idéia básica é a de obter a transformada de Fourier sob janelas locais, de modo que posições distintas do sinal possam ser analisadas separadamente. A forma geral de uma transformada de Fourier janelada é a seguinte:

$$\tilde{I}(\omega, x) = \int_{-\infty}^{\infty} e^{-i\omega\xi} \rho(\xi - x) I(\xi) d\xi$$
(2.1)

onde $\rho(\xi)$ é uma função espacial bem localizada em torno da origem, ou seja, em torno de $\xi = 0$. A transformada $\tilde{I}(\omega, x)$ é portanto a transformada de Fourier do produto $\rho(\xi - x)I(\xi)$. Para cada valor de x, $\rho(\xi - x)$ está centrada em $\xi = x$, de forma que a função ρ , na Equação (2.1), agirá como uma janela deslizante, permitindo que a transformada de Fourier analise diferentes posições do sinal sucessivamente.

A transformada de Gabor [17] é o caso particular da transformada janelada em que a

janela é uma Gaussiana. A sua expressão geral pode ser dada como

$$G(\omega, x) = \int_{-\infty}^{\infty} e^{-i\omega\xi} e^{-\frac{(\xi-x)^2}{2\sigma^2}} I(\xi) d\xi$$
(2.2)

onde σ é o desvio-padrão da Gaussiana. A expressão acima deixa claro que os núcleos analisadores da transformada, neste caso, são funções de Gabor, ou seja, senóides moduladas por Gaussianas. Em contraste com a transformada de Fourier, cujos núcleos são funções de frequência definida mas de extensão ilimitada, a de Gabor analisa cada sinal por meio de funções limitadas tanto no espaço quanto na frequência. Aliás, é sabido que as funções complexas de Gabor são aquelas que possibilitam máxima localização simultânea no espaço e na frequência [4].

Na prática, um problema que surge com a transformada de Gabor e, de forma geral, com todas as transformadas de Fourier janeladas, é o da escolha da largura da janela, pois não é possível determinar *a priori* em que resolução se encontram as características relevantes de um sinal. Assim, o tamanho da janela é um parâmetro livre, que deve ser ajustado a cada experimento. A utilização, como núcleos de transformadas de Gabor, das funções de representação introduzidas no capítulo anterior pode representar uma contribuição importante para a solução deste problema, já que os parâmetros dessas funções são intrinsicamente relacionados a cada sinal analisado.

No que se segue, introduziremos duas versões deste tipo de transformada sintonizada ao sinal: uma para a análise de sinais espaciais, baseada na primeira função de representação, $\psi_{\omega}(x)$, e a outra para análise de sinais no domínio da frequência, baseada na segunda função de representação, $\psi_x(\omega)$, conforme definidas no Capítulo 1 (Equações (1.5) e (1.18), respectivamente). Tomadas em conjunto, essas duas transformadas apresentam propriedades semelhantes às da transformada de Wigner (considerada ideal para a análise espaço-frequência, e que será definida na seção abaixo), evitando porém o chamado problema dos termos cruzados [2], característico desta última.

2.2 A Transformada de Gabor Sintonizada

2.2.1 Primeira Forma da Transformada: $T(\omega, x)$

Considerando um sinal espacial arbitrário, I(x), iremos analisá-lo sob o conjunto das funções $\psi_{\omega}(x)$, da Equação (1.5), definindo a transformada de Gabor sintonizada, $T(\omega, x)$,

como

$$T(\omega, x) = \frac{1}{(2\pi)^{3/2}} \int_{-\infty}^{\infty} e^{-i[\omega\xi + \varphi_{\tilde{I}}(\omega)]} e^{-\frac{(\xi-x)^2}{2\sigma^2(\omega)}} I(\xi) d\xi$$
(2.3)

Para o cálculo dessa transformada, primeiro aplicamos a transformada de Fourier sobre o sinal e obtemos os parâmetros $\varphi_{\tilde{I}}(\omega) \in \sigma(\omega)$, das funções de representação (neste caso também chamadas de funções analisadoras), para cada frequência ω . Relembrando: $\varphi_{\tilde{I}}(\omega)$ é a fase da transformada de Fourier do sinal analisado, e $\sigma(\omega)$ é o módulo desta transformada.

Deve-se notar, na Equação (2.3), que, similarmente ao que ocorre na transformada S[18], a parte senoidal da função analisadora não varia com a posição¹, mas aqui incorpora um deslocamento de fase dependente da frequência (mantemos este fator no integrando, para salientar a relação entre a função analisadora e a função codificadora do sinal analisado). Portanto, a transformada sintonizada $T(\omega, x)$ pode ser interpretada como a correlação "com fase corrigida"[18] entre o sinal analisado e a sua função codificadora na frequência ω . Esta transformada apresenta diversas propriedades interessantes. Pode-se demonstrar, por exemplo, que a densidade de energia em frequência do sinal analisado, $|\tilde{I}(\omega)|^2$, é obtida como

$$\left|\tilde{I}(\omega)\right|^{2} = 2\pi \int_{-\infty}^{\infty} T(\omega, x) dx$$
(2.4)

de onde também resulta

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} T(\omega, x) d\omega dx = \frac{1}{2\pi} \int_{-\infty}^{\infty} |\tilde{I}(\omega)|^2 d\omega = \int_{-\infty}^{\infty} |I(x)|^2 dx$$
(2.5)

usando-se o teorema de Parseval, para a segunda igualdade.

Lembrando que $|\tilde{I}(\omega)|^2$ é a transformada de Fourier da auto-correlação espacial, definida como

$$(I \star I)(x) = \int_{-\infty}^{\infty} I^*(\xi) I(\xi + x) d\xi$$
 (2.6)

onde o asterisco $<^*>$ denota o conjugado complexo, também obtemos, a partir da Equação (2.4),

$$(I \star I)(x) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{i\omega x} T(\omega, \xi) d\omega d\xi$$
(2.7)

¹A transformada S de um sinal I(x) fica definida como

$$S(\omega, x) = \frac{|\omega|}{(2\pi)^{3/2}} \int_{-\infty}^{\infty} e^{-i\omega\xi} e^{-\frac{(\xi-x)^2\omega^2}{2(2\pi)^2}} I(\xi) d\xi$$

e corresponde, portanto, à transformada de Fourier sob uma janela Gaussiana cuja largura é inversamente proporcional à frequência. A ausência da posição x na exponencial complexa faz com que a fase da transformada S não seja localmente referenciada, o que se interpreta como uma "correção de fase".

As demonstrações dos resultados acima encontram-se no Apêndice B, Equações (B.13) e (B.14).

Como vimos, a transformada de Gabor sintonizada, num certo sentido, utiliza o próprio sinal como sua função analisadora, o que lhe confere propriedades semelhantes às da transformada de Wigner [3],

$$W(\omega, x) = \int_{-\infty}^{\infty} I^*(x - \xi/2) I(x + \xi/2) e^{-i\omega\xi} d\xi$$
 (2.8)

A transformada de Wigner é considerada ideal sob o ponto de vista da análise espaçofrequência [2], mas, devido ao fato de ser quadrática no sinal, ela gera termos de frequência espúrios, quando da análise de sinais multicomponentes — é o chamado problema dos termos cruzados, que não ocorre com a transformada de Gabor sintonizada. Isto pode ser ilustrado pela análise de um sinal simples, composto pela adição de duas senóides:

$$I_1(x) = \cos\left(\frac{8\pi}{100}x\right) + \cos\left(\frac{16\pi}{100}x\right)$$
(2.9)

O sinal $I_1(x)$ é apresentado na Figura 2.1a, na Figura 2.1b é apresentado o espectrograma $|T(\omega, x)|$ e, em 2.1c, a magnitude da sua transformada de Wigner, $|W(\omega, x)|$. A transformada de Gabor sintonizada detecta as duas componentes de frequência com boa resolução, e evita o problema dos termos cruzados, que é evidente na Figura 2.1c: frequências que não existem no sinal são detectadas pela transformada de Wigner.

A transformada de Gabor sintonizada, assim como a transformada S [18], analisa os sinais de entrada por meio de funções de Gabor cuja largura depende da frequência. A transformada S mantém a largura inversamente proporcional à frequência [18], ao passo que, na Gabor sintonizada, a largura está associada à magnitude de Fourier de cada sinal analisado. Como consequência, a transformada $T(\omega, x)$ apresenta a vantagem de reduzir os efeitos de *self-aliasing* [18], o que pode ser observado quando reescrevemos a sua magnitude sob a forma (a demonstração se encontra no Apêndice B, Equação (B.4)):

$$|T(\omega, x)| = \frac{1}{(2\pi)^2} \left| \int_{-\infty}^{\infty} \tilde{I}(\Omega + \omega) W_{\omega}(\Omega) e^{i\Omega x} d\Omega \right|$$
(2.10)

onde

$$W_{\omega}(\Omega) = |\tilde{I}(\omega)| e^{-\frac{|\tilde{I}(\omega)|^2}{2}\Omega^2}$$
(2.11)

A ocorrência de *self-aliasing* torna-se evidente quando consideramos a Equação (2.10) para sinais discretizados: a amostragem espacial de tais sinais produzirá espectros discretos e periódicos, ocasionando a sobreposição da janela W_{ω} na porção de frequência negativa de



Figura 2.1: Em (a), o sinal $I_1(x)$. Em (b), o seu espectrograma $|T(\omega, x)|$. Em (c), a magnitude da sua transformada de Wigner, $|W(\omega, x)|$.

 \tilde{I} . No entanto, de acordo com a equação (2.11), uma janela $W_{\omega}(\Omega)$ larga estará associada a uma baixa amplitude, e, inversamente, grandes amplitudes estarão associadas apenas a janelas estreitas. Assim, a contribuição dos termos de sobreposição para o cômputo da transformada de Gabor sintonizada será reduzida. O mesmo não sucede no caso da transformada S, em que o *self-aliasing* pode se tornar grave [18], especialmente nas frequências mais altas.²

Para alguns sinais simples, é possível obter a transformada $T(\omega, x)$ analiticamente. Apresentaremos, a seguir, alguns exemplos, cuja demonstração aparece no Apêndice B:

Transformada da delta de Dirac:

Quando $I(x) = \delta(x - x_0),$

$$T(\omega, x) = \frac{1}{(2\pi)^{3/2}} e^{-\frac{1}{2}(x-x_0)^2}$$
(2.12)

ou seja, a transformada $T(\omega, x)$ provoca um pequeno espalhamento do impulso, que se torna uma Gaussiana com desvio-padrão unitário. A Figura 2.2 ilustra esse comportamento para o sinal $I_2(x) = \delta(x - 200)$, ou seja, uma delta centrada na posição 200.



Figura 2.2: Em (a), sinal $I_2(x) = \delta(x - 200)$. Em (b), o seu espectrograma $|T(\omega, x)|$.

²Isto se explica porque, no caso da transformada S, vale uma relação semelhante à Equação (2.10), mas com a janela obtida como $W_{\omega}(\Omega) = e^{-\frac{8\pi^4 \Omega^2}{\omega^2}}$ (ver [18])

Transformada do Trem de Impulsos:

Seja

$$I(x) = \sum_{k} \delta(x - x_k) \tag{2.13}$$

um trem de impulsos, com $x_k = kL$, para k inteiro $\in \{0, N\}$, e N grande o suficiente, tal que I(x) possa ser considerado um sinal periódico de período L e frequência fundamental $\omega_0 = 2\pi/L$. Obtém-se, para a sua transformada de Gabor sintonizada,

$$T(\omega, x) = \begin{cases} 0, & \omega \neq \omega_n, x \neq x_m \\ \frac{1}{(2\pi)^{3/2}} e^{-i\omega x_m}, & \omega \neq \omega_n, x = x_m \\ \frac{1}{(2\pi)^{3/2}} \sum_k 1, & \omega = \omega_n, \forall x \end{cases}$$
(2.14)

onde $\omega_n = n\omega_0$, para qualquer inteiro *n*. Isto significa que a transformada de Gabor sintonizada é capaz de detectar as singularidades do sinal, tanto no espaço quanto na frequência, conforme ilustrado na Figura 2.3 para o sinal $I_3(x) = \Sigma_k \delta(x - kN)$, com $k \in \{0, 33\}$ e N = 30.



Figura 2.3: Em (a), sinal $I_3(x) = \sum_k \delta(x - kL)$. Em (b), o seu espectrograma $|T(\omega, x)|$.

Transformada da exponencial complexa:

Quando $I(x) = e^{i\omega_0 x}$,

$$T(\omega, x) = \delta(\omega - \omega_0) \equiv \frac{1}{2\pi} \tilde{I}(\omega)$$
(2.15)

ou seja, a transformada $T(\omega, x)$ da exponencial complexa é proporcional à transformada de Fourier, não ocorrendo espalhamento em frequência. A Figura 2.4 ilustra este comportamento para o sinal $I_4(x) = e^{i(\pi/50)x}$.



Figura 2.4: Espectrograma $|T(\omega, x)|$ do sinal $I_4(x) = e^{i(\pi/50)x}$.

É importante salientar também que a transformada de Gabor sintonizada admite inversa, ou seja, a partir da transformada de Gabor sintonizada é possível recuperar a transformada de Fourier do sinal, desde que se conheça a fase desta última em uma dada frequência. Demonstramos no Apêndice B que, para uma determinada frequência ω_0 ,

$$\tilde{I}(\omega) = 2\pi \frac{e^{i\varphi_{\tilde{I}}(\omega_0)} \mathcal{F}[T(\omega_0, x)](\omega - \omega_0)}{|\tilde{I}(\omega_0)| e^{-\frac{|\tilde{I}(\omega_0)|^2}{2}(\omega - \omega_0)^2}}$$
(2.16)

onde

$$|\tilde{I}(\omega_0)|^2 = 2\pi \mathcal{F}[T(\omega_0, x)](\omega = 0)$$
(2.17)

Finalmente, verifica-se que é possível obter uma expressão alternativa para a transformada de Gabor sintonizada $T(\omega, x)$, fazendo uso do teorema de Parseval:

$$\int_{-\infty}^{\infty} f(x)h^*(x)dx = \frac{1}{2\pi} \int_{-\infty}^{\infty} \tilde{F}(\omega)\tilde{H}^*(\omega)d\omega$$
(2.18)

Comparando a relação acima com a Equação (2.3), esta pode ser reescrita como:

$$T(\omega, x) = \frac{1}{\sqrt{2\pi}} |\tilde{I}(\omega)| e^{-i\omega x} \int_{-\infty}^{\infty} e^{i[\Omega x - \varphi_{\tilde{I}}(\omega)]} e^{-\frac{|\tilde{I}(\omega)|^2}{2}(\Omega - \omega)^2} \tilde{I}(\Omega) d\Omega$$
(2.19)

Neste caso, a $T(\omega, x)$ fica expressa como uma transformada no domínio da frequência.

A seguir, nós apresentaremos uma série de outras propriedades analíticas da $T(\omega, x)$ que têm correspondência com propriedades da transformada de Wigner. As demonstrações de todas essas propriedades aparecem no apêndice B.

2.2.2 Propriedades analíticas da $T(\omega, x)$:

2.2.2.1 Momentos da frequência:

A partir da Equação (2.4), obtemos

$$\int_{-\infty}^{\infty} \omega^n |\tilde{I}(\omega)|^2 d\omega = 2\pi \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \omega^n T(\omega, x) d\omega dx$$
(2.20)

2.2.2.2 Translação Espacial:

Se $I_2(x) = I(x - x_0)$,

$$T_{I_2}(\omega, x) = T(\omega, x - x_0) \tag{2.21}$$

onde T_{I_2} denota a Transformada de Gabor Sintonizada do sinal $I_2(x)$.

2.2.2.3 Translação na Frequência:

Se
$$I_2(x) = e^{i\omega_0 x} I(x),$$

 $T_{I_2}(\omega, x) = T(\omega - \omega_0, x)$

$$(2.22)$$

Quando o sinal de entrada é modulado por uma exponencial complexa, o efeito sobre a transformada é apenas uma translação na frequência.

2.2.2.4 Primeiro momento condicional do espaço:

Por definição [3], o primeiro momento condicional do espaço é dado por

$$\langle x \rangle_{\omega} = \frac{1}{|\tilde{I}(\omega)|^2} \int_{-\infty}^{\infty} x \ T(\omega, x) dx$$
 (2.23)

o que resulta em

$$\langle x \rangle_{\omega} = \frac{1}{2\pi} \frac{d}{d\omega} \left[-\varphi_{\tilde{I}}(\omega) + i \log |\tilde{I}(\omega)| \right]$$
 (2.24)

onde $\varphi_{\tilde{i}}(\omega)$ é a fase da transformada de Fourier do sinal.

Portanto, $\langle x \rangle_{\omega}$ fornece o *retardo de grupo* do sinal, $\varphi'_{\tilde{I}}(\omega) \equiv -\frac{d}{d\omega}\varphi_{\tilde{I}}(\omega)$, como

$$-\varphi_{\tilde{I}}'(\omega) = 2\pi Re \langle x \rangle_{\omega} \tag{2.25}$$

2.2.2.5 Momento cruzado do espaço e da frequência:

Este é definido como

$$\langle x\omega \rangle = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \omega x \ T(\omega, x) d\omega dx$$
 (2.26)

de onde se obtém

$$\langle x\omega \rangle = \int_{-\infty}^{\infty} \xi \varphi_I'(\xi) |I(\xi)|^2 d\xi + i \int_{-\infty}^{\infty} \xi |I(\xi)| \frac{d}{d\xi} |I(\xi)| d\xi$$
(2.27)

e logo,

$$\langle x\varphi_I'(x) \rangle \equiv \int_{-\infty}^{\infty} \xi \varphi_I'(x) |I(\xi)|^2 d\xi = \operatorname{Re} \langle x\omega \rangle$$
 (2.28)

onde $\varphi_I(x)$ é a fase do sinal – ou seja, $I(x) = |I(x)|e^{i\varphi_I(x)}$ –, e, portanto, $\varphi'_I(x) = \frac{d}{dx}\varphi_I(x)$ é a sua frequência instantânea.

A grandeza $\langle x \varphi'_I(x) \rangle$ está relacionada à covariância do sinal, definida como [3]

$$\operatorname{Cov}_{x\omega} = \langle x\varphi'_{\tilde{I}}(x) \rangle - \langle x \rangle \langle \omega \rangle$$
(2.29)

onde < x > e < ω > representam a posição e a frequência médias do sinal, definidas, respectivamente, como:

$$\langle x \rangle = \int_{-\infty}^{\infty} |I(x)|^2 x dx \tag{2.30}$$

е

$$<\omega>=\int_{-\infty}^{\infty}|\tilde{I}(\omega)|^{2}\omega d\omega$$
 (2.31)

A abordagem sintonizada de Gabor pode ser estendida para permitir a análise de um sinal de entrada pelas funções de codificação da sua derivada. Relembrando que $\mathcal{F}\left[\frac{d}{dx}I(x)\right](\omega) = i\omega\tilde{I}(\omega)$, a nova transformada fica definida por

$$\hat{T}(\omega, x) = \frac{1}{(2\pi)^{3/2}} \int_{-\infty}^{\infty} e^{-i[\omega\xi + \varphi_{\bar{I}}(\omega)]} e^{-\frac{(\xi - x)^2}{2\omega^2 \sigma^2(\omega)}} I(\xi) d\xi$$
(2.32)

desconsiderando-se o fator constante de fase, $e^{i\frac{\pi}{2}}$. A abordagem da derivada pode se mostrar útil quando se deseja realçar componentes de alta frequência e baixa energia no sinal³.

³Observe-se que a forma $\hat{T}(\omega, x)$ constitui uma transformada sintonizada cruzada: o sinal de entrada é analisado pelas funções de representação de um outro sinal — neste caso, a sua derivada.

A partir da $\hat{T}(\omega, x)$, obtemos a densidade de energia em frequência sob a forma

$$\left|\tilde{I}(\omega)\right|^{2} = 2\pi\omega \int_{-\infty}^{\infty} \hat{T}(\omega, x) dx$$
(2.33)

e é possível mostrar que a transformada $\hat{T}(\omega, x)$ da função $\delta(x - x_0)$ torna-se $\frac{1}{(2\pi)^{3/2}}e^{-\frac{1}{2\omega^2}(x-x_0)^2}$, e a da exponencial complexa, $e^{i\omega_0 x}$, fornece $\omega_0\delta(\omega - \omega_0)$. Portanto, um impulso espacial se espalha numa Gaussiana cuja largura aumenta linearmente com a magnitude da frequência, enquanto um impulso em frequência é preservado pela transformada, porém com intensidade também proporcional à frequência. A transformada $\hat{T}(\omega, x)$ também admite inversa, que se obtém de uma forma semelhante à Eq.(2.16), substituindo-se $\tilde{I}(\omega_0)$ por $\omega_0 \tilde{I}(\omega_0)$ ali.

2.2.3 Segunda Forma da Transformada: $T^{(2)}(\omega, x)$

Utilizando-se a segunda forma de função de representação apresentada no Capítulo 1 $(\psi_x(\omega), \text{ conforme a Eq. (1.18)})$, é possível introduzir uma transformada sintonizada de Gabor adequada à análise de sinais definidos no domínio da frequência. Esta é dada por

$$T^{(2)}(\omega, x) = \frac{1}{(2\pi)^2} \int_{-\infty}^{\infty} e^{i[\Omega x - \varphi_I(x)]} e^{-\frac{(\Omega - \omega)^2}{2\Sigma^2(x)}} \tilde{I}(\Omega) d\Omega$$
(2.34)

onde os parâmetros da função analisadora são agora funções da posição, obtidas a partir da magnitude e da fase da transformada inversa do sinal de entrada. Explicitamente, $\varphi_I(x)$ é a fase da transformada de Fourier inversa de $\tilde{I}(\omega)$ – ou seja, I(x) –, e $\Sigma(x)$ relaciona-se ao módulo desta como $\Sigma(x) = (2\pi)^{3/2} |I(x)|$.

A transformada $T^{(2)}(\omega, x)$ também apresenta uma série de propriedades comuns à transformada de Wigner, mas que não são compartilhadas pela $T(\omega, x)$. De modo geral, verifica-se que metade das propriedades analíticas da transformada de Wigner são replicadas por uma das transformadas de Gabor sintonizadas, enquanto a metade restante é satisfeita pela outra transformada. Por exemplo, a transformada $T^{(2)}(\omega, x)$ da função $\delta(x - x_0)$ é esta própria função, e a integral de $T^{(2)}(\omega, x)$ sobre todas frequências fornece a densidade espacial de energia do sinal, $|I(x)|^2$ (ver abaixo).

Na Figura 2.5, apresentamos o espectrograma $|T^{(2)}(\omega, x)|$ obtido a partir do sinal $\tilde{I}_1(\omega) = \mathcal{F}\{I_1(x)\}$, onde $I_1(x)$ é dado pela Equação (2.9). Verifica-se um padrão sobreposto às retas horizontais correspondentes a cada uma das frequências do sinal (comparar ao espectro obtido com a primeira transformada sintonizada, $T(\omega, x)$, na Fig. 2.1b, e com a transformada de Wigner, na Fig. 2.1c). Esse comportamento deve-se à grande variação em amplitude de $I_1(x)$: como as frequências são analisadas por funções cuja largura é proporcional à magnitude do sinal espacial (Eq.(1.16)), a resolução da transformada será tanto pior quanto maior for esta. Os picos sobre as linhas horizontais correspondem, portanto, aos máximos do sinal $I_1(x)$. Ao preço de uma menor resolução em frequência, a transformada $T^{(2)}(\omega, x)$ preserva informação espacial do sinal analisado.

A Figura 2.6 apresenta o espectrograma $|T^{(2)}(\omega, x)|$ para a função delta da Figura 2.2a. Neste caso, a resolução espacial é superior à obtida com a transformada T (Figura 2.1b). Já a Figura 2.7 apresenta o espectrograma obtido, com a $T^{(2)}$, para o sinal exponencial complexo da Figura 2.4, podendo-se observar que a resolução em frequência é inferior à conseguida com a transformada T. Este comportamento é típico das duas transformadas: a forma T em geral proporciona uma melhor resolução espectral, e a $T^{(2)}$, uma melhor



Figura 2.5: Espectrograma $|T^{(2)}(\omega, x)|$ do sinal $I_1(x)$.

resolução espacial.



Figura 2.6: Espectrograma $|T^{(2)}(\omega, x)|$ do sinal $I_2(x)$.



Figura 2.7: Espectrograma $|T^{(2)}(\omega, x)|$ do sinal $I_4(x)$.

Assim como a transformada T, a transformada sintonizada $T^{(2)}$ também admite uma inversa. Dada uma posição x_0 qualquer, verifica-se que

$$I(x) = \frac{\mathcal{F}^{-1}[e^{-i[\omega x_0 - \varphi_I(x_0)]}T^{(2)}(\omega, x_0)](x)}{|I(x_0)|e^{-4\pi^3|I(x_0)|^2(x-x_0)^2}}$$
(2.35)

onde

$$|I(x_0)|^2 = \mathcal{F}^{-1}[T^{(2)}(\omega, x_0)](x=0)$$
(2.36)

Este resultado indica que é possível recuperar o sinal I(x), a partir dos valores da sua transformada de Gabor sintonizada $T^{(2)}$ para uma posição fixa qualquer, desde que se
conheça a fase do sinal naquela posição.

Usando o teorema de Parseval, também é possível reescrever a $T^{(2)}$ como uma transformada no domínio espacial:

$$T^{(2)}(\omega,x) = (2\pi)^2 |I(x)| e^{i\omega x} \int_{-\infty}^{\infty} e^{-i[\omega\xi + \varphi_I(x)]} e^{-4\pi^3 |I(x)|^2(\xi-x)^2} I(\xi) d\xi$$
(2.37)

No que se segue, apresentaremos uma série de propriedades da transformada sintonizada $T^{(2)}$ que são correlatas a propriedades da Transformada de Wigner. As demonstrações aparecem no Apêndice B.

2.2.4 Propriedades analíticas da $T^{(2)}(\omega, x)$:

2.2.4.1 Transformada da delta de Dirac:

Quando $I(x) = \delta(x - x_0),$

$$T^{(2)}(\omega, x) = \delta(x - x_0)$$
(2.38)

Ou seja, não há espalhamento do impulso, pela transformada $T^{(2)}(\omega, x)$.

2.2.4.2 Transformada do trem de impulsos:

Seja

$$I(x) = \sum_{k} \delta(x - x_k) \tag{2.39}$$

um trem de impulsos, com $x_k = kL$, para k inteiro $\in \{0, N\}$ e N grande o suficiente tal que I(x) possa ser considerado um sinal periódico de período L e frequência fundamental $\omega_0 = 2\pi/L$. Obtém-se

$$T^{(2)}(\omega, x) = \begin{cases} 0, & \omega \neq \omega_n, x \neq x_m \\ \frac{1}{(2\pi)L} e^{i\omega_n x}, & \omega = \omega_n, x \neq x_m \\ \frac{1}{(2\pi)L} \sum_k 1, & \forall \omega, x = x_m \end{cases}$$
(2.40)

onde $\omega_n = n\omega_0$, para qualquer inteiro n. Isso significa que a transformada de Gabor sintonizada $T^{(2)}$ também é capaz de detectar as singularidades do sinal, tanto no espaço como na frequência.

2.2.4.3 Transformada da exponencial complexa:

Quando $I(x) = e^{i\omega_0 x}$,

$$T^{(2)}(\omega, x) = \frac{1}{(2\pi)} e^{-\frac{(\omega-\omega_0)^2}{16\pi^3}}$$
(2.41)

Ou seja, há um espalhamento Gaussiano na frequência, com desvio-padrão de $(2\pi)^{\frac{3}{2}} = 15,75.$

2.2.4.4 Densidade de Energia:

A densidade de energia no espaço, $|I(x)|^2$, pode ser obtida como

$$|I(x)|^{2} = \frac{1}{2\pi} \int_{-\infty}^{\infty} T^{(2)}(\omega, x) d\omega$$
 (2.42)

de onde também resulta

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} T^{(2)}(\omega, x) d\omega dx = 2\pi \int_{-\infty}^{\infty} |I(x)|^2 dx = \int_{-\infty}^{\infty} |\tilde{I}(\omega)|^2 d\omega$$
(2.43)

usando-se o teorema de Parseval, para a segunda igualdade.

2.2.4.5 Correlação em frequência:

Com
o $(2\pi)|I(x)|^2$ é a transformada inversa de Fourier da correlação em frequência,

$$(\tilde{I} \star \tilde{I})(\omega) = \int_{-\infty}^{\infty} \tilde{I}^*(\Omega)\tilde{I}(\Omega + \omega)d\Omega$$
(2.44)

resulta, a partir da Equação (2.42),

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-i\omega x} T^{(2)}(\Omega, x) d\Omega dx = (\tilde{I} \star \tilde{I})(\omega)$$
(2.45)

2.2.4.6 Momentos da posição:

Ainda usando a Equação (2.42), obtemos

$$\int_{-\infty}^{\infty} x^n |I(x)|^2 dx = \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x^n T^{(2)}(\omega, x) d\omega dx$$
(2.46)

2.2.4.7 Translação Espacial:

Se $I_2(x) = I(x - x_0),$

$$T_{I_2}^{(2)}(\omega, x) = T^{(2)}(\omega, x - x_0)$$
(2.47)

onde $T_{I_2}^{(2)}$ é a transformada de Gabor sintonizada, $T^{(2)}$, para o sinal I_2 .

2.2.4.8 Translação na Frequência:

Se $I_2(x) = e^{i\omega_0 x} I(x),$

$$T_{I_2}^{(2)}(\omega, x) = T^{(2)}(\omega - \omega_0, x)$$
(2.48)

2.2.4.9 Primeiro momento condicional da frequência:

Por definição [3], o primeiro momento condicional da frequência é dado por

$$\langle \omega \rangle_x = \frac{1}{|I(x)|^2} \int_{-\infty}^{\infty} \omega \ T^{(2)}(\omega, x) d\omega$$
 (2.49)

o que resulta em

$$\langle \omega \rangle_x = 2\pi \frac{d}{dx} \left[\varphi_I(x) - i \log |I(x)| \right]$$
 (2.50)

onde $\varphi_{\tilde{I}}(\omega)$ é a fase da transformada de Fourier do sinal.

Portanto, < $\omega>_x$ fornece a frequência instantânea do sinal, $\varphi_I'(x),$ como

$$\varphi_I'(x) = \frac{1}{2\pi} \operatorname{Re}\{\langle \omega \rangle_x\}$$
(2.51)

2.2.4.10 Momento cruzado da posição e da frequência:

Este é definido como

$$\langle x\omega \rangle = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \omega x \ T^{(2)}(\omega, x) d\omega dx$$
 (2.52)

de onde se obtém

$$\langle x\omega \rangle = -\int_{-\infty}^{\infty} \Omega \varphi_{\tilde{I}}'(\Omega) |\tilde{I}(\Omega)|^2 d\Omega - i \int_{-\infty}^{\infty} \Omega |\tilde{I}(\Omega)| \frac{d}{d\Omega} |\tilde{I}(\Omega)| d\Omega$$
(2.53)

e logo,

$$\langle \omega \varphi_{\tilde{I}}'(\omega) \rangle \equiv \int_{-\infty}^{\infty} \Omega \varphi_{\tilde{I}}'(\Omega) |\tilde{I}(\Omega)|^2 d\Omega = -\text{Re}\{\langle x\omega \rangle\}$$
 (2.54)

onde $\varphi'_{\tilde{I}}(\omega)$ é o retardo de grupo do sinal.

A grandeza $\langle \omega \varphi'_{\tilde{\iota}}(\omega) \rangle$ está relacionada à covariância do sinal, definida como [3]

$$\operatorname{Cov}_{x\omega} = - \langle \omega \varphi'_{\tilde{I}}(\omega) \rangle - \langle x \rangle \langle \omega \rangle$$

$$(2.55)$$

Conforme já mencionado, as duas versões das transformadas de Gabor sintonizadas reproduzem algumas das propriedades mais relevantes da função de distribuição de Wigner [2]. Por exemplo, enquanto a $T(\omega, x)$ satisfaz a propriedade de projeção no domínio da frequência (Equação (2.4)), a $T^{(2)}(\omega, x)$ satisfaz a propriedade correspondente no domínio espacial (Equação (2.42)). A primeira satisfaz a propriedade da média condicional no espaço (Equação (2.25)), enquanto a segunda satisfaz a propriedade da média condicional na frequência (Equação (2.51)). O mesmo se aplica para as propriedades de momento (Equações (2.20) e (2.46)), e as propriedades para a recuperação do sinal (Equações $(2.16) \in (2.35)$). De forma geral, uma abordagem da transformada de Gabor sintonizada irá satisfazer uma propriedade da distribuição da Wigner no domínio espacial, enquanto a outra irá satisfazer a propriedade correspondente no domínio da frequência. Como, do ponto de vista teórico, a transformada da Wigner é considerada ideal para a análise espaçofrequência, os nossos resultados indicam que a abordagem de Gabor sintonizada pode se tornar uma ferramenta importante para análise de sinais não estacionários, levando-se em conta, sobretudo, que ela não apresenta o problema dos termos cruzados, característico da transformada da Wigner.

É fácil verificar, também, que permanecem válidas, no caso da $T^{(2)}(\omega, x)$, as nossas observações sobre a questão do *self-aliasing*, e a extensão da transformada para a análise de um sinal a partir das funções de representação da sua derivada. Nesta tese, no entanto, nos concentraremos no estudo da primeira forma da transformada de Gabor sintonizada, a forma $T(\omega, x)$. No que se segue, apresentamos um estudo experimental sobre a sua aplicação a diferentes tipos de sinais, enquanto, no Capítulo 3, consideramos mais detidamente o seu emprego à análise de sinais eletro-encefalográficos, visando ao diagnóstico de epilepsias.

Antes de apresentar os nossos experimentos, fornecemos abaixo o pseudocódigo dos procedimentos empregados para o cálculo das duas versões da Transformada Sintonizada de Gabor.

Cálculo da Primeira Versão da Transformada Sintonizada de Gabor T:

$ \begin{array}{c} 1. \\ 2. \\ 3. \\ 4. \\ 5. \\ 6. \\ 7. \\ \end{array} $	Entre com o sinal I_1 ; Crie uma variável valor Maximo para o valor máximo absoluto do sinal; Crie uma variável ω_{max} para o número máximo de frequências; Crie um vetor modulo para o módulo da transformada de Fourier; Crie um vetor fase para a fase da transformada de Fourier; Crie um vetor resultado _{ωx} para auxiliar no cálculo da T; Crie uma matriz modulo TS para o módulo da T;
8.	Para cada pixel x do sinal de entrada Faça
9.	valorMaximo := 0.0;
10.	Se $I(x) > valorMaximo$
11.	Então $valorMaximo := I(x);$
12.	Fim_Se
13.	Fim_Para
14.	Para cada pixel x do sinal de entrada Faça
15.	$I(x) := \frac{I(x)}{valor Maximo};$
16.	Fim_Para
17.	Obtenha a transformada de Fourier do sinal I normalizado;
 18. 19. 20. 21. 22. 23. 	Para cada frequência ω de 0 a ω_{max} Faça Obtenha o módulo da transformada de Fourier de $I \text{ em } \omega - \tilde{I}(\omega) ;$ $modulo(\omega) := módulo de \tilde{I}(\omega);$ Obtenha a fase da transformada de Fourier de $I \text{ em } \omega - \varphi_{\tilde{I}}(\omega);$ $fase(\omega) := fase de \tilde{I}(\omega);$ Fim_Para;
24.	Para cada posição x de I Faça { $Centro \ da \ janela \ Gaussiana.$ }
25.	Para cada frequência ω de 0 a ($\omega_{max} - 1$) Faça
26.	Para cada posição n de I Faça
27.	$resultado_{\omega x}(n) := rac{1}{(2\pi)^{3/2}} * I(n) * e^{-i*fase(\omega)} * e^{-rac{(n-x)^2}{2*sigma^2(\omega)}};$
28.	Fim_Para
29.	Transformada de Fourier para gerar a Transformada Sintonizada T;
30.	Obtenha a transformada de Fourier de $resultado_{\omega x}$;
31.	T := Resultado da transformada de Fourier;
32.	$M \acute{o} dulo \ da \ Transformada \ Sintonizada \ T \ na \ frequência \ \omega \ e \ posição \ x;$
33.	$moduloTS(\omega, x) :=$ módulo da $T(\omega, x);$
34.	Desenha o espectrograma com as informações de $x, \ \omega \ e \ moduloTS(\omega, x):$
35.	Fim_Para $\{\omega\}$
36.	Fim_Para $\{x\}$

Cálculo da Segunda Versão da Transformada Sintonizada de Gabor $T^{(2)}$:

- Entre com o sinal I_1 ; 1.
- 2.Crie uma variável valor Maximo para o valor máximo absoluto do sinal;
- 3. Crie uma variável ω_{max} para o número máximo de frequências;
- Crie um vetor Sigma para o módulo da transformada de Fourier; 4.
- Crie um vetor fase para a fase da transformada de Fourier; 5.
- Crie um vetor $resultado_{\omega x}$ para auxiliar no cálculo da $T^{(2)}$; 6.
- Crie uma matriz modulo $T2S(\omega, x)$ para o módulo da $T^{(2)}$; 7.

Para cada pixel x do sinal de entrada Faça 8.

- 9. valorMaximo := 0.0;
- 10. Se I(x) > valor Maximo11.
 - Então valorMaximo := I(x);
- 12. Fim Se
- 13.Fim Para

Para cada pixel x do sinal de entrada **Faça** 14.

- $I(x) := \frac{I(x)}{valor Maximo};$ 15.
- 16. Fim Para
- 17.Para cada pixel x do sinal de entrada I normalizado Faça

 $Sigma(x) := (2\pi)^{3/2} *$ módulo do sinal I(x);18.

- fase(x) := fase de I(x);19.
- Fim Para; 20.

24.

```
21.
        Para cada frequência \omega de 0 a (\omega_{max} - 1) Faça {Centro da janela Gaussiana.}
22.
           Para cada posição x de I Faça
23.
              Para cada frequência m de 0 a (\omega_{max} - 1) Faça
```

$$resultado_{\omega x}[m] := rac{1}{(2\pi)^2} * \widetilde{I}[m] * e^{-i*fase[x]} * e^{-rac{(m-\omega)}{2*Sigma^2[x]}};$$

25.	Fim_Para
26.	Transformada de Fourier inversa para gerar a Transformada Sintonizada $T^{(2)}$;
27.	Obtenha a transformada de Fourier inversa de $resultado_{\omega x}$;
28.	T2 := Resultado da transformada de Fourier inversa;
29.	Módulo da Transformada Sintonizada $T^{(2)}$ na frequência ω e posição x;
30.	$moduloT2S(\omega, x) := m$ ódulo da $T^{(2)}(\omega, x);$
31.	Desenha o espectrograma com as informações de x, $\omega \in moduloT2S(\omega, x)$;
32.	Fim_Para $\{x\}$
33.	Fin $\overline{\mathbf{Para}} \{\omega\}$

2.3 Estudo Experimental da $T(\omega, x)$

A transformada $T(\omega, x)$ é calculada da seguinte forma (ver pseudocódigo acima): a fim de impor um limite superior à largura das funções analisadoras, trabalhamos com sinais amostrados uniformemente e normalizados, $I_{norm}(n) = I(n)/I_{max}$, onde I_{max} é o máximo valor absoluto de $I(n), n \in \{0, 1, ..., N-1\}$. Primeiramente, nós obtemos a transformada de Fourier discreta (DFT) do sinal $I_{norm}(n)$, o que nos fornece $\sigma(\omega_k)$ e $\varphi_{\tilde{I}}(\omega_k)$, para as frequências discretas $\omega_k = \frac{2\pi k}{N}, k \in \{0, 1, ..., N-1\}$. Para cada ω_k , nós então construímos as funções $I_{mk}(n) = e^{-i\varphi_{\tilde{I}}(\omega_k)}e^{\frac{-(n-m)^2}{2\sigma^2(\omega_k)}}I_{norm}(n)$, com $m \in \{0, 1, ..., N-1\}$. De acordo com a Equação (2.3), a DFT de $I_{mk}(n)$, na frequência ω_k , corresponde a $T(\omega_k, m)$. Para o cálculo das DFTs, usamos o algoritmo FFTW (http://www.fftw.org/). A linguagem de programação empregada nesta e nas demais implementações descritas neste documento foi o C++.

Vamos, inicialmente, analisar o sinal da Figura 2.8a, formado por quatro componentes: duas ondas senoidais de frequências idênticas, no início e no final, uma região intermediária de *chirp*, onde a frequência aumenta linearmente com a posição, e uma singularidade do tipo impulso. A sua expressão matemática é dada por

$$I_{5}(x) = 2\delta(x - 1043) + \cos\left[\frac{2\pi}{50}x\right] [u(x) - u(x - 215)] + \\ + \cos\left[\frac{2\pi x}{72000}(x - 215) + \pi\right] [u(x - 215) - u(x - 1815)] + \\ + \cos\left[\frac{2\pi}{50}(x - 1815)\right] u(x - 1815)$$
(2.56)

onde $\delta(x)$ é o impulso unitário, e u(x) é a função degrau unitário, definida como:

$$u(x) = \begin{cases} 1, \text{ para } x > 0\\ 0, \text{ para } x < 0 \end{cases}$$
(2.57)

Na Figura 2.8a, à direita, plotamos a magnitude de cada função de representação associada ao sinal (todas elas centradas em x = 0), em termos da frequência. Este gráfico indica que as frequências de maior peso na composição do sinal (que dão origem a magnitudes de Fourier maiores) serão analisadas por funções de Gabor espacialmente mais largas, e portanto mais estreitas na frequência, o que significa que as componentes dominantes do sinal serão analisadas com maior precisão. Na Figura 2.8b, à esquerda, apresentamos o espectrograma $|T(\omega, x)|^{1/4}$ (a raiz quarta foi usada para realçar os va-



Figura 2.8: Em (a), sinal $I_5(x)$ e o mapa de magnitude de suas funções de representação. Em (b), à esquerda, é mostrado o espectrograma $|T(\omega, x)|^{1/4}$ e, à direita, o espectrograma $|G(\omega, x)|^{1/4}$.

lores menores). Observa-se que as duas regiões senoidais e o *chirp* linear são facilmente identificados, bem como as descontinuidades entre essas três regiões, e aquela associada ao impulso. Na Figura 2.8b à direita, apresentamos o espectrograma $|G(\omega, x)|^{1/4}$, obtido com uma transformada de Gabor tradicional, da forma⁴

$$G(\omega, x) = \int_{-\infty}^{\infty} e^{\frac{\pi}{10^4} (\xi - x)^2} e^{-i\omega\xi} I(\xi) d\xi$$
(2.58)

Embora as componentes senoidais e o *chirp* possam ser facilmente identificados, as descontinuidades do sinal são muito mal definidas, neste caso.

Consideremos agora o sinal na Figura 2.9a, que consiste na combinação de uma senóide, um pacote de onda de Gabor, e uma descontinuidade em impulso. A sua expressão analítica é dada por

$$I_6(x) = 2\delta(x - 800) + \sin\left[\frac{2\pi}{5}x\right] + e^{-\frac{\pi(x - 300)^2}{10000}}\cos\left[\frac{\pi}{5}(x - 300)\right]$$
(2.59)

⁴Nos demais experimentos reportados aqui, esta mesma forma da transformada de Gabor tradicional foi utilizada. Deve-se ter em mente que a escolha dos parâmetros da função analisadora de Gabor é, em larga medida, arbitrária (o que não acontece com a transformada sintonizada), e que parâmetros diferentes levam a resoluções diferentes no espaço e na frequência. Como veremos, a forma da transformada de Gabor escolhida aqui mostra-se adequada para alguns experimentos, e inadequada para outros.

O espectrograma $|T(\omega, x)|^{1/4}$ para este sinal aparece na Figura 2.9b, novamente mostrando a nítida separação das suas componentes, tanto no espaço quanto na frequência. O espectrograma $|G(\omega, x)|^{1/4}$, na Figura 2.9c, identifica as componentes senoidais, mas não define corretamente a descontinuidade do sinal.



Figura 2.9: Em (a), sinal $I_6(x)$. Em (b), o seu espectrograma $|T(\omega, x)|^{1/4}$. Em (c), o seu espectrograma $|G(\omega, x)|^{1/4}$ obtido com a transformada de Gabor.

O próximo sinal a ser analisado consiste na combinação de três pacotes de onda de Gabor bem próximos, e é definido analiticamente como

$$H_7(x) = e^{-\frac{(x-420)^2}{200}} \cos\left[\frac{\pi}{5}(x-400)\right] + e^{-\frac{(x-425)^2}{600}} \cos\left[\frac{\pi}{(7,2)}(x-425)\right] + e^{-\frac{(x-440)^2}{400}} \cos\left[\frac{\pi}{5,3}(x-440)\right]$$
(2.60)

O gráfico do sinal aparece na Figura 2.10a, o seu espectrograma $|T(\omega, x)|$ na Figura 2.10b, e o seu espectrograma $|G(\omega, x)|$ na Figura 2.10c. Neste caso, a fim de evidenciar a resolução obtida para as componentes do sinal, apresentamos duas vistas do gráfico de $|T(\omega, x)|$, com os máximos da superfície plotada identificando nitidamente as componentes espaço-frequência dos três pacotes de onda de Gabor. O mesmo não é observado no espectrograma de Gabor (Figura 2.10c).

De fato, a análise do sinal $I_7(x)$ é particularmente difícil, devido à proximidade dos três pacotes de onda que o compõem. Em [19], uma wavelet de Morlet da forma

$$e^{-\frac{\beta^2 x^2}{2}}\cos(\pi x)$$
 (2.61)

foi empregada, e verificou-se que a escolha apropriada do parâmetro β é crucial, para a correta identificação das componentes. A Figura 2.11 ilustra isto, apresentando os espectros obtidos para diferentes valores de β : $\beta = 0,2$ permite distinguir dois máximos ao longo do eixo vertical (parâmetro de dilatação da wavelet), enquanto $\beta = 1,2$ distingue dois máximos ao longo do eixo horizontal (posição). Apenas $\beta = 0,6$ – valor obtido com o emprego de um procedimento de minimização da entropia dos coeficientes da wavelet [19] – permite discriminar adequadamente as três componentes do sinal.

O próximo sinal a ser analisado é o sinal $I_8(x)$ da Figura 2.12a, apresentado inicialmente em [18]. O sinal $I_8(x)$ consiste numa combinação de dois *chirps* cruzados e dois *bursts* de alta frequência, e é definido analiticamente como

$$I_8(x) = \cos[2\pi(10 + x/7)x/256] + \cos[2\pi(256/2.8 - x/6)x/256],$$

para $x \in \{0, 255\}, e$

$$I_8(x) = I_8(x) + \cos[2\pi(0, 42x)],$$

para $x \in \{114, 122\} e$ para $x \in \{134, 142\}.$ (2.62)

Conforme verificado em [18], os espectrogramas obtidos com as transformadas de Wigner e de Gabor com janela fixa não se mostram adequados para a análise desse sinal. Na Figura 2.12b, à esquerda, plotamos o espectro de magnitude de Fourier, e, à direita, o espectrograma $|T(\omega, x)|$. A detecção dos *chirps* e dos *bursts* é evidente, embora o espectrograma pareça sujo, já que a sucessão muito próxima de picos e vales na transformada de Fourier leva à análise de frequências vizinhas por funções de Gabor com larguras muito distintas. Para gerar um espectro mais limpo, nós suavizamos as magnitudes de Fourier (Figura 2.12c, à esquerda), obtendo assim uma representação



Figura 2.10: Em (a), sinal $I_7(x)$. Em (b), são apresentadas duas vistas do gráfico tridimensional do seu espectrograma $|T(\omega, x)|$. Em (c), é apresentado o espectrograma $|G(\omega, x)|$, obtido com a transformada de Gabor.

espaço-frequência nítida para o sinal original (Figura 2.12c, à direita), que preserva todas as informações relevantes. Essa suavização foi obtida iterando-se cinco vezes a operação de média $[|\tilde{I}(\omega_{k-1})|+2|\tilde{I}(\omega_k)|+|\tilde{I}(\omega_{k+1})|]/4$, para cada frequência ω_k , e depois ajustando a escala do resultado ao mesmo máximo do espectro de Fourier original. Para comparação, o espectrograma obtido com a transformada S é apresentado na Figura 2.12d, e podese verificar que ele perde resolução nas frequências mais altas, o que não ocorre com o $|T(\omega, x)|$ (observe-se como neste último distinguem-se os dois máximos locais nos extremos das frequências do espectro de Fourier). O espectrograma obtido com a transformada de Gabor, apresentado na Figura 2.12e, não representa de forma adequada as componentes do sinal.

A seguir, analisamos o sinal $I_9(x)$ da Figura 2.13a, composto pela combinação linear



Figura 2.11: Transformada wavelet (Equação (2.61)) para o sinal $I_7(x)$. Parâmetros: (a) $\beta = 0,2$. (b) $\beta = 0,6$. (c) $\beta = 1,2$.

de um trem de impulsos, um *chirp* linear, e um *chirp* quadrático inverso, cada uma destas componentes normalizada de modo a comportar energia unitária na duração do sinal. A sua expressão analítica é

$$I_{9}(x) = 25z_{1}(x) + 25z_{2}(x) + 10z_{3}(x),$$

$$\operatorname{com} z_{i}(x) = \frac{y_{i}(x)}{\sqrt{\Sigma_{x}y_{i}^{2}(x)}}, \quad i = 1 \text{ a } 3,$$

$$y_{1}(x) = \operatorname{cos} \left[\frac{\pi x^{2}}{2000}\right],$$

$$y_{2}(x) = \operatorname{cos} \left[\frac{1000\pi \left(\frac{x}{1000-1}\right)^{3}}{3}\right],$$

$$y_{3}(x) = \Sigma_{k}\delta[x - (140 + 200k)],$$

(2.63)

onde $k \in \{0,4\}$ e $x \in \{0,1000\}$. Um espectro suavizado de Fourier, obtido de forma semelhante à do exemplo anterior, foi empregado no cálculo do espectrograma $|T(\omega, x)|$ para este sinal. Conforme podemos observar na Figura 2.13b, todas as suas três componentes são devidamente detectadas. O espectrograma $|G(\omega, x)|$, na Figura 2.13c, detecta devidamente as componentes de frequência, mas os trens de impulso são muito mal definidos.

Agora consideraremos um exemplo em que se torna interessante a utilização da abordagem sintonizada baseada na derivada do sinal, isto é, a transformada $\hat{T}(\omega, x)$, da Eq. (2.32). Trata-se, neste caso, de um sinal real de voz: a emissão de um 'ah', registrada no arquivo $ah_lrr.wav$, disponível em http://cronos.rutgers.edu /~lrr/speech%20recognition% 20course.html (ver Fig. 2.14a). A magnitude de Fourier do sinal (Fig. 2.14b) exibe harmônicos de baixa energia em torno das frequências $\omega = 2,3$ e $\omega = 2,8$. O espectrograma $|\hat{T}(\omega, x)|$ (Fig. 2.14c, à esquerda) evidencia melhor essas frequências, quando comparado ao espectrograma $|T(\omega, x)|$ (Fig. 2.14c, à direita). O espectrograma $|G(\omega, x)|$, na Fig. 2.14c, representa adequadamente as componentes do sinal.

A Figura 2.15a apresenta um outro sinal de voz, representando a frase **This is a Test**, obtido da página http://cronos.rutgers.edu/~lrr/speech%20recognition%20course.html. O espectrograma $|T(\omega, x)|$ aparece na Fig. 2.15b, onde é possível distinguir as diferentes componentes do sinal, como os impulsos no início das palavras *this* e *test* (em torno de t=750 e t=12000), as assinaturas de baixa frequência associadas às vogais, e as assinaturas de média para alta frequência associadas à sibilante *s*. O mesmo se pode dizer do espectrograma $|G(\omega, x)|$, apresentado na Fig. 2.15c.

Finalmente, a Figura 2.16b apresenta o espectrograma $|T(\omega, x)|$ do sinal *Glockenspiel* (arquivo *Glock.wav* [20]), apresentado na Fig. 2.16a. Nós trabalhamos com vinte mil amostras do áudio original, com duração aproximada de 0,32s a 0,78s. O espectrograma $|T(\omega, x)|$ representa de forma satisfatória tanto os ataques bruscos do início das notas como as subsequentes ressonâncias quase-senoidais, um resultado que não pode ser obtido com transformadas de Gabor de janela fixa, conforme demonstrado em [21]. A Fig. 2.16c ilustra este fato, mostrando como o espectrograma $|G(\omega, x)|$ apresenta resolução inferior tanto no espaço como na frequência, quando comparado ao $|T(\omega, x)|$.



Figura 2.12: Em (a), sinal $I_8(x)$. Em (b), à esquerda, é mostrada a magnitude da transformada de Fourier, e, à direita, o espectrograma $|T(\omega, x)|$. Em (c), à esquerda, é mostrado o espectro de Fourier suavizado, e, à direita, o espectrograma $|T(\omega, x)|$ obtido a partir deste. Em (d), é apresentada a magnitude da Transformada S. Em (e), é apresentado o espectrograma $|G(\omega, x)|$, obtido a partir da transformada de Gabor.



Figura 2.13: Em (a), sinal $I_9(x)$. Em (b), o espectrograma $|T(\omega, x)|$ obtido a partir do espectro de Fourier suavizado. Em (c), o espectrograma $|G(\omega, x)|$, obtido a partir da transformada de Gabor.



Figura 2.14: Em (a), à esquerda, sinal $I_{10}(x)$ e, à direita, a magnitude da transformada de Fourier. Em (b), à esquerda, o espectrograma $|\hat{T}(\omega, x)|$ e, à direita, o espectrograma $|T(\omega, x)|$. Em (c), o espectrograma $|G(\omega, x)|$ obtido a partir da transformada de Gabor.



Figura 2.15: Em (a), sinal **This is a Test** e, em (b), o seu espectrograma $|T(\omega, x)|$. Em (c), o espectrograma $|G(\omega, x)|$, obtido a partir da transformada de Gabor.



Figura 2.16: Em (a), sinal *Glockenspiel* e, em (b), o seu espectrograma $|T(\omega, x)|$. Em (c), o espectrograma $|G(\omega, x)|$, obtido a partir da transformada de Gabor.

Capítulo 3

Aplicação à Análise de Sinais de EEG para a Caracterização de Epilepsias

3.1 Introdução

O eletroencefalograma (EEG) é um exame que registra a atividade elétrica cerebral por meio de eletrodos posicionados em pontos específicos da cabeça, sendo uma técnica muito utilizada para detectar uma série de distúrbios neurológicos, entre eles a epilepsia. Com o emprego do EEG, torna-se possível identificar os potenciais elétricos anormais que acompanham diversos tipos de crises epilépticas [22] [23]. Uma crise (conhecida em inglês por *seizure*) representa um episódio repentino marcado por atividade motora involuntária, alterações nas emoções, nos sentidos, comportamento, memória ou consciência, devido às descargas elétricas anormais no cérebro. As crises devem ser espontâneas e recorrentes para caracterizar uma epilepsia [24].

As ondas cerebrais dos pacientes epilépticos apresentam uma morfologia diferente daquelas dos pacientes normais, caracterizando-se pelos chamados *paroxismos epileptiformes*, geralmente de amplitude mais elevada. De acordo com [25], "Os *paroxismos epileptiformes* podem se manifestar como pontas ou espículas (do inglês *spikes*), ou como ondas agudas (do inglês *sharp waves*). Outras variedades de *paroxismos epileptiformes* são as combinações de pontas com ondas lentas, de ondas agudas com ondas lentas, as polipontas e as polipontas-ondas lentas", conforme observado na Figura 3.1.

Diversos tipos de análise podem ser efetuados sobre os sinais de EEG, visando ao diagnóstico de epilepsias, e a transformada tempo-frequência é uma das ferramentas utilizadas [26]. No presente capítulo, mostraremos como a transformada de Gabor sintonizada pode contribuir para a análise espectral desse tipo de sinais, apresentando vantagens compa-



Figura 3.1: Em (a), é mostrado um exemplo de sinal de EEG para um paciente normal. Em (b), são mostrados exemplos de *paroxismos epileptiformes interictais*, e, em (c), exemplos de *paroxismos epileptiformes ictais* (ver texto) [25].

Ondas do EEG	Frequências Associadas	Estados Mentais Associados
DL	(em Hz)	
Delta	0,5 a 3,5	As ondas delta sao ondas de baixa frequên-
		cia e de alta amplitude. Elas caracterizam
		estados de sono profundo. Além disso, as os-
		cilações delta estão relacionadas a diferentes
		patologias, dependendo de suas morfologias,
		localizações e ritmos.
Teta	3,5 a 7,5	As ondas teta estão associadas ao sono, e
		desempenham um importante papel em cri-
		anças. Nos adultos, a presença de alta ativi-
		dade nessas ondas é considerada anormal, e
		pode estar relacionada a diversos distúrbios
		cerebrais, tais como ansiedade, epilepsia, e
		lesão cerebral traumática.
Alfa	7,5 a 12,5	As ondas alfa aparecem espontaneamente,
		nos adultos normais, em condições de rela-
		xamento e inatividade mental, e são me-
		lhor visualizadas estando o paciente de olhos
		fechados. A amplitude das ondas alfa varia
		de 10 a 50 mV.
Beta	12,5 a 30	As ondas beta são ondas de alta frequência
		e de baixa amplitude (inferior à das ondas
		alfa). As oscilações beta estão associadas a
		estados de alerta, estimulação, resolução de
		problemas, e concentração. Elas são tradi-
		cionalmente divididas em oscilações beta 1 e
		beta 2 [26].

Tabela 3.1: Estados mentais associados às bandas de frequência.

rativas com relação a outras técnicas comumente empregadas [22], como a transformada de Gabor [17] e a transformada S [18]. Iniciamos com uma breve exposição, na seção seguinte, sobre as características gerais dos sinais de EEG de pacientes epilépticos.

3.2 Coleta de sinais de EEG e caracterização dos sinais de epilepsia

Os profissionais da área médica registram a atividade dos sinais de EEG num conjunto de bandas de frequência conhecidas tradicionalmente como delta, teta, alfa e beta [27]. A Tabela 3.1 descreve as particularidades de cada uma delas [27][28].

A Figura 3.2 apresenta as formas de onda associadas a essas bandas de frequência [28].



Figura 3.2: Bandas de frequência usadas na análise de sinais de EEG. Figura extraída de [28].



Figura 3.3: Esquema de localização dos eletrodos de superfície de acordo com o sistema internacional 10-20, como visto a partir da esquerda (a), e acima da cabeça (b). A nomenclatura das posições dos eletrodos deriva da sua localização anatômica, a saber: F - lobo Frontal, T - lobo Temporal, C - lobo Central, P - lobo Parietal, O - lobo Occipital. Z se refere ao eletrodo localizado na linha medial. Figura extraída de [30].

Os sinais de EEG podem ser obtidos extra-cranialmente ou intra-cranialmente. O registro extra-cranial segue um esquema padronizado de colocação de eletrodos de superfície, de acordo com o sistema internacional 10-20 [29], apresentado na Figura 3.3. A localização dos eletrodos acompanha as diferentes áreas do córtex cerebral. Os índices 10 e 20 referem-se à distância entre os eletrodos adjacentes [30].

A Figura 3.4 apresenta uma amostra dos sinais de EEG obtidos através de eletrodos de superfície [27].

As crises epilépticas podem ser classificadas como *generalizadas* ou *parciais*. As crises generalizadas envolvem os dois hemisférios do cérebro, já as parciais envolvem apenas um hemisfério cerebral e podem ser classificadas como simples (sem perda de consciên-

sense a functional and allow against the price of applications of the second state of the second state of the	1101110111111111110100
million was were and when the second of the second	man the
mathematica Manuscratting and the manuscratting	man have
12-10	
13-14 Martin Ma	
13-19 water where the man water have a second and the second seco	www.
mound mention when when when we have the second sec	man frage
and the second s	Long Long Long
N-M	
19-11	
	montant
El-manus and the providence of	and the second s
and when my property was a service of the service o	
in the second day and the second many property in the second day and the second day of the second day	
Ci-e	1.00 -
Here was a second was a second was a second was a second of the second o	
many many many many many many many many	
19-11	Tim 7
May	and the second s
mannenallyman and man provide the second of t	monter
and a second day a second second a second se	Terra Internet Internet
M-BI	1
SS-SU MARCAN AND MARCAN AND MARCAN AND AND AND AND AND AND AND AND AND A	
mapara margarette and the second and the second s	manimum
01-01	
and manufactures the second se	- manufacture

Figura 3.4: Amostra de sinais de EEG obtidos via eletrodos de superfície em várias áreas do córtex cerebral. Figura extraída de [27].

cia) ou complexas (com perda de consciência). A classificação das diversas crises pode ser encontrada em [26][31]. Uma crise epiléptica parcial pode evoluir com generalização secundária, principalmente sob a forma de crise tônico-clônica generalizada (grande-mal), que é a forma mais conhecida, caracterizada por violentas contrações musculares, com duração aproximada de 40 a 90 segundos, em que uma fase tônica inicial — envolvendo tensão muscular extrema, mas sem movimento — é seguida alguns segundos depois pela fase clônica, com contrações rítmicas de todo o corpo, perdurando até o final da crise [26][31]. As crises tônico-clônicas são dominadas inicialmente por ondas alfa e teta, características da fase tônica, que decaem ao longo do tempo para as ondas delta, de frequências menores e maior amplitude, características da fase clônica [27][32]. A referência [33], por exemplo, reporta uma fase tônica caracterizada pela frequência de 10Hz (banda alfa), com duração aproximada de 10 segundos, e uma fase clônica caracterizada pelo aumento progressivo das baixas frequências (5-6Hz). As faixas de frequências registradas podem sofrer influência, por exemplo, do fato de os registros serem intra-craniais ou extra-craniais, do tipo da medicação administrada ou da utilização de métodos de filtragem [26].

Os paroxismos epileptiformes podem ser classificados como interictais (Fig. 3.1b) ou ictais (Fig. 3.1c). Os paroxismos interictais se apresentam no período entre as crises clínicas, são assintomáticos e de duração limitada (em torno de 70 a 80 ms), e não estão associados a sinais de crises epilépticas. Os paroxismos ictais estão relacionados a manifestações clínicas das crises e têm, como algumas de suas características, uma du-



Figura 3.5: Esquema de implantação de eletrodos intra-craniais, para avaliação précirúrgica de pacientes epilépticos. Em (a), eletrodos de profundidade implantados simetricamente nas formações do hipocampo. Em (b) e (c), faixas de eletrodos implantados nas regiões do neocórtex (áreas mais evoluídas do córtex) [13].

ração superior a 10 segundos, atividades rítmicas em diferentes bandas de frequência, e mudanças de amplitude e morfologia durante o fenômeno crítico [30].

O registro intra-cranial é utilizado nas epilepsias focais (parciais), quando há a necessidade de uma avaliação pré-cirúrgica, com o intuito de se localizar com precisão a área do córtex cerebral causadora da crise, chamada de *zona epileptogênica* (foco epiléptico). Nessa região, em ausência de crise, registram-se ocorrências intermitentes de atividades epileptiformes *interictais* e, durante as crises epilépticas (atividade *ictal*), o sinal é quase periódico e de amplitude elevada [13]. É importante ressaltar que apenas as crises tônicoclônicas com generalização secundária apresentam foco epiléptico, por serem originadas em uma área específica do cérebro, sendo, portanto, passíveis de uma intervenção cirúrgica. Em contraste, as crises tônico-clônicas com generalização primária, que aparentemente têm início em toda a extensão cerebral, não são candidatas a cirurgia, devido à impossibilidade de se detectar o local de origem da crise (foco) [24]. A Figura 3.5 apresenta um esquema de localização de eletrodos intra-craniais.

3.3 Análise Tempo-Frequência de sinais de EEG

A literatura descreve várias técnicas para a análise quantitativa de sinais de EEG, tais como a análise em frequência (análise espectral), o mapeamento topográfico (mapeamento cerebral), e outras técnicas analíticas complexas [26]. Clinicamente, o método mais difundido é o da análise espectral, juntamente com a inspeção visual. Embora esta última ainda represente a principal ferramenta clínica, e constitua o ponto de referência para as demais abordagens, provou-se interessante relacioná-la à análise matemática baseada numa descrição tempo-frequência [22][26][34][35]. A inspeção visual é muito subjetiva e dificilmente permite uma análise estatística, logo, várias técnicas foram criadas a fim de quantificar a informação do EEG. Dentre elas, a transformada de Fourier vem sendo empregada para a inspeção das componentes em frequência dos sinais eletroencefalográficos, uma atividade que é difícil de executar visualmente, pela presença de várias componentes rítmicas simultâneas. Como se sabe, no entanto, a transformada de Fourier clássica não provê resolução temporal, ao passo que a descrição adequada dos sinais de EEG requer localização conjunta tanto no tempo quanto na frequência. As transformadas tempofrequência tornam-se, assim, ferramentas adequadas para detecção e análise desse tipo de sinais [26].

A análise tempo-frequência é importante sobretudo para o tratamento de sinais de EEG de pacientes que apresentam crises epilépticas do tipo tônico-clônicas, em atividade ictal, principalmente quando a coleta dos sinais é feita por meio de eletrodos de superfície (como os mostrados na Fig. 3.3). Estes sinais estão sujeitos a ruídos de alta frequência (na banda beta) introduzidos por artefatos musculares. Entende-se por artefato qualquer alteração do EEG proveniente de eventos involuntários, como, por exemplo, movimento da cabeça, piscar dos olhos ou atividade muscular. Devido às baixas amplitudes dos sinais de EEG, os artefatos geralmente contaminam os registros, comprometendo a sua análise. Eles causam variações anormais nos potenciais elétricos, o que torna insuficiente uma inspeção meramente visual (um exemplo deste efeito pode ser visto em [27]), e, como consequência, em muitos casos a análise do EEG fica restrita à interpretação dos sinais elétricos que precedem ou sucedem a atividade tônico-clônica, desprezando-se a fase *ictal* [36]. Uma transformada tempo-frequência torna-se aqui vantajosa, por permitir observar os padrões escondidos sob os artefatos musculares. Isto possibilita a análise apenas das frequências relevantes, características da crise, desconsiderando-se os artefatos introduzidos durante a fase *ictal*. Consegue-se assim obter mais informação por um meio não-invasivo. Já na análise intra-cranial, a influência dos artefatos é menor, mas, como vimos, esta técnica costuma ser indicada apenas nos casos mais graves, quando uma intervenção cirúrgica está sendo cogitada [26].

A referência [22] analisa diversas abordagens tempo-frequência, e conclui que elas podem ser úteis tanto no estudo da dinâmica dos eventos de epilepsia, quanto no monitoramento para avaliação pré-cirúrgica ou na detecção de ataques epilépticos, a escolha da transformada dependendo do tipo de informação que se deseja obter. No entanto, algumas técnicas apresentam características desfavoráveis, como, por exemplo, uma baixa resolução temporal ou em frequência, ou a introdução de termos espúrios. A transformada de Gabor sintonizada, conforme observado anteriormente, possui algumas vantagens para

Classes	Descrição
A e B	Segmentos obtidos a partir de registros de EEG de super-
	fície em cinco voluntários saudáveis, utilizando o esquema
	padronizado de colocação de eletrodos da Figura 3.3. Os vo-
	luntários se encontravam relaxados, em estado desperto, de
	olhos abertos (A) ou fechados (B) .
$C \in D$	Segmentos obtidos de cinco indivíduos epilépticos, registrados
	na ausência da crise, em diferentes regiões do cérebro. Esses
	segmentos foram coletados por eletrodos de profundidade,
	conforme o esquema de colocação intra-cranial da Figura 3.5a.
E	Segmentos obtidos de indivíduos em estado epiléptico, de na-
	tureza ictal. São pacientes com crises parciais, que podem ter
	evoluído em crises secundárias generalizadas tônico-clônicas.
	Esse tipo de sinal registra contrações musculares nas ban-
	das de frequência beta, que interferem em sua análise. Os
	segmentos desta classe foram obtidos por meio de eletrodos
	posicionados conforme as Figuras 3.5b e 3.5c.

Tabela 3.2: Descrição das cinco classes de sinais de EEG consideradas [13].

a análise de sinais não-estacionários. Diferentemente da transformada de Wigner, por exemplo, ela não apresenta o problema dos termos cruzados, e graças à sua boa resolução tanto temporal quanto espectral, ela permite evidenciar as componentes de frequência relevantes nos sinais de EEG, ao mesmo tempo em que acompanha as suas componentes temporais transientes, como as espículas e as ondulações características das atividades *interictais* e *ictais*. Como também já constatado, a nossa abordagem reduz o problema de *self-aliasing*, relevante no caso da transformada S, que pode interferir na análise das componentes de frequência mais elevada. No que se segue, nós apresentamos um estudo sobre a aplicação da transformada de Gabor sintonizada à análise de EEGs, visando à caracterização da epilepsia.

3.4 Experimentos

Em nossa pesquisa, foram utilizadas cinco classes de dados, disponíveis online [13], apresentando registros de EEG tanto de pacientes saudáveis como de epilépticos. Essas classes, descritas por A, B, C, $D \in E$, contêm, cada uma, 100 segmentos de EEG de um único canal (diferença de potencial entre eletrodos adjacentes [37]), com 23,6 segundos de duração. A Tabela 3.2 apresenta uma breve descrição das cinco classes. A classe Erepresenta o interesse primordial do nosso estudo, mas as demais classes também serão analisadas. As classes $C, D \in E$ são originárias de arquivos de EEG referentes a diagnóstico précirúrgico, constando que todos os pacientes envolvidos conseguiram controle completo das crises epilépticas, após ressecção de uma das formações do hipocampo — logo, essa região foi diagnosticada corretamente como sendo a zona epileptogênica (Figura 3.5). Os segmentos da classe D foram obtidos internamente na zona epileptogênica, e os segmentos da classe C foram registrados a partir da formação do hipocampo do hemisfério oposto a esta. Enquanto os segmentos das classes $C \in D$ registram apenas atividades medidas durante um intervalo de ausência de crise (atividade *interictal*), os segmentos da classe Eforam selecionados de todos os locais em que se manifestou atividade *ictal* [13].

Cada uma das classes apresenta diferentes características morfológicas. A Figura 3.6 mostra exemplos de seus segmentos. O sinal de EEG registrado extra-cranialmente durante o estado relaxado de pacientes saudáveis, com olhos fechados, apresenta um ritmo predominante na banda alfa [23], atividade que é mais evidente na parte posterior da cabeça (classe B). Em contrapartida, uma gama mais ampla de frequências é observada para pacientes saudáveis e de olhos abertos (classe A). Os sinais das classes $A \in B$ têm aparência aleatória, sem demonstrar um padrão visível ao longo do seu desenvolvimento [34]. As amplitudes dos registros de EEG de superfície costumam ser da ordem de alguns microvolts. Já os sinais das classes $C \in D$, obtidos a partir de registros intra-craniais, apresentam padrões de natureza interictal, com paroxismos isolados em sua composição. As amplitudes desses registros variam em torno de $100\mu V$. Os sinais da classe E exibem uma natureza periódica característica, bem como a morfologia dos paroxismos epileptiformes, com a presença de espículas e ondulações. As suas amplitudes podem ultrapassar os $1000\mu V$. Os padrões determinados pelas diversas formas de paroxismo ajudam a identificar, classificar e localizar as crises epilépticas [13]. Aqui nós apresentamos trechos correspondentes aos três primeiros segundos do sinal, com o intuito de obter uma melhor visualização do aspecto das suas ondas. Na Figura 3.6e, por exemplo, é possível identificar os paroxismos epileptiformes presentes no sinal. Se nós utilizamos toda a duração do sinal (23,6 segundos), este tipo de verificação torna-se difícil, como ilustrado na Figura 3.7. Conforme constatamos, o padrão de cada sinal não sofre alteração visível ao longo de toda a sua duração.

Em [26], foi introduzida uma ferramenta de análise, baseada na transformada de Gabor, para verificação da relação entre a intensidade espectral de cada banda de frequência e a intensidade total do sinal em cada instante, a partir de registros de EEG de pacientes em crise tônico-clônica, durante as fases *pré-ictal*, *ictal* e *pós-ictal*. A ideia é a de verificar



Figura 3.6: Exemplos de segmentos de EEG de cada uma das cinco classes. De cima para baixo, classes A a E. Nos eixos verticais, as amplitudes dos sinais estão em μV [13].



Figura 3.7: O mesmo sinal da Fig. 3.6e, apresentado em sua duração total (23,6 segundos).

o peso de cada banda de frequência em cada instante, associando-o aos eventos de crise. A seguir nós mostramos como se obtém essa medida da *intensidade relativa de banda* (*RIR*).

A intensidade espectral da banda de frequência b (b = delta, teta, alfa, beta), definida no intervalo $\omega_{min}^{(b)} \leq \omega \leq \omega_{max}^{(b)}$, com $\omega_{min}^{(b)}$ e $\omega_{max}^{(b)}$ conforme as faixas especificadas na Tabela 3.1, é calculada da seguinte forma:

$$I^{(b)}(t_1) = \int_{\omega_{\min}^{(b)}}^{\omega_{\max}^{(b)}} |G_{\sigma}(\omega, t_1)|^2 d\omega$$
(3.1)

onde $G_{\sigma}(\omega, t_1)$ é a transformada de Gabor definida por

$$G_{\sigma}(\omega, t_1) = \int_{-\infty}^{\infty} e^{-i\omega t'} g_{\sigma}(t' - t_1) I(t') dt'$$
(3.2)

com $g_{\sigma}(t'-t_1)$ sendo uma janela gaussiana deslizante de largura σ e centro em t_1 . A intensidade espectral total é obtida como

$$I_T(t_1) = \Sigma_b I^{(b)}(t_1)$$
(3.3)

e a intensidade relativa de banda (RIR) fica definida, para a banda de frequência b, como a razão

$$RIR^{(b)}(t_1) = \frac{I^{(b)}(t_1)}{I_T(t_1)} \times 100\%$$
(3.4)

Intensidades relativas de banda podem ser obtidas também a partir da transformada sintonizada de Gabor, e os resultados são comparáveis aos que resultam da abordagem acima, conforme a Figura 3.8. A Figura 3.8a apresenta a *RIR* fornecida pela transformada sintonizada, para o sinal de EEG da Figura 3.6e. As três ondas (delta, teta, alfa) coexistem ao longo de toda a duração do sinal, observando-se que, quanto maior a influência das ondas teta, menor a influência das ondas delta, e vice-versa, o que concorda com [36]



Figura 3.8: Intensidade relativa de banda (RIR), para cada banda de frequência, obtida a partir da transformada de Gabor sintonizada, (a) para o segmento de sinal apresentado na Fig. 3.6e, e (b), para o mesmo sinal, mas na duração total de 23,6 segundos (Fig. 3.7).

(nessa referência, assim como aqui, as ondas beta foram excluídas da análise, visto estarem associadas a artefatos musculares). É importante destacar que, como o segmento analisado é de apenas 3 segundos, não é possível a caracterização completa de uma crise clônicotônica, já que a duração desse tipo de evento é de 40 a 90 segundos. De qualquer modo, este exemplo serve para mostrar a importância variável que cada banda de frequências assume em cada instante, no desenrolar do sinal. A Figura 3.8b apresenta a curva correspondente à duração total (23,6 segundos) do mesmo EEG (Figura 3.7), e, assim com a Figura 3.8a, ilustra o comportamento típico de uma crise clônica, em que há o predomínio das ondas delta e, às vezes, também, das ondas teta [33]. É importante salientar que cada paciente apresenta especificidades, dentro do que é esperado para cada tipo de crise, e torna-se interessante a análise conjunta da morfologia do sinal e do gráfico da RIR, para evidenciar possíveis discrepâncias no padrão observado. Tanto a transformada de Gabor tradicional quanto a sua versão sintonizada mostram-se igualmente adequadas para este tipo de análise.

A Figura 3.9 ilustra o comportamento geral da transformada de Gabor sintonizada quando aplicada a sinais de EEG. A figura exibe os espectrogramas $|T(\omega, x)|$ obtidos para os sinais da Figura 3.6 (a)-(d) (sinais das classes de A a D). Conforme anteriormente

mencionado, os sinais da classe A apresentam um espectro distribuído numa ampla gama de frequências (Figura 3.9a). Os espectros da classe B, característicos de indivíduos saudáveis quando monitorados com os olhos fechados, apresentam destaque na faixa alfa de frequências, conforme a Figura 3.9b. Já os espectros dos sinais das classes $C \in D$ (Figuras 3.9c e 3.9d) apresentam baixas frequências concentradas nas bandas delta e teta. Os espectrogramas permitem identificar os instantes associados aos eventos transientes mais importantes desses sinais de natureza *interictal*.

A seguir, apresentamos um estudo comparativo da transformada de Gabor sintonizada e de outras transformadas utilizadas na análise de sinais de EEG associados a epilepsia. As Figuras 3.10 e 3.11 mostram segmentos de EEG da classe E, com duração de 10 segundos, escolhidos nos instantes de 5 a 15 segundos (Figura 3.10), e nos instantes de 0 a 10 segundos (Figura 3.11), juntamente com os respectivos espectrogramas obtidos com a transformada sintonizada, a transformada de Gabor tradicional, e a transformada S.

Na Figura 3.10b, pode-se observar a ótima resolução em frequência e a ótima resolução temporal proporcionadas pela transformada sintonizada, que permite discriminar os paroxismos característicos do sinal (indicados pelo conjunto de linhas verticais paralelas que surgem nos instantes associados à crise), e as suas principais componentes espectrais. As demais transformadas não demonstram igual capacidade. Por exemplo, na Figura 3.10c, o espectrograma obtido com a transformada de Gabor de janela $0,25s^1$ reflete bem a dinâmica do EEG, ao mostrar as suas variações em frequência, mas apresenta resolução tempo-frequência inferior à da transformada sintonizada, dificultando a detecção da ocorrência de crises. Por outro lado, a transformada S (Figura 3.10d), embora tenha proporcionado resolução temporal suficiente para a detecção dos paroxismos, mostrou-se inconsistente em sua resolução tempo-frequência, não permitindo discriminar com segurança as faixas de frequência mais importantes do sinal em cada instante. Na região de baixas frequências, associadas às bandas delta e teta, obteve-se uma resolução espectral razoável (comportamento verificado também em [22]), mas, nas frequências superiores, como as das ondas alfa, a resolução mostrou-se crítica.

Conclusões semelhantes advêm da análise do experimento ilustrado na Figura 3.11. Enquanto no sinal da Figura 3.10a se verifica uma predominância das frequências na banda alfa (de 7,5 a 12,5 Hz) — característica que se reflete bem na transformada sintonizada —, no EEG da Figura 3.11a, a faixa de frequências predominante é a da banda teta — ou seja, de 3,5 a 7,5 Hz —, o que se pode observar tanto na transformada sintonizada quanto

 $^{^1 \}rm N$ ós verificamos experimentalmente que este é o tamanho de janela que proporciona a melhor resolução conjunta tempo-frequência.



Figura 3.9: Espectrograma $|T(\omega, x)|$ dos sinais da Figura 3.6, referentes às classes de A a D (de cima para baixo).



Figura 3.10: (a) Segmento de um sinal de EEG de um paciente com diagnóstico de epilepsia. Amplitude em μV no eixo vertical. (b) Espectrograma da transformada sintonizada $|T(\omega, x)|$ (c) Espectrograma da transformada de Gabor $|G(\omega, x)|$, utilizando janela de 0,25 segundo. (d) Espectrograma da transformada S.

na transformada de Gabor. Em ambos os casos, devido à baixa resolução espectral (que acarreta a superposição das diversas faixas de frequência), a transformada S se mostrou incapaz de proporcionar precisão comparável.

Em experimentos com 30 sinais da base de dados disponível em [13], o tipo de comportamento reportado acima se repetiu. Na comparação entre a transformada sintonizada, a transformada de Gabor tradicional e a transformada *S*, a primeira se mostrou capaz de proporcionar a melhor resolução simultânea no tempo e na frequência, permitindo tanto a discriminação temporal dos paroxismos epilépticos quanto a identificação das suas componentes espectrais. A visualização adequada dos paroxismos nestes dois domínios constitui uma vantagem importante, pois ajuda a identificar as frequências associadas a cada crise epiléptica, bem como a intensidade da crise em um dado instante, tudo isto com o emprego de uma mesma representação gráfica, sem que seja preciso recorrer paralelamente à própria curva do EEG. Infelizmente, só nos foi possível trabalhar com amostras de registros de EEG de curta duração (23,6 segundos), disponibilizadas publicamente (os outros dados a que nós tivemos acesso estavam armazenados em formatos proprietários, o que impediu o seu uso). O ideal seria a utilização de segmentos mais longos, que permitiriam uma caracterização mais precisa das crises epilépticas (uma crise tônico-clônica, por exemplo, somente fica bem caracterizada em sinais de duração superior a 90 segundos).

O trabalho aqui apresentado foi desenvolvido em parceria com a Dra. Glenda Lacerda, do setor de epilepsia do Hospital Universitário Antônio Pedro, que nos forneceu suporte técnico e teórico. Como desdobramento futuro dessa colaboração, nós pretendemos estender a nossa análise para sinais obtidos, de um mesmo paciente, em todos os canais de EEG, o que permitiria determinar, em instantes relevantes (por exemplo, na instalação de uma crise), quais são as frequências preponderantes nos diversos canais, e a que regiões do córtex cerebral elas se associam. Este tipo de informação pode constituir um auxílio importante na detecção da zona epileptogênica.



Figura 3.11: (a) Segmento de um sinal de EEG de um paciente com diagnóstico de epilepsia. Amplitude em μV no eixo vertical. (b) Espectrograma da transformada sintonizada $|T(\omega, x)|$ (c) Espectrograma da transformada de Gabor $|G(\omega, x)|$, utilizando janela de 0,25 segundo. (d) Espectrograma da transformada S.

Capítulo 4

Aplicação à Modelagem Neuronal

Neste capítulo, nós aplicaremos os modelos matemáticos de representação de sinais, apresentados anteriormente, à modelagem de alguns tipos de neurônios visuais [8]. Conforme já mencionado no Capítulo 1, a representação espaço-frequência baseada nas funções sintonizadas de Gabor possui características que parecem torná-la adequada à modelagem de propriedades das células simples do córtex visual dos mamíferos [38]. Além disso, nós verificamos que uma representação sintonizada alternativa — não mais baseada em funções de Gabor, mas agora em funções codificadoras circularmente simétricas — permite modelar células da retina e do *núcleo geniculado lateral* (NGL) que apresentam tal simetria.

Antes de iniciarmos a apresentação destas aplicações da abordagem sintonizada à modelagem neuronal, nós faremos uma breve introdução aos conceitos básicos da neurofisiologia da visão.

4.1 Neurofisiologia da Visão: O Caminho Visual

O cérebro é uma estrutura altamente complexa, constituída por um grande número de células nervosas (os neurônios) densamente conectadas entre si (da ordem de 10^{12} células e 10^5 conexões – ou *sinapses* –, por célula) [8]. Em um dado neurônio, tipicamente identificam-se três estruturas distintas: o corpo celular, os dendritos e o axônio (Figura 4.1). Os dendritos atuam como os terminais receptores do neurônio, e o axônio, como o seu terminal transmissor. O funcionamento do cérebro consiste em um processo coletivo de recepção e emissão de impulsos elétricos (os *potenciais de ação*) que se propagam rapidamente entre conjuntos específicos de neurônios [8, 12]. Aqui nós estaremos interessados nos neurônios do chamado *caminho visual* (Figura 4.2), que segue desde os olhos até
o córtex visual, situado na parte posterior do cérebro, onde a maior parte da informação visual relevante é extraída e codificada.



Figura 4.1: Estrutura de um neurônio [8].



Figura 4.2: Caminho Visual [8].

O estímulo luminoso incide inicialmente na retina, parte neural dos olhos, onde as células fotorreceptoras fazem a transdução da luz em sinais nervosos (potenciais de ação), que são transmitidos para dois outros tipos principais de neurônios da própria retina, as células *bipolares* e as células *ganglionares*. As células ganglionares são dotadas de axônios muito longos, que se unem numa cavidade próxima ao centro da retina, formando o chamado *eixo óptico*, que se prolonga para o interior do cérebro e faz conexão com o núcleo geniculado lateral. O NGL é geralmente comparado a um centro de relé no caminho

visual, localizado a meio caminho do córtex, no interior de cada hemisfério cerebral¹. A partir dele, a informação visual ascende ao córtex visual, que corresponde a cerca de um quinto da área total do córtex cerebral, a camada mais externa do nosso cérebro, com cerca de 1 a 4 mm de espessura, onde tem lugar a atividade neuronal mais sofisticada. Diversos tipos de neurônios já foram identificados nas diferentes sub-áreas em que se divide o córtex visual. Aqui nós estaremos principalmente interessados nos chamados *neurônios simples* e *neurônios complexos* da região conhecida como V1, ou córtex visual primário.

4.1.1 Campos Receptivos

Cada uma das células no caminho visual responde a estímulos luminosos apenas em áreas restritas do campo visual do observador. Este fato leva ao conceito de *campo receptivo*, que é primordial para a descrição da atividade neuronal. O campo receptivo de um neurônio visual é tradicionalmente definido pela região do espaço onde um estímulo luminoso evoca a resposta neuronal, e também pela natureza dessa resposta [40]. Uma resposta é dita excitatória se o estímulo luminoso faz com que a célula emita potenciais de ação a uma taxa superior à sua taxa de repouso; ela é dita inibitória, se a taxa de disparo for inferior à de repouso.

Os campos receptivos das células ganglionares da retina, das células do NGL e das células do primeiro estágio do córtex visual têm *simetria circular* – ou seja, um dado estímulo luminoso produz a mesma resposta, independentemente de sua orientação. Estes campos apresentam uma organização *centro-periferia*, podendo ser do tipo centro-*on* ou centro-*off*. As células centro-*on*, produzem resposta excitatória (*on*) quando o estímulo luminoso incide na parte central do campo receptivo, e uma resposta inibitória (*off*), quando o estímulo é apresentado em sua periferia. As células centro-*off* apresentam o comportamento inverso. A Figura 4.3 apresenta esses dois tipos de campos receptivos.

Nos estágios superiores do córtex visual, as células vão se tornando cada vez mais seletivas, e já não aparecem campos receptivos com simetria circular. No córtex visual primário (V1), distinguem-se dois tipos principais de células, ambas sensíveis à orientação do estímulo: as *células simples* e as *células complexas* [8, 41]. Nas células simples, os domínios excitatórios ou inibitórios são bem definidos, apresentando-se como faixas alongadas adjacentes, cuja direção vai definir a orientação preferencial da célula: estímulos com esta mesma orientação tenderão a evocar maiores respostas [42]. As células com-

¹O NGL é, obviamente, bem mais do que isso. Ele recebe a maior parte de suas entradas, não dos olhos, mas do córtex cerebral, e a sua resposta é afetada por nossa atividade de alto nível, como os estados de alerta ou atenção [39].



Figura 4.3: Tipos de campos receptivos *centro-periferia*. À esquerda, um campo centro-*on* e à direita, um campo centro-*off*. O símbolo "+"denota regiões excitatórias, enquanto o "-"denota regiões inibitórias [8].

plexas são as células mais comuns no córtex cerebral, correspondendo a cerca de 70% do total. Ao contrário das células simples, elas não apresentam uma clara subdivisão do seu campo receptivo em regiões excitatórias e inibitórias. Por isso, as células complexas respondem independentemente da posição do estímulo no campo receptivo, desde que a sua orientação seja apropriada.

4.2 Modelagem das Células Corticais Simples por Funções Sintonizadas

A descrição clássica do campo receptivo assume uma organização espacial fixa, com a resposta neuronal a um estímulo invariante no tempo sendo modelada pela filtragem deste por um campo receptivo pré-determinado. No caso das células corticais simples, o modelo tradicional para o campo receptivo são as funções de Gabor [43]. Essa visão clássica de campos receptivos estáticos tem, contudo, sido contestada por experimentos neurofisiológicos recentes, que mostram que a estrutura do campo receptivo se altera de acordo com a entrada neural [5, 6, 7]. Motivados por este resultado, nós propomos utilizar as funções de representação 2-D introduzidas no Capítulo 1, como modelo para os campos receptivos das células corticais simples.

Esses campos ficariam então definidos por funções de Gabor da forma (Equação (1.14)):

$$\psi_{\omega_x,\omega_y}(x,y) = e^{i[\omega_x x + \omega_y y + \varphi_{\tilde{I}}(\omega_x,\omega_y)]} e^{-\frac{x^2 + y^2}{2\sigma^2(\omega_x,\omega_y)}}$$
(4.1)

onde $\varphi_{\tilde{I}}(\omega_x, \omega_y)$ é a fase da transformada de Fourier do sinal, $\tilde{I}(\omega_x, \omega_y)$, e onde $\sigma(\omega_x, \omega_y)$

é obtido a partir do módulo desta transformada, como

$$\sigma(\omega_x, \omega_y) = \frac{1}{(2\pi)^{3/2}} \sqrt{|\tilde{I}(\omega_x, \omega_y)|}$$
(4.2)

Na Figura 4.4, são mostrados exemplos das funções de representação acima, obtidas de um fragmento 16×16 de uma imagem natural.

O espectro de Fourier das imagens típicas é fortemente concentrado nas baixas frequências, de modo que as funções $\psi_{\omega_x,\omega_y}(x,y)$ tendem a se tornar mais estreitas – e, consequentemente, as suas larguras de banda tendem a aumentar – com o aumento da frequência. Esse tipo de comportamento tem sido observado para a largura de banda de células corticais simples do macaco [44], e, em um trabalho mais recente [45], campos receptivos com perfil passa-baixa e pouco sintonizados em orientação, semelhantes à função de codificação DC da Figura 4.4, também foram encontrados. Esse tipo de campo receptivo não é previsto pelos modelos usuais de codificação cortical [43, 45].



Figura 4.4: Exemplos de funções de representação da Equação (4.1). As frequências consideradas, (ω_x, ω_y) , de cima para baixo e da esquerda para a direita, são: (0,0), (0,1), (1,0), (0,2), (1,1) e (1,2).

Com os campos receptivos definidos como acima, a representação de imagens pelos neurônios corticais simples seria modelada como (ver Equação (1.12)):

$$I(x,y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{i[\omega_x x + \omega_y y]} * \psi_{\omega_x,\omega_y}(x,y) d\omega_x d\omega_y$$
(4.3)

Esta expressão é um resultado matemático exato, válido para qualquer sinal de quadrado integrável definido sobre um domínio infinito, mas ele permanece aproximadamente válido sobre janelas finitas, com diferentes valores para os parâmetros das funções $\psi_{\omega_x,\omega_y}(x,y)$ computados em cada janela. Isto permite a interpretação destas funções de representação locais como campos receptivos.

Na Figura 4.5, nós apresentamos as representações obtidas com a Equação (4.3) para um conjunto de imagens naturais (adquiridas da base de van Hateren [46]), considerando apenas as seis componentes de mais baixa frequência, computadas sobre janelas 3×3 . O erro médio da representação, sobre todo o conjunto, é de apenas 3,9%.



Figura 4.5: Em cima: imagens originais, de tamanho 128×192 . Embaixo: representação obtida a partir do nosso modelo, computada usando janelas 3×3 , para as frequências $(0,0), (0,1), (1,0), (0,2), (1,1) \in (1,2)$.

4.3 Modelagem das Células Centro-Periferia por Funções Sintonizadas

A abordagem apresentada acima, para modelagem de campos receptivos por funções sintonizadas, pode ser estendida para células com organização centro-periferia, como encontradas na retina e no núcleo geniculado lateral. Este tipo de organização é explicado, com base em princípios da teoria da informação, assumindo-se que o objetivo dos estágios iniciais do caminho visual é o de produzir uma versão descorrelacionada do sinal de entrada, que é então propagada ao córtex cerebral para posterior processamento [10, 47]. Os neurônios da retina e do núcleo geniculado lateral teriam, assim, campos receptivos otimizados para *branquear* imagens naturais, cujo espectro de Fourier decai, aproximadamente, com o inverso da magnitude da frequência – ou seja, $\approx (\omega_x^2 + \omega_y^2)^{-1/2}$ [48].

Coerentes com esta interpretação, nós introduzimos funções de codificação circularmente simétricas, e dependentes do estímulo, em termos das quais uma representação semelhante à da Equação (4.3) pode ser definida para o sinal branqueado:

$$I_{branq}(x,y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{i(\omega_x x + \omega_y y)} * \psi(r;\omega_x,\omega_y) d\omega_x d\omega_y$$
(4.4)

onde $I_{branq}(x, y)$ é uma imagem branqueada, obtida pela convolução da imagem de entrada com um filtro branqueador de fase nula, B(x, y),

$$I_{branq}(x,y) = B(x,y) * I(x,y)$$

$$(4.5)$$

e onde $\psi(r; \omega_x, \omega_y)$ é uma função de representação circularmente simétrica – com $r = \sqrt{x^2 + y^2}$ –, definida como

$$\psi(r;\omega_x,\omega_y) = -\frac{e^{i\varphi_{\bar{I}}(\omega_x,\omega_y)}}{\pi r} \{1 - \cos[\sigma(\omega_x,\omega_y)\pi r] - \sin[\sigma(\omega_x,\omega_y)\pi r]\}$$
(4.6)

Na equação acima, $\varphi_{\tilde{I}}$ é a fase da transformada de Fourier do sinal de entrada, e $\sigma(\omega_x, \omega_y)$ se relaciona à magnitude desta transformada, conforme demonstrado no Apêndice D, como

$$\sigma(\omega_x, \omega_y) = \frac{\rho}{\pi} \sqrt{1 - \left[1 + \frac{\rho \tilde{B}(\omega_x, \omega_y) |\tilde{I}(\omega_x, \omega_y)|}{4\pi}\right]^{-2}}$$
(4.7)

onde $\tilde{B}(\omega_x, \omega_y)$ é a transformada de Fourier do filtro de branqueamento, e onde ρ é definido como a magnitude da frequência: $\rho = \sqrt{\omega_x^2 + \omega_y^2}$.

O módulo da função $\psi(r; \omega_x, \omega_y)$ atinge o valor máximo de σ , para r = 0, qualquer que seja a frequência. O seu primeiro zero ocorre para $r = r_0 = 1/2\sigma$, e o segundo, para $r = r_1 = 2/\sigma$, o intervalo entre r_0 e r_1 definindo um lobo lateral de tamanho $3/2\sigma$. Em $r > r_1$, aparecem outros zeros, e portanto outros lobos laterais, mas agora com profundidades bem menores – da ordem de um terço, ou menos, da do primeiro. A função $\psi(r; \omega_x, \omega_y)$ apresenta, portanto, um lobo lateral dominante, que definirá a periferia do campo receptivo. O seu tamanho depende do conteúdo espectral da imagem de entrada, e da escolha do filtro de branqueamento (ver Equação (4.7)).

Seguindo a interpretação proposta em [9, 10], nós assumimos que a estrutura dos campos receptivos centro-periferia é tal que eles tendem a equalizar o espectro das imagens naturais — que decai com $1/\rho$, como já mencionado —, ao mesmo tempo suprimindo ruído de alta frequência. Nós modelamos este comportamento por meio de um filtro B(x, y)com espectro

$$\tilde{B}(\omega_x, \omega_y) = \frac{\rho}{1 + \kappa \rho^2} \tag{4.8}$$

onde κ é um valor arbitrário. Outros filtros com as mesmas propriedades espectrais

poderiam ter sido escolhidos, sem alterar significativamente o nosso modelo.

Na Figura 4.6 são apresentados exemplos das funções de representação obtidas com este modelo, a partir de um fragmento da imagem natural apresentada na Figura 4.7a. A figura apresenta o módulo de $\psi(r; \omega_x, \omega_y)$ dividido por σ , de modo que todas as funções atingem o mesmo valor máximo de 1 (quando o fator de fase $e^{i\varphi_{\tilde{I}}(\omega_x, \omega_y)}$ é considerado, obtemos tanto estruturas centro-*on* quanto centro-*off*).

O nosso modelo não responde a uma entrada uniforme, já que, quando $\rho = 0$, σ também se anula, e a função de representação desaparece. Por outro lado, para baixas frequências, como ilustrado na Figura 4.6a, os campos receptivos apresentam uma organização centro-periferia bem marcada. À medida em que ρ cresce, a periferia tende a se tornar menos saliente (Figura 4.6b), quase desaparecendo nas frequências mais altas (Figura 4.6c). Tal comportamento se manifesta independentemente do valor escolhido para κ (em nossos experimentos, usamos $0.05 < \kappa \leq 1$) — o que está de acordo com o nosso modelo de branqueamento: o peso relativo das baixas frequências médias deve ser aumentado, e o ruído da alta frequência deve ser eliminado. Portanto, os campos receptivos de baixa frequência tenderiam a ter um perfil passa-faixa, enquanto os de alta frequência tenderiam a ter um perfil passa-faixa, enquanto os de alta Figura 4.6.

Apresentamos, a seguir, o pseudocódigo para o branqueamento de imagens utilizando funções de codificação circularmente simétricas.

Processo de branqueamento de imagens pelas funções de codificação circularmente simétricas:

1. Entre com a imagem original I_1 ;

2.Defina o tamanho jan das janelas deslizantes;

3. Defina o número máximo de frequências a serem utilizadas nas direções x e y - $\omega_x max$ e $\omega_y max;$

4. Defina o valor do parâmetro de suavização κ do filtro de branqueamento;

5.Crie uma matriz *modulo* para o módulo das transformadas;

6. Crie uma matriz *fase* para a fase das transformadas:

Crie uma matriz sigma para a largura das funções de representação conforme a Equação 7. (4.7);

Crie uma matriz funcRep para as funções de representação conforme a Equação (4.6); 8.

Crie uma matriz funcRepTF para a transformada de Fourier das funções de represen-9. tação, em janelas, conforme a Equação (D.2);

10. **Para cada** pixel (x, y) da imagem de entrada **Faça**

Crie uma subimagem I formada pelos pixeis de I_1 correspondentes a uma janela de 11. tamanho $jan \times jan$, centrada em (x, y);

Obtenha a transformada de Fourier da subimagem I; 12.

13.Variando a frequência (ω_x, ω_y) de (0,0) a $(\omega_x max - 1, \omega_y max - 1)$

Obtenha o módulo da transformada de Fourier de I em (ω_x, ω_y) - $|I(\omega_x, \omega_y)|$; 14.

15.Faça $modulo(\omega_x, \omega_y) := |I(\omega_x, \omega_y)|;$

Obtenha a fase da transformada de Fourier de I em (ω_x, ω_y) - $\varphi_{\tilde{I}}(\omega_x, \omega_y)$; 16.

17. **Faça** $fase(x, y, \omega_x, \omega_y) := \varphi_{\tilde{I}}(\omega_x, \omega_y);$

Obtenha sigma $(x, y, \omega_x, \omega_y)$ em $(x, y), (\omega_x, \omega_y)$, utilizando o modulo (ω_x, ω_y) , e o 18. κ do filtro de branqueamento;

Construa as funções de representação $funcRep(x, y, \omega_x, \omega_y)$, em (x, y), (ω_x, ω_y) , 19.utilizando a $fase(x, y, \omega_x, \omega_y)$ e o $sigma(x, y, \omega_x, \omega_y)$; 20. **Fim_Variando**

Fim Para 21.

22.**Para cada** pixel (x, y) da imagem de entrada **Faça**

- Crie uma janela de tamanho $jan \times jan$, iniciando nas coordenadas (x, y); 23.
- 24. Variando a frequência (ω_x, ω_y) de (0,0) a $(\omega_x max - 1, \omega_y max - 1)$

Construa $funcRepTF(\omega_x, \omega_y)$, em (ω_x, ω_y) , utilizando a $fase(x, y, \omega_x, \omega_y)$ e o 25.

 $sigma(x, y, \omega_x, \omega_y);$ 26. Fim Variando

27.Obtenha a transformada inversa de Fourier de FuncRepTF para obter a imagem branqueada janelada IBranqJan, de tamanho $jan \times jan$ (Equação (D.1));

- Obtenha a imagem branqueada em cada pixel (x, y) a partir de IBranqJan; 28.
- 29.Fim Para



Figura 4.6: Magnitude das funções de representação da Equação (4.6), obtidas de um fragmento 3×3 da imagem natural na Figura 4.7a. As frequências representadas (ω_x, ω_y) , de cima para baixo, são: (0,1), (0,2), e (3,1).

As Figuras 4.7 a 4.10 mostram exemplos de imagens naturais (adquiridas da base de van Hateren [46]) codificadas pelas funções de representação $\psi(r; \omega_x, \omega_y)$. Para cada figura, apresentamos a imagem original — a partir da qual obtemos os valores de $\sigma(\omega_x, \omega_y)$ e da função de representação $\psi(r; \omega_x, \omega_y)$ — e a representação branqueada resultante, obtida computando-se a Equação (4.4) sobre janelas finitas. Para fins comparativos, nós também geramos os espectros log-log dos sinais de entrada e dos sinais branqueados (nos gráficos, o eixo vertical é a média rotacional do logaritmo da magnitude das transformadas de Fourier, e o eixo horizontal é log ρ). O que se observa é que a representação tende a equalizar os espectros de entrada, resultando em imagens com a aparência de mapas de bordas, que codificam tanto o 'valor' da borda (a variação de intensidade observada ao cruzá-la) quanto a sua 'polaridade' (o sinal daquela variação).

Nós apresentamos resultados obtidos com janelas 3×3 e 5×5 , e para κ assumindo os valores extremos da faixa considerada ($\kappa = 0.05$ e $\kappa = 1$). A janela menor fornece uma melhor definição para bordas estreitas, porém ao custo de se perderem as mais largas. O efeito de κ não é marcante, mas é consistente com o seu papel como uma medida de nível de ruído, valores maiores desse parâmetro geralmente realçando as componentes de baixa frequência dos espectros.

Além de branquear as imagens naturais, os campos receptivos do nosso modelo apresentam as seguintes propriedades neurofisiologicamente plausíveis: eles aparecem com organização tanto centro-*on* quanto centro-*off*; eles dependem da frequência, e não respondem a imagens uniformes; eles apresentam periferias bem marcadas nas baixas frequências e menos definidas nas altas. Todas estas propriedades já foram experimentalmente verificadas, seja para células ganglionares da retina, seja para células do NGL.



Figura 4.7: Em (a), à esquerda, a imagem original e à direita a imagem branqueada. b) Os espectros log-log correspondentes (ver o texto). Foram utilizadas janelas 5×5 , com $\kappa = 0.05$. Essas imagens, e também as das Figuras 4.8 a 4.10, são 192×192 .

Como forma de validar a nossa proposta de modelagem dos campos receptivos dependentes do estímulo, nós a testamos em um algoritmo de inspiração biológica para estimação de disparidades estereoscópicas, conforme descrito a seguir.



Figura 4.8: Em (a), à esquerda, imagem original e à direita a imagem branqueada. b) Os espectros log-log correspondentes (ver o texto). Foram utilizadas janelas 3×3 , com $\kappa = 0.05$.



Figura 4.9: Em (a), à esquerda, imagem original e à direita, a imagem branqueada. b) Os espectros log-log correspondentes (ver o texto). Foram utilizadas janelas 5×5 , com $\kappa = 0.05$.



Figura 4.10: Em (a), à esquerda, imagem original e à direita, a imagem branqueada. b) Os espectros log-log correspondentes (ver o texto). Foram utilizadas janelas 3×3 , com $\kappa = 1$.

Capítulo 5

Estereoscopia

5.1 Introdução

Muitos neurônios visuais recebem entrada apenas de um olho, sendo portanto chamados de neurônios monoculares, mas em V1 encontram-se também células simples e células complexas que são binoculares — ou seja, que respondem a estimulação tanto do olho esquerdo quanto do olho direito [8]. Estes neurônios estão envolvidos no mecanismo da visão estereoscópica, ou estereoscopia (do grego, 'visão sólida'), que permite a estimação de profundidades a partir das duas projeções, ligeiramente diferentes, produzidas pela cena nas retinas dos dois olhos. O cérebro funde essas duas imagens, e, no processo, obtém informação quanto à profundidade, distância, posição e tamanho dos objetos, gerando a sensação de tridimensionalidade.

A Figura 5.1 ilustra a geometria da visão estereoscópica. Quando um observador fixa seus olhos em um ponto P no espaço, as imagens deste são projetadas nas duas *fóveas* - regiões onde o objeto aparece em foco perfeito -, indicadas pela letra F, na figura. Qualquer ponto no espaço que se situe à mesma distância que o ponto P, como o ponto Q, se projetará em posições correspondentes nas duas retinas (pontos $Q_L e Q_R$), enquanto pontos a diferentes distâncias gerarão projeções não correspondentes. Quanto mais próximo estiver o ponto na cena, mais afastadas horizontalmente estarão as imagens formadas nas duas retinas (pontos $Q'_L e Q'_R$), e quanto mais longe ele estiver, mais próximas estarão as imagens. A essa ausência de correspondência chama-se *disparidade binocular*.

Em [11], empregando funções de Green de equações de casamento¹, foi introduzido um

¹Equações de casamento têm a forma geral $I_1(x + U(x)) = I_2(x)$, onde U(x) é um campo de deslocamentos. Tomando-se uma expansão em série de Taylor do lado esquerdo da relação, obtémse uma equação diferencial que pode ser resolvida pelo método da função de Green [11], fornecendo



Figura 5.1: Geometria da visão estereoscópica [8].

modelo computacional para a estimação de disparidades que se mostrou compatível com abordagens biologicamente plausíveis para a estereoscopia [49, 50, 51, 53]. Tais abordagens, fundadas no chamado *modelo da energia* para os neurônios binoculares, partem da representação dos campos receptivos das células corticais simples em termos de funções de Gabor, e descrevem as respostas (monoculares) de um par em quadratura de tais células, dado um par de imagens estereoscópicas, como

$$\int_{-\infty}^{\infty} dx \ e^{i(\omega x + \phi)} e^{-\frac{(x + \alpha)^2}{2\sigma^2}} I_r(x)$$
(5.1)

onde $I_r(x)$ é a imagem direita, e

$$\int_{-\infty}^{\infty} dx \ e^{i(\omega x + \psi)} e^{-\frac{(x+\delta)^2}{2\sigma^2}} I_l(x)$$
(5.2)

onde $I_l(x)$ é a imagem esquerda, com $\omega \in \sigma$ denotando, respectivamente, a frequência espacial e a largura comuns aos dois campos receptivos, que podem ter parâmetros de fase $(\phi \in \psi)$ e centros ($\alpha \in \delta$) distintos. De acordo com o modelo da energia, a resposta estereoscópica do par de células simples é obtida somando-se as duas integrais acima, enquanto a magnitude quadrada dessa soma constitui a resposta da célula complexa alimentada pelo par de células simples. Esta última resposta fornece a estimativa de disparidade estereoscópica do modelo [50].

Por outro lado, a abordagem de [11] baseou-se no fato de que, filtrando-se a imagem vista sob o campo receptivo *esquerdo* de um célula binocular simples (na representação de Gabor) pela função de Green de uma equação de casamento à *direita* — ou seja, modelando uma disparidade positiva —, pode-se obter uma estimativa do que seria a mesma imagem

 $I_1(x) = \int G(x, x')I_2(x')dx'$, onde G(x, x'), o núcleo de Green da equação considerada, é a sua solução quando $I_2(x) = \delta(x - x')$.

vista, com um certo deslocamento, sob o campo receptivo direito. E vice-versa: filtrandose a imagem vista sob o campo receptivo direito, pela função de Green de uma equação de casamento à esquerda — isto é, modelando uma disparidade negativa —, pode-se obter uma estimativa da mesma imagem vista sob o campo receptivo esquerdo. Computandose a diferença quadrática entre as estimativas da imagem esquerda e direita, geradas com diferentes parâmetros das equações de casamento, pode-se obter o deslocamento (disparidade) em que elas melhor coincidem. Com base nisso, conforme demonstrado em [11], a abordagem tradicional de estimação de disparidades via casamento de imagens fica reformulada em termos compatíveis com os do modelo da energia (ver Apêndice E).

No entanto, seja na formulação da função de Green, seja na do modelo da energia tradicional, a escolha dos parâmetros que vão descrever os campos receptivos das células simples – que constituem a base de ambos os processos – permanece empírica e um tanto arbitrária. Não existe um critério fundamentado que guie a determinação das faixas de valores em que os parâmetros das funções de Gabor ($\omega, \sigma, \phi, \psi, \alpha \in \delta$) devem ser escolhidos. Tampouco existe a garantia de que a faixa de valores adequada para um dado par estereoscópico o seja também para um outro. Motivados por este resultado, nós propomos utilizar, na estereoscopia, as funções de representação sintonizadas, como modelo para os campos receptivos das células corticais. Com isso, os parâmetros desses campos não são mais arbitrariamente escolhidos, já que eles passam a ser determinados pela transformada de Fourier local das imagens de entrada. Como veremos, os núcleos de Gabor do nosso algoritmo para estimação de disparidades tornam-se aqueles que permitem representar, de acordo com a Equação (1.10), cada imagem do par estereoscópico. Ademais, visando a tornar a nossa abordagem ainda mais compatível com as condições da visão biológica, nós decidimos trabalhar com imagens de entrada branqueadas segundo o modelo para os campos receptivos centro-periferia introduzido na Seção 4.3. Dessa forma, o nosso algoritmo de estimação estereoscópica, a ser apresentado a seguir, passa a incorporar características da representação neurofisiológica dos estímulos, tanto no nível cortical (células simples e complexas do córtex visual) quanto no sub-cortical (células centroperiferia da retina e do NGL). O algoritmo em si constitui uma modificação da abordagem da função de Green [11], que a um só tempo a simplifica e a torna mais compatível com técnicas tradicionais de modelagem dos sistemas neurais.

5.2 Estimação de Disparidades Biologicamente Inspirada

Na abordagem tradicional da função de Green [11], a estimação de disparidades procede da seguinte forma (ver Apêndice E): as imagens de entrada, $I_l \in I_r$, são multiplicadas por funções de Gabor, produzindo sinais complexos que são então filtrados pelos núcleos de Green de equações diferenciais que modelam um casamento com disparidades *lineares* – ou seja,

$$I_l(x + U(x), y) = I_r(x, y)$$
(5.3)

com U(x) = u + vx, para $u \in v$ constantes. O sinal complexo associado à imagem esquerda é filtrado pelo núcleo de Green de uma equação de casamento à direita (U(x) > 0), e aquele associado à imagem direita é filtrado pelo núcleo correspondente a um casamento à esquerda (U(x) < 0). Variando-se os parâmetros das funções de Green, diferentes deslocamentos relativos podem ser induzidos sobre os sinais esquerdo e direito, e o deslocamento espacial ótimo em cada ponto nas imagens pode ser obtido avaliando-se a correspondência local entre os sinais deslocados. Isto fornece uma estimativa do mapa de disparidades codificado pelo par estereoscópico.

Como anteriormente citado, em tal abordagem os parâmetros dos filtros de Gabor (assim como os das funções de Green) devem ser empiricamente determinados caso a caso. Nós verificamos, porém, que com o uso das funções de representação sintonizadas é possível eliminar este problema, ao mesmo tempo simplificando a abordagem de Green original: em lugar de trabalhar com os núcleos de Green que modelam campos de disparidade lineares, nós podemos nos restringir àqueles, mais simples, associados a disparidades uniformes. Embora esta última classe seja evidentemente menos rica do que a dos campos lineares – que incorporam variabilidade espacial — esta deficiência é compensada pela informação local, e dependente do sinal, que se ganha com o uso da representação sintonizada. A abordagem proposta é a seguinte: para a estimação do campo de disparidades associado a um dado par de imagens estereoscópicas, nós computamos, em cada posição x, a medida

$$R(x) = |Y_r(x) - Y_l(x)|^2$$
(5.4)

com

$$Y_r(x) = \int K_k(x - x')\hat{I}_l(x')I_r(x')dx'$$
(5.5)

$$Y_l(x) = \int K_k^{(-)}(x - x')\hat{I}_r(x')I_l(x')dx'$$
(5.6)

onde

$$K_k(x - x') = \begin{cases} 2ke^{-(1-i)k(x-x')}, & \text{se } x > x' \\ 0, & \text{se } x < x' \end{cases}$$
(5.7)

representa o núcleo de Green para uma equação de casamento à direita, com campo de disparidades uniforme, U(x) = 1/k = constante (ver Apêndice E), e $K_k^{(-)}(x - x') = K_k(x' - x)$ representa o núcleo de Green correspondente, para um casamento à esquerda². Nas equações (5.5) e (5.6), \hat{I}_l e \hat{I}_r denotam as representações sintonizadas das imagens de entrada esquerda e direita, respectivamente, ou seja (ver Capítulo 1),

$$\hat{I}_l(x) = \frac{1}{\left(2\pi\right)^{3/2}} \int \int e^{i[\omega x + \varphi_l(\omega)]} e^{-\frac{(x-b)^2}{2\sigma_l^2(\omega)}} d\omega db$$
(5.8)

onde $\sigma_l(\omega)$ e $\varphi_l(\omega)$ são, respectivamente, a magnitude e a fase de $I_l(x)$. Uma expressão similar vale para $\hat{I}_r(x)$.

Empregando a representação acima na Equação (5.5), nós obtemos

$$Y_r(x) = \frac{1}{(2\pi)^{3/2}} \int d\omega \int db \int dx' K_k(x - x') e^{i[\omega x' + \varphi_l(\omega)]} e^{-\frac{(x' - b)^2}{2\sigma_l^2(\omega)}} I_r(x')$$
(5.9)

onde se pode identificar, como na abordagem de Green original, o produto de uma das imagens do par estereoscópico por uma função de Gabor, e a filtragem deste produto por um núcleo de Green — neste caso, aquele de uma equação de casamento *uniforme*. Também em contraste com a abordagem original, aqui os parâmetros da função de Gabor não são arbitrários, mas estão associados, por intermédio da representação sintonizada, à transformada de Fourier da segunda imagem do par estereoscópico. De forma semelhante, $Y_l(x)$ é obtido como

$$Y_l(x) = \frac{1}{(2\pi)^{3/2}} \int d\omega \int db \int dx' K_k^{(-)}(x - x') e^{i[\omega x' + \varphi_r(\omega)]} e^{-\frac{(x'-b)^2}{2\sigma_r^2(\omega)}} I_l(x')$$
(5.10)

Nos experimentos reportados a seguir, nós empregamos, como sinais de entrada, versões branqueadas das imagens "retinais", $I_l \in I_r$, obtidas conforme a Seção 4.3. Desta forma, nós modelamos o casamento estereoscópico como se processando sobre a saída dos estágios pré-corticais do caminho visual. A estrutura geral do nosso algoritmo é semelhante àquela da abordagem tradicional da função de Green [11], sumariada no Apêndice

²Como nós admitimos valores tanto positivos quanto negativos para k, é fácil verificar que a nossa abordagem também contempla a situação em que o sinal I_l é modulado pelo núcleo $K_k^{(-)}$ – ou seja, sofre um deslocamento para a esquerda – e o sinal I_r é modulado pelo K_k , sofrendo um deslocamento para a direita.

E. No entanto, o uso do núcleo de Green do casamento uniforme, e a obtenção dos parâmetros das funções de Gabor a partir da tranformada de Fourier dos sinais de entrada simplificam bastante o processo. Como veremos, com um algoritmo consideravelmente mais simples que o de [11], nós conseguimos obter resultados equivalentes, e mesmo superiores, aos ali reportados. Antes de passarmos aos experimentos, porém, é interessante observar que a medida R(x) pode ser expressa sob a forma de uma resposta de Volterra de segunda ordem [12], conforme demonstrado em seguida.

A série de Volterra constitui um análogo, para funcionais, da série de Taylor, e tem sido utilizada como modelo para as respostas não-lineares de sistemas neurais, expressando a atividade de um neurônio em termos de potências do estímulo recebido³. Uma série de Volterra associada à resposta de um neurônio binocular envolveria termos da forma

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} D_2(\xi, \xi') I_l(x-\xi) I_r(x-\xi') d\xi d\xi'$$
(5.11)

onde $D_2(\xi, \xi')$ denota um núcleo de Volterra de segunda ordem. Efetuando a mudança de variáveis $\xi' \to x'$, com $x' = x - \xi'$, isto pode ser reescrito como

$$\int_{-\infty}^{\infty} \left[\int_{-\infty}^{\infty} D_2(\xi, x - x') I_l(x - \xi) d\xi \right] I_r(x') dx'$$
(5.12)

Assumindo que $D_2(\xi, x - x')$ possa ser expresso como $D_2(\xi, x - x') = K_k(\xi)\delta(\xi - x + x')$, nós então obteríamos

$$\int_{-\infty}^{\infty} K_k(x - x') I_l(x') I_r(x') dx'$$
(5.13)

que se torna a mesma expressão proposta acima para $Y_r(x)$ (Equação (5.5)), quando a representação sintonizada, $\hat{I}_l(x')$, é utilizada em lugar de $I_l(x')$. Uma análise similar vale para $Y_l(x)$, e portanto para a medida de correspondência, R(x). A nossa abordagem estereoscópica pode ser assim interpretada como um modelo de Volterra de segunda ordem, tendo a função de Green como núcleo.

Sobre a função de Green, um último comentário é importante. Conforme já observado em [11], esta função constitui um elemento da nossa abordagem estereoscópica que não encontra paralelo no modelo da energia para os neurônios binoculares. No entanto, se levarmos em conta os estágios pré-corticais do caminho visual — como fazemos aqui é possível sugerir uma interpretação neurofisiológica também para os núcleos de Green. Como visto na Seção 4.1, os sinais provenientes das células ganglionares das duas retinas se

³A resposta temporal de um neurônio seria expressa como $r(t) = r_0 + \int d\tau_1 D_1(\tau_1) s(t - \tau_1) + \int d\tau_1 d\tau_2 D_2(\tau_1, \tau_2) s(t - \tau_1) s(t - \tau_2) + \int d\tau_1 d\tau_2 d\tau_3 D_3(\tau_1, \tau_2, \tau_3) s(t - \tau_1) s(t - \tau_2) s(t - \tau_3) + \dots$, onde D_n denota o núcleo de Volterra de ordem n.

transmitem, por intermédio dos eixos ópticos, aos núcleos geniculados laterais de cada um dos hemisférios cerebrais. Antes de atingir os NGLs, os impulsos neuronais segregam-se de uma maneira característica, quando os eixos ópticos se cruzam para formar o chamado quiasma óptico (ver Figura 5.2). No quiasma, as fibras ópticas se dispõem de tal modo que a informação correspondente aos campo visuais direitos dos dois olhos se projeta no NGL do hemisfério cerebral esquerdo, enquanto os campos visuais esquerdos se projetam sobre o NGL direito. Cada NGL, portanto, só processa informação proveniente de uma metade dos campos visuais. Se considerarmos, por exemplo, que um dos olhos esteja fixado num ponto de coordenada horizontal x no espaço – ou seja, este ponto se projeta sobre a fóvea – os dois hemicampos visuais definidos pelas condições $\{x' : x' < x\}$ e $\{x' : x' > x\}$ se projetarão sobre NGLs distintos. Efetivamente, é como se, para um dos geniculados, as entradas provenientes de x' > x fossem zeradas, o mesmo ocorrendo, com o outro geniculado, para aquelas provenientes de x' < x.



Figura 5.2: Vista superior do Caminho Visual [52].

Agora, se considerarmos as integrais que definem as respostas $Y_l(x)$ e $Y_r(x)$, do nosso modelo, é fácil verificar que este incorpora um mecanismo semelhante. A forma funcional do núcleo de Green (Equação (5.7)) implica em que, para a obtenção de $Y_r(x)$ (Equação (5.5)), apenas as entradas com x' < x sejam consideradas, enquanto somente aquelas com x' > x contribuem para a obtenção de $Y_l(x)$ (essas duas situações se invertem quando k <0). O núcleo de Green cumpre portanto o papel de incorporar, no nosso modelo, o dado neurofisiológico da segregação dos campos visuais no quiasma óptico. Evidentemente, este papel se explica tão-somente pelo fato de o núcleo se anular em metade do seu domínio, não servindo para justificar os detalhes da sua forma funcional na outra metade. Entretanto, do ponto de vista neurofisiológico, a forma do $K_k(x - x')$ na Equação (5.7) não parece inadequada. Trata-se de um filtro cuja saída se concentra em torno de x = x', e que poderia descrever, por exemplo, um *pooling*⁴ das células vizinhas à posição x, para produzir a resposta neuronal em x. Ademais, como as oscilações senoidais geram lobos negativos em K_k , à medida em que x' se afasta de x, o núcleo de Green implementaria simultaneamente um mecanismo de inibição lateral (ver [54], por exemplo). O estudo de possíveis alternativas para o filtro de Green empregado aqui é uma das propostas a serem exploradas no seguimento das nossas pesquisas.

Antes de apresentar os nossos experimentos, fornecemos a seguir o pseudocódigo do procedimento empregado para gerar o mapa de disparidades a partir de um par de imagens retinais branqueadas.

⁷⁷

 $^{^4\}mathrm{Como}$ proposto, por exemplo, em [53].

Procedimento para gerar o mapa de disparidades a partir de uma par de imagens retinais branqueadas:

- 1. Entre com a imagem original direita branqueada I_r ;
- 2. Entre com a imagem original esquerda branqueada I_l ;
- 3. Defina o número máximo de frequências ω_{max} na direção x;
- Crie um vetor freq para armazenar as frequências utilizadas, na direção x; 4.
- Crie um vetor $I_r linha$ para cada linha da imagem direita; 5.
- Crie um vetor $I_l linha$ para cada linha da imagem esquerda; 6.
- 7.
- 8.
- Crie um vetor $sigma_Ir$ para o módulo da transformada de Fourier de $I_r linha;$ Crie um vetor $fase_Ir$ para a fase da transformada de Fourier de $I_r linha;$ Crie um vetor $sigma_Il$ para o módulo da transformada de Fourier de $I_l linha;$ Crie um vetor $fase_Il$ para a fase da transformada de Fourier de $I_l linha;$ 9.
- 10.

Crie um vetor $I_r linha\dot{S}$ para acumular as representações sintonizadas das linhas da i-11. magem direita, inicializado com zero;

12.Crie um vetor $I_l linhaS$ para acumular as representações sintonizadas das linhas da imagem esquerda, inicializado com zero;

Crie um vetor GreenDir para os núcleos de Green de uma equação de casamento à direita; 13.

14. Crie um vetor GreenEsq para os núcleos de Green de uma equação de casamento à esquerda;

15.Defina o número de iterações *itera* para gerar a faixa de disparidades;

16.Defina o intervalo *passo* entre as iterações;

Defina o desvio padrão sigmaX da Gaussiana na direção x; 17.

18. Defina o desvio padrão sigmaY da Gaussiana na direção y;

19. Se a disparidade das imagens de entrada for muito alta Então

20.

25.

Fazer alinhamento inicial: Implementar casamento uniforme em toda a extensão 21.da imagem, com núcleos de Green como os da Equação (E.3).

22.Defina a variação de disparidades *dispNum*;

23.24.

Defina o intervalo passoC entre as iterações; Para cada iteração k de -dispNum a dispNum Faça

Obtenha a disparidade em comprimento u para k com passoC;

Filtre a imagem de entrada esquerda por um núcleo de Green à direita, uti-26.lizando a constante u, para gerar um deslocamento uniforme positivo;

Filtre a imagem de entrada direita por um núcleo de Green à esquerda, uti-27.lizando a constante u, para gerar um deslocamento uniforme negativo;

Calcule a diferença absoluta das imagens filtradas (e também das entradas 28.originais: deslocamento zero); 29.

- Escolha o valor do deslocamento que minimiza essa medida;
- 30. Fim Para
- 31. Fim Se

32.Considera as dez menores frequências e um subconjunto das demais frequências em passos de dez em dez, nas duas imagens.

Para cada linha y da imagem direita I_r ($I_r linha$) e da imagem esquerda I_l ($I_l linha$) 33. Faça

34. Obtenha a transformada de Fourier de $I_r linha$;

35.Obtenha a transformada de Fourier de $I_l linha$;

- 36. Para cada ω de freq(0) a $freq(\omega_{max})$ Faça
- 37. Parâmetros de I_rlinha em cada frequência ω :
- 38. Obtenha o módulo da transformada de Fourier de $I_r linha em \omega - |I_r linha(\omega)|$;

39. sigma $Ir(\omega) := |\tilde{I}_r linha(\omega)|;$

Obtenha a fase da transformada de Fourier de $I_r linha$ em ω - $\varphi_{\tilde{I}_r linha}(\omega)$; 40.

- 41. $fase_Ir(\omega) := \varphi_{\tilde{I}_r linha}(\omega);$
- 42. Parâmetros de I_llinha em cada frequência ω :
- Obtenha o módulo da transformada de Fourier de $I_l linha em \omega |I_l linha(\omega)|$; 43.
- 44. $sigma_Il(\omega) := |I_llinha(\omega)|;$
- Obtenha à fase da transformada de Fourier de $I_l linha$ em ω $\varphi_{\tilde{I}_l linha}(\omega)$; 45.

46. $fase_Il(\omega) := \varphi_{\tilde{I}_llinha}(\omega);$

47. **Fim** Para $\{\omega\}$

```
48.
           Para cada pixel x de I_r linha e I_l linha Faça
49.
              Obtenha a representação sintonizada de I<sub>r</sub>linha;
50.
              Obtenha a representação sintonizada de I<sub>l</sub>linha;
51.
              Para cada \omega de freq(0) a freq(\omega_{max}) Faça
                 Para cada posição b da imagem de entrada Faça
52.
                      I_r linha(x) := função de Gabor com os parâmetros de sigma Ir(\omega) e
53.
fase_Ir(\omega) com centro em b e frequência \omega;
                    I_r linhaS(x) := I_r linhaS(x) + I_r linha(x);
54.
                       I_l linha(x) := função de Gabor com os parâmetros de sigma Il(\omega) e
55.
fase Il(\omega) com centro em b e frequência \omega;
56.
                    I_l linhaS(x) := I_l linhaS(x) + I_l linha(x);
57.
                 Fim Para \{b\}
58.
              Fin Para \{\omega\}
           Fim Para \{x\}
59.
60.
           Construa as Equações (5.9) e (5.10).
           Para cada pixel x de I_r linha e I_l linha Faça
61.
              Para cada disparidade k de (\frac{-itera}{2} * passo) a (\frac{-itera}{2} * passo) Faça
62.
                 Para cada posição x' de I_r linha \in I_l linha Faça
63.
                    Construa os núcleos de Green para uma equação de casamento à direita (Eq.
64.
5.7):
65.
                    GreenDir := K_k(x - x'), para x > x'
                    Acumule em Y_r(x) o produto da imagem direita em x' com a representação
66.
da imagem esquerda em x' filtrado pelo núcleo de Green GreenDir.
67.
                    Construa os núcleos de Green para uma equação de casamento à esquerda:
                    GreenEsq := K_k(x - x'), para x < x';
68.
69.
                    Acumule em Y_l(x) o produto da imagem esquerda em x' com a representação
da imagem direita em x' filtrado pelo núcleo de Green GreenEsq;
70.
                 Fim Para \{x'\}
71.
                  Construa a estimativa do campo de disparidades em cada posição x (Equação
(5.4))
72.
                 R(x) = |Y_l(x) - Y_r(x)|^2;
73.
                  Pondere o R(x) por uma Gaussiana bidimensional de parâmetros sigmaX e
sigmaY;
74.
                  Obtenha a medida de casamento com a resposta inversamente proporcional a
R(x);
              Fim Para \{k\}
75.
76.
              Obtenha a média das respostas locais R(x);
77.
           Fim Para \{x\}
78.
        Fim Para \{y\}
```

5.3 Experimentos

Assim como em [11], o casamento estereoscópico se faz linha por linha, assumindo-se geometria epipolar horizontal. Dado um par de imagens estereoscópicas, os sinais $I_l(x)$ e $I_r(x)$, das Equações (5.9) e (5.10), são pares de linhas correspondentes nas versões branqueadas daquelas imagens (para o branqueamento, nós usamos o parâmetro $\kappa =$ 0,05 - ver Equação (4.8)), cujas transformadas de Fourier, por sua vez, fornecerão os parâmetros $\sigma_l(\omega)$, $\varphi_l(\omega)$, $\sigma_r(\omega)$, e $\varphi_r(\omega)$, das funções moduladoras de Gabor. O número de frequências ω sobre as quais nós devemos somar os integrandos em (5.9) e (5.10), de modo a implementar as versões discretizadas daquelas equações, está associado ao tamanho das linhas das imagens de entrada. Nós verificamos que, em geral, são as frequências mais baixas que definem o mapa de disparidades resultante, e portanto, nos experimentos aqui reportados, nós empregamos todas as dez menores frequências, preservando apenas um subconjunto das demais, escolhidas em passos de dez em dez. Variações sobre este esquema não produziram alteração apreciável sobre os resultados. Os valores do parâmetro b, que determina o centro da janela Gaussiana sob a qual os sinais são considerados, foram incrementados em passos de um píxel.

Variando-se o parâmetro k dos núcleos de Green em (5.9) e (5.10), diferentes disparidades (de valor 2/k) são induzidas sobre o par de sinais modulados por Gabor. Em cada ponto do plano imagem, e para cada valor de disparidade, a medida de casamento é então obtida como a média, ponderada por uma Gaussiana bidimensional, das respostas locais R(x). Assim como em [11], em vez de escolher como valor da disparidade local aquele que minimiza a medida de casamento, nós o tomamos como a soma de todas as disparidades consideradas, cada uma delas ponderada pelo inverso da resposta R(x) correspondente. Isto nos fornece um mapa de disparidades denso, sem requerer suavização adicional. Ainda seguindo [11], como forma de lidar com disparidades muito altas, nós efetuamos um pré-alinhamento de cada par estereoscópico, implementando um casamento uniforme baseado em toda a extensão das imagens, e não linha por linha (ver Apêndice E). Os únicos parâmetros livres do nosso algoritmo são, portanto, as faixas de disparidades consideradas (no pré-alinhamento e no próprio casamento), e as dimensões do filtro Gaussiano utilizado para ponderar as respostas locais.

Experimentos com Imagens Sintéticas

As Figuras 5.3 e 5.4 apresentam os resultados obtidos a partir de estereogramas sintéticos, de dimensão 128×128 . As imagens de entrada, originais e branqueadas, aparecem



Figura 5.3: Estereograma de pontos aleatórios denso. Em (a), são mostradas as imagens originais, em (b), as imagens branqueadas, e, em (c), os mapas de disparidades real e estimado.

juntamente com os mapas de disparidades real e estimado. A Figura 5.3 refere-se ao experimento com um estereograma de pontos aleatórios binário, de densidade 50%. A porção central das imagens, de 96 × 96 píxeis, encontra-se em registro, exceto pela região quadrada em seu interior, de tamanho 64×64 , onde a disparidade é positiva, de um píxel. Já a borda externa apresenta disparidade de -1 píxel. O nosso mapa de disparidades foi estimado considerando-se nove possíveis valores para o parâmetro k, correspondentes a disparidades entre -3,2 e 3,2 píxeis. A função de ponderação Gaussiana utilizada tinha desvio-padrão, em x, de 2 píxeis, e, em y, de 3,4 píxeis.

A Figura 5.4 corresponde a um estereograma com disparidades contínuas. O nosso mapa de disparidades foi estimado em 11 iterações (11 valores de k), correspondendo a disparidades entre -3 e 3 píxeis. A função de ponderação da Gaussiana tinha desviopadrão de 1 píxel, tanto em x como em y.

A Figura 5.5 mostra o resultado do experimento com o estereograma sintético Sawtooth, de dimensão 95×108 píxeis, comumente utilizado para benchmarking de algoritmos



Figura 5.4: Estereograma de disparidade contínua. Em (a), são mostradas as imagens originais, em (b), as imagens branqueadas, e, em (c), os mapas de disparidades real e estimado.

de estereoscopia. O mapa de disparidades foi estimado em 15 iterações, correspondentes a disparidades entre -5,95 e 5,95 píxeis. A função de ponderação Gaussiana tinha desviopadrão de 1,8 píxeis, tanto em x quanto em y. Neste experimento realizou-se um alinhamento inicial considerando-se disparidades entre -1,38 e 1,38 píxeis.

Exceto pelos erros nas bordas de disparidade, as estimativas para os estereogramas denso e de disparidade contínua estão bem próximas ao real. A reconstrução do *Sawtooth* também apresentou um mapa de disparidades qualitativamente correto.

Experimentos com Imagens Reais

Nas Figuras 5.6 a 5.10, nós apresentamos experimentos com estereogramas reais, cujos mapas de disparidade não são conhecidos a priori. Todos os nossos resultados são qualitativamente corretos, e, assim como os dos experimentos com as imagens sintéticas, são em geral comparáveis — e, em certos casos, superiores — aos fornecidos tanto pela abordagem de Green tradicional [11] como pela versão modificada introduzida em [62]. Nesta, que representa uma proposta anterior à aqui apresentada, o mesmo algoritmo de



Figura 5.5: Estereograma *Sawtooth*. Em (a), são mostradas as imagens originais, em (b), as imagens branqueadas, e, em (c), os mapas de disparidades real e estimado.

[11] foi implementado sobre sinais de entrada branqueados.

Na Figura 5.6, nós observamos o resultado para o estereograma *Motherboard*, de dimensão 97×128 . O seu mapa de disparidades foi calculado em 7 iterações, abrangendo as disparidades de -9 a +9 píxeis. A função de ponderação utilizada tinha desvio-padrão de 2 píxeis, tanto em x como em y.

A Figura 5.7 mostra o resultado para o estereograma *Pentagon*, de resolução 128×128 píxeis. O seu mapa de disparidades foi calculado em 21 iterações, cobrindo as disparidades entre -5 e +5 píxeis. A função de ponderação Gaussiana tinha um desvio-padrão de 2,5 píxeis, em x e em y. Para comparação, nós apresentamos também os resultados obtidos com as duas abordagens de Green anteriores.

A Figura 5.8 apresenta o resultado para o estereograma *Tree*, de dimensão 117×128 . O seu mapa de disparidades foi estimado em 21 iterações, correspondendo às disparidades de -4 a +4 píxeis. A função de ponderação Gaussiana tinha desvio-padrão, em x, de 1,92 píxeis, e em y, de 3,2 píxeis. Aqui também, nós apresentamos os resultados obtidos com as duas abordagens de Green anteriores.

O resultado para o estereograma Shrub, de 120×128 píxeis, é apresentado na Figura 5.9. O seu mapa de disparidades foi estimado em 7 iterações do nosso algoritmo, correspondendo a disparidades entre -2,7 e +2,7 píxeis. A função Gaussiana utilizada tinha desvio-padrão de 2 píxeis, tanto em x como em y. Foi implementado um alinhamento inicial, considerando valores de disparidade entre -1,38 e 1,38 píxeis. Os resultados obtidos com as duas abordagens de Green anteriores também são apresentados.

Finalmente, na Figura 5.10, nós mostramos o resultado para o estereograma *Meter*, de dimensão 120×128 . O mapa de disparidades foi obtido em 11 iterações, correspondendo a disparidades de -3,5 a 3,5 píxeis. A função de ponderação utilizada tinha desvio-padrão, em x, de 1,92 píxeis, e em y, de 3,2 píxeis. Foi feito um alinhamento inicial, considerando valores de disparidades entre -0,9 e 0,9 píxeis. Neste caso, o resultado obtido com a nova versão da abordagem de Green é claramente superior aos fornecidos pelas versões anteriores.



Figura 5.6: Estereograma *Motherboard*. Em (a), são mostradas as imagens originais, em (b), as imagens branqueadas, e, em (c), o mapa de disparidades.



Figura 5.7: Estereograma *Pentagon*. Em (a), são mostradas as imagens originais, em (b), as imagens branqueadas, em (c), o nosso mapa de disparidades, e, em (d), o mapa de disparidades obtido com as abordagens de Green anteriores: à esquerda, aquela de [62]; à direita, aquela de [11].



Figura 5.8: Estereograma *Tree*. Em (a), são mostradas as imagens originais, em (b), as imagens branqueadas, em (c), o nosso mapa de disparidades, e, em (d), o mapa de disparidades obtido com as abordagens de Green anteriores: à esquerda, aquela de [62]; à direita, aquela de [11].



Figura 5.9: Estereograma *Shrub*. Em (a), são mostradas as imagens originais, em (b), as imagens branqueadas, em (c), o nosso mapa de disparidades, e, em (d), o mapa de disparidades obtido com as abordagens de Green anteriores: à esquerda, aquela de [62]; à direita, aquela de [11].



Figura 5.10: Estereograma *Meter*. Em (a), são mostradas as imagens originais, em (b), as imagens branqueadas, em (c), o nosso mapa de disparidades e, em (d), o mapa de disparidades obtido com as abordagens de Green anteriores: à esquerda, aquela de [62]; à direita, aquela de [11].

Capítulo 6

Conclusões e Trabalhos Futuros

A presente Tese versou sobre as funções de Gabor sintonizadas e suas aplicações à análise e à síntese de sinais. Em resumo, as contribuições apresentadas foram as seguintes:

 Extensão, para o domínio da frequência, do modelo de representação de sinais espaciais introduzido em [1] (Capítulo 1).

2. Introdução da *transformada de Gabor sintonizada*, cujas duas versões — no domínio espacial e no domínio da frequência — apresentam, em conjunto, as propriedades da transformada de Wigner, considerada ideal para a análise espaço-frequência (Capítulo 2).

3. Aplicação da primeira forma da transformada sintonizada à análise de sinais nãoestacionários de diferentes naturezas (Capítulo 2), e, em especial, à análise de eletroencefalogramas (EEGs) de pacientes com diagnóstico de epilepsia (Capítulo 3).

4. Emprego da extensão bidimensional das funções sintonizadas de Gabor, para a modelagem dos campos receptivos das células corticais simples (Capítulo 4).

5. Proposta de funções sintonizadas rotacionalmente simétricas, para a modelagem do processo de *branqueamento* de imagens pelas células da retina e do núcleo geniculado lateral (Capítulo 4).

6. Utilização dos resultados de modelagem neuronal, apresentados no Capítulo 4, para a definição de uma abordagem de inspiração biológica para a estimação de disparidades estereoscópicas, incorporando características de mecanismos neurofisiológicos corticais e pré-corticais (Capítulo 5).

Do ponto de vista da análise de sinais, a abordagem de Gabor sintonizada parece encontrar um campo de aplicação promissor no auxílio ao diagnóstico da epilepsia, conforme evidenciado pelos experimentos aqui reportados. Em nosso trabalho futuro, tencionamos seguir explorando este campo, em particular visando estender a nossa análise de EEGs para sinais multicanais. Com este objetivo, ampliaremos a nossa colaboração com a Dra. Glenda Lacerda, e com o setor de epilepsia do Hospital Universitário Antônio Pedro.

Em uma vertente mais teórica, pretendemos explorar o conceito expandido de representação sintonizada de Gabor, que se obtém com o emprego da transformada fracionária de Fourier [55, 56]. A transformada fracionária de Fourier de parâmetro α é definida como

$$I_{\alpha}(u) \equiv \text{FRFT}_{\alpha}\{I(x)\}(u) = \sqrt{\frac{1 - i \cot \alpha}{2\pi}} e^{i\frac{\cot \alpha}{2}u^2} \int_{-\infty}^{\infty} dx I(x) e^{i\frac{x^2}{2}\cot \alpha - ixu \csc \alpha}$$
(6.1)

onde $x \in u$ aqui são adimensionais

Verifica-se que, para k inteiro, quando $\alpha = 2k\pi$, $I_{\alpha}(u) = I(u)$, e quando $\alpha = (2k+1)\pi$, $I_{\alpha}(u) = I(-u)$. Assim, o sinal original, I(x), pode ser identificado como $I_0(x)$. Por outro lado, para $\alpha = \frac{\pi}{2}$, obtém-se a transformada de Fourier,

$$I_{\frac{\pi}{2}}(u) \equiv \frac{\tilde{I}(u)}{\sqrt{2\pi}} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dx I(x) e^{-ixu}$$
(6.2)

A FRFT equivale a uma representação intermediária entre espaço e frequência, correspondendo a uma rotação de ângulo α no domínio espaço-frequência. A aplicação de duas transformadas fracionárias de Fourier sucessivas, com ângulos α e β , equivale à aplicação de uma única transformada com ângulo $\alpha + \beta$, e a inversa de uma FRFT com ângulo α é a FRFT com ângulo $-\alpha$.

Em [57], demonstrou-se que é possível generalizar a abordagem sintonizada de Gabor, definindo funções de representação cujos parâmetros são obtidos a partir da FRFT do sinal de entrada. As duas funções de representação sintonizada consideradas aqui — no domínio espacial e no domínio da frequência — resultam então como casos particulares desta abordagem mais geral.

Toda a análise teórica e experimental da representação sintonizada estendida permanece aberta à exploração, assim como a possibilidade de uma extensão ainda mais ampla do mesmo modelo, com base na transformada canônica linear [58], que tem a transformada fracionária de Fourier como caso particular.

No que se refere a aplicações de síntese de sinais, continuaremos explorando as possibilidades de emprego da abordagem sintonizada na modelagem neural. Em particular, pretendemos consolidar o nosso modelo para os campos receptivos dependentes do estímulo, tentando obter resultados quantitativos mais concludentes (quanto às dimensões e ao comportamento das larguras de banda em orientação e frequência, por exemplo), que permitam comprovar a adequação das nossas funções de representação.

No que diz respeito à aplicação à estereoscopia, continuaremos perseguindo o desenvolvimento de uma abordagem neurofisiologicamente plausível para a estimação de disparidades. A abordagem biologicamente inspirada apresentada no Capítulo 5, embora constitua uma pioneira e bem-sucedida proposta de estimação de disparidades fundada na emulação dos estágios cortical e pré-cortical do caminho visual, pode ser aperfeiçoada em vários aspectos. Em primeiro lugar, em seu formato atual, o nosso algoritmo constitui uma abordagem do tipo Linear/Não-linear [59], com a saída do estágio linear do pré-alinhamento alimentando a etapa não-linear do casamento propriamente dito. Uma proposta mais elegante seria a de substituir a etapa de pré-alinhamento por termos lineares adicionais de Volterra, no cálculo das respostas $Y_r(x)$ e $Y_l(x)$ – Equações (5.5) e (5.6) –, do tipo

$$\int D_1^r(x-x')I_r(x')dx' + \int D_1^l(x-x')I_l(x')dx'$$
(6.3)

onde se poderia assumir, como primeira aproximação, que os núcleos $D_1^r \in D_1^l$ seriam as próprias funções de Green, $K_k \in K_k^{(-)}$. De forma semelhante, termos de auto-interação não-linear das imagens de entrada poderiam ser incluídos, sob a forma

$$\int D_2^{rr}(x-x')I_r(x')I_r(x')dx' + \int D_2^{ll}(x-x')I_l(x')I_l(x')dx'$$
(6.4)

onde, como antes, um dos sinais, em cada integrando, seria substituído por sua representação sintonizada, \tilde{I}_r ou \tilde{I}_l . Também numa primeira aproximação, poderíamos identificar $D_2^{rr} \in D_2^{ll}$ a $K_k \in K_k^{(-)}$, respectivamente. Uma outra possibilidade seria empregar, para todas as interações de segunda ordem, as funções de Green associadas ao campo de disparidades linear (ver Apêndice E), reservando as do casamento uniforme para os núcleos de Volterra lineares (Equação 6.3).

Finalmente, devemos observar que a nossa interpretação da função de Green como emulando a segregação dos sinais no quiasma óptico requer a suposição de que os olhos foquem consecutivamente em cada coordenada x do espaço, para a estimação da disparidade naquele ponto.¹ Um modelo mais realístico, talvez, consideraria a estimação da disparidade em todos as coordenadas x, assumindo os olhos focados num ponto fixo x_0 , que depois se faria varrer todo o plano-imagem. Neste caso, a Equação (5.5), por exemplo,

¹A consideração das Equações (5.5) a (5.7), e da Figura 5.2, deve conduzir facilmente a esta conclusão.
seria substituída pela forma

$$Y_r(x) = \int \left[\int K_k(x - x')u(x_0 - x')\tilde{I}_l(x')I_r(x')dx' \right] dx_0$$
(6.5)

onde $\boldsymbol{u}(\boldsymbol{x})$ denota a função degrau unitário.

APÊNDICE A

Neste apêndice, demonstraremos algumas das identidades apresentadas no Capítulo 1, para as duas abordagens de representação de sinais.

Primeira abordagem de representação

A identidade na Equação (1.3) pode ser verificada, mudando-se a variável de integração para ω' e tomando a transformada de Fourier em ambos os lados da equação, para obter

$$\tilde{I}(\omega) = \frac{1}{(2\pi)^{3/2}} \int_{-\infty}^{\infty} \mathcal{F}\left\{e^{i\omega'x} * e^{i[\omega'x + \varphi_{\tilde{I}}(\omega')]} e^{\frac{-x^2}{2\sigma^2(\omega')}}\right\} d\omega'$$
(A.1)

e, usando a propriedade da transformada da convolução,

$$\tilde{I}(\omega) = \frac{1}{(2\pi)^{3/2}} \int_{-\infty}^{\infty} \mathcal{F}\left\{e^{i\omega'x}\right\} \times \mathcal{F}\left\{e^{i[\omega'x+\varphi_{\tilde{I}}(\omega')]}e^{\frac{-x^2}{2\sigma^2(\omega')}}\right\} d\omega'$$
(A.2)

Ambas as transformadas no integrando são conhecidas, resultando em

$$\tilde{I}(\omega) = \int_{-\infty}^{\infty} \delta(\omega - \omega') \sigma(\omega') e^{i\varphi_{\tilde{I}}(\omega')} e^{-\frac{\sigma^2(\omega')}{2}(\omega - \omega')^2} d\omega'$$
(A.3)

onde δ é a função delta de Dirac (sinal impulso unitário). Finalmente, usando a propriedade de filtragem da delta, obtemos $\tilde{I}(\omega) = \sigma(\omega)e^{i\varphi_{\tilde{I}}(\omega)}$, o que está consistente com as nossas definições para $\sigma(\omega) \in \varphi_{\tilde{I}}(\omega)$.

Explicitando a operação de convolução na Equação (1.3), esta pode ser reescrita como

$$I(x) = \int_{-\infty}^{\infty} e^{i\omega x} d\omega \int_{-\infty}^{\infty} \frac{1}{(2\pi)^{3/2}} e^{-i\omega\xi} e^{i[\omega\xi + \varphi_{\tilde{I}}(\omega)]} e^{-\frac{\xi^2}{2\sigma^2(\omega)}} d\xi$$
(A.4)

onde a segunda integral é identificada como o produto interno entre a função de base exponencial complexa e a função analisadora de Gabor, demonstrando assim a Equação (1.6). A fim de validarmos a Equação (1.10), nós a reescrevemos na forma:

$$I(x) = \frac{1}{(2\pi)^{3/2}} \int_{-\infty}^{\infty} e^{i\omega x} e^{\varphi_{\tilde{I}}(\omega)} d\omega \int_{-\infty}^{\infty} e^{-\frac{(x-\xi)^2}{2\sigma(\omega)^2}} d\xi$$
(A.5)

e como a segunda integral resulta em $\sqrt{(2\pi)}|\tilde{I}(\omega)|$, ao calcularmos a transformada inversa de Fourier de $\tilde{I}(\omega)$, obtemos o próprio sinal I(x).

Segunda abordagem de representação

Para demonstrar a identidade na Equação (1.15), mudamos a variável de integração para x' e tomamos a transformada inversa de Fourier em ambos os lados da equação, obtendo

$$I(x) = \frac{1}{(2\pi)^2} \int_{-\infty}^{\infty} \mathcal{F}^{-1} \left\{ e^{-i\omega x'} * e^{-i[\omega x' - \varphi_I(x')]} e^{\frac{-\omega^2}{2\Sigma^2(x')}} \right\} dx'$$
(A.6)

Novamente usando a propriedade da transformada da convolução, resulta

$$I(x) = \frac{1}{(2\pi)^2} \int_{-\infty}^{\infty} \mathcal{F}^{-1} \left\{ e^{-i\omega x'} \right\} \times \mathcal{F}^{-1} \left\{ e^{-i[\omega x' - \varphi_I(x')]} e^{\frac{-\omega^2}{2\Sigma^2(x')}} \right\} dx'$$
(A.7)

e, como ambas as transformadas são conhecidas,

$$I(x) = \frac{1}{(2\pi)^{3/2}} \int_{-\infty}^{\infty} \delta(x - x') \Sigma(x') e^{i\varphi_I(x')} e^{-\frac{\Sigma^2(x')}{2}(x - x')^2} dx'$$
(A.8)

Finalmente, pela propriedade de filtragem da delta, obtemos $I(x) = \frac{1}{(2\pi)^{3/2}} \Sigma(x) e^{i\varphi_I(x)}$, o que é consistente com as nossas definições para $\Sigma(x)$ e $\varphi_I(x)$.

Explicitando-se a operação de convolução, a Equação 1.15 pode ser reescrita como:

$$\tilde{I}(\omega) = \int_{-\infty}^{\infty} e^{-i\omega x} dx \int_{-\infty}^{\infty} \frac{1}{(2\pi)^2} e^{i\Omega x} e^{-i[\Omega x - \varphi_I(x)]} e^{-\frac{\Omega^2}{2\Sigma^2(x)}} d\Omega$$
(A.9)

onde a segunda integral é identificada como o produto interno entre a função de base exponencial complexa e a função analisadora de Gabor, dessa maneira demonstrando a Equação (1.19).

APÊNDICE B

Neste apêndice demonstraremos as identidades apresentadas no Capítulo 2, para a primeira forma da Transformada de Gabor sintonizada, $T(\omega, x)$.

Primeiramente, demonstraremos a Equação (2.10), obtida observando-se inicialmente que a $T(\omega, x)$ pode ser reescrita da forma:

$$T(\omega, x) = \frac{1}{(2\pi)^{3/2}} \int_{-\infty}^{\infty} I(\xi) g^*(\xi) d\xi$$
(B.1)

para $g(\xi) = e^{i[\omega\xi + \varphi_{\tilde{I}}(\omega)]} e^{-\frac{(\xi-x)^2}{2\sigma^2(\omega)}}$. E, de acordo com o teorema de Parseval [60],

$$T(\omega, x) = \frac{1}{(2\pi)^{5/2}} \int_{-\infty}^{\infty} \tilde{I}(\Omega) G^*(\Omega) d\Omega$$
(B.2)

onde $G(\Omega) \equiv \mathcal{F}\{g(\xi)\}$ pode ser alcançado, conforme observado nos resultados anteriores, como

$$\sqrt{2\pi}\sigma(\omega)e^{i\varphi_{\tilde{I}}(\omega)}e^{-i(\Omega-\omega)x}e^{-\frac{\sigma^2(\omega)}{2}(\Omega-\omega)^2}$$
(B.3)

Logo, obtemos

$$|T(\omega, x)| = \frac{1}{(2\pi)^2} \left| \int_{-\infty}^{\infty} \tilde{I}(\Omega) \sigma(\omega) e^{-\frac{\sigma^2(\omega)}{2}(\Omega-\omega)^2} e^{i(\Omega-\omega)x} d\Omega \right|$$
(B.4)

e, verificando-se que $\sigma(\omega) = |\tilde{I}(\omega)|$ e alterando-se a variável $\Omega \to \Omega + \omega$, obtemos a Equação (2.10).

A seguir, apresentaremos as demonstrações das propriedades da transformada sintonizada.

1. Transformada da delta de Dirac

Como a transformada de Fourier do sinal $I(x) = \delta(x - x_0)$ é $\tilde{I}(\omega) = e^{-i\omega x_0}$, obtemos $\sigma(\omega) = 1$ e $\varphi_{\tilde{I}}(\omega) = -\omega x_0$. A Equação (2.3) pode ser reescrita como

$$T(\omega, x) = \frac{1}{(2\pi)^{3/2}} \int_{-\infty}^{\infty} e^{i[\omega(\xi - x_0)]} e^{\frac{-(\xi - x)^2}{2}} \delta(\xi - x_0) d\xi = \frac{1}{(2\pi)^{3/2}} e^{-\frac{1}{2}(x - x_0)^2}$$
(B.5)

usando-se a propriedade de filtragem da delta de Dirac.

2. Transformada do trem de impulsos

A Equação (2.3) pode ser reescrita na forma

$$T(\omega, x) = \frac{1}{(2\pi)^{3/2}} \sum_{k} \int_{-\infty}^{\infty} e^{-i\omega\xi} e^{-\frac{(\xi-x)^2}{2\sigma^2(\omega)}} \delta(\xi - x_k) d\xi$$
(B.6)

e, novamente usando a propriedada de filtragem da delta,

$$T(\omega, x) = \frac{1}{(2\pi)^{3/2}} \sum_{k} e^{-i\omega x_k} e^{-\frac{(x-x_k)^2}{2\sigma^2(\omega)}}$$
(B.7)

I(x) pode ser aproximado como um sinal periódico, e assim,

$$\sigma^{2}(\omega) \equiv |\tilde{I}(\omega)|^{2} = \left(\frac{2\pi}{L}\right)^{2} |\sum_{l} \delta(\omega - \omega_{l})|^{2}$$
(B.8)

onde $\omega_l = l\omega_0 = 2\pi l/L$.

A partir da equação acima, nós verificamos que $\sigma^2(\omega) = 0$, sempre que ω não for um múltiplo inteiro de ω_0 . Neste caso, todas a Gaussianas na Equação (B.7), e também $T(\omega, x)$, se anularão, exceto se $x = x_m$, para qualquer inteiro m, quando obtemos:

$$T(\omega, x_m) = \frac{1}{(2\pi)^{3/2}} e^{-i\omega x_m}, \omega \neq \omega_n$$
(B.9)

Por outro lado, sempre que ω for um múltiplo inteiro de $\omega_0, \sigma^2(\omega) \to \infty$. Neste caso, as Gaussianas na Equação (B.7) irão todas tender a um, e iremos obter

$$T(\omega_n, x) = \frac{1}{(2\pi)^{3/2}} \sum_k e^{-i\omega_n x_k} = \frac{1}{(2\pi)^{3/2}} \sum_k 1, \quad \forall x$$
(B.10)

assim finalizando a prova da Equação (2.14).

3. Transformada da exponencial complexa

A transformada de Fourier de $I(x) = e^{i\omega_0 x}$ é $\tilde{I}(\omega) \equiv \sigma(\omega) = 2\pi\delta(\omega - \omega_0)$, com $\varphi_{\tilde{I}}(\omega) = 0$. A Equação (2.3) então se torna:

$$T(\omega, x) = \frac{1}{(2\pi)^{3/2}} \int_{-\infty}^{\infty} e^{-i[(\omega - \omega_0)\xi]} e^{-\frac{(\xi - x)^2}{2\sigma^2(\omega)}} d\xi = \frac{1}{(2\pi)^{3/2}} \mathcal{F}\left[e^{-\frac{(\xi - x)^2}{2\sigma^2(\omega)}}\right] (\omega - \omega_0)$$
(B.11)

e, ao computar a transformada de Fourier indicada, obtemos

$$T(\omega, x) = \frac{1}{(2\pi)^{3/2}} \sqrt{2\pi} \sigma(\omega) e^{-\frac{\sigma^2(\omega)}{2}(\omega-\omega_0)^2} e^{-i(\omega-\omega_0)x} = \delta(\omega-\omega_0)$$
(B.12)

4. Densidade de Energia, Correlação Espacial, Momentos da Frequência

Substituindo $T(\omega, x)$ no lado direito da Equação (2.4), obtemos

$$\frac{1}{\sqrt{(2\pi)}} \int_{-\infty}^{\infty} e^{-i[\omega\xi + \varphi_{\bar{I}}(\omega)]} I(\xi) d\xi \int_{-\infty}^{\infty} e^{\frac{-(\xi - x)^2}{2\sigma^2(\omega)}} dx$$
(B.13)

onde a segunda integral resulta em $\sqrt{2\pi}|\tilde{I}(\omega)|$. A fórmula acima pode então ser reescrita como:

$$|\tilde{I}(\omega)|e^{-i\varphi_{\tilde{I}}(\omega)}\int_{-\infty}^{\infty}e^{-i\omega\xi}I(\xi)d\xi$$
(B.14)

e, identificando-se a integral que restou como a transformada de Fourier de $I(\xi)$, a equação fica demonstrada. As Equações (2.5), (2.7) e (2.20) são imediatamente verificadas.

5. Translação Espacial

A transformada de Fourier de $I_2(x) \equiv I(x - x_0)$ é $\tilde{I}_2(\omega) = e^{-i\omega x_0} \tilde{I}(\omega)$. Logo, $\sigma_2(\omega) = \sigma(\omega)$ e $\varphi_{\tilde{I}_2}(\omega) = \varphi_{\tilde{I}}(\omega) - \omega x_0$.

A Equação (2.3) pode ser então reescrita como

$$T_{I_2}(\omega, x) = \frac{1}{(2\pi)^{3/2}} \int_{-\infty}^{\infty} e^{-i[\omega(\xi - x_0) + \varphi_{\bar{I}}(\omega)]} e^{-\frac{(\xi - x)^2}{2\sigma^2(\omega)}} I(\xi - x_0) d\xi$$
(B.15)

e, fazendo-se a mudança de variável $\xi' = \xi - x_0$,

$$T_{I_2}(\omega, x) = \frac{1}{(2\pi)^{3/2}} \int_{-\infty}^{\infty} e^{-i[\omega\xi' + \varphi_{\bar{I}}(\omega)]} e^{-\frac{(\xi' + x - x_0)^2}{2\sigma^2(\omega)}} I(\xi') d\xi' = T(\omega, x - x_0)$$
(B.16)

6. Translação na Frequência

A transformada de Fourier de $I_2(x) \equiv e^{i\omega_0 x} I(x)$ é $\tilde{I}_2(\omega) = \tilde{I}(\omega - \omega_0)$. Logo, $\sigma_2(\omega) = \sigma(\omega - \omega_0)$ e $\varphi_{\tilde{I}_2}(\omega) = \varphi_{\tilde{I}}(\omega - \omega_0)$. A Equação (2.3) então se torna:

$$T_{I_2}(\omega, x) = \frac{1}{(2\pi)^{3/2}} \int_{-\infty}^{\infty} e^{-i[(\omega-\omega_0)\xi + \varphi_{\tilde{I}}(\omega-\omega_0)]} e^{-\frac{(\xi-x)^2}{2\sigma^2(\omega-\omega_0)}} I(\xi) d\xi = T(\omega-\omega_0, x)$$
(B.17)

7. Primeiro momento condicional do espaço

Usando a definição de $T(\omega, x)$ na Equação (2.23), obtemos:

$$< x >_{\omega} = \frac{1}{(2\pi)^{3/2} |\tilde{I}(\omega)|^2} \int_{-\infty}^{\infty} e^{-i[\omega\xi + \varphi_{\bar{I}}(\omega)]} I(\xi) d\xi \int_{-\infty}^{\infty} x e^{\frac{-(\xi-x)^2}{2\sigma^2(\omega)}} dx$$
(B.18)

de onde obtemos

$$\langle x \rangle_{\omega} = \frac{1}{2\pi \tilde{I}(\omega)} \int_{-\infty}^{\infty} \xi e^{-i\omega\xi} I(\xi) d\xi$$
 (B.19)

com o uso da identidade

$$\int_{-\infty}^{\infty} x e^{\frac{-(\xi-x)^2}{2\sigma^2(\omega)}} dx = \xi \int_{-\infty}^{\infty} e^{\frac{-(\xi-x)^2}{2\sigma^2(\omega)}} dx = \xi \sqrt{2\pi} |\tilde{I}(\omega)|$$
(B.20)

Identificando a integral no lado direito da Equação (B.19) como

$$\int_{-\infty}^{\infty} \xi e^{-i\omega\xi} I(\xi) d\xi = i \frac{d}{d\omega} \int_{-\infty}^{\infty} e^{-i\omega\xi} I(\xi) d\xi = i \frac{d}{d\omega} \tilde{I}(\omega)$$
(B.21)

finalmente obtemos

$$\langle x \rangle_{\omega} = \frac{i}{2\pi} \frac{d}{d\omega} \log \tilde{I}(\omega)$$
 (B.22)

e a Equação (2.24) fica demonstrada.

8. Momento cruzado do espaço e da frequência

Usando a definição de $T(\omega, x)$ na Equação (2.26), obtemos:

$$\langle x\omega \rangle = \frac{1}{2\pi} \int_{-\infty}^{\infty} \xi I(\xi) d\xi \int_{-\infty}^{\infty} \omega e^{-i\omega\xi} \tilde{I}^*(\omega) d\omega$$
 (B.23)

onde a Equação (B.20) foi utilizada. A segunda integral é semelhante à da Equação (B.21), e pode ser facilmente identificada como

$$i\frac{d}{d\xi}\left\{\int_{-\infty}^{\infty}e^{i\omega\xi}\tilde{I}(\omega)d\omega\right\}^{*} = i2\pi\frac{d}{d\xi}I^{*}(\xi)$$
(B.24)

Portanto,

$$\langle x\omega \rangle = i \int_{-\infty}^{\infty} \xi I(\xi) \left[e^{-i\varphi_I(\xi)} \frac{d}{d\xi} |I(\xi)| - i\varphi_I'(x)I^*(\xi) \right] d\xi$$
(B.25)

e, através de uma manipulação direta, demonstramos a Equação (2.27).

9. Inversa

A transformada de $T(\omega_0, x)$, em uma frequência ω , é obtida como

$$\mathcal{F}\left[T(\omega_0, x)\right](\omega) = \frac{1}{(2\pi)^{3/2}} \int_{-\infty}^{\infty} e^{-i[\omega_0\xi + \varphi_{\tilde{I}}(\omega_0)]} I(\xi) \mathcal{F}\left[e^{\frac{-(\xi - x)^2}{2\sigma^2(\omega_0)}}\right](\omega) d\xi$$
(B.26)

e, usando

$$\mathcal{F}\left[e^{\frac{-(\xi-x)^2}{2\sigma^2(\omega_0)}}\right](\omega) = \sqrt{2\pi} |\tilde{I}(\omega_0)| e^{-i\omega\xi} e^{-\frac{|\tilde{I}(\omega_0)^2|}{2}\omega^2}$$
(B.27)

obtemos

$$\mathcal{F}\left[T(\omega_0, x)\right](\omega) = \frac{1}{2\pi} \tilde{I}^*(\omega_0) e^{\frac{-|\tilde{I}(\omega_0)^2|}{2}\omega^2} \int_{-\infty}^{\infty} e^{-i(\omega+\omega_0)\xi} I(\xi) d\xi \tag{B.28}$$

onde a integral é identificada como $\tilde{I}(\omega+\omega_0).$

Considerando a equação acima na frequência $\omega' = \omega - \omega_0$, obtemos

$$\tilde{I}(\omega) = 2\pi \frac{\mathcal{F}[T(\omega_0, x)](\omega - \omega_0)}{\tilde{I}^*(\omega_0)e^{-\frac{|\tilde{I}(\omega_0)|^2}{2}(\omega - \omega_0)^2}}$$
(B.29)

que é a mesma relação da Equação (2.16).

APÊNDICE C

Neste apêndice demonstraremos as identidades apresentadas no Capítulo 2, para a segunda forma da Transformada de Gabor sintonizada, $T^{(2)}(\omega, x)$.

1. Transformada da delta de Dirac

Como a transformada de Fourier do sinal $I(x) = \delta(x - x_0)$ é $\tilde{I}(\omega) = e^{-i\omega x_0}$, podemos reescrever a Equação (2.34) como:

$$T^{(2)}(\omega, x) = \frac{1}{(2\pi)^2} e^{-i\varphi_I(x)} \int_{-\infty}^{\infty} e^{i\Omega(x-x_0)} e^{\frac{-(\Omega-\omega)^2}{2\Sigma^2(x)}} d\Omega$$
(C.1)

Considerando que $\Sigma(x) = \delta(x - x_0)(2\pi)^{3/2}$ e $I^*(x) = |I(x)|e^{-i\varphi_I(x)}$, e usando a propriedade de amostragem da delta, podemos verificar que:

$$T^{(2)}(\omega, x) = \frac{1}{(2\pi)^2} e^{-i\varphi_I(x)} \sqrt{2\pi} \Sigma(x) e^{-i(x-x_0)\omega} e^{\frac{-\Sigma^2(x)(x-x_0)^2}{2}} = \delta(x-x_0)$$
(C.2)

2. Transformada do trem de impulsos

A Equação (2.34) pode ser reescrita na forma:

$$T^{(2)}(\omega,x) = \frac{1}{(2\pi)^2} \sum_{k} \int_{-\infty}^{\infty} e^{i\Omega x} e^{-\frac{(\Omega-\omega)^2}{2\Sigma^2(x)}} \frac{2\pi}{L} \delta(\Omega-\omega_k) d\Omega$$
(C.3)

e, usando a propriedade de filtragem da delta,

$$T^{(2)}(\omega, x) = \frac{1}{(2\pi)L} \sum_{k} e^{i\omega_k x} e^{-\frac{(\omega-\omega_k)^2}{2\Sigma^2(x)}}$$
(C.4)

I(x) pode ser aproximado como um sinal periódico, e então

$$\Sigma^{2}(x) = |I(x)|^{2} (2\pi)^{3} = |\sum_{l} \delta(x - x_{l})|^{2} (2\pi)^{3}$$
(C.5)

A partir da equação acima, verificamos que $\Sigma^2(x) = 0$, sempre que x for diferente de x_m , para qualquer inteiro m. Se, além dessa condição, ω não for um inteiro múltiplo de

 ω_0 , todas as Gaussianas na Equação (C.4), e também $T^2(\omega, x)$, se anularão, exceto se ω for um inteiro múltiplo de ω_0 , quando obtemos

$$T^{2}(\omega_{n}, x) = \frac{1}{(2\pi)L} e^{i\omega_{n}x}, \quad \omega = \omega_{n}, x \neq x_{m}$$
(C.6)

Por outro lado, sempre que $x = x_m$, para qualquer inteiro $m, \Sigma^2(x) \to \infty$. Neste caso, as Gaussianas na Equação (C.4) irão todas tender a um, e iremos obter:

$$T^{2}(\omega, x_{m}) = \frac{1}{(2\pi)L} \sum_{k} e^{i\omega x_{k}} = \frac{1}{(2\pi)L} \sum_{k} 1, \quad \forall \omega$$
 (C.7)

assim finalizando a prova da Equação (2.40).

3. Transformada da exponencial complexa

Como a transformada de Fourier de $I(x) = e^{i\omega_0 x}$ é $\tilde{I}(\omega) = 2\pi\delta(\omega - \omega_0)$, e $\varphi_I(x) = \omega_0 x$, a Equação (2.34) pode ser reescrita como:

$$T^{(2)}(\omega, x) = \frac{1}{(2\pi)} \int_{-\infty}^{\infty} e^{i[x(\Omega - \omega_0)]} e^{\frac{-(\Omega - \omega)^2}{2\Sigma^2(x)}} \delta(\Omega - \omega_0) d\Omega$$
(C.8)

e, usando a propriedade de filtragem da delta, considerando também que |I(x)| = 1, obtemos

$$T^{(2)}(\omega, x) = \frac{1}{(2\pi)} e^{-\frac{(\omega - \omega_0)^2}{16\pi^3}}$$
(C.9)

4. Densidade de Energia, Correlação Espacial, Momentos da Frequência

Substituindo $T^{(2)}(\omega, x)$ no lado direito da Equação (2.42), obtemos:

$$\frac{1}{(2\pi)^3} \int_{-\infty}^{\infty} e^{\frac{-(\Omega-\omega)^2}{2\Sigma^2(x)}} d\omega \int_{-\infty}^{\infty} e^{i[\Omega x - \varphi_I(x)]} \tilde{I}(\Omega) d\Omega$$
(C.10)

onde a segunda integral resulta em $2\pi |I(x)|$. A fórmula acima pode então ser reescrita como:

$$\frac{1}{(2\pi)^2}|I(x)|\int_{-\infty}^{\infty}e^{\frac{-(\Omega-\omega)^2}{2\Sigma^2(x)}}d\omega$$
(C.11)

e, identificando-se a integral que restou como $\sqrt{2\pi}\Sigma(x)$, a equação fica demonstrada. As Equações (2.43), (2.45) e (2.46) são imediatamente verificadas.

5. Translação Espacial

A transformada de Fourier de $I_2(x) \equiv I(x-x_0)$ é $\tilde{I}_2(\omega) = e^{-i\omega x_0} \tilde{I}(\omega)$. Como $\Sigma_{I_2}(x) =$

 $\Sigma(x-x_0)$ e $\varphi_{I_2}(x) = \varphi_I(x-x_0)$, a Equação (2.34) pode ser então reescrita como

$$T_{I_2}^{(2)}(\omega, x) = \frac{1}{(2\pi)^2} \int_{-\infty}^{\infty} e^{i[\Omega(x-x_0)-\varphi_I(x-x_0)]} e^{-\frac{(\Omega-\omega)^2}{2\Sigma^2(x-x_0)}} \tilde{I}(\Omega) d\Omega = T^{(2)}(\omega, x-x_0) \quad (C.12)$$

6. Translação na Frequência

A transformada de Fourier de $I_2(x) \equiv e^{i\omega_0 x} I(x)$ é $\tilde{I}_2(\omega) = \tilde{I}(\omega - \omega_0)$. Como $\Sigma_{I_2}(x) = \Sigma(x)$ e $\varphi_{I_2}(x) = \varphi_I(x) + \omega_0 x$, a Equação (2.34) pode ser reescrita como:

$$T_{I_2}^{(2)}(\omega, x) = \frac{1}{(2\pi)^2} \int_{-\infty}^{\infty} e^{i[(\Omega - \omega_0)x - \varphi_I(x)]} e^{\frac{-(\Omega - \omega)^2}{2\Sigma^2(x)}} \tilde{I}(\Omega - \omega_0) d\Omega$$
(C.13)

e, fazendo-se a mudança de variável $\Omega'=\Omega-\omega_0$

$$T_{I_2}^{(2)}(\omega, x) = \frac{1}{(2\pi)^2} \int_{-\infty}^{\infty} e^{i[\Omega' x - \varphi_I(x)]} e^{\frac{-(\Omega' + \omega - \omega_0)^2}{2\Sigma^2(x)}} \tilde{I}(\Omega') d\Omega' = T^{(2)}(\omega - \omega_0, x)$$
(C.14)

7. Primeiro momento condicional da frequência

Empregando a definição de $T^{(2)}(\omega, x)$ na Equação (2.49), obtemos:

$$<\omega>_{x}=\frac{1}{(2\pi)^{2}|I(x)|^{2}}\int_{-\infty}^{\infty}e^{i[\Omega x-\varphi_{I}(x)]}\tilde{I}(\Omega)d\Omega\int_{-\infty}^{\infty}\omega e^{\frac{-(\omega-\Omega)^{2}}{2\Sigma^{2}(x)}}d\omega$$
(C.15)

e, usando a identidade

$$\int_{-\infty}^{\infty} \omega e^{-\frac{(\omega-\Omega)^2}{2\Sigma^2(x)}} d\omega = \Omega \int_{-\infty}^{\infty} e^{-\frac{(\omega-\Omega)^2}{2\Sigma^2(x)}} d\omega = \Omega \sqrt{2\pi} \Sigma(x) = (2\pi)^2 \Omega |I(x)|$$
(C.16)

obtemos

$$\langle \omega \rangle_x = \frac{1}{|I(x)|^2} \int_{-\infty}^{\infty} e^{i\Omega x} \tilde{I}(\Omega) \Omega |I(x)| e^{-i\varphi_I(x)} d\Omega = \frac{I^*(x)}{|I(x)|^2} \int_{-\infty}^{\infty} e^{i\Omega x} \Omega \tilde{I}(\Omega) d\Omega \quad (C.17)$$

Identificando a integral no lado direito da equação acima como

$$\int_{-\infty}^{\infty} e^{i\Omega x} \Omega \tilde{I}(\Omega) d\Omega = -i \frac{d}{dx} \int_{-\infty}^{\infty} e^{i\Omega x} \tilde{I}(\Omega) d\Omega = -i2\pi \frac{d}{dx} I(x)$$
(C.18)

finalmente obtemos

$$\langle \omega \rangle_x = \frac{-i(2\pi)}{I(x)} \frac{d}{dx} I(x) = -(2\pi)i \frac{d}{dx} log I(x) = -(2\pi)i \frac{d}{dx} \{ log |I(x)| + i\varphi(x) \}$$
 (C.19)

e a Equação (2.50) fica demonstrada.

8. Momento cruzado do espaço e da frequência

Aplicando a definição de $T^{(2)}(\omega, x)$ na Equação (2.52), e utilizando a Equação (C.16),

obtemos:

$$\langle x\omega \rangle = \int_{-\infty}^{\infty} \Omega \tilde{I}(\Omega) d\Omega \int_{-\infty}^{\infty} x I^*(x) e^{i\Omega x} dx$$
 (C.20)

A segunda integral é semelhante à da Equação (C.18), e pode ser facilmente identificada como:

$$\left\{i\frac{d}{d\Omega}\int_{-\infty}^{\infty}e^{-i\Omega x}I(x)\right\}^{*}dx = \left\{i\frac{d}{d\Omega}\tilde{I}(\Omega)\right\}^{*}$$
(C.21)

Logo,

$$\langle x\omega \rangle = -i \int_{-\infty}^{\infty} \Omega \tilde{I}(\Omega) \left[e^{-i\varphi_I(\Omega)} \left(\frac{d}{d\Omega} |\tilde{I}(\Omega)| \right) - i\varphi_I'(\Omega) \tilde{I}^*(\Omega) \right] d\Omega$$
(C.22)

e, através de uma manipulação direta, demonstramos a Equação (2.53).

9. Inversa

A transformada de Fourier inversa de $T^{(2)}(\omega, x_0)$, em uma posição x, é obtida como

$$\mathcal{F}^{-1}[e^{-i[\omega x_0 - \varphi_I(x_0)]}T^{(2)}(\omega, x_0)](x) = \frac{1}{(2\pi)^3} \int_{-\infty}^{\infty} e^{i\Omega x_0} \tilde{I}(\Omega) d\Omega \int_{-\infty}^{\infty} e^{i\omega x} e^{-i\omega x_0} e^{\frac{-(\omega - \Omega)^2}{2\Sigma^2(x_0)}} d\omega$$
(C.23)

Resolvendo a segunda integral, obtemos

$$\mathcal{F}^{-1}[e^{-i[\omega x_0 - \varphi_I(x_0)]}T^{(2)}(\omega, x_0)](x) = \frac{1}{(2\pi)^3} \int_{-\infty}^{\infty} \tilde{I}(\Omega)e^{i\Omega x} |I(x_0|(2\pi)^2 e^{-4\pi^3|I(x_0)|^2(x-x_0)^2} d\Omega$$
(C.24)

e, finalmente,

$$\mathcal{F}^{-1}[e^{-i[\omega x_0 - \varphi_I(x_0)]}T^{(2)}(\omega, x_0)](x) = I(x)|I(x_0)|e^{-4\pi^3|I(x_0)|^2(x-x_0)^2}$$
(C.25)

que é a mesma relação da Equação (2.35).

APÊNDICE D

Neste apêndice demonstraremos a Equação (4.7) do Capítulo 4, Seção 4.3: Modelagem das Células Centro-Periferia por Funções Sintonizadas.

Tomando a transformada de Fourier de ambos os lados da Equação (4.4) (com as variáveis de integração devidamente mudadas para $\omega'_x \in \omega'_y$), fazendo uso da Equação (4.5) e aplicando a propriedade de filtragem da delta, obtemos

$$\tilde{B}(\omega_x, \omega_y)\tilde{I}(\omega_x, \omega_y) = 2\pi \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \delta(\omega'_x - \omega_x, \omega'_y - \omega_y)\tilde{\psi}(\rho; \omega'_x, \omega'_y) d\omega'_x d\omega'_y = 2\pi \tilde{\psi}(\rho; \omega_x, \omega_y)$$
(D.1)

onde $\rho = \sqrt{\omega_x^2 + \omega_y^2}$, e com \tilde{B} e $\tilde{\psi}$ denotando, respectivamente, as transformadas de Fourier do filtro de branqueamento e da função de representação. Esta última pode ser obtida, a partir da Equação (4.6), como

$$\tilde{\psi}(\rho;\omega_x,\omega_y) = -2e^{i\varphi_{\tilde{I}}(\omega_x,\omega_y)} \left\{ \frac{1}{\rho} - \frac{1 - \operatorname{cyl}(\rho/2\pi\sigma)}{\sqrt{\rho^2 - \pi^2\sigma^2}} - \frac{\operatorname{cyl}(\rho/2\pi\sigma)}{\sqrt{\pi^2\sigma^2 - \rho^2}} \right\}$$
(D.2)

onde $\operatorname{cyl}(\rho/2\pi\sigma) = 1$, caso $0 \le \rho \le \pi\sigma$, e zero, caso contrário.

Usando o resultado acima na Equação (D.1), podemos resolvê-la para $\sigma(\omega_x, \omega_y)$, obtendo duas soluções distintas: uma se assumimos $\sigma < \rho/\pi$, e a outra para $\sigma \ge \rho/\pi$. Como já mencionado no Capítulo 3, a dimensão radial da periferia do campo receptivo modelado por $\psi(r; \omega_x, \omega_y)$ é dada por $3/2\sigma$. Considerando que as células da retina e do núcleo geniculado lateral apresentam, experimentalmente, periferias largas, escolhemos a solução para $\sigma < \rho/\pi$ (verificamos que o uso da outra solução não acarreta alterações significativas em nossos resultados). Neste caso, as Equações (D.1) e (D.2) conduzem a

$$\tilde{B}(\omega_x, \omega_y) |\tilde{I}(\omega_x, \omega_y)| = -4\pi \left[1/\rho - \frac{1}{\rho^2 - \pi^2 \sigma^2} \right]$$
(D.3)

onde identificamos $\tilde{I}(\omega_x, \omega_y) = |\tilde{I}(\omega_x, \omega_y)| e^{i\varphi_{\tilde{I}}(\omega_x, \omega_y)}$. A equação acima pode ser resolvida para σ , resultando na Equação (4.7).

É importante salientar que o modelo mais utilizado para a descrição dos campos receptivos centro-periferia — a diferença de Gaussianas [61] — não foi utilizado neste trabalho, já que esse modelo requer dois parâmetros (a largura das duas Gaussianas) para a definição das funções de codificação, enquanto que a nossa abordagem provê uma única equação para esse fim.

APÊNDICE E

Aqui sumariamos o processo de estimação de disparidades via funções de Green de equações de casamento, conforme introduzido em [11].

Em estereoscopia, assumindo-se que a geometria epipolar é conhecida, uma equação de casamento assume a forma unidimensional

$$I_l(x + U(x), y) = I_r(x, y)$$
 (E.1)

onde U(x) denota o campo de disparidades, e onde I_l e I_r representam as imagens de entrada esquerda e direita. Tipicamente, uma equação como a (E.1), é empregada para a estimação de U(x), dado o par estereoscópico, mas em [11] ela foi tomada como uma restrição sobre as imagens a serem casadas, de forma que, dada I_r , o seu par perfeito de casamento deve ser encontrado, assumindo-se formas simples para o campo de disparidades.

Por exemplo, assumindo-se disparidades uniformes, U(x) = u, onde u é uma constante, podemos tomar a expansão de Taylor em segunda ordem da Equação (E.1), para obter

$$\frac{u^2}{2}I_l''(x) + uI_l'(x) + I_l(x) = I_r(x)$$
(E.2)

onde as *plicas* denotam diferenciação com relação a x (por simplicidade, passamos a omitir as dependências em y). A Equação (E.2) pode então ser resolvida pela abordagem da função de Green [62], fornecendo

$$I_l(x) \approx I_r(x-u) = \int_{-\infty}^{\infty} G_k(x, x_0) I_r(x_0) dx_0$$
 (E.3)

onde

$$G_k(x, x_0) \equiv G_k(x - x_0) = \begin{cases} 2k \sin [k(x - x_0)]e^{-k(x - x_0)}, & \text{se } x > x_0\\ 0, & \text{se } x < x_0 \end{cases}$$
(E.4)

com k = 1/u, é a função de Green da Eq. (E.2), que corresponde à solução daquela

equação quando o seu lado direito é a função impulso, $\delta(x - x_0)$ (estamos assumindo o domínio ilimitado $x \in [-\infty, \infty]$).

Pela Equação (E.3) verificamos que, até segunda ordem em u, a convolução da imagem direita, $I_r(x)$, com a função de Green $G_k(x)$ provoca um deslocamento naquela imagem, para a direita, de magnitude u. Similarmente, a convolução com $G_k(-x)$ (mantendo ucomo um parâmetro positivo) provocaria um deslocamento similar para a esquerda. De forma mais genérica, podemos considerar um núcleo de Green complexo, introduzindo o par de quadratura de $G_k(x - x_0)$, dado por

$$H_k(x - x_0) = \begin{cases} 2k \cos \left[k(x - x_0)\right] e^{-k(x - x_0)}, & \text{se } x > x_0 \\ 0, & \text{se } x < x_0 \end{cases}$$
(E.5)

que fornece uma solução homogênea para a Equação (E.2) — isto é, uma solução daquela equação quando o seu lado direito é identicamente zero.

Assim, sendo

$$K_k(x - x_0) = G_k(x - x_0) + iH_k(x - x_0)$$
(E.6)

o núcleo de Green complexo, obtemos, como a versão de $I_r(x)$ deslocada para a direita,

$$I_r(x-u) = \int_{-\infty}^{\infty} K_k(x-x_0) I_r(x_0) dx_0$$
 (E.7)

uma expressão que é correta até segunda ordem em u.

A abordagem de estimação de disparidades de [11] se baseia num núcleo de Green similar ao da Equação (E.6), mas para uma equação diferencial que aproxima uma restrição de casamento linear, ou seja, uma para a qual o campo de disparidades assume a forma U(x) = u + vx, onde $u \in v$ são constantes. Mais especificamente, um deslocamento para a direita é neste caso obtido com o núcleo complexo

$$K(x, x_0) = \begin{cases} 2ke^{[ik(x-x_0)]}e^{-\left[\frac{(x+a)^2 - (x_0+a)^2}{2\sigma^2}\right]}, & \text{se } x > x_0\\ 0, & \text{se } x < x_0 \end{cases}$$
(E.8)

onde σ e *a* são constantes positivas, e $k = a/\sigma^2$. O núcleo $K(x, x_0)$ é a função de Green complexa da equação

$$\frac{1}{2}\left(\frac{\sigma^2}{a}\right)^2 I_l''(x) + \left(\frac{\sigma^2}{a} + \frac{\sigma^2}{a^2}x\right) I_l'(x) + \left(1 + \frac{x^2 + 2ax + \sigma^2}{2a^2}\right) I_l(x) = I_r(x)$$
(E.9)

cuja solução homogênea é a função de Gabor $e^{ikx}e^{-\frac{(x+a)^2}{2\sigma^2}}$.

Quando $|x| \in \sigma$ são ambos muito menores que a, a Equação (E.9) pode ser aproximada como

$$\frac{1}{2} \left(\frac{\sigma^2}{a}\right)^2 I_l''(x) + \left(\frac{\sigma^2}{a} + \frac{\sigma^2}{a^2}x\right) I_l'(x) + I_l(x) = I_r(x)$$
(E.10)

que, até primeira ordem em σ^2/a , corresponde a uma equação de casamento para um campo de deslocamentos linear

$$U(x) = \frac{\sigma^2}{a} + \frac{\sigma^2}{a^2}x \tag{E.11}$$

Similarmente ao caso uniforme, um deslocamento para a esquerda seria efetuado, neste modelo linear de disparidades, pelo núcleo $K^{(-)}(x, x_0) = K(-x, -x_0)$.

Para a estimação das disparidades, inicialmente as imagens de entrada, I_l and I_r , são multiplicadas por funções de Gabor, e o resultado é então submetido à filtragem, respectivamente, pelos núcleos $K(x, x_0)$ e $K^{(-)}(x, x_0)$. Estes têm o efeito de produzir deslocamentos espaciais e mudanças de fase nas entradas moduladas por Gabor (ver abaixo). O deslocamento espacial ótimo, em cada posição, pode assim ser obtido avaliando-se a correspondência dos sinais gerados por funções de Green com diferentes parâmetros. Isto fornecerá uma estimativa do mapa de disparidades codificado pelo par estereoscópico.

A relação da abordagem de Green com os modelos neurofisiológicos de estereoscopia advém da seguinte propriedade dos núcleos de Green: ao filtrar um sinal modulado por Gabor, eles produzem saídas similarmente moduladas, mas com versões espacialmente deslocadas do sinal. Por exemplo, vamos considerar o resultado da filtragem do sinal complexo

$$I_1(x) = e^{i\kappa x} e^{-\frac{x^2}{2\sigma^2}} I_r(x)$$
 (E.12)

pelo núcleo $K(x, x_0)$. Mostra-se que este é [11]

$$I_2(x) = e^{i(\kappa x + \psi)} e^{-\frac{x^2}{2\sigma^2}} I_r(x - u)$$
(E.13)

onde $I_r(x-u)$ é dado pela Equação (E.7), com $u = \sigma^2/a$, e onde $\psi = \kappa u$. Assim, filtrar o sinal $I_1(x)$ pelo núcleo $K(x, x_0)$ essencialmente preserva o fator modulador de Gabor (introduzindo uma fase), mas desloca espacialmente a imagem modulada.

Assumindo que, localmente — isto é, sob a janela Gaussiana da função de Gabor —, a disparidade entre as imagens direita e esquerda seja bem aproximada por u, de modo que $I_l(x) \approx I_r(x-u)$, poderíamos reescrever a Equação (E.13) como

$$I_2(x) = e^{i(\kappa x + \psi)} e^{-\frac{x^2}{2\sigma^2}} I_l(x)$$
 (E.14)

Juntamente com a Equação (E.12), isto significaria que $I_1(x)$ e $I_2(x)$ correspondem, respectivamente, a imagens estereoscópicas direita e esquerda, vistas sob os campos receptivos de um par em quadratura de células corticais simples, de acordo com o chamado modelo de deslocamento de fase para as respostas estereoscópicas [63].

Assim, operando sobre o equivalente à entrada cortical do *olho direito (viz.*, sobre a imagem *retinal* do olho direito como vista sob o campo receptivo de uma célula simples associada àquele olho), o núcleo de Green $K(x, x_0)$ produz o equivalente à entrada cortical do *olho esquerdo (viz.*, a imagem *retinal* do olho esquerdo como vista sob o campo receptivo de uma célula simples associada àquele olho).

Se trocamos direito por esquerdo e vice-versa, no desenvolvimento acima, ele se torna igualmente válido para o núcleo de Green $K^{(-)}(x, x_0)$. Especificamente, sendo

$$I_1^{(-)}(x) = e^{i\kappa x} e^{-\frac{x^2}{2\sigma^2}} I_l(x)$$
(E.15)

a imagem modulada por Gabor a ser filtrada por $K^{(-)}(x, x_0)$, obtemos

$$I_2^{(-)}(x) = e^{i(\kappa x - \psi)} e^{-\frac{x^2}{2\sigma^2}} I_l(x + u)$$
(E.16)

onde, novamente, $u = \sigma^2/a$. Do mesmo modo, a Equação (E.16) pode também ser reescrita, similarmente à (E.14), como

$$I_2^{(-)}(x) = e^{i(\kappa x - \psi)} e^{-\frac{x^2}{2\sigma^2}} I_r(x)$$
(E.17)

Assim, operando sobre o equivalente à entrada cortical do *olho esquerdo* (*viz.*, sobre a imagem *retinal* do olho esquerdo como vista sob o campo receptivo de uma célula simples associada àquele olho), o núcleo de Green $K^{(-)}(x, x_0)$ produz o equivalente à entrada cortical do *olho direito* (*viz.*, a imagem *retinal* do olho direito como vista sob o campo receptivo de uma célula simples associada àquele olho).

Portanto, se agora computamos a correspondência local entre os sinais $I_2(x) \in I_2^{(-)}(x)$ - por exemplo, através da medida

$$|I_2(x) - I_2^{(-)}(x)|^2$$
(E.18)

- estaremos efetivamente avaliando o grau de correspondência entre duas imagens corticais, esquerda e direita, sob a hipótese de que a disparidade local entre elas seja 2u(recordando as Equações (E.13) and (E.16)). Diferentes valores de u podem ser obtidos variando-se o parâmetro a dos núcleos de Green, enquanto $\sigma \in \kappa$ se mantêm fixos, como proposto em [11]. Isto permite a estimação da disparidade local, por exemplo, como

$$d(x) = \frac{2\sigma^2}{\overline{a}(x)} \tag{E.19}$$

onde \overline{a} é o valor de a para o qual a medida de correspondência (Equação E.18) é minimizada.

Referências

- [1] J.L. Fernandes, Shape from shading: novas abordagens a partir de movimento e de foco, Tese de Doutorado, Instituto de Computação, UFF. Disponível em http://www.ic.uff.br/PosGraduacao. 2006.
- [2] E.P. Wigner, On the quantum correction for thermodynamic equilibrium, Phys. Rev. 40, 749-759, 1932.
- [3] L. Cohen, Time-Frequency Analysis, Prentice Hall PTR, Upper Saddle River, 1995.
- [4] S. Marcelja, Mathematical description of the responses of simple cortical cells, J. Opt. Soc. Am. A 70 1297-1300, 1980.
- [5] J. Allman, F. Miezin, and E. McGuinness, Stimulus specific responses from beyond the classical receptive field, Annu. Rev. Neuroscience 8 407-430, 1985.
- [6] W. Bair, Visual receptive field organization, Current Opinion in Neurobiology 15(4) 459-464, 2005.
- [7] S.V. David, W.E. Vinje, and J.L. Gallant, Natural stimulus statistics alter the receptive field structure of V1 neurons, The Journal of Neuroscience 24(31) 6991-7006, 2004.
- [8] D. H. Hubel, Eye, brain and vision. Scientific American Library series, 22, 1988. 234(1), 135, (1995).
- [9] J.J. Atick and A.N. Redlich, What does the retina know about natural scenes, Neural Comp. 4, 196-210, 1992.
- [10] Y. Dan, J.J. Atick and R.C. Reid, Efficient coding of natural scenes in the lateral geniculate nucleus: experimental test of a computational theory, J. Neurosci. 16(10), 3351-3362, 1996.
- [11] J.R. Torreão, Disparity estimation through Green's functions of matching equations, Biol. Cybernetics 97, 307-316, 2007.
- [12] P. Dayan and L.F. Abbott, Theoretical Neuroscience, MIT Press, 2001.
- [13] R.G. Andrzejak, K. Lehnertz, F. Mormann, C. Rieke, P. David, and C.E. Elger, Indications of nonlinear deterministic and finite-dimensional structures in time series of brain electrical activity: Dependence on recording region and brain state. Physical Review E, Volume 64, 061907, 2001. Disponível em http://www.meb.unibonn.de/epileptologie/science/physik/eegdata.html.

- [14] M. Porat and Y.Y. Zeevi, The generalized Gabor scheme of image representation in biological and machine vision, IEEE Trans. PAMI 10(4), 452-468, 1988.
- [15] R.L. Allen and D.W. Mills, Signal Analysis, Wiley-Interscience, 2004.
- [16] K. Gröchenig, Foundations of Time-Frequency Analysis, Birkhäuser, 2001.
- [17] H. Feichtinger, T. Strohmer (Eds.), Advances in Gabor Analysis, Birkhäuser, Boston, 2002.
- [18] R.G. Stockwell, L. Mansinha and R. P. Lowe, Localization of the Complex Spectrum: The S Transfom. IEEE Transactions on Signal Processing, Vol.44, No. 4, 998-1001, 1996.
- [19] J. Lin and L. Qu, Feature extraction based on the Morlet wavelet and its application for mechanical fault diagnosis, J. Sound Vibr. 234, 135-148, 2000.
- [20] F. Jaillet, B. Torrésani, Time-Frequency jigsaw puzzle: Adaptive multiwindow and multilayered Gabor expansions, Int. J. Wavelets, Multiresolut. Inf. Process.5(2), 293-315, 2007.
- [21] F. Jaillet, P. Balazs, M. Dorfler, Non-stationary Gabor frames, in: 8th Int. Conf. on Sampling Theory and Applications, May 18-22, Marseille, France, 2009. Disponível em http://www.latp.univ-mrs.fr/SAMPTA09/.
- [22] C.G. Mosquera, A.M. Trigueros, J.I. Franco, Time-Frequency EEG analysis in epilepsy: What is more suitable?, Signal Processing and Information Technology. Proceedings of the Fifth IEEE International Symposium, 2005. ISBN: 0-7803-9313-9.
- [23] G.V. Hoeym, W. Philips, I. Lemahieu, Time-Frequency Analysis of EEG Signals, Proceedings of the ProRISC Workshop on Circuits, Systems and Signal Processing, 1997.
- Saul, Stanford |24| R. Fisher М. Overview Epilepsy, Neurology, and of Comprehensive Epilepsy Center, copyright 1997, 2006.Disponível em http://neurology.stanford.edu/divisions/e_handout.html.
- [25] R.M.F. Fernandes, O Eletroencefalograma na Caracterização das Síndromes Epilépticas. Departamento de Neurologia, Psiquiatria e Psicologia Médica da Faculdade de Medicina de Ribeirão Preto, Universidade de São Paulo. Disponível em http://lasse.med.br/mat_didatico/lasse1/textos/regina01.html. Acessado em 01/10/2011.
- [26] S. Blanco, R.Q. Quiroga, O.A. Rosso, and S. Kochen, Time-Frequency analysis of electroencephalogram series, Physical Review E, Volume 51, Number 3, 1995.
- [27] R.Q. Quiroga, Quantitative analysis of EEG signals: Time-frequency methods and Chaos theory, Institute of Physiology - Medical University Lubeck and Institute of Signal Processing - Medical University Lubeck, Thesis, 1998.
- [28] D. Starling, Temporal Analysis of EEG patterns in a biofeedback based Brain Computer Interface, BSc Computer Science and Cybernetics, Brain computer Interface, 2003.

- [29] J. Malmivuo and R. Plonsey, Bioelectromagnetism Principles and Applications of Bioelectric and Biomagnetic Fields, Oxford University PRESS, 1995.
- [30] E. Niedermeyer and F. Lopes da Silva, Electroencephalography: Basic Principles, Clinical Applications, and Related Fields, page 140. Lippincott Williams & Wilkins, ISBN 0-7817-5126-8, 2004.
- [31] S.S. Gusmão, G.B. Campos, e A.L. Teixeira, Exame Neurológico, Bases Anatomofuncionais, Revinter, Segunda Edição, 2007.
- [32] R.Q. Quiroga, H. Garcia, and A. Rabinowicz, Frequency Evolution during tonicclonic seizures, 1999.
- [33] H. Gastaut and R. Broughton, Epileptic Seizures. Thomas, Springfield, 1972.
- [34] W. Abdulla and L. Wong, Neonatal EEG signal characteristics using time frequency analysis, Physica A 390, 1096-1110, 2011.
- [35] A.T. Tzallas, M.G. Tsipouras, and D.I. Fotiadis, Automatic Seizure Detection Based on Time-Frequency Analysis and Artificial Neural Networks, Research Article, Hindawi Publishing Corporation, Computational Intelligence and Neuroscience, Article ID 80510, 2007.
- [36] R.Q. Quiroga, S. Blanco, O.A. Rosso, H. Garcia, and A. Rabinowicz, Searching for hidden information with Gabor Transform in generalized tonic-clonic seizures, Eletroencephalography and clinical Neurophysiology 103, 434-439, Elsevier, 1997.
- [37] Disponível em http://en.wikipedia.org/wiki/Electroencephalography. Acessado em 15/01/2012.
- [38] J.R.A. Torreão, J.L. Fernandes, and S.M.C. Victer, A model for neuronal signal representation by stimulus-dependent receptive fields, Procs. 19th International Conference on Artificial Neural Networks - ICANN 2009, vol.I, 356-362, 2009.
- [39] M.F. Bear, B.W. Connors, and M.A. Paradiso, Neuroscience: Exploring the Brain, Lippincott Williams & Wilkins, Baltimore, 2001.
- [40] D.H. Hubel and T. Wiesel, Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. J. Physiol. 160, 106-154, 1962.
- [41] B. A. Wandell, Foundations of Vision, 1995.
- [42] A. Hyvärinen, J. Hurri, and P.O. Hoyer, Natural Image Statistics, A probabilistic approach to early computational vision, Springer, February 27, 2009.
- [43] B.A. Olshausen and D.J. Field, Emergence of simple-cell receptive field properties by learning a sparse code for natural images, Nature 381, 607-609, 1996.
- [44] D.G. Albrecht and L.G. Thorell, Spatial frequency selectivity of cells in macaque primary visual cortex, Vision Res. 22, 545-559, 1987.
- [45] D.L. Ringach, Spatial structure and symmetry of simple-cell receptive fields in macaque primary visual cortex, J. Neurophysiol. 88 455-463, 2002.

- [46] J.H. van Hateren and A. van der Schaaf, Independent component filters of natural images compared with simple cells in primary visual cortex, Proc. R. Soc. Lond. B265 359-366, 1998.
- [47] J.J. Atick and A.N. Redlich, What does the retina know about natural scenes, Neural Comp. 4, 196-210, 1992.
- [48] D.L. Ruderman and W. Bialek, Statistics of natural images: Scaling in the woods, Phys. Rev. Lett. 73(6), 814-817, 1994.
- [49] E.H. Adelson and J.R. Bergen, Spatiotemporal energy models for the perception of motion. J. Opt. Soc. Am. A2 284-299, 1985.
- [50] I. Ohzawa, G.C. DeAngelis and R.D. Freeman, Stereoscopic depth discrimination in the visual cortex: Neurons ideally suited as disparity detectors. Science 249, 1037-1041, 1990.
- [51] I. Ohzawa, G.C. DeAngelis and R.D. Freeman, Encoding of binocular disparity by complex cells in the cat's visual cortex. J. Neurophysiol. 77, 2879-2909, 1997.
- [52] Disponível em http://www.positscience.com/human-brain/image-gallery/visionimages?page=1. Acessado em 10/01/2012.
- [53] Y. Chen and N. Qian, A coarse-to-fine disparity energy model with both phase-shift and position-shift receptive field mechanisms, Neural Computation 16, 1545-1577, 2004.
- [54] E.T. Rolls and G. Deco, Computational Neuroscience of Vision, Oxford University Press, New York, 2003.
- [55] L.B. Almeida, The fractional Fourier transform and time-frequency representations, IEEE Trans. Signal Processing 42, 3084-3091, 1994.
- [56] A.I. Zayed, A convolution and product theorem for the fractional Fourier transform, IEEE Signal Processing Letters 5(4), 101-103, 1998.
- [57] J.R.A. Torreão, Generalizando a Transformada de Gabor Sintonizada, Prépublicação D078/2010 do IMPA, disponível em http://www.preprint.impa.br/cgibin/MMMsearch.cgi.
- [58] K.B. Wolf, Integral transforms in science and engineering, (Plenum Press, New York), 1979.
- [59] J.A. Movshon, I.D. Thompson and D.J. Tolhurst. Spatial summation in the receptive fields of simple cells in the cat's striate cortex. J. Physiology 293, 101-120, 1978.
- [60] S.S. Soliman, M.D. Srinath, Continuous and Discrete Signals and Systems, 2nd ed., Prentice-Hall, Upper Saddle River, 1990.
- [61] C. Enroth-Cugell, J.G. Robson, D.E. Schweitzer-Tong and A.B. Watson, Spatiotemporal interactions in cat retinal ganglion cells showing linear spatial summation, J. Physiol. 341, 279-301, 1983.

- [62] J.R.A. Torreão, S.M.C. Victer, Towards a Biologically Plausible Stereo Approach, in J.R.A. Torreão, editor. Advances in Stereo Vision, Intech, Rijeka, Croácia, 57-70, 2011.
- [63] A.D. Jepson, and M.R.M. Jenkin, Phase-based disparity measurement, CVGIP: Image Understanding 53(2), 198-210, 1991.