

*Uma Nova Representação  
Espaço-Frequência,  
com Aplicações à Análise de Sinais  
e à Visão Computacional*

**José R.A. Torreão**

**Instituto de Computação, UFF**

**Com a colaboração de**

***João Luiz Fernandes e Silvia Mara Victer (IC-UFF)***

## ***Introdução***

---

### *Transformada de Fourier*

➔ Representação de sinais em termos de exponenciais complexas

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

➤ Os pesos são funções apenas da *frequência*

$$\tilde{I}(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dx e^{-i\omega x} I(x)$$

### *Expressão Alternativa para a TF*

$$I(x) = \int_{-\infty}^{\infty} d\omega e^{i\omega x} * e^{i\omega(x-a)} e^{-\frac{x^2}{2\sigma^2}}$$

onde

$$\sigma(\omega) = \frac{1}{(2\pi)^{3/2}} |\tilde{I}(\omega)| \quad \text{e} \quad a(\omega) = -\frac{\varphi_{\tilde{I}}(\omega)}{\omega}, \quad \omega \neq 0$$

➤ Pesos são funções localizadas da *posição* e da *frequência*

$$\text{Funções de Gabor: } \psi_{\omega}(x) = e^{i\omega(x-a)} e^{-\frac{x^2}{2\sigma^2}}$$

$$\tilde{I}(\omega) = \int_{-\infty}^{\infty} d\omega' \mathcal{F}\{e^{i\omega'x}\} \times \mathcal{F}\{e^{i\omega'(x-a)} e^{-\frac{(x-a)^2}{2\sigma^2}}\}$$

$$\tilde{I}(\omega) = (2\pi)^{3/2} \int_{-\infty}^{\infty} d\omega' \delta(\omega - \omega') \sigma(\omega') e^{-i\omega a} e^{-\frac{\sigma^2(\omega')}{2}(\omega - \omega')^2}$$

$$\tilde{I}(\omega) = (2\pi)^{3/2} \sigma(\omega) e^{-i\omega a}$$

Obs.: A função de representação

$$e^{i\omega(x-a)} e^{-\frac{(x-a)^2}{2\sigma^2}}$$

também funciona

*A expressão alternativa pode ser reescrita como*

$$I(x) = \int_{-\infty}^{\infty} d\omega e^{i\omega x} \langle \psi_{\omega}(x), e^{i\omega x} \rangle$$

onde  $\langle \cdot, \cdot \rangle$  denota um produto interno

Comparando com a expressão para uma expansão na base

$$I(x) = \int_{-\infty}^{\infty} d\omega e^{i\omega x} \langle I(x), e^{i\omega x} \rangle$$

➤ *Concluimos que, no que refere à frequência  $\omega$ ,*

$$\psi_{\omega}(x) \equiv I(x)$$

➤ Dado o sinal  $I(x)$

A função

$$\psi_{\omega}(x) = e^{i\omega(x-a)} e^{-\frac{x^2}{2\sigma^2}}$$

com

$$\sigma(\omega) = \frac{1}{(2\pi)^{3/2}} |\tilde{I}(\omega)| \quad \text{e} \quad a(\omega) = -\frac{\varphi_{\tilde{I}}(\omega)}{\omega}, \quad \omega \neq 0$$

corresponde a uma *representação local oscilatória do conteúdo do sinal na frequência  $\omega$*

## Relação com a Representação de Gabor

*Representação de Gabor tradicional*

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

onde  $\sigma$  não depende da frequência

*A nova representação pode ser reescrita como*

$$I(x) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} d\omega d\xi e^{i\omega[x-a(\omega)]} e^{-\frac{(x-\xi)^2}{2\sigma^2(\omega)}}$$

➤ Representação com coeficientes 1 ou 0

## Aplicação: ‘Signal-Tuned Gabor Analysis’

➔ *Transformada de Gabor baseada na nova representação:*

$$T(\omega, b) = \int_{-\infty}^{\infty} dx e^{-i\omega[x-a(\omega)]} e^{-\frac{(x-b)^2}{2\sigma^2(\omega)}} I(x)$$

➤ pode ser interpretada como uma *correlação* entre o sinal e a *sua representação* na frequência  $\omega$

De fato, 
$$|\tilde{I}(\omega)|^2 = 2\pi \int_{-\infty}^{\infty} db T(\omega, b)$$

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

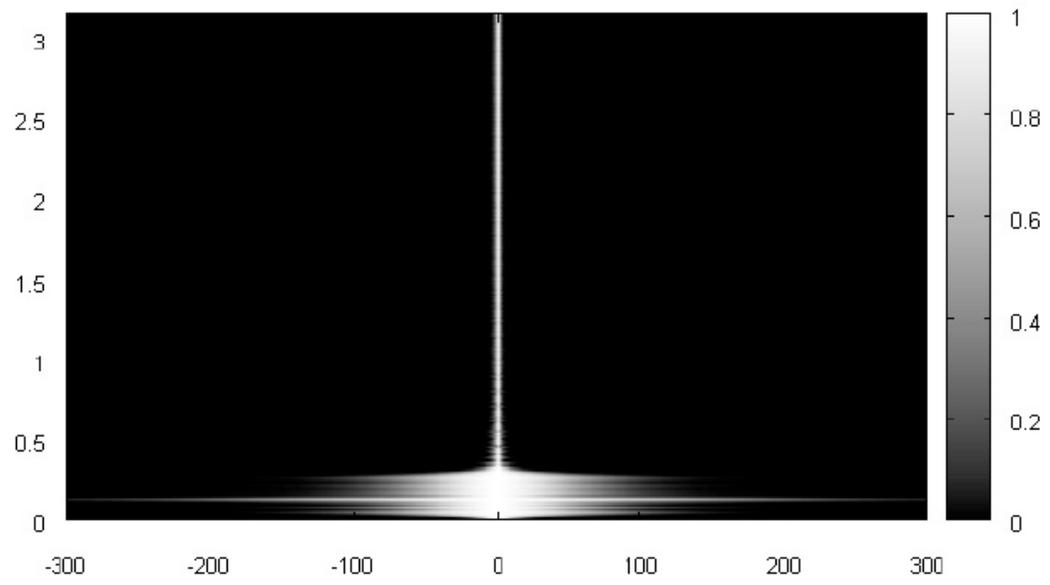
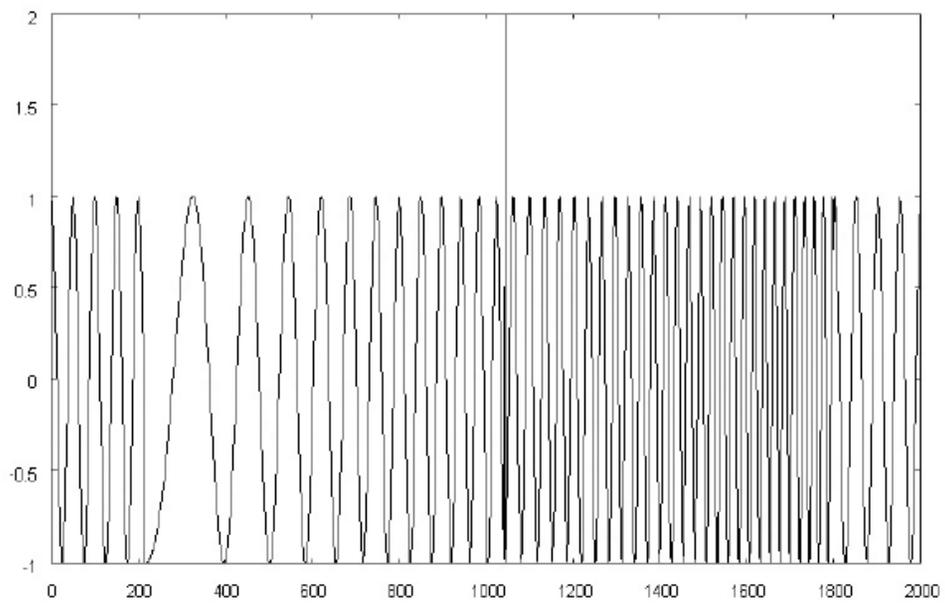
➤ Transformada de Gabor *sintonizada* ao sinal

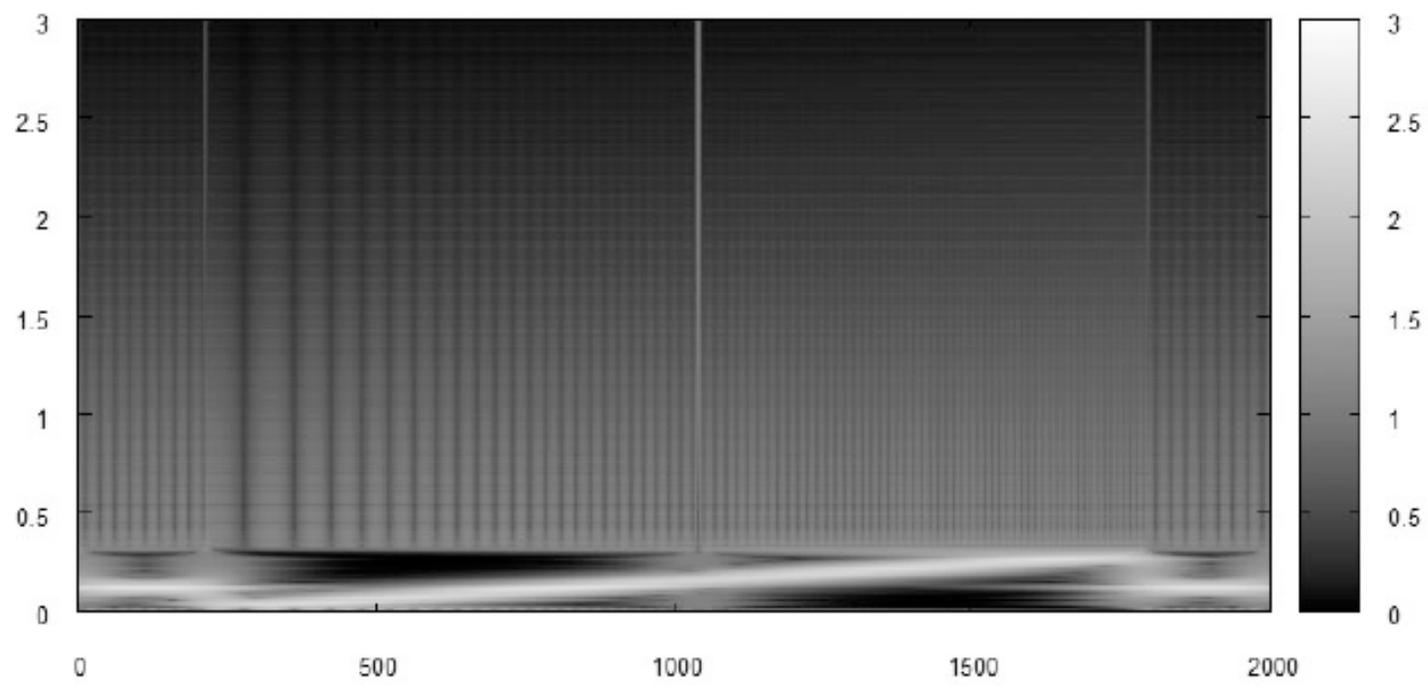
*Largura e fase das funções analisadoras relacionadas ao conteúdo em frequência do sinal analisado*

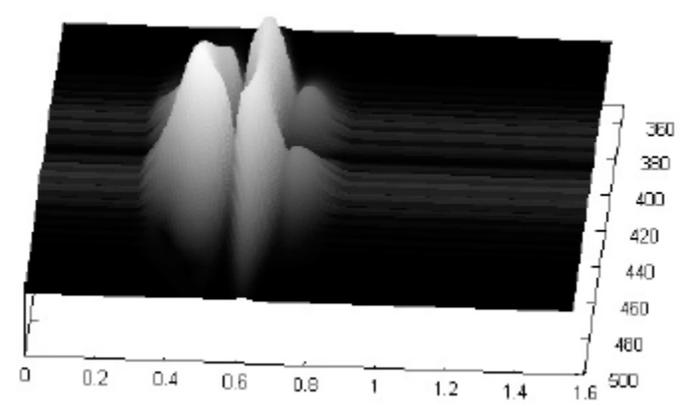
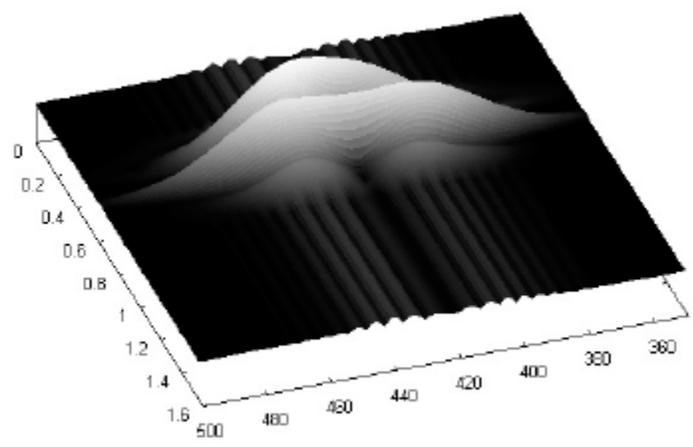
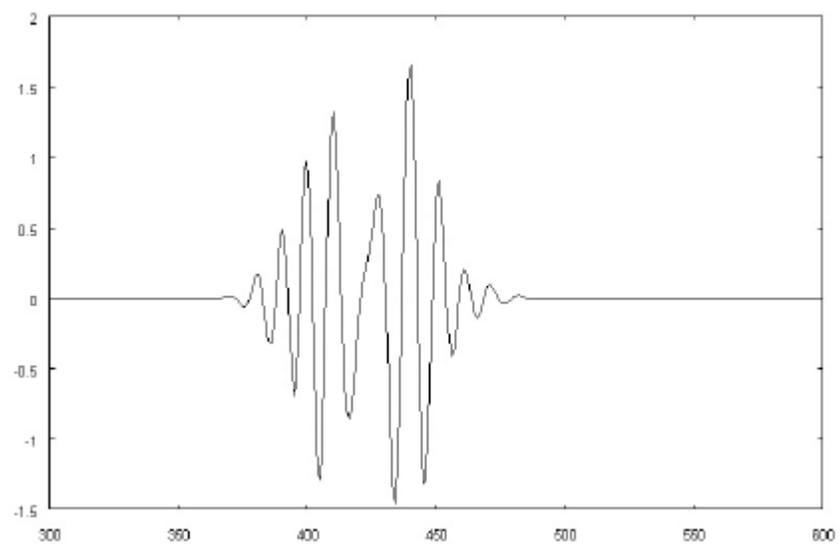
Propriedades semelhantes às da transformada *wavelet*

- ⇒ Características transientes analisadas por funções de *menor escala e maior largura de banda*
- ⇒ Características estacionárias analisadas por funções de *maior escala e menor largura de banda*

*Sem requerer ajuste de parâmetros*







# Representação de Imagens

⇒ *Extensão para 2D*

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

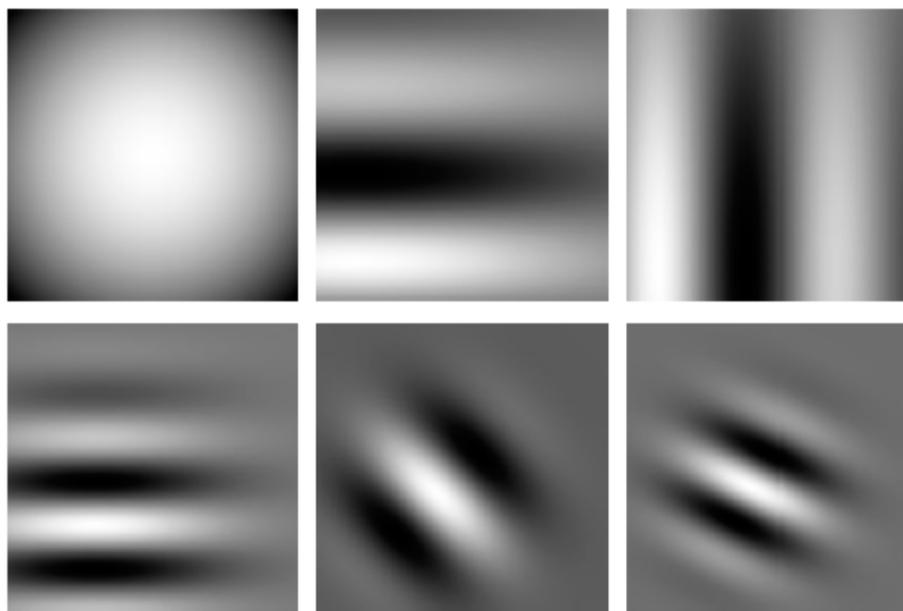
$$\sigma(\omega_x, \omega_y) = \frac{1}{(2\pi)^{3/2}} \sqrt{|\tilde{I}(\omega_x, \omega_y)|}$$

$$a(\omega_x, \omega_y) = -\frac{\varphi_{\tilde{I}}(\omega_x, \omega_y)}{\omega_x + \omega_y}, \quad (\omega_x, \omega_y) \neq (0, 0)$$

⇒ Funções de representação

$$e^{i[\omega_x(x-a) + \omega_y(y-a)]} e^{-\left(\frac{x^2 + y^2}{2\sigma^2}\right)}$$

Exemplos:



⇒ *Janela 3x3*

## ➔ *Representações em Janelas*

- Diferentes  $\sigma$  e  $a$  em cada janela

### *Originais X Representações*



➔ *Erro médio de 3.9%*

➔ *Interpretação 'física' da nova representação*

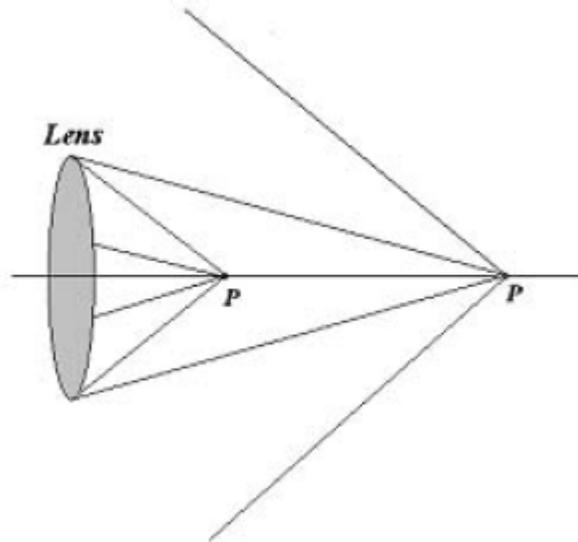
$$I(x, y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} d\omega_x d\omega_y e^{i(\omega_x x + \omega_y y)} * e^{i[\omega_x(x-a) + \omega_y(y-a)]} e^{-\left(\frac{x^2 + y^2}{2\sigma^2}\right)}$$

➤ *Mantendo apenas o termo DC:*

$$I(x, y) \approx \exp\left[-\frac{x^2 + y^2}{2\sigma^2(0, 0)}\right] * 1$$

$$\tilde{I}(\omega_x, \omega_y) \approx (2\pi)^3 \sigma^2(0, 0) \delta(\omega_x, \omega_y)$$

➤ *Desfocamento de uma superfície uniforme distante, por uma câmera com foco no infinito*

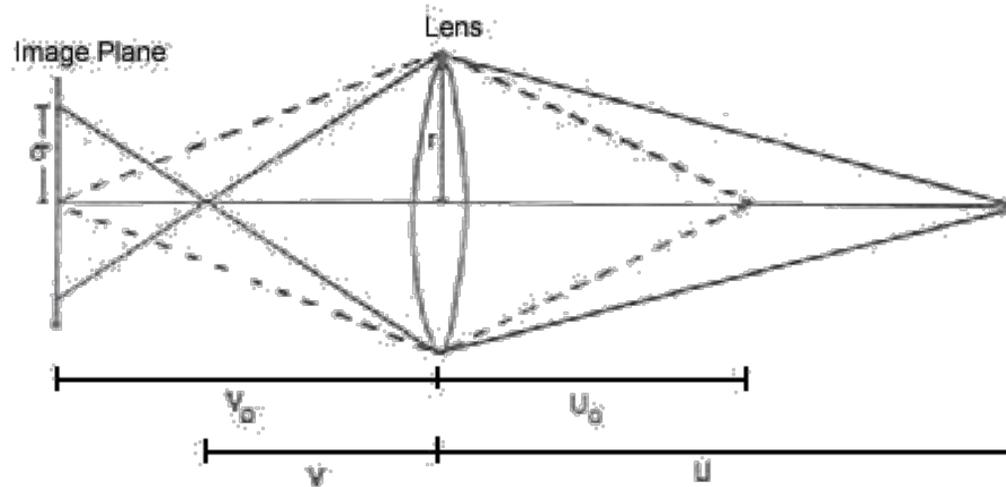


- O ângulo sólido abrangido pelos raios emanando de  $P$  decresce com  $Z^2$
- Se vamos modelar o desfocamento pela convolução com uma gaussiana, devemos ter

$$\frac{S}{2\pi(\sigma Z)^2} \exp \left[ - \left( \frac{x^2 + y^2}{2\sigma^2} \right) \right]$$

onde  $S$  é a área da lente

*Pentland:*



$$Z(x, y) = \frac{F v_0}{(v_0 - F) - f \sigma(x, y)} \quad \text{e} \quad \sigma Z = \frac{(v_0 - F) Z}{f} - \frac{F v_0}{f}$$

➤  $\sigma Z$  torna-se independente de  $Z$ , se

$$(v_0 - F) \approx 0 \quad \text{e} \quad Z \rightarrow \infty$$

➤ Nestas condições, geralmente assumidas em SFS, o desfocamento pode ser modelado por uma gaussiana *não-normalizada*

Voltando à representação:

- Para uma frequência qualquer (a menos da fase),

$$\exp[i(\omega_x^0 x + \omega_y^0 y)] \exp\left[-\frac{x^2 + y^2}{2\sigma^2(\omega_x^0, \omega_y^0)}\right] * \exp[i(\omega_x^0 x + \omega_y^0 y)]$$

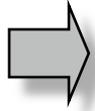
- No domínio da frequência

$$(2\pi)^3 \sigma^2(\omega_x^0, \omega_y^0) \delta(\omega_x - \omega_x^0, \omega_y - \omega_y^0)$$

- Efeito semelhante ao do termo DC:

*‘Desfocamento’ de uma superfície senoidal distante, por uma câmera com foco no infinito*

⇒ *Imagens: Combinação de ‘desfocamentos generalizados’ em várias frequências*



Propusemos um *Shape from Shading* via *Shape from Defocus*:

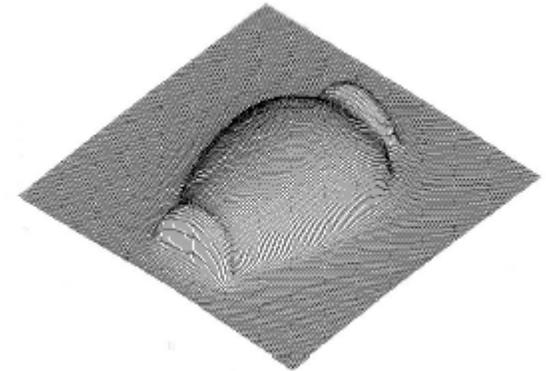
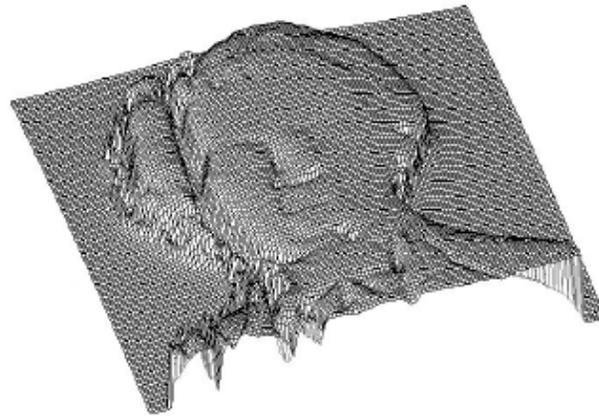
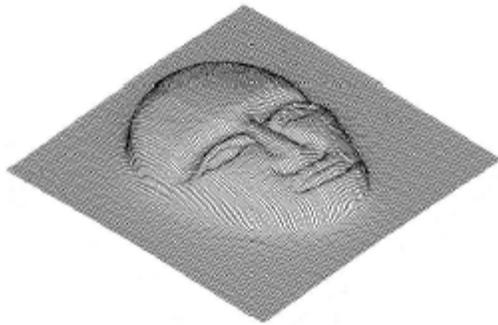
*Fórmula de Pentland:*

$$Z(x, y) = \frac{Fv_0}{(v_0 - F) - f\sigma(x, y)}$$

No nosso caso, para cada frequência:

$$Z_{\omega_x, \omega_y}(x, y) = \alpha \frac{1}{\phi - \sqrt{|\tilde{I}(\omega_x, \omega_y)|}}$$

$$Z(x, y) = \sum_{\omega_x} \sum_{\omega_y} \cos[\varphi_{\tilde{I}}(\omega_x, \omega_y)] Z_{\omega_x, \omega_y}(x, y)$$

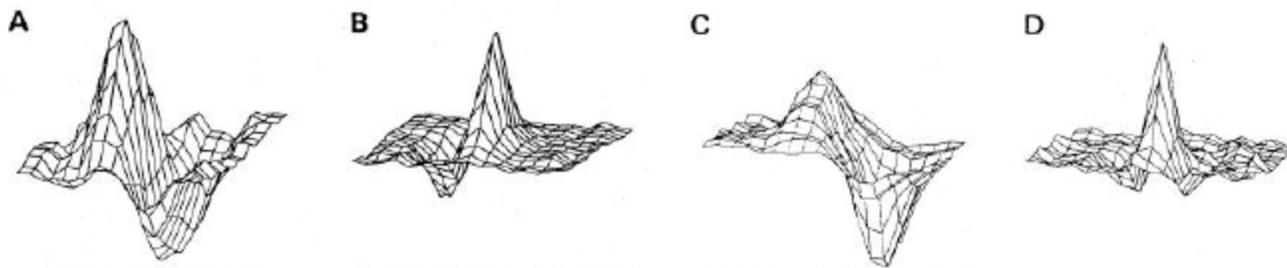


## Possível Aplicação à Neurociência

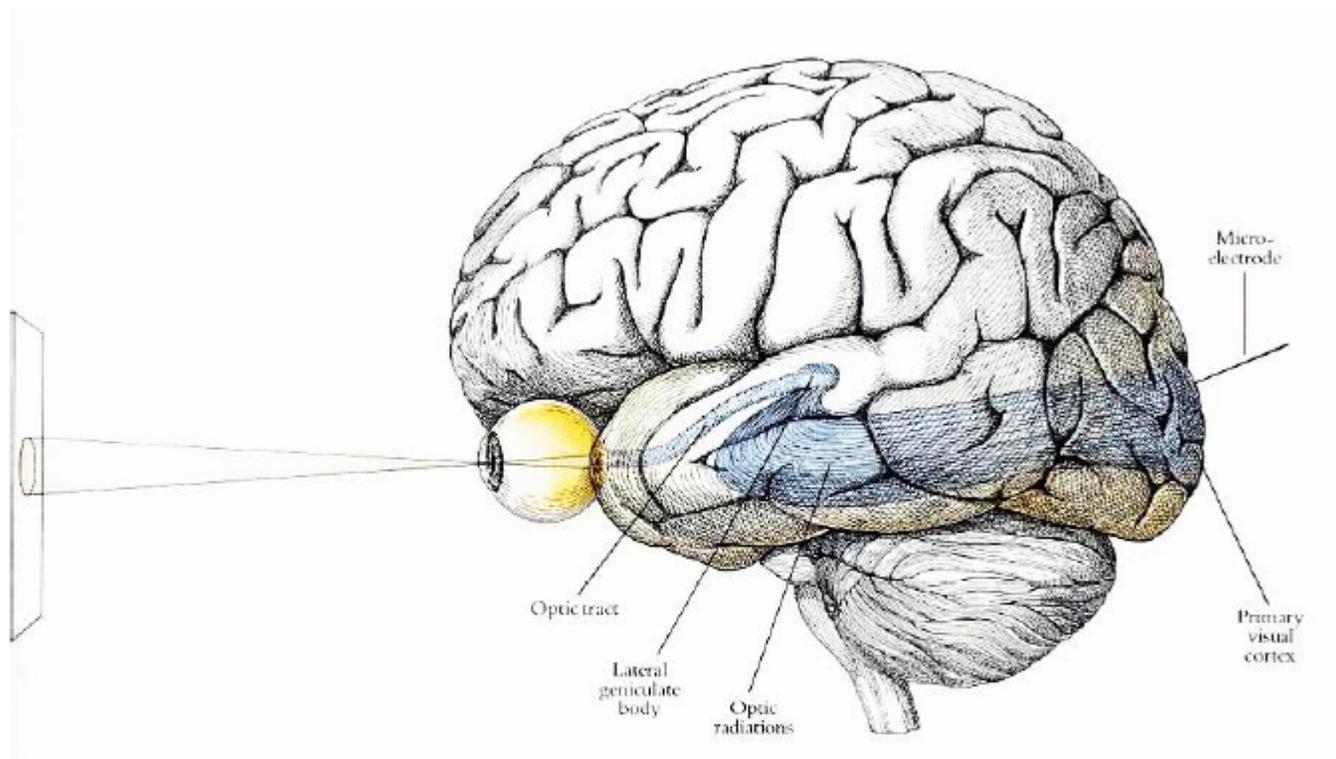
### ➔ *Modelagem do Campo Receptivo de Células Corticais*

Campo Receptivo Clássico: *Independente do Sinal*

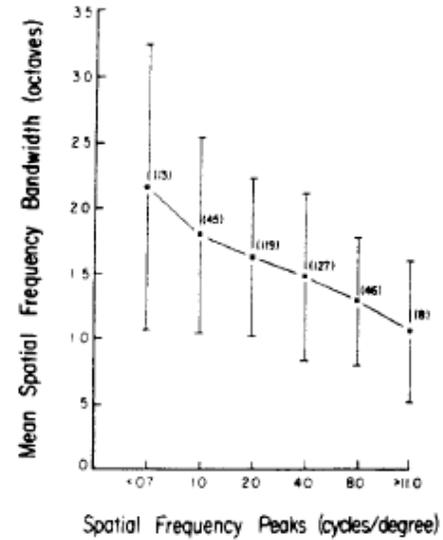
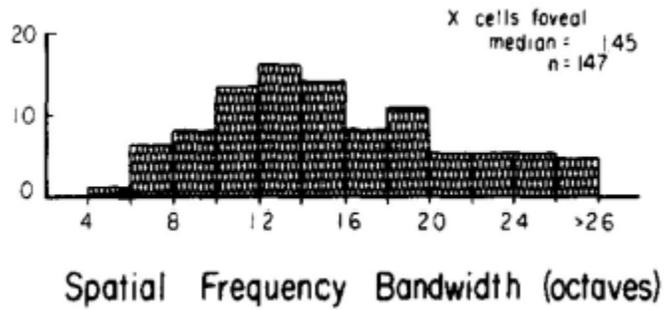
Células Corticais Simples: *CR ~ Funções de Gabor*



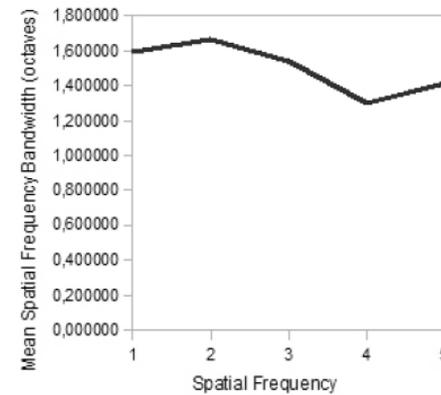
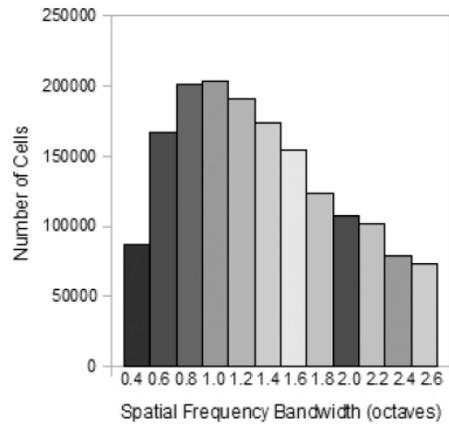
### ➔ *Evidência Recente: CR depende da entrada*



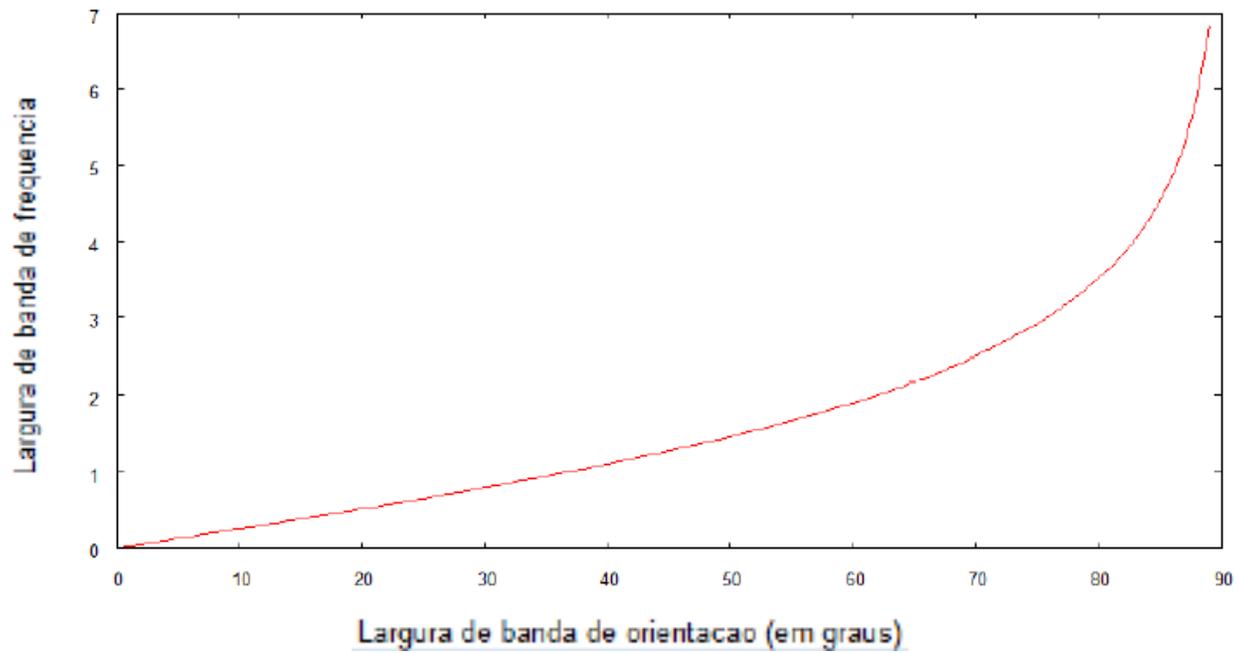
*Neurofisiologia:*



*Nova Representação (Imagens Naturais):*



*Correlação entre as larguras de banda em orientação e frequência*



## **Outras Possíveis Aplicações**

---



*Abordagens biologicamente plausíveis  
em Visão Computacional*

➤ *Estereoscopia via Funções de Green*

➤ *Modelo Linear-Não-Linear para o SFS*

## Estéreo via Funções de Green

### *Equações de Casamento*

$$I_l(x + U) = I_r(x)$$

e.g., Casamento Uniforme:  $U(x) = u$

$$\frac{u^2}{2} I_l'' + u I_l' + I_l = I_r$$

$$\Rightarrow I_l(x) \equiv I_r(x - u) = \int_{-\infty}^{\infty} G_u(x, x_0) I_r(x_0) dx_0$$

$$G_u(x, x_0) = \frac{2}{u} \sin\left(\frac{x - x_0}{u}\right) \exp\left[-\left(\frac{x - x_0}{u}\right)\right]$$

para  $x > x_0$

## *Equação considerada*

$$\frac{1}{2} \left( \frac{\sigma^2}{a} \right)^2 I_l'' + \left( \frac{\sigma^2}{a^2} \right) (x + a) I_l' + \left[ 1 + \frac{(x + a)^2 + (\sigma^2 - a^2)}{2a^2} \right] I_l = I_r$$

- Soluções homogêneas são funções de Gabor
- Aproxima equação de casamento *afim*, i.e., com

$$U(x) = \frac{\sigma^2}{a} + \frac{\sigma^2}{a^2} x$$

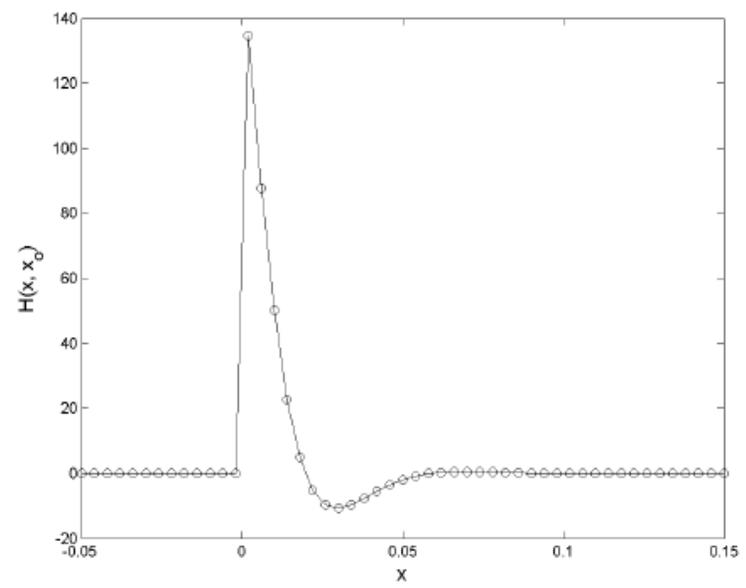
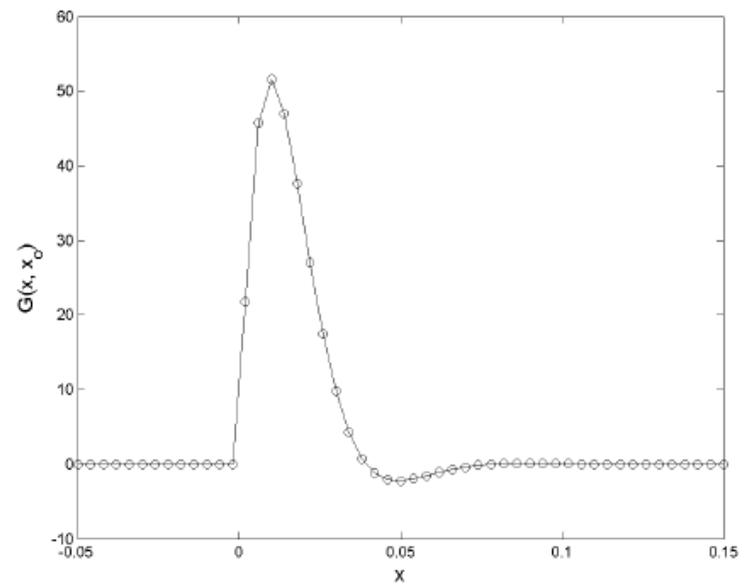
para  $x$ ,  $\sigma^2/a$  pequenos

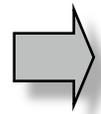
➔ *Núcleo de Green complexo*

$$K(x, x_0) = 2k \exp[ik(x - x_0)] \exp \left[ -\frac{(x + a)^2 - (x_0 + a)^2}{2\sigma^2} \right]$$

para  $x > x_0$

com  $u \equiv 1/k \equiv \sigma^2/a$





*Paralelo com Estéreo Neurofisiológico ('Modelo da Energia')*

- Imagem direita *vista sob CR*, modelado por Gabor

$$I_1(x) = \exp[i(\kappa x + \phi)] \exp\left[-\frac{(x + \alpha)^2}{2\sigma^2}\right] I_r(x)$$

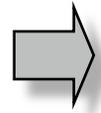
- Filtragem pelo núcleo de Green

$$I_2(x) = \int_{-\infty}^{\infty} K(x, x_0) I_1(x_0) dx_0$$

- Imagem esquerda *vista sob CR*, modelado por Gabor:  $\tilde{I}_1(x)$
- Filtragem pelo núcleo de Green:  $\tilde{I}_2(x)$

- Obtenção da disparidade que minimiza o erro

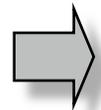
$$R(x) = |I_2(x) - \tilde{I}_2(x)|^2$$



*Nesta abordagem, assim como no Modelo da Energia, os campos receptivos são modelados por funções de Gabor empíricas*

$$\exp[i(\kappa x + \phi)] \exp\left[-\frac{(x + \alpha)^2}{2\sigma^2}\right]$$

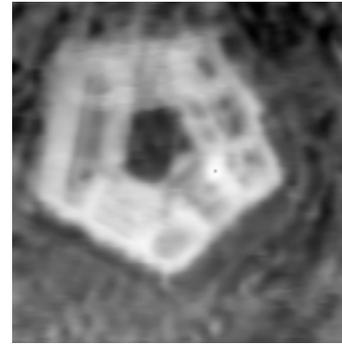
➤ *Parâmetros escolhidos ‘com a mão’*



*Utilizando as funções de Gabor obtidas com o novo modelo de representação, conseguimos resultados semelhantes ou superiores aos obtidos anteriormente*



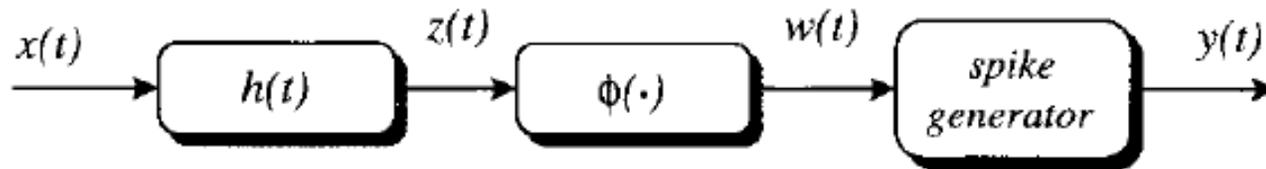
*Sem necessidade de ajuste de parâmetros*



## Modelo LN para o Shape from Shading

---

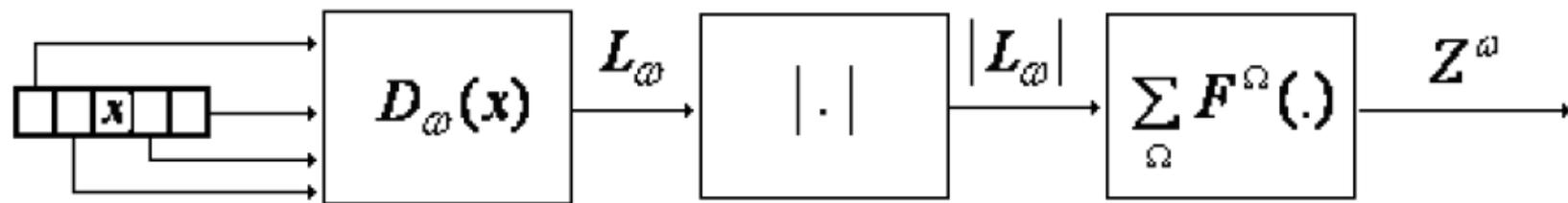
➔ *Modelo LN: Cascata de estágios Linear e Não-Linear*



- Utilizado na modelagem de diversos tipos de neurônios
- Não-Linearidade escolhida sob diferentes formas: função *limiar*, *gaussiana*, *sigmóide*, etc.

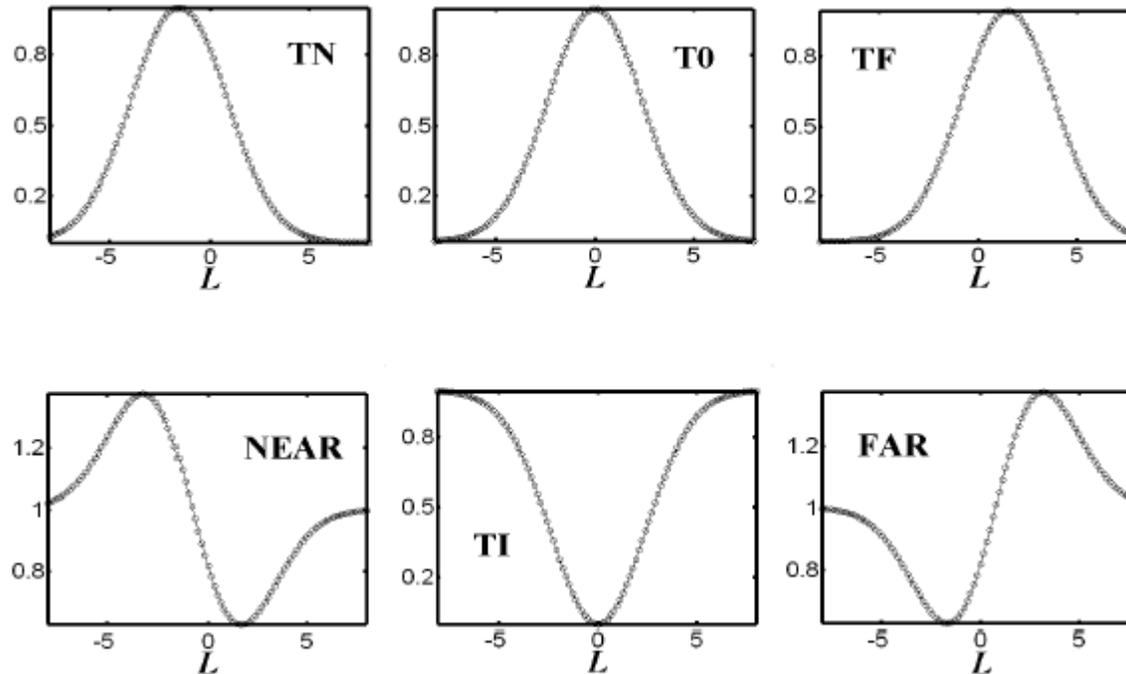
*SFS: Não existe uma fundamentação neurofisiológica sólida*

⇒ Propusemos um modelo *LN* para o SFS



- Filtros lineares são funções de Gabor de diferentes frequências
  - Estágio não-linear é uma combinação de funções modeladas nos *perfis de sintonia à disparidade* dos neurônios binoculares
- ⇒ Estabelecemos um paralelo entre SFS e Estereoscopia

Modelo para as curvas de sintonia a disparidade:



➔ Saída do estágio linear tem dimensão de deslocamento ( $I$  adimensional)

$$L_{\omega} = \int dx e^{i\omega x} e^{-\frac{(x-x_0)^2}{2\sigma^2}} I(x, y)$$

➤ Respostas *sintonizadas* modeladas por gaussianas:

*Tuned Near, Tuned Far*

$$R_{T_n, T_f}^\Omega(L) = \exp \left[ -\frac{(L \pm |L_\Omega|)^2}{\Sigma^2} \right]$$

*Tuned Zero*

$$R_{T_n}^\Omega(L) = R_{T_f}^\Omega(L) \equiv R_{T_0}(L) = \exp \left[ -\frac{L^2}{\Sigma^2} \right]$$

*Tuned Inhibitory*

$$R_{T_i}(L) = 1 - \exp \left( -\frac{L^2}{\Sigma^2} \right)$$

➤ Nossa resposta não-linear:

$$F^\Omega(|L_\omega|) = [R_{T_0} \delta(L_\Omega) + R_{T_i} + R_{T_n}^\Omega + R_{T_f}^\Omega](|L_\omega|)$$

➡ Quando  $L_\Omega$  é diferente de zero, isto gera os perfis *Near e Far*

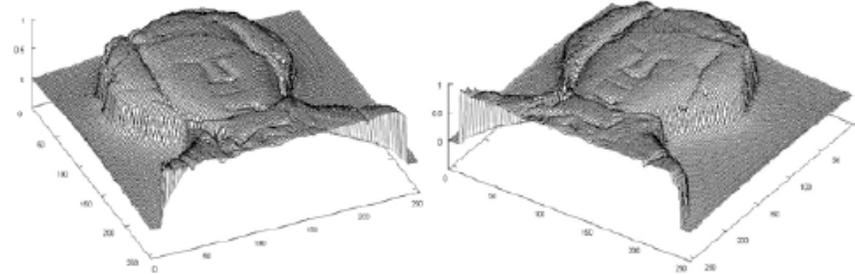
## Reconstruções:

$$Z = \sum_{\omega} Z^{\omega} = \sum_{\omega, \Omega} F^{\Omega}(|L_{\omega}|)$$

- *Independentes do Mapa de Reflectância*
- *Mais próximas de um SFS biológico*
- *Porém: Filtros lineares são funções de Gabor empíricas*
- *Proposta: Usar as funções da nova representação espaço-freqüência*



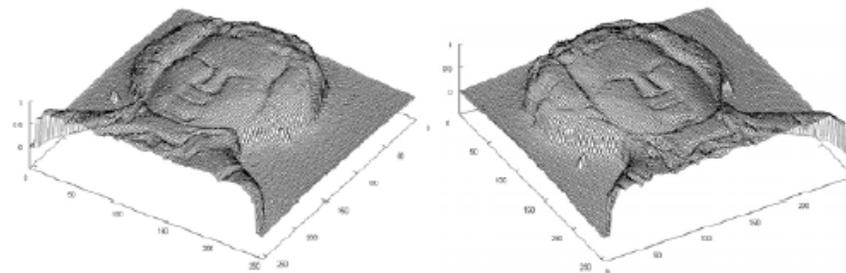
a



b



a



b

## Outras Possíveis Aplicações

---

- *Todas em que funções de Gabor são usadas*
- *Também: Variantes da representação*
  - *Outras formas de janela: e.g., pillbox, PSF física*
  - *Funções de base também na forma de Gabor*
  - *Outras transformadas: e.g., Cosseno*

# Conclusão

---

?

*(Em Julho de 2010...)*