

UNIVERSIDADE FEDERAL DO MARANHÃO
CENTRO DE CIÊNCIAS EXATAS E TECNOLOGIA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA DE ELETRICIDADE

ANDRÉ BORGES CAVALCANTE

Campos receptivos similares às wavelets de Haar são gerados a
partir da codificação eficiente de imagens urbanas

São Luís - MA

2008

ANDRÉ BORGES CAVALCANTE

Campos receptivos similares às wavelets de Haar são gerados a partir da codificação eficiente de imagens urbanas

Dissertação apresentada ao Programa de Pós-Graduação em Engenharia de Eletricidade da UFMA, como requisito para a obtenção parcial do grau de MESTRE em Engenharia de Eletricidade.

Orientador: Prof. Allan Kardec Duailibe Barros Filho, Dr.

(Orientador)

São Luís - MA

2008

Cavalcante, Andr Borges

Campos receptivos similares às wavelets de Haar são gerados a partir da codificação eficiente de imagens urbanas. / André Borges Cavalcante. - 2008.

54f

Impresso por computador (fotocópia).

Orientador: Allan Kardec Duailibe Barros Filho.

Dissertação (Mestrado) - Programa de Pós-Graduação em Engenharia de Eletricidade, Universidade Federal do Maranhão, São Luís, 2007.

1. Codificação eficiente. 2. Imagens naturais. 3. Imagens urbanas. 4. Campos receptivos. 5. Córtex visual primrio (V1) I. Orientador: Barros Filho, Allan Kardec Duailibe. II. Título.

CDU : 004.627

ANDRÉ BORGES CAVALCANTE

Campos receptivos similares às wavelets de Haar são gerados a partir da codificação eficiente de imagens urbanas

Dissertação apresentada ao Programa de Pós-Graduação em Engenharia de Eletricidade da UFMA, como requisito para a obtenção parcial do grau de MESTRE em Engenharia de Eletricidade.

Aprovada em

BANCA EXAMINADORA

Prof. Allan Kardec Duailibe Barros Filho, Dr.

(Orientador)

Prof. Hani Camille Yehia, Dr.

(Membro da Banca Examinadora)

Prof. João Viana da Fonseca Neto, Dr.

(Membro da Banca Examinadora)

Aos meus pais e minha irmã.

Resumo

A codificação eficiente de imagens naturais gera filtros similares às wavelets de Gabor que relembram os campos receptivos de células simples do córtex visual primário. No entanto, imagens naturais e urbanas tem características estatísticas diferentes. Será mostrado que uma simples análise do espectro de potência em um modelo eficiente sugere que imagens naturais e urbanas requerem filtros específicos para cada grupo. De fato, aplicando codificação eficiente à imagens urbanas, encontramos filtros similares às wavelets de Gabor e de Haar. Além disso, observou-se que imagens urbanas quando projetadas nesses filtros geraram um menor erro médio quadrático do que quando projetadas somente em filtros de similares a Gabor. Desta forma, como imagens naturais e urbanas requerem filtros diferentes para serem representadas de forma eficiente, estes resultados sugerem que além de Gabor, o córtex visual primário também deve possuir células com campos receptivos similares às wavelets de Haar.

PALAVRAS-CHAVE: codificação eficiente, imagens naturais, imagens urbanas, campos receptivos, córtex visual primário, V1

Abstract

Efficient coding of natural images yields filters similar to the Gabor-like receptive fields of simple cells of primary visual cortex. However, natural and man-made images have different statistical proprieties. Here we show that a simple theoretical analysis of power spectra in a sparse model suggests that natural and man-made images would need specific filters for each group. Indeed, when applying sparse coding to man-made scenes, we found both Gabor and Haar wavelet-like filters. Furthermore, we found that man-made images when projected on those filters yielded smaller mean squared error than when projected on Gabor-like filters only. Thus, as natural and man-made images require different filters to be efficiently represented, these results suggest that besides Gabor, the primary visual cortex should also have cells with Haar-like receptive fields.

KEYWORDS: efficient coding, natural images, man-made images, receptive fields, primary visual cortex, V1

Agradecimentos

Ao Prof. Allan Kardec Barros pelo trabalho e atenção dedicada no desenvolvimento de um ambiente de pesquisa como o PIB - Laboratório de Processamento da informação Biológica. Pelo entusiasmo em melhorar a situação universitária e da pesquisa no estado do Maranhão e no Nordeste brasileiro. Pelo carinho e confiança depositados em seus alunos. Além disso, queria expressar minha gratidão ao Prof. Allan por orientar minha formação como pessoa e profissional.

A Enio Aguiar pelas excelentes discussões e atenção dedicada no desenvolvimento deste trabalho.

Aos todos os meus amigos do PIB por todas as excelentes horas do café: Ewaldo Eder, Ewaldo Jr., Fábio, Ricardo Robson, Carlos Magno, Daniel, Deusdete, Márcio, Cristiane, Aurea, Sidcley, Flávio, Mendes.

Aos novos amigos pelo carinho e amizade: Bebela, Euler, Lucas, Anderson, Khalil.

Aos amigos de trabalho que moram em lugares distantes: Fausto Lucena, Denner Guilhon e Lúcio Flávio.

Aos amigos do peito por tudo: Duda, Marquinhos, Mila, Manu, Wilson, Rafael, Danilo, Ago, Diego, Kakau, Karla.

A Cylia Oliveira pelo carinho e companheirismo em todas as horas.

Aos meus pais e minha irmã pelo lar construído.

A Deus pela chance dada.

“Nada é permanente, senão a mudança.”

(Heráclito)

Sumário

Lista de Figuras	viii
Lista de Abreviações	x
1 Introdução	1
1.1 Organização do trabalho	1
2 Córtex visual	3
2.1 Córtex visual primário (V1)	3
3 Codificação eficiente ou esparsa	8
4 Análise de componentes independentes	12
4.1 O que é independência?	12
4.2 Restrições, limitações e ambiguidades de ICA	12
4.3 Modelo de ICA	13
4.4 ICA por maximização de não-gaussianidade	13
4.5 Curtose como medida de não-gaussianidade	14
4.6 Maximização de não-gaussianidade pelo método do gradiente	15
5 Imagens naturais e urbanas	16
6 Métodos	21
6.1 Aprendizado	22
6.2 Projeção	23

	vii
7 Resultados	25
7.1 Base de dados de imagens	25
7.2 Filtros A_{nat} e A_{urb}	25
7.3 Curvas de erro	27
8 Discussão	29
9 Conclusão	31
Referências	33

Lista de Figuras

2.1	Córtex Visual. (a) As setas indicam as áreas V1, V2, V3 e V4 do córtex visual. (b) Organização hierárquica e fluxo de informação bi-direcional do córtex visual primário.	3
2.2	Exemplos de campos receptivos de células simples de V1. Os marcadores χ e $\triangle$ representam áreas que evocam respostas excitatórias e inibitórias, respectivamente [HUBEL, WIESEL (1962)].	4
2.3	Campos receptivos mapeados através de correlação reversa. Os contornos sólidos representam áreas excitatórias enquanto que as regiões com contornos tracejados são inibitórias. A representação tridimensional dos campos receptivos é mostrada ao lado [DEANGELIS et al(1993)].	5
2.4	Aproximação de campos receptivos através de wavelets de Gabor. Os parâmetros associados são $\sigma_x = 1$, $\sigma_y = 2$, $1/k = 0.56$ e (a) $\phi = 1 - \pi/2$ ou (b) $\phi = 1 - \pi$ [DAYAN, ABBOTT (2001)]	6
2.5	Representação de estímulos no córtex visual. (a) Primeiramente, um estímulo visual ativa ou inibi as células da retina. Tal informação é transmitida ao córtex visual primário através do nervo óptico. (b) Em V1, os campos receptivos que correspondem as características do estímulo são utilizados para geração de uma representação neural.	7
3.1	Comparação entre a distribuição esparsa (azul) e Gaussiana (vermelha). Uma distribuição esparsa é caracterizada por possuir um pico em zero e longas extremidades.	9
3.2	Filtros gerados utilizando codificação eficiente [OLSHAUSEN, FIELD (1996)].	10
3.3	Filtros organizados a partir de suas orientações [HYVÄRINEN et al (2001)]. A organização topográfica é gerada a partir das correlações.	11

5.1	Assinaturas espectrais médias de imagens (a) naturais e (b) urbanas. Cada grupo utilizado consiste de 40 imagens. O grupo de imagens naturais inclui cenas distantes e próximas de florestas, campos, montanhas e objetos naturais como árvores e flores. O grupo de imagens urbanas inclui cenas em cidades sem restrição à presença de vegetação. Os contornos de cada assinatura correspondem a 65 % e 75 % da energia total no espectro de potência.	16
5.2	Assinaturas espectrais de imagens (a) naturais e (b) urbanas analisadas isoladamente [TORRALBA, OLIVA (2003)].	17
5.3	Espectro de potência médio de imagens (a) naturais e (b) urbanas. 6000 imagens foram utilizados para cada grupo [TORRALBA, OLIVA (2003)].	18
5.4	Representação de uma borda (a) fractal e (b) coerente.	18
5.5	Característica do alinhamento em fase para uma borda coerente de uma dimensão [OLSHAUSEN, FIELD (1996)]. As componentes de frequência devem estar todas em fase para formação de uma borda coerente.	19
6.1	Diagrama de blocos do modelo utilizado. O sistema consiste de duas fases: (a) <i>aprendizado</i> e (b) <i>projeção</i> . Na fase de aprendizado, nós utilizamos ICA para aprender filtros a partir de grupos de imagens naturais e urbanas. Na fase de projeção, foi calculado o erro médio quadrático (MSE) gerado pela representação ou projeção de imagens naturais e urbanas nos filtros gerados. . .	21
7.1	Filtros (a) $\mathbf{A}_{\text{nat}}$ e (b) $\mathbf{A}_{\text{urb}}$ gerados pelo algoritmo FastICA a partir de imagens naturais e urbanas, respectivamente.	26
7.2	Wavelets de (a) Gabor e (b) Haar.	27
7.3	Erro médio quadrático para a projeção de imagens (a) naturais e (b) urbanas nos filtros gerados por ICA.	28

Lista de Abreviações

ICA - *Independent component analysis* (análise de componentes independentes)

MSE - *Mean squared error* (erro quadrático médio)

1 Introdução

As células do córtex visual primário (V1) de mamíferos respondem preferencialmente para estruturas visuais tais como barras e bordas orientadas [HUBEL, WIESEL (1962)]. De acordo com critérios estatísticos [SHANNON (1951)], estas estruturas são classificadas como as partes informativas ou não redundantes de imagens [ATTNEAVE (1954)].

Dessa forma, foi proposto [BARLOW (1961)] [BARLOW (1989)] [FIELD (1994)] que a função das células de V1 é produzir uma representação eficiente de estímulos visuais. De fato, a codificação eficiente de imagens naturais gera filtros similares aos campos receptivos das células simples de V1 [DEANGELIS et al(1993)]. Tais campos receptivos são caracterizados por serem similares às wavelets de Gabor.

No entanto, imagens naturais e urbanas possuem diferentes propriedades estatísticas [OLIVA, TORRALBA (2001)]. Como exemplo, o espectro de potência pode ser utilizado para classificar uma imagem como natural ou urbana [TORRALBA, OLIVA (2003)]. Isto pode implicar que a codificação eficiente de imagens urbanas podem gerar filtros não semelhantes às wavelets de Gabor.

Neste trabalho propomos analisar a codificação eficiente de imagens urbanas. Para o processo de codificação, utilizaremos a análise de componentes independentes. Dessa forma, a partir de exemplos de imagens urbanas, filtros especializados serão gerados. Como avaliação, vamos quantificar, em termos de erro médio quadrático, a fidelidade da representação gerada por esses filtros.

1.1 Organização do trabalho

Este trabalho está organizado da seguinte forma:

No **capítulo 2**, são apresentadas características funcionais e fisiológicas do córtex visual, em especial V1.

O **capítulo 3** é dedicado à análise estatística de imagens naturais e urbanas.

O modelo utilizado para análise da codificação eficiente de imagens naturais e urbanas é apresentado no **capítulo 4**.

Os resultados obtidos são descritos e discutidos no **capítulo 5** e **capítulo 6**, respectivamente.

Os problemas em aberto e os trabalhos futuros são encontrados na conclusão do trabalho no **capítulo 7**.

2 Córtex visual

O córtex visual é localizado no lóbulo occipital, na parte posterior do cérebro de acordo com a Figura 2.1.a. O córtex visual primário (V1) e áreas de alta ordem V2, V3 e V4 são indicadas pelo conjunto de setas na figura.

O córtex visual possui uma organização hierárquica que realiza análise dos estímulos visuais em diferentes níveis de processamento ainda não completamente compreendidos. Esta organização é representada na Figura 2.1.b em conjunto com o fluxo de informação bi-direcional do córtex visual. O

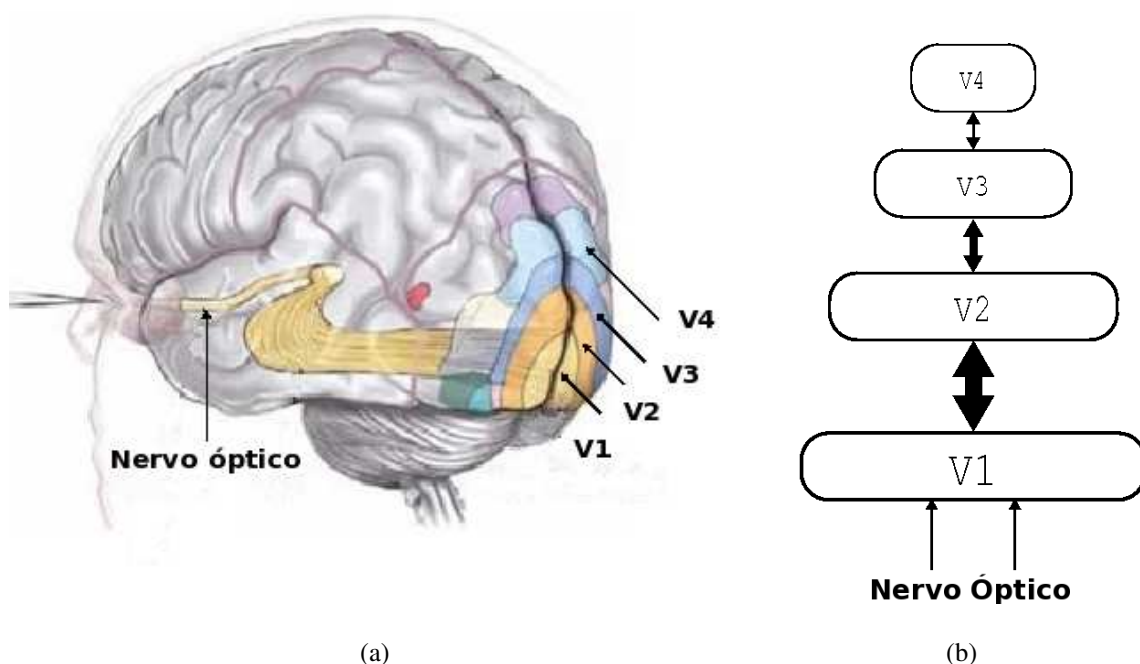


Figure 2.1: Córtex Visual. (a) As setas indicam as áreas V1, V2, V3 e V4 do córtex visual. (b) Organização hierárquica e fluxo de informação bi-direcional do córtex visual primário.

2.1 Córtex visual primário (V1)

O padrão de resposta das células do córtex visual primário está relacionado à parâmetros de estímulos visuais, tais como, orientação de contraste, frequência espacial,

posição, fase, direção de movimento, etc [HUBEL, WIESEL (1962)].

A partir desses parâmetros, foi possível mapear características ou estruturas em estímulos visuais que inibem ou excitam uma determinada célula do cortex visual. O mapa de inibição/excitação resultante é definido como o campo receptivo de uma célula do córtex.

Inicialmente, o processo de mapeamento dos campos receptivos foi realizado analisando a resposta das células de V1 à estimulação visual através de barras de luz orientadas [HUBEL, WIESEL (1962)]. Iluminando certas regiões do espaço foi possível observar o aumento (excitação) ou decaimento (inibição) da taxa de ativação das células estudadas. Mudando os parâmetros (orientação, posição, etc) associados com as barras de luz, foi possível definir a estrutura dos campos receptivos.

A partir dessa estrutura, células de V1 foram caracterizadas como simples ou complexas. Os campos receptivos simples podem ser definidos por possuírem regiões discretas de inibição e excitação como também por serem altamente seletivos aos parâmetros anteriormente citados. Na Figura 2.2.a são mostrados exemplos de campos receptivos de células simples mapeados a partir de estimulação visual utilizando barras de luz orientadas.

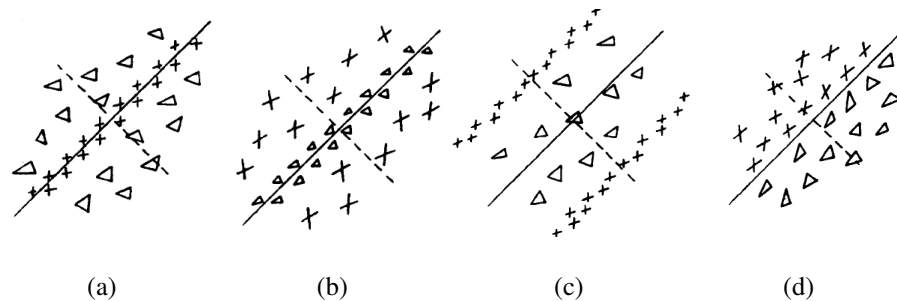


Figure 2.2: Exemplos de campos receptivos de células simples de V1. Os marcadores χ e Δ representam áreas que evocam respostas excitatórias e inibitórias, respectivamente [HUBEL, WIESEL (1962)].

Na Figura 2.2.a, o campo receptivo representado possui uma área de excitação central acompanhada por áreas inibitórias em cada lado. Isto significa que estímulo posicionado e orientado de acordo com esse campo receptivo irá aumentar ou inibir a atividade da célula correspondente. Um campo receptivo com estrutura complementar é mostrado na Figura 2.2.b. O campo receptivo da Figura 2.2.c possui áreas com dimensões diferentes em relação aos dois anteriormente citados. Interessantemente, o campo receptivo mostrado na Figura 2.2.d apresenta apenas duas regiões, uma excitatória e outra inibitória. Para este campo receptivo, o

estímulo ótimo é uma borda ou duas regiões em diferentes níveis de brilho.

Uma forma alternativa para mapeamento dos campos receptivos é o método da correlação reversa. O objetivo deste método é estimar a “resposta ao impulso” espaço-temporal $h(x, y, t)$ de uma célula do córtex. Primeiramente, o processo de mapeamento consiste mensurar e gravar a resposta $y(t)$ de um neurônio a um estímulo momentâneo $s(x, y, t)$ apresentado repetidamente. A correlação cruzada

$$C(x, y, t) = \int_0^T y(\tau) s(x, y, \tau - t) d\tau, \quad (2.1)$$

entre a resposta da célula e a sequência utilizada como estímulo visual é utilizada como uma relação “entrada-saída” do neurônio analisado. Considerando que o estímulo $s(x, y, t)$ é aproximado como ruído branco, a função de correlação cruzada $C(x, y, t)$ é igual à resposta ao impulso $h(x, y, t)$ da célula analisada. A resposta ao impulso é então utilizada como uma estimação espaço-temporal do campo receptivo. Como exemplo, a estrutura espacial de dois campos receptivos estimada em instantes t específicos é mostrada na Figura 2.3.

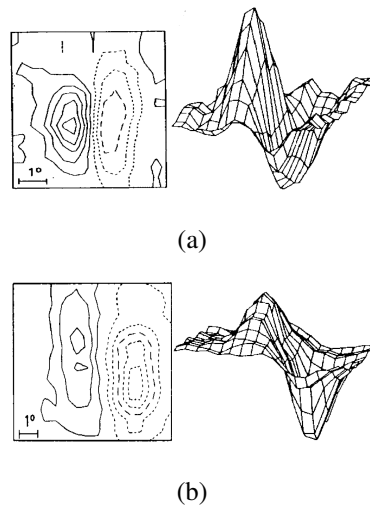


Figure 2.3: Campos receptivos mapeados através de correlação reversa. Os contornos sólidos representam áreas excitatórias enquanto que as regiões com contornos tracejados são inibitórias. A representação tridimensional dos campos receptivos é mostrada ao lado [DEANGELIS et al(1993)].

Uma aproximação matemática da estrutura espacial mostrada na Figura 2.3 pode ser realizada através de wavelets de Gabor [MALLAT (1989)]. De acordo com a Eq. 2.2, uma wavelet de Gabor $G(x, y)$ pode ser definida como um produto entre uma função Gaussiana e

uma função senoidal.

$$G(x, y) = \frac{1}{2\pi\sigma_x\sigma_y} \exp\left(-\frac{x^2}{2\sigma_x^2} - \frac{y^2}{2\sigma_y^2}\right) \cos(kx - \phi). \quad (2.2)$$

Os parâmetros de $G(x, y)$ determinam as propriedades da estrutura espacial de um campo receptivo. As extensões do campo receptivo nas direções x e y são determinadas por σ_x e σ_y , respectivamente. O parâmetro k determina a frequência espacial preferida, ou seja, a forma das oscilações entre áreas excitatórias e inibitórias. A fase das oscilações é determinada por ϕ . Exemplos de campos receptivos aproximados a partir de wavelets de Gabor são mostrados na Figura 2.4.

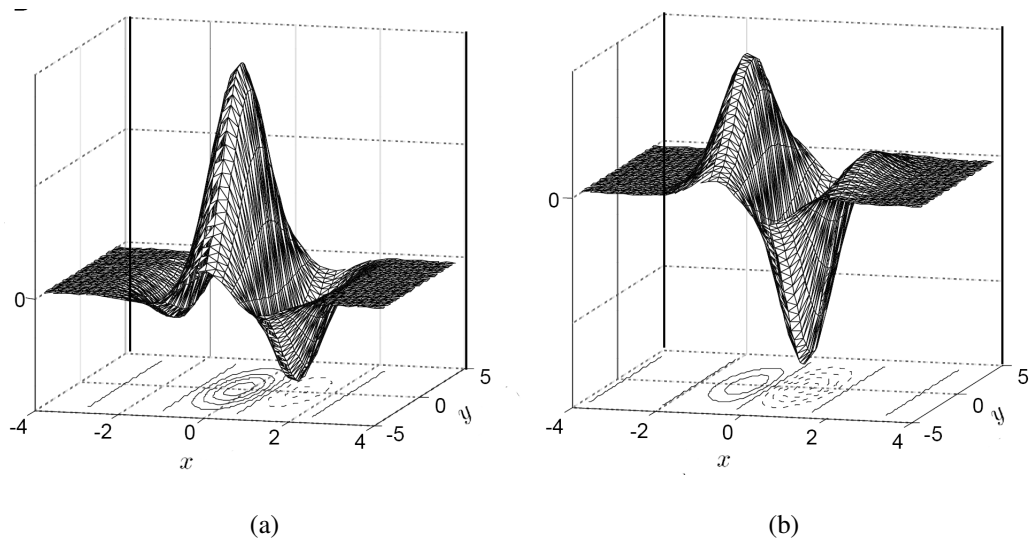


Figure 2.4: Aproximação de campos receptivos através de wavelets de Gabor. Os parâmetros associados são $\sigma_x = 1$, $\sigma_y = 2$, $1/k = 0.56$ e (a) $\phi = 1 - \pi/2$ ou (b) $\phi = 1 - \pi$ [DAYAN, ABBOTT (2001)]

Os parâmetros dos campos receptivos das células simples de V1 variam ao longo da população. Dessa forma, cada célula é específica, respondendo apenas aos estímulos visuais que correspondam aos seus respectivos parâmetros.

Por exemplo, é possível observar que todos os campos receptivos mostrados na Figura 2.2 apresentam orientação oblíqua. Assim, estímulos em mesma orientação deverão excitar ou inibir a ativação das células, enquanto que estímulos em diferentes orientações não deverão influenciar o padrão de ativação. Por outro lado, se estes estímulos possuírem uma outra característica que corresponda aos campos receptivos da célula, então eles deverão influenciar o padrão de ativação, indiferentemente à orientação.

Dessa forma, as células de V1 funcionam como um detector de características sendo ativadas por estruturas que correspondam à informação representada pelo seus respectivos campos receptivos.

Este processo é representado na Figura 2.5.

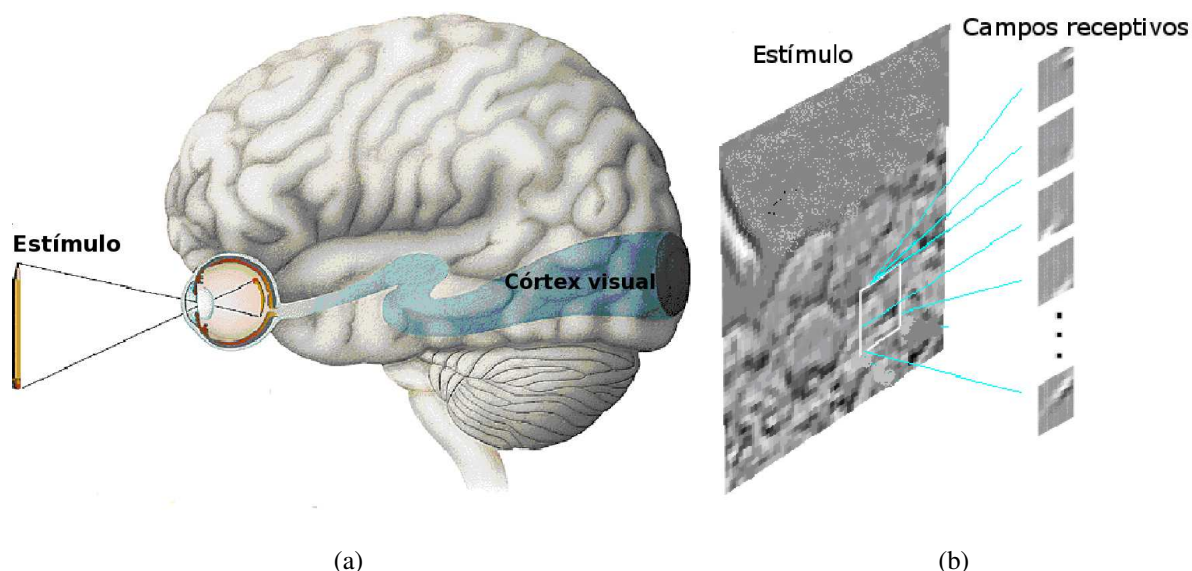


Figure 2.5: Representação de estímulos no córtex visual. (a) Primeiramente, um estímulo visual ativa ou inibi as células da retina. Tal informação é transmitida ao córtex visual primário através do nervo óptico. (b) Em V1, os campos receptivos que correspondem as características do estímulo são utilizados para geração de uma representação neural.

Um dos problemas mais antigos em neurociência visual é entender os princípios e o objetivo dessa representação neural criada pelos campos receptivos de V1. Uma hipótese é que células sensoriais tem como objetivo extrair a maior quantidade de informação do ambiente para sobrevivência do organismo. No entanto para isso, as células devem gerar uma representação não redundante ou eficiente do ambiente.

3 Codificação eficiente ou esparsa

Supõe-se que uma das estratégias utilizadas pelo cérebro para representar informação seja o princípio da codificação eficiente [BARLOW (1961)]. Este conceito foi proposto por Horace Barlow, em 1961, como um modelo teórico para a codificação das informações sensoriais pelo sistema nervoso. Para ele, um modelo eficiente seria aquele que minimizasse a quantidade de impulsos nervosos utilizados para transmitir a informação desejada.

No desenvolvimento de sua teoria, Barlow foi inspirado por conceitos da Teoria da Informação. Ele definiu que os caminhos neurais percorridos por informações sensoriais são similares a canais de comunicação. Através de conceitos como capacidade de canal e redundância, Barlow sugeriu que a codificação neural é realizada de forma a maximizar a capacidade de canal e reduzir a redundância na informação transmitida.

Biologicamente, o conceito de codificação eficiente pode ser descrito em duas perspectivas: para uma única célula e para múltiplas células. Na primeira, é considerado que o padrão de resposta de um neurônio deve ser esparsamente distribuído [OLSHAUSEN, FIELD (2004)]. Uma distribuição de probabilidade esparsa pode ser caracterizada por possuir um pico em zero e longas extremidades. Uma curva que pode caracterizar a distribuição de resposta de um neurônio é mostrada na Figura 3.1. As longas extremidades fazem com que as probabilidades de resposta da célula sejam pequenas. Dessa forma, para um dado estímulo, é provável que apenas um pequeno conjunto de células respondam.

De fato, é mostrado [BADDELEY et al (1998)] que células do córtex visual primário de primatas apresentam distribuições esparsas quando estimuladas com seqüências de imagens. Este comportamento também se estende ao córtex auditivo primário (A1) [DEWEESE et al (2003)]. As células de A1 podem produzir um único disparo em resposta a um estímulo sonoro. Este comportamento foi caracterizado como uma “codificação binária” porque tais células produzem “0 ou 1” em resposta a um estímulo. No entanto, a probabilidade de disparo é muito pequena ao longo do tempo, o que é consistente com

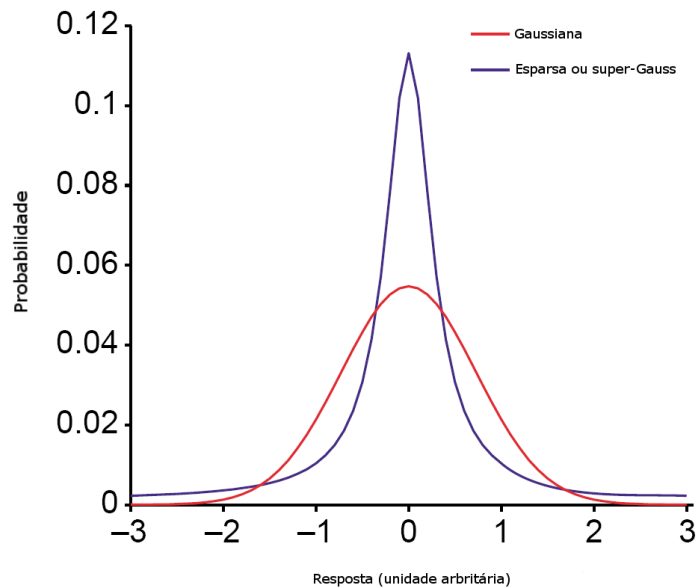


Figure 3.1: Comparação entre a distribuição esparsa (azul) e Gaussiana (vermelha). Uma distribuição esparsa é caracterizada por possuir um pico em zero e longas extremidades.

a idéia de esparsifisidade. Pequenas probabilidades de disparo também são observadas em células do córtex somatosensory de roedores [BRECHT et al (2002)]. Além disso, neurônios do sistema olfativo de insetos disparam somente até duas vezes para odores utilizados como estímulos [PEREZ-ORIVE et al (2002)]. Dessa forma, sparsifisidade parece ser um dos princípios utilizados na codificação neural.

Na segunda perspectiva, a representação de informação sensorial é eficiente caso as respostas neurais sejam estatisticamente independentes. Dessa forma, não existe “informação redundante” entre células da população.

De fato, células de V1 emitem fortes respostas quando estimuladas com estruturas não redundantes, tais como, barras e bordas [HUBEL, WIESEL (1962)]. Isto sugere que estímulos visuais devam ser representados de forma eficiente no córtex visual primário.

O conceito de codificação eficiente também pode ser definido a partir de um modelo matemático para geração de imagens. Neste modelo, uma imagem $\hat{x}$ é gerada a partir de um conjunto de filtros $\mathbf{A} = [\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n]$ ativados independentemente por coeficientes $\mathbf{s} = [s_1, s_2, \dots, s_n]$. O modelo matemático é dado por

$$\hat{x} = \sum_{i=1}^n \mathbf{a}_i s_i. \quad (3.1)$$

[OLSHAUSEN, FIELD (1996)] mostrou que utilizando imagens naturais, este modelo é capaz de gerar filtros $\mathbf{A}$ similares aos campos receptivos das células simples de V1. A Figura 3.2 mostra os filtros gerados.

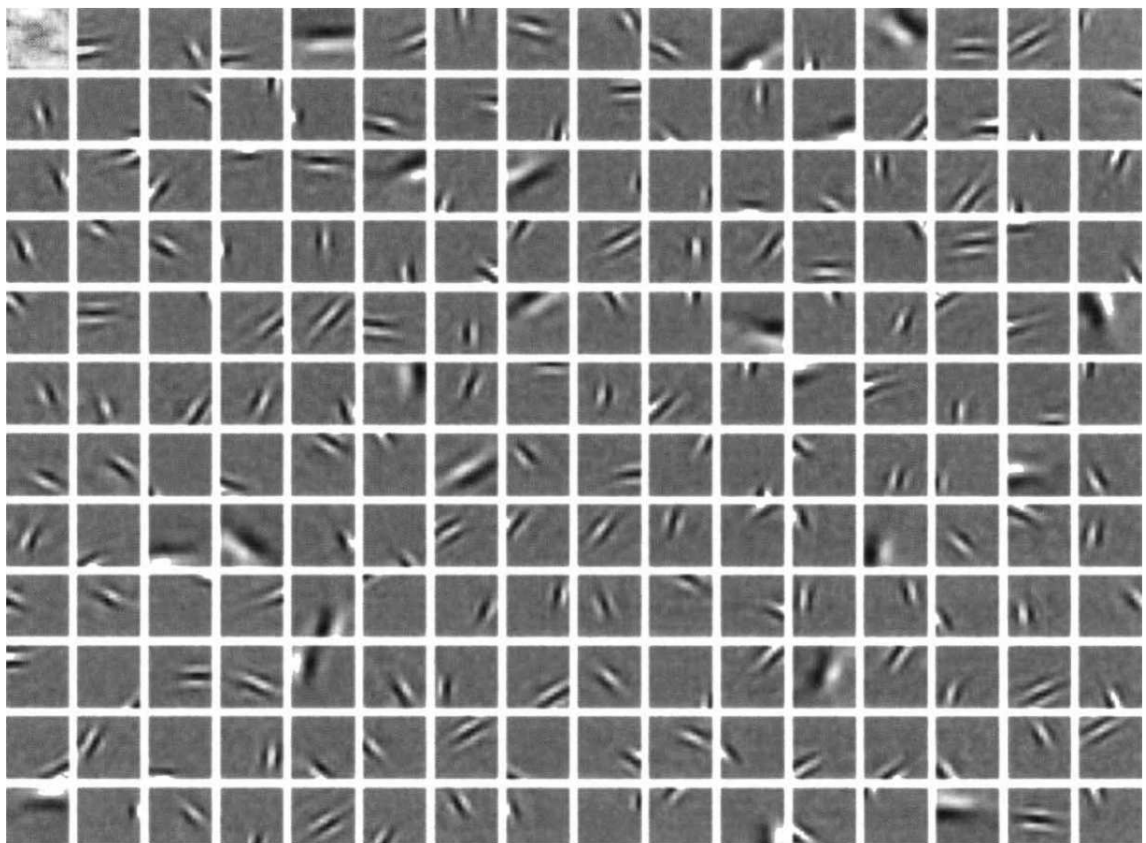


Figure 3.2: Filtros gerados utilizando codificação eficiente [OLSHAUSEN, FIELD (1996)].

É possível observar que os filtros mostrados na Figura 3.2 apresentam características similares às wavelets de Gabor que correspondem aos campos receptivos de V1. Na Figura 3.3 são mostrados outros exemplos de filtros gerados por codificação eficiente. No entanto, estes são apresentados como uma topologia de acordo com suas orientações.

A partir desses resultados, foi proposto [OLSHAUSEN, FIELD (1996)] que os campos receptivos das células simples de V1 são gerados a partir de um processo neural baseado em codificação eficiente.

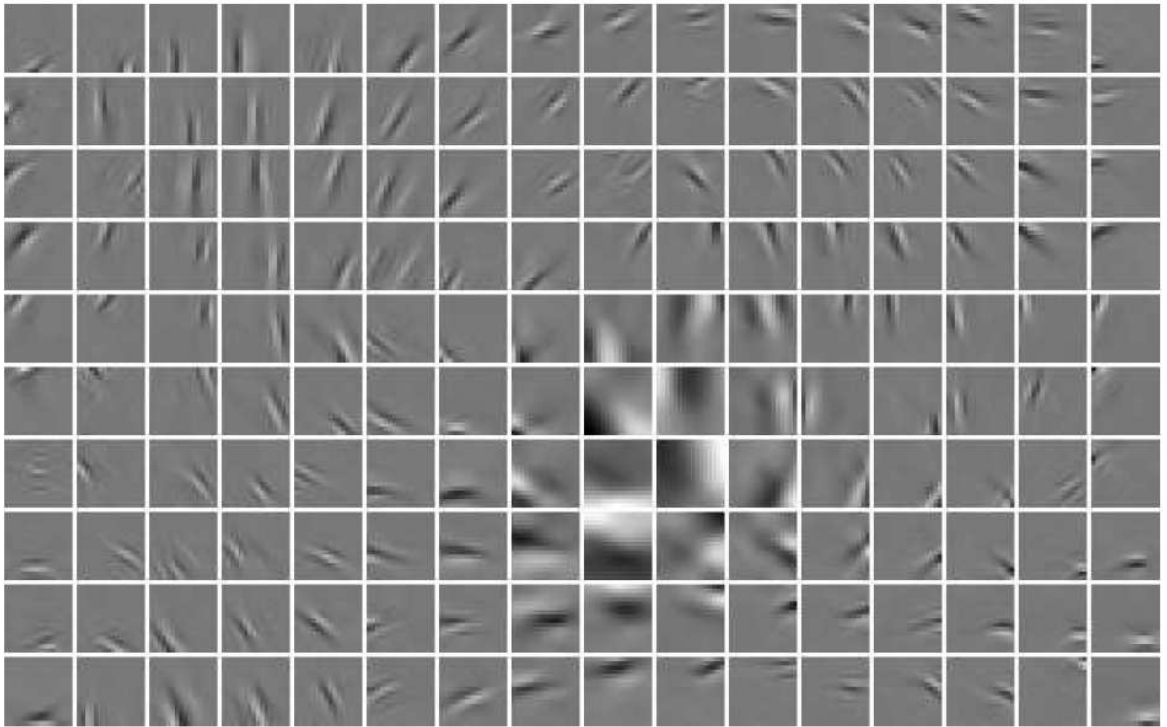


Figure 3.3: Filtros organizados a partir de suas orientações [HYVÄRINEN et al (2001)]. A organização topográfica é gerada a partir das correlações.

No entanto, é importante citar que campos receptivos similares às wavelets de Gabor são adaptados às imagens naturais. Para um conjunto de imagens com características estatísticas diferentes, os campos receptivos gerados por codificação eficiente podem ser diferentes.

4 Análise de componentes independentes

A análise de componentes independentes (ICA, do inglês *independent component analysis*) é uma técnica baseada no modelo de codificação eficiente. ICA foi introduzida como solução para o problema de separação cega de fontes [COMON (1994)]. A suposição básica no modelo de ICA é que as fontes devem ser mutuamente estatisticamente independentes. Dessa forma, para entender o modelo de ICA, é necessário primeiramente definir independência estatística.

Por motivos de simplicidade, iremos nos referir à independência estatística apenas como independência.

4.1 O que é independência?

Duas variáveis aleatórias s_i e s_j são independentes se a partir de s_i não é possível estimar ou inferir algum valor ou informação de s_j . Exemplos bem conhecidos de variáveis aleatórias independentes são um sinal de fala e o ruído de fundo originado a partir de um sistema de ventilação.

Matematicamente, s_i e s_j são independentes se e somente se

$$p_{s_i, s_j}(s_i, s_j) = p_{s_i}(s_i)p_{s_j}(s_j). \quad (4.1)$$

Dessa forma, a probabilidade conjunta de duas variáveis independentes pode ser calculada como apenas o produto das marginais.

4.2 Restrições, limitações e ambiguidades de ICA

Além do conceito de independência, é necessário definir as restrições e problemas associados com o modelo de ICA.

Primeiramente, ICA não pode ser aplicado quando as variáveis aleatórias têm distribuições Gaussianas. Isto é devido ao fato que a distribuição conjunta de misturas dessas variáveis também é Gaussiana. Dessa forma, como esta distribuição é rotacionalmente simétrica, a informação da rotação da mistura é perdida.

Outros dois problemas associados com o modelo de ICA são:

- A informação de variância das componentes independentes é perdida no processo de estimação;
- Não se pode estabelecer uma ordem para as componentes independentes.

4.3 Modelo de ICA

O modelo de ICA é similar ao modelo de codificação eficiente da Equação 3.1. No entanto, generalizando, vamos considerar que $\mathbf{x} = [x_1, x_2, \dots, x_n]$ e $\mathbf{s} = [s_1, s_2, \dots, s_n]$ são vetores aleatórios tal que cada elemento x_i é uma mistura dos elementos de $\mathbf{s}$.

É importante citar que, além de independentes, as componentes de $\mathbf{s}$ são latentes. Isto significa que estas variáveis não podem ser observadas diretamente.

Dessa forma, utilizando uma notação matricial, temos que

$$\mathbf{x} = \mathbf{A}\mathbf{s}, \quad (4.2)$$

onde $\mathbf{A} = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}$ é a matriz de mistura.

4.4 ICA por maximização de não-gaussianidade

O problema central de ICA é estimar as componentes latentes s_i a partir de $\mathbf{x}$, utilizando a suposição básica de independência. Como será discutido, um conceito chave neste processo de estimação é o de não-gaussianidade.

O teorema do limite central implica que a distribuição de soma de componentes independentes é sempre mais próxima de uma gaussiana do que qualquer uma das distribuições das componentes isoladas.

Assim, para estimar uma componente independente s_i , vamos considerar a seguinte soma ou combinação linear dos vetores x_i ,

$$y = \mathbf{b}^T \mathbf{x}, \quad (4.3)$$

onde $\mathbf{b}$ é um vetor a ser determinado. É fácil ver que se $\mathbf{b}$ for uma das linhas da inversa de $\mathbf{A}$, y será igual a uma das componentes independentes s_i .

Agora, vamos denotar o vetor $\mathbf{b}^T \mathbf{A}$ apenas como $\mathbf{q}$. Assim temos que

$$y = \mathbf{b}^T \mathbf{x} = \mathbf{q}^T \mathbf{s} = \sum_i q_i s_i. \quad (4.4)$$

A partir da Equação 4.4, é possível observar que y também é uma soma das componentes independentes s_i . Assim, pelo teorema do limite central podemos concluir que a distribuição de y é mais gaussiana do que a distribuição de qualquer s_i . Dessa forma, quando a distribuição de y for maximamente não-gaussiana, então

$$y \approx s_i. \quad (4.5)$$

No entanto, existe dois problemas a serem resolvidos:

- Como medir a não-gaussianidade de y ?
- Como calcular o vetor $\mathbf{b}$ que maximiza essa medida de não-gaussianidade?

4.5 Curtose como medida de não-gaussianidade

Curtose ou cumulante de quarta-ordem é uma medida clássica de Gaussianidade. A curtose de uma variável y é definida como

$$curt(y) = E\{y^4\} - 3(E\{y^2\})^2. \quad (4.6)$$

É possível observar que a curtose é uma versão normalizada do quarto momento $E\{y^4\}$. Se y é gaussiana, o quarto momento é igual a $3(E\{y^2\})^2$. Assim, a curtose é zero para variáveis gaussianas.

Dessa forma, podemos medir o grau de não-gaussianidade de uma variável y a partir da distância do valor absoluto de sua curtose para zero.

4.6 Maximização de não-gaussianidade pelo método do gradiente

O processo de maximização de não-gaussianidade pode ser realizado através do método do gradiente. Este método consiste em “direcionar” o vetor $\mathbf{b}$ na direção de maior crescimento do valor absoluto da curtose de y .

Dessa forma, a direção de $\mathbf{b}$ deve corresponder a direção do vetor gradiente da curtose de y . Este vetor é definido como

$$\frac{\partial |\text{curt}(\mathbf{b}^T \mathbf{x})|}{\partial \mathbf{b}} = 4 \text{sign}(\text{curt}(\mathbf{b}^T \mathbf{x})) [E\{\mathbf{x}(\mathbf{b}^T \mathbf{x})^3\} - 3\mathbf{b}\|\mathbf{b}\|^2]. \quad (4.7)$$

A partir da Equação 4.7 é possível obter um algoritmo gradiente para estimação do vetor $\mathbf{b}$. Este algoritmo é dado por

$$\Delta \mathbf{b} \propto \frac{\partial |\text{curt}(\mathbf{b}^T \mathbf{x})|}{\partial \mathbf{b}}. \quad (4.8)$$

5 Imagens naturais e urbanas

Uma vez que os campos receptivos simples de V1 são similares às wavelets de Gabor, foi proposto que as células do córtex foram adaptadas às imagens naturais [SIMONCELLI, OLSHAUSEN (2001)]. No entanto, imagens naturais e urbanas tem características estatísticas distintas, o que pode implicar na formação de campos receptivos diferentes de wavelets de Gabor.

Primeiramente, vamos analisar o espectro de potência das duas classes de imagens. Na figura 5.1 é mostrado a assinatura espectral média para grupos de imagens naturais e urbanas.

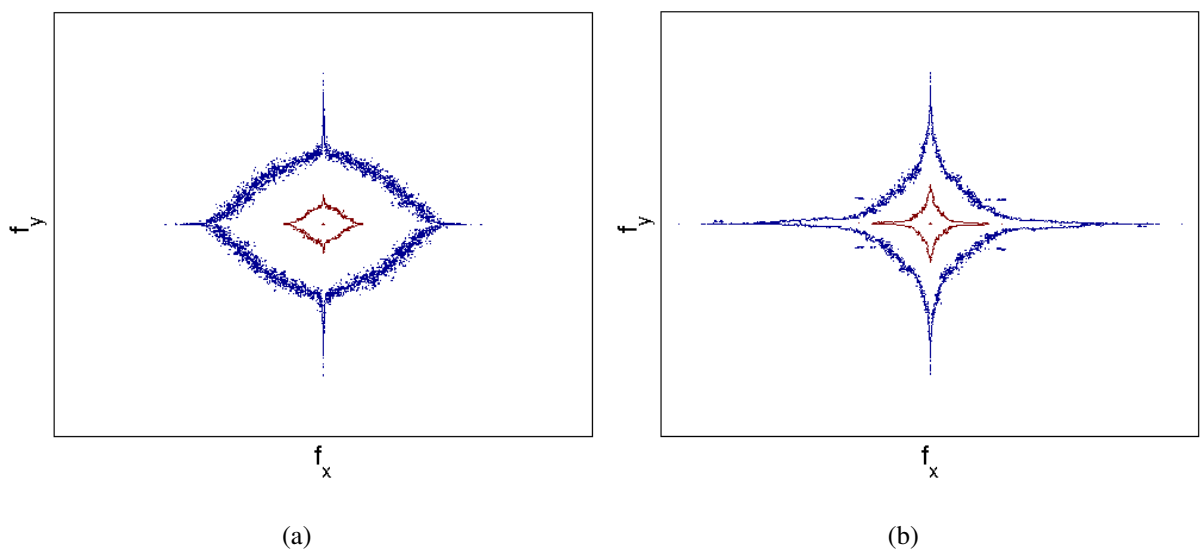


Figure 5.1: Assinaturas espectrais médias de imagens (a) naturais e (b) urbanas. Cada grupo utilizado consiste de 40 imagens. O grupo de imagens naturais inclui cenas distantes e próximas de florestas, campos, montanhas e objetos naturais como árvores e flores. O grupo de imagens urbanas inclui cenas em cidades sem restrição à presença de vegetação. Os contornos de cada assinatura correspondem a 65 % e 75 % da energia total no espectro de potência.

O formato de cruz da assinatura de imagens urbanas indica que orientações verticais e horizontais dominam fortemente o espectro de potência enquanto que orientações oblíquas possuem baixa energia. Por outro lado, a assinatura natural é mais isotrópica, ou seja, todas as orientações exibem quase a mesma energia, embora orientações verticais e horizontais sejam

levemente mais poderosas.

As características dos espectros de potência dessas duas classes também podem ser observadas para imagens isoladas. Na figura abaixo são mostradas imagens naturais e urbanas e suas respectivas assinaturas espectrais.

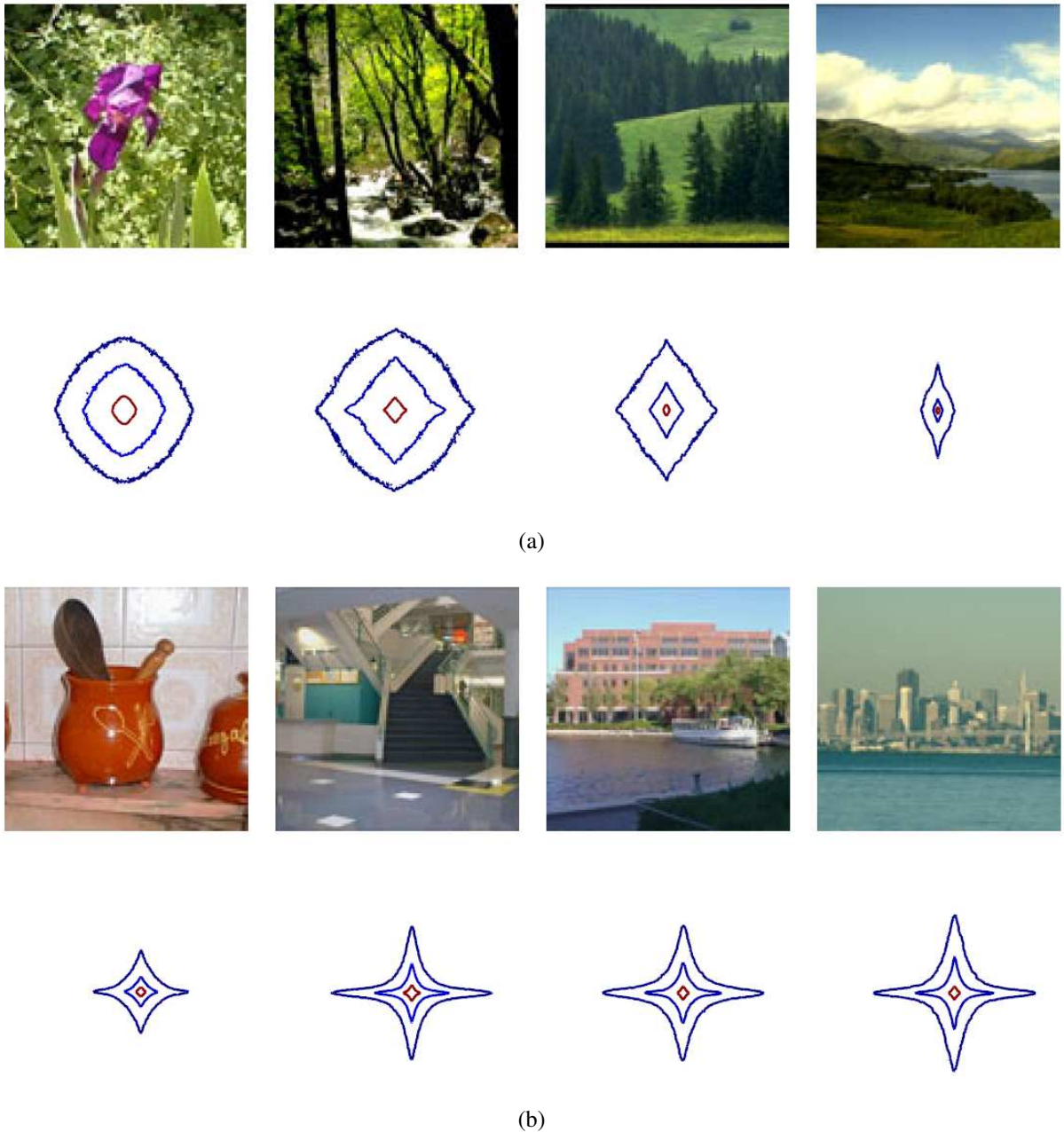


Figure 5.2: Assinaturas espectrais de imagens (a) naturais e (b) urbanas analisadas isoladamente [TORRALBA, OLIVA (2003)].

Os espectros de potência completos de imagens naturais e urbanas podem ser observados na Figura 5.3.

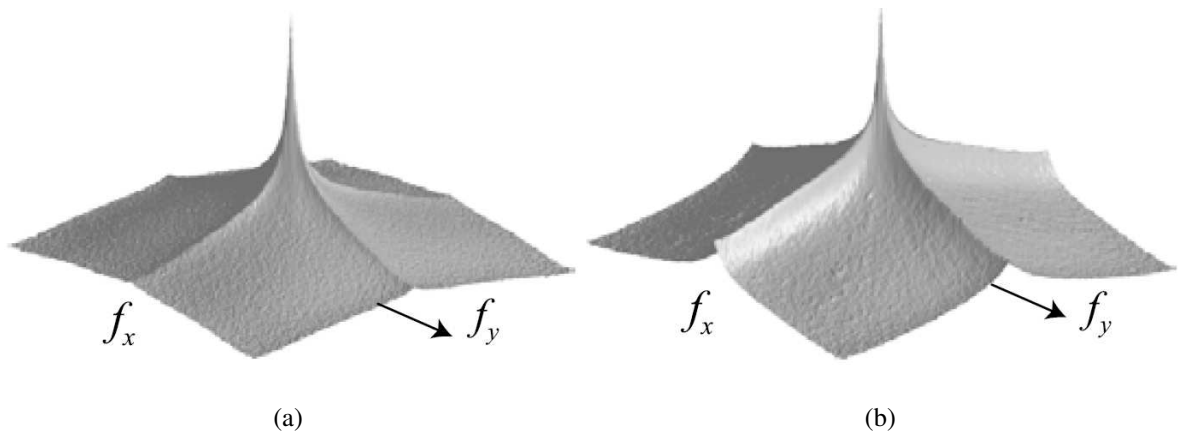


Figure 5.3: Espectro de potência médio de imagens (a) naturais e (b) urbanas. 6000 imagens foram utilizados para cada grupo [TORRALBA, OLIVA (2003)].

As características destes espectros de potência podem ser devido a forte presença de bordas fractais em imagens naturais e bordas coerentes em imagens urbanas. Estas duas estruturas são representadas na Figura 5.4.

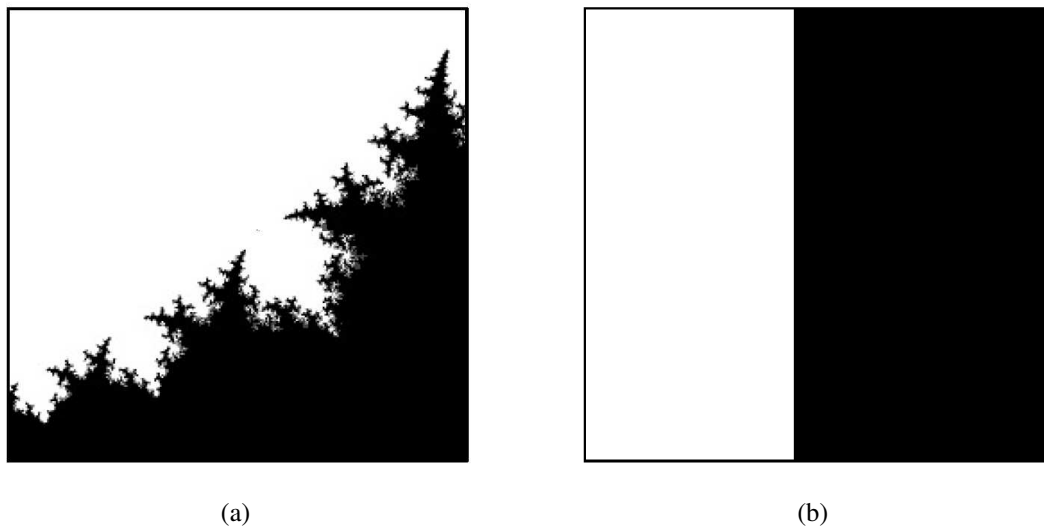


Figure 5.4: Representação de uma borda (a) fractal e (b) coerente.

Uma borda fractal exibe continuidade ao longo do espaço e frequência [FIELD (1994)]. Isto significa que a orientação da borda muda ao longo de seu comprimento e frequência. De fato, em uma única borda fractal, várias orientações podem ser encontradas. Uma borda coerente deve, no entanto, apresentar uma única orientação. É reportado [VAILAYA (1998)] que bordas em imagens urbanas apresentam um significativo grau de coerência enquanto bordas em cenas naturais não são coerentes.

A presença dessas estruturas também implicam que imagens naturais e urbanas devem exibir informações diferentes em seus espectros de fase. De fato, bordas fractais e coerentes possuem diferentes características de fase [OLSHAUSEN, FIELD (1996)]. Enquanto bordas fractais exibem somente alinhamentos locais ao longo da frequência espacial, uma borda coerente produz um alinhamento global no espectro de fase. Na Figura 5.5, este comportamento é representado para uma borda coerente de uma dimensão.

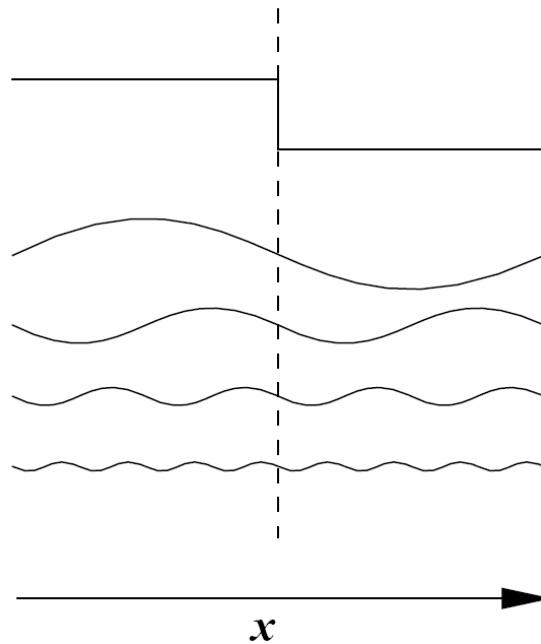


Figure 5.5: Característica do alinhamento em fase para uma borda coerente de uma dimensão [OLSHAUSEN, FIELD (1996)]. As componentes de frequência devem estar todas em fase para formação de uma borda coerente.

Características estatísticas distintas, podem implicar em diferenças no processo codificação de imagens naturais e urbanas. Primeiramente, vamos reconsiderar um modelo baseado em codificação eficiente, definido na Equação 5.1.

$$\mathbf{I}(x, y) = \sum_{i=1}^n \mathbf{a}_i(x, y) s_i. \quad (5.1)$$

Agora, vamos considerar o espectro de potência de ambos os lados da Equação 5.1. Uma vez que $\mathbf{s}$ é esparsamente distribuído, seu espectro de potência é constante. Então, pode-se concluir que a imagem $\mathbf{I}(x, y)$ tem uma distribuição espectral proporcional aos filtros $\mathbf{A}$.

Então, vamos considerar que $\mathbf{A}_{\text{nat}}$ e $\mathbf{A}_{\text{urb}}$ são filtros gerados a partir da codificação eficiente de imagens naturais e urbanas, respectivamente. Como estes grupos de imagens têm espectros de potência diferentes, concluí-se que $\mathbf{A}_{\text{nat}}$ e $\mathbf{A}_{\text{urb}}$ também devem ser diferentes.

6 Métodos

Para analisar o processo de codificação eficiente de imagens naturais e urbanas, utilizamos o modelo representado na Figura 6.1 que permite o aprendizado e a avaliação de filtros baseados em codificação eficiente.

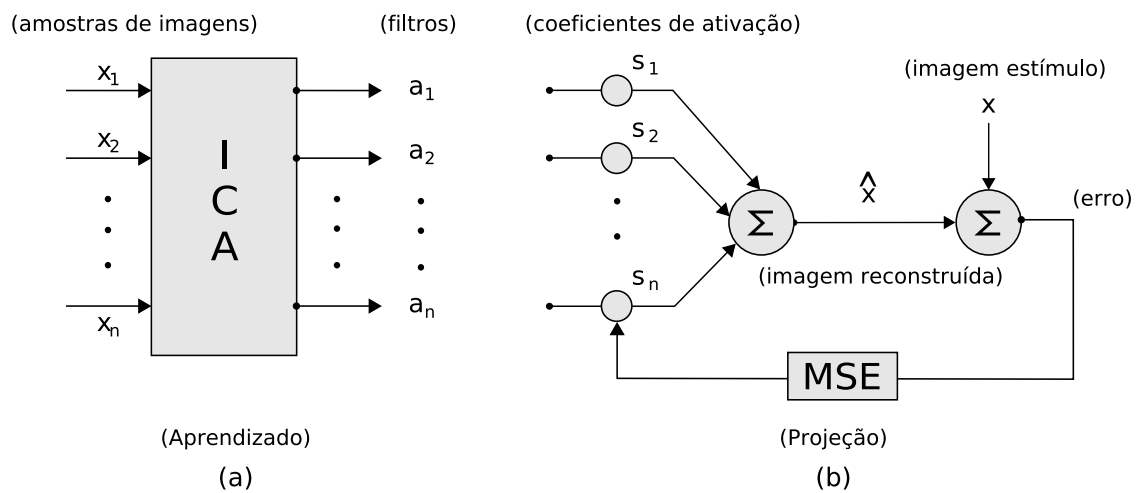


Figure 6.1: Diagrama de blocos do modelo utilizado. O sistema consiste de duas fases: (a) *aprendizado* e (b) *projeção*. Na fase de aprendizado, nós utilizamos ICA para aprender filtros a partir de grupos de imagens naturais e urbanas. Na fase de projeção, foi calculado o erro médio quadrático (MSE) gerado pela representação ou projeção de imagens naturais e urbanas nos filtros gerados.

O modelo proposto consiste de duas fases: *aprendizado* e *projeção*. Na primeira fase, utilizamos ICA para gerar filtros A_{nat} e A_{urb} especializados em imagens naturais e urbanas, respectivamente. Na fase de projeção, estes filtros são avaliados a partir do erro médio quadrático (MSE do inglês *mean square error*) gerado pela projeção de imagens naturais e urbanas.

As seções seguintes seções fornecem uma descrição detalhada de cada fase do modelo proposto.

6.1 Aprendizado

Os modelos de codificação eficiente e ICA são ambos baseados em redução de redundância. Dessa forma, os filtros estimados por ICA também podem ser utilizados em um modelo eficiente.

Neste trabalho utilizamos o algoritmo FastICA [HYVÄRINEN (1999)]. Este algoritmo é baseado em interação de ponto-fixo derivada a partir do algoritmo gradiente. A regra de ponto fixo consiste em assumir que os vetores $\mathbf{b}$ são iguais ao vetor gradiente. Dessa forma, no caso da curtose, temos que

$$\mathbf{b} \leftarrow \frac{\partial |\text{curt}(\mathbf{b}^T \mathbf{x})|}{\partial \mathbf{b}}. \quad (6.1)$$

No entanto, na versão do FastICA utilizada neste trabalho, a curtose foi substituída pela sua versão mais robusta, a negentropia [HYVÄRINEN, KARHUNEN, OJA (2001)]. O algoritmo FastICA pode ser resumido nos seguintes passos:

1. Centralizar $\mathbf{x}$, ou $\mathbf{x} = \mathbf{x} - E\{\mathbf{x}\}$;
2. Branquear $\mathbf{x}$, ou $\mathbf{x} = \mathbf{V}\mathbf{x}$, onde $\mathbf{V}$ é matriz de braqueamento.
3. Escolher n , o número de componentes independentes a estimar;
4. Inicializar a matrix $\mathbf{B} = [\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n]^T$. Normalizar os vetores $\mathbf{b}_i$;
5. Ortogonalizar a matriz $\mathbf{B}$ como no passo 7 abaixo;
6. Para cada $i = 1, \dots, n$, seja

$$\mathbf{b}_i \leftarrow E\{\mathbf{x}g(\mathbf{b}_i^T \mathbf{x})\} - E\{g'(\mathbf{b}_i^T \mathbf{x})\}\mathbf{b}_i,$$

sendo g uma função não-linear;

7. Ortogonalizar simetricamente a matriz $\mathbf{B}$ fazendo

$$\mathbf{B} \leftarrow (\mathbf{B}\mathbf{B}^T)^{-1/2}\mathbf{B}.$$

8. Caso o algoritmo não convirja, repetir o passo 6.

Como os vetores $\mathbf{b}_i$ estimam as componentes independentes no modelo de ICA, podemos afirmar que $\mathbf{B}$ é uma estimação da matrix inversa de $\mathbf{A}$. Dessa forma, podemos gerar os filtros $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n$ do modelo eficiente da Equação 3.1 a partir de ICA.

6.2 Projeção

Nesta fase, os filtros gerados são avaliados. A avaliação é realizada a partir do cálculo do MSE gerado pela projeção de imagens naturais e urbanas nos filtros gerados.

O processo de projeção, representado na Figura 6.1, consiste em estimar os coeficientes de projeção ou ativação para cada filtro a_i . Para isto, utilizaremos uma estimação de erro médio quadrático mínimo.

Primeiramente, vamos definir o erro ϵ entre a imagem original x que foi utilizada como “estímulo” e a imagem $\hat{x}$ reconstruída a partir dos filtros.

$$\epsilon = x - \hat{x} \quad (6.2a)$$

$$= x - \mathbf{A}\mathbf{s} = x - \mathbf{s}^T \mathbf{A}. \quad (6.2b)$$

tomando o MSE ou valor esperado de ϵ^2 , temos que

$$\xi = E[\epsilon^2] = E[(x - \mathbf{s}^T \mathbf{A})^2] \quad (6.3a)$$

$$= E[(x^2 - \mathbf{s}^T \mathbf{A} \mathbf{A}^T \mathbf{s} - 2x \mathbf{A}^T \mathbf{s})^2] \quad (6.3b)$$

$$= E[x^2] - \mathbf{s}^T E[\mathbf{A} \mathbf{A}^T] \mathbf{s} - 2E[x \mathbf{A}^T] \mathbf{s}. \quad (6.3c)$$

Definindo

$$\mathbf{R} = E[\mathbf{A} \mathbf{A}^T] \quad (6.4)$$

$$\mathbf{P} = E[x \mathbf{A}^T], \quad (6.5)$$

sendo $\mathbf{R}$ a matriz de autocorrelação dos filtros $\mathbf{A}$ e $\mathbf{P}$ a matriz de correlação cruzada entre os filtros e a imagem estímulo. Assim, reescrevendo a Equação 6.3c temos que

$$\xi = E[x^2] - \mathbf{s}^T \mathbf{R} \mathbf{s} - 2\mathbf{P}^T \mathbf{s}. \quad (6.6)$$

Tomando o gradiente de ξ , temos que

$$\nabla(\xi) = 2\mathbf{R}\mathbf{s} - 2\mathbf{P}. \quad (6.7)$$

O erro médio quadrático mínimo é obtido quando o gradiente $\nabla(\xi)$ é nulo:

$$\nabla(\xi) = 0 = 2\mathbf{R}\mathbf{s}^* - 2\mathbf{P} \quad (6.8)$$

Para esse valor temos que

$$\mathbf{s}^* = \mathbf{R}^{-1}\mathbf{P}. \quad (6.9)$$

O vetor $\mathbf{s}^*$ corresponde aos coeficientes de ativação que geram o MSE mínimo para a projeção da imagem estímulo x nos filtros $\mathbf{A} = [\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n]$.

Neste trabalho, analisamos o comportamento do MSE a medida que aumentamos o número de filtros $\mathbf{a}_i$ utilizados na projeção. Para ordenar os filtros, utilizamos o seguinte algoritmo:

- **1.** Criar um espaço vazio Φ que irá conter os filtros ordenados a partir $\mathbf{A}$.
- **2.** Fazer $i = 1 : n$
- **4.** Encontrar a imagem reconstruída $\hat{x}_i$ projetando x no espaço formado por $[\Psi, \mathbf{a}_i]$ através da Equação 6.9.
- **5.** Após $i=n$, selecionar o filtro $\mathbf{a}_i$ de acordo com o critério abaixo:

$$\mathbf{a}_i = \operatorname{argmin}(x - \hat{x}_i) \quad (6.10)$$

- **6.** Inserir o filtro $\mathbf{a}_i$ selecionado no espaço Φ :

$$\Phi \leftarrow [\Phi, \mathbf{a}_i].$$

- **7.** Retirar o filtro $\mathbf{a}_i$ de $\mathbf{A}$ fazendo com que $n = n - 1$.
- **8.** Voltar ao passo 2 até $n = 0$.

7 Resultados

7.1 Base de dados de imagens

Neste trabalho, utilizamos imagens naturais e urbanas da base de dados de imagens McGill Calibrated Colour Image Database [OLMOS et al (2004)]. Esta base contém 850 imagens coloridas de 768 x 756 pixels escalonadas a partir de originais de 1920 X 2560 pixels.

Estas imagens estão divididas em classes como florestas, flores, montanhas, cenas urbanas, etc. Todas as imagens foram transformadas para escala de cinza de forma a reduzir o processamento computacional.

7.2 Filtros A_{nat} e A_{urb}

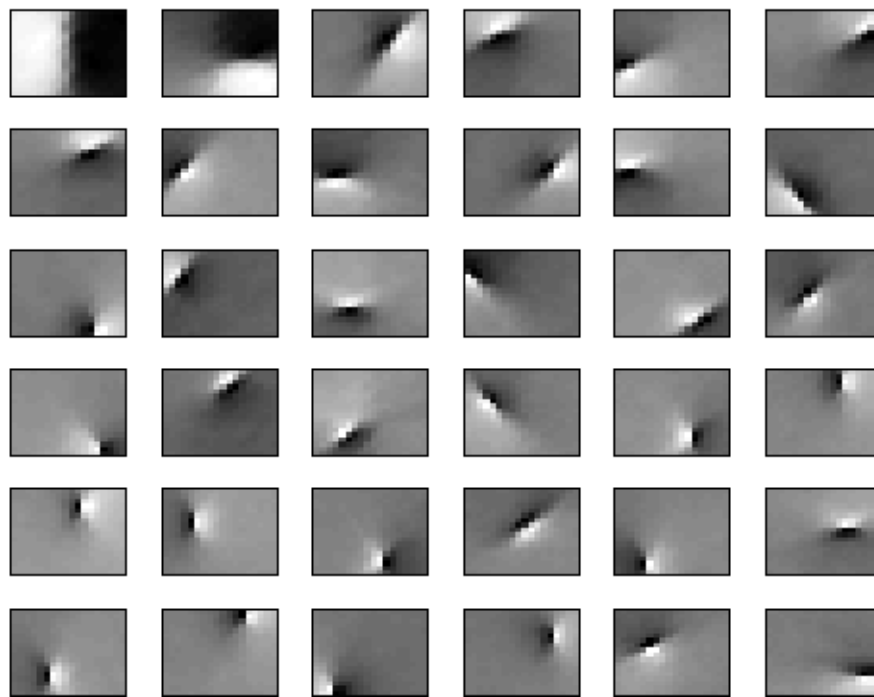
Para aprender filtros A_{nat} e A_{urb} a partir de imagens naturais e urbanas, escolhemos dez imagens de cada grupo. O grupo de imagens naturais incluiu cenas distantes e próximas de florestas, campos, montanhas e objetos naturais como árvores e flores. Novamente, para as imagens urbanas, nenhuma restrição à presença de vegetação foi feita.

Estas imagens foram processadas em blocos de 18 x 18 pixels. Para dimensão das imagens usadas, este tamanho de bloco garante que estruturas como bordas fractais e coerentes nas imagens sejam capturadas.

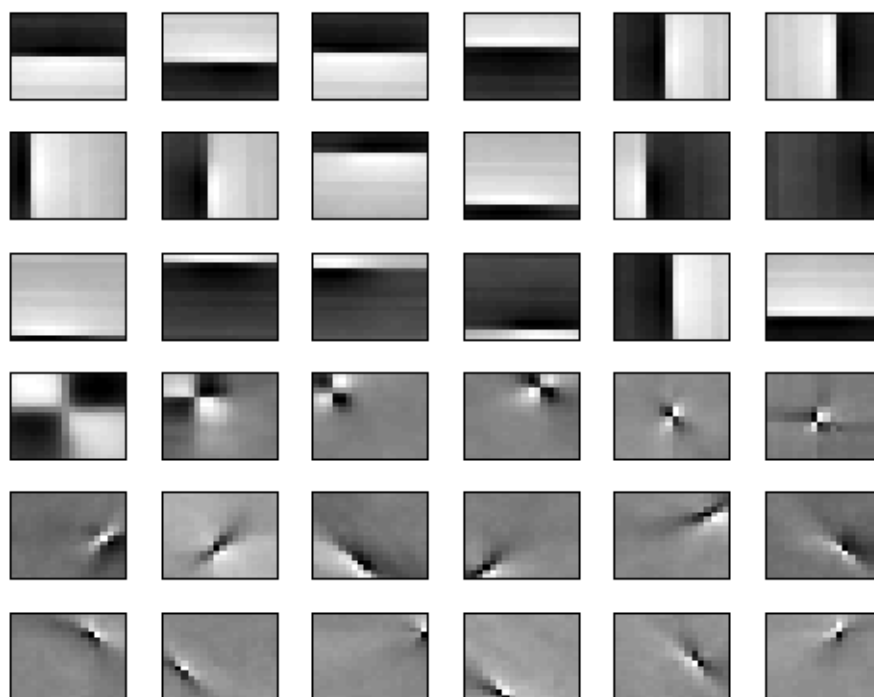
Para formar o conjunto de treinamento para o algoritmo FastICA, para cada bloco, duas amostras foram geradas. A primeira amostra foi obtida realizando uma leitura dos pixels ao longo das colunas do bloco e a segunda através de uma leitura ao longo das linhas. No total, extraímos 100.000 amostras de cada grupo de imagens.

Os filtros A_{nat} e A_{urb} gerados pelo algoritmo FastICA são mostrados na Figura 7.1.

Os filtros gerados a partir de imagens naturais exibem características similares às wavelets de Gabor, o que corresponde com resultados de pesquisas fisiológicas e



(a)



(b)

Figure 7.1: Filtros (a) $\mathbf{A}_{\text{nat}}$ e (b) $\mathbf{A}_{\text{urb}}$ gerados pelo algoritmo FastICA a partir de imagens naturais e urbanas, respectivamente.

teóricas [DEANGELIS et al(1993)], [OLSHAUSEN, FIELD (1996)]. Além disso, estes filtros não apresentam nenhuma preferência por orientações específicas como sugerido pelo espectro de potência de imagens naturais.

Por outro lado, cenas urbanas produziram filtros similares a Gabor e às wavelets de Haar [MALLAT (1989)]. Os filtros similares a Haar são bem caracterizados pela presença de descontinuidades, as quais só aparecem em nas orientações vertical, horizontal ou ambas. Para comparação, wavelets de Gabor e Haar são mostradas na Figura 7.2.

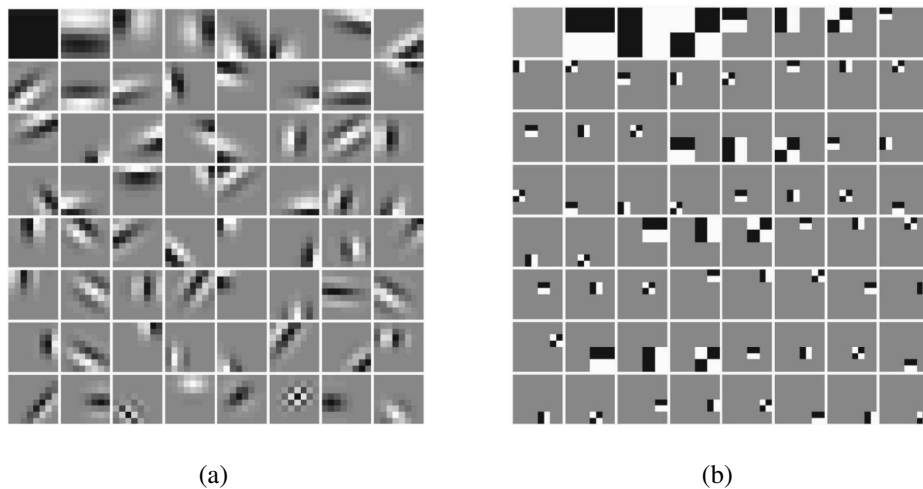
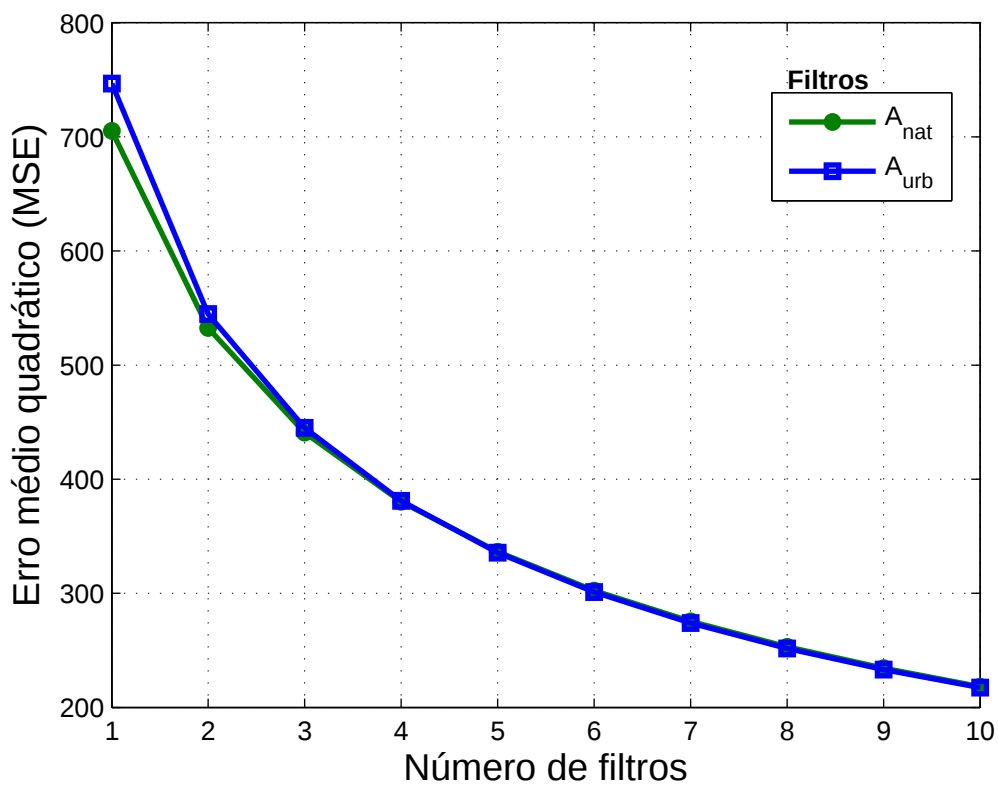


Figure 7.2: Wavelets de (a) Gabor e (b) Haar.

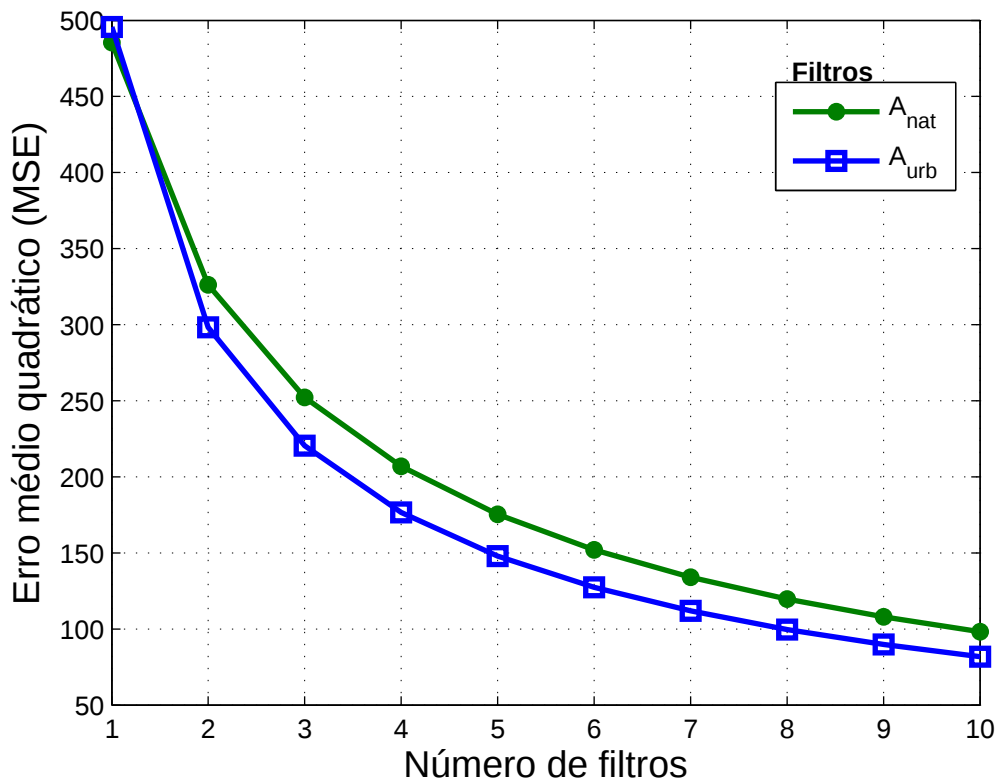
7.3 Curvas de erro

Para avaliar a importância dos filtros $\mathbf{A}_{\text{nat}}$ e $\mathbf{A}_{\text{urb}}$ para representação de imagens, a Figura 7.3 mostra MSE gerado a partir da projeção de 40 imagens naturais e 40 imagens urbanas. Estas imagens não foram utilizadas na fase de aprendizado.

De acordo com a Figura 7.3a, os filtros $\mathbf{A}_{\text{nat}}$ e $\mathbf{A}_{\text{urb}}$ produziram erros similares para representação de imagens naturais. Isto pode ser devido a presença de filtros similares a wavelets de Gabor em $\mathbf{A}_{\text{urb}}$. No entanto, a partir da Figura 7.3b, é possível observar que os filtros $\mathbf{A}_{\text{urb}}$ geraram erros menores na representação de imagens urbanas. Isto sugere que estas imagens não podem ser representadas com erro mínimo a partir somente de filtros similares às wavelets de Gabor, tais como $\mathbf{A}_{\text{nat}}$.



(a)



(b)

Figure 7.3: Erro médio quadrático para a projeção de imagens (a) naturais e (b) urbanas nos filtros gerados por ICA.

8 Discussão

Os resultados encontrados indicam que filtros com características das wavelets de Haar possuem um papel fundamental na representação de imagens urbanas. De fato, observando a Figura 7.1b, estes filtros capturam tanto a preferência de orientação do espectro de potência de imagens urbanas quanto as características de alinhamento em fase de bordas coerentes. Além disso, o error de representação na Figura 7.3b foi sempre menor quando filtros similares às wavelets de Haar foram utilizados.

Estes resultados tem implicações imediatas para área de compressão de imagens. De fato, vários esquemas de compressão utilizam dicionários formados apenas com wavelets de Gabor [FISCHER et al (2001)][DAUGMAN (2002)] [LEE (1996)]. De acordo com os nossos resultados, um dicionário completo para codificação de imagens também deve conter wavelets de Haar assim como $\mathbf{A}_{urb}$.

Por outro lado, $\mathbf{A}_{urb}$ também contém filtros com características das wavelets de Gabor. Isto pode ser devido ao fato que não houve restrição a vegetação nas cenas urbanas utilizadas nos testes. Além disso, objetos presentes nessas cenas podem também apresentar estruturas como bordas fractais. Sob estas condições, imagens urbanas requerem filtros similares a Gabor e a Haar para serem representadas eficientemente.

Estes resultados introduzem uma perspectiva mais ampla para modelos do córtex visual primário. Nesta perspectiva, parte da células da população teriam campos receptivos definidos por descontinuidades ao contrário aos “lóbulos” ou “sidelobes” de wavelets de Gabor. De fato, campos receptivos similares às wavelets de Gabor correspondem somente aos de células simples. Células complexas e hipercomplexas que possuem respostas menos seletivas às características de estímulos visuais não podem ser caracterizadas por uma simples wavelet de Gabor.

Finalmente, foi proposto [OLSHAUSEN, FIELD (2004)] que aproximadamente 60 % da população de células de V1 ainda não foram analisadas devido limitações nos experimentos. Das 40 % estudadas, apenas 30-40 % podem ser descritas pela perspectiva

atual de V1 como “banco de filtros de Gabor”. Dessa forma, levando em consideração que somos extremamente expostos às imagens urbanas, é plausível sugerir que as células de V1 não-descritas por experimentos possuem campos receptivos similares às wavelets de Haar.

9 Conclusão

A teoria atual sobre campos receptivos de V1 descreve células somente como wavelets de Gabor, o que pode ser devido a limitações em experimentos. Os resultados encontrados neste trabalho mostram que, de acordo com o conceito de codificação eficiente, imagens urbanas geram campos receptivos fundamentalmente diferentes dos estabelecidos atualmente.

Os campos receptivos gerados a partir de imagens urbanas são caracterizados pela presença de descontinuidades sendo similares às wavelets de Haar. Observou-se também que o erro associado com a representação de imagens urbanas é sempre menor quando campos receptivos similares a Haar são utilizados.

Um impacto imediato é que esses resultados forcem pesquisadores a reconsiderar o modelo atual de V1. De fato, foi estimado que apenas 15 % das células de V1 já são caracterizadas por tal modelo. O que permanece um problema é desenvolver métodos mais confiáveis para investigar as respostas desconhecidas de V1 em relação aos estímulos.

A partir dos resultados encontrados neste trabalho, é possível propor aplicações para diversas áreas, como por exemplo:

- Classificação de imagens naturais e urbanas a partir dos filtros gerados;
- Análise e segregação de imagens naturais e urbanas;
- Dicionários para compressão de imagens.

Artigos publicados pelo autor

1. Sousa, C.M., Cavalcante, A. B., Guilhon, D., Barros, A. Image compression by redundancy reduction. 7th International Conference on Independent Component Analysis and Signal Separation, London. Proc. Berlin Heidelberg: Springer-Verlag 4666, 422-429 (2007).
2. Santana, E.E., Cavalcante, A. B., Barros, A., Freire, R.C. On-line algorithm for extraction of specific signals with temporal structure. 14th International Conference on Neural Information Processing - ICONIP, Kitakyushu, Proc. (2007).
3. Cavalcante, A. B., Mandic, D. P., Rutkowski, T., Barros, A. Speech enhancement based on the response features of facilitated EI neurons. 6th International Conference on Independent Component Analysis and Blind Signal Separation, Charleston. Proc. Berlin Heidelberg: Springer Berlin 3889, 585-592 (2006).
4. Robson, R., Cavalcante, A. B., Catunda, S. Y., Freire, R.C., Medeiros, E., Barros, A. Low cost electrocardiogram system based on mobile platforms for telemedicine applications. Cairo International Biomedical Engineering Conference - CIBEC, Cairo. Proc. (2006).
5. Sauaia, E.N., Barros, A., Silva, I.B., Soares, L.V., Cavalcante, A. B., Araújo, E.A. A interdisciplinaridade na biomedicina: ciências da computação, engenharia da computação, engenharia elétrica, física, matemática, farmácia/bioquímica e medicina. 45º Congresso Brasileiro de Educação Médica, Minas Gerais. Proc. (2007).
6. Da Silva, M.G., Rodrigues A. B., de Castro, A. C., Moutinho, E. A., Neto, N.S.A., Cavalcante, A. B. Reliability assessment of distribution networks in Brazil's northeast. 8th IEEE International Conference on Probabilistic Methods Applied to Power Systems, Ames-Iowa. IEEE Proc. 94-96, (2004).
7. Da Silva, M. G., Rodrigues, A. B., de Castro, C. L. C., Neto, A. C., Moutinho, E. A., Cavalcante, A. B. An application of predictive reliability analysis techniques in Brazil's northeast distribution networks. International Journal of Electrical Power and Energy Systems 29, 155-162, (2007).

Referências

- [ATICK (1992)] Atick, J. Could information theory provide an ecological theory of sensory processing? *Network: Computation in Neural Systems* 3, 213-251 (1992).
- [ATTNEAVE (1954)] Attneave, F. Some informational aspects of visual perception. *Psychological Review* 61, No 3, 183-193 (1954).
- [BADDELEY et al (1998)] Baddeley R., Abbott, L.F., Booth, M.C., Sengpiel, F., Freeman, T., Wakeman E.A., Rolls E.T. Responses of neurons in primary and inferior temporal visual cortices to natural scenes. *Proc R Soc Lond B Biol Sci* 264, 1775-1783 (1998).
- [BARLOW (1961)] Barlow, H. Possible principles underlying the transformations of sensory messages. Chapter 13. In: *Sensory Communication*, W.Rosenblith (Ed.), M.I.T. Press, 217-234 (1961).
- [BARLOW (1989)] Barlow, H. Unsupervised learning. *Neural Computation*. 1, 295-311 (1989).
- [BRECHT et al (2002)] Brecht, M., Sakmann B. Dynamic representation of whisker deflection by synaptic potentials in spiny stellate and pyramidal cells in the barrels and septa of layer 4 rat somatosensory cortex. *J Physiol* 543, 49-70 (2002).
- [COMON (1994)] Comon, P. Independent component analysis, A new concept? *Signal Processing* 36, 287-314, (1994).
- [DAYAN, ABBOTT (2001)] Dayan, P. e ABBOTT, L. F. *Theoretical Neuroscience - Computational and Mathematical Modeling of Neural Systems*. MIT Press. (2001).
- [DAUGMAN (2002)] Daugman, J. G. Complete discrete 2-D Gabor transforms by neural networks for image analysis and compression. *IEEE transactions on acoustics, speech, and signal processing* 36, 1169-1179 (1988).
- [DEANGELIS et al(1993)] DeAngelis, G, C. Ohzawa, I e Freeman, R.D. Spatiotemporal organization of simple-cell receptive fields in the cat's striate cortex. I. General characteristics and postnatal development. *J. Neurophysiology*. 69(4), 1091-1117 (1993).
- [DEWEESE et al (2003)] DeWeese M., Wehr M., Zador A. Binary spiking in auditory cortex. *J Neurosci* 23, 7940-7949 (2003).
- [FELSEN et al (2005)] Felsen, G., Touryan, J., Han, F. e Dan, Y. Cortical sensitivity to visual features in natural scenes. *PLoS Biol.* 3, e342 (2005).
- [FIELD (1994)] Field, D. J. What is the goal of sensory coding? *Neural Computation* 6, 559-601 (1994).
- [FISCHER et al (2001)] Fischer, S. e Cristobal, G. Minimum entropy transform using Gabor wavelets for image compression. *Image Analysis and Processing Proc.* 428-433 (2001).
- [HYVÄRINEN (1999)] Hyvärinen, A. Fast and Robust Fixed-Point Algorithms for Independent Component Analysis. *IEEE Transactions on Neural Networks* 10(3), 626-634, (1999).
- [HYVÄRINEN et al (2001)] Hyvärinen, A., Hoyer P.O. e Inki. M. Topographic independent component analysis. *Neural Computation* 1, 1527-1558 (2001).
- [HYVÄRINEN et al (2005)] Hyvärinen, A., Gutmann M. e Hoyer P. O. Statistical model of natural stimuli predict edge-like pooling of spatial frequency channels in V2. *BMC Neuroscience* 6:12 (2005).

- [HYVÄRINEN, KARHUNEN, OJA (2001)] Hyvärinen, A., Karhunen, J. e Oja, E. Independent component analysis. John Wiley and Sons (2001).
- [HUBEL, WIESEL (1962)] Hubel, D. H. e Wiesel, T. N. Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *J. Physiol.* 160, 106-154 (1962).
- [LEE (1996)] Lee, T.S. Image representation using 2D Gabor wavelets. *IEEE transactions on pattern analysis and machine intelligence* 18, 959-971 (1996).
- [MALLAT (1989)] Mallat, S. G. A theory for multiresolution signal decomposition: The wavelet representation. *IEEE Trans. Pattern Anal. Mach. Intell.* 7, 674-693 (1989).
- [OLIVA, TORRALBA (2001)] Oliva, A. e Torralba, A. Modeling the shape of the scene: A holistic representation of the spatial envelope. *International Journal of Computer Vision* 42(3), 145-175 (2001).
- [OLMOS et al (2004)] Olmos, A. e Kingdom, F. A. A. McGill Calibrated Colour Image Database, <http://tabby.vision.mcgill.ca>. (2004)
- [OLSHAUSEN, FIELD (1996)] Olshausen, B. A. e Field, D. J. Natural image statistics and efficient coding. *Network* 7, 333-339 (1996).
- [OLSHAUSEN, FIELD (2004)] Olshausen, B. A. e Field, D. J. What is the other 85 % of V1 doing? *Problems in Systems Neuroscience* (2004).
- [OLSHAUSEN, FIELD (1996)] Olshausen, B. A. e Field, D. J. Emergence of simple-cell receptive field proprieties by learning a sparse code for natural images. *Nature* 381, 607-609 (1996).
- [OLSHAUSEN, FIELD (2004)] Olshausen, B. A. e Field, D. J. Sparse coding of sensory inputs. *Current Opinion in Neurobiology* 14, 481-487 (2004).
- [PEREZ-ORIVE et al (2002)] Perez-Orive, J., Mazor O., Turner G.C., Cassenaer S., Wilson R.I., Laurent G.: Oscillations and sparsening of odor representations in the mushroom body. *Science* 297, 359-365 (2002).
- [RIESENHUBER, POGGIO (1999)] Riesenhuber, M. e Poggio, T. Hierarchical models of object recognition in cortex. *Nature Neuroscience.* 2, 1019-1025 (1999).
- [SIMONCELLI, OLSHAUSEN (2001)] Simoncelli, E. P. e Olshausen, B. Natural image statistics and neural representation. *Annu. Rev. Neurosci.* 24, 1193-1216 (2001).
- [SHANNON (1951)] Shannon, C.E. Prediction and entropy of printed English. *Bell Syst. Tech. J.* 30, 50-64 (1951).
- [TORRALBA, OLIVA (2003)] Torralba, A. e Oliva, A. Statistics of natural image categories. *Network: Comput. Neural Syst.* 14, 391-412 (2003).
- [VAILAYA (1998)] Vailaya, A., Jain, A.K. e Zhang, H.J. On image classification: City images vs. landscapes. *Pattern Recognition* 31, 1921-1936 (1998).