

Compressão Seletiva de Imagens Coloridas com Base em Redes Neurais

Diego de Miranda Gomes¹, Wedson Torres de Almeida Filho², Adrião Duarte Dória Neto¹

¹*Departamento de Engenharia da Computação e Automação
Centro de Tecnologia
Universidade Federal do Rio Grande do Norte (UFRN)
59072-970 – Natal/RN, Brasil.*

²*Departamento de Engenharia da Computação e Automação Industrial
Faculdade de Engenharia Elétrica e da Computação
Universidade Estadual de Campinas (UNICAMP)
13083-970 – Campinas/SP, Brasil.*

dmg@dca.ufrn.br, wedson@dca.fee.unicamp.br, adriao@dca.ufrn.br

Abstract

This paper proposes a selective compression method of digital color images where regions of the image, called regions of interest (RoI), are compressed totally lossless and the rest of the image in a lossy manner. This method can be used in applications where the loss of information in some regions is not tolerable, as for example, the medical images for telemedicine use. This way, high compressions ratios are reached without compromising the use of the image. The method uses techniques as artificial neural networks, discrete wavelet transform, mathematical morphology, Huffman adaptive coding, and run-length encoding.

1. Introdução

As técnicas de compressão de imagens objetivam reduzir a quantidade de dados necessários para representar uma imagem digital [1]. Esta redução é possível através da identificação e remoção das redundâncias contidas na imagem. A compressão pode se dar sem perda e com perda de informação. No primeiro caso, o resultado em geral leva a baixas taxas de compressão e no segundo é possível alcançar elevadas taxas de compressão. A decisão quanto ao uso de compressão com perdas ou sem perdas é função da aplicação da imagem. Isto é, existem aplicações em que uma perda da informação é aceitável, mas em outras uma perda da informação não é tolerável. As imagens médicas, para fins de diagnóstico são comprimidas sem perdas [2]. Com a expansão da telemedicina através da internet, verificou-se a necessidade de transmissão de um grande volume de imagens através de meios de comunicação de baixa velocidade inviabilizando algumas aplicações, como o diagnóstico cooperativo. Uma solução pode ser o uso de técnicas de compressão

mistas, onde regiões de interesse são selecionadas e comprimidas sem perdas e o restante da imagem é comprimido com perdas. Obtendo-se assim elevadas taxas de compressão sem comprometer a análise da imagem por parte do profissional médico [3]. Isto naturalmente pode ser aplicado a outras classes de imagens.

O JPEG 2000 *still image compression standard* permite que uma região de interesse possa ser escolhida. Porém, este método não garante uma reconstrução perfeita da região de interesse. Garante apenas que a região de interesse será comprimida com uma melhor qualidade que o resto da imagem [4].

Almeida Filho *et al.* [5-6], propuseram um método com base na transformada wavelet, redes neurais e morfologia matemática, que garante a reconstrução sem perdas da região selecionada. Entretanto, o método se restringia a imagens monocromáticas. Neste artigo é proposto um método de compressão seletiva de imagens coloridas, onde uma ou mais regiões de interesses (RoI) podem ser escolhidas e é garantida uma reconstrução perfeita dessas regiões.

O método tem como base o trabalho desenvolvido anteriormente para imagens monocromáticas, utilizando a transformada wavelet para decorrelacionar os pixels de cada canal, uma rede neural competitiva, (cujo algoritmo de treinamento foi especialmente desenvolvido para esta aplicação) para realizar a quantização vetorial do mapa de significância, morfologia matemática para criar clusters de coeficientes wavelets significantes, código de Huffman adaptativo [8] para reduzir a redundância de codificação e codificação por comprimento de corrida [1] para codificar o mapa das regiões de interesse.

O método utiliza-se do espaço de cor YUV com o intuito de passar a imagem do espaço de cor RGB para um espaço mais decorrelacionado e, dessa forma, conseguir uma maior compressão [1].

Este artigo está organizado da seguinte forma: na seção 2 é feita uma análise sobre espaço de cor e espaço de cor YUV. Na seção 3 é descrito o método de compressão seletiva de imagens coloridas. As seções 4 e 5 abordam as etapas de compressão – com perdas e sem perdas. A seção 6 mostra os resultados obtidos utilizando o método de compressão para algumas imagens e na seção 7 são feitas as considerações finais.

2. Espaços de cores

Os espaços de cores são sistemas de coordenadas tri-dimensionais onde uma cor é representada por um ponto. O RGB é o mais comum, sendo utilizado em monitores de vídeo, por exemplo. Porém neste espaço de cor suas camadas são muito correlacionadas e, para obter vantagens na compressão, se deve converter a imagem para um espaço de cor mais descorrelacionado [1]. O espaço de cor YUV possui uma excelente descorrelação e foi escolhido para o método proposto.

2.1 Espaço de cor YUV

O espaço de cor YUV [1] é, basicamente, uma recodificação do espaço de cor RGB. Cada cor aparece como uma componente de luminância, Y, e duas componentes de crominância, U e V. Luminância, a intensidade percebida, é desacoplada das componentes de crominância, dessa forma a intensidade pode ser alterada sem afetar a cor.

Este espaço de cor foi criado para ser usado na transmissão de TV colorida. Dessa forma uma menor banda é necessária para transmissão em comparação com RGB (no caso de transmissão digital, menos bits). A conversão de RGB para YUV é definida como:

$$\begin{bmatrix} Y \\ U \\ V \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ -0.146 & -0.288 & 0.434 \\ 0.617 & -0.517 & -0.100 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (1)$$

Para fazer a transformação de YUV para RGB é necessário apenas utilizar a matriz inversa.

Esse espaço de cor foi utilizado para esse método, pois, dessa forma, as componentes U e V, que são de menor importância, podem utilizar o mapa de coeficientes significantes obtido da camada Y.

3. Compressão seletiva de imagens coloridas

O diagrama de blocos do método desenvolvido é apresentado na figura 1. Considere I a imagem original em RGB e M o mapa da região de interesse. Os passos do método são os seguintes:

1) Mapear a imagem no espaço RGB para o YUV.

2) Proceder a compressão com perdas detalhada na seção 4 observando que o codificador para os canais U e V difere um pouco da do canal Y.

3) Mapear a imagem comprimida para o espaço RGB, gerando a imagem I' . Obter a imagem de erro, ∇I , como a diferença entre as imagens I e I' .

4) Fornecer a imagem de erro ∇I assim como o mapa da RoI M ao compressor sem perdas.

5) Gerar um único stream de dados a partir das três imagens obtidas no compressor com perdas juntamente com o mapa e os pixels de erro da RoI.

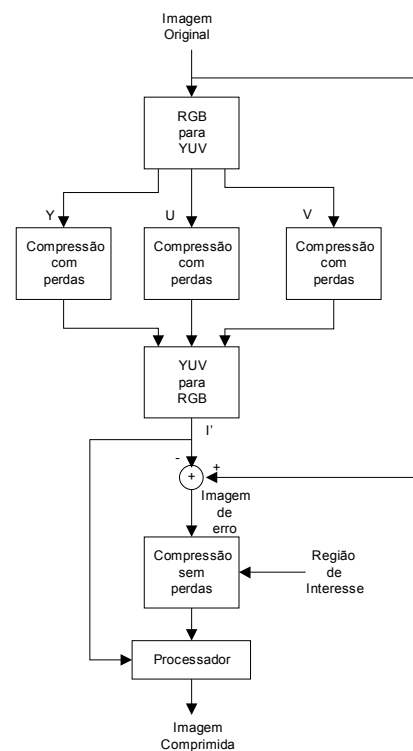


Figura 1: Diagrama de blocos do método de compressão seletiva de imagens coloridas

O método possibilita a escolha de uma ou mais regiões de interesse e de qualquer formato, basta defini-las no mapa. No processo de descompressão, a reconstrução perfeita delas é garantida. O armazenamento dos pixels da imagem de erro é mais vantajoso, pois diminui a abrangência dos valores dos pixels (que normalmente é de 0-255), diminuindo a entropia.

A descompressão, figura 2, é feita de maneira direta. Primeiramente são obtidas do stream de bits informações necessárias para a recuperação da compressão com perdas. Após isso, as três camadas resultantes da transformada YUV com perdas são obtidas e então é feita a transformada YUV inversa para obter a imagem equivalente à original com perdas

I' . Então o mapa da RoI é descomprimido e em seguida os pixels da RoI da imagem de erro também são descomprimidos para serem somados a imagem I' nos locais indicados pelo mapa. Ao fim deste processo, a imagem comprimida resultante é obtida.

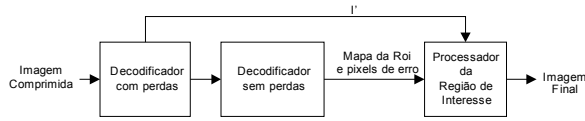


Figura 2: Diagrama de blocos da descompressão

4. Compressor com perdas

A compressão com perdas é aplicada para cada uma das camadas resultantes da transformada YUV, os passos são os seguintes: (1) Mapear para o domínio da transformada, através da transformada wavelet discreta, os pixels do canal de cor. (2) Estabelecer através de um limiar os coeficientes significantes. (3) Gerar clusters de coeficientes significantes através dos operadores de morfologia matemática. (4) Quantizar o mapa de significância através de uma rede neural competitiva com algoritmo de treinamento especialmente desenvolvido para este objetivo. (5) Quantizar os coeficientes significantes. (6) Aplicar a codificação de Huffman [8] tanto no mapa como nos coeficientes significantes.

Os canais U e V, como são de menor importância, utilizam o mesmo mapa de coeficientes significantes obtido do canal Y. Por isso, para as camadas U e V, não é feita nem uma operação sobre o mapa.

4.1 Transformada wavelet discreta

Neste método foram utilizadas as wavelets diádicas, que possuem relação com banco de filtros de reconstrução perfeita [7]. O sinal é decomposto utilizando dois filtros: um passa alta e outro passa baixas frequências. A seguir, uma subamostragem é feita nos dois sinais. O mesmo processo é repetido para o sinal de mais baixa frequência até que seja atingida a escala desejada.

Neste trabalho o filtro que possibilitou um melhor resultado foi o biortogonal [5] e os seus coeficientes são mostrados na tabela 1.

Tabela 1: Coeficientes do filtro wavelet

Biortogonal SWE (13,7)	
h_0	-0.001953125, 0, 0.03515625, -0.03125, -0.123046875, 0.28125, 0.6796875, 0.28125, -0.123046875, -0.03125, 0.03515625, 0, -0.001953125
g_0	0.0625, 0, -0.5625, 1, -0.5625, 0, 0.0625

4.2 Morfologia matemática

Morfologia matemática é um conjunto de técnicas que lida com o formato dos elementos contidos na imagem. Os operadores básicos da morfologia matemática que são invariantes à translação são a dilatação e a erosão por um elemento estruturante A , denotados como $\delta_A(.)$ e $\varepsilon_A(.)$ respectivamente [9]. Banon e Barrera [10] mostraram que qualquer operador morfológico invariante à translação pode ser decomposto em dilatações e erosões. Neste trabalho operador utilizado foi

$$I = \varepsilon_A \{ \delta_A [\delta_A (I)] \}, \quad (2)$$

onde A é um elemento estruturante circular de tamanho 3. A primeira dilatação provoca uma fusão de pontos significantes próximos. Em seguida são realizadas uma dilatação e depois uma erosão (conhecido também como *fechamento*) que funde pequenos buracos nos aglomerados e também provoca a fusão de aglomerados separados por pequenos espaços, sem alterar significativamente a forma dos mesmos [11].

4.3 Rede neural competitiva

Para um algoritmo de treinamento de uma rede neural possa ser considerado competitivo, três elementos básicos devem existir [12]:

- Um conjunto de neurônios iguais exceto pela distribuição randômica dos pesos sinápticos, e que, portanto respondem de maneira diferente para um dado conjunto de padrões de entrada;
- Um limite impondo a força de cada neurônio;
- Um mecanismo que permita que os neurônios possam competir entre si para que, dado um conjunto de entrada, apenas um deles possa estar ativo.

A arquitetura da rede neural usada é mostrada na figura 2. Os neurônios da camada de saída são totalmente conectados aos nós de entrada, e, também, uns aos outros, através de arcos inibidores.

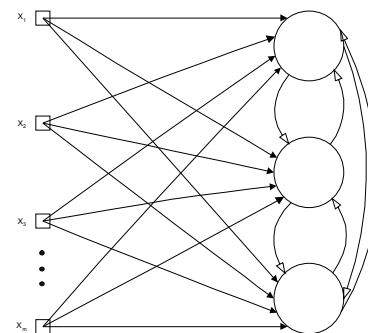


Figura 2: Arquitetura da rede neural utilizada

Como na rede neural competitiva tradicional, o neurônio k considerado vencedor é aquele cujo campo local induzido v_k é o maior dentre todos os neurônios [13]. A saída do neurônio vencedor é um e a saída de todos os outros é zero. O campo local induzido v_k é a distância Euclidiana do ponto representado pelo vetor de entrada e o vetor dado pelos pesos sinápticos do neurônio k , dado por:

$$v_k = \sqrt{\sum_{i=1}^m (w_{ki} - x_i)^2} \quad (2)$$

onde w_{ki} é o peso sináptico do nó de entrada i ao neurônio k .

A rede neural usada neste método possui esses três elementos citados anteriormente. Os neurônios são os mesmos exceto pelos seus pesos sinápticos. Como ele codifica um mata em preto e branco os pesos sinápticos devem assumir valores em um intervalo $\{min, \dots, max\}$, e o limite imposto é quando os pesos sinápticos formam uma imagem totalmente branca. O mecanismo que permite que os neurônios possam competir entre si é o mesmo usado nas redes neurais competitivas tradicionais. Porém o modo como os pesos sinápticos dos neurônios são atualizados é feito de uma forma diferente, de acordo com:

$$\Delta w_{ki}(n) = \eta \cdot \text{sgn}[x_i(n) - w_{ki}(n)], \quad (3)$$

onde $\text{sgn}(\cdot)$ é a função sinal e assume valor -1 se o argumento for negativo, 0 caso seja zero e $+1$ caso seja positivo. O fator η é dado por

$$\eta = \frac{max - min}{k}, \quad (4)$$

onde k é o número de passos (escolhido arbitrariamente) necessário para modificar um peso sináptico do min para o max .

Esta nova regra possui algumas características que a tornam mais vantajosa que a tradicional. Primeiramente, é possível controlar a quantidade de passos necessários para um peso sináptico mudar completamente o seu estado. Se k for escolhido de tal forma que η seja um número inteiro (i.e., $max - min$ seja divisível por k), todas as operações matemáticas realizadas pela rede neural podem ser executadas com aritmética inteira, fazendo com que a sua execução em *software* seja consideravelmente mais rápida e sua implementação em *hardware* mais simples. Experimentos mostram que esta regra tem uma convergência em um número de iterações menor quando comparada com a regra tradicional.

5. Compressor sem perdas

A função do compressor sem perdas é comprimir o mapa da região de interesse e os pixels da imagem de erro para cada uma das camadas de cor R, G e B.

5.1 Codificação do mapa da região de interesse

Na fase de descompressão da imagem, o mapa da região de interesse deve ser descomprimido de forma exata. O método de compressão escolhido foi o por comprimento de corrida [1].

Este método é indicado para os casos em que existe uma repetição muito grande de padrões. É exatamente o que ocorre no mapa de região de interesse. Ele é composto basicamente de muitos zeros (regiões que não são de interesse) e por uns (regiões de interesse).

Ele consiste em substituir repetições seguidas de símbolos pelo símbolo e em seguida de quantas vezes ele se repete. Para a compressão do mapa da RoI este método atinge taxas de compressões muito boas para um método de compressão sem perdas.

5.2 Codificação da região de interesse da imagem de erro.

É mais útil utilizar a diferença entre os pixels da imagem original com os pixels da imagem comprimida ao invés dos pixels da imagem original porque diminui a abrangência dos valores possíveis, pois geralmente a diferença entre a imagem original e a comprimida é pequena.

O método escolhido para fazer a compressão da região de interesse foi a codificação adaptativa de Huffman [8]. Este método consiste em atribuir a símbolos que ocorrem com mais frequência códigos menores, dessa forma diminui o número médio de bits, conseqüentemente, diminuindo o tamanho da imagem. Este método, assim como o por comprimento de corrida, também é sem perdas.

6. Resultados experimentais

Para realizar os testes foram utilizadas algumas imagens e nelas foram escolhidas algumas regiões de interesse. Para avaliar a qualidade das imagens obtidas foi utilizada a taxa pico sinal-ruído (PSNR) que é definida como

$$PSNR(I, \hat{I}) = 10 \cdot \log_{10} \frac{255^2}{e_{ms}^2(I, \hat{I})}, \quad (6)$$

onde 255 é o nível máximo de cada camada de cor da imagem original e e_{ms}^2 é dado por

$$e_{ms}^2(I, \hat{I}) = \frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N [I(i, j) - \hat{I}(i, j)]^2, \quad (7)$$

onde $I(\cdot)$ representa a imagem original $M \times N$ e $\hat{I}(\cdot)$ a imagem reconstruída.

A figura 3 mostra alguns resultados obtidos para o compressor com perdas com as imagens *Lena Color*, *Airplane (F-16)*, *Earth from space* e *House*. Todas elas foram obtidas no *USC-SIPI Image Database* [14]. Para determinar a compressão é necessário escolher alguns fatores como níveis da transformada wavelet, limiar, número de épocas do treinamento, etc. Por este motivo não é possível determinar uma taxa de compressão prévia e é possível obter diferentes taxas de PSNR para uma mesma imagem e uma mesma taxa de compressão.

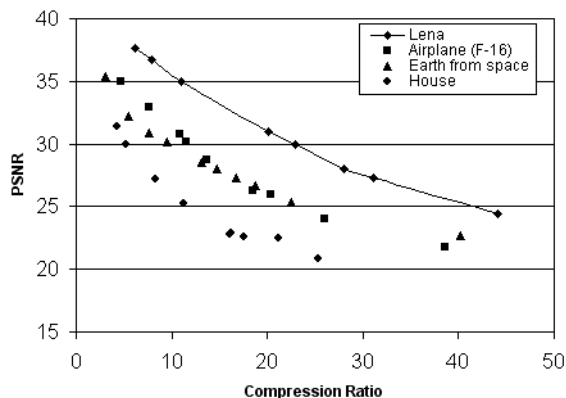


Figura 3: Resultados obtidos para o método de compressão com perdas.



Figura 4: Lena com olhos e boca comprimidos sem perdas.

A perfeita reconstrução das regiões de interesse foi confirmada, pois para os casos em que foram selecionadas regiões de interesse elas possuíam PSNR infinito em comparação com as mesmas regiões da imagem original.

A figura 4 mostra o resultado de uma compressão seletiva feita na imagem Lena onde as regiões de interesse selecionadas foram a região dos olhos e a região da boca com uma taxa de compressão de 17,78. Já a figura 6 mostra o resultado obtido na compressão

de uma imagem médica (*Lewy bodies* – imagem obtida no *Bristol Biomed Image Archive* [15]) onde a original é mostrada na figura 5. Neste caso as regiões de interesse foram as regiões mais escuras. A taxa de compressão foi de 9,69.

Quanto maior forem as regiões de interesse, menor será a compressão final e, em um caso onde a região de interesse for toda a imagem, o método de compressão seletiva terá taxas de compressão equivalentes, ou menores, devido as informações adicionais, aos métodos de compressão sem perdas.

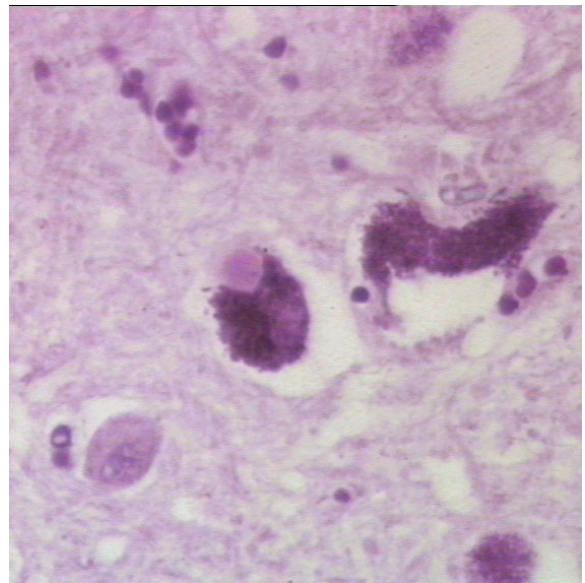


Figura 5: *Lewy bodies* original.

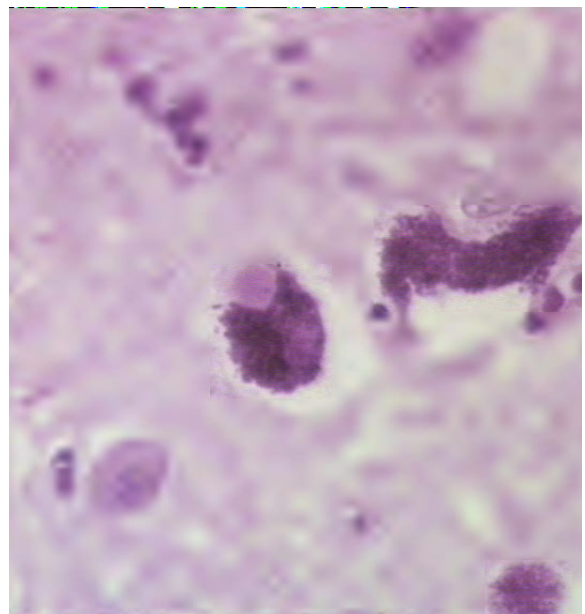


Figura 6: *Lewy bodies* cin as regiões escuras comprimidas sem perdas.

7. Conclusões

Comprimir imagens de maneira seletiva indicando regiões de interesse é interessante, pois permite obter

altas taxas de compressão e mesmo assim possibilita o uso da imagem em aplicações onde não é permitida a perda de informações de certas regiões da mesma.

Os resultados obtidos com o compressor com perdas foram satisfatórios para várias classes de imagem e, as regiões de interesse quando selecionadas, são comprimidas totalmente sem perdas. Já a utilização de apenas um mapa de significância torna o método mais rápido, sem a necessidade de fazer a quantização vetorial de três mapas.

O compressor sem perdas não depende da classe da imagem e como o compressor sem perdas produz bons resultados para uma ampla classe de imagens então o método de compressão seletiva pode ser usado para uma ampla classe de imagens.

Agradecimentos

Os autores gostariam de agradecer ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) pelo apoio financeiro.

Referências

- [1] R. C. Gonzales and R.E. Woods, *Digital Image Processing*. Addison-Wesley, 2000.
- [2] S. Wong, L. Zaremba, D. Gooden, and H. K. Huang, "Radiologic image compression — a review," *Proceeding of the IEEE* 83(2): 194-219, 1995.
- [3] A. Bruckmann and A. Uhl, "Selective medical image compression technique for telemedical and archiving applications," *Computer in Biology and Medicine*, vol. 30, pp. 153-169, 2000.
- [4] C. Christopoulos, A. Skodras and T. Ebrahimi, "The JPEG2000 still image coding system: An Overview," *IEEE Transactions on Consumer Electronics*, Vol. 46, No. 4, pp. 1103-1127, November 2000.
- [5] W. T. Almeida Filho, A. D. Dória Neto and A. M. Brito Júnior, "A neural and morphological method for wavelet-based image compression," *Proc. of the International Joint Conference On Neural Networks*, 2002.
- [6] W. T. Almeida Filho, A. D. Dória Neto and A. M. Brito Júnior, "A selective compression method based on wavelets and competitive neural networks," *Congresso Brasileiro de Automática*, Setembro 2002.
- [7] N. J. Fliege, *Multirate Digital Signal Processing*. Wiley, 1994.
- [8] D. A. Huffman, "A Method for construction of minimum redundancy codes," in *Proc. of the Institute of Electronics and Radio Engineers*, vol. 40, pp. 1098-1101, September 1952.
- [9] H. J. A. M. Heijmans, *Morphological Image Operators*. Academic Press, 1994.
- [10] G. J. F. Banon and J. Barrera, "Minimal representations for translation-invariant set mappings by mathematical morphology," *SIAM Journal on Applied Mathematics*, vol. 51, no. 6, pp. 1782-1798, 1991.
- [11] E. Morales and F. Y. Shih, "Wavelet coefficients clustering using morphological operations and pruned quadrees," *Pattern Recognition*, vol. 33, pp. 1611-1620, 2000.
- [12] D. E. Rumelhart and D. Zipser, "Feature discovery by competitive learning," *Cognitive Science*, vol. 9, pp. 75-112, 1985.
- [13] S. Haykin, *Neural Networks*. Prentice-Hall, 1999.
- [14] USC-SIPI Image Database. Disponível na World Wide Web: <<http://sipi.usc.edu/services/database/Database.html>>
- [15] Bristol Biomed Image Archive. Disponível na World Wide Web: <<http://www.brisbio.ac.uk>>