

## CLASSIFICAÇÃO DE IMAGENS DE ALTA RESOLUÇÃO INTEGRANDO VARIÁVEIS ESPECTRAIS E FORMA UTILIZANDO REDES NEURAIAS ARTIFICIAIS

ANDRÉA FARIA ANDRADE<sup>1</sup>  
MOSAR FARIA BOTELHO<sup>1</sup>  
JORGE ANTONIO SILVA CENTENO<sup>2</sup>

<sup>1</sup>CPGCG – Curso de Pós-Graduação em Ciências Geodésicas - UFPR  
Caixa Postal 515 – 12201-970 - Curitiba - PR, Brasil  
andreaifar@ig.com.br  
mfbotelho@yahoo.com.br

<sup>2</sup>Departamento de Geomática – UFPR  
Caixa Postal 515 - 12201-970 - Curitiba - PR, Brasil  
centeno@geoc.ufpr.br

**Abstract.** The classification of remote sensing images has been an important process for thematic mapping, and one of the main sources for geographic information systems. And has been improved as a result of the development of new sensors technology. Because of that, the efficiency of spectral classification algorithms was put in doubt and new methods are needed to extract information. Artificial Neural Nets arises as a powerful choice to the statistics based classifiers, since no previous information or hypotheses about data distribution is required and input data can have different scales. Shape and size analysis is a powerful tool for discrimination of objects with the same spectral aspect. The aim of this work is the study of the use of auxiliary input information based in the shape of the objects (especially buildings) in the classification of an IKONOS II image through Artificial Neural Nets.

**Keywords:** remote sensing, image processing, neural network, shape parameters.

### 1. Introdução

Com o surgimento dos sensores de alta resolução espacial como o sistema IKONOS II de  $1 m$  ou o QUICKBIRD de  $0.61 m$ , abriu-se a possibilidade da identificação de objetos da paisagem urbana como casas e ruas, provocando uma revolução nos processos de geração e manutenção de bases cartográficas, como também nos processos clássicos de classificação de imagens: Nishida (1998).

De acordo com Meyer e Werth (1990), cerca de 95% de toda classificação procura realizar a tarefa de fotointerpretação (feita por um analista com inúmeras variáveis), usando apenas uma variável, a cor ou tons de cinza. Portanto não é surpreendente que hajam erros em mapas temáticos provenientes da classificação. Vários cientistas sabendo desta condição tentaram aumentar a acurácia e a qualidade da classificação derivada do sensoriamento remoto através da incorporação de dados auxiliares no processo de classificação: Hutchinson (1982), Franklin e Wilson (1992).

Os algoritmos de classificação baseados apenas nos dados puramente espectrais de um *pixel* isolado perderam o espaço para outras abordagens que incluem outros atributos, espaciais e de contexto, no processo decisório. Dentre elas, a segmentação da imagem e a posterior classificação dos segmentos como unidades da mesma é uma opção interessante, pois os segmentos resultantes correspondem a objetos, ou partes deles na imagem, e podem ser descritos em termos de suas características espaciais (forma, área, orientação, etc.) além de seus atributos espaciais. As técnicas de classificação passam então a ser *orientadas ao objeto*, ou seja, baseadas no princípio da análise não apenas de um *pixel* isoladamente, mas sim da

análise dos segmentos obtidos nas imagens através de um processo de segmentação: Ribeiro et al. (2002).

A análise de forma e tamanho é uma poderosa ferramenta para a discriminação de objetos que possuem a mesma aparência espectral. Assim como a análise espectral baseia-se nos valores digitais das bandas, é necessário dispor de um parâmetro para descrever a forma dos objetos. Entre os parâmetros utilizados como descritores da variação espacial do objeto encontram-se: Área ( $A$ ), Perímetro ( $P$ ), Coeficiente de Compacidade ( $cc$ ), razão entre os eixos da elipse que mais se aproxima do objeto, entre outros.

Neste trabalho alguns destes parâmetros serão analisados a fim de se obter o que melhor auxiliará no processo de classificação. Pretende-se, com isto, determinar qual parâmetro de forma é mais apropriado para discriminação de prédios em imagens de alta resolução. Como elemento integrador, as redes neurais artificiais são utilizadas. A utilização de Redes Neurais Artificiais emerge como uma alternativa vantajosa aos classificadores baseados em conceitos estatísticos, uma vez que nenhuma hipótese prévia sobre a distribuição dos dados a serem classificados é exigida e os dados de entrada podem ter escalas diferentes.

## 2. Materiais e Métodos

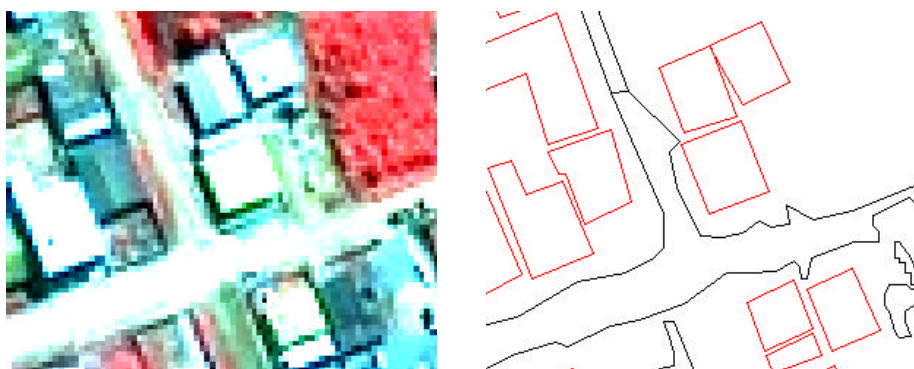
A imagem utilizada para este estudo foi um recorte da imagem híbrida IKONOS II da região do Município de Matinhos, região do Litoral Paranaense, resultante da fusão da banda pancromática com as bandas 2, 3 e 4.

### 2.1. Extração de Regiões – Segmentação

O processo inicia com a segmentação da imagem, identificando neste passo regiões espectralmente similares. De acordo com Mascarenhas e Velasco (1989), região é um conjunto de pontos "*conectados*", ou seja, de qualquer ponto da região pode-se chegar a qualquer outro ponto por um caminho inteiramente contido na região. As regiões que se deseja detectar, geralmente são regiões "*homogêneas*", ou seja, apresentam alguma propriedade local aproximadamente constante em toda sua extensão. Algumas propriedades locais comumente utilizadas são nível de cinza, forma, textura e cor.

A segmentação é um processo em que uma imagem é subdividida em partes ou objetos constituintes. Os algoritmos de segmentação são geralmente baseados em uma das seguintes propriedades básicas de valores de cinza: descontinuidade e similaridade. Na primeira categoria a abordagem é dividir a imagem procurando mudanças bruscas nos níveis de cinza. As principais áreas de interesse nessa categoria são a detecção de pontos isolados e detecção de linhas e bordas na imagem. As principais abordagens da segunda categoria baseiam-se em limiarização, crescimento de regiões e divisão e fusão de regiões: Gonzáles e Woods (2000).

Para a obtenção da imagem temática, a fim de se discriminar os vários segmentos de cada padrão da imagem, foi realizada uma segmentação *orientada a regiões* utilizando o *software E-COGNITION*. Um exemplo do resultado da segmentação pode ser visto na **Figura 1**. Neste exemplo, é visível a separação de regiões espectralmente uniformes. Estas regiões são, a seguir, classificadas considerando seus atributos espectrais e sua forma.



**Figura 1.** Imagem resultante da segmentação realizada no *software E-COGNITION*.

## 2.2. Análise da Forma

Os agrupamentos resultantes de pixels segmentados são usualmente representados e descritos em um formato apropriado para o processamento subsequente. A representação de uma região envolve duas escolhas: através da representação de sua fronteira (forma); ou em termos dos pixels que compõem a região (textura).

Entre os parâmetros utilizados como descritores da variação espacial do objeto encontram-se: área ( $A$ ); perímetro ( $P$ ); coeficiente de compactidade ( $cc$ ); distância polar ao centro de gravidade; relação entre os eixos menor e maior da elipse que melhor se aproxima do objeto; descritores de *Fourier*; sinuosidade dos contornos e a dimensão Fractal ( $D$ ). Há várias outras maneiras de representar e descrever a variação da forma dos objetos, como através da teoria do código da cadeia, teoria das assinaturas, dos números de formas, circularidade, entre outras: Gonzáles e Woods (2000).

Após a segmentação, onde as regiões foram identificadas de acordo com as classes espectralmente similares, foram calculados os parâmetros de forma para estes grupos. E, finalmente, os dados a respeito da forma, junto com a informação espectral média dos segmentos, serviram de entrada numa rede neural, que discriminou os objetos presentes na imagem. Todos estes cálculos forma realizados através de um programa junto ao aplicativo MATLAB, sendo o resultado obtido foi em forma de uma tabela (entrada de dados da rede neural), mostrada abaixo através da **Tabela 1**.

Segmento	R	G	B	Razão Eixos
1.0000000e+000	2.8311856e+001	1.5906959e+002	1.7910309e+002	8.0748987e+000
2.0000000e+000	2.1534615e+002	2.5104327e+002	2.5401442e+002	3.4158753e+000
3.0000000e+000	1.9185417e+002	1.9577083e+002	2.1485417e+002	1.7093830e+000
4.0000000e+000	2.1851744e+002	1.5504070e+002	1.5643023e+002	1.4918898e+000
5.0000000e+000	2.3487019e+002	1.3537019e+002	1.4165865e+002	1.6518829e+000
6.0000000e+000	1.9376364e+002	1.2937273e+002	1.3171818e+002	1.1523692e+000
...				
9.6000000e+002	1.2831111e+002	2.1822222e+001	2.1333333e+001	8.3356224e+000
9.6100000e+002	1.4245455e+002	1.0454545e+001	9.6363636e+000	9.7987274e+000
9.6200000e+002	1.5360000e+002	5.2800000e+001	4.5900000e+001	2.5000000e+000
9.6300000e+002	1.8492857e+002	1.2950000e+002	1.3192857e+002	3.5000000e+000
9.6400000e+002	1.9680000e+002	2.2950000e+002	2.3470000e+002	3.0000000e+000

**Tabela 1.** Tabela de entrada para a classificação por Redes Neurais Artificiais, com os valores de nível de cinza para cada banda, e os parâmetros de forma correspondentes para cada segmento da imagem.

### 2.2.1. Área e Perímetro

São os mais simples descritores utilizados, a área é obtida contando todos os pixels que formam a região, e o perímetro é calculado considerando o número de *pixels* conexos (em função do predicado de conectividade escolhido) que constituem o contorno da região.

### 2.2.2. Distância Polar ao Centro de Gravidade (Momentos)

Assumindo que regiões representadas por valores de *pixel* diferentes de zero, *momentos* podem ser utilizados para descrever regiões binárias ou em nível de cinza. Um momento de ordem  $(p + q)$  é dependente de escala, translação, rotação e em uma imagem em nível de cinza a equação do momento é dada por: Facon (1993):

$$m_{pq} = \sum_i \sum_j x_i^p y_j^q f(i, j)$$

onde  $f(i, j)$  representa o nível de cinza do ponto  $P(i, j)$ .

No caso de uma imagem binária (dois níveis de cinza, ou seja,  $f(i, j) = 0$  ou  $1$ ), os momentos permitem definir algumas propriedades de base de uma região:

- área: essa propriedade comumente usada corresponde ao momento  $m_{00}$ . Esse momento representa o número de *pixels* constituindo a região:

$$m_{00} = \sum_i \sum_j f(i, j)$$

- centro de gravidade: definido da seguinte maneira:

$$x_g = m_{10} / m_{00}$$

$$y_g = m_{01} / m_{00}$$

- eixos principais: eles são os vetores próprios da matriz variância covariância obtida usando os *pixels* constituindo a região como variáveis aleatórias. Os dois vetores próprios mostram as direções da expansão máxima da região. A medida do degrau da expansão é dado pelos valores próprios correspondentes. Então o primeiro eixo principal corresponde à direção do vetor definido pelo maior valor próprio. O segundo eixo principal é ortogonal ao primeiro. A orientação  $v$  do grande eixo de inércia pode ser determinada a partir dos momentos centrados  $\mu_{01}, \mu_{10}, \mu_{11}, \mu_{02}$  e  $\mu_{20}$ :

$$v_{pq} = \frac{\mathbf{m}_{pq}}{(\mathbf{m}_{00})^g}$$

onde  $\mu_{pq}$  representam os momentos centrados de ordem  $p + q$  cuja fórmula é dada por:

$$\mathbf{m}_{pq} = \sum_i \sum_j (x_i - x_g)^p (y_i - y_g)^q f(i, j)$$

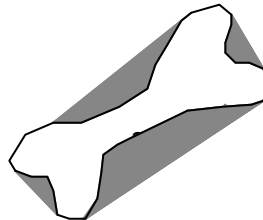
De acordo com Sonka (1998), este tipo de descrição é invariável pela translação e mudança de escala. Para eliminar o efeito de rotação, aplicam-se os 7 momentos normalizados, que são obtidos a partir dos momentos  $\mu_{pq}$  e estes são invariantes a translação, rotação e mudança de escala, e são dados por:

$$\begin{aligned}
 \mathbf{f}_1 &= v_{20} + v_{02} \\
 \mathbf{f}_2 &= (v_{20} + v_{02})^2 + 4v_{11}^2 \\
 \mathbf{f}_3 &= (v_{30} - 3v_{12})^2 + (3v_{21} - v_{03})^2 \\
 \mathbf{f}_4 &= (v_{30} - v_{12})^2 + (v_{21} + v_{03})^2 \\
 \mathbf{f}_5 &= (v_{30} - 3v_{12})(v_{30} + v_{12})[(v_{30} + v_{12})^2 - 3(v_{21} + v_{03})^2] \\
 &+ (3v_{21} - v_{03})(v_{21} + v_{03})[3(v_{30} + v_{12})^2 - (v_{21} + v_{03})^2] \\
 \mathbf{f}_6 &= (v_{20} - v_{02})[(v_{30} + v_{12})^2 - (v_{21} + v_{03})^2] + 4v_{11}(v_{30} + v_{12})(v_{21} + v_{03}) \\
 \mathbf{f}_7 &= (3v_{21} - v_{03})(v_{30} + v_{12})[(v_{30} + v_{12})^2 - 3(v_{21} + v_{03})^2] \\
 &- (v_{30} - 3v_{12})(v_{21} + v_{03})[3(v_{30} + v_{12})^2 - (v_{21} + v_{03})^2]
 \end{aligned}$$

### 2.2.3. Deficiência Convexa

A decomposição de fronteiras em segmentos é freqüentemente útil. Tais decomposições reduzem a complexidade da fronteira, simplificando o processo de descrição. Essa abordagem é atrativa quando a fronteira estiver uma ou mais concavidades significativas que carreguem informação sobre a forma. Neste caso, o uso do fecho convexo da região delimitada pela fronteira é uma ferramenta poderosa para a decomposição robusta da fronteira.

O *fecho convexo*  $H$  de um conjunto arbitrário  $S$  é o menor conjunto convexo que contiver  $S$ . A diferença entre os conjuntos  $H - S$  é chamada *deficiência convexa*  $D$  do conjunto  $S$ . Para imaginar como estes conceitos podem ser utilizados, considere a **Figura 2**, que mostra um objeto (conjunto  $S$ ) e sua deficiência convexa (regiões sombreadas): Gonzales e Woods (2000).



**Figura 2.** Região ( $S$ ) e sua deficiência convexa. Fonte: Adaptado de Gonzáles e Woods (2000).

### 2.3. Classificação dos segmentos usando Redes Neurais Artificiais

De acordo com Haykin (2001), o trabalho em Redes Neurais Artificiais (RNA's), tem sido motivado desde o começo pelo reconhecimento de que o cérebro humano processa informações de uma forma inteiramente diferente do computador digital convencional. O cérebro é um sistema de processamento de informação altamente complexo, não-linear e paralelo. Ele tem a capacidade de organizar seus constituintes estruturais, conhecidos por neurônios, de forma a realizar certos processamentos (p. ex., reconhecimento de padrões, percepção e controle motor) muito mais rápido que o mais ágil computador hoje existente.

Nelson e Illingworth (1990), definem as RNAs como sendo técnicas inovadoras de processamento de informações, que têm como objetivo básico, simular, em computadores, o funcionamento do sistema nervoso biológico.

Conforme Dyminski (2000), uma rede neural é constituída por diversos neurônios dispostos em camadas. Usualmente as camadas são classificadas em três grupos: camada de entrada: onde os padrões são apresentados à rede; camadas escondidas: onde é feita a maior parte do processamento; camada de saída: onde o resultado final é concluído e apresentado. O número de neurônios de entrada é função do número de dados considerados na classificação, no caso do presente trabalho o número de entradas será a soma das entradas espectrais (RGB) e dos dados espaciais (descritores de forma utilizados).

No presente trabalho utilizou-se uma rede do tipo *feedforward* com o algoritmo *backpropagation*, já que os mesmos são os mais recomendados pela bibliografia no reconhecimento de padrões. Para os primeiros experimentos para a classificação dos segmentos, utilizou-se área, perímetro, coeficiente de compacidade, razão dos eixos, deficiência convexa e momentos como parâmetros descritores da forma. Utilizou-se de 21 segmentos para cada padrão como amostras de treinamento e foram utilizadas duas camadas escondidas 138 e 46 neurônios respectivamente. O número de neurônios das camadas escondidas como o número de amostras forma encontrados através de experimentos realizados com imagens sintéticas e posteriormente com imagens IKONOS (porém de menos complexidade para a classificação).

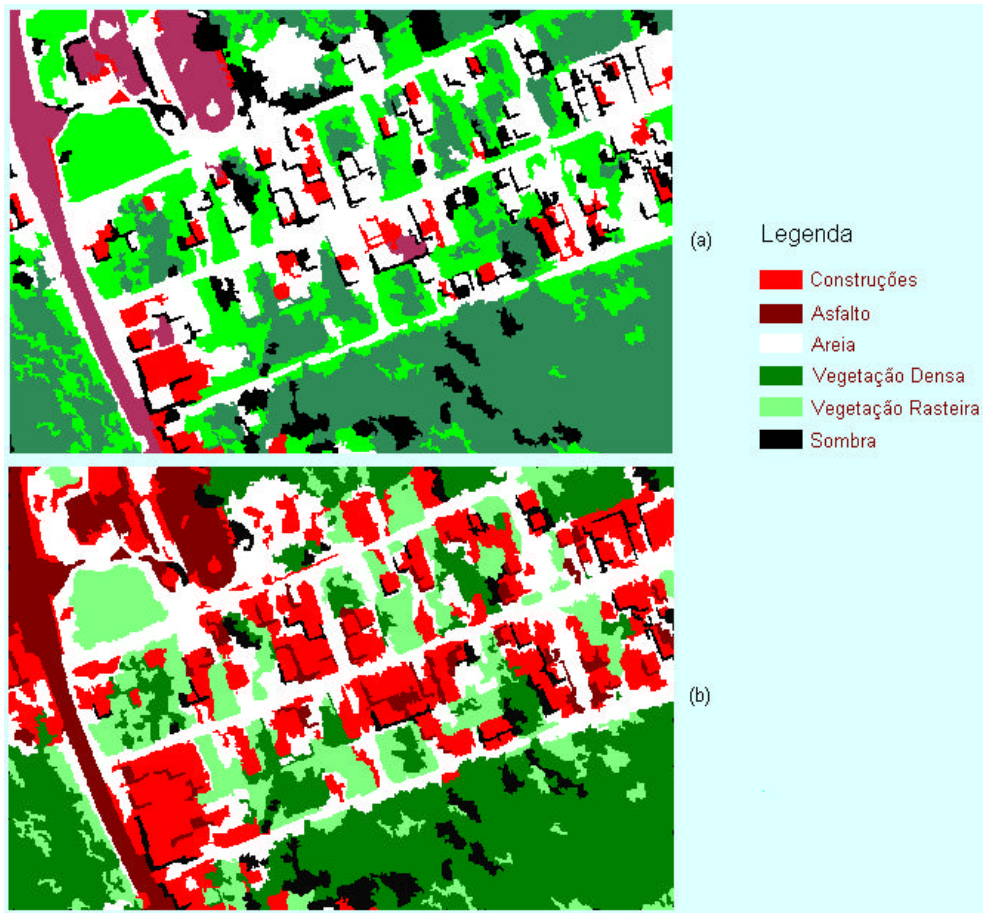
Nos experimentos realizados, as entradas da rede neural foram variadas para avaliar a melhor combinação de variáveis. No primeiro experimento, apenas as variáveis espectrais foram usadas. Numa segunda série de experimentos, junto às variáveis espectrais, os parâmetros de forma foram incluídos, um de cada vez. A acurácia da classificação foi analisada através da comparação entre as imagens classificadas e uma imagem considerada como “verdade de campo” obtida através da interpretação visual. O sucesso de cada conjunto de variáveis foi avaliado calculando a correlação entre a imagem temática resultante e a imagem de referência. Os resultados foram resumidos na **Tabela 2**.

Método	Correlação
RGB *	38%
RGB + Área	81%
RGB + Perímetro	84%
RGB + Deficiência Convexa	82%
RGB + Momento	83%

\* Executado no Envi

**Tabela 2.** Correlação entre as imagens temáticas e a imagem de referência (análise visual) para a classe “construções”.

A **Figura 3** mostra uma comparação entre o resultado da classificação puramente espectral (**Figura 3a**) e outro obtido usando parâmetros de forma (**Figura 3b**). No resultado da classificação puramente espectral, observa-se uma certa confusão entre as classes “areia” e “construções”, pois ambos possuem características espectrais similares. A imagem temática apresentada na **Figura 3b** foi obtida usando o descritor “*perímetro*” e nela pode ser visto que estas classe foram melhor discriminados.



**Figura 3.** (a) Classificação *pixel-a-pixel* puramente espectral; (b) Classificação *orientada-a-regiões* com parâmetros espectrais e espaciais.

### 3. Considerações Finais

A análise dos resultados obtidos na classificação espectral *pixel-a-pixel* comprova a limitação desta abordagem para imagens de alta resolução. A principal causa é o fato dos objetos visíveis na imagem serem espectralmente muito próximos e sua discriminação visual ser possível usando outros atributos além dos espectrais. Os dados providos pelos sistemas sensores de alta resolução apresentam maior detalhe, e os métodos de classificação *orientados ao objeto* passam a ter maior importância, já que os mesmos consideram além dos atributos espectrais as relações espaciais entre os *pixels*. No caso particular de uma cena urbana, a análise da forma ganha significado, pois os objetos presentes na imagem são construídos pelo homem, com um predomínio de formas geométricas.

O ganho obtido com a inclusão de parâmetros de forma é significativo. No experimento, a acurácia estimada para a classe “construções” subiu de 38% até 84%. Embora o melhor parâmetro seja “perímetro”, os outros, “deficiência convexa” e “momento” também possibilitaram um aumento da qualidade do mapa temático. O aumento da acurácia usando estes três parâmetros é praticamente igual.

A utilização de Redes Neurais Artificiais emerge como uma alternativa vantajosa aos classificadores baseados em conceitos estatísticos, já que se podem obter resultados satisfatórios utilizando-se de um número mínimo de amostras para que a mesma faça a generalização. A flexibilidade dos algoritmos de classificação baseados em redes neurais permite integrar variáveis diferentes na identificação de objetos, como por exemplo de características espectrais e espaciais.

A precisão do resultado da classificação realizada é dependente da segmentação utilizada para a extração das regiões da imagem a serem analisadas, isto porque uma má formação das regiões (por exemplo, segmentos muito pequenos ou muito grandes), ocasionará em um péssimo resultado da classificação.

Dos experimentos realizados com os descritores citados, pôde-se concluir que o perímetro apesar de ser um descritor simples é considerado relevante ao resultado da classificação.

## Referências

- Crosta, A. P. *Processamento Digital de Imagens de Sensoriamento Remoto*. Campinas: IG/Unicamp, 1992. 170p.
- Dyminski, A. S. *Análise de Problemas Geotécnicos Através de Redes Neurais*. Rio de Janeiro. (Tese) - Universidade Federal do Rio de Janeiro, 2000.
- Facon, J. *Processamento e Análise de Imagens*. Argentina: Universidad Nacional de Córdoba, 1993. 198p.
- Franklin, S.E.; Wilson, B.A A Three Stage Classifier for Remote Sensing of Mountain Environment. *Photogrammetric Engineering and Remote Sensing*. v. 58, n. 4, p. 449-454, 1992.
- Gonzales, R.C.; Woods, R.E. *Processamento de Imagens Digitais*. São Paulo: Edgard Blücher Ltda, 2000. 509p.
- Haykin, S. *Redes Neurais. Princípios e prática*. Porto Alegre: Bookman, 2001. 900p.
- Hutchinson, F.C. Techniques for Combining Landsat and Ancillary Data for Digital Classification Improvement. *Photogrammetric Engineering and Remote Sensing*. v. 48, n. 1, p. 123-130, 1982.
- Mascarenhas, N.D.A; Velasco, F.R.D. *Processamento Digital de Imagens*. Buenos Aires: Kapelusz, 1989.
- Meyer, M. Werth. Satellite Data: Management Panacea or Potencial Problem? *Journal of Forestry*. v. 88, n. 09, p. 10-13, 1990.
- Nelson, M.M.; Illingworth, W.T. *A Pratical Guide to Neural Networks*. USA: Addison-Wesley Publishing Company, 1991.
- Nishida, W. *Uma Rede Neural Artificial para Classificação de Imagens Multiespectrais de Sensoriamento Remoto*. Florianópolis. (Dissertação) - Universidade Federal de Santa Catarina, 1998.
- Ribeiro, S. R. A; Santos, D. R.; Centeno, J. S. Aplicação da Metodologia de Dados Orientado a Objeto na Classificação de uma Área Urbanizada, Utilizando uma Imagem Digital Obtida por Meio da Tecnologia do Laser Scanner. Presidente Prudente, jul. 2002. *Anais: Presidente Prudente: UNESP – Simpósio Brasileiro de Geomática*, 2002. p. 144-149.
- Sonka, M. *Image processing, analysis, and machine vision*. USA: Brooks/Cole, 1999.