

IDENTIFICAÇÃO DE EDIFICAÇÕES PRESENTES EM IMAGEM DE ALTA RESOLUÇÃO UTILIZANDO REDES NEURAIS E DADOS DO LASER SCANNER

Mosar Faria Botelho ¹
Jorge Antonio Silva Centeno ²

¹ Universidade Federal do Paraná – UFPR – Programa de Pós-Graduação em Ciências Geodésicas – mfbotelho@yahoo.com.br

² Universidade Federal do Paraná – UFPR – Departamento de Geomática –centeno@ufpr.br

RESUMO

O objetivo deste estudo está em mostrar uma alternativa mais viável para a identificação de edificações através da classificação de imagem de alta resolução utilizando dados do laser scanner e imagens do QUICKBIRD II. Para tanto foi utilizado dados de intensidade e altimetria do laser, integrando-os com a banda pan-cromática do sensor QUICKBIRD II, por meio do algoritmo de redes neurais e uma abordagem orientada a regiões. O trabalho justifica-se por utilizar tecnologias recentes (laser scanner e imagem QUICKBIRD II) e um algoritmo integrador de variáveis de diferentes origens (as redes neurais artificiais), na elaboração de mapas temáticos com custos menores.

Palavras-chave: Sensoriamento Remoto, Redes Neurais Artificiais, Laser Scanner

IDENTIFICATION OF PRESENT CONSTRUCTIONS IN IMAGE OF HIGH SPACE RESOLUTION USING NEURAL NETWORK IS GIVEN OF THE LASER SCANNER

ABSTRACT

The goal of this work is to show a viable option for building identification through high resolution image classification using laser scanner data and images of QUICKBIRD II. For doing so it was used intensity and altimetry laser data integrated with the pan-chromatic band of the QUICKBIRD II sensor, by means of neural network algorithms and a region oriented approach. The work is justified because it makes use of recent technologies (laser scanner and QUICKBIRD II images) and it can reduce the production costs of a thematic map.

Keywords: Remote Sensing, Neural NetWork, Laser Scanner.

1. INTRODUÇÃO

Uma tecnologia recente é o laser scanner. Esta surge como uma tecnologia interessante para a extração de informações dos objetos da superfície do solo. Resumidamente, o método envolve a emissão de um pulso eletromagnético em direção a superfície tendo como produto as coordenadas plano-altimétricas e a intensidade do feixe refletido pelo objeto atingido. Esta informação, a altura dos pontos na superfície da terra, pode também ser incluída no processo de classificação.

O surgimento dos sistemas de Sensoriamento Remoto de alta resolução espacial, como o sistema IKONOS II, lançado em meados 1999, com resolução espacial de 1m e o QUICKBIRD II com 61 cm, possibilita a análise de objetivos menores, como ruas e edificações, fato que vêm provocando uma revolução nos processos de geração e manutenção de bases cartográficas, como também nos processos de classificação de imagens. Porém, estas imagens, possuem um custo elevado para sua aquisição.

A classificação de imagens digitais tem por objetivo extrair feições espaciais de forma automática visando à compreensão humana. O processo de classificação de imagens pode ser realizado por dois processos distintos: “pixel” a “pixel” ou por regiões. O processo “pixel” a “pixel” consiste na atribuição de cada “pixel” a uma determinada classe analisando, na maioria dos casos, apenas as propriedades espectrais dos mesmos. O processo por regiões consiste em primeiro agrupar “pixels” similares e próximos em regiões e em atribuir cada região a uma das classes. Esse processo tem como vantagem a utilização de descritores podendo ser forma, tamanho e textura da região. Sendo que a definição de regiões é dada por um grupo de “pixels” que são similares e adjacentes. Observa-se também que o aumento de variáveis de entrada no processo de classificação gera problemas como o aumento do tempo de processamento e dificuldade da proposta do modelo mais adequado.

Para classificar dados de diferentes origens as Redes Neurais Artificiais (RNA's) são atualmente uma opção emergente. A utilização RNA's emerge como uma alternativa vantajosa aos classificadores de imagens baseados em conceitos estatísticos (Ex.: Máxima Verossimilhança), 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.

O uso de RNA's, aliado ao aumento de variáveis derivadas da imagem, que oferece a abordagem de classificação baseado em regiões, é uma alternativa viável e interessante no sensoriamento remoto, especialmente na classificação de imagens de alta resolução espacial. No intuito de amenizar os custos de aquisição da imagem, o objetivo deste estudo está em avaliar a classificação de imagem digital utilizando apenas a banda PAN do satélite QUICKBIRD, a altimetria e a Intensidade fornecida pelo laser scanner aplicando a estas uma abordagem orientada a regiões.

2. REVISÃO LITERÁRIA

Para um melhor entendimento deste estudo serão abordados, nas próximas seções, os conceitos de laser scanner, redes neurais e classificação digital de imagem.

2.1. LASER SCANNER

De acordo com CENTENO (2003), o laser scanner é um sistema ativo usado para medir a distância entre o sensor e a superfície dos objetos. Seu funcionamento baseia-se na utilização de um feixe de laser que é disparado em direção aos objetos, refletido por eles e novamente captados pelo sistema. A partir do sinal de retorno e o tempo decorrido entre a emissão e o registro do retorno, informações como a distância entre o sensor e o objeto podem ser calculadas, assim como a intensidade da reflectância dos objetos podem ser captadas. Este tipo de sistema também é chamado de LIDAR (Light Detection And Ranging) ou LADAR (Laser Detection And Ranging).

2.2. REDES NEURAS ARTIFICIAIS

Uma rede neural pode ser definida como um arranjo de unidades (neurônios interconectados) que trabalham em paralelo para classificar e generalizar dados de entrada em classes de saída (RIBEIRO e CENTENO, 2000).

As redes neurais são organizadas em camadas conectadas, podendo ser a “camada de entrada”, a(s) camada(s) escondida(s) ou oculta(s) e a camada de saída. Na camada de entrada, ou simplesmente nomeada “entrada” devido não ocorrer processamento, os dados são apresentados à rede. A camada de entrada se comunica a uma ou mais camadas ocultas onde ocorre o processamento por meio de um sistema de conexões ponderadas. A camada de saída apresenta o resultado do processamento. O funcionamento da rede é determinado pela conexão entre as unidades de processamento (GERMAN e GAHEGAN, 1996), definida pelos pesos W_i de cada ligação.

De acordo com JARVIS e STUART (1996), o número de neurônios de entrada é função do número de dados considerados na classificação, no caso do presente trabalho o número de entradas é a soma das entradas espectrais, espaciais e altitude fornecida pelo laser scanner. A saída da rede, para este trabalho, é a soma das classes que se deseja classificar.

2.3. CLASSIFICAÇÃO DIGITAL

Um modo de tornar a interpretação de imagens mais simples consiste em separar grupos de elementos da imagem (“pixels” ou regiões) com características similares em classes, geralmente associadas a tipos de uso ou cobertura do solo. Tomando como hipótese que os elementos da imagem que cobrem um mesmo alvo têm características similares, algoritmos computacionais podem ser usados para dividi-los em grupos. Este processo é denominado classificação de imagem.

Os métodos de classificação podem ser “pixel” a “pixel” ou por regiões, dependendo do nível de abstração. O método “pixel” a “pixel” admite como o menor elemento a ser classificado o “pixel” e o método por regiões admite como menor elemento o segmento, ou seja, a região que pode representar um objeto na cena ou parte deste, permitindo a descrição de sua forma.

De acordo MASHARU e HASEGAWA (2000), um método eficiente para extrair dados de prédios de alta elevação, obtidos por um laser scanner, é aplicando o método de segmentação de regiões. A idéia básica é que cada prédio tem uma altura diferente do vizinho. Entretanto pode-se distinguir cada prédio segmentando o DEM, com a condição de que um “pixel”, com diferença de altura, comparado com seu vizinho, pertença a mesma região se estiverem dentro de uma área pré-determinada. O estudo constatou alguns problemas como: os objetos extraídos não são apenas prédios ou outros objetos diferentes mas por exemplo árvores, também são incluídos. O estudo, apesar do problema mencionado aponta que a segmentação dos dados do laser scanner é um bom método para gerar modelo de imagem em 3D, no sentido que o método pode gerá-la através de um processo automático.

3. MATERIAL E MÉTODO

3.1. MATERIAL

3.1.1. Área de estudo

Foi realizado um recorte de banda pan-cromática pertencentes a imagem QUICKBIRD II correspondentes à área de estudo. Este recorte apresenta vias asfaltadas, casas, árvores e vegetação branda (gramíneas, pequenos arbustos, pastagens etc), destaca-se pela grande aglomeração entre edificações, representando um ambiente urbano realista. Este recorte pode ser observado na Figura 1



Figura 1 - Imagem pan-cromática do QUICKBIRD II, da região em estudo

3.1.2. Dados do laser scanner

O levantamento laser scanner da região, foi realizado com o sistema OPTECH / INS GPS APPLANIX, pertencente ao Instituto de Tecnologia para o Desenvolvimento (LACTEC - UFPR). O sistema de referência de coordenadas utilizado para o rastreamento foi o WGS84 (World Geodetic System, 1984).

A imagem intensidade foi fornecida pelo LACTEC (Figura 2b). Nesta foi realizado uma reamostragem, utilizando o método bilinear, com o objetivo de obter a mesma resolução espacial (0.7cm) da imagem pan-cromática do sensor QUICKBIRD II.

Utilizando os dados altimétricos disponíveis pelo laser é gerada, através do programa SURFER, utilizando o método de interpolação inversa da distância, uma grade altimétrica regular de 0.7m com as mesmas dimensões da resolução espacial da imagem, bem como as mesmas coordenadas. No intuito de obter apenas as alturas dos objetos em relação ao solo, presentes na cena, foram realizados os seguintes passos:

1. Inicialmente, são coletados pontos (aproximadamente 70 pontos para cada recorte), a partir da grade regular original. Pontos representativos que modelem a superfície do solo, desprezando assim as elevações das edificações, árvores, etc.
2. De posse deste conjunto de pontos, que descrevem a superfície do solo, é gerada uma segunda grade desses pontos.
3. Finalmente, em um terceiro momento, é realizada a subtração das grades. Esta operação resulta em uma terceira grade com apenas a altura dos objetos em relação ao solo, contidos na cena. Este produto final pode ser observado na da Figura 2a e é chamado de MDS (modelo digital de superfície normalizado (WEIDNER E FÖRSTNER, 1995)).



Figura 2 – Imagens oriundas dos dados do laser scanner

3.2. MÉTODO

O objetivo destes experimentos é classificar uma imagem de alta resolução espacial, utilizando dados do laser scanner como variáveis auxiliares no processamento de aprendizado da rede neural, orientados a região.

3.2.1. Segmentação

O conceito de evolução de rede fractal (FNEA: fractal net evolution approach), considera a imagem de sensoriamento remoto como de natureza fractal. Ao segmentar-se a imagem simultaneamente de forma fina e de forma grosseira, dá-se origem a diferentes escalas de segmentação. O procedimento de segmentação é similar ao crescimento de regiões que agrupa “pixels” ou sub-regiões em regiões maiores. A segmentação da imagem resulta em segmentos ou objetos, que baseados em parâmetros espectrais e espaciais, podem ainda ser reagrupados em objetos maiores denominados superobjetos ou regiões. Estas regiões possuem relações com os segmentos que as constituem num nível hierárquico mais baixo (ANTUNES, 2003).

No mercado existem alguns programas que vem utilizando o algoritmo de segmentação FNEA. Cita-se como exemplo, o programa “eCognition”, onde o usuário pode representar o conhecimento utilizando relações derivadas experimentalmente e modificar tais relações de maneira a satisfazer os fins da análise da imagem. Neste programa, os objetos resultantes da segmentação representam a informação da imagem de uma forma abstrata. Além das informações espectrais outras informações ou atributos adicionais, como forma, textura relacionamento entre objetos e sua vizinhança, podem ser usados para a classificação (ECOGNITION, 2001).

Para representar o conhecimento do usuário, o “eCognition” dispõe de redes conectadas, chamadas de classes hierárquicas de operadores. Estas classes são capazes de modificar suas conexões ao sofrerem alterações nas informações fornecidas pelo usuário. Regiões representativas de cada classe são fornecidas como amostras e, a partir delas, uma descrição das classes é obtida. O usuário pode selecionar a informação útil, interagindo com o programa através da escolha dos atributos relevantes, derivados da imagem, ou informações advindas de mapas. Quanto mais informações relevantes o usuário introduzir no sistema, maior é o conhecimento armazenado nas classes hierárquicas.

3.2.2. Análise Estatística

Para a verificação do melhor classificador, foi utilizado o índice “kappa”, proposto por COHEN (1960), que retrata o quanto o algoritmo acertou em uma determinada área de verificação. Esta área foi obtida através de uma restituição fotogramétrica, realizada na escala de 1:2000, de uma foto aérea na escala de 1:8000.

No intuito de verificar a qualidade da classificação, será comparado o produto obtido pela rede neural com o produto obtido pelo algoritmo mínima distância. O método da mínima distância atribui cada elemento desconhecido à classe cuja média é mais próxima, ou seja, a atribuição de uma classe ao elemento depende da distância em relação à classe a qual este foi atribuído.

3.2.3. Redes Neurais

Para este trabalho é adotada a arquitetura multicamadas com alimentação adiante (“feed-forward”) com o aprendizado de retro-propagação (“back-propagation”); possui “bias” que tem influência externa ao neurônio, e a função de transferência adotada é a sigmóide.

4. EXPERIMENTO

Os experimentos foram executados na seguinte ordem: segmentação da imagem, amostragem e treinamento da rede neural e finalmente a classificação da imagem e sua verificação.

4.1. SEGMENTAÇÃO

O resultado da segmentação no “eCognition”, que utiliza o método FNEA, é uma imagem temática, onde a cada região é atribuído um valor digital, e uma tabela com os descritores de cada região, disponíveis neste sistema. Com estes segmentos foi realizada a amostragem.

O recorte mencionado, foi segmentado utilizando os mesmos parâmetros de escala, cor, forma, compacidade e suavidade disponíveis no “eCognition”. Quanto maior o valor da escala, maior será as regiões que descrevem o objeto na cena, ou seja, menos detalhes é observado na imagem.

Para a realizar a segmentação no “eCognition”, foram introduzidas a banda pan-cromática, imagem altimétrica e imagem intensidade, todas com o mesmo peso, originando uma nova imagem segmentada (Figura 3). Cada região desta imagem é codificada, ou seja, a cada região é dado um valor digital. Os valores utilizados para esta segmentação pode ser observado pelo Quadro 1.

Quadro 1 - Fornece os parâmetros utilizados em cada nível de segmentação no ecognition.

Nível	Escala	Escala		Forma	
		Cor	Forma	Suavidade	Compacidade
1	5	0.7	0.3	0.7	0.3
2	5	0.3	0.7	0.7	0.3
3	10	0.3	0.7	0.7	0.3

Como pode ser observado no Quadro 1, foram utilizados 3 níveis de segmentação. No primeiro nível foi utilizada uma escala de valor igual a 7, parâmetro de cor igual a 0.7 e forma igual a 0.3 isto devido ao fato dos objetos da cena estarem caracterizados mais pela cor (informação espectral). À suavidade e compacidade, após vários testes, foram dados os mesmos pesos de 0.7 e 0.3 valores “default” do “eCognition”. No segundo nível foi acrescida a imagem altimétrica gerada a partir da grade regular (Figura 2). O “eCognition” gera os segmentos do segundo nível a partir dos segmentos criados pelo primeiro nível. Portanto, com base no primeiro nível, a configuração do segundo nível utilizou uma escala com valor igual a 5, na tentativa de aglomerar mais os segmentos de um objeto, para representá-lo melhor, parâmetro de cor igual a 0.3 e forma 0.7 isto devido, neste nível, ter sido incluído a imagem altimétrica produzida a partir dos dados laser scanner. Foi mantida a mesma configuração do nível 1 para os parâmetros de suavidade e compacidade. E finalmente, um terceiro nível, com o valor de escala maior, sendo este baseado no segundo, na tentativa de aglomerar mais “pixels” e aumentar o tamanho e representatividade da região.



Figura 3 - Imagem segmentada no eCognition

4.2. AMOSTRAGEM E TREINAMENTO DOS ALGORITMOS DE CLASSIFICAÇÃO UTILIZADOS

Foram encontradas quatro classes diferentes para esta área em estudo, como já mencionado anteriormente.

Segundo GERMAN e GAHEGAN (1996), a escolha da arquitetura da rede é totalmente experimental. Porém para iniciar o processo, em um “perceptron” de multicamadas, pode-se adotar o critério da segunda camada intermediária ser 1/3 da primeira camada intermediária. Portanto foram executados testes que ajudaram na escolha da arquitetura para a melhor classificação desta imagem. Foram coletadas 15 amostras de cada classe para o treinamento deste algoritmo.

Para o treinamento do algoritmo de mínima distância foi utilizada as mesmas classes porém aumentando o número de amostras por classe, sendo 25 amostras por classe. Isso ocorre devido este algoritmo exigir um número de amostras maior para seu aprendizado.

4.3. CLASSIFICAÇÃO ORIENTADA A REGIÃO E SUA VERIFICAÇÃO

Três variáveis foram utilizadas para o treinamento dos classificadores (Redes Neurais e Distância Mínima), sendo banda pan-cromática, altimetria e intensidade. No Quadro 1 pode ser observado os resultados obtidos pelos classificadores utilizando a composição entre Altimetria, Intensidade do laser e a banda pan-cromática.

Quadro 1 – Resultado da rede neural e mínima distância para o recorte em estudo

Variável	Método	
	Redes Neurais Artificiais	Distancia Mínima
P+I+A	58.1	56.1
I+A	52.2	50.3

P: pan-cromática

I: Intensidade do laser scanner

A: Altimetria fornecida pelo laser scanner

Observando-se o Quadro1, nota-se que a rede neural apresenta um melhor resultado comparado com o método de distância mínima. Observando-se a influência da banda pan-cromática, nota-se um ganho de 5.9% quando adicionado esta banda na composição. O mesmo ocorre com o algoritmo distância mínima. Na Figura 4 pode-se observar a imagem temática produzida pelo melhor algoritmo de classificação, para esta área de estudo, a rede neural artificial.

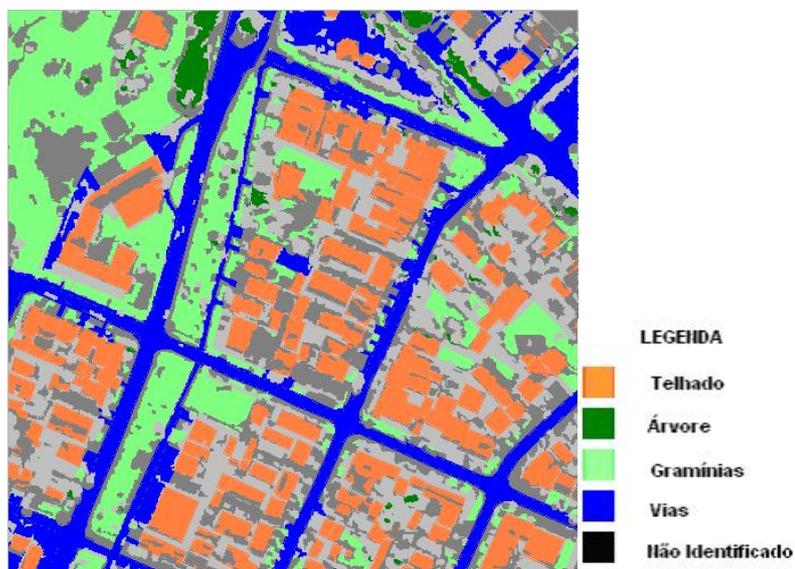


Figura 4 – Imagem temática produzida pela rede neural

6. CONSIDERAÇÕES FINAIS

Conclui-se que o uso da altimetria e intensidade derivada do laser como variável auxiliar na classificação de imagem de alta resolução espacial contribuem na produção da imagem temática final. Através da análise dos resultados obtidos por meio dos dois métodos de classificação empregados, pode-se concluir que a composição banda pan-cromática fornecida pelo sensor QUICKBIRD II, a altimetria e a intensidade fornecida pelo laser scanner possibilita a obtenção de uma boa imagem temática.

As duas metodologias de classificação empregadas tiveram um bom desempenho na classificação de imagem de alta resolução, sendo a melhor delas, a que empregou as redes neurais artificiais. Em seu treinamento, o algoritmo considera implicitamente a escala e dispersão das classes e exige poucas amostras para seu treinamento, apesar de ser muito trabalhosa em seu treinamento devido a escolha da quantidade de neurônios a ser utilizado.

Os índices encontrados do “kappa” são baixos. Como a avaliação da classificação a ser realizada compara o resultado com a digitalização dos contornos dos objetos, obtidos por restituição fotogramétrica, uma parte dos erros apontados pelo índice “kappa” corresponde a “pixels” localizados nas fronteiras dos objetos. Tais fronteiras dependem do sucesso da segmentação. A avaliação da imagem temática, portanto inclui erros na segmentação e na classificação e não pode diferencia-los.

Como a classificação orientada a regiões depende de uma prévia segmentação, a mesma deve ser a mais rigorosa possível. Percebe-se que os resultados das classificações ficaram limitados ao resultado obtido pela segmentação. Não foi possível corrigir os erros vindos da segmentação na classificação, oriundo do “eCognition”.

2. REFERÊNCIAS BIBLIOGRÁFICAS

- ANTUNES, A. F. B., **Classificação de Ambiente Ciliar Baseada em Orientação a Objeto em Imagens de Alta Resolução Espacial**. Tese de Doutorado pela Universidade Federal do Paraná, Curitiba – PR, 145 p., 2003.
- CENTENO, J. A. S. **Sensoriamento Remoto e Processamento de Imagens Digitais**. Departamento de Geomática, UFPR, 2003.
- COHEN, J. **A coefficient of agreement for nominal scales**. Educational and Psychological Measurement, v.20, n.1, p.37-46. 1960.
- GERMAN, G. W. H. e GAHEGAN, M. N. **Neural Network Architectures for the Classification of Temporal Image Sequences**. Computers & Geosciences, v.22, n.09, p.969-979, 1996.
- JARVIS, C. H.; STUART, N. **The Sensitivity of a Neural Network for Classifying Remotely Sensed Imagery**. Computers & Geosciences. V.22, n.9, p.959-967, 1996.

- MASAHARU, H; HASEGAWA, H.. **Three-dimensional city modeling from laser scanner data by extracting building polygons using region segmentation method.** International Archives fo Photogrammetry and Remote Sensing. Vol. XXXIII, Part B3. P.556-562, Amsterdam 2000.
- RIBEIRO, S. R. A.; CENTENO J. A. S. **Clasificación del uso del suelo utilizando redes neurales y el algoritmo maxver.** Innovación, Año 14 n.1, p.55-60. Antofagasta – Chile, 2002.
- WEIDNER, U.; FÖRSTNER, W. **Towards automatic building reconstruction from hilgh-resolution digital elevation models.** ISPRS Journal of Photogrammetry and Remote Sensing 50(4), p.38-49. 1995.