

INTEGRAÇÃO DE IMAGEM AÉREA DE ALTA RESOLUÇÃO E DADOS DE VARREDURA A LASER NA CLASSIFICAÇÃO DE CENAS URBANAS PARA DETECTAR REGIÕES DE VIA

*Integration of high resolution aerial image and laser scanner data in the
classification of the urban scenes to detect regions of road*

TATIANA SUSSEL GONÇALVES MENDES¹
ALUIR PORFÍRIO DAL POZ²

¹ Departamento de Engenharia Ambiental - ICT/UNESP
Av. Eng. Francisco José Longo, 777
12245-000 – São José dos Campos - SP.

² Departamento de Cartografia - FCT/UNESP
Rua Roberto Simonsen, 350
19060-900 – Presidente Prudente - SP.
tatiana.mendes@ict.unesp.br; aluir@fct.unesp.br

RESUMO

O problema de extração automática da malha viária urbana é extremamente complexo, uma vez que em cenas urbanas as vias apresentam forte interação com os outros objetos da cena (vegetação, edificações, veículos etc.). Esse problema pode ser simplificado se regiões correspondente às vias forem previamente isoladas. Na sequência, a malha viária urbana pode ser extraída baseando-se apenas nessas regiões, reduzindo a área de busca e o esforço computacional. A classificação de imagens pode ser usada no intuito de isolar as regiões de via, mas em cenas urbanas complexas a utilização de somente dados espectrais pode não ser suficiente para separar com confiabilidade classes com comportamento espectral similar. Para contornar esse problema, é proposta a integração dos dados geométricos e radiométricos de varredura a laser com imagem aérea RGB de alta resolução numa classificação por Redes Neurais Artificiais, tendo por foco principal o isolamento de regiões de via. O benefício desta integração foi verificado usando diferentes combinações de dados de entrada na rede. Os experimentos mostraram que a combinação que integra diferentes fontes de dados permitiu separar a classe via com

melhor acurácia e que problemas relacionados com as respostas espectrais similares foram minimizados.

Palavras-chave: Redes Neurais Artificiais; Modelo Digital de Superfície Normalizado; Imagem de Intensidade do Pulso Laser.

ABSTRACT

The problem of automated urban road network extraction is extremely complex, as roads in urban scenes interact strongly with other scene objects (vegetation, buildings and vehicles). This problem can be simplified if regions corresponding to roads were previously isolated using a classification procedure. Next, the urban road network can be extracted from these regions previously detected, leading to a geometric and semantic description. The image classification procedure can be used in order to isolate regions of road, but in complex urban scenes, the use of only spectral data may not be sufficient for a reliable separation of classes with similar spectral characteristics. To overcome the problem, it is proposed the integration of laser scanner geometric and radiometric data with high-resolution RGB aerial images in the classification using Artificial Neural Network, with main focus to isolate regions of road. The benefit of this integration was checked using different combinations of input data. The experiments showed that the combination that integrates different sources of data allows the separations of the road class with better accuracy and the problems related to spectral similar responses were minimized.

Keywords: Artificial Neural Network; Normalized Digital Surface Model; Laser Pulse Intensity Image.

1. INTRODUÇÃO

A extração de feições cartográficas em ambientes urbanos a partir de imagens aéreas ou orbitais possui limitações, uma vez que as cenas urbanas apresentam objetos dispostos de forma aglomerada e que se relacionam diretamente entre si. As vias, por exemplo, apresentam obstruções provocadas por veículos, copas de árvores, projeções de edificações e sombras de objetos que estão ao longo da mesma. Uma forma de simplificar o problema é detectar primeiramente regiões que correspondem às vias. Essas regiões podem servir de base para uma descrição semântica detalhada das vias, tendo como consequência, a diminuição da área de busca e do esforço computacional no subsequente processo de extração. A classificação de imagens destaca-se como importante método, no sentido de detectar previamente essas regiões, como pode ser visto em Schackelford e Davis (2003), Benediktsson et al. (2003), Chanussot et al. (2006) e Bellens et al. (2008), cujos trabalhos atuam no contexto de cenas urbanas usando imagens de alta resolução.

Entretanto, como destacam Bellens et al. (2008) e Alonso e Malpica (2010), a utilização de apenas informação espectral para classificar imagens de cenas urbanas podem não ser suficientes para separar classes com características espectrais

semelhantes, como por exemplo, vias e edificações de telhado cinza. Em vista desta limitação, alguns métodos têm incorporado informações adicionais no processo de classificação, como por exemplo, informações provenientes dos sistemas de varredura a laser. Trabalhos que combinam informações proveniente desse sistema com imagens multiespectrais para melhorar o resultado da classificação podem ser encontrados em Haala e Walter (1999), Zeng et al. (2002), Botelho e Centeno (2005), Alonso e Malpica (2010), Trinder e Salah (2011) e Makarau et al. (2011).

Vários métodos usam a classificação de imagens para detectar regiões de via ou como primeiro passo para extrair a malha viária, principalmente quando se trata de áreas urbanas complexas (Grote & Rottensteiner, 2010; Shackelford & Davis, 2003). Exemplos são apresentados na sequência. Gao e Wu (2004) usaram imagens Ikonos numa classificação não-supervisionada, cuja classe correspondente às vias passa por um processo de remoção de ruídos, ligação de segmentos e esqueletização. Zhang e Wang (2004) classificaram imagens *pan-sharpened* QuickBird e a imagem binária que representa a classe via foi segmentada usando informações de bordas de modo a desconectar regiões de não-via. Essas regiões foram removidas baseando-se em critérios de forma e por meio de um algoritmo de filtragem de segmentos. Mokhtarzade e Valadan Zoj (2007) apresentam uma estratégia de detecção de vias em imagens Ikonos e Quickbird por meio de uma classificação por Redes Neurais Artificiais (RNA), com a verificação de influência de vários parâmetros (textura, informações de bordas etc.) como entrada na rede, resultando em duas classes: vias e não vias. Em Benkouider et al. (2011) foi desenvolvido um método para separar a classe via em imagens SPOT RGB, usando características espectrais de vias em um processo de classificação por RNA, com pós-processamento morfológico para regularizar as vias previamente classificadas. Das et al. (2011) apresentaram um método de detecção de vias em imagens multiespectrais de alta resolução baseado em dois estágios. No primeiro, regiões candidatas às vias foram detectadas usando uma classificação por SVM e atributos salientes de vias (propriedades de contraste e de trajetória linear). No segundo, módulos de pós-processamento foram usados para refinar a malha viária, de modo a conectar segmentos e remover falsos positivos. A integração de dados de varredura a laser e imagens multiespectrais, num processo de classificação baseado no algoritmo de aprendizado de máquina AdaBoost (abreviação de *Adaptive Boosting* – estímulo adaptativo), pode ser encontrada em Mancine et al. (2009), resultando em 4 classes. Os autores usaram a classe vias para extrair rotatórias e vias lineares usando, respectivamente, a transformada de *Hough* e um método de segmentação por crescimento de regiões modificado.

Como é notável nos trabalhos relacionados, pesquisas que incorporam dados do sistema de varredura a laser no processo de extração de informações a partir de imagens tem se intensificado nos últimos anos, principalmente devido à evolução tecnológica do sistema. A imagem de intensidade do pulso laser é, em geral, pouco utilizada no processo de extração de feições, fato decorrente de sua característica ruidosa e de algumas anomalias nas respostas de alguns alvos, quando compara-se a

imagem de intensidade com uma imagem do infravermelho próximo (Jensen, 2009). As vias nesse tipo de dado apresentam características peculiares que podem favorecer o processo de extração da malha viária.

A hipótese central deste artigo é que as informações geométricas e radiométricas dos dados de varredura a laser e alta definição nas imagens aéreas de alta resolução fornecem informação complementar de uma cena, e quando estes são usados conjuntamente podem aumentar a acurácia e a confiabilidade da detecção e extração da malha viária, sobretudo em ambientes urbanos complexos. O objetivo é avaliar essa hipótese por meio de uma verificação usando diferentes combinações de dados de entrada numa classificação por RNA, tendo por foco principal o isolamento da classe via. As imagens usadas como entradas na rede são: R, G, B, Modelo Digital de Superfície normalizado (MDSn), MDSn limiarizado, imagem de intensidade, imagem de intensidade suavizada pela mediana e imagem de intensidade suavizada com Morfologia Matemática (MM).

Neste trabalho, a classificação por RNA é utilizada justamente pela sua capacidade de integrar dados de diferente natureza, que geralmente apresentam-se em escalas diferentes e não necessariamente apresentam natureza espectral. Essa integração de dados de diferente natureza no processo de classificação de imagens por RNA é também tema de alguns trabalhos (ANDRADE & CENTENO, 2003; BOTELHO & CENTENO, 2005; MOKHTARZADE & VALADAN ZOEJ, 2007; FERREIRA & GALO, 2010; E BENKOUIDER et al., 2011) que visam obter uma maior qualidade no resultado da classificação, em comparação quando se usa somente uma fonte de informação.

Este artigo está organizado em quatro seções principais. A Seção 2 apresenta os fundamentos teóricos necessários para o desenvolvimento do método proposto. A Seção 3 descreve o material e métodos. Os resultados experimentais são apresentados e discutidos na Seção 4. A Seção 5 apresenta as conclusões e as perspectivas futuras.

2. FUNDAMENTOS TEÓRICOS

2.1 Dados de Varredura a Laser

O sistema de varredura a laser é um sistema de sensor ativo capaz de coletar uma grande quantidade de pontos com coordenadas tridimensionais de elevada acurácia altimétrica. O sistema integra basicamente três tecnologias: o INS (*Inertial Navigation System* - Sistema de Navegação Inercial), o GNSS (*Global Navigation Satellite System* – Sistemas Globais de Navegação por Satélite) e o instrumento de varredura a laser (EL-SHEIMY et al., 2005), que consiste de um gerador de pulsos laser, um sistema de controle e armazenamento e um conjunto óptico de transmissão e recepção do detector de sinais.

Esse sistema, instalado numa aeronave ou helicóptero, emite com o auxílio de um espelho de varredura, um feixe de pulso laser em direção à superfície terrestre. A energia emitida interage com os alvos da superfície e parte dela é retornada em

direção ao sistema, registrando pelo receptor a intensidade e o tempo decorrido entre a emissão e a recepção do pulso laser (JENSEN, 2009). O tempo registrado é usado para calcular a distância entre o sensor e alvo irradiado pelo pulso laser (DALMOLIN & SANTOS, 2004).

No instante em que cada feixe de pulso laser é emitido e recebido, a posição e a orientação desse feixe também são determinadas. A posição é determinada mediante um sistema de GPS Diferencial (DGPS) e a orientação por um sistema de unidade de medida inercial (IMU – *Inertial Measurement Unit*).

Com base nos dados geométricos obtidos pelo sistema de varredura a laser (distância e parâmetros de posição e orientação do feixe de laser) são estimadas as coordenadas tridimensionais de cada ponto sobre a superfície irradiada pelo sensor. O sistema também pode registrar a intensidade da energia retornada (dados radiométricos) após a iteração do pulso de laser com o alvo irradiado. O valor da intensidade está relacionado com a capacidade do alvo de refletir ou absorver o pulso laser.

Os pontos de elevação obtidos pelo sistema de varredura a laser apresentam-se irregularmente distribuídos. Esses pontos podem ser organizados de duas maneiras. Uma é através de uma estrutura de dados composta por uma rede irregular de triângulos (TIN – *Triangular Irregular Network*) e a outra é a partir de uma grade regular usando algum método de interpolação. Como cada ponto de elevação possui o registro da intensidade do retorno do pulso laser, uma grade regular também pode ser gerada, permitindo a sua visualização como uma imagem.

2.2 Classificação por Redes Neurais Artificiais

Dentre as diversas técnicas de classificação usadas para análise e extração de informações a partir de imagens de Sensoriamento Remoto, as RNA tem se destacado nos últimos anos, tendo como uma das vantagens a não necessidade de que as propriedades estatísticas das classes de informação sejam estimadas. Outra vantagem dessa técnica é que permite integrar dados de diferente natureza no processo de classificação.

As RNA são compostas de unidades simples de processamento, os chamados neurônios artificiais, os quais foram inicialmente construídos com base no funcionamento biológico de um neurônio (GONZALEZ & WOODS, 2010). O modelo de rede *Multilayer Perceptron* consiste de um conjunto de neurônios arranjados em múltiplas camadas, as quais são: camada de entrada, uma ou mais camadas ocultas e camada de saída. Com exceção da primeira camada, onde os padrões de entrada são inseridos na rede, cada neurônio é conectado a todos os neurônios da próxima camada através de pesos, produzindo um somatório. Assim, a entrada total de cada neurônio é uma soma ponderada das saídas dos neurônios da camada anterior.

Para o treinamento da rede, o algoritmo mais comum é o de retropropagação (*Backpropagation*), o qual é designado para minimizar o erro entre a saída calculada por uma arquitetura e a saída desejada da rede (SCHOWENGERDT, 2007).

A função erro é apresentada na Equação 1 (GONZALEZ & WOODS, 2010).

$$E = \frac{1}{2} \sum_{j=1}^L (d_j - o_j)^2 \quad (1)$$

sendo que E é o erro quadrático entre as saídas desejadas (d_j) e as saídas calculadas (o_j) nos neurônios da camada de saída. A apresentação de um padrão de treinamento na entrada da rede permite saber qual deveria ser a resposta desejada d_j de cada neurônio na camada de saída. L é o número de neurônios na camada de saída.

O processo de treinamento se inicia com um conjunto arbitrário de pesos para as conexões entre os neurônios e envolve duas fases. Na primeira, um vetor de treinamento com a saída desejada (d_j) é apresentado à rede e propagado através de suas camadas para obter a saída calculada (o_j) para cada neurônio. As saídas o_j dos neurônios na camada de saída são então comparadas com as saídas desejadas d_j , para serem obtidos os termos de erros. A segunda fase envolve um caminho de volta pelas camadas da rede durante o qual o erro é repassado para cada neurônio, alterando os pesos das conexões entre os neurônios. Essa fase de treinamento termina quando os pesos estiverem devidamente ajustados de modo que o valor da função erro (Equação 1) seja minimizada. Assim, num treinamento bem sucedido, o erro diminui com o número de iterações e o procedimento converge para um conjunto estável de pesos (GONZALEZ & WOODS, 2010).

Uma vez a rede treinada, é possível classificar os padrões não conhecidos utilizando o conjunto de pesos ajustados na fase de treinamento (GALO, 2000). Nesta etapa, um padrão de entrada é propagado pelas camadas da rede e é classificado como pertencente à classe cujo neurônio de saída apresentar resposta mais alta, enquanto que os demais neurônios apresentam respostas mais baixa (GONZALEZ & WOODS, 2010).

2.3. Morfologia Matemática

Gonzalez & Woods (2010) definem Morfologia Matemática (MM) como uma ferramenta para extrair componentes das imagens que são úteis na representação e na descrição da forma de uma região. Ainda, segundo os autores, a MM também pode ser usada como técnica de pré-processamento e pós-processamento, por exemplo, a filtragem morfológica.

Os operadores fundamentais na MM são erosão e dilatação, os quais são aplicados em uma imagem usando um conjunto de tamanho e forma conhecida, denominado elemento estruturante. Outros operadores morfológicos podem ser expressos em termos de erosão e dilatação, como por exemplo, a abertura e o fechamento (SOILLE, 1999).

Nas imagens em níveis de cinza, a operação de abertura elimina padrões claros da imagem, menores que o elemento estruturante. A operação dual é o fechamento, cujo operador remove padrões escuros. No entanto, as operações de abertura e fechamento podem causar deformações dos objetos presentes na cena (BELLENS et

al., 2008). Para evitar esse problema pode-se utilizar a abertura e o fechamento por reconstrução, que permite remover detalhes espúrios da imagem preservando a forma dos objetos mais relevantes.

A reconstrução morfológica envolve duas imagens e um elemento estruturante. Uma imagem, denominada de marcadora, contém os pontos de partida para a transformação, enquanto que a outra, denominada de máscara, restringe a transformação. O elemento estruturante é usado para definir a conectividade. Na reconstrução morfológica são usados os conceitos de dilatação e erosão geodésicas (GONZALEZ & WOODS, 2010).

A abertura por reconstrução consiste em uma erosão da imagem de entrada, cujo resultado é usado como imagem marcadora, seguida de uma reconstrução morfológica por dilatação. Ao contrário de uma abertura elementar, a abertura por reconstrução preserva a forma dos objetos que não foram removidos na erosão. Os objetos que permaneceram após a erosão são então usados como sementes para a reconstrução da imagem original. De forma dual, o fechamento por reconstrução consiste em uma dilatação da imagem original seguida por uma reconstrução morfológica por erosão.

Nas imagens em níveis de cinza, os máximos (detalhes mais claros) e os mínimos (detalhes mais escuros) revelam detalhes importantes na imagem (SOILLE, 1999). Em MM, os máximos e os mínimos são usados no sentido regional, ou seja, não é necessariamente restrito a um único pixel. Uma forma de filtrar esses extremos da imagem é usando as transformações $hmax$ e $hmin$, que removem os máximos e mínimos entre regiões planas, unindo as regiões adjacentes que possuem níveis de cinza similares. Estas transformações são baseadas em reconstrução morfológica por dilatação e erosão, sendo que, quando se usa um elemento estruturante com conectividade 4 (por exemplo, em forma de cruz) tem-se mais regiões planas do que quando se usa com conectividade 8 (por exemplo, em forma de máscara 3x3).

A transformação $hmax$ suprime todos os máximos cuja profundidade é menor ou igual a um determinado limiar h e isso se consegue a partir de uma reconstrução por dilatação de f (R_f^δ) a partir de $f - h$, como apresenta a Equação 2 (SOILLE, 1999).

$$hmax_h(f) = R_f^\delta(f - h), \quad (2)$$

onde f representa a imagem original.

Da mesma forma, uma transformação $hmin$ suprime todos os mínimos que apresentam uma profundidade inferior ao limiar h e é definida por (Soille, 1999):

$$hmin_h(f) = R_f^\varepsilon(f + h), \quad (3)$$

onde R_f^ε representa a reconstrução por erosão de f .

3. MATERIAL E MÉTODOS

3.1 Material

A área de estudo corresponde a uma região urbana da cidade de Curitiba-PR, a qual se caracteriza por vias asfaltadas, estacionamentos, edificações e vegetações de diferentes tamanhos (áreas gramadas, arbustos e árvores de grande porte). Para esta área tem-se disponível uma imagem aérea RGB de alta resolução. A Figura 1 apresenta um recorte da imagem RGB, destacando algumas dificuldades a serem contornadas na classificação.

Figura 1 - Recorte da imagem RGB.



- A) Projeção de sombra de edificações e árvores nas vias modificando o comportamento espectral nessas áreas;
- B) Edificações com comportamento espectral similar ao das vias; e
- C) Copas de árvores causando obstrução total das vias.

Para a mesma região da imagem RGB tem-se os dados geométricos e radiométricos de varredura a laser, composto de um arquivo com as coordenadas (E,N,H) , no sistema de referência WGS 84, e com o atributo I correspondente ao valor da intensidade registrada pelo sensor. O registro das diferentes fontes de dados é realizado a partir da projeção dos dados de varredura a laser na imagem RGB de alta resolução, através da equação de colinearidade, adição das distorções relacionadas ao sistema de lentes e em seguida, a transformação entre referenciais no espaço imagem.

Para o desenvolvimento do trabalho são usados os seguintes *softwares*: SURFER da *Golden Software*, para a interpolação dos dados 3D de varredura a laser; HALCON da *MVTec Software GmbH* e SDC *Morphology Toolbox* para Matlab da *SDC Information Systems*, para os processamentos das imagens

resultantes dos dados de varredura a laser; e IDRISI Andes desenvolvido pela *Clark University*, para a classificação por RNA.

3.2 Pré-Processamento dos Dados Laser

O processo de geração do MDSn é realizado a partir da diferença entre o MDS e o Modelo Digital de Terreno (MDT), conforme descrito em Botelho e Centeno (2005). O MDS é gerado a partir da interpolação do conjunto de pontos 3D que correspondem aos dados geométricos de varredura a laser, resultando numa grade regular. A grade é gerada com espaçamento de 30 cm (mesma resolução espacial da imagem RGB) e a partir do método de interpolação Vizinheiro mais próximo, devido a sua característica de não causar suavização das bordas dos objetos, mantendo-as bem definidas.

A partir do MDS, pontos que representam a superfície do terreno são coletados manualmente e usando o método de interpolação Inverso da distância é gerada outra grade regular, que representa o MDT. Este método de interpolação possui a característica de suavizar a superfície, que é mais adequada para representar o MDT.

O MDSn é então obtido subtraindo o MDT do MDS, resultando numa representação dos objetos elevados em uma superfície plana. A grade que representa o MDSn é convertida em uma imagem digital, cujos níveis de cinza são proporcionais à altitude dos pontos. O MDSn é apresentado na Figura 2(a), onde observa-se que os objetos mais elevados (árvores e edificações) aparecem em tonalidade mais clara, enquanto que os objetos próximos do nível do terreno (vias, calçadas, estacionamentos, praças e áreas gramadas) apresentam-se em tonalidade bem escura.

A partir da análise do histograma da imagem que representa o MDSn é selecionado um limiar T , cujo valor separa os objetos elevados dos restantes, resultando em uma imagem em dois níveis denominada, no presente trabalho, de MDSn limiarizado. Assim, sendo $f(x, y)$ o nível de cinza num pixel (x, y) pode-se definir uma imagem limiarizada $G(x, y)$ como (Gonzalez & Woods, 2010):

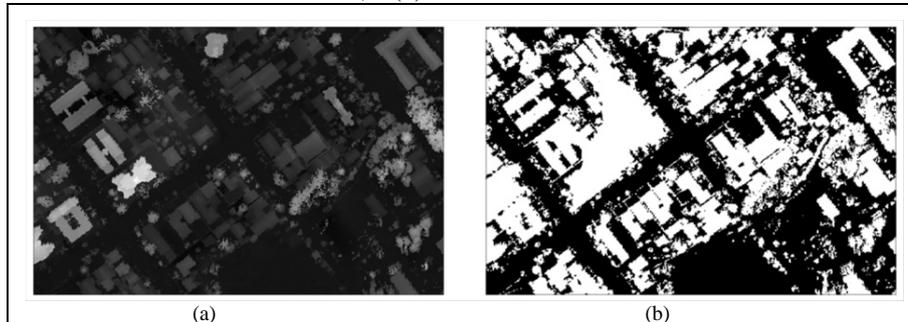
$$G(x, y) = \begin{cases} 255 & \text{se } f(x, y) > T \\ 0 & \text{se } f(x, y) \leq T \end{cases} \quad (4)$$

A Figura 2(b) mostra o MDSn limiarizado, com o valor do limiar T igual a 30, equivalente à aproximadamente 2m de altura. Esse limiar foi escolhido considerando que, em geral, uma residência apresenta um pé-direito mínimo de 2,5m.

Os dados radiométricos correspondem à imagem de intensidade de pulso laser, onde as vias são facilmente identificadas, uma vez que possuem resposta espectral homogênea e aparecem em tonalidade escura devido ao asfalto apresentar baixa reflectividade (em torno de 17%, segundo Wehr e Lohr (1999)) em relação ao pulso laser. A Figura 3(a) apresenta a imagem de intensidade do pulso laser. Na região em

destaque observam-se ruas estreitas bem identificáveis, as quais na imagem RGB (Figura 1) não podem ser visualizadas, pois estão obstruídas por árvores. Outro aspecto importante é em relação às sombras dos objetos, que principalmente nas vias, não causam perturbações. Neste sentido acredita-se que a imagem de intensidade pode contribuir significativamente na complementação da informação no processo de classificação de imagens.

Figura 2 - Imagens geradas a partir dos dados geométricos de varredura a laser: (a) MDSn; e (b) MDSn limiarizado.



No entanto, a imagem de intensidade apresenta alguns pontos fracos. A presença de objetos que apresentam resposta radiométrica similar às das vias, como algumas árvores e edificações, é um deles. Outro problema é em relação à característica ruidosa que afeta diretamente o resultado da classificação. Para minimizar esse problema é aplicado o filtro de mediana usando uma máscara de dimensão 5x5. O filtro de mediana substitui o valor de um pixel pela mediana dos valores de intensidade na vizinhança desse pixel. A vizinhança é definida pela dimensão da máscara (Gonzalez & Woods, 2010). A Figura 3(b) mostra a imagem de intensidade suavizada pela mediana.

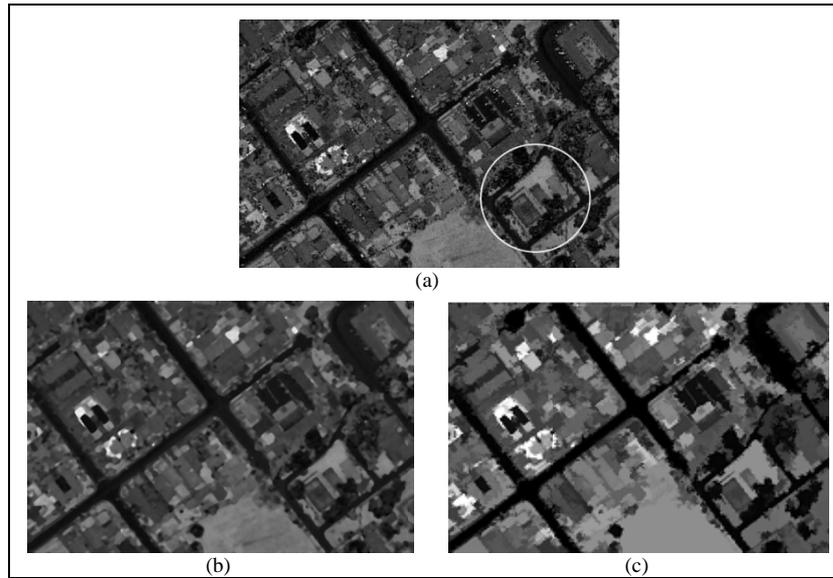
Além da aplicação do filtro mediana, a imagem de intensidade também é suavizada usando MM, de modo a preservar e realçar as regiões mais escuras, principalmente aquelas correspondentes às vias, e suavizar regiões mais claras da imagem. A SDC *Morphology Toolbox* para Matlab é usada nesse processo, cuja sequência de operações é descrita a seguir.

- Aplicação do operador 'mmhmax' na imagem original, usando um limiar h igual a 70. Nesta operação, as regiões mais escuras não são alteradas e as regiões mais claras são escurecidas e os detalhes reduzidos, ou seja, apresentaram-se mais homogêneas.
- Aplicação de fechamento com reconstrução morfológica usando o operador 'mmcloserec', com o elemento estruturante disco de raio igual a 5 pixels. Neste caso, pequenas regiões escuras (menores que o elemento estruturante) são eliminadas.

- Aplicação do operador ‘mmhmin’ com limiar h igual a 20. Nesta operação as regiões mais claras não são alteradas e as regiões mais escuras tornaram-se mais homogêneas e levemente mais claras, devido a usar um valor baixo para o limiar h .
- Aplicação de abertura com reconstrução, usando o operador ‘mmopenrec’, com elemento estruturante disco com raio igual a 5 pixels. Esta operação escurece pequenas regiões mais claras.

A Figura 3(c) apresenta a imagem de intensidade suavizada com MM, após as etapas descritas anteriormente, resultando em vias mais homogêneas e redução de detalhes no interior das quadras.

Figura 3 – Dados radiométricos de varredura a laser: (a) Imagem de intensidade; (b) Imagem de intensidade suavizada com filtro mediana; e (c) Imagem de intensidade suavizada usando MM.



3.3 Classificação Supervisionada por RNA

Primeiramente são definidas as classes de informação, as quais são: ‘via asfaltada’, ‘grama’, ‘árvore’, ‘edificação de telhado cinza’, ‘edificação de telhado vermelho’ e ‘concreto’. Para cada uma dessas classes são coletados elementos amostrais de treinamento sobre áreas representativas da classe, caracterizando uma classificação supervisionada.

Para a classificação por RNA é usado o algoritmo de treinamento *Backpropagation*, disponível no *software* IDRISI. Para a taxa de aprendizado é adotado o valor igual a 0,01, com inserção do fator momentum com valor igual a

0,5, para otimizar a convergência da rede. Como dados de entrada na rede, são propostas diversas combinações de dados, como mostra a Tabela 1. As imagens RGB são combinadas com imagens resultantes dos dados de varredura a laser (Seção 3.2). Todos os dados de entrada na rede foram escalonados entre 0 e 255.

Conforme a Tabela 1, a primeira combinação usa as imagens RGB e visa analisar e comparar a contribuição da posterior adição dos dados de varredura a laser no processo de classificação. Na segunda e terceira combinação, os dados geométricos de varredura a laser são adicionados de duas formas: uma usando a imagem que representa o MDSn (combinação RGB+MDSn) e a outra usando o MDSn limiarizado (combinação RGB+MDSnLim), verificando qual delas melhor contribui para discriminar os objetos elevados daqueles que se encontram no mesmo nível do terreno. A imagem de intensidade é adicionada às imagens RGB de diferentes formas: sem processamento (combinação RGB+I), suavizada com filtro de mediana (combinação RGB+IS) e suavizada com MM (combinação RGB+ISM). Nestas combinações, o alto contraste das vias nas imagens de intensidade influencia o resultado da classificação para a classe 'via asfaltada'. As duas últimas combinações, apresentadas na Tabela 1, são propostas de modo a integrar as imagens RGB e ambos os dados de varredura a laser (combinações RGB+MDSnLim+IS e RGB+MDSnLim+ISM).

Tabela 1 – Combinações de dados de entrada na classificação por RNA.

| Combinações | Dados de entrada |
|---------------------|--|
| RGB | imagens RGB |
| RGB + MDSn | imagens RGB imagem MDSn |
| RGB + MDSnLim | imagens RGB imagem MDSn limiarizado (MDSnLim) |
| RGB + I | imagens RGB imagem de intensidade (I) |
| RGB + IS | imagens RGB imagem de intensidade suavizada usando filtro de mediana (IS) |
| RGB + ISM | imagens RGB imagem de intensidade suavizada usando MM (ISM) |
| RGB + MDSnLim + IS | imagens RGB imagem MDSn limiarizado imagem de intensidade suavizada usando filtro de mediana |
| RGB + MDSnLim + ISM | imagens RGB imagem MDSn limiarizado imagem de intensidade suavizada usando MM |

A arquitetura da rede para cada combinação é testada variando o número de camadas ocultas e o número de neurônios dentro dessas camadas. Após o processo

de treinamento, foram escolhidas as arquiteturas que apresentaram: 1) menor diferença entre o valor do Erro Médio Quadrático (EMQ) de treinamento e de teste; 2) valor alto para a acurácia atingida; e 3) diminuição das oscilações entre as curvas que representam o EMQ de treinamento e de teste, conforme aumenta o número de iterações.

Para cada arquitetura escolhida, a combinação de dados de entrada é classificada usando os parâmetros da rede treinada e a partir da análise visual da imagem classificada é verificado se a classificação foi satisfatória ou não, principalmente em relação à classe ‘via asfaltada’.

3.4 Avaliação dos Resultados da Classificação

Para avaliar numericamente os resultados obtidos da classificação é necessário obter dados de referência (verdade terrestre). Esses dados são obtidos a partir da digitalização de polígonos sobre áreas correspondentes a cada classe de informação, obtendo elementos amostrais de validação. A estratégia de amostragem aleatória estratificada com 100 pontos é usada de modo a definir, para cada ponto gerado, o local onde as amostras devem ser coletadas. Esta estratégia permite a aleatoriedade de coleta de amostras em estratos que cobrem toda a área.

A validação da classificação é realizada a partir da tabulação cruzada entre os dados de referência (pixels de referência) e o resultado da classificação, cujo procedimento é realizado pontualmente, ou seja, os pixels dos dados de referência são comparados com os seus correspondentes, resultantes da classificação. Esse procedimento é realizado no próprio *software* IDRISI, gerando a matriz de confusão para cada classificação. A partir dessa matriz obtém-se o coeficiente kappa, que envolve toda a matriz em seu cálculo e mede a concordância entre os dados classificados e os dados de referência. Quanto maior o valor do coeficiente kappa, melhor o desempenho da classificação. Se todas as classes forem corretamente classificadas, o valor de kappa seria igual a 1 (TSO & MATHER, 2009).

Também a partir da matriz de confusão se obtém para cada classe de informação os índices de acurácias do usuário e do produtor e os erros de inclusão e de omissão. Desta forma, é possível avaliar separadamente a classe ‘via asfaltada’, foco do presente trabalho, e analisar a contribuição das demais classes para degradar o resultado para esta classe de interesse.

4. RESULTADOS E DISCUSSÕES

A classificação por RNA foi realizada testando diferentes arquiteturas de rede para cada uma das combinações de dados de entrada descritas na Tabela 1. A Tabela 2 mostra, para cada combinação, a arquitetura que apresentou melhor desempenho, juntamente com o coeficiente kappa para a classificação geral e as acurácias do usuário (AU) e do produtor (AP) para a classe ‘via asfaltada’, todos obtidos na etapa de validação da classificação.

A descrição da arquitetura apresentada na Tabela 2 indica que, por exemplo, para a combinação RGB+MDSn tem-se 4 neurônios na camada de entrada (imagens

R, G, B e MDSn), duas camadas ocultas com 24 e 40 neurônios, respectivamente, na 1ª e na 2ª camada e 6 neurônios na camada de saída (classes de informação). A partir da Tabela 2 verifica-se que os maiores valores para o coeficiente kappa ocorreram para as combinações RGB+MDSnLim, RGB+MDSnLim+IS e RGB+MDSnLim+ISM. Destas combinações, aquelas que integram as imagens RGB com ambos os dados de varredura a laser atingiram valores maiores que 90% para AU e AP para a classe 'via asfaltada'. O resultado de cada combinação é discutido a seguir, analisando as imagens geradas a partir da classificação, a matriz de confusão e os dados apresentados na Tabela 2.

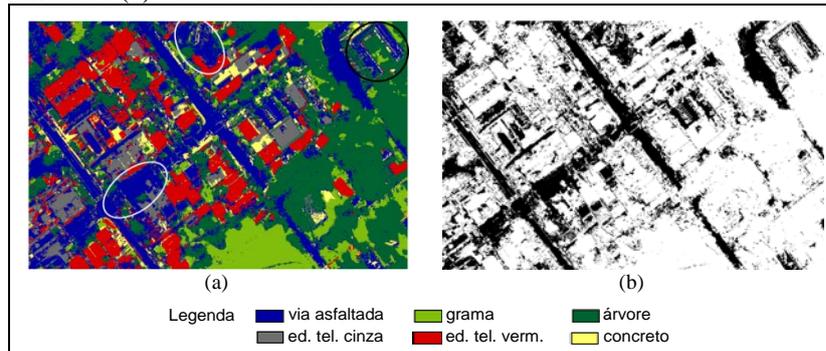
Tabela 2 – Arquitetura com melhor desempenho, coeficiente kappa e as acurácias do usuário e do produtor para a classe 'via asfaltada' para cada combinação.

| Dados de entrada | Arquitetura | Kappa | Classe 'via asfaltada' | |
|---------------------|-----------------|-------|------------------------|-------|
| | | | AU | AP |
| RGB | 3 – 21 – 40 – 6 | 0,67 | 68,71 | 88,76 |
| RGB + MDSn | 4 – 24 – 40 – 6 | 0,75 | 79,56 | 96,01 |
| RGB+MDSnLim | 4 – 23 – 46 – 6 | 0,80 | 83,01 | 96,03 |
| RGB + I | 4 – 15 – 45 – 6 | 0,75 | 85,99 | 93,88 |
| RGB + IS | 4 – 45 – 15 – 6 | 0,74 | 89,37 | 95,13 |
| RGB+ISM | 4 – 37 – 45 – 6 | 0,76 | 91,06 | 96,25 |
| RGB + MDSnLim + IS | 5 – 30 – 50 – 6 | 0,80 | 90,69 | 96,17 |
| RGB + MDSnLim + ISM | 5 – 20 – 27 – 6 | 0,79 | 94,53 | 94,56 |

O resultado da classificação usando a combinação RGB é mostrado na Figura 4(a). A partir da análise visual verifica-se que áreas registradas como sombras na imagem colorida foram classificadas como 'via asfaltada' (exemplos destacados em branco). Algumas edificações de telhado cinza foram classificadas como 'via asfaltada', como por exemplo, a edificação em destaque em preto. Analisando somente a classe 'via asfaltada' (Figura 4(b)), observa-se muitas falhas e grande quantidade de pixels classificados erroneamente no interior das quadras. Esta combinação obteve os menores valores para as AU e AP para a classe 'via asfaltada', respectivamente, iguais a 68,71% e 88,76%, conforme apresenta a Tabela 2.

A partir da matriz de confusão (Tabela 3), tem-se que os maiores erros de inclusão ocorreram para as classes 'ed. telhado cinza', 'concreto' e 'via asfaltada', justamente porque estas classes apresentam comportamentos espectrais similares.

Figura 4 – Resultado da classificação usando a combinação RGB. (a) Imagem classificada e (b) Classe ‘via asfaltada’.



O erro de inclusão para a classe ‘via asfaltada’ apresentou maior contribuição justamente da classe ‘ed. telhado cinza’ com 1145 pixels (16,2%). Os maiores erros de omissão ocorreram para as classes ‘ed. telhado cinza’ e ‘concreto’. Para a classe ‘ed. telhado cinza’, 1145 pixels (43%), pertencentes a essa classe, foram incorretamente classificados como ‘via asfaltada’.

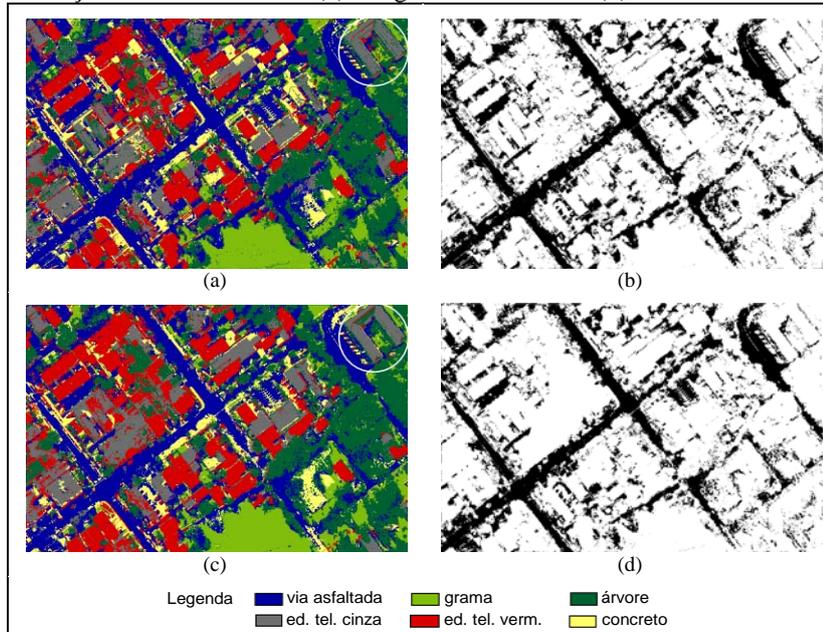
Tabela 3 - Matriz de confusão resultante da combinação RGB.

| Classificação por RNA | Dados de referência | | | | | | | TOTAL | Erro inclusão |
|-----------------------|---------------------|-------|--------|-------------------|-------------------|----------|-------|-------|---------------|
| | via asfaltada | grama | árvore | ed. telhado cinza | ed. telhado verm. | concreto | | | |
| via asfaltada | 4847 | 79 | 582 | 1145 | 192 | 209 | 7054 | 0,31 | |
| grama | 13 | 928 | 183 | 8 | 4 | 5 | 1141 | 0,19 | |
| árvore | 275 | 109 | 3092 | 204 | 7 | 66 | 3753 | 0,18 | |
| ed. tel. cinza | 318 | 0 | 0 | 901 | 58 | 244 | 1521 | 0,41 | |
| ed. tel. verm. | 8 | 38 | 41 | 250 | 2617 | 59 | 3013 | 0,13 | |
| concreto | 0 | 0 | 0 | 154 | 56 | 386 | 596 | 0,35 | |
| TOTAL | 5461 | 1154 | 3898 | 2662 | 2934 | 969 | 17078 | | |
| Erro omissão | 0,11 | 0,20 | 0,21 | 0,66 | 0,11 | 0,60 | | | |

A Figura 5 apresenta os resultados da classificação usando as combinações RGB+MDSn e RGB+MDSnLim. Fazendo uma comparação visual entre os resultados apresentados na Figura 5 e o resultado da Figura 4 (combinação RGB), é possível observar a diminuição das edificações classificadas como ‘via asfaltada’ e um aumento dos pixels classificados como ‘concreto’. As imagens MDSn e MDSnLim contribuem para diminuir a confusão entre as classes ‘via asfaltada’ e ‘ed. telhado cinza’, separando possíveis elementos com respostas espectrais semelhantes e que estão em níveis de alturas diferentes, por exemplo, a edificação

de telhado cinza em destaque branco nas Figuras 5(a) e 5(c). De uma forma geral, verifica-se a diminuição dos pixels classificados como 'via asfaltada' no interior das quadras para ambas as combinações (Figuras 5(b) e 5(d)), sendo que agora, a maior parte desses pixels representam regiões de sombras dos objetos projetadas sobre a superfície no nível do terreno. A melhora do resultado da classificação pode ser confirmada a partir dos valores do coeficiente kappa (Tabela 2), que aumentou de 0,67 (combinação RGB) para 0,75 (combinação RGB+MDSn) e para 0,80 (combinação RGB+MDSnLim) e o aumento dos valores para as AU e AP para a classe 'via asfaltada'. As AP apresentaram valores próximos para ambas as combinações, já a AU apresentou valor maior para a combinação RGB+MDSnLim.

Figura 5 – Resultado da classificação adicionando os dados geométricos do laser. Combinação RGB+MDSn: (a) Imagem classificada e (b) Classe 'via asfaltada'; e Combinação RGB+MDSnLim: (c) Imagem classificada e (d) Classe 'via asfaltada'.



Como na imagem MDSn, cada pixel possui um nível de cinza relativo à sua altura real, os objetos mais elevados são mais propícios de serem classificados corretamente. A partir das análises visuais e numéricas (Tabela 2) verificou-se que a combinação RGB+MDSnLim apresentou melhor resultado quando comparado com a combinação RGB+MDSn e, por isso a imagem MDSnLim representará os dados geométricos de varredura a laser, nas futuras combinações que integram ambos os dados de varredura a laser.

Comparando a matriz de confusão resultante da combinação RGB (Tabela 3) com a da combinação RGB+MDSnLim (Tabela 4), verifica-se que os erros de inclusão diminuíram significativamente, com exceção da classe 'concreto'. Em relação ao erro de inclusão para a classe 'via asfaltada', tem-se maior confusão com a classe 'árvore', com 677 pixels (10,2%) classificados erroneamente. Isso pode ser em decorrência da variação da dimensão das copas das árvores registradas na imagem colorida e nos dados de varredura a laser, principalmente porque os dados não correspondem à mesma data. O maior erro de omissão ocorreu para a classe 'concreto', com 94 pixels (9,7%) pertencentes a essa classe e que foram incluídos erroneamente na classe 'via asfaltada'.

Tabela 4 - Matriz de confusão resultante da combinação RGB+MDSnLim.

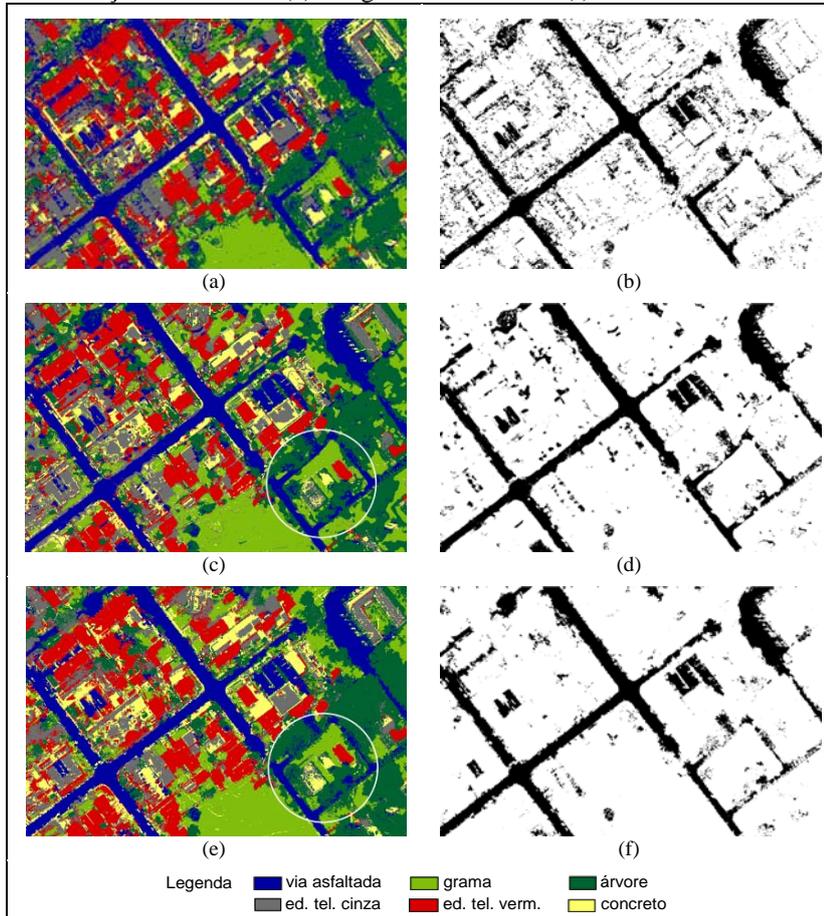
| | | Dados de referência | | | | | | | |
|-----------------------|----------------|---------------------|-------------|-------------|-------------------|-------------------|------------|--------------|---------------|
| Classificação por RNA | | via asfaltada | grama | árvore | ed. telhado cinza | ed. telhado verm. | concreto | TOTAL | Erro Inclusão |
| | | via asfaltada | 5244 | 118 | 677 | 174 | 10 | 94 | 6317 |
| | grama | 3 | 920 | 95 | 6 | 0 | 5 | 1029 | 0,11 |
| | árvore | 21 | 47 | 2958 | 90 | 1 | 26 | 3143 | 0,06 |
| | ed. tel. cinza | 42 | 0 | 94 | 1894 | 68 | 208 | 2306 | 0,18 |
| | ed. tel. verm. | 18 | 48 | 70 | 367 | 2832 | 81 | 3416 | 0,17 |
| | concreto | 133 | 21 | 4 | 131 | 23 | 555 | 867 | 0,36 |
| | TOTAL | 5461 | 1154 | 3898 | 2662 | 2934 | 969 | 17078 | |
| | Erro omissão | 0,04 | 0,20 | 0,24 | 0,29 | 0,03 | 0,45 | | |

A Figura 6 apresenta os resultados da classificação usando as combinações RGB+I, RGB+IS e RGB+ISM. A partir da análise visual dos resultados verifica-se que na combinação RGB+I (Figuras 6(a) e 6(b)), o resultado apresentou classes com aparência ruidosa, justamente devido aos ruídos inerentes à imagem de intensidade. Alguns elementos, que apresentam respostas radiométricas similares às das vias nessa imagem (por exemplo, regiões de sombra, árvores e algumas edificações) foram classificadas como 'via asfaltada' e, como consequência, essa classe apresentou-se com muitos pixels no interior das quadras (Figura 6(b)). Nas Figuras 6(c) e 6(d) é apresentado o resultado da classificação usando a combinação RGB+IS. Comparando com o resultado da classificação anterior (combinação RGB+I), verifica-se que os ruídos foram minimizados. A região em destaque na Figura 6(c) mostra que as vias obstruídas por árvores nas imagens RGB, puderam ser identificadas, entretanto, algumas árvores causam obstrução total das vias tanto na imagem de intensidade como na imagem RGB, causando uma falha na sua continuidade. Analisando isoladamente a classe 'via asfaltada' (Figura 6(d)), observa-se a diminuição significativa de pixels no interior das quadras classificados como 'via asfaltada'. As Figuras 6(e) e 6(f) apresentam os resultados da

classificação usando a combinação RGB+ISM, onde verifica-se pouca diferença em relação ao resultado da classificação anterior (combinação RGB+IS). Ao comparar somente a classe 'via asfaltada' (Figura 6(f)) com a classificação anterior (Figura 6(d)), as modificações tornam-se mais nítida, primeiramente observando o efeito serrilhado na borda da classe 'via asfaltada'. Outro aspecto é a diminuição dos pixels no interior das quadras. Vale notar também, que as vias situadas na parte inferior direita do recorte (área em destaque na Figura 6(e)) apresentaram-se com mais falhas.

Figura 6 - Resultado da classificação adicionando os dados radiométricos do laser.

Combinação RGB+I: (a) Imagem classificada e (b) Classe 'via asfaltada';
 Combinação RGB+IS: (c) Imagem classificada e (d) Classe 'via asfaltada'; e
 Combinação RGB+ISM: (e) Imagem classificada e (f) Classe 'via asfaltada'.



Em todos os resultados apresentados na Figura 6 foi possível a identificação das vias, mesmo em áreas mais obstruídas por árvores. Outro fato importante é em relação às sombras projetadas sobre as vias, que não afetaram o resultado da classificação para a classe 'via asfaltada'. Para as combinações RGB+I, RGB+IS e RGB+ISM, verificam-se valores muito próximos para o coeficiente kappa (Tabela 2) e dentre essas combinações, aquelas com a imagem suavizada (RGB+IS e RGB+ISM) foram as que apresentaram maiores valores para a AU e AP para a classe via. Pela análise visual, a combinação RGB+IS apresentou melhor resultado, principalmente, na área crítica em destaque, obtendo-a sem muitas falhas. A partir da Tabela 2 constata-se que a combinação RGB+ISM apresentou valores mais altos para as acurácias.

A matriz de confusão para a combinação RGB+IS é apresentada na Tabela 5, onde é possível constatar que os erros de inclusão e de omissão para a classe 'via asfaltada' apresentam-se baixos. A classe 'árvore' contribuiu para o erro de inclusão da classe 'via asfaltada', com 415 pixels (7,13%) e a classe 'ed. telhado cinza' contribuiu para o erro de omissão da classe 'via asfaltada' com 198 pixels (3,62%).

Tabela 5 - Matriz de confusão resultante da combinação RGB+IS

| Classificação por RNA | Dados de referência | | | | | | | TOTAL | Erro Inclusão |
|-----------------------|---------------------|-------|--------|-------------------|-------------------|----------|-------|-------|---------------|
| | via asfaltada | grama | árvore | ed. telhado cinza | ed. telhado verm. | Concreto | | | |
| via asfaltada | 5195 | 0 | 415 | 106 | 16 | 81 | 5813 | 0,11 | |
| grama | 2 | 972 | 790 | 24 | 3 | 5 | 1796 | 0,46 | |
| árvore | 49 | 73 | 2326 | 104 | 14 | 86 | 2652 | 0,12 | |
| ed. tel. cinza | 198 | 41 | 170 | 1759 | 226 | 110 | 2504 | 0,30 | |
| ed. tel. verm. | 1 | 41 | 74 | 256 | 2655 | 58 | 3085 | 0,14 | |
| concreto | 16 | 27 | 123 | 413 | 20 | 629 | 1228 | 0,49 | |
| TOTAL | 5461 | 1154 | 3898 | 2662 | 2934 | 969 | 17078 | | |
| Erro omissão | 0,05 | 0,16 | 0,40 | 0,34 | 0,10 | 0,35 | | | |

A combinação RGB+ISM também apresentou valores baixos para os erros de inclusão e de omissão para a classe 'via asfaltada' (Tabela 6). O erro de inclusão para essa classe apresentou contribuição de 279 pixels (4,83%) da classe 'árvore' e 121 pixels (2,78%) da classe 'concreto'. Já o erro de omissão para a classe 'via asfaltada' apresentou contribuição muito pequena da classe 'concreto' (menos de 2%).

Com base nas análises realizadas para as combinações RGB+I, RGB+IS e RGB+ISM conclui-se que imagem de intensidade adicionada às imagens RGB melhora significativamente o resultado da classificação para a classe 'via asfaltada'.

Usando essas combinações foi possível classificar corretamente pixels correspondentes às vias que não aparecem na imagem aérea colorida por problemas de obstruções, mas que foram registradas na imagem de intensidade. O alto contraste das regiões de vias na imagem de intensidade “força” o resultado da classificação, como foi observado na Figura 6. No entanto, verificou-se que não foi possível manter um bom resultado para as demais classes.

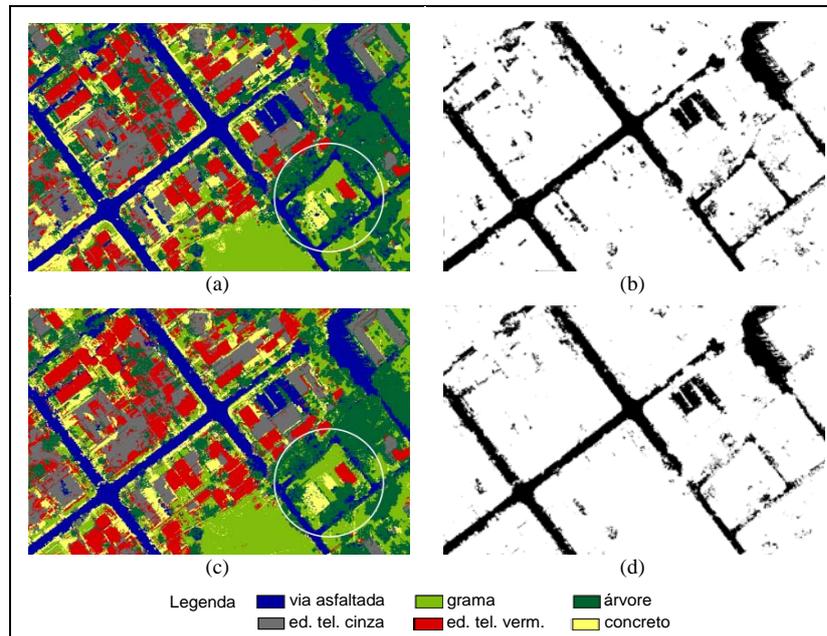
Tabela 6 - Matriz de confusão resultante da combinação RGB+ISM.

| | | Dados de referência | | | | | | | TOTAL | Erro Inclusão |
|-----------------------|----------------|---------------------|-------|--------|-------------------|-------------------|----------|-------|-------|---------------|
| | | via asfaltada | grama | árvore | ed. telhado cinza | ed. telhado verm. | concreto | | | |
| Classificação por RNA | via asfaltada | 5256 | 0 | 279 | 116 | 0 | 121 | 5772 | 0,09 | |
| | grama | 4 | 853 | 455 | 19 | 0 | 7 | 1338 | 0,36 | |
| | árvore | 36 | 174 | 2707 | 62 | 1 | 54 | 3034 | 0,11 | |
| | ed. tel. cinza | 48 | 33 | 364 | 1515 | 82 | 93 | 2135 | 0,29 | |
| | ed. tel. verm. | 14 | 69 | 71 | 399 | 2826 | 52 | 3431 | 0,18 | |
| | concreto | 103 | 25 | 22 | 551 | 25 | 642 | 1368 | 0,53 | |
| | TOTAL | 5461 | 1154 | 3898 | 2662 | 2934 | 969 | 17078 | | |
| Erro omissão | 0,04 | 0,26 | 0,31 | 0,43 | 0,04 | 0,34 | | | | |

Uma forma de melhorar o resultado da classificação é combinar os dados radiométricos e geométricos de varredura a laser com a imagem aérea colorida. Com base nos resultados das classificações anteriores, as combinações RGB+MDSnLim+IS e RGB+MDSnLim+ISM são propostas, conforme apresentado na Tabela 1, e os resultados da classificação são apresentados na Figura 7.

A partir da análise visual da imagem classificada usando a combinação RGB+MDSnLim+IS (Figura 7(a)), verifica-se que não houve edificações classificadas como ‘vias asfaltadas’ e as vias obstruídas por árvores (com destaque em branco) puderam ser identificadas corretamente. Na Figura 7(b), é nítida a diminuição da quantidade de pixels no interior das quadras correspondentes aos outros objetos, mas permanecem alguns deles ainda relacionados às sombras. No entanto, a quantidade é bem menor do que os resultados anteriores apresentaram. O resultado da classificação usando a combinação RGB+MDSnLim+ISM, pode ser visto na Figura 7(c), onde observa-se que a influência da sombra diminuiu ainda mais, e consequentemente diminuiu os pixels no interior das quadras classificados incorretamente como ‘via asfaltada’ (Figura 7(d)). No entanto, na região onde as vias são obstruídas por árvores, quando comparado com o resultado anterior, verifica-se que elas foram detectadas com mais falhas, como pode ser visualizado em destaque em branco na Figura 7(c).

Figura 7 - Resultado da classificação adicionando ambos os dados de varredura a laser. Combinação RGB+MDSnLim+IS: (a) Imagem classificada e (b) Classe 'via asfaltada'; e Combinação RGB+MDSnLim+ISM: (c) Imagem classificada e (d) Classe 'via asfaltada'.



Analisando a matriz de confusão para a classificação usando a combinação RGB+MDSnLim+IS (Tabela 7), tem-se que a classe 'via asfaltada' apresentou valores muito pequenos para os erros de inclusão e de omissão, confirmando um resultado mais acurado para esta classe. No caso do erro de inclusão, 306 pixels (5,3%) foram classificados como 'via asfaltada', mas deveriam ser atribuídos à classe 'árvore'. Analisando as imagens de entrada e o resultado da classificação, pode-se dizer que, provavelmente esses pixels correspondem àquelas árvores que na imagem de intensidade possuem a mesma característica das vias, e que não puderam ser separados pela informação de altura, devido aos dados não serem da mesma data. Os maiores erros de inclusão e de omissão ocorreram para a classe 'concreto', com a classe 'via asfaltada' contribuindo com 108 pixels (9,6%) para o erro de inclusão e com 83 pixels (8,6%) para o erro de omissão.

Os erros de inclusão e de omissão também obtiveram valores baixos para a classe 'via asfaltada' na classificação usando a combinação RGB+MDSnLim+ISM,

como mostra a Tabela 8. O maior erro de inclusão ocorreu para a classe 'concreto', com contribuição de 145 pixels (13,5%) que deveriam ser classificados como 'via asfaltada'.

Tabela 7 - Matriz de confusão resultante da combinação RGB+MDSnLim+IS.

| | | Dados de referência | | | | | | | | |
|-----------------------|----------------|---------------------|------------|-------------|-------------------|-------------------|------------|--------------|---------------|--|
| Classificação por RNA | | via asfaltada | grama | árvore | ed. telhado cinza | ed. telhado verm. | concreto | TOTAL | Erro Inclusão | |
| | via asfaltada | 5252 | 0 | 306 | 144 | 6 | 83 | 5791 | 0,09 | |
| | grama | 0 | 998 | 228 | 25 | 38 | 21 | 1310 | 0,24 | |
| | árvore | 50 | 96 | 2808 | 105 | 26 | 101 | 3186 | 0,12 | |
| | ed. tel. cinza | 51 | 0 | 270 | 1992 | 102 | 208 | 2623 | 0,24 | |
| | ed. tel. verm. | 0 | 0 | 101 | 212 | 2728 | 4 | 3045 | 0,10 | |
| | concreto | 108 | 60 | 185 | 184 | 34 | 552 | 1123 | 0,51 | |
| | TOTAL | 5461 | 1154 | 3898 | 2662 | 2934 | 969 | 17078 | | |
| Erro omissão | 0,04 | 0,14 | 0,28 | 0,25 | 0,07 | 0,43 | | | | |

A classe 'concreto' também apresentou maior erro de omissão, mas sem contribuição significativa da classe 'via asfaltada'. Apesar da classe 'via asfaltada' apresentar um resultado mais limpo, com diminuição de pixels no interior das quadras, verifica-se na Tabela 2, que o acurácia do produtor diminuiu em relação ao resultado da classificação anterior (combinação RGB+MDSnLim+IS), enquanto que o coeficiente kappa permaneceu igual. A imagem suavizada usando MM ajudou a diminuir pixels relacionados às sombras, no entanto, a área onde as vias são obstruídas por árvores foi detectada com mais falhas.

Tabela 8 - Matriz de confusão resultante da combinação RGB+MDSnLim+ISM.

| | | Dados de referência | | | | | | | | |
|-----------------------|----------------|---------------------|-------------|-------------|-------------------|-------------------|------------|--------------|---------------|--|
| Classificação por RNA | | via asfaltada | grama | árvore | ed. telhado cinza | ed. telhado verm. | concreto | TOTAL | Erro Inclusão | |
| | via asfaltada | 5164 | 0 | 173 | 77 | 0 | 49 | 5463 | 0,05 | |
| | grama | 20 | 1060 | 574 | 15 | 13 | 37 | 1719 | 0,38 | |
| | árvore | 84 | 69 | 2743 | 69 | 1 | 45 | 3011 | 0,09 | |
| | ed. tel. cinza | 47 | 0 | 255 | 1975 | 166 | 209 | 2652 | 0,26 | |
| | ed. tel. verm. | 1 | 12 | 90 | 315 | 2720 | 38 | 3176 | 0,14 | |
| | concreto | 145 | 13 | 63 | 211 | 34 | 591 | 1057 | 0,44 | |
| | TOTAL | 5461 | 1154 | 3898 | 2662 | 2934 | 969 | 17078 | | |
| Erro omissão | 0,05 | 0,08 | 0,30 | 0,26 | 0,07 | 0,39 | | | | |

Conforme mostra a Tabela 2, a classe 'via asfaltada' apresentou um valor para a acurácia do produtor igual a 96,17%, considerado alto, mas um pouco menor que quando usado a combinação RGB+ISM, ou seja, somente com dados radiométricos do laser. O coeficiente kappa apresentou valor igual a 0,79, só alcançado com a combinação usando dados geométricos do laser (MDSnLim). Por isso, a combinação de ambos os dados laser ajuda a melhorar a acurácia tanto para a classe 'via asfaltada' como para as demais classes.

7. CONCLUSÃO

O presente trabalho apresentou um estudo experimental combinando imagem aérea RGB de alta resolução com imagens geradas a partir dos dados geométricos e radiométricos de varredura a laser, num processo de classificação supervisionada por RNA, cujo objetivo foi isolar a classe 'via asfaltada' com maior acurácia. Diversas arquiteturas de rede foram testadas para cada combinação de dados de entrada, apresentando aquelas que obtiveram um melhor resultado, o qual foi analisado visual e numericamente. A análise numérica foi realizada por meio das acurácias do produtor e do usuário para cada classe, o coeficiente kappa e a matriz de confusão.

Com base nos experimentos realizados verificou-se que as combinações RGB+MDSnLim, RGB+MDSnLim+IS e RGB+MDSnLim+ISM apresentaram valores muito próximos para o coeficiente kappa e acurácia do produtor. No entanto, a combinação RGB+MDSnLim apresentou valor inferior para a AU, justamente devido à maior quantidade de pixels incorretamente classificados como 'via asfaltada', os quais representam regiões de sombra e estão no nível do terreno. As combinações que integram ambos os dados de varredura a laser (RGB+MDSnLim+IS e RGB+MDSnLim+ISM) apresentaram valores maiores que 90% para as acurácias do produtor e do usuário para a classe 'via asfaltada'. Logo, conclui-se que os dados radiométricos e geométricos de varredura a laser complementam a imagem aérea RGB no sentido de melhorar o isolamento da classe 'via asfaltada' das demais classes. A característica radiométrica das vias na imagem de intensidade beneficia diretamente o resultado para a classe 'via asfaltada', permitindo classificar corretamente pixels correspondentes às vias que não aparecem na imagem RGB por problemas de obstruções, mas que foram registradas na imagem de intensidade. O MDSn limiarizado permite separar os pixels da imagem em dois conjuntos de classes: 1) classes que representam objetos da cena na altura do terreno (vias, áreas gramadas etc.) e; 2) classes que representam objetos elevados (edificações e árvores). Nestas combinações, problemas relacionados com as respostas espectrais similares foram minimizados.

Os pixels no interior das quadras, classificados erroneamente como 'vias asfaltadas', são encontrados em todos os resultados. Alguns deles em menor quantidade, sem a deterioração da classe de interesse. O pós-processamento da classe 'via asfaltada' resultante é proposto como trabalhos futuros, cujas tarefas são: 1) eliminação de falsos positivos a partir de informações de área, forma e

compacidade; 2) ligação de vias desconectadas a partir do esqueleto das regiões formadas; e 3) regularização das bordas, envolvendo processamento de morfologia matemática.

AGRADECIMENTOS

Ao Programa de Pós-Graduação em Ciências Cartográficas – FCT/UNESP, ao CNPq, por concessão de bolsa de doutorado do primeiro autor (Edital MCT/CNPq nº 27/2007) e ao LACTEC pelo fornecimento dos dados de varredura a LASER.

REFERÊNCIAS BIBLIOGRÁFICAS

- ALONSO, M. C.; MALPICA, J. A. Satellite imagery classification with LiDAR data. In: *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Science*, Volume XXXVIII, Part 8, Japan, p. 730 – 735, 2010.
- ANDRADE A. F.; CENTENO, J. A. S. Integração de informações espectrais e de forma na classificação de imagens com redes neurais. *Boletim de Ciências Geodésicas*, v. 9, n. 2, p. 217-231, 2003.
- BELLENS, R.; GAUTAMA, S.; MARTINEZ-FONTE, L.; PHILIPS, W.; CHAN, J. C.; CANTERS, F. Improved Classification of VHR Images of Urban Areas using Directional Morphological Profiles. *IEEE Transactions on Geoscience and Remote Sensing*, v. 46, n. 10, p. 2803 - 2813, 2008.
- BENEDIKTSSON J. A.; PESARESI, M.; ARNASON, K. Classification and Feature Extraction for Remote Sensing Images from Urban Areas based on Morphological Transformations. *IEEE Transactions on Geoscience and Remote Sensing*, v. 41, n. 9, p.1940 - 1949, 2003.
- BENKOUIDER, F.; HAMAMI, L.; ABDELLAOUI, A. Use of the Neural Net for Road Extraction from Satellite Images, Application in the City of Laghouat (Algeria). *PIERS ONLINE*, v. 7, n. 2, p. 146-150, 2011.
- BOTELHO, M. F.; CENTENO, J. A. S. Uso integrado de imagem de alta resolução espacial e altura derivada do LASER Scanner na escolha do classificador orientado a região. *Boletim de Ciências Geodésicas*, Curitiba, v. 11, n. 1, p. 71-87, 2005.
- CHANUSSOT, J.; BENEDIKTSSON, J. A.; FAUVEL, M. Classification of Remote Sensing Images From Urban Areas Using a Fuzzy Possibilistic Model. *IEEE Geoscience and Remote Sensing Letters*, v. 3, n. 1, p. 40 - 44, 2006.
- DALMOLIN, Q.; SANTOS, D. R. Sistema *LASERSCANNING: Conceitos e princípios de funcionamento*. Curitiba: UFPR, 97 p., 2004.
- DAS, S.; MIRNALINEE, T.T.; VARGHESE, K. Use of Salient Features for the Design of a Multistage Framework to Extract Roads From High-Resolution Multispectral Satellite Images. *IEEE Transactions on Geoscience and Remote Sensing*, 2011.

- EL-SHEIMY, N.; VALEO, C.; HABIB, A. *Digital Terrain Modeling: Acquisition, manipulation, and applications*. Artech House, 257 p., 2005.
- FERREIRA, M. S.; GALO, M. L. B. T. Influência dos dados de entrada na classificação a partir de Rede Neural Artificial. In: *III Simpósio Brasileiro de Ciências Geodésicas e Tecnologias da Geoinformação*. Recife, 2010.
- GALO, M. L. B. T. *Aplicação de redes neurais artificiais e sensoriamento remoto na caracterização ambiental do Parque Estadual Morro do Diabo*. 2000. Tese (Doutorado em Ciências da Engenharia Ambiental) - Escola de Engenharia de São Carlos, Universidade de São Paulo, São Carlos.
- GAO, J.; WU, L. Automatic Extraction of Road Networks in Urban Areas from Ikonos Imagery Based on Spatial Reasoning. In: *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Science*, Istanbul, v.35, Comm. III/3, p. 331 – 337, 2004
- GONZALEZ, R. C.; WOODS, R. E. *Processamento Digital de Imagens*. 3ª ed. Pearson, 624 p., 2010.
- GROTE, A.; ROTTENSTEINER, F. Automatic road network extraction in suburban areas from high resolution aerial images. In: *Photogrammetric Computer Vision and Image Analysis*, Proceedings, France, 2010.
- HAALA, N.; WALTER, V. Automatic classification of urban environments for database revision using LiDAR and color aerial imagery. In: *International Archives of Photogrammetry and Remote Sensing*, Vol. 32, Part 7-4-3 W6, 3-4, Spain, 1999.
- JENSEN, J. R. *Sensoriamento Remoto do Ambiente: Uma Perspectiva em Recursos Terrestres*. Tradução da 2ª ed. por (pesquisadores do INPE): José Carlos N. Epiphânio (Coordenador); Antonio R. Formaggio; Athos R. Santos; Bernardo F. T. Rudorff; Cláudia M. Almeida; e Lênio S. Galvão. Editora Parêntese, 598 p., 2009.
- MAKARAU, A.; PALUBINSKAS, G.; REINARTZ, P. Multi-sensor data fusion for urban area classification. In: *Urban Remote Sensing Event 2011*, Munique, p. 21 - 24, 2011.
- MANCINE, A.; FRONTONI, E.; ZINGARETTI, P. Automatic extraction of urban objects from multi-source aerial data. In: *CMRT09. IAPRS*, Vol. XXXVIII, Part 3/W4 – Paris, p. 13 - 18, 2009.
- MOKHTARZADE, M.; VALADAN ZOEJ, M. J. Road detection from high resolution satellite images using artificial neural networks. *International Journal of Applied Earth Observation and Geoinformation*, 9(1), p. 32 - 40, 2007.
- SCHOWENGERDT, R. A. *Remote Sensing: Models and Methods for Image Processing*. 3rd Edition. San Diego: Academic Press, 515 p., 2007.
- SHACKELFORD, A. K.; DAVIS, C. H. Urban Road Network Extraction from High-Resolution Multispectral Data. In: *2nd GRSS/ISPRS Joint Workshop on Remote Sensing and Data Fusion over Urban Areas*. p. 142 - 146, 2003.
- SOILLE, P. *Morphological Image Analysis*. Berlin: Springer-Verlag, 316 p., 1999.

- TRINDER, J. C.; SALAH, M. Support Vector Machines: Optimization and Validation for Land Cover Mapping Using Aerial Images and Lidar Data. In: *34th ISRSE Symposium*, Australia, p. 10 - 15, 2011.
- TSO, B.; MATHER, P. M. Classification Methods for Remotely Sensed Data. CRC Press, 356 p., 2009.
- WEHR, A.; LOHR, U. Airborne LASERscanning – an introduction and overview. *ISPRS Journal of Photogrammetry and Remote Sensing*. v. (54) 2-3. p. 68 - 82, 1999.
- ZENG Y.; ZHANG J.; WANG G.; LIN Z. Urban land-use classification using integrated Airborne Laser Scanning data and high resolution multi-spectral satellite imagery. *Pecora 15/Land Satellite Information IV/ISPRS Commission I/FIEOS*, 2002.
- ZHANG, Y.; WANG, R. Multi-resolution and multi-espectral image fusion for urban object extraction. In: *XXth ISPRS CONGRESS GEO-IMAGERY BRIDGING CONTINENTS*, Vol. XXXV, Part B3, Commission 3, Istanbul, p. 960 - 966, 2004.

(Recebido em dezembro de 2012. Aceito em maio de 2013.)