

DETECÇÃO AUTOMÁTICA DE CONTORNOS DE EDIFICAÇÕES UTILIZANDO IMAGEM GERADA POR CÂMARA DIGITAL DE PEQUENO FORMATO E DADOS LIDAR

*Automatic Detection of Building Contours using a small format digital camera and
LIDAR data*

ALVARO MURIEL LIMA MACHADO

EDSON A. MITISHITA

Universidade Federal do Paraná

Curso de Pós-Graduação em Ciências Geodésicas

Caixa Postal 515 – Centro Politécnico – Jardim das Américas

CEP 12201-970 – Curitiba – PR – Brasil

{[alvaroml.mitihista](mailto:alvaroml.mitihista@ufpr.br)}@ufpr.br

RESUMO

Este trabalho apresenta uma metodologia para detectar contornos de edificações automaticamente, a partir da integração de imagem tomada por câmara digital não fotogramétrica com dados de varredura LASER. Assume-se o conhecimento da orientação exterior da imagem e orientação interior da câmara. Como espaço comum de trabalho, das informações provenientes da imagem e dos dados LASER, usa-se o espaço-imagem, isento de distorções relativas à câmara digital não fotogramétrica. Para a remoção dos erros sistemáticos, contidos na imagem original, emprega-se o método iterativo de Newton-Raphson para sistemas de equações, corrigindo a imagem. Essa imagem é segmentada, segundo suas cores, no espaço de cores CIELUV, via algoritmo de deslocamento pela média (*mean shift*). Realiza-se, a seguir, uma avaliação dos segmentos gerados, segundo propriedades como tons de verde (removendo vegetação), altitude média de regiões frente a regiões vizinhas (preservando regiões altas), e grau de simplificação do polígono envolvente (usando o algoritmo de Douglas-Peucker), sempre com objetivo de identificar feições retilíneas (edificações potenciais). Diversos experimentos realizados sobre área do Centro Politécnico - UFPR, em Curitiba/PR, são discutidos, com resultados promissores para edificações isoladas.

Palavras-chave: Monorestituição, erros sistemáticos, processamento digital de imagens.

ABSTRACT

This work presents a methodology to automatically extract 2D building contours from the integration between a small format digital camera image and laser scanning raw data. The external image orientation and internal camera orientation parameters are assumed to be known. As a common space to integrate information from both sources it is used the image space without distortion. To remove systematic errors in the original image, the Newton-Raphson iterative method for simultaneous equation is used in a resampling process to this space. At this step, all the pixels have color and altitude associated information. The RGB color values are transformed to (almost isotropic) CIELUV color space and the image is segmented according to these values, by mean shift algorithm. The forward steps filter the gotten segments, using colors (to remove green vegetation), medium altitude of a region comparing it with the altitude of its neighbours (to preserve high regions), and polygonal simplification degree (Douglas-Peucker), attempting to detect linear features. Several experiments realized over a Centro Politécnico - UFPR area, in Curitiba/PR, are discussed, with good results for isolated buildings.

Keywords: Monoplotting, systematic errors, digital image processing.

1. INTRODUÇÃO

Segundo BALTSAVIAS (1999), o LIDAR (*Light Detection And Ranging*) se constituiu em um dos desenvolvimentos tecnológicos mais importantes do final do século passado. A automatização de procedimentos, aliada à precisão, acurácia e densidade das coordenadas tridimensionais dos pontos adquiridos, de forma rápida, torna a varredura pelo sistema LIDAR revolucionária.

Infelizmente, devido à forma amostral com que as informações provenientes do sistema LIDAR são geradas, elas carecem de qualidade morfológica. Isto quer dizer que as linhas de quebra das feições existentes em uma região analisada não se encontram explicitamente presentes nos dados. Por conseguinte, os contornos de edificações são mal delineados (ACKERMANN, 1999; AXELSSON, 1999; HAALA et al., 1997; KRAUS e PFEIFER, 1998). Assim, o modelo digital de superfície (MDS) gerado pelo sistema LIDAR tende a ser suavizado, e perde informação importante sobre uma feição de terreno (KRAUS e PFEIFER, 1998). Segundo BALTSAVIAS (1999), medições fotogramétricas ainda representam o melhor método de obtenção de informação de alta qualidade geomorfológica.

HABIB et al. (2004) observa que dados de Fotogrametria e de LIDAR têm características distintas que tornam uma metodologia ou outra preferencial para certas aplicações. Uma desvantagem nas medições fotogramétricas convencionais é contrastada por uma forte vantagem no sistema LIDAR, e vice-versa. O LIDAR fornece dados de altimetria com melhor qualidade do que planimetria, e na Fotogrametria acontece exatamente o contrário. Desta forma, uma integração entre ambos os métodos se torna desejável. Na busca de informação fotogramétrica, para complementar as informações provenientes da varredura de uma região pelo sistema

LIDAR, as imagens obtidas por uma câmara digital de pequeno formato são convenientes.

A proposta deste trabalho é a de apresentar uma metodologia capaz de extrair contornos bidimensionais de edificações, agregando as informações espaciais obtidas pela varredura LIDAR de uma região urbana qualquer, com imagens adquiridas por uma câmara digital de pequeno formato da mesma região. O processo é automático, em toda sua extensão. A região utilizada para testes foi a do campus da Universidade Federal do Paraná, no Centro Politécnico. Os parâmetros de orientação interior da câmara, e exterior da imagem obtida são supostos conhecidos.

A literatura apresenta diversas metodologias para se extrair contornos de edificações, considerando-se informações provenientes de sistema LIDAR, sobretudo as sustentadas por modelos digitais de superfície normalizada. Essa sistemática utiliza tão somente os dados oriundos de perfilamento por LIDAR, que constituem o modelo digital de superfície (MDS). A partir do MDS, por interpolação, obtém-se o modelo digital de terreno (MDT). Nesta empreitada usam-se procedimentos de eliminação de desníveis abruptos. A diferença entre o MDS e o MDT compreende edificações, vegetações, e outros resíduos.

Neste trabalho procura-se fugir desta abordagem, tendo em vista, que se dispõe de mais informação: a imagem da região trabalhada. As distâncias amostrais de terreno coberto pela imagem são menores (melhor resolução) que as existentes no perfilamento por LIDAR. Para a identificação das edificações emprega-se, sobretudo, as feições retilíneas presentes na maioria delas.

2. MATERIAL E MÉTODOS

A metodologia desenvolvida emprega o espaço-imagem isento de distorções como sendo o espaço comum de integração, entre as informações provenientes da imagem digital (cores e métrica), e dos pontos obtidos pelo sistema LIDAR (planialtimetria). A imagem isenta de distorções é segmentada, segundo suas cores no espaço CIELUV, pelo algoritmo de deslocamento pela média (FUKUNAGA e HOSTETLER, 1975; COMANICIU e MEER, 1999; COMANICIU e MEER, 2002). A seguir, aplica-se o algoritmo de perseguição de contornos por varredura radial, para a geração dos contornos das regiões segmentadas, definindo-os pixel por pixel.

Complementarmente à segmentação por cores deve-se efetuar um trabalho de análise e filtragem dos segmentos encontrados, eliminando-se as regiões não condizentes com edificações. Para se validar os segmentos, três filtros são aplicados sucessivamente, diminuindo passo a passo a quantidade de regiões avaliadas: filtro para o padrão VERDE (eliminação de vegetação), filtro altimétrico (preservação de regiões altas), e filtro Douglas-Peucker (detecção de feições lineares). As regiões que passaram pelo crivo dos filtros apresentados são avaliadas quanto à possibilidade de fusão entre regiões vizinhas, estabelecendo-se os contornos

definitivos (nível de pixel) do que se acredita serem edificações.

A última etapa transforma as coordenadas de imagem dos pontos mais significativos do contorno de uma edificação presumida, para coordenadas de terreno, vetorizando-os. Esta transformação é realizada via equações de colinearidade, em sua forma inversa, mas sem iterações.

Detalha-se, a seguir, a sistemática elaborada, com a apresentação de alguns dos resultados obtidos, demonstrando a aplicabilidade da metodologia desenvolvida.

2.1. Materiais

O levantamento fotográfico da região do campus do Centro Politécnico da Universidade Federal do Paraná (UFPR), em Curitiba, foi realizado no dia 27/06/2003, em vôo com altitude aproximada de 1650m. As imagens tomadas são coloridas.

A câmara empregada, Sony DSC-F717 de propriedade da Agritec, permite a aquisição de imagens com 2560 X 1920 pixels (5,2Mb), possuindo CCD com dimensões 8,8mm X 6,6mm. Esta câmara foi calibrada por DELARA (2003). Os parâmetros de orientação interior (IOP) significativos no ajustamento MMQ (Método dos Mínimos Quadrados) realizado são mostrados no Quadro 1, sendo: c = constante da câmara (ou distância focal), em mm; x_0 , y_0 = coordenadas do ponto principal no referencial fiducial, em mm; k_1 , k_2 = parâmetros da distorção radial simétrica.

Quadro 1. Parâmetros de orientação interior (IOP)

DSC-F717	PARÂMETRO S	DESVIO PADRÃO
c	10,078mm	4,58E-3mm
x_0	-0,246mm	5,24E-3mm
y_0	-0,142mm	5,26E-3mm
k_1	-2,25E-03mm ⁻²	8,53E-5 mm ⁻²
k_2	2,49E-05mm ⁻⁴	1,16E-5 mm ⁻⁴

Fonte: DELARA (2003)

O levantamento por LIDAR, da mesma região do campus, foi realizado com o sistema OPTECH ALTM (Airborne LASER Terrain Mapper) 2050 / INS GPS APPLANIX, pertencente ao Instituto de Tecnologia para o Desenvolvimento (LACTEC - UFPR). Seis sobrevôos foram realizados em 09/05/2003, com altitude aproximada de 1000m. As varreduras são em formato de zig-zag transversal ao vôo.

A imagem escolhida para este trabalho, apresentada na Figura 1, tem seus parâmetros de orientação exterior (EOP) indicados no Quadro 2. Estes parâmetros foram obtidos através de uma recessão espacial empregando um conjunto de pontos de apoio observados na imagem.

Sabendo-se que a altitude média da região visualizada é de aproximadamente

900m, pode-se calcular o valor da distância amostral no terreno ($GSD = \text{Ground Sample Distance}$) como sendo 25cm (por pixel). Este valor pode ser obtido através da escala da foto:

$$GSD = \text{Tamanho do Pixel} * \frac{\text{Altura de Vôo}}{\text{Constante da Câmara}} \quad (01)$$

onde:

$$\text{Tamanho do Pixel} = \frac{\text{Dimensão CCD}}{\text{Qtde de Pixels}} = \frac{8,8}{2560} = \frac{6,6}{1920} = 3,4375 \mu\text{m/pixel} \quad (02)$$

Quadro 2. Parâmetros de orientação exterior de imagem (EOP)

IDENTIFICAÇÃO	PARÂMETROS	DESVIO PADRÃO
X_0	677607,705 m	8,456E-1 m
Y_0	7183723,953 m	6,759E-1 m
Z_0	1651,643 m	4,337E-1 m
KAPA	1,951547 rad	2,347E-4 rad
PHI	0,008552 rad	5,888E-4 rad
OMEGA	-0,015407 rad	4,037E-4 rad

Figura 1. Imagem selecionada



A densidade média de pontos do LIDAR observáveis na imagem resultou 4,0698 pontos/m² ou, linearmente, 2,017 pontos/m. É importante salientar-se que se trata de uma aproximação, pois existem várias regiões, dentro da imagem, cobertas por mais do que uma faixa. Os cálculos apresentados enfatizam a melhor resolução da imagem, em relação aos dados do LIDAR, no presente trabalho.

2.2. Metodologia

Dispõe-se de um conjunto de coordenadas tridimensionais de pontos obtidos por LIDAR (espaço objeto) e de uma imagem colorida tomada por câmara não fotogramétrica. A metodologia aplicada busca, inicialmente, a equiparação com o processo natural empregado pelo ser humano: a diferenciação de uma edificação pela sua coloração e a detecção de formas retilíneas. A altura da edificação, em relação à sua vizinhança, também é verificada.

Tendo em vista a necessidade de análise da forma das feições, deve-se eliminar as distorções presentes na imagem, reamostrando-se a mesma. Para não haver a criação de novas cores, opta-se pelo emprego da amostragem segundo o vizinho mais próximo. Nesse processo, utiliza-se o método iterativo de Newton-Raphson para a resolução do sistema de equações para correção dos erros sistemáticos, como discutido por MACHADO et al. (2004).

A nuvem de pontos do LIDAR, projetada no espaço-imagem isento de distorções, adiciona a informação de altimetria em alguns pixels que apresentam correspondência. O espaço de cores RGB da imagem reamostrada, convertido para o espaço "isotrópico" CIELUV, informa a cor. A adoção do espaço de cores CIELUV é justificada pela minimização de parâmetros de definição da forma e tamanho da região circunvizinha (espaço de cores "isotrópico") a um dado pixel.

Com a projeção dos pontos do LIDAR sobre o espaço-imagem isento de distorções cria-se um referencial 3D (C, L, Z) com coordenadas planimétricas de imagem Coluna (C) e Linha (L), e altitude (Z) proveniente do LIDAR. A colorimetria advém da foto tomada. Estas informações constituem um banco de dados trabalhado nos procedimentos que se seguem.

2.2.1. Segmentação por Cores (Algoritmo de Deslocamento pela Média)

É sabido que a etapa de segmentação é fundamental para se obter um bom resultado no processo de extração de edificações. Segundo BRADSKI e THRUN (2003) o algoritmo de deslocamento pela média talvez seja o melhor segmentador existente na atualidade. Dadas essas características, empregou-se o procedimento de COMANICIU e MEER (1997) como base, neste trabalho, para a etapa de segmentação da imagem colorida.

O algoritmo de deslocamento pela média (*mean-shift*) é uma técnica robusta de classificação de agrupamentos que não exige conhecimento a priori sobre a quantidade de grupos, nem restringe o formato destes agrupamentos. Trata-se de um método iterativo que estima a moda (regiões mais densas) de distribuições multivariadas obtidas no espaço de atributos. A quantidade de agrupamentos é obtida automaticamente pela descoberta dos centros destas modas.

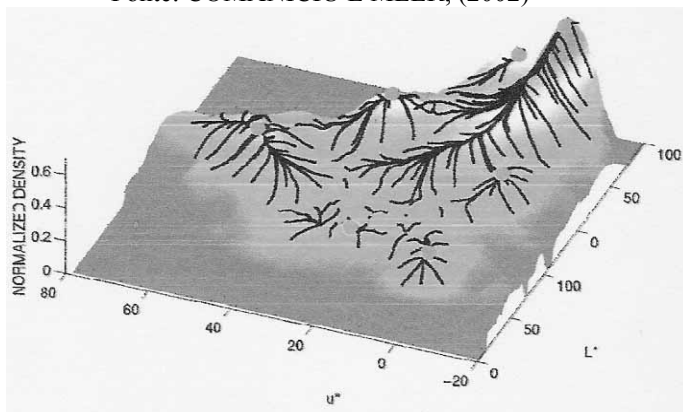
A Equação (03) define o vetor de deslocamento pela média, sendo primeiramente apresentada por FUKUNAGA e HOSTETLER (1975), como um algoritmo de classificação não-paramétrica, e mostra que a estimativa do gradiente

de densidade ($\hat{\nabla}f(x)$) normalizado ($\hat{\nabla}f(x) / \hat{f}(x)$) pode ser calculada através do deslocamento amostral pela média em um *kernel* uniforme centrado em x . O deslocamento pela média é um excelente estimador não supervisionado do gradiente de densidade não-paramétrico, e o vetor de deslocamento pela média representa a diferença entre uma média local, calculada em uma janela com raio (banda de passagem) h no espaço d -dimensional, e o centro desta janela.

$$\vec{M}_h(x) = \frac{h^2}{(d+2)} \frac{\hat{\nabla}f(x)}{\hat{f}(x)} \quad (03)$$

O vetor de deslocamento pela média aponta sempre na direção do maior incremento na densidade, isto é, descobre o caminho que leva a um máximo local de densidade (moda da densidade). O algoritmo trabalha em forma iterativa, deslocando as amostragens sucessivamente em direção ao máximo local mais adequado. A Figura 2 ilustra esta tendência.

Figura 2. Tendência do algoritmo de deslocamento pela média
Fonte: COMANICIU E MEER, (2002)



O algoritmo de deslocamento pela média pode ser estendido para processamento de imagens levando em consideração a vizinhança espacial de cada pixel selecionado (COMANICIU e MEER, 2002). Estes autores introduzem o domínio conjunto: toda amostra passa a ser descrita pela união do domínio da imagem (características da imagem: tons de cinza, componentes RGB, etc.) com o domínio espacial (localização de um pixel dentro da imagem). Dado $z_i^m = \{s_i^m, a_i^m\}, \forall i \in [1..n]$, domínio conjunto, onde s representa a informação espacial e a a informação contida na imagem, COMANICIU e MEER (2002) criaram o *kernel* definido pelo produto de dois *kernels*: K_A para o domínio da

imagem, e K_S para o domínio espacial.

A estimativa de densidade (suavizada) pode ser, então, escrita como:

$$\hat{f}(z_j^m) = \frac{1}{n} \sum_{i=1}^n \frac{1}{h_A h_S} K_A \left(\frac{a_j^m - a_i^m}{h_A} \right) K_S \left(\frac{s_j^m - s_i^m}{h_S} \right) \quad (04)$$

onde h_A e h_S são as bandas de passagem (janelas) no domínio da imagem e no domínio espacial, respectivamente. Somente as amostras que recaem dentro dos limites impostos por estas bandas de passagem e z_j^m é que contribuem para a estimativa da densidade no ponto z_j^m . O deslocamento das médias é realizado em ambos os espaços separadamente, isto é:

$$a_j^{m+1} = a_j^m + M_{h_A, K_A}(a_j^m) \quad \text{e} \quad s_j^{m+1} = s_j^m + M_{h_S, K_S}(s_j^m) \quad (05)$$

e a convergência ocorre simultaneamente nos dois domínios, quando ambos os vetores de deslocamento pela média ficarem abaixo de um limiar. O algoritmo classifica os pontos segundo os agrupamentos existentes nesses domínios.

O algoritmo de deslocamento pela média tem este nome porque, no fundo, não passa de um calculador de média ponderada, deslocando-se segundo os valores desta média. Os pesos dos pixels advêm do *kernel* empregado. Neste trabalho empregou-se o *kernel* uniforme, isto é, as médias são médias aritméticas.

A definição do tamanho da janela de pesquisa h_A , no espaço de atributos de cores, resulta de medida da "atividade visual" da imagem, isto é, deriva da matriz global de variância e covariância de cores. O tamanho da janela de pesquisa no domínio de atributos espaciais h_S deve ser obtido empiricamente, ocorrendo um compromisso entre os efeitos indesejáveis de divisão de uma mesma região (tamanho de janela pequeno) ou união de diferentes regiões (tamanho de janela grande). Neste trabalho adotou-se o valor único de 21 pixels por 21 pixels como sendo o tamanho da janela de pesquisa no âmbito de atributos espaciais.

As janelas de pesquisa, quer no domínio das cores (três dimensões), quer no domínio espacial (duas dimensões), definem uma hiperesfera de cinco dimensões. As três dimensões do espaço de atributos das cores são iteradas, no algoritmo de deslocamento pela média, assumindo novos valores e, por conseguinte alterando o posicionamento da hiperesfera. A janela de pesquisa no domínio espacial é fixa durante todo o processo. Todo o algoritmo de deslocamento pela média é executado dentro desta hiperesfera semi-móvel (parte móvel, parte fixa), com resultado válido para o pixel central da hiperesfera final.

2.2.2. Perseguição de Contornos por Varredura Radial

Para se distinguir as fronteiras das regiões segmentadas executa-se uma

varredura de toda a imagem, segundo qualquer ordem pré-estabelecida, com marcação dos pixels que pertencem a uma região com vizinhos em outra região. O conceito de vizinhança é o de vizinhança 4, isto é, os vizinhos da direita, esquerda, de cima e embaixo. O processo de marcação dá origem a uma matriz bidimensional, com dimensões idênticas à da imagem original, e que só tem dois valores por pixel (bi-valorado): ou marcado, ou não marcado.

Os pixels marcados são usados pelo algoritmo de varredura radial visando a geração dos traçados dos contornos das regiões, compondo polígonos. Este algoritmo pode ser encontrado em diversos livros sobre algoritmos e processamento de imagens, tais como PAVLIDIS (1982), SEUL et al. (2001).

Ao se executar este algoritmo de perseguição de contorno, memoriza-se a cada visita de pixel, as suas coordenadas, gerando como produto diversas listas de pixels, sendo que cada lista constitui um polígono.

Como subproduto dessa etapa relaciona-se todo pixel da imagem com um dos polígonos constituídos, via identificador de região.

2.2.3. Refinamento

Esta etapa é responsável pela filtragem dos objetos segmentados, pela identificação, e extração das informações que caracterizam um objeto ou feição. Diversas fases são aqui executadas.

Reunindo as informações de cor originais (CIELUV) e altitude de cada pixel com o identificador de região, obtém-se a cor média de cada região, e a altitude média de cada região. Para o cálculo de altitude média de uma região, usa-se apenas a informação relativa aos pixels que apresentam correspondência com pontos de LIDAR, não se levando em consideração nenhuma altitude interpolada. Nos casos em que não se tem nem uma altitude para se calcular a altitude média de uma região, despreza-se a região. Esta tarefa pode ser executada em uma única busca por toda a imagem, gerando informações que são usadas adiante, em uma sequência de filtrações.

2.2.3.1. Filtro para padrão VERDE

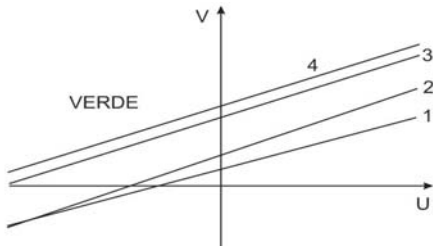
Verifica-se que os dados oriundos do sistema LIDAR, através da diferença de altitude entre seus primeiro e último pulsos, confundem vegetação com bordas de edificações, não fornecendo informação confiável sobre o que é um e o que é outro. Este fato origina a existência desse filtro.

Aproveitando-se das características isotrópicas do espaço de cores CIELUV, e analisando-se as cores formadas neste espaço tridimensional, em função de suas variáveis independentes, estabeleceu-se um subconjunto do domínio destas variáveis para caracterizar o padrão de cor VERDE (com todas suas nuances), na tentativa de identificar de forma simples e inequívoca a vegetação existente na imagem digital utilizada.

Após diversos testes realizados, definiu-se que a região do padrão VERDE, no

espaço de cores CIELUV, poderia ser delimitada através dos parâmetros de quatro retas ($V = A*U + B$). Cada reta atenderia uma faixa do contradomínio de luminosidade. Todo ponto que estivesse acima e a esquerda das retas (no espaço bidimensional $U \times V$) poderia ser considerado como pertencente ao padrão VERDE. O Quadro 3 apresenta estes parâmetros estabelecidos empiricamente, e a Figura 3 fornece uma idéia melhor de sua distribuição espacial, com a identificação das retas especificadas no Quadro 3.

Figura 3. Localização do padrão VERDE adotado



O filtro VERDE estabelecido funciona, então, da seguinte forma: Avalia-se a luminosidade (L), e dependendo de seu valor, seleciona-se uma dentre quatro retas. Feito isto, os valores de U e V (espaço de cores CIELUV) são checados para verificar se o ponto (U,V) se encontra acima ou abaixo da reta selecionada, resultando em VERDE ou não, respectivamente.

Este procedimento é realizado com todas as regiões obtidas, classificando a região como pertencente ao padrão VERDE ou não. As regiões com padrão VERDE são eliminadas dos processamentos consecutivos, com todos os seus pixels, e seus polígonos de fronteira.

Quadro 3. Luminosidade no modelo CIELUV e parâmetros de retas

RETA	FAIXA DE LUMINOSIDADE (L)	COEFICIENTE ANGULAR (A)	COEFICIENT E LINEAR (B)
1	[15 a 50]	1,172291	23,402487
2	(50 a 70]	1,497580	63,752759
3	(70 a 85]	1,382398	124,997724
4	(85 a 90]	1,383178	145,080374

2.2.3.2. Filtro altimétrico

Todas as regiões têm altitude média calculada. Para o processamento, classificam-se essas regiões em ordem crescente, segundo suas altitudes médias. O objetivo deste procedimento é o de atribuir uma classificação como edificação potencial ou não para cada região, em função de exame em todas as regiões

vizinhas. Assume-se que as vegetações (regiões) já foram eliminadas do processo.

O processamento consiste em atribuir o rótulo de edificação potencial (região ALTA) ou não (região BAIXA). Devido à ordem de classificação para o processamento, quando se está processando uma dada região, ela representa a região com menor altitude média sem atribuição de rótulo. Como resultado do processamento obtêm-se classificações binárias para as regiões como ALTA ou BAIXA.

As identificações de edificação potencial ou não são baseadas em desníveis abruptos com regiões vizinhas (principal característica), ou semelhança de altitude com regiões vizinhas previamente rotuladas. No processamento de uma região corrente, todas as suas regiões vizinhas são examinadas. A lógica comparativa emprega limiares (histerese), com adoção, no presente trabalho, de 0,5m e 2,5m. O processamento de uma região corrente, sem rotulação após o exame de todas as regiões vizinhas, faz com que essa região receba a rotulação de BAIXA, porque deste ponto em diante ela é mais baixa do que todas as posteriores.

2.2.3.3. Filtro Douglas & Peucker

De um modo geral, os objetos antrópicos apresentam linhas geométricas, sobretudo linhas retas. Neste filtro pretende-se explorar estas características das edificações, buscando-se segmentos de retas nos contornos das regiões obtidas até o momento. Para se fazer isto, foi utilizado o algoritmo de simplificação de poligonais descrito por Douglas-Peucker (DOUGLAS e PEUCKER, 1973). A simplificação de poligonais é dependente do limiar de tolerância empregado, apresentando mais pontos à medida que a tolerância diminui.

Com este filtro objetiva-se identificar com robustez a presença de linhas retas no polígono envolvente de região (edificação em potencial). Aplica-se, inicialmente, uma tolerância bem baixa, de forma a oferecer boas garantias da presença das feições retas procuradas, e finalmente uma tolerância mais relaxada, para analisar se ainda persiste uma simplificação razoável.

O critério empregado para a tomada de decisão, se um polígono pertence (é mantido) ou não a uma edificação, é dado pelo tamanho da simplificação resultante. Na tolerância inicial (1,5 pixels), se um polígono simplificar em algum trecho de seu contorno mais do que vinte (20) pixels, então ele é mantido. Na tolerância final (3,0 pixels), se um polígono obtiver, em sua simplificação, menos vértices do que 5% dos vértices originais, então ele também é mantido, representando um bom candidato a uma edificação.

2.2.3.4. Fusão de polígonos

Prosseguindo com o refinamento de polígonos envolventes de regiões, observa-se a existência de regiões vizinhas que não foram eliminadas por nenhum dos filtros anteriores. Isto significa que elas passaram nos teste de vegetação, altimetria (altitudes médias em um mesmo patamar), e feições retilíneas.

Aplica-se, neste caso, a seguinte regra: se duas regiões vizinhas apresentam as mesmas características altimétricas, então elas devem pertencer a uma mesma edificação. Surge, então, a necessidade de fundir ambas as regiões em uma única. O processo de fusão elimina a maior fronteira comum entre duas regiões vizinhas, mantendo todos os demais pixels. Isto acontece desde que um segmento não envolva totalmente o outro, caso em que o segmento envolvido é simplesmente eliminado.

2.2.3.5. Vetorização

Nesta etapa estão incluídos os processos para ajustar, os polígonos aceitos como válidos (edificações potenciais), a segmentos de retas, e determinação das intersecções entre essas retas, ou seja, os vértices que definem as feições, visando a definição geométrica de seus contornos.

Para a realização desta tarefa, necessita-se, inicialmente, dividir os vértices de cada polígono, atribuindo-os a diferentes segmentos de retas. Se dois vértices são atribuídos a um mesmo segmento de reta, todos os seus vértices intermediários, dentro da ordem de geração do polígono, devem ser atribuídos ao mesmo segmento de reta. Em suma, é preciso saber quais vértices serão ajustados a quais segmentos de retas.

Para a resolução deste problema usa-se novamente o algoritmo de Douglas-Peucker. Busca-se determinar, dentro de uma tolerância grosseira, quais são os vértices mais críticos dentro do polígono. Neste trabalho, adota-se uma tolerância igual a dez (10,0) pixels. Como resultado, divide-se o polígono em partes, segundo mudanças acentuadas na direção, e todos os vértices intermediários são ajustados a uma mesma reta.

Determinados os coeficientes angulares e lineares das retas que melhor se ajustam aos pontos formadores de um polígono, resolvem-se sistemas de equações (cada duas retas consecutivas compõem um sistema) para a identificação dos pontos de intersecção destas retas entre si. Obtidas tais coordenadas de intersecções de retas, tem-se a vetorização completa de um polígono.

2.2.4. Monorestituição sem iterações

Os algoritmos utilizados fizeram uso de uma altitude associada a cada pixel no espaço-imagem isento de distorções. Estas altitudes foram sempre distinguidas, ao longo de todo o processo, como sendo interpoladas ou originais (provenientes do LIDAR, via equações de colinearidade e parâmetros de orientação exterior).

As coordenadas finais de um trabalho cartográfico devem ser fornecidas no espaço-objeto, em coordenadas referenciadas a um sistema de projeção. Para se obter tais coordenadas, a partir de uma localização específica dentro de uma imagem, deve-se empregar as equações de colinearidade, em sua forma inversa. Estas equações necessitam da altitude Z , previamente à obtenção das coordenadas planimétricas. Sem a altitude não se pode ter X e Y , coordenadas planimétricas. Por

outro lado, tendo-se as mesmas, pode-se encontrar a altitude via MDT.

MAKAROVIC (1973) e MITISHITA (1997) apresentam o processo de monorestituição, para superfícies contínuas, que consiste em adotar uma altitude inicial (geralmente uma altitude média da região trabalhada) e calcular, a partir daí, as coordenadas planimétricas. Encontradas as mesmas, obtém-se uma nova altitude, mais refinada, que pode ser usada novamente nas equações de colinearidade, em sua forma inversa. Este processo é repetido até que as diferenças entre as coordenadas novas e antigas sejam menores que uma tolerância prefixada.

A imagem tomada por uma câmara segue as características da projeção central, com colinearidade (eliminadas as distorções do processo) entre um ponto no terreno, o centro de perspectiva e a imagem do mesmo ponto. Uma linha reta no espaço-objeto é imageada como uma linha reta no espaço-imagem, mantendo todas as proporções. Como conclusão importante tem-se que a triangulação com interpolação linear (tipo de interpolação) apresenta os mesmos resultados quer seja executada no espaço-imagem ou no espaço-objeto. Mais, qualquer interpolação que se faça no espaço-imagem pode ser vista como mais um tipo de interpolação (uma nova função) apresentando um novo resultado, diferente ou não (como qualquer algoritmo de interpolação).

Com este embasamento propõe-se associar, a todo pixel do espaço-imagem, uma altitude original ou interpolada, procurando representar o relevo do terreno, como se fosse uma maquete (a escala de Z é real 1:1).

Para se transformar as coordenadas de imagem (do polígono vetorizado) para o espaço-objeto usa-se a altitude associada ao pixel do qual se deseja as coordenadas de terreno. Neste caso não se precisa realizar iterações para a obtenção das coordenadas finais. Uma vez que a altitude é conhecida, as coordenadas planimétricas são determinadas diretamente.

3. RESULTADOS

Algumas imagens intermediárias obtidas durante o processamento foram colocadas nesta seção para ilustrar melhor o processo. Mais experimentos, com maiores detalhes, podem ser vistos em MACHADO (2006).

Para um dos experimentos escolheu-se uma imagem próxima da situação ideal: uma única edificação isolada, sem árvores na vizinhança da mesma, bem separada de qualquer outra edificação. As Figuras 4 a 9 retratam toda a seqüência de imagens obtidas neste processamento, desde a segmentação até a vetorização.

A fase de segmentação (Figura 4) distinguiu claramente quaisquer diferenças maiores na coloração da imagem. Todos os objetos semelhantes foram agrupados. Prosseguindo-se, algumas áreas verdes mais evidentes foram eliminadas (Figura 5). No filtro altimétrico (Figura 6), tendo em vista o limiar imposto de 2,5m, acabou-se eliminando uma pequena cobertura (anexa à edificação principal) existente na parte superior da imagem, encostada na edificação.

Figura 4. Segmentação

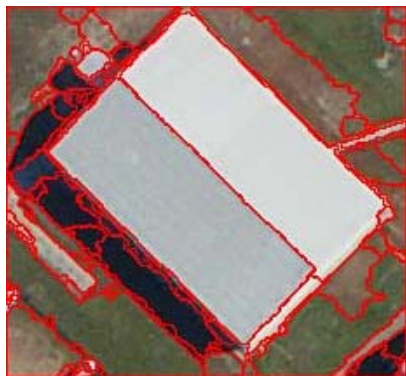


Figura 6. Filtro altimétrico



Figura 7. Douglas e Peucker



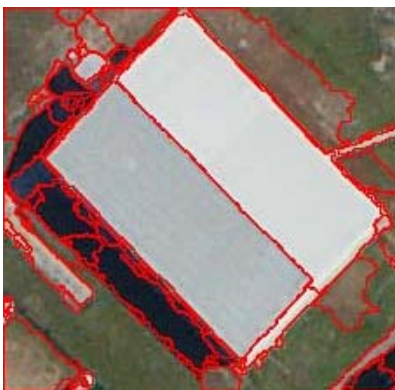
Figura 8. Fusão de segmentos



Figura 9. Vetorização



Figura 5. Filtro VERDE CIELUV



Na Figura 7 apresenta-se as regiões ainda existentes após a passagem do filtro Douglas-Peucker, evidenciando a eliminação das pequenas regiões esparsas, restando somente a edificação. A seguir, fundiram-se as regiões vizinhas existentes, sobrando uma única região (Figura 8).

Um detalhe importante neste experimento, e que acaba se constituindo em uma vantagem da metodologia proposta, é a identificação da linha de quebra da cobertura da edificação (cumeeira), que pode ser vista na Figura 7, mas não na Figura 8. O algoritmo de fusão de polígonos, de forma correta, fez com que esta linha desaparecesse. Chama-se atenção porque a detecção desta linha é muito interessante para os processos de reconstrução tridimensional das edificações, tópicos de grande importância na Fotogrametria, mas não tratada nesta pesquisa.

A fase de vetorização, devido a limiares com alta tolerância que impõem uma generalização, corrigiu, de certa forma, o contorno da edificação que foi prejudicado na segmentação (observar as diferenças entre as Figuras 8 e 9). A rotina de paralelismo entre segmentos de reta opostos, igualando o coeficiente angular para desvios de até cinco graus, contribuiu para a melhoria do aspecto final da vetorização.

As Figuras 10a e 10b retratam as vetorizações obtidas em outros dois experimentos, podendo-se verificar, na Figura 10b, os efeitos causados pela sombra de árvores.

Figura 10. Exemplos de vetorização



A avaliação da qualidade fotogramétrica dos resultados obtidos foi feita selecionando-se alguns pontos, dentre aqueles cujas vetorizações apresentaram correção aproximada de traçados. Tendo em vista esta abordagem, esta análise tem por objetivo verificar a correção dos procedimentos fotogramétricos.

Uma amostra com setenta e um (71) pontos, distribuídos pela imagem vista na Figura 1, foi observada. O processamento final da vetorização fornece as coordenadas de terreno de todos os seus pontos, polígono por polígono. Para se comparar tais coordenadas executou-se uma monorestituição manual sobre as mesmas feições.

Admitindo-se a monorestituição manual como referência, os desvios encontrados apresentaram as médias próximas de zero. As médias amostrais são:

$$\Delta E = 0,093 \pm 0,498\text{m} \quad (06)$$

$$\Delta N = -0,048 \pm 0,552\text{m} \quad (07)$$

$$\Delta H = -0,192 \pm 0,628\text{m} \quad (08)$$

Aplicando-se o teste de Kolmogorov-Smirnov para grandes amostras (mais que quarenta pontos) pode ser facilmente constatado que os resultados encontrados pela metodologia desenvolvida apresentam distribuição estatística semelhante à dos resultados da monorestituição manual, como se tivessem sido extraídos de uma mesma população.

Sendo estendidos os resultados da amostra para todas as medidas, segue-se que aproximadamente 68,26% das observações apresentarão uma discrepância da ordem de 0,5m (1 σ), em cada direção, e 95,45% da ordem de 1,0m (2 σ).

4. CONCLUSÕES

Este trabalho versou sobre a extração automática de contornos de edificações a partir de imagens tomadas por câmara digital de pequeno formato (não fotogramétrica) e dados LIDAR. As pesquisas centraram-se na diferenciação dos padrões de cor da imagem, buscando com isto equiparação aos procedimentos humanos. Para a avaliação deve-se atentar no aspecto de automação de todos os procedimentos propostos, realizados sem intervenção humana.

Os experimentos comprovaram que a metodologia proposta é capaz de distinguir edificações em ambientes urbanos, onde os dados LIDAR, por si só, não têm condições de fornecer elementos para a classificação.

Uma vantagem evidenciada nos experimentos é a capacidade de detecção da linha de quebra presente no topo de algumas edificações, informação inexistente (diretamente) nos dados LIDAR. Esta característica é muito importante nos trabalhos de reconstrução tridimensional das edificações, tópico de grande interesse em Fotogrametria, mas não deste trabalho.

Como desvantagem da metodologia mostrada pode-se citar a baixa performance obtida no traçado das edificações, isto é, em áreas urbanas densamente edificadas mais da metade das edificações não têm seus contornos adequadamente

vetorizados. Esta incapacidade enseja diversas alterações, agregando conhecimento aos algoritmos, que poderão ou não melhorar os resultados.

Não se deve esquecer que todos os experimentos foram efetuados com dados reais. As edificações existentes em alguns recortes, dentre os experimentados (ver MACHADO, 2006), especialmente em áreas com maior densidade de edificações, podem apresentar problemas de inconsistência entre si, se restituídas por operadores diferentes, ou com pouca experiência.

Ainda como desvantagem, os algoritmos apresentados dependem fortemente da etapa de segmentação. Se houver alguma segmentação errada, isto é, misturando-se regiões de topo de edificação com solo vizinho à edificação, como ocorreu com frequência nos experimentos, os demais algoritmos não tem condições de corrigir, ocasionando distorções na vetorização final.

Em edificações isoladas, sem sombras ou árvores pelas imediações da cobertura da mesma, os resultados foram animadores. As pesquisas continuam buscando melhorar os resultados da metodologia em situações com casas geminadas, árvores perto de casas, sombras, etc. Os limiares empregados estão sendo convertidos para distância de terreno. No tocante à segmentação está-se desenvolvendo uma adaptação envolvendo o espaço 4-dimensional formado pelas três dimensões do espaço de cores CIELUV mais uma dimensão dada pela altitude dos pontos do LIDAR.

AGRADECIMENTOS

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - CAPES, pelo suporte financeiro, em forma de bolsa de estudo e estágio no exterior (Calgary, Canadá).

À empresa AGRITEC, pelo fornecimento de imagens tomadas por câmara digital de pequeno formato Sony DSC-F717, de sua propriedade; e ao Instituto de Tecnologia para o Desenvolvimento (LACTEC - UFPR), pela cessão de levantamento efetuado por sistema LIDAR (OPTECH / INS GPS APPLANIX), ambos recursos indispensáveis, sem os quais não se poderia executar o presente trabalho.

BIBLIOGRAFIA

- ACKERMANN, F. Airborne Laser Scanning – Present Status and Future Expectations. *ISPRS Journal of Photogrammetry and Remote Sensing*. v. 54 (1), 1999. p. 64-67.
- AXELSSON, P. Processing of Laser Scanner Data – Algorithms and Applications. *ISPRS Journal of Photogrammetry and Remote Sensing*. Vol. 54 (1), 1999. p. 138-147.
- BALTSAVIAS, E. P. A Comparison between Photogrammetry and Laser Scanning. *ISPRS Journal of Photogrammetry & Remote Sensing*. v. 54 (1), 1999. p. 83-94.

- BRADSKI, G. R.; THRUN, S. Lecture 11: Segmentation and Grouping. Notas de aula da Universidade de Stanford. Disponível em: <<http://robots.stanford.edu/cs223b04/CS%20223-B%20L11%20Segmentation.ppt>>, 2003. Acesso em: 08 jan. 2006.
- COMANICIU, D.; MEER, P. Robust Analysis of Feature Spaces: Color Image Segmentation. In Proceedings IEEE International Conference on Computer Vision and Pattern Recognition, 1997.
- COMANICIU, D.; MEER, P. Mean Shift Analysis and Applications. In Proceedings IEEE International Conference on Computer Vision, 1999.
- COMANICIU, D.; MEER, P. Mean Shift: A Robust Approach Toward Feature Space Analysis. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2002. v. 24, n. 5, p. 603-619.
- DELARA, R. Calibração de Câmaras Digitais Não Métricas de Pequeno Formato Utilizando o Pixel como Unidade no Espaço-Imagem. Seminário apresentado no Curso de Pós-Graduação em Ciências Geodésicas da UFPR. Setembro, 2003.
- DOUGLAS, D. H.; PEUCKER, T. K. Algorithms for the Reduction of the Number of Points Required to Represent a Line or Its Caricature. The Canadian Cartographer. v. 10 (2), 1973.
- FUKUNAGA, K.; HOSTETLER, L. The Estimation of the Gradient of a Density Function with Applications in Pattern Recognition. IEEE Transactions on Information Theory, 1975. v. 21, n. 1, p. 32-40.
- HAALA, N.; BRENNER, C.; ANDERS, K.-H. Generation of 3D City Models from Digital Surface Models and 2D GIS. International Archives of Photogrammetry and Remote Sensing. 32 (3-4W2), 1997. p. 68-75.
- HABIB, A. F.; GHANMA, M.; MORGAN, M.; MITISHITA, E. Integration of Laser and Photogrammetric Data for Calibration Purposes. XX ISPRS Congress, Istanbul, Turkey, PS WG II/2 Systems for SAR and LIDAR Processing, 12-23 July 2004.
- KRAUS, K.; PFEIFER, N. Determination of Terrain Models in Wooded Areas with Airborne Laser Scanner Data. ISPRS Journal of Photogrammetry and Remote Sensing. v. 53 (4), 1998. p. 193-203.
- MACHADO, A. M. L.; MITISHITA, E. A.; DELARA, R. Estudo de Caso: Correção de Erros Sistemáticos em Câmaras Digitais Não Métricas. In.: Série em Ciências Geodésicas: Cartografia, Instrumento de Renovação Política e Inovação Tecnológica, v. 3, p. 130-144, 2004.
- MACHADO, A. M. L. Extração Automática de Contornos de Edificações Utilizando Imagem Gerada por Câmara Digital de Pequeno Formato e Dados LIDAR. Curitiba, 2006. Tese (Doutorado em Ciências Geodésicas) – Setor de Ciências da Terra, Universidade Federal do Paraná.
- MAKAROVIC, B. Digital Mono-Plotters. I.T.C. Journal, v. 1, p. 101-122, 1973.

- MITISHITA, E. A. Monorestituição Digital de Aerofotos, Associada com Sistema de Computação Gráfica C.A.D., para fins de Mapeamento na Área Florestal. Curitiba, 1997. Tese (Doutorado em Engenharia Florestal) – Setor de Ciências Agrárias, Universidade Federal do Paraná.
- PAVLIDIS, T. Algorithms for Graphics and Image Processing. Computer Science Press, 1982. ISBN 0-914894-65-X.
- SEUL, M.; O’GORMAN, L.; SAMMON, M. J. Practical Algorithms for Image Analysis: Description, Examples, and Code. Cambridge University Press, 2001. ISBN 0-521-66065-3.

(Recebido em Setembro de 2006. Aceito em novembro de 2006)

