

Uso Integrado de Mapas e Dados de Laser Scanner para o Reconhecimento Automático de Áreas Construídas

Jorge Silva Centeno
IPF - Karlsruhe University

✉ centeno@ipf.bau-verm.uni-karlsruhe.de

Hans-Peter Bähr
IPF - Karlsruhe University

✉ baehr@ipf.bau-verm.uni-karlsruhe.de

Regina T. Kishi
ISWW - Karlsruhe University

✉ kishi@iswws1.bau-verm.uni-karlsruhe.de

Conteúdo	<ol style="list-style-type: none"> 1. Introdução 2. Dados 3. Combinação de diferentes fontes de informação 4. Metodologia 5. Resultados e discussão 6. Conclusões 7. Agradecimentos 8. Referências
-----------------	--

Resumo: O presente artigo descreve um método que utiliza duas fontes de informação, mapas e dados de laser scanner, para o reconhecimento automático de áreas construídas. A segmentação inicial, utilizando exclusivamente informação de mapas, fornece indícios sobre a presença de prédios na imagem. Os dados do laser scanner servem para otimizar esta primeira segmentação na fase de pós-processamento. Morfologia matemática e classificação das áreas reconhecidas quanto a seu tamanho servem para definir melhor os contornos dos prédios. Um exemplo baseado numa carta 1:5000 serve para ilustrar o processo e uma discussão a respeito da exatidão dos resultados é apresentada.

Abstract: This paper describes an approach which combines two different information sources, scanned maps and laser scanner data, for automated recognition of buildings. An initial segmentation of the scanned map provides clues about the location of the buildings, which are later optimized using the information of the other source using post-classification sorting. Finally, mathematical morphology and a size filter are applied to delineate the contours of the buildings. Different steps of the methodology are illustrated and discussed using an example.

1. Introdução

Para a implantação de sistemas de apoio ao planejamento no âmbito urbano é necessário o conhecimento da geometria de áreas construídas que, na maioria das vezes, somente se encontra disponível sob forma de mapas. Atualmente, com o crescente uso de sistemas de informações geográficas (SIG), o uso de algoritmos de reconhecimento de padrões no processo de entrada de dados em um SIG tem levantado grande interesse, com a finalidade de acelerar a fase de digitalização de informações.

Uma das maneiras de abordar este problema é através do reconhecimento de documentos gráficos disponíveis, onde os mapas em formato analógico são convertidos ao formato raster e submetidos a um processo de reconhecimento de padrões (Förstner e Pallaske, 1993). Duas grandes correntes podem ser reconhecidas neste particular. A primeira transforma a imagem raster do mapa ao formato vetorial, como primeiro passo. A segunda baseia-se nos dados em formato raster, evitando com isso perdas de informação decorrentes de erros na vetorização. Além disso, o processamento de dados no formato raster é simples, pois rotinas de baixo nível podem ser implementadas com facilidade.

Novas tecnologias, como laser scanner, também podem ser utilizadas para a obtenção dos contornos de prédios. Estes dados permitem a análise de uma cena urbana a partir da elevação dos objetos nela localizados. Contudo, somente a elevação dos objetos não é suficiente para distinguir um prédio de outros objetos que possuem elevação similar.

O presente artigo descreve uma maneira de integrar estas duas fontes de informação para a obtenção automática da projeção horizontal de edificações em uma região urbana.

2. Dados

Duas fontes de dados são empregadas na metodologia aqui descrita. A primeira consiste em documentos cartográficos na escala 1:5000, nos quais as edificações estão desenhadas com detalhe. Uma imagem raster foi obtida a partir da carta utilizando um scanner. O tamanho do pixel foi fixado em 63,5 micrometros e a imagem foi obtida com 256 níveis de cinza. Ela foi binarizada para obter apenas duas classes: a primeira agrupa os pixels que correspondem às linhas ou símbolos da carta (aqui denominados FORE) e a segunda é formada pelos pixels que representam o fundo branco da imagem (denominados BACK).

A representação de prédios numa carta varia em função das normas especificadas para sua elaboração. No exemplo aqui apresentado, os prédios são desenhados usando polígonos hachurados, como é mostrado na Figura 1. Embora simples, esta representação pode ser confusa se se analisa apenas a imediata vizinhança do pixel para avaliar a presença de áreas hachuradas. Por exemplo, a região identificada como "A" na Figura 1 pode ser confundida com uma região hachurada, pois linhas equidistantes formam-se em função da superposição de caracteres sobre as linhas da rua. Interpretações ambíguas também ocorrem em regiões onde existem várias linhas paralelas como resultado da representação gráfica de elementos alheios aos prédios.

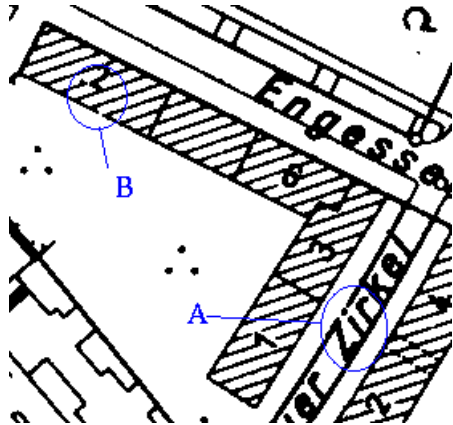


Fig. 1: Representação de prédios em mapas

A segunda fonte de informação é composta por dados de altitude levantados por um laser scanner. Estes dados são obtidos utilizando um sistema ativo de imageamento a bordo de uma aeronave que sobrevoa a área de estudo. O funcionamento do sistema se baseia na medição precisa do tempo entre a emissão de um sinal laser, sua reflexão na superfície dos objetos no terreno e sua chegada ao sensor. A posição da aeronave no momento da medição é determinada através de medições com GPS. Estes dados são combinados com informações auxiliares que descrevem a localização do sensor para calcular a distância entre o avião e o solo, a partir da qual se deduz a altitude dos pontos na superfície reflectora (Gutelius, 1998). A resolução espacial destes dados é elevada. Atualmente, levantamentos com resolução no terreno de 0,5m e precisão na vertical em torno de 0,1m são oferecidos comercialmente (Haala e Brenner, 1997). Um exemplo destes dados é mostrado na Figura 2, onde prédios e árvores aparecem na cena. Como pode ser visto, estes objetos apresentam altura e textura semelhantes, tornando difícil sua distinção somente com informação a respeito de sua altura.

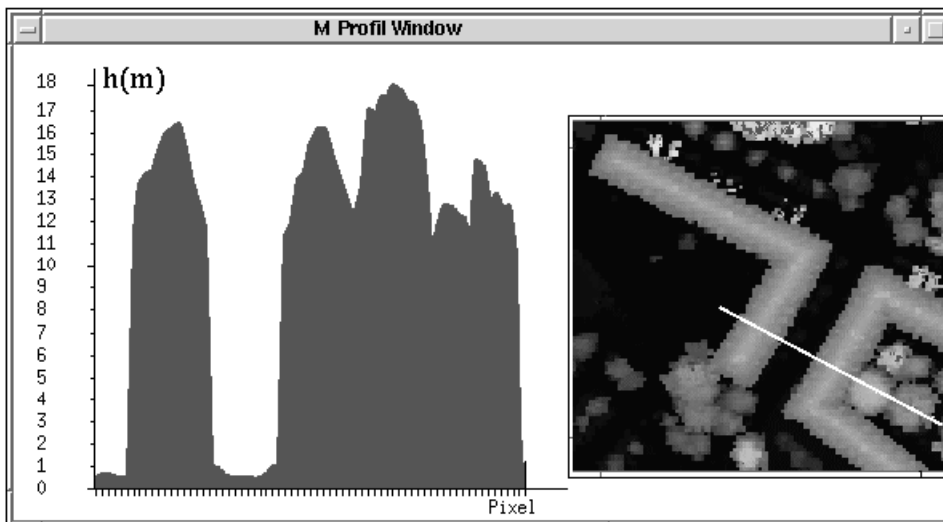


Fig. 2: Dados de altitude de laser scanner, em metros (Resolução horizontal 1m)

3. Combinação de diferentes fontes de informação

A natureza dos dados digitais provenientes de laser scanner e mapas digitalizados é diferente e o uso combinado de ambos demanda uma especial atenção. Do ponto de vista do processamento de dados, o tema pode ser abordado de diferentes maneiras. A primeira opção é integrar os dados em uma única análise. A segunda é realizar a análise usando apenas uma fonte de dados e otimizar os resultados usando para isto os dados da segunda fonte. Isto pode ser executado antes ou depois da análise (pré ou pós-processamento) (Hutchinson:82).

A integração de toda a informação num único vetor para descrever os pixels é aparentemente a melhor solução, pois possibilita uma descrição mais detalhada de cada pixel. No entanto, isto faz com que a dimensão do vetor de estado aumente em proporção igual ao número de variáveis, tornando a estrutura dos dados mais complexa. O aumento do volume de informação implica também o aumento do grau de dificuldade e tempo de processamento necessários. Sua grande vantagem, no entanto, reside no fato de a ordem, na qual os dados são introduzidos na análise, não influenciar o resultado, além de não depender da tomada de decisões subjetivas por parte do operador. A utilização de métodos estatísticos nesta análise é restrita em função da diferente natureza dos dados. Por exemplo, a

suposição de que os dados seguem uma distribuição Gaussiana não se verifica em se tratando de mapas, onde a informação é representada de maneira binária.

As informações auxiliares podem ser usadas na fase de pré ou pós-processamento para efetuar uma estratificação dos dados. No caso do pré-processamento, isto consiste em separar os dados de entrada em camadas de mais fácil interpretação. Cada camada pode ser assim processada separadamente, o que simplifica sua análise e melhora a qualidade dos resultados.

O uso de um conjunto de dados como informação auxiliar no pré-processamento ou no pós-processamento pode ser feito através da estratificação. No pré-processamento, isto consiste em separar os dados de entrada em camadas de mais fácil interpretação. Cada camada pode ser assim processada separadamente, o que simplifica sua análise e melhora a qualidade dos resultados. No caso do pós-processamento (Figura 3), trata-se de refinar a análise dos dados usando uma fonte de informação complementar. As classe que não puderem ser separadas na análise inicial são reavaliadas separadamente e um novo particionamento é obtido a partir dos dados auxiliares. Para efetuar a estratificação é necessário um conjunto de dados auxiliar, onde a separação das classes seja evidente.

A estratificação é um processo simples, mas demanda um bom conhecimento da estrutura dos dados, pois a escolha dos parâmetros usados tem grande influência no resultado final. As fronteiras entre camadas são definidas explícita e exatamente, motivo pelo qual deve-se levar em consideração que erros na sua fixação provocam erros na classificação dos pixels nas fronteiras entre classes. O método é altamente dependente da experiência do operador que determina a estratificação e eventualmente pode não ser sensível a diferenças sutis nos dados.

4. Metodologia

Para a integração das diferentes fontes de informação, mapas e dados do laser scanner, foi escolhida a metodologia da estratificação no pós-processamento. A estratégia utilizada para localizar os prédios se baseia na detecção no mapa de áreas hachuradas e a posterior otimização com a informação do laser scanner. Um esquema do processo é mostrado na Figura 3.

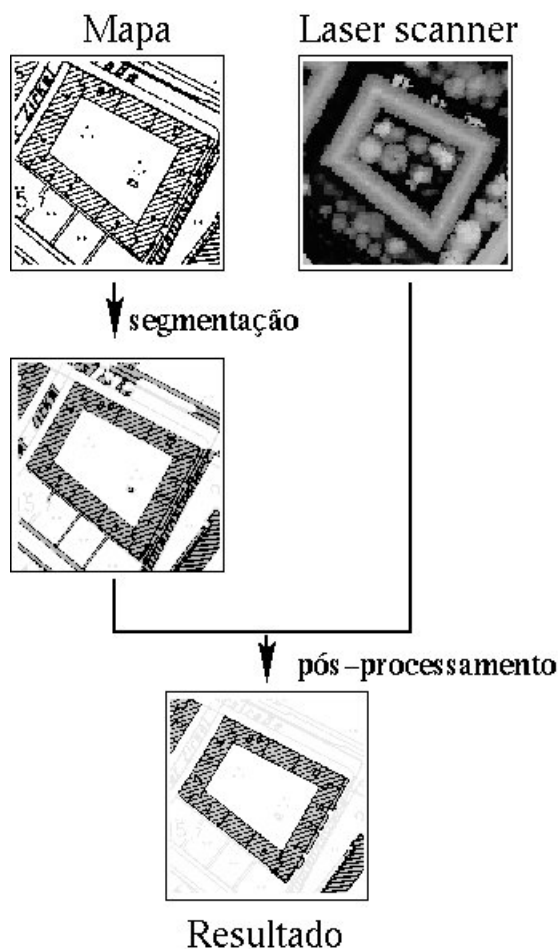


Fig. 3: Esquema do processo de identificação automática de prédios

A equidistância entre linhas é uma característica particular de áreas hachuradas. Ela cria espaços intermediários estreitos e de largura aproximadamente constante, cuja identificação pode ser usada como indício da presença de prédios (Brügelmann, 1996).

Para detectar estes espaços, utilizamos um processo similar àquele desenvolvido para simular o escoamento sobre modelos numéricos do terreno (Jenson e Domingue 1988). Este algoritmo define caminhos para o escoamento em função da variação do gradiente local calculado dentro de uma janela 3x3 em torno de cada pixel. O escoamento é simplificado, assumindo que ele pode ocorrer apenas em uma única direção. A direção preferencial do escoamento é escolhida como sendo aquela associada ao maior gradiente, dentre aqueles calculados para cada um dos oito vizinhos do pixel. Assim, ao ser estimada uma direção preferencial para cada pixel, define-se o caminho que o escoamento seguiria sobre o modelo numérico do terreno. A presença de depressões no modelo digital impede a escolha de uma direção preferencial nesses pontos. Para evitar estas indefinições, as depressões são previamente enchidas num processo iterativo. A determinação da cota de enchimento da depressão requer a estimativa da bacia contribuinte do ponto com escoamento indeterminado (mínimo local).

Embora este algoritmo tenha sido originalmente desenvolvido e aplicado para delinear bacias em modelos numéricos do terreno (Jenson e Domingue, 1988, Donker, 1992), aplicações do mesmo no processamento digital de imagens são conhecidas (Vincent e

Soille, 1991).

Se, na imagem obtida da carta, um valor máximo (255) é atribuído aos pixels da classe FORE e mínimo (0) aos da classe BACK e o algoritmo de Jenson e Domingue (1988) é aplicado, verifica-se que as regiões encerradas por pixels da classe FORE formam depressões que deverão ser preenchidas. As áreas entre as linhas das hachuras estão incluídas entre elas. No entanto, junto com as áreas intermediárias procuradas encontram-se outras que não pertencem às hachuras.

Uma variação mais suave dos valores dos pixels é gerada, se lhes são atribuídos valores de maneira inversamente proporcional à sua distância até a classe FORE, dentro de um raio pré-determinado. Quando a distância entre os pixels da classe FORE é pequena em relação ao raio especificado, mínimos locais com valores acima de zero surgem dentro das áreas internas. Pelo contrário, quando os pixels FORE encontram-se afastados, os mínimos continuam com valor zero. Esta propriedade serve para diferenciar as regiões quanto a sua largura. As bacias contribuintes dos mínimos locais de cota maior que zero satisfazem a condição da distância máxima aceita para considerar que a região em questão encontra-se dentro de uma área hachurada.

A **Figura 4** mostra um esquema simplificado de uma linha de uma imagem binária, onde a altura das barras representam os valores dos níveis de cinza dos pixels (NC). Nesta figura, 6 pixels correspondem à classe FORE (valores máximos) e os restantes à classe BACK (valores nulos).

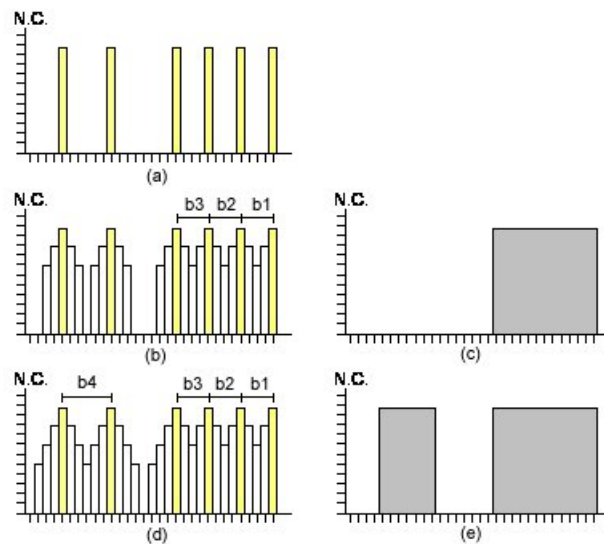


Fig. 4: Exemplo simplificado da identificação de hachuras. (a) linha original; (b) níveis de cinza gerados em função da distância com raio de 2 pixels; (c) regiões após o agrupamento das bacias com distância máxima de 2 pixels; (d) níveis de cinza gerados em função da distância com raio de 3 pixels; (e) regiões após o agrupamento das bacias com distância máxima de 3 pixels.

Atribuindo aos pixels valores decrescentes em função da distância à classe FORE, dentro de um raio de 2 pixels, é obtida a distribuição mostrada na Figura 4b, com 3 mínimos locais maiores que zero. Estes mínimos locais devem ser preenchidos para garantir que o escoamento seja definido ao longo de toda a linha. As bacias contribuintes destes mínimos estão identificadas como b1, b2 e b3. Se, por outro lado, é utilizado um raio de 3 pixels (**figura 4d**), surge, além dos 3 anteriores, mais um mínimo local, cuja bacia contribuinte é identificada como b4. Nota-se que, o aparecimento de mínimos locais maiores que zero pode ser controlado pela variação do raio fixado. Juntando estas bacias, obtém-se um agrupamento dos pixels em função da distância entre os pixels da classe FORE, como mostram as **figuras 4c e 4e**.

Aplicando este algoritmo à imagem do mapa, as regiões hachuradas foram caracterizadas, embora muitas outras, que não correspondem às áreas construídas, também estejam presentes por satisfazerem o critério de distância entre pixels FORE (**Figura 5**).

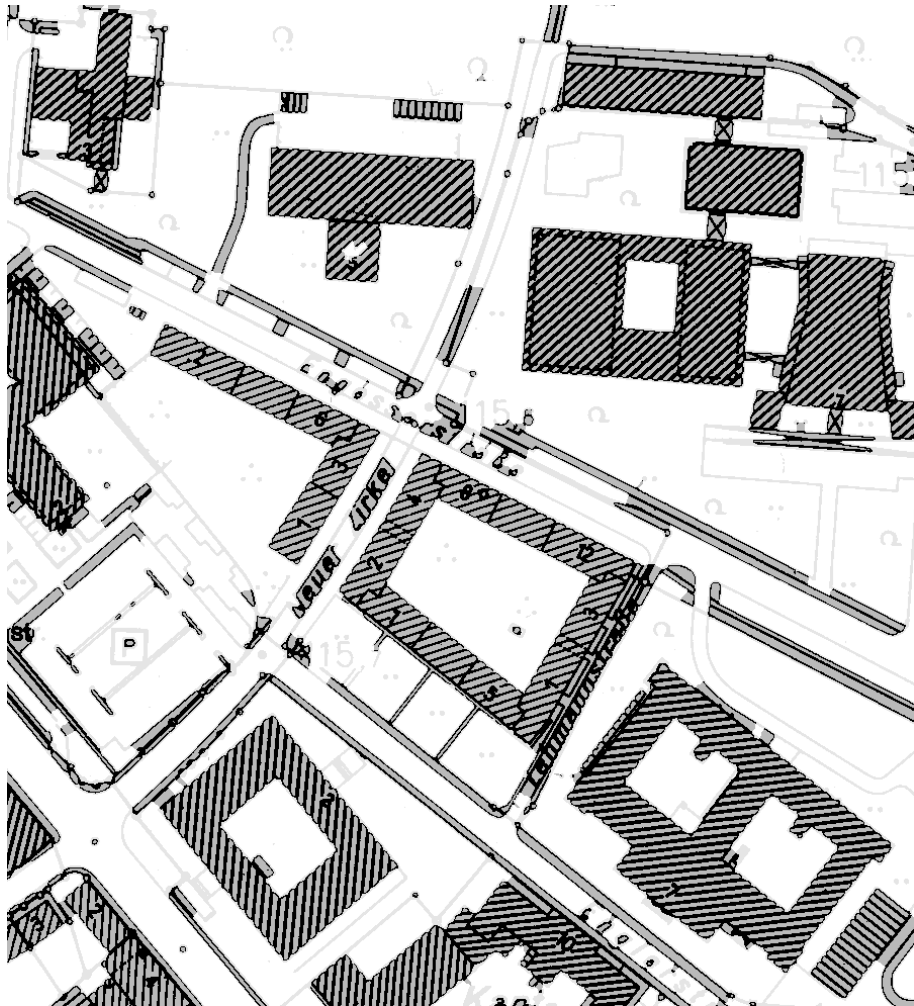


Fig. 5: Resultado da segmentação inicial do mapa, superposto à imagem original

A informação da altitude, derivada dos dados do laser scanner, é usada para eliminar regiões que não podem ser prédios por apresentarem baixa altitude. Isto é feito fixando uma altura mínima para os píxeis. No exemplo apresentado, este valor é de 2,80m. Esta restrição reduz a quantidade de píxeis pré-selecionados de maneira significativa, pois eliminam aqueles localizados entre linhas muito próximas que dão aparência de hachuras, mas que não tem a altura pré-definida e portanto não fazem parte de prédios. O resultado deste passo é mostrado na [figura 6](#).

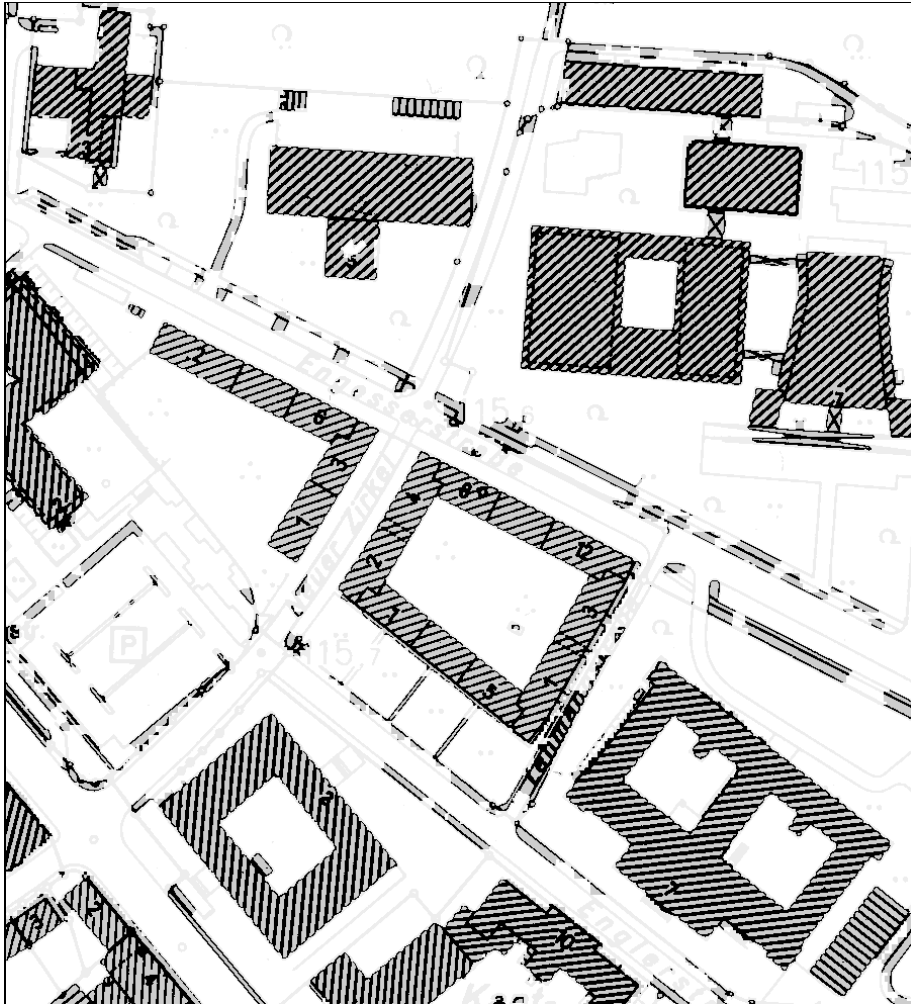


Fig. 6: Imagem após a inclusão da informação do laser scanner, superposta à imagem original

Finalmente, operadores de morfologia matemática (opening and closing) (Haralick e Shapiro, 1997) são utilizados para eliminar perturbações nas bordas dos prédios e definir melhor seus contornos. Regiões muito pequenas, que não podem ser prédios, são facilmente eliminadas nesta fase utilizando como critério o tamanho de cada região.

5. Resultados e discussão

A figura 7 mostra o resultado final do processo, bem como as áreas que foram eliminadas no pós-processamento. O tom cinza foi usado para representar os prédios identificados e o branco para as regiões descartadas. Os prédios foram bem identificados, embora algumas irregularidades em suas bordas ainda estejam presentes. Os erros remanescentes nos contornos dos edifícios não puderam ser reduzidos, porque eles correspondem a píxeis com altura maior que o limiar fixado e se encontram dentro de regiões circundadas por píxeis da classe FORE no mapa. Assim sendo, essas regiões não podem ser descartadas nem pelo critério da altura mínima, nem pelos operadores de morfologia matemática.



Fig. 7: Resultado final. As áreas indetificadas aparecem em tom cinza e em branco as áreas eliminadas no pós-processamento

Um detalhe de um prédio corretamente identificado é apresentado na **figura 8**. O afastamento entre os pontos que formam os polígonos das regiões identificadas e polígonos digitalizados manualmente foi calculado para cada prédio, com a finalidade de complementar a avaliação do resultado. A estatística deste desvio é mostrada na **tabela 1**. Os desvios nas direções x e y (S_x e S_y), bem como o desvio total (S_{xy}), são estimados a partir das distâncias d_x e d_y entre cada ponto e sua projeção no polígono digitalizado, considerado como verdadeiro.

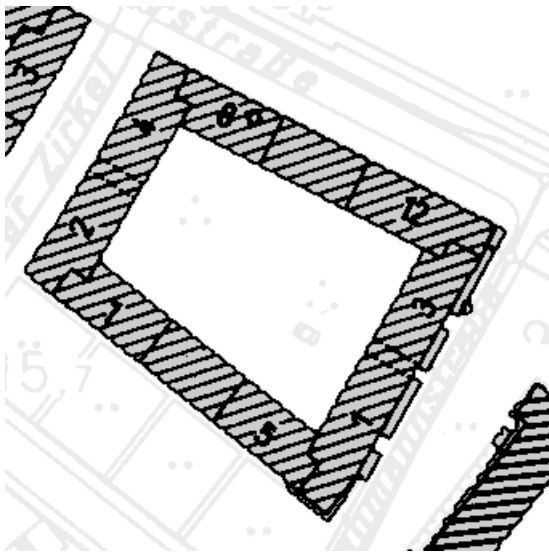


Fig. 8: Exemplo de prédio identificado superposto à imagem original

Através desta estatística, uma circunferência, que representa o erro médio, pode ser estimada (Scheuring, 1995). Diferentes estimativas do raio desta circunferência, CEP (desvio circular provável), RMS (desvio quadrático médio) e CES (desvio padrão médio), são dadas na mesma tabela. Os valores do desvio na **tabela 1** estão próximos de 1m, o que equivale a 1 pixel na resolução dos dados do laser scanner. Quanto ao mapa, eles ficam abaixo da dimensão mínima recomendada para representar graficamente elementos em documentos cartográficos, fixada em 0,3mm na Alemanha, que corresponde a 1,5m na escala 1:5000 (Hake e Grünreich, 1994).

Tabela 1: Estatística do erro.

Estatística	Píxeis	Metros
$S_x = \text{SQRT}\{ \text{SUM}(d_x^2) / n \}$	3,09	0,98
$S_y = \text{SQRT}\{ \text{SUM}(d_y^2) / n \}$	2,67	0,85
$S_{xy} = \text{SUM}(d_x * d_y) / n$	1,25	0,40
$\text{RMS} = \text{SQRT}\{ S_x^2 + S_y^2 \}$	4,08	1,29
$\text{CEP} = 0,59 * (S_x + S_y)$	3,40	1,08
$\text{CES} = 0,5 * (S_x + S_y)$	2,88	0,91

6. Conclusões

Uma metodologia para a identificação de prédios utilizando uma combinação de informação derivada de mapas e dados de laser scanner foi apresentada e discutida. A abordagem utiliza apenas operações de processamento digital de baixo nível, o que a torna de fácil implementação. Isto, por outro lado, faz com que erros ainda estejam presentes no resultado final, pois os critérios empregados são relativamente simples. No entanto, os resultados obtidos são bons, pois os prédios foram corretamente localizados e caracterizados.

Os dados de entrada, mapas e dados de altitude, quando analisados isoladamente, apresentam dificuldades para a identificação dos prédios, pois interpretações ambíguas surgem em ambos os casos. Mas, devido à diferente natureza dos dados, as ambiguidades de cada um deles não se repetem na outra fonte de informação e isto possibilita aproveitar as vantagens que cada um deles oferece. Assim sendo, seu uso combinado reduz substancialmente os erros na segmentação, compensando mutuamente suas deficiências.

A aplicação desta metodologia é útil no planejamento de áreas urbanas usando processamento digital e SIG. Há um ganho significativo na substituição da fase de digitalização manual de edificações por rotinas de processamento digital de imagens, liberando o analista da mesa digitalizadora e estendendo seu tempo disponível para a análise.

7. Agradecimentos

Os autores agradecem o apoio do CNPq-Brasil e DAAD -Alemanha.

8. Referências

Brügelmann, R.: *Recognition of Hatched Cartographic Patterns*. In: The International Archives of Photogrammetry and Remote Sensing, ISPRS Commission III, Vol.31-B3, ISPRS, pp.92-97, 1996.

Donker, N: *Automatic Extraction of Catchment Hydrological Properties from Digital Elevation Data*. ITC Journal, pp.257-265, 1992.

Förstner, W. e Pallaske, R.: *Mustererkennung und 3D-Geoinformationssysteme*. Zeitschrift für Photogrammetrie und Fernerkundung 5, pp.167-177, 1993.

Gutelius, B.: *Engineering Applications of Airborne Scanning Lasers: Report from the Field*. Photogrammetric Engineering and Remote Sensing 64(4), pp.246-253, 1998.

Haala, G. e Brenner, C.: *Generation of 3D City Models from Airborne Laser Scanning Data*. In: EARSEL Workshop on LIDAR Remote Sensing of Land and Sea, Tallinn, Estonia, pp.105-112, 1997.

Hake, G. e Grünreich, D.: *Kartographie*. 7ma. edição. Ed. Springer, Berlin, 1994.

Haralick, R. e Shapiro, L. : *Computer and Robot Vision*. Vol.1, Addison-Wesley Publishing Company Inc., 1997.

Hutchinson, C.: *Techniques for Combining Landsat and Ancillary Data for Digital Classification Improvement*. Photogrammetric Engineering and Remote Sensing 48(1), pp.123-130. 1982.

Jenson, S. and Domingue, J.: *Extraction of Topographic Structure from Digital Elevation Data for Geographic Information System Analysis*. Photogrammetric Engineering and Remote Sensing 54(11), pp.1593-1600, 1988.

Scheuring, R.: *Zur Qualität der Basisdaten von Landinformationssystemen*. Dissertação, Univ. der Bundeswehr, Munique. 1995.

Vincent, L. and Soille, P.: *Watershed in Digital Spaces: An Efficient Algorithm on Immersion Simulation*. IEEE Transactions on Pattern Analysis and Machine Intelligence 13(6), pp.583-589, 1991.