

**UNIVERSIDADE ESTADUAL PAULISTA**

**FACULDADE DE CIÊNCIAS E TECNOLOGIA**

*Programa de Pós-Graduação em Ciências Cartográficas*

**MICHELLE SAYURI YANO YWATA**

**EXTRAÇÃO AUTOMÁTICA DE CONTORNOS DE TELHADOS DE EDIFÍCIO NO  
ESPAÇO-OBJETO INTEGRANDO UM ESTÉREO PAR DE IMAGENS AÉREAS DE  
ALTA RESOLUÇÃO E MODELOS 3D DE TELHADO**



PRESIDENTE PRUDENTE

2019

**MICHELLE SAYURI YANO YWATA**

**EXTRAÇÃO AUTOMÁTICA DE CONTORNOS DE TELHADOS DE EDIFÍCIO NO  
ESPAÇO-OBJETO INTEGRANDO UM ESTÉREO PAR DE IMAGENS AÉREAS DE  
ALTA RESOLUÇÃO E MODELOS 3D DE TELHADO**

Tese de doutorado apresentada ao Programa de Pós-Graduação em Ciências Cartográficas da Universidade Estadual Paulista “Júlio de Mesquita Filho” – Faculdade de Ciências e Tecnologia – Campus de Presidente Prudente, como parte dos requisitos para a obtenção do título de Doutor em Ciências Cartográficas.

Orientador: Prof. Dr. Aluir P. Dal Poz

Coorientador: Prof. Dr. Milton H. Shimabukuro

Y99e

Ywata, Michelle Sayuri Yano

Extração automática de contornos de telhados de edifício no espaço-objeto integrando um estéreo par de imagens aéreas de alta resolução e modelos 3D de telhado / Michelle Sayuri Yano Ywata. -- Presidente Prudente, 2019

122 p.

Tese (doutorado) - Universidade Estadual Paulista (Unesp), Faculdade de Ciências e Tecnologia, Presidente Prudente

Orientador: Aluir Porfírio Dal Poz

Coorientador: Milton Hirokazu Shimabukuro

1. Extração de contornos de telhado. 2. Imagem aérea. 3. Dados LASER. 4. Snakes. 5. Programação dinâmica. I. Título.

Sistema de geração automática de fichas catalográficas da Unesp. Biblioteca da Faculdade de Ciências e Tecnologia, Presidente Prudente. Dados fornecidos pelo autor(a).

Essa ficha não pode ser modificada.

**CERTIFICADO DE APROVAÇÃO**

TÍTULO DA TESE: Extração automática de contornos de telhados de edifícios no espaço-objeto integrando um estéreo par de imagens aéreas de alta resolução e modelos 3D de telhados

**AUTORA: MICHELLE SAYURI YANO**

**ORIENTADOR: ALUIR PORFIRIO DAL POZ**

**COORIENTADOR: MILTON HIROKAZU SHIMABUKURO**

Aprovada como parte das exigências para obtenção do Título de Doutora em CIÊNCIAS CARTOGRÁFICAS, área: Aquisição, Análise e Representação de Informações Espaciais pela Comissão Examinadora:

  
Prof. Dr. ALUIR PORFIRIO DAL POZ  
Departamento de Cartografia / Faculdade de Ciências e Tecnologia de Presidente Prudente

  
Prof. Dr. MAURÍCIO GALO  
Departamento de Cartografia / Faculdade de Ciências e Tecnologia de Presidente Prudente

  
Prof. Dr. ADILSON BERVEGLIERI  
Departamento de Estatística / Câmpus de Presidente Prudente

Prof. Dr. EDSON APARECIDO MITISHITA  
Departamento de Geomática / Universidade Federal do Paraná

Prof. Dr. HENRIQUE CANDIDO DE OLIVEIRA  
Departamento de Geotecnia e Transportes / UNICAMP

VIDEOCONFERÊNCIA

VIDEOCONFERÊNCIA

Presidente Prudente, 26 de agosto de 2019

## DEDICATÓRIA

*À minha batiam, Mitsuko Tomiyoshi, por todos os ensinamentos  
e boas lembranças deixadas em meu coração.*

ありがとうございました。

## **AGRADECIMENTOS**

Ao meu grande amor, Fuyu, por todo o apoio, ajuda, ensinamentos e companheirismo durante todos esses anos.

Aos meu pais, Celso e Beth, pela educação e oportunidades a mim proporcionadas durante toda a minha vida. Ao meu irmão, Gabriel, pelo apoio e confiança de um irmão caçula.

Ao meu orientador Prof. Dr. Aluir Dal Poz, pela orientação, ensinamentos e confiança depositada em mim desde a graduação. Ao meu coorientador Prof. Dr. Milton Shimabukuro, pelos ensinamentos e contribuições dadas para o desenvolvimento deste trabalho. Ao Prof. Dr. Ayman Habib, pela recepção na Purdue University e por todo apoio durante o meu doutorado sanduíche.

Aos professores do Departamento de Cartografia e do PPGCC, que contribuíram para minha formação acadêmica, desde a graduação até o doutorado.

Aos amigos do PPGCC, principalmente aqueles que estão comigo desde o primeiro ano da graduação e aqueles que fizeram parte do mesmo grupo de pesquisa.

Ao meu amigo Prof. Dr. Érico Fernando de Oliveira Martins pelo fornecimento do código de Programação Dinâmica utilizado como base no desenvolvimento do programa implementado neste trabalho.

Ao meu amigo Prof. Dr. Henrique Candido de Oliveira pela geração dos mapas de visibilidade utilizados neste trabalho.

À FCT/UNESP, ao Programa de Pós-Graduação em Ciências Cartográficas (PPGCC) e à Purdue University por proporcionar todos os meios para o desenvolvimento deste trabalho.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código de Financiamento 001.

為せば成る  
為さねば成らぬ何事も  
成らぬは人の為さぬなりけり

*“Se você tentar, você pode ter sucesso.*

*Se você não tentar, não terá sucesso.*

*Isso é verdade para todas as coisas.*

*Não ter sucesso é o resultado de não tentar.”*

(Uesugi Harunori)

## RESUMO

Neste trabalho foi proposta uma metodologia para a extração de contornos de telhados de edifícios no espaço-objeto, a partir da integração de um estéreo par de imagens aéreas de alta resolução e modelos 3D aproximados de telhado obtidos a partir de dados de varredura a LASER. Um modelo matemático considerando as propriedades radiométricas e geométricas dos telhados foi formulado a fim de representar o contorno do telhado no espaço-imagem, tendo como base o modelo de contorno ativo *Snake*. Esse modelo foi então adaptado para descrever os contornos no espaço-objeto considerando um estéreo par de imagens aéreas. Finalmente, o polígono ótimo que representa um dado contorno do telhado foi determinado a partir da otimização, via Programação Dinâmica, da função de energia criada. A solução obtida é uma representação mais acurada para o correspondente contorno do modelo 3D do telhado. O método desenvolvido apresenta também mecanismos para realizar a compensação automática de três tipos de problemas comuns em ambientes urbanos e que podem prejudicar a extração automática de telhados: obstruções perspectivas causadas por edifícios elevados, obstruções diretas causadas por vegetação que se eleva acima do telhado e sombras adjacentes aos telhados, as quais podem ser confundidas com as bordas do telhado. Os experimentos foram realizados utilizando imagens aéreas com GSD  $\approx 0,10$  m e nuvem de pontos LASER com densidade média de 6 pontos/m<sup>2</sup>. Os resultados mostraram que o método funciona adequadamente na extração dos contornos de telhados em áreas contendo obstruções e sombras, apresentando valores médios de completeza e correção acima de 90%, e valores médios de RMSE abaixo de 0,5 m para as componentes E e N, e abaixo de 1 m para a componente H.

**Palavras-chave:** Extração de telhados, dados LASER, imagens aéreas, *Snake*, Programação Dinâmica.



## ABSTRACT

In this work a methodology was proposed for extracting building roof contours in the object-space, by integration of a high-resolution aerial images stereo pair and 3D roof models reconstructed from LASER scanning data. A mathematical model considering the radiometric and geometric properties of roofs was developed in order to represent the roof contour in the image-space, based on the Snake active contour model. Then, the model was adapted to represent the contours in the object space considering a stereo pair of aerial images. Finally, the optimal polygon representing a selected roof contour was obtained by optimizing the proposed energy function using Dynamic Programming algorithm. The solution obtained, i.e., a polygon representing each 3D roof contour, will be a higher accurate representation for the correspondent contour of the 3D roof model. The proposed method also presents mechanisms to perform the compensation of three types of common problems in urban environment and which can disturb the automatic roof extraction: perspective occlusions caused by high buildings, occlusions caused by vegetation that covers the roof and shadows that are adjacent to the roofs which can be misinterpreted as roof edges. The experiments were performed using aerial images with GSD  $\approx 0,10$  m and LASER point cloud with average density of 6 points/m<sup>2</sup>. The results showed that the proposed method works properly in contour extraction of roofs with occlusion and shadows areas, presenting completeness and correctness average values above 90%, RMSE average values below 0,5 m for E and N components, and below 1 m for H component.

**Keywords:** roof extraction, LiDAR data, aerial images, Snake, Dynamic Programming.

## LISTA DE FIGURAS

Figura 1 – Sistema fotogramétrico. ....	22
Figura 2 – Sistema fotográfico de coordenadas do negativo.....	22
Figura 3 – Sistema de imagem. ....	23
Figura 4 – Condição de colinearidade entre o CP, o ponto imagem $p$ e o ponto no espaço-objeto $P$ .....	25
Figura 5 – Distorção radial simétrica. ....	26
Figura 6 – Efeitos da distorção radial simétrica. ....	26
Figura 7 – Distorção descentrada. ....	27
Figura 8 – Perfil dos tipos de bordas existentes em uma imagem.....	29
Figura 9 – Detecção de bordas por operadores de derivação. ....	30
Figura 10 – Máscara de gradiente. (a) Região de tamanho 3x3 em uma imagem; (b) Operadores de Sobel; (c) Operadores de Prewitt.....	32
Figura 11 – Linha do tempo da criação dos principais operadores de interesse. ....	33
Figura 12 – Janelas 3x3 e os vetores nas direções 0°, 45°, 90° e 135°. ....	34
Figura 13 – (a) Visada em perspectiva central. (b) Visada ortogonal. ....	37
Figura 14 – Ilustração do efeito de duplo mapeamento.....	38
Figura 15 – (a) Imagem original em projeção perspectiva. (b) Efeito de duplo mapeamento. (c) Área de oclusão (região em vermelho). ....	39
Figura 16 – Situações em que ocorrem múltiplos retornos. ....	41
Figura 17 – Componentes de um sistema de VLA e suas relações geométricas.....	43
Figura 18 – Perfil gerado pela conexão de pontos LASER sequenciais numa região de vegetação. ....	49
Figura 19 – Exemplos de elementos estruturantes. ....	50
Figura 20 – Efeitos da erosão utilizando o EE octógono de tamanho 21 pixels. ....	52
Figura 21 – Efeitos da dilatação utilizando o EE octógono de tamanho 21 pixels. ....	54
Figura 22 – Exemplo da transformação <i>top-hat</i> por abertura em um sinal unidimensional. ...	56
Figura 23 – Exemplo da transformação <i>top-hat</i> por fechamento em um sinal unidimensional. .....	56
Figura 24 – Representação geométrica de uma curva <i>Snake</i> .....	61
Figura 25 – Problema clássico de Programação Dinâmica. ....	66
Figura 26 – Fluxograma do método proposto. ....	72

Figura 27 – (a) Amostra do mapa de visibilidade, em azul a região oclusa. (b) Imagem aérea correspondente.....	74
Figura 28 – Etapas da classificação da nuvem de pontos LASER.....	75
Figura 29 – Pontos de vegetação.....	76
Figura 30 – Ilustração do elemento estruturante octógono de tamanho 3.....	77
Figura 31 – Exemplo de imagem de vegetação gerada pelo método proposto.....	77
Figura 32 – Exemplo de imagem de sombra gerada a partir de uma imagem sem suavização. (a) Detalhes de uma imagem aérea de alta resolução. (b) Sombras detectadas (em branco).....	78
Figura 33 – Imagem de sombra gerada pelo método proposto.....	79
Figura 34 – Fator de compensação de sombra.....	82
Figura 35 – Ideia utilizada na criação do termo CS para o cálculo da resposta de quina.....	84
Figura 36 – Princípio da ambiguidade.....	87
Figura 37 – Criação do espaço de busca 1D no espaço 3D.....	89
Figura 38 – Criação do espaço de busca 2D no espaço 3D.....	90
Figura 39 – Criação do espaço de solução no espaço 3D.....	92
Figura 40 – Aplicação da PD.....	95
Figura 41 – Gráficos do número total de combinações obtidas em cada algoritmo a partir de diferentes quantidades de pontos no espaço de solução.....	101
Figura 42 – Áreas utilizadas no Experimento 1 contendo telhados cobertos por vegetação.....	102
Figura 43 – Resultado visual do Experimento 1.....	104
Figura 44 – Área utilizada no Experimento 2 contendo sombras adjacente aos telhados.....	106
Figura 45 – Resultado visual do Experimento 2.....	108
Figura 46 – Áreas utilizadas no Experimento 3 contendo obstruções perspectivas.....	110
Figura 47 – Resultado visual do Experimento 3.....	112
Figura 48 – Detalhe da mancha branca no telhado que atrapalhou o delineamento correto da borda, contorno detectado (verde) e contorno correto (rosa tracejado).....	113

## LISTA DE TABELAS

Tabela 1 – Tabela da matriz de “custos parciais”.....	67
Tabela 2 – Tabela da matriz de "custos parciais" $g_2(e_2, e_3)$ .....	68
Tabela 3 – Tabela da matriz de "custos máximos" $f_1(e_2)$ .....	68
Tabela 4 – Tabela da matriz de "custos acumulados" $f_1(e_2) + g_2(e_2, e_3)$ .....	68
Tabela 5 – Tabela da matriz de "custos máximos".....	69
Tabela 6 – Última tabela da matriz de "custos máximos".....	69
Tabela 7 – Caminho inverso com a superposição de 1 nó.....	69
Tabela 8 – Parâmetros e limiares utilizados.....	98
Tabela 9 – Testes de comparação dos algoritmos.....	100
Tabela 10 – Comparação entre a quantidade de pontos no espaço de solução e número de combinações resultantes em cada algoritmo.....	101
Tabela 11 – Parâmetros de qualidade dos contornos iniciais do Experimento 1.....	105
Tabela 12 – Parâmetros de qualidade dos contornos extraídos pelo Experimento 1.....	105
Tabela 13 – Parâmetros de qualidade dos contornos iniciais do Experimento 2.....	109
Tabela 14 – Parâmetros de qualidade dos contornos extraídos pelo Experimento 2.....	109
Tabela 15 – Parâmetros de qualidade dos contornos iniciais do Experimento 3.....	113
Tabela 16 – Parâmetros de qualidade dos contornos extraídos pelo Experimento 3.....	113

## LISTA DE SIGLAS

1D	Unidimensional
2D	Bidimensional
3D	Tridimensional
BTH	<i>Black Top-hat</i>
CCD	<i>Charge Couple Device</i>
CP	Centro perspectivo
EE	Elemento estruturante
EI	Espaço-imagem
RMSE	Raiz do erro médio quadrático
EO	Espaço-objeto
FCT	Faculdade de Ciências e Tecnologia
GNSS	<i>Global Navigation Satellite System</i>
GSD	<i>Ground Sample Distance</i>
IFOV	<i>Instantaneous Field of View</i>
IMU	<i>Inertial Measurement Unit</i>
INS	<i>Inertial Navigation System</i>
ISPRS	<i>International Society for Photogrammetry and Remote Sensing</i>
LASER	<i>Light Amplification by Stimulated Emission of Radiation</i>
LPS	<i>Leica Photogrammetry Suite</i>
MDS	Modelo digital de superfície
MDT	Modelo digital de terreno
MM	Morfologia matemática
PD	Programação Dinâmica
POE	Parâmetros de orientação exterior
POI	Parâmetros de orientação interior
PP	Ponto principal
SGBM	<i>Surface-Gradient-Based Method</i>
TIM	<i>Triangulated Irregular Network</i>
TM	Transverso de Mercator
UNESP	Universidade Estadual Paulista
UTM	Universal Transverso de Mercator
VLA	Varredura a LASER aerotransportado
WGS84	<i>World Geodetic System 1984</i>
WTH	<i>White Top-hat</i>

## SUMÁRIO

<b>1. INTRODUÇÃO .....</b>	<b>14</b>
1.1. Considerações iniciais.....	14
1.2. Hipótese e Objetivos .....	17
1.3. Justificativa .....	18
1.4. Estrutura do trabalho.....	19
<b>2. FUNDAMENTAÇÃO TEÓRICA .....</b>	<b>21</b>
2.1. Sistemas de coordenadas .....	21
2.2. Conceito de borda .....	28
2.3. Conceito de quina .....	33
2.4. Conceito de sombra .....	35
2.5. Áreas de oclusão e mapa de visibilidade .....	36
2.6. Sistema de varredura a LASER aerotransportado (SVLA) .....	40
2.7. Morfologia Matemática .....	50
2.8. Modelo de contorno ativo: <i>Snakes</i> .....	57
2.9. Programação Dinâmica .....	65
<b>3. MATERIAL E MÉTODOS .....</b>	<b>70</b>
3.1. Material .....	70
3.2. Métodos .....	71
3.3. Forma de análise dos resultados .....	95
<b>4. RESULTADOS .....</b>	<b>97</b>
4.1. Introdução .....	97
4.2. Parâmetros e limiares utilizados .....	97
4.3. Resultados e análise .....	98
<b>5. CONCLUSÕES E RECOMENDAÇÕES .....</b>	<b>114</b>
<b>REFERÊNCIAS .....</b>	<b>117</b>

## **1. INTRODUÇÃO**

### **1.1. Considerações iniciais**

Atualmente a humanidade está em uma era na qual as informações espaciais desempenham um importante papel para o público em geral. No entanto, o grande desafio continua sendo fornecer informações de alta qualidade, para permitir o processamento geoespacial avançado e apoiar a solução colaborativa de problemas (CHEN et al., 2016a). Diversas questões urbanas como a modelagem e visualização da poluição, atualizações cadastrais, estimativa de volume predial, planejamento urbano e projeto de redes de comunicação tem utilizado os modelos tridimensionais das cidades como base de estudo (BENCIOLINI et al., 2018).

A reconstrução tridimensional dos edifícios presentes nos modelos de cidades envolve as etapas de detecção, extração do contorno e reconstrução do telhado. Embora muitos esforços tenham sido feitos nas últimas décadas para o desenvolvimento de novos métodos de extração de edifícios, a extração exata e automática ainda continua sendo um desafio para as comunidades de Sensoriamento Remoto e Visão Computacional (HUANG et al., 2019). Isso se deve às condições de complexidade e variabilidade das cenas, que torna desafiadora a questão de extração de telhados no âmbito do reconhecimento de objetos. Dessa forma, se torna interessante o desenvolvimento de metodologias que explorem a sinergia de diferentes fontes de dados para realizar essa tarefa, visto que cada fonte possui as suas vantagens e desvantagens. Diversos autores (ROTTENSTEINER et al., 2014; GERKE e XIAO, 2014; ZHOU e ZHOU, 2014; WU et al., 2015; OLIVEIRA, 2016a; LARI et al., 2017) têm apresentado a integração entre imagem e dados de varredura a LASER (*Light Amplification by Stimulated Emission of Radiation*) aerotransportado (VLA) para a extração de edifícios. Essa combinação é atrativa, uma vez que esses dois tipos de dados possuem informações complementares. Quando se trata da obtenção de planos de telhados e sua orientação, os dados de VLA fornecem melhores resultados, pois eles apresentam informações altimétricas mais precisas. Já com relação à extração de contornos de telhados, as imagens ópticas fornecem melhores resultados se comparados com dados de VLA de baixa ou média densidade, uma vez que elas possuem melhores informações posicionais ao longo das linhas de quebra (DAL POZ et al., 2009).

No entanto, a questão de como integrar de forma eficiente essas duas fontes de dados para a extração de edifícios, ainda precisa ser resolvida e continua sendo estudada por diversos autores (CHEN et al., 2016a). No trabalho de Gerke e Xiao (2014) uma nova abordagem para o processo de segmentação de áreas urbanas é apresentada, integrando a qualidade geométrica dos dados de VLA e as informações espectrais presentes em imagens, para detectar edifícios, árvores e terreno. Os atributos de geometria e textura da imagem são relacionados a cada ponto da nuvem de pontos LASER, formando um conjunto de voxels. Esse conjunto é então classificado a partir de duas abordagens diferentes (supervisionada e não supervisionada) com o objetivo de detectar edifícios, árvores, terrenos cobertos por vegetação e terrenos descobertos. Zhou e Zhou (2014) apresentam uma metodologia para a extração de edifícios através da elaboração de um grafo de atributos baseado nas características dos edifícios (geometria, estrutura, forma), e construído a partir da combinação das imagens aéreas e dos dados de VLA. Lari et al. (2017) propõem um método para melhorar a interpretabilidade das superfícies planares extraídas de dados de VLA, a partir do uso de informações de textura provenientes de imagens coletadas através de sistemas de mapeamento moderno de baixo custo. Primeiro as superfícies planares são extraídas dos dados de VLA através de um novo processo de segmentação. Em seguida, as partes não oclusas das superfícies planares extraídas são decompostas em segmentos e um processo de renderização é realizado a fim de texturizar esses segmentos usando as imagens disponíveis.

Como pode ser observado nos exemplos apresentados anteriormente, as metodologias de extração de edifícios existentes na literatura exploram diversas técnicas para alcançar o objetivo desejado. Embora existam poucas abordagens baseadas na modelagem *Snake*, essa teoria vem ganhando espaço no campo da extração de feições. Yan et al. (2015) realiza a construção de modelos 3D de edifícios a partir de dados de VLA com base em um novo modelo para o algoritmo *Snake*, denominado “algoritmo *Snake 2D*”, o qual minimiza a função de energia associada à topologia planar através da técnica de redução de grafos. Essa técnica consiste em um conjunto de operações gráficas capazes de reduzir uma topologia planar bidimensional a vértices isolados, enquanto mantém a energia mínima da topologia. No trabalho de Chen et al. (2016b), um modelo melhorado de *Snake* é proposto para refinar os contornos iniciais de edifícios extraídos de dados de VLA. Primeiro, um modelo *Snake* aprimorado é construído com as restrições de ângulo, gradiente e área. Em seguida, os nós do contorno são movidos para encontrar o melhor resultado utilizando o algoritmo guloso (ZIVIANI, 2011). Griffiths e Boehm (2019) apresentam um método para aumentar a



qualidade de grandes conjuntos de dados com baixa precisão, para permitir o treinamento de redes de segmentação. Os pontos centrais de cada edifício são calculados e usados como ponto semente para a aplicação do modelo *Snake*. O modelo é utilizado para determinar as bordas dos edifícios e realizar uma segmentação precisa. A imagem segmentada é então utilizada como dado de treinamento para a aplicação em redes neurais convolucionais.

De uma forma geral, os métodos de extração dos contornos de telhados existentes na literatura, tanto aqueles que usam imagem ou dados de VLA quanto aqueles que usam os dois tipos de dados, podem trabalhar de três diferentes formas. No primeiro caso, têm-se os métodos que extraem os contornos de telhados diretamente na imagem aérea (COTE e SAEEDI, 2013). O segundo caso envolve métodos que extraem contornos tridimensionais de telhados utilizando nuvem de pontos em 3D de VLA (SEO et al., 2014; CHEN et al., 2014). E, no terceiro caso, têm-se os métodos que combinam de alguma forma ambas fontes de dados: imagem e nuvem de pontos LASER (GERKE e XIAO, 2014; ZHOU e ZHOU, 2014; GILANI et al., 2015; OLIVEIRA, 2016a).

Este trabalho inspira-se inicialmente na ideia proposta em Fazan (2014), o qual realiza a extração dos contornos de telhados diretamente no espaço-objeto (3D), integrando numa única função (função de energia *Snake*) uma imagem aérea e dados de VLA. Esse método de extração combina as propriedades radiométricas do telhado, obtidas a partir da imagem aérea, e as propriedades geométricas do telhado, obtidas a partir dos dados de VLA. A presente proposta diferencia-se principalmente por usar: 1) um estéreo par de imagens aéreas, em vez de apenas uma imagem; 2) modelos tridimensionais de telhado previamente reconstruídos de dados de VLA, em vez de dados brutos dessa mesma fonte de dados; e 3) informações contextuais da cena, obtidas a partir das imagens e dos dados de VLA. Além desses aspectos diferenciadores, apresenta-se um método totalmente automático, contrapondo-se à estratégia semi-automática proposta em Fazan (2014). Vale ressaltar que embora não sejam utilizados dados de VLA em sua forma bruta, mas sim como modelos 3D de telhados, o método proposto para o delineamento do contorno do telhado é considerado totalmente automático visto que não há necessidade de interferência humana em nenhuma das etapas do método de extração.

Uma grande vantagem do método proposto é a possibilidade de tratar obstruções perspectivas, que são dependentes do ponto de vista de tomada das imagens e que não podiam ser tratadas pelo método proposto em Fazan (2014). Outro diferenciador é a incorporação de mecanismos para superar problemas comuns quando se trata de extração de contornos de

telhados em ambientes urbanos, sendo eles: obstruções causadas por vegetação que se eleva acima dos telhados; e sombras adjacentes aos telhados, as quais podem ser confundidas com as bordas dos telhados. O método proposto visa aproveitar a sinergia entre dados de VLA e imagens ópticas no contexto de extração de edifícios, isto é: planos de telhados são melhor definidos em dados de VLA e contornos de edifícios são melhor definidos em imagens aéreas de alta resolução (*Ground Sample Distance (GSD)*  $\approx 10$  cm). Isso será originalmente realizado através de uma função de energia baseada na formulação *Snake*, cuja solução (polígonos representando contornos refinados de edifícios) será obtida por otimização via Programação Dinâmica (PD).

## **1.2. Hipótese e Objetivos**

### **1.2.1. Hipótese**

Antes da apresentação da hipótese que fundamenta este trabalho, é necessário salientar que o alvo de estudo são contornos de telhados formados por segmentos de retas conectados em  $90^\circ$ . A partir dessa premissa, apresenta-se a seguinte hipótese:

*“A partir da otimização via PD de uma função Snake com parâmetros adaptativos, é possível, de forma robusta, melhorar os contornos 3D de telhados previamente obtidos de dados de VLA, utilizando imagens aéreas de alta resolução e informações contextuais extraídas das imagens e dados de VLA, atingindo melhor acurácia dos contornos de telhado no espaço-objeto.”*

### **1.2.2. Objetivo Geral**

Esse trabalho tem por objetivo principal desenvolver um método para a extração de contornos de telhados de edifícios no espaço-objeto (3D), a partir do uso integrado de um estéreo par de imagens aéreas de alta resolução ( $GSD \approx 10$  cm), modelos tridimensionais aproximados de telhado obtidos a partir de dados de VLA, e informações contextuais obtidas previamente das imagens e dados de VLA, utilizando a modelagem *Snake* com parâmetros adaptativos e solução em 3D do modelo matemático via otimização por PD.

### **1.2.3. Objetivos Específicos**

- 1) Desenvolver uma função de energia baseada na teoria de contorno ativo (*Snake*) para modelar contornos de telhado de edifícios no espaço-objeto (3D), tendo como base os atributos radiométricos e geométricos de telhados derivados de um estéreo par de imagens aéreas, as relações matemáticas entre os espaços imagem e objeto, e as propriedades geométricas presentes em modelos tridimensionais de telhados;
- 2) Desenvolver um mecanismo para a compensação automática de obstruções perspectiva sobre telhados, causadas por edifícios vizinhos elevados;
- 3) Desenvolver um mecanismo para a compensação automática de obstruções diretas sobre telhados, causadas por partes da vegetação que se eleva acima dos telhados;
- 4) Desenvolver um mecanismo para a compensação automática de sombreamento nas adjacências do edifício, causados por ele próprio e que podem ser confundidos com as bordas de telhados;
- 5) Desenvolver uma solução para eficientemente otimizar por PD a função de energia desenvolvida;
- 6) Avaliar experimentalmente a metodologia proposta através da comparação entre os resultados obtidos pelo método e os dados obtidos por um operador humano.

### **1.3. Justificativa**

A relevância científica e tecnológica deste trabalho é evidenciada pela importância dada ao tema pela *International Society for Photogrammetry and Remote Sensing* (ISPRS), na Comissão II (*Photogrammetry*) que possui como um de seus objetivos realizar de forma integrada os processos de correspondência (*matching*) e delineamento de objetos. Segundo Chen et al. (2016a), essa integração ainda precisa ser explorada com mais profundidade, uma vez que, atualmente, esses processos são desenvolvidos de forma independente uma da outra, e assim, os erros na solução de uma das tarefas se propaga para a próxima. Dessa forma, a solução simultânea, como é proposta neste trabalho, pode ajudar a superar algumas dessas limitações existentes nos métodos atuais. Como também apontam Chen et al. (2016a), a fusão de dados de diferentes sensores, como dados ópticos e não ópticos, que também é o caso desse trabalho, é atualmente um tópico de grande relevância.

O estudo dos processos de extração automática de objetos em áreas urbanas é de grande relevância tecnológica, visto que, atualmente, a extração ainda depende da participação de um operador humano, o que demanda muito tempo e elevados custos no processo de aquisição dos dados. Sendo assim, este trabalho se insere no contexto de desenvolvimento de novas metodologias para a eficiente aquisição de informações espaciais a partir do uso de dados provenientes de diferentes sensores.

O ineditismo deste trabalho está no fato de se desconhecer na literatura relacionada, metodologias de extração de contornos de telhados de edifícios no espaço-objeto (3D) que utilizem um estéreo par de imagens aéreas juntamente com o modelo tridimensional aproximado do telhado. O grande desafio dessa proposta está relacionado principalmente ao Objetivo Específico 1, o qual propõe o desenvolvimento de uma função de energia baseada na modelagem *Snake* para a representação de contornos de telhados no espaço-objeto, utilizando a integração de um estéreo par de imagens aéreas, modelo 3D do telhado e informações contextuais da cena. Além disso, constitui também um importante desafio (Objetivo Específico 5) o desenvolvimento de uma estratégia para eficientemente otimizar esta função de energia usando PD. Outros desafios de difícil solução estão relacionados com a presença de vários tipos de obstrução (Objetivos Específicos 2, 3 e 4), os quais requerem estratégias especiais para compensação de seus efeitos. Como será descrito na proposta do método, esta compensação fará parte do próprio processo de otimização, caracterizando um método com solução simultânea e não hierárquica, sendo outro aspecto original desta proposta.

#### **1.4. Estrutura do trabalho**

Este trabalho está dividido em cinco capítulos. O Capítulo 1 apresentou uma contextualização sobre o assunto tratado nesta tese, bem como a hipótese, objetivos e justificativa do mesmo.

O Capítulo 2 traz a fundamentação teórica necessária para o desenvolvimento da metodologia proposta, onde são apresentados os sistemas de coordenadas do espaço-imagem e espaço-objeto; os conceitos de quina, borda e sombra; o conceito de ortorretificação, áreas de oclusão e mapa de visibilidade; os fundamentos de um sistema de varredura a LASER aerotransportado; os conceitos básicos de morfologia matemática; o modelo de contorno ativo *Snake*; e o algoritmo de PD.

No Capítulo 3 são apresentados os dados e recursos de hardware e software utilizados, bem como o método proposto para a extração dos contornos de telhados, o qual foi dividido em três partes. Primeiramente é apresentado o processo de geração das imagens auxiliares contendo as informações contextuais da cena de interesse. Após isso, é apresentada a formulação da função de energia *Snake* considerando o estéreo par de imagens. Por fim, é descrita a estratégia utilizada para a solução da função de energia via PD. Nesse capítulo é apresentado também a forma de análise dos resultados.

O Capítulo 4 apresenta os parâmetros e limiares empregados na implementação do método e os resultados experimentais obtidos pelo método proposto. Os experimentos foram divididos em três grupos, onde cada um teve o objetivo de demonstrar a eficiência do método em tratar os problemas de obstruções e sombreamento. Também foram realizados testes para analisar a eficiência do algoritmo de PD na solução do método proposto.

Por fim, o Capítulo 5 apresenta as conclusões do trabalho realizado e algumas recomendações para trabalhos futuros.

## **2. FUNDAMENTAÇÃO TEÓRICA**

O presente capítulo apresenta uma revisão teórica sobre conceitos e técnicas fundamentais para o desenvolvimento do estudo proposto. A Seção 2.1 apresenta os sistemas de coordenadas do espaço-imagem e espaço-objeto, bem como o procedimento para a transformação de coordenadas entre os dois sistemas. Na Seção 2.2 é apresentada a definição de borda e os principais conceitos sobre o cálculo do gradiente em pontos de borda. A Seção 2.3 aborda o conceito de quina e o operador de Moravec. A Seção 2.4 trata sobre as sombras presentes em imagens aéreas e sobre os métodos de detecção existentes. Na Seção 2.5 é apresentado o conceito de ortorretificação, áreas de oclusão e mapas de visibilidade. A Seção 2.6 trata sobre os fundamentos de um sistema de VLA, o seu funcionamento e os processos de filtragem e classificação de dados de VLA, apresentando as características básicas de um algoritmo de filtragem e a ideia fundamental do funcionamento de um algoritmo de classificação baseado em curvatura. A Seção 2.7 traz os conceitos básicos da morfologia matemática e dos operadores de erosão, dilatação, abertura e fechamento. A Seção 2.8 apresenta os fundamentos do modelo de contorno ativo *Snake*. E, por fim, a Seção 2.9 trata dos fundamentos do algoritmo de otimização de Programação Dinâmica.

### **2.1. Sistemas de coordenadas**

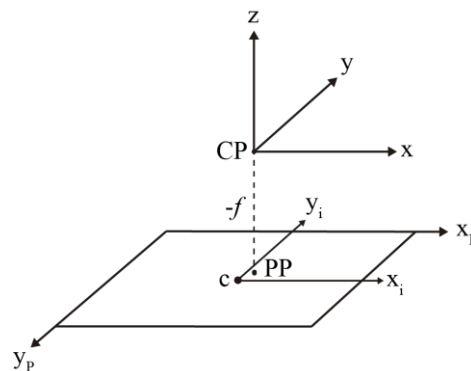
#### **2.1.1. Sistemas de coordenadas do espaço-imagem**

O espaço-imagem é considerado como sendo o espaço que envolve o nodo posterior do sistema de lentes da câmara e o plano do negativo (LUGNANI, 1987). A seguir são apresentados os principais sistemas de coordenadas do espaço imagem: sistema com origem no centro da imagem, sistema fotogramétrico e sistema de imagem.

- 1) Sistema com origem no centro da imagem: é um sistema bidimensional definido da seguinte forma:
  - Origem coincidente com o centro da imagem (intersecção entre as diagonais da imagem), cuja posição é dada por  $(C_x, C_y)$ ;
  - O eixo  $x_c$  tem a direção horizontal mais próxima da linha de voo, sendo as coordenadas positivas nesse sentido;

- O eixo  $y_c$  tem sentido positivo a  $90^\circ$  (sentido anti-horário) do eixo  $x_c$ .
- 2) Sistema Fotogramétrico: é um sistema tridimensional definido como segue (Figura 1):
- Origem no Centro Perspectivo (CP) da câmara;
  - Os eixos  $x$  e  $y$  são paralelos e de mesmo sentido dos eixos  $x_c$  e  $y_c$  do Sistema com origem no centro da imagem;
  - O eixo  $z$  é perpendicular ao plano da imagem, de forma a tornar o sistema dextrogiro. Nesse sistema a coordenada  $z$  de qualquer ponto possui o valor de  $-f$ .

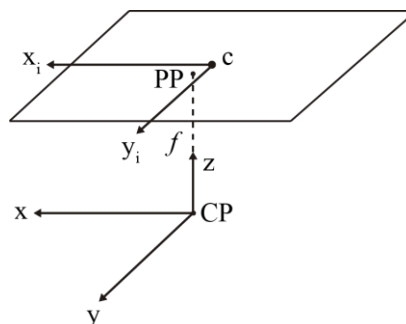
Figura 1 – Sistema fotogramétrico.



Fonte: Lugnani (1987)

Para o caso do negativo, a origem desse sistema coincide com o CP, os eixos  $x$  e  $y$  sofrem uma reflexão em relação ao sistema anterior e o eixo  $z$ , é tal que, torna o sistema dextrogiro (Figura 2). Neste caso, a coordenada  $z$  de qualquer ponto situado no plano do negativo assumirá o valor de  $f$ .

Figura 2 – Sistema fotográfico de coordenadas do negativo.

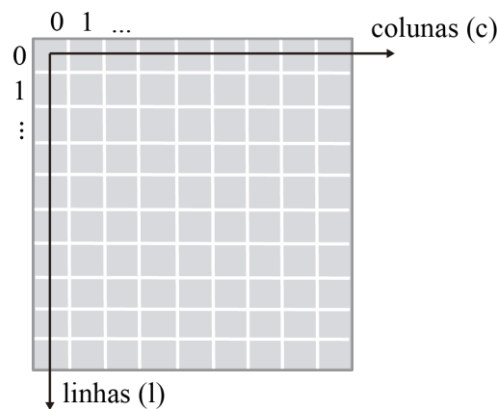


Fonte: Lugnani (1987)

3) Sistema de Imagem: é um sistema bidimensional associado à imagem digital (Figura 3). Uma imagem digital é constituída por um conjunto de elementos de imagem, denominado pixels, espacialmente ordenados em forma de matriz. Cada pixel possui um atributo que indica o nível de cinza ( $g(c,l)$ ), e a sua posição é dada por um par ordenado (coluna, linha) ou  $(c, l)$ . A definição desse sistema é apresentada a seguir:

- Origem no centro do pixel situado no canto superior esquerdo;
- O eixo  $l$  coincide com a primeira coluna da imagem;
- O eixo  $c$  coincide com a primeira linha da imagem.

Figura 3 – Sistema de imagem.



### **2.1.2. Sistemas de coordenadas do espaço-objeto – Sistema Universal Transverso de Mercator (UTM)**

O sistema UTM é obtido a partir da projeção dos pontos do elipsoide de revolução sobre um cilindro secante, que por sua vez é desenvolvido em um plano. Esse sistema faz parte da categoria de projeções conformes, cuja propriedade principal é a preservação da forma de pequenas áreas.

Nessa projeção as coordenadas de um ponto são representadas por  $(E, N)$  e podem ser calculadas a partir das coordenadas no sistema Transverso de Mercator (TM) através das seguintes equações:



$$\begin{bmatrix} N \\ E \end{bmatrix} = \begin{bmatrix} 0,9996 & 0 \\ 0 & 0,9996 \end{bmatrix} \begin{bmatrix} X_{TM} \\ Y_{TM} \end{bmatrix} + \begin{bmatrix} 10.000.000 \\ 500.000 \end{bmatrix} \text{ (Hemisfério sul)} \quad (2.1)$$

ou

$$\begin{bmatrix} N \\ E \end{bmatrix} = \begin{bmatrix} 0,9996 & 0 \\ 0 & 0,9996 \end{bmatrix} \begin{bmatrix} X_{TM} \\ Y_{TM} \end{bmatrix} + \begin{bmatrix} 0 \\ 500.000 \end{bmatrix} \text{ (Hemisfério norte)} \quad (2.2)$$

Como as coordenadas no Sistema UTM são referenciadas a um sistema cartesiano, é necessária uma terceira componente para se ter a determinação unívoca de um ponto sobre a superfície física. Essa terceira componente pode ser a altitude geométrica ( $h$ ), gerando a combinação UTM-h, ou a altitude ortométrica ( $H$ ), gerando a combinação UTM-H.

Essas combinações constituem sistemas caracterizados como híbrido, por serem compostos por superfícies de referência de diferentes características. As coordenadas  $E$  e  $N$  são coordenadas cartesianas sobre o cilindro, que é desenvolvível no plano, e  $h$  e  $H$  são obtidas a partir das distâncias dos pontos da superfície física a uma superfície não plana. Em alguns procedimentos essas combinações provocam erros sistemáticos, principalmente no caso de extensas áreas.

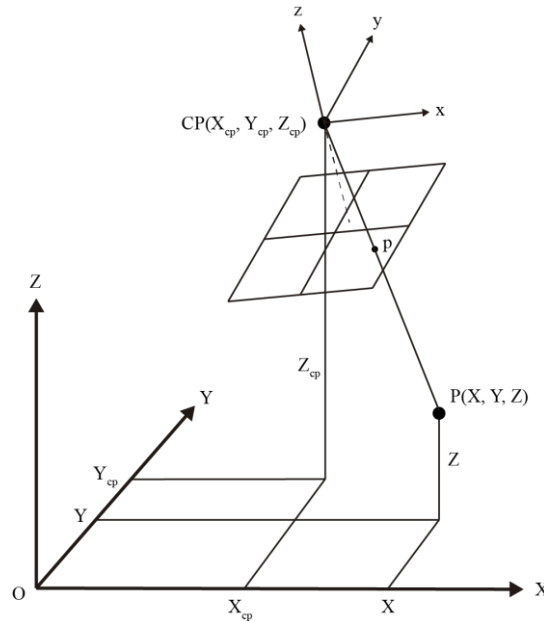
### 2.1.3. Transformação de um ponto no espaço-objeto para o espaço-imagem

O modelo matemático comumente utilizado para relacionar pontos no espaço-objeto com os respectivos pontos no espaço-imagem são as Equações de Colinearidade. Este modelo consiste basicamente em duas equações (Equação 2.3), as quais consideram que um ponto no espaço-objeto  $P(X, Y, Z)$ , o seu correspondente no espaço-imagem  $p(x, y)$  e o centro perspectivo  $CP(X_{CP}, Y_{CP}, Z_{CP})$  são colineares no instante da tomada da imagem (Figura 4). Sendo assim, a transformação das coordenadas do espaço-objeto para o espaço-imagem se inicia transformando o ponto no espaço-objeto para o Sistema Fotogramétrico, a partir da seguinte equação (ANDRADE, 1998):

$$\begin{aligned} x &= -f \frac{r_{11}(X_L - X_{CP}) + r_{12}(Y_L - Y_{CP}) + r_{13}(Z_L - Z_{CP})}{r_{31}(X_L - X_{CP}) + r_{32}(Y_L - Y_{CP}) + r_{33}(Z_L - Z_{CP})} \\ y &= -f \frac{r_{21}(X_L - X_{CP}) + r_{22}(Y_L - Y_{CP}) + r_{23}(Z_L - Z_{CP})}{r_{31}(X_L - X_{CP}) + r_{32}(Y_L - Y_{CP}) + r_{33}(Z_L - Z_{CP})} \end{aligned} \quad (2.3)$$

onde  $(x, y)$  são as coordenadas de um ponto  $p$  no Sistema Fotogramétrico;  $f$  é a distância focal da câmara;  $r_{ij}$ ,  $1 \leq i \leq 3$  e  $1 \leq j \leq 3$ , são os elementos da matriz de rotação;  $(X_L, Y_L, Z_L)$  são as coordenadas do ponto  $P$  no sistema do espaço-objeto;  $(X_{CP}, Y_{CP}, Z_{CP})$  são as coordenadas do CP no sistema do espaço-objeto.

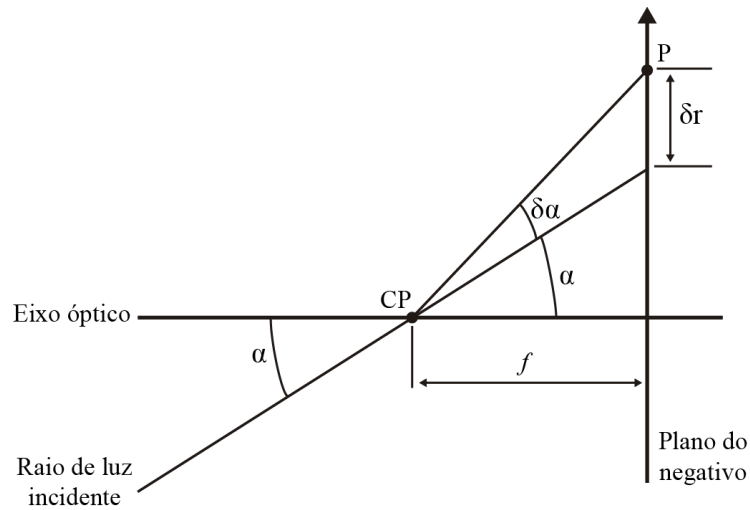
Figura 4 – Condição de colinearidade entre o CP, o ponto imagem  $p$  e o ponto no espaço-objeto  $P$ .



Após isso, o ponto  $p(x, y)$  no Sistema Fotogramétrico é transformado para o referencial da imagem  $(c, l)$ , sendo necessário primeiro introduzir as distorções relacionadas ao sistema de lentes. Isso é necessário visto que a imagem digital original, onde geralmente as operações de análise de imagem são realizadas, é afetada por tais distorções. Em se tratando de imagens capturadas por câmaras digitais, as distorções mais relevantes a se considerar são as distorções das lentes, sendo elas: distorção radial simétrica e distorção descentrada.

A distorção radial simétrica é provocada pela dificuldade em fabricar superfícies de lentes com paraboloides de revolução perfeitos. Isso provoca um desvio na trajetória do raio de luz emergente no ponto nodal posterior, em relação ao raio incidente no ponto nodal anterior. Mais formalmente, a distorção radial simétrica pode ser encarada como sendo a componente radial indesejável da refração sofrida por um raio de luz ao atravessar uma lente ou um sistema de lentes (ANDRADE, 1998). Dessa forma, um raio de luz ao atravessar um sistema de lentes com um ângulo de incidência  $\alpha$  com o eixo óptico, refrata-se com valor  $\alpha + \delta\alpha$ , provocando um deslocamento  $\Delta r$  no ponto imagem (Figura 5).

Figura 5 – Distorção radial simétrica.

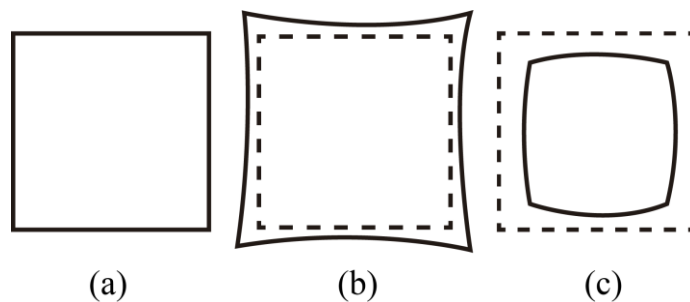


Fonte: adaptado de Andrade (1998).

Essa distorção provoca um deslocamento radial e simétrico com relação ao ponto principal, por isso é denominada de distorção radial simétrica. Vale salientar, que esse efeito deforma geometricamente as imagens. A Figura 6 ilustra as deformações que essa distorção causa nas imagens. A Figura 6(a) apresenta um quadrado sem distorções, a Figura 6(b) ilustra a distorção positiva chamada de distorção em almofada e a Figura 6(c) mostra a distorção negativa, chamada de distorção em barrilete.

Figura 6 – Efeitos da distorção radial simétrica.

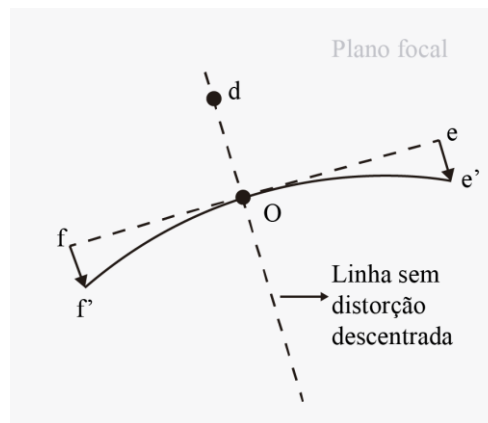
(a) Quadrado sem distorções; (b) Distorção em almofada; e (c) Distorção barrilete.



A distorção descentrada é causada pela impossibilidade do fabricante em alinhar perfeitamente os eixos ópticos das lentes que compõe o sistema de lentes. Ela provoca o deslocamento de um ponto pertencente à imagem, normal à linha radial passante pelo ponto (Figura 7).

Como pode ser visto na Figura 7, a imagem  $d$  está em sua posição correta tangencialmente. Já o ponto  $e$  foi deslocado no sentido horário de sua posição correta em  $e'$ , e o ponto  $f$  foi deslocado no sentido anti-horário de sua posição correta  $f'$ . Como exemplo, se a imagem de uma reta  $e'f'$  passa por  $O$  (ponto principal), a distorção fará com que a reta seja uma linha curva, e isto só não ocorrerá quando a linha coincidir com a linha de distorção nula.

Figura 7 – Distorção descentrada.



O modelo matemático que permite corrigir ambas as distorções é dado por (LUGNANI, 1987):

$$\begin{aligned} x &= x' - (k_1 r^2 + k_2 r^4 + k_3 r^6)x' - P_1(r^2 + 2x'^2) + 2P_2 x' y' \\ y &= y' - (k_1 r^2 + k_2 r^4 + k_3 r^6)y' - P_2(r^2 + 2y'^2) + 2P_1 x' y' \end{aligned} \quad (2.4)$$

onde  $k_1$ ,  $k_2$ ,  $k_3$ ,  $P_1$ ,  $P_2$  são parâmetros conhecidos de calibração;  $r = \sqrt{(x' - x_0)^2 + (y' - y_0)^2}$ ;  $(x', y')$  são as coordenadas do ponto imagem  $p$  afetadas pelas distorções da lente e  $(x, y)$  são as coordenadas do ponto  $p$  livre das distorções da lente, sendo essas coordenadas obtidas pelas Equações de Colinearidade (Equação 2.3).

A Equação 2.4 permite o cálculo das coordenadas não afetadas pelos erros sistemáticos  $(x, y)$  em função das coordenadas afetadas pelos erros sistemáticos  $(x', y')$ . No entanto, como esta equação não admite a inversa analítica, é necessário utilizar alguma solução numérica para encontrar os valores de  $(x', y')$  que satisfaçam a Equação 2.4. Isso pode ser eficientemente resolvido através de um método numérico para encontrar raízes de funções, com por exemplo o Newton-Raphson (SULI e MAYERS, 2003).

O ponto resultante  $p(x', y')$  da introdução das distorções das lentes necessita agora ser transformado para o referencial central da imagem  $(x_c, y_c)$ . O referencial central da imagem é paralelo ao referencial fotogramétrico bidimensional, deferindo apenas na origem. A origem do sistema fotogramétrico bidimensional é o ponto principal (PP) e o sistema central possui origem no centro da imagem ( $C_i$ ). As coordenadas do PP( $x_0, y_0$ ) no referencial central da imagem são determinadas em processos de calibração de câmeras aéreas. Portanto, sendo  $(x_c, y_c)$  as coordenadas do ponto  $p(x', y')$  no referencial central, tem-se:

$$\begin{aligned}x_c &= x' - x_0 \\y_c &= y' - y_0\end{aligned}\tag{2.5}$$

Finalmente, o último passo consiste em transformar o ponto  $p(x_c, y_c)$  para seu correspondente ponto  $p(c, l)$  no referencial digital da imagem, como mostrado a seguir:

$$\begin{aligned}c &= \frac{x_c}{DC} + \frac{DW}{2} \\l &= -\frac{y_c}{DL} + \frac{DH}{2}\end{aligned}\tag{2.6}$$

onde:

- $DW = \text{número\_colunas} - 1$ ;
- $DH = \text{número\_linhas} - 1$ ;
- $DC$  e  $DL$  são as dimensões horizontal e vertical (em milímetros) de um pixel.

## **2.2. Conceito de borda**

### **2.2.1. Definição de borda**

As bordas oferecem importantes informações dos objetos imageados, já que as mesmas correspondem a descontinuidades fotométricas e geométricas. Segundo Gonzalez e Woods (2000), a borda é o limite entre duas regiões com propriedades distintas de nível de cinza.

A Figura 8 apresenta o perfil dos tipos de bordas presentes em uma imagem, sendo elas (JAIN et al., 1995):

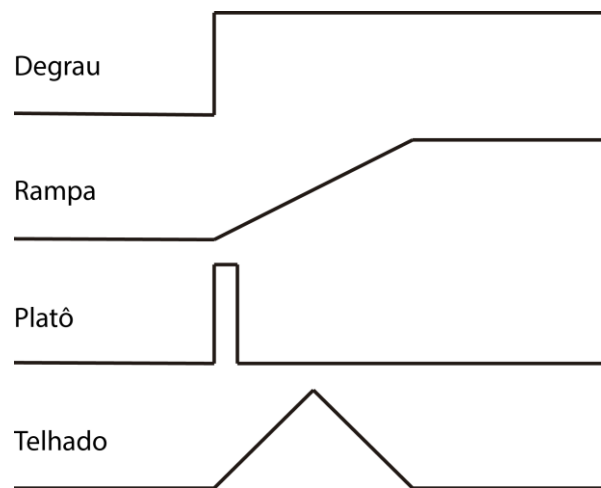
Borda degrau: ocorre quando a intensidade muda abruptamente de um valor em um dos lados da descontinuidade para outro valor no outro lado da descontinuidade.

Borda rampa: é um caso especial da borda degrau, onde a mudança de intensidade não é instantânea, ou seja, acontece de forma gradual em uma distância finita.

Borda platô: decorre de uma mudança abrupta de intensidade, voltando à intensidade anterior após um uma curta distância.

Borda telhado: é um caso especial da borda platô, onde a mudança de intensidade não é simultânea, ocorrendo em uma distância finita.

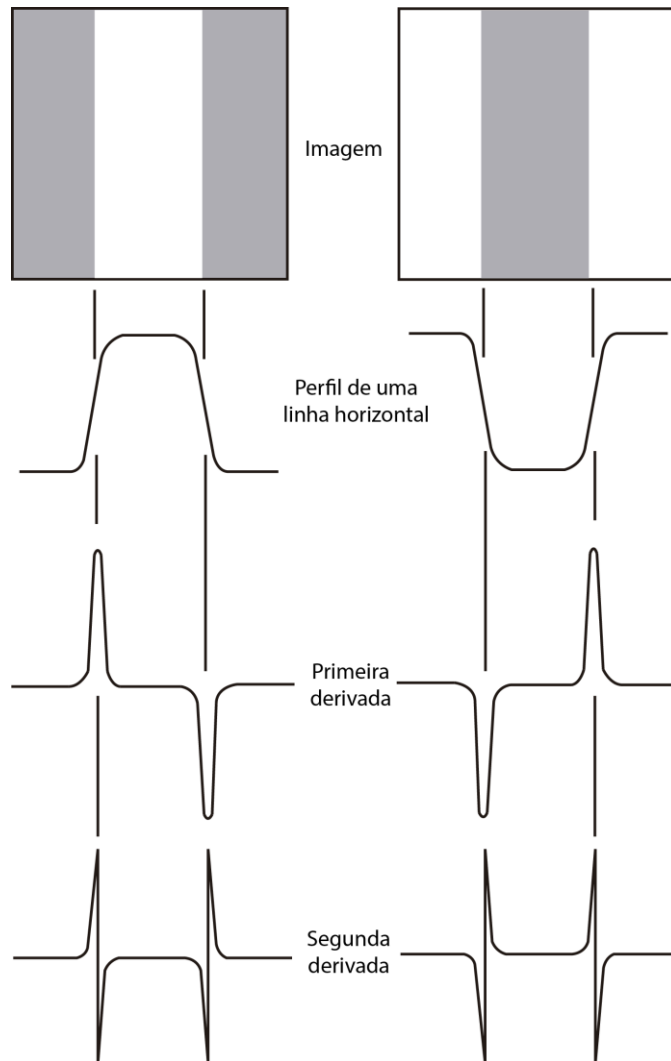
Figura 8 – Perfil dos tipos de bordas existentes em uma imagem.



Fonte: Adaptado de JAIN et al., 1995.

A partir da Figura 9 é possível observar que as bordas estão localizadas nos máximos ou mínimos da primeira derivada, e nos pontos de cruzamento por zero da segunda derivada. Além disso, nota-se que a segunda derivada é positiva na parte da transição associada ao lado escuro da borda e negativa na parte da transição associada ao lado claro da borda. Sendo assim, a magnitude da primeira derivada pode ser utilizada para detectar a presença de uma borda na imagem, enquanto o sinal da segunda derivada poder ser utilizado para indicar se um pixel de borda se encontra no lado escuro ou claro da mesma (GONZALEZ e WOODS, 2000).

Figura 9 – Detecção de bordas por operadores de derivação.



Fonte: GONZALEZ e WOODS, 2000.

### 2.2.2. Gradientes de borda

Segundo Gonzalez e Woods (2000), a maioria das técnicas de detecção de bordas são baseadas em um operador local diferencial. O método mais comum de diferenciação é o do gradiente. Considerando uma função  $f(x,y)$ , por definição, sabe-se que o vetor gradiente dessa função irá apontar na direção onde há uma mudança mais rápida de  $f$  na posição  $(x,y)$ . Sendo assim, em regiões de bordas o vetor gradiente irá apontar na direção transversal a essas bordas (GONZALEZ e WOODS, 2000).

O gradiente de  $f$  nas direções  $(x,y)$  é dado pelo vetor:

$$\nabla f = \begin{bmatrix} G_x \\ G_y \end{bmatrix} = \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix} \quad (2.7)$$

A magnitude desse vetor, denotada por  $\|\nabla f\|$ , é uma quantidade muito importante na detecção de bordas. Ela equivale à maior taxa de aumento de  $f(x,y)$  por unidade de distância na direção de  $\nabla f$ , e é dada por:

$$\|\nabla f\| = \text{mag}(\nabla f) = [G_x^2 + G_y^2]^{1/2} \quad (2.8)$$

Geralmente aproxima-se a magnitude do vetor gradiente com valores absolutos:

$$\|\nabla f\| \approx |G_x| + |G_y| \quad (2.9)$$

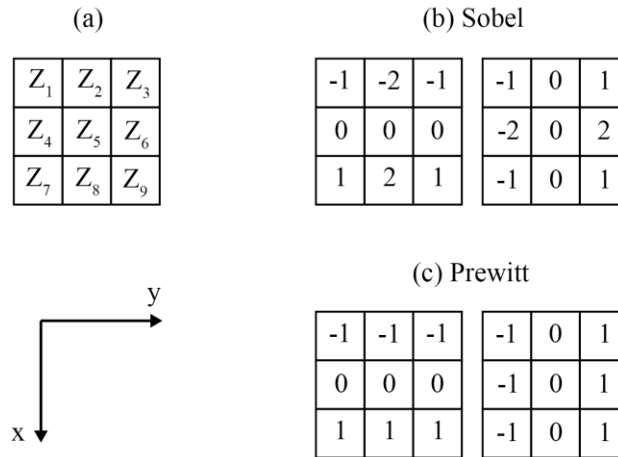
Outra quantidade importante é a direção do vetor gradiente. Seja  $\alpha(x,y)$  o ângulo da direção do vetor  $\nabla f$  na posição  $(x,y)$  e medido em relação ao eixo  $x$ , tem-se que no primeiro quadrante:

$$\alpha(x,y) = \tan^{-1}\left(\frac{G_x}{G_y}\right) \quad (2.10)$$

Considerando a Figura 10(a), onde os  $z$ 's são os valores dos níveis de cinzas, a Equação 2.8, por exemplo, pode ser escrita no ponto  $z_5$  de várias maneiras.



Figura 10 – Máscara de gradiente. (a) Região de tamanho 3x3 em uma imagem;  
 (b) Operadores de Sobel; (c) Operadores de Prewitt.



Fonte: Adaptado de GONZALEZ e WOODS, 2000.

Utilizando os operadores de Sobel (Figura 10(b)), as derivadas baseadas nas máscaras são:

$$G_x = (z_7 + 2z_8 + z_9) - (z_1 + 2z_2 + z_3) \quad (2.11)$$

$$G_y = (z_3 + 2z_6 + z_9) - (z_1 + 2z_4 + z_7)$$

Então a magnitude do vetor gradiente (Equação 2.9) resulta em:

$$\|\nabla f\| \approx |(z_7 + 2z_8 + z_9) - (z_1 + 2z_2 + z_3)| + |(z_3 + 2z_6 + z_9) - (z_1 + 2z_4 + z_7)| \quad (2.12)$$

Para os operadores de Prewitt (Figura 10(c)) as derivadas baseadas nas máscaras são:

$$G_x = (z_7 + z_8 + z_9) - (z_1 + z_2 + z_3) \quad (2.13)$$

$$G_y = (z_3 + z_6 + z_9) - (z_1 + z_4 + z_7)$$

A magnitude do vetor gradiente (Equação 2.9) para esse caso resulta em:

$$\|\nabla f\| \approx |(z_7 + z_8 + z_9) - (z_1 + z_2 + z_3)| + |(z_3 + z_6 + z_9) - (z_1 + z_4 + z_7)| \quad (2.14)$$

O cálculo da magnitude do vetor gradiente é realizado em cada pixel da imagem para a obtenção da imagem de magnitude do gradiente.

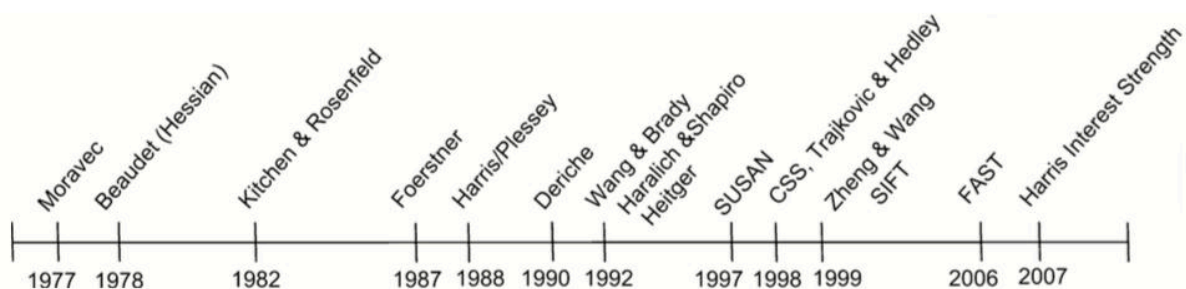
### 2.3. Conceito de quina

Segundo Mehrotra et al. (1990), uma quina é definida como um ponto de intersecção entre duas ou mais bordas. Uma importante concepção na área de visão computacional é a definição de “ponto de interesse”, o qual é considerado como sendo um ponto que tem posição bem definida e pode ser detectado de forma robusta. Isso significa que esse ponto pode ser uma quina, mas também pode ser, por exemplo, um ponto isolado com intensidade local máxima ou mínima, final de linhas, ou um ponto em uma curva onde a curvatura é localmente máxima. Dessa forma, quando se trata da detecção de quinas, um detector de quinas deve possuir as seguintes características (MEHROTRA et al., 1990):

- Todos os pontos de quina devem ser detectados;
- O detector de quina deve ser insensível aos ruídos;
- Os pontos de falsas quinas não devem ser detectados;
- O detector de quina deve fornecer o ângulo e a orientação da quina.

A Figura 11 apresenta a linha do tempo com os anos de criação dos principais detectores de pontos de interesse, também conhecidos como operadores de interesse.

Figura 11 – Linha do tempo da criação dos principais operadores de interesse.



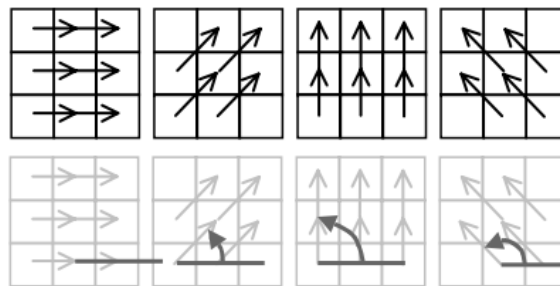
Fonte: Jazayeri e Fraser (2010).

O primeiro operador de interesse foi desenvolvido por Hans P. Moravec em 1977 (MORAVEC, 1977), que introduziu também o conceito de ponto de interesse. Desde então o operado de Moravec vem sendo usado em várias aplicações na área de fotogrametria, incluindo a reconstrução 3D de objetos. No entanto, esse operador possui algumas limitações devido à sua baixa taxa de repetibilidade e sensibilidade a ruídos (JAZAYERI e FRASER,

2010). A baixa repetibilidade é consequência da natureza anisotrópica do algoritmo e qualquer rotação na imagem gera resultados diferentes na detecção de pontos de interesse.

O operador de Moravec se baseia na medida de variância direcional, considerando quatro direções. A resposta desse operador considera os valores dos quadrados das diferenças dos tons de cinza nas direções  $0^\circ$ ,  $45^\circ$ ,  $90^\circ$  e  $135^\circ$ , como mostra a Figura 12 com janelas de dimensão  $3 \times 3$  e conectividade 8.

Figura 12 – Janelas  $3 \times 3$  e os vetores nas direções  $0^\circ$ ,  $45^\circ$ ,  $90^\circ$  e  $135^\circ$ .



Fonte: Galo e Tozzi (2002).

Para o caso de janelas  $n \times n$ , com  $n$  ímpar e origem do sistema de referência no centro da janela, as diferenças quadráticas são calculadas da seguinte maneira (GALO e TOZZI, 2002):

$$V_{0^\circ} = \frac{1}{n(n-1)} \sum_{l=-N}^N \sum_{c=-N+1}^N (g(l, c) - g(l, c-1))^2 \quad (2.15)$$

$$V_{45^\circ} = \frac{1}{(n-1)^2} \sum_{l=-N}^{N-1} \sum_{c=-N+1}^N (g(l, c) - g(l+1, c-1))^2 \quad (2.16)$$

$$V_{90^\circ} = \frac{1}{n(n-1)} \sum_{l=-N}^{N-1} \sum_{c=-N}^N (g(l, c) - g(l+1, c))^2 \quad (2.17)$$

$$V_{135^\circ} = \frac{1}{(n-1)^2} \sum_{l=-N}^{N-1} \sum_{c=-N}^{N-1} (g(l, c) - g(l+1, c+1))^2 \quad (2.18)$$

onde  $N = (n-1)/2$ . O uso de janelas de dimensão ímpar não é restrição do operador de Moravec, mas se deve ao fato de se ter escolhido como origem o pixel central da janela, sendo a janela simétrica em relação a esse ponto.

A partir dos resultados obtidos com as Equações (2.15) à (2.18), a resposta do operador de Moravec para o pixel  $(l, c)$  é calculada por:

$$OI(l, c) = \min(V_{0^\circ}, V_{45^\circ}, V_{90^\circ}, V_{135^\circ}) \quad (2.19)$$

A Equação (2.19) é usada em cada ponto da imagem, e a partir dos resultados obtidos, os máximos locais de  $OI(l, c)$  podem ser encontrados. Sendo assim, os pixels nos quais os valores de  $OI(l, c)$  forem superiores a um limiar preestabelecido, corresponderão aos pontos de quina.

#### **2.4. Conceito de sombra**

Em imagens aéreas de alta resolução é comum a presença de sombras quando a cena envolvida compreende regiões urbanas complexas, uma vez que o principal motivo para a ocorrência de sombras é o bloqueio direto da luz solar por meio de objetos altos como árvores e edifícios. Consequentemente, as superfícies afetadas são fracamente iluminadas e aparecem escuras nas imagens. Os efeitos negativos das sombras afetam a qualidade visual das imagens, causando perturbações severas na análise dessas imagens. Isso dificulta tarefas como a correspondência de imagens para a extração de elevações, detecção de alterações cadastrais, extração automática de edifícios e da malha viária (LI et al., 2004; MADHAVAN et al., 2004; MASSALABI et al., 2004).

As principais características consideradas em análise de imagem para detectar regiões de sombras ou compensar seus efeitos negativos são (MASSALABI et al., 2004):

- Sombras são regiões formadas por pixels com baixo valor de intensidade;
- A forma de uma sombra é função da forma do objeto que a produz;
- Uma ou mais extremidades de uma região sombreada são orientadas na direção do azimute solar (no instante de tomada da imagem);
- A área de uma sombra depende da altura solar (no instante de tomada da imagem) e da altura do objeto que a produz;
- A sombra não modifica a saturação da cor de um objeto;
- Alguns elementos da textura de uma superfície são invariantes à sombra, ou seja, a textura da superfície praticamente não é afetada pelo sombreamento.

Em geral, os métodos de detecção regiões de sombras podem ser classificados em dois grupos. No primeiro as regiões de sombras geralmente são obtidas utilizando-se dados de imagens combinados com outras fontes de informação. Em geral são utilizadas imagens combinadas com informações solares e de altura dos objetos. A estratégia adotada é semelhante em praticamente todos os métodos encontrados na literatura: as regiões de

sombras são determinadas através de uma simulação (predição) empregando dados de elevação, geralmente derivados de um Modelo Digital de Superfície (MDS), e dados de posição solar (azimute e altura), estimados para o instante do imageamento. Fazan e Dal Poz (2008) apresentam um método para a predição de regiões afetadas por sombras de edifícios em uma imagem aérea. Os contornos de telhado de edifícios obtidos a partir de um MDS e as informações sobre a posição solar no instante de aquisição da imagem são utilizados como base para a obtenção das regiões de sombra, que posteriormente são registradas na imagem aérea.

O segundo grupo emprega dados de imagem em conjunto com técnicas convencionais de análise de imagem (limiarização, segmentação, etc.), explorando principalmente as propriedades radiométricas das sombras. No trabalho de Azevedo et al. (2015) a detecção das regiões de sombra em imagens é realizada através de operações de morfologia matemática. O primeiro passo consiste em realçar os padrões escuros da imagem através da transformação *top-hat* por fechamento de área. A Seção 2.7.2.1 apresenta mais detalhes sobre essa transformação. Como as sombras em imagens aéreas não obedecem um padrão específico, a operação de fechamento algébrico por área é utilizado para que não haja a necessidade de estabelecer uma forma particular para o elemento estruturante.

O próximo passo consiste em realizar uma limiarização na imagem resultante da transformação *top-hat*, a fim de detectar as sombras realçadas. O histograma da imagem resultante da transformação *top-hat* possui um comportamento bimodal, ou seja, os alvos de interesse se destacam dos demais alvos da imagem. Dessa forma, o método de Otsu é utilizado para realizar a limiarização, pois ele consegue determinar automaticamente o limiar mais adequado de acordo com uma medida de separação de classes. O resultado é uma imagem binária onde os pixels de sombra possuem o valor 255 e os demais pixels possuem o valor 0. Por fim, a operação de abertura morfológica por área é aplicada na imagem binária para eliminar regiões de sombra pequenas, as quais podem ser consideradas ruídos.

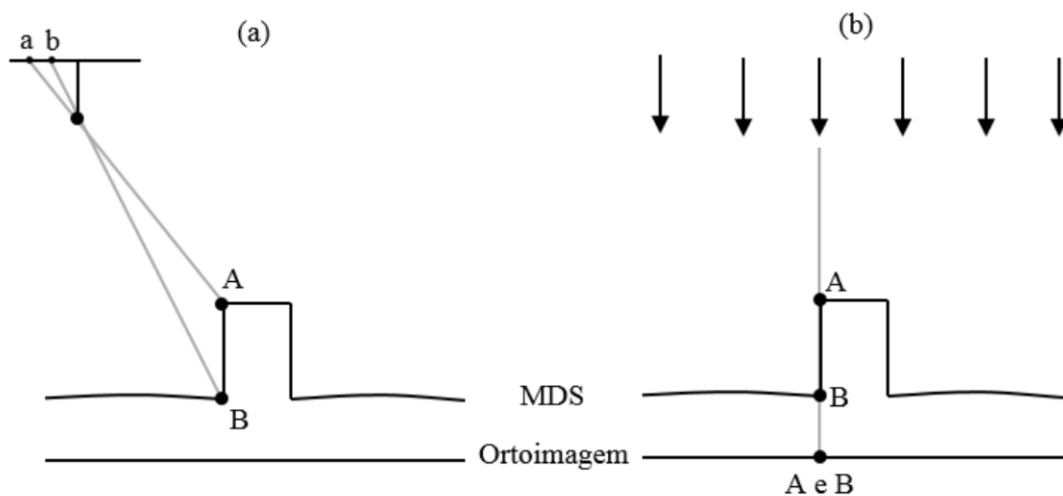
## **2.5. Áreas de oclusão e mapa de visibilidade**

As imagens aéreas apresentam a projeção das feições do mundo real em um plano por meio da projeção via perspectiva central (Figura 13(a)). Nas projeções de perspectiva central os objetos com diferentes altitudes, mas com a mesma posição planimétrica, são projetados em diferentes posições na imagem, causando a não uniformidade da escala da

imagem. A Figura 13(a) ilustra um exemplo da projeção perspectiva central, onde os pontos A e B estão na mesma posição planimétrica, no entanto, são projetados em diferentes posições na imagem devido a suas diferenças de altitude (OLIVEIRA, 2016b).

O deslocamento devido ao relevo presente na projeção de perspectiva central não ocorre na projeção ortogonal (Figura 13(b)), onde a escala é mantida e as feições são representadas em suas posições planimétricas corretas na imagem (MIKHAIL et al., 2001).

Figura 13 – (a) Visada em perspectiva central. (b) Visada ortogonal.



Fonte: Oliveira (2016b)

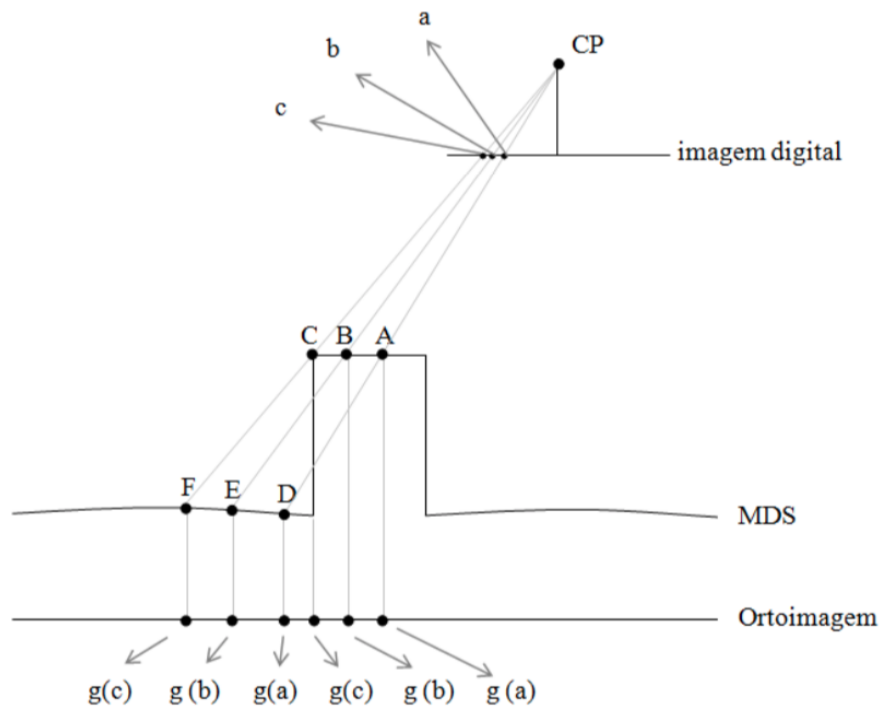
Segundo Mikhail et al. (2001), a partir do processo denominado ortorretificação é possível transformar uma imagem, que está em projeção de perspectiva central, para a projeção ortogonal. Como resultado tem-se uma imagem onde todos os elementos são representados ortogonalmente em um plano de referência, assim como acontece em um mapa.

Essa transformação é realizada por meio de um Modelo Digital de Superfície (MDS) e dos parâmetros de orientação interior e exterior do sensor utilizado na aquisição das imagens. O MDS é uma representação digital das informações de elevação da superfície, o qual inclui tanto o terreno quanto os elementos acima dele, como por exemplo, vegetação e feições antrópicas (WOLF et al., 2014). A utilização do MDS na ortorretificação tem como objetivo eliminar o deslocamento causado pela variação de altura dos objetos acima do terreno, permitindo assim a construção de uma nova imagem de forma ortogonal (NIELSEN, 2004).

No entanto, como é apresentado por Nielsen (2004), Habib et al. (2007) e Oliveira (2013), o uso do MDS no processo de ortorretificação causa o efeito denominado duplo

mapeamento. A Figura 14 ilustra um caso de duplo mapeamento. Nesse exemplo, nota-se que cada um dos pixels *a*, *b* e *c* relaciona-se com mais de um *groundel*. O pixel *a* possui relação com os *groundels* A e D, o pixel *b* com B e E, o pixel *c* com C e F. Entretanto, como D, E e F pertencem ao terreno e não são visíveis ao CP, durante o processo de ortorretificação eles são preenchidos com os mesmos valores de tons de cinza de A, B e C, respectivamente. Isso gera uma ambiguidade, causando o efeito denominado duplo mapeamento.

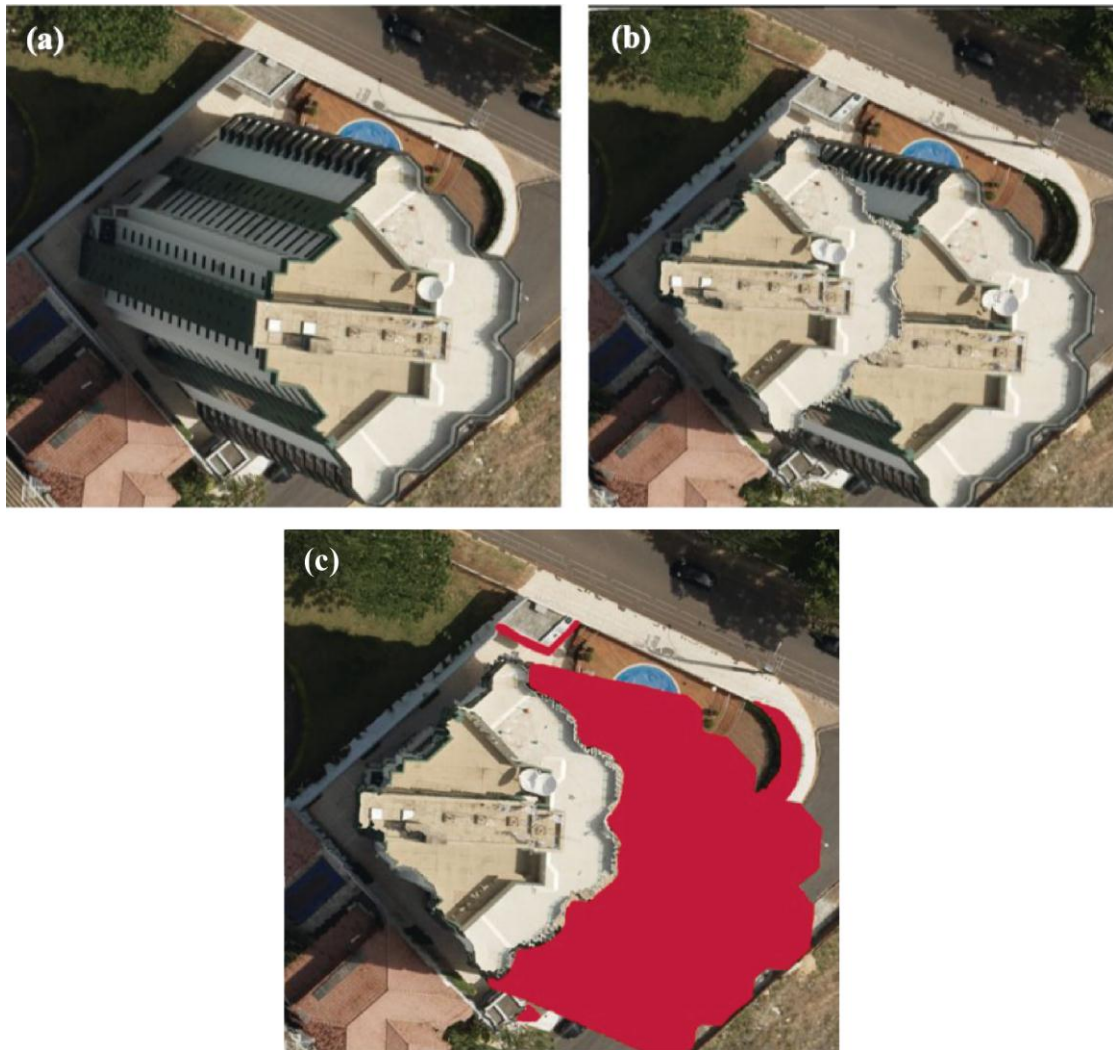
Figura 14 – Ilustração do efeito de duplo mapeamento.



Fonte: Oliveira (2016b)

A Figura 15(a) apresenta a imagem de um edifício elevado tomado em perspectiva central, onde ele parece estar “deitado” sobre a superfície do terreno. Após a ortorretificação o topo do edifício é projetado em sua posição correta, mas é criado o efeito de duplo mapeamento (Figura 15(b)). A Figura 15(c) apresenta em vermelho a região afetada pelo duplo mapeamento, essa região é denominada área de oclusão. Ao se aplicar um processo de detecção de oclusão, os tons de cinza da área duplicada podem ser convertidos para uma cor sólida, como mostra a Figura 15(c), gerando os mapas de visibilidade.

Figura 15 – (a) Imagem original em projeção perspectiva. (b) Efeito de duplo mapeamento. (c) Área de oclusão (região em vermelho).



Fonte: Oliveira (2016b)

Segundo Oliveira (2013), os métodos de detecção de oclusão mais utilizados são: método *Z-buffer*, foi o primeiro a ser utilizado para realizar a detecção de oclusão na geração de ortoimagens verdadeiras e é muito utilizado em computação gráfica (AMHAR et al., 1998); método baseado em ângulos proposto por Habib et al. (2007), o qual foi criado para evitar algumas imprecisões resultantes do método *Z-buffer*; método proposto por Volotão (2001), que utiliza a projeção das bordas de edifícios para identificar as células de oclusão. Oliveira (2016b) apresenta também um novo método de detecção das áreas de oclusão, denominado *Surface-Gradient-Based Method* (SGBM), o qual se baseia na análise de gradientes de altura negativos sobre um MDS.



## **2.6. Sistema de varredura a LASER aerotransportado (SVLA)**

Devido à grande necessidade de aquisição rápida e eficaz de dados de elevação, o uso da tecnologia de VLA tem sido foco de pesquisas desde a década de 1990, uma vez que esses sistemas permitem coletar, em um curto intervalo de tempo, uma alta densidade de pontos tridimensionais com precisão e acurácia equivalentes às técnicas tradicionais de levantamento *in situ* e Fotogrametria. Isso atribui aos dados de VLA uma variedade de aplicações como, reconstrução de superfície, extração de feições, modelagem urbana, mapeamento de zonas costeiras, entre outros.

O funcionamento do sistema de VLA é baseado na utilização de pulsos LASER que são disparados em direção à superfície terrestre com o auxílio de um sistema de varredura. Esses pulsos interagem com a superfície e parte do sinal é refletido de volta em direção ao sensor. O sensor mede a intensidade do sinal retornado e usa o tempo decorrido entre a emissão e a recepção do pulso para calcular a distância entre o sensor e a superfície amostrada. Essas medidas de distâncias, em conjunto com as informações de ângulo de varredura, das posições e atitudes da plataforma e medidas de calibração entre todos os componentes do sistema de VLA, permitem a determinação das coordenadas tridimensionais de diversos pontos no terreno, formando um conjunto de pontos denominado nuvem de pontos LASER.

Os sistemas de VLA usados para mapeamento topográfico operam na região do infravermelho, que compreende o intervalo de comprimento de onda entre 1040 nm e 1060 nm (BALTSAVIAS, 1999).

A qualidade e a acurácia dos dados podem ser afetadas por alguns fatores, tais como a superfície do material, altura de voo, integração GNSS (*Global Navigation Satellite System*)/INS (*Inertial Navigation System*), ângulo de observação, tipo de sensor utilizado, entre outros.

Cada pulso LASER cobre uma área aproximadamente circular, determinada pelo IFOV (*Instantaneous Field of View* – campo de visada instantânea), que é a área da superfície irradiada pelo sensor ativo. Dentro desta área circular (*footprint*), o pulso pode gerar um retorno ou múltiplos retornos, caso o pulso LASER encontre um ou mais objetos (por exemplo, folhas e galhos de árvores ou extremidades de outras feições elevadas) antes de atingir a superfície do terreno (Figura 16). Em geral, o primeiro retorno apresenta a reflexão

de objetos mais elevados e o último retorno apresenta a reflexão da superfície do terreno (JENSEN, 2009).

Figura 16 – Situações em que ocorrem múltiplos retornos.



### 2.6.1. Posição e orientação do feixe LASER

A determinação das coordenadas dos pontos amostrados pela VLA é realizada combinando as medidas oriundas de cada componente que integra o sistema de VLA e os parâmetros que relacionam esses componentes. A unidade LASER é reponsável por realizar a medida da distância entre o emissor do pulso e um ponto na superfície atingida. As coordenadas desse ponto são matematicamente associadas a um sistema de referência global relacionando o ângulo de varredura do sistema de VLA, a posição (receptores GNSS) e a atitude (INS) da plataforma no instante da emissão do pulso LASER. Esse procedimento é realizado para todos os pontos amostrados durante o levantamento, obtendo-se assim um conjunto de dados tridimensionais da superfície.

A Figura 17 apresenta os componentes de um sistema de VLA (BALTSAVIAS, 1999; PETRIE E TOTH, 2008):

Unidade LASER: contém o conjunto óptico de emissão e recepção do LASER, detector de sinal, amplificador (necessário para emissão da luz), sistema de contagem de

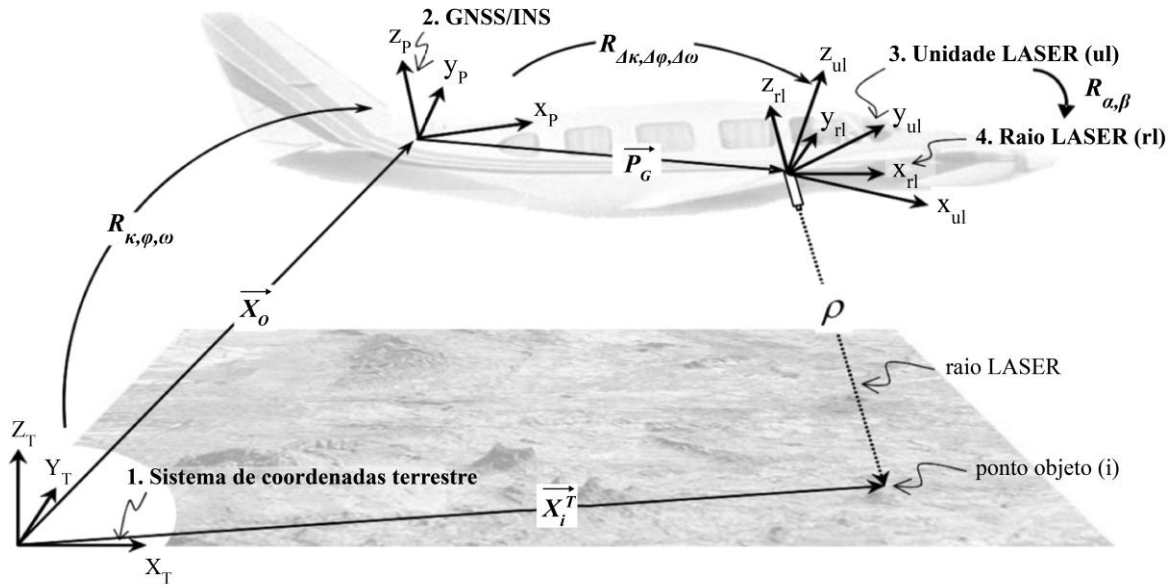
tempo e outros componentes eletrônicos. Após ser estimulado, o pulso LASER é direcionado ao espelho de varredura, e o conjunto óptico orienta o pulso à direção desejada. Ao chegar na superfície, o pulso interage com os objetos, retorna ao sistema de VLA e é coletado por meio de um receptor. O sinal analógico captado é convertido para digital e filtrado para a eliminação de possíveis ruídos. Por fim, esse sinal é utilizado para o cálculo da distância entre o emissor do pulso LASER e a superfície amostrada.

INS: é composto por computadores, instrumentos eletrônicos de apoio e a IMU (*Inertial Measurement Unit* – Unidade de Medição Inercial). A IMU é responsável pela mensuração das acelerações lineares e velocidades angulares da plataforma, o que possibilita a determinação das posições e da atitude da plataforma ao longo da trajetória do levantamento. Essas observações não são coletadas ao mesmo tempo em que são determinadas as posições espaciais estimadas pelo sistema GNSS, criando a necessidade de um sincronismo entre os dois sistemas.

Receptor GNSS: responsável pela coleta das observações necessárias para determinação das posições espaciais da plataforma em que se encontra o sistema de VLA, em intervalos de tempo pré-determinados. Na Figura 17 é considerado que o centro de fase da antena do receptor GNSS coincide com a origem do sistema de coordenadas do INS. Essa simplificação pode ser considerada desde que o vetor ligando a origem do sistema de coordenadas do INS e do receptor GNSS seja conhecido ou determinado a priori com alta precisão.

Estações GNSS de referência: são utilizadas para realizar um pós-processamento e obter posições acuradas para a trajetória da plataforma.

Figura 17 – Componentes de um sistema de VLA e suas relações geométricas.



Fonte: Adaptado de Habib et al. (2008)

Segundo Habib et al. (2008), a posição de um ponto  $i$  no terreno ( $\vec{X}_i^T$ ) pode ser determinada por meio da soma de três vetores ( $\vec{X}_O, \vec{P}_G, \vec{\rho}$ ) e aplicando-se as rotações necessárias ( $R_{\kappa, \varphi, \omega}, R_{\Delta\kappa, \Delta\varphi, \Delta\omega}, R_{\alpha, \beta}$ ):

$$\vec{X}_i^T = \vec{X}_O(t) + R_{\kappa, \varphi, \omega}(t) \cdot \vec{P}_G + R_{\kappa, \varphi, \omega}(t) \cdot R_{\Delta\kappa, \Delta\varphi, \Delta\omega} \cdot R_{\alpha, \beta}(t) \begin{bmatrix} 0 \\ 0 \\ -\rho_i \end{bmatrix} \quad (2.20)$$

onde:

$\vec{X}_i^T$  – coordenadas de um ponto genérico  $i$  no sistema de coordenadas terrestre;

$\vec{X}_O$  – coordenadas da origem do sistema de coordenadas do GNSS/INS, no sistema de coordenadas terrestre;

$\vec{P}_G$  – vetor translação entre a origem do sistema de coordenadas GNSS/INS e a origem do sistema de coordenadas da unidade LASER. Esse vetor translação também é conhecido como *lever arm offset*;

$R_{\kappa, \varphi, \omega}$  – matriz de rotação entre o sistema de coordenadas GNSS/INS e o sistema de coordenadas terrestre;

$R_{\Delta\kappa, \Delta\varphi, \Delta\omega}$  – matriz de rotação entre o sistema de coordenadas GNSS/INS e o sistema de coordenadas da unidade LASER. Os ângulos de rotação que compõem essa matriz são conhecidos como ângulos de desalinhamento (*boresight angles*);

$R_{\alpha, \beta}$  – matriz de rotação entre o sistema de coordenadas da unidade LASER e o sistema de coordenadas do raio LASER. Os ângulos que compõem essa matriz são conhecidos como ângulos de varredura;

$\rho_i$  – medida de distância entre a origem do sistema de coordenadas da unidade LASER e o ponto objeto  $i$  no terreno;

$t$  – instante de recepção do pulso LASER.

A posição ( $\vec{X}_o$ ) e a orientação ( $R_{\kappa, \varphi, \omega}$ ) da plataforma durante o levantamento são determinadas pela integração do GNSS e INS. Em virtude do problema de deriva, inerente ao INS, as posições e velocidades obtidas pelo sistema GNSS podem ser usadas como medidas externas para atualizar a informação gerada pelo INS, melhorando sua estabilidade e precisão ao longo do tempo. Em contrapartida, quando o sinal de recepção do GNSS é interrompido e/ou a geometria dos satélites não está dentro dos limites aceitáveis, o sistema de navegação inercial pode proporcionar as informações necessárias para a navegação (EL-SHEIMY e NIU, 2007).

Durante o deslocamento as coordenadas obtidas pelo sistema GNSS e as observações do sistema inercial não necessariamente coincidem com os instantes de emissão e recepção dos pulsos LASER. Por essa razão é fundamental que o sincronismo entre estes sistemas (GNSS, INS e emissor do pulso LASER) seja estabelecido, para que todos os pontos amostrados possuam coordenadas e valores de atitudes relativos à sua real posição no momento da coleta. Outro fator importante é que os elementos de translação ( $\vec{P}_G$ ) e desalinhamento angular ( $\Delta\kappa, \Delta\varphi, \Delta\omega$ ) devem estar referenciados entre si, ou seja, eles devem ser determinados a priori por meio de um processo de calibração do sistema (HABIB et al., 2008).

### **2.6.2. Filtragem e classificação dos dados de VLA**

Ao fim da aquisição, o sistema de VLA fornece um conjunto de pontos (nuvem de pontos) distribuídos irregularmente ao longo da linha de voo, e com coordenadas

tridimensionais em relação ao sistema de referência WGS 84 (*World Geodetic System 1984*) determinadas através do processo apresentado anteriormente (JENSEN, 2009).

Dependendo da finalidade do trabalho, a nuvem de pontos originada pela varredura a LASER pode passar por um processo de tratamento a partir da filtragem e classificação dos pontos. Esse processo auxilia na diminuição de dados irrelevantes, como por exemplo, a eliminação da vegetação em estudos focados na extração de edifícios e topografia.

Esse pré-processamento dos dados LASER consiste em separar o terreno de objetos altos através do procedimento de filtragem e, em seguida, analisar individualmente as regiões contendo os objetos altos a fim de classificá-las como vegetação ou edifício, segundo um algoritmo específico de classificação, podendo então, eliminar as regiões indesejadas (vegetação ou edifício). Vale ressaltar que a etapa de filtragem, para a separação do terreno e dos objetos altos, também é necessária e precede o processo de geração do MDT.

A seguir são apresentados alguns conceitos sobre cada um desses processos.

#### *2.6.2.1. Filtragem*

Segundo Sithole e Vosselman (2004), a filtragem de dados LASER é fundamentada na combinação de diferentes elementos. Alguns deles são apresentados a seguir.

- 1) Estrutura dos dados: a nuvem de pontos produzida pelo sistema de VLA é composta por pontos tridimensionais irregularmente espaçados. Alguns algoritmos de filtragem (AXELSSON, 1999; PFEIFER *et al.*, 1998; SITHOLE, 2001; SOHN e DOWMAN, 2002; ROGGERO, 2001) trabalham diretamente com a nuvem de pontos LASER original. Entretanto, para aproveitar as ferramentas da área de processamento de imagens, alguns algoritmos de filtragem (BROVELLI *et al.*, 2002; ELMQVIST, 2001; WACK e WIMMER, 2002) reamostram a nuvem de pontos original para uma grade regular de pontos, antes de realizar a filtragem.
- 2) Definição da vizinhança a ser testada: no que diz respeito à vizinhança utilizada pelo algoritmo de filtragem para classificar os pontos em terreno ou objeto, três possíveis formas de classificação podem ser listadas (SITHOLE e VOSSELMAN, 2004):
  - Ponto contra ponto: nesses algoritmos (SITHOLE, 2001; ROGGERO, 2001) dois pontos são comparados ao mesmo tempo. Uma função discriminante é resolvida com base na posição desses dois pontos e, se a saída da função discriminante estiver acima de certo limiar, então um dos pontos é assumido

como pertencente ao objeto. Em outras palavras, somente um ponto é classificado de cada vez.

- Ponto contra pontos: nesses algoritmos (AXELSSON, 1999; SOHN e DOWMAN, 2002) os pontos vizinhos a um ponto de interesse são usados para resolver uma função discriminante. Com base na saída da função discriminante o ponto de interesse pode ser classificado. Apenas um ponto é classificado a cada vez.
  - Pontos contra pontos: nesses algoritmos (ELMQVIST, 2001; PFEIFER *et al.*, 1999, BROVELLI *et al.*, 2002; WACK e WIMMER, 2002) vários pontos são usados para resolver uma função discriminante, que é posteriormente utilizada para classificar os vários pontos. Mais de um ponto é classificado nesse método.
- 3) Medida de descontinuidade: as descontinuidades dos objetos (por exemplo, edifícios) em relação ao terreno são utilizadas pelos algoritmos de filtragem para separar os pontos do terreno e de objetos elevados. Alguns exemplos de medidas de descontinuidade comumente utilizadas são: diferença de altura, declividade, menor distância em relação às faces do TIN (*Triangulate Irregular Network*), e menor distância de pontos a uma superfície.
- 4) Princípios de filtragem: geralmente os algoritmos de filtragem baseiam-se em algum modelo geométrico do terreno em uma vizinhança local. Quatro princípios de filtragem são apresentados por Sithole e Vosselman (2004):
- Baseado na declividade: nesses algoritmos (SITHOLE, 2001; ROGGERO, 2001) a declividade ou a diferença de altura entre dois pontos é medida. Se o desnível exceder certo limiar, então o ponto mais alto é assumido como pertencente a um objeto elevado. Neste caso é assumida a hipótese básica de que o terreno varia suavemente.
  - Baseado no bloco mínimo: como apresentado por Wack e Wimmer (2002), a função discriminante neste princípio visa delimitar uma região em 3D (se assemelharia a um paralelepípedo horizontal ou vertical), tendo por referência um segmento local de plano horizontal a partir do qual, acima ou abaixo, os pontos de terreno devem ser encontrados.
  - Baseado numa superfície: nesse caso é utilizada uma superfície que modela globalmente a superfície do terreno. Por exemplo, tendo em vista que o terreno é suave, Elmqvist (2001) utiliza contorno ativo para reter apenas pontos do

terreno. Já Axelsson (1999) refina progressivamente uma representação poliédrica grosseira do relevo.

- Baseado em segmentação/agrupamento: estes métodos (BROVELLI *et al.*, 2002) levam em conta que pontos que se agrupam e que estão acima de um outro agrupamento adjacente, pertencem a um objeto.
- 5) Mecanismo da filtragem: o mecanismo do processo de filtragem pode ser iterativo ou não iterativo. Os algoritmos não iterativos (SITHOLE, 2001; ROGGERO, 2001) realizam a filtragem numa única passagem, enquanto os iterativos (ELMQVIST, 2001; AXELSSON, 1999; BROVELLI *et al.*, 2002; PFEIFER *et al.*, 1998; SOHN e DOWMAN, 2002; WACK e WIMMER, 2002) classificam pontos através de múltiplas passagens. Em geral, os métodos não iterativos são computacionalmente mais atrativos devido a sua velocidade na execução. No entanto, em contrapartida, os iterativos são geralmente mais acurados, baseando-se na justificativa de que a cada passagem mais informações sobre a vizinhança de um ponto é incorporada e, assim, uma classificação mais confiável pode ser obtida.
- 6) Natureza da filtragem: os algoritmos de filtragem podem remover os pontos filtrados do conjunto de dados, ou recolocar os pontos filtrados no conjunto de dados. Os métodos que removem os pontos filtrados (SITHOLE, 2001; ROGGERO, 2001; AXELSSON, 1999; PFEIFER *et al.*, 1998; SOHN e DOWMAN, 2002) normalmente operam sobre a nuvem de pontos LASER original, com os pontos irregularmente espaçados. Já os métodos de recolocação (ELMQVIST, 2001; BROVELLI *et al.*, 2002; WACK e WIMMER, 2002) retornam os pontos filtrados para o conjunto de dados, mas com diferentes alturas, tendo por base alguma estratégia de interpolação a partir de pontos vizinhos. Esses métodos geralmente operam sobre malhas regularmente espaçadas.

#### 2.6.2.2. *Classificação*

Na etapa da classificação das regiões altas, a propriedade do pulso LASER de penetrar entre a vegetação é o elemento chave para a distinção entre vegetação e edifício (DAL POZ, 2013). Ao transpassar a vegetação, o pulso LASER pode atingir o terreno, gerando regiões com superfícies irregulares e rugosas. Já no caso dos edifícios, que possuem telhados planos, o retorno é de apenas um pulso, gerando superfícies regulares. Com base



nessa diferença, algoritmos podem ser criados para a classificação de vegetação e edifício, como pode ser visto em Axelsson (1999).

Os edifícios geralmente possuem telhados planos, o que origina uma rede triangulada de pontos descrita por triângulos conectados com orientações similares em cada face do telhado. Dessa forma, desconsiderando irregularidades geométricas (rugosidade do próprio telhado e detalhes do telhado) e erros grosseiros, é possível modelar um telhado através de segmentos de planos justapostos.

Considerando inicialmente o caso unidimensional, ao longo de linhas de varredura, pode-se modelar um perfil de telhado através de sucessivas retas conectadas (DAL POZ, 2013). Para cada segmento de reta, pode-se escrever:

$$z = f(x) = ax + b \quad (2.21)$$

onde  $z$  é a elevação e  $x$  a distância ao longo do perfil.

Pela Equação 2.21 tem-se que:

$$\frac{\partial^2 z}{\partial x^2} = 0 \quad (2.22)$$

A Equação 2.22 mostra que idealmente bastaria procurar pontos com derivada segunda nula para detectar os pontos pertencentes ao plano considerado (AXELSSON, 1999). Mas em problemas práticos tem-se que levar em consideração o conceito de planaridade, que está associado a dois fatores básicos: 1) o fato do próprio telhado não ser rigorosamente plano, apresentando uma pequena rugosidade; e 2) o fato dos pontos apresentarem incertezas decorrentes do processo de medida do sistema de VLA. Assim, a derivada segunda precisa ser considerada nula em relação a um certo limiar que reflita essas incertezas, ou seja, um desvio-padrão possível de ser estimado a partir de incertezas associadas com a rugosidade do telhado e da precisão associada a cada ponto da nuvem obtida por posicionamento LASER.

Os limites dos planos também podem ser identificados com base no critério de curvatura. Os pontos pertencentes a essas discontinuidades devem apresentar (AXELSSON, 1999):

$$\frac{\partial^2 z}{\partial x^2} \neq 0 \quad (2.23)$$

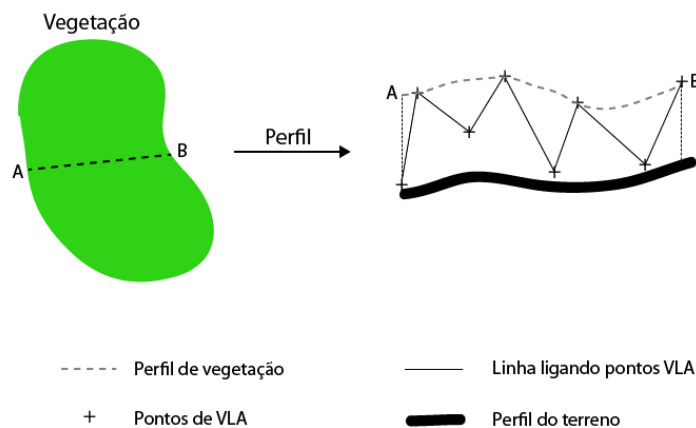
Os princípios apresentados anteriormente podem ser estendidos para o caso bidimensional, em que a análise é realizada sobre uma representação poliédrica obtida por uma rede triangular de pontos. Neste caso considera-se que as normais às faces triangulares, que compõem o plano em estudo, devem ser aproximadamente paralelas.

Assim, a detecção de pontos pertencente a uma região plana definida por pontos LASER estruturados numa rede triangulada, pode ser feita utilizando o seguinte procedimento (DAL POZ, 2013):

- 1) Para cada ponto  $P$ , identificar os triângulos que compartilham o vértice  $P$ .
- 2) Determinar a curvatura da superfície em  $P$  através da média ( $m$ ) dos ângulos entre os vetores normais aos triângulos identificados anteriormente.
- 3) A curvatura em  $P$  pode ser considerada nula (plano) se  $m$  for menor que um dado limiar, estimado a partir do conceito de planaridade. Numa sub-região englobando pontos praticamente coplanares, os vetores normais aos triângulos dessa região serão praticamente paralelos, assim, quanto mais plana for a sub-região em análise, mais próximo de zero será  $m$ .
- 4) Por fim, os pontos com curvatura considerada nula podem ser agrupados (como num processo de crescimento de regiões) para formar a região de telhado.

No caso de uma região de vegetação, a varredura a LASER pode gerar um perfil do tipo apresentado na Figura 18.

Figura 18 – Perfil gerado pela conexão de pontos LASER sequenciais numa região de vegetação.



Fonte: Adaptado de Dal Poz (2013).

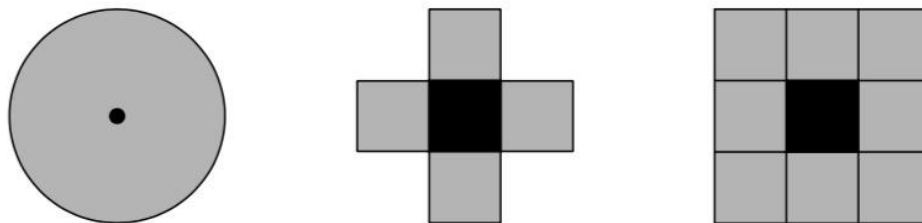
Como pode ser notado, o perfil gerado pela conexão dos pontos sequenciais é bastante irregular. Sendo assim, ao longo desse perfil a segunda derivada normalmente é diferente de zero e varia randomicamente (AXELSSON, 1999). Essa variação não é só local, mas se estende por um longo segmento, o que a diferencia das descontinuidades locais típicas de telhado (bordas, cumeeiras e terreno).

Na prática, uma região estruturada na forma de uma rede triangular apresentaria os vetores normais às faces dos triângulos variando bastante em direção e de forma randômica. Isso possibilita que pontos com essas características possam ser agrupados para formar regiões de vegetação.

## **2.7. Morfologia Matemática**

A Morfologia Matemática (MM) é definida como uma ferramenta para extrair informações relativas à geometria e à topologia de uma imagem através de um subconjunto que possui forma e tamanho pré-definidos, denominado elemento estruturante (EE). O subconjunto conhecido do EE é comparado com o conjunto desconhecido da imagem através de uma transformação, com o intuito de testar e quantificar de que maneira o EE está ou não contido na imagem (FACON, 1996). Sendo assim, a forma e o tamanho do EE devem ser estabelecidos de acordo com a geometria dos objetos de interesse na imagem. A Figura 19 apresenta três exemplos de EE, sendo a origem do EE representada pelo pixel preto.

Figura 19 – Exemplos de elementos estruturantes.



Fonte: Adaptado de Soille (2004).

A linguagem morfológica se baseia na teoria dos conjuntos para representar os objetos presentes na imagem. Em uma imagem binária, por exemplo, o conjunto de todos os pixels brancos pode descrever completamente a imagem, uma vez que os demais pixels só podem ser pretos. Assim, cada elemento do conjunto é um vetor bidimensional  $(x, y)$

contendo, por convenção, as coordenadas dos pixels brancos da imagem. Já as imagens em níveis de cinza podem ser representadas por conjuntos de vetores tridimensionais. Nesse caso, os vetores têm três elementos, sendo  $(x, y)$  as coordenadas do pixel e o terceiro elemento o seu nível de cinza (SOILLE, 2004).

Os principais operadores morfológicos são a dilatação e erosão. A partir da combinação desses dois operadores, outras operações mais complexas podem ser realizadas, como por exemplo a abertura e o fechamento. A seguir serão apresentados os conceitos básicos desses operadores.

### **2.7.1. Erosão e Dilatação**

Para o caso de imagens binárias, a erosão ( $\varepsilon$ ) do conjunto  $X$  pelo elemento estruturante  $B$  é dada por (SOILLE, 2004):

$$\varepsilon_B(X) = \{x | B_x \subseteq X\} \quad (2.24)$$

onde  $B_x$  é o elemento estruturante centrado no pixel  $x \in X$ . Nessas condições, o elemento estruturante  $B$  desliza sobre a imagem  $X$ , comparando cada pixel com a vizinhança de  $x$ . O resultado serão todos os pixels  $x$  quando  $B$  estiver inteiramente contido em  $X$ .

Os efeitos da erosão em uma imagem binária são (Figura 20(b)):

- Eliminação de objetos com tamanho inferior ao elemento estruturante;
- Diminuição dos objetos presentes na imagem;
- Aumento dos buracos no interior dos objetos;
- Separação de objetos próximos.

Para uma imagem  $h$  em níveis de cinza, a erosão de  $h$  pelo elemento estruturante  $B$ , quando a origem de  $B$  estiver em  $x$ , será o valor mínimo da imagem ( $h$ ) contida na região definida por  $B$ , ou seja:

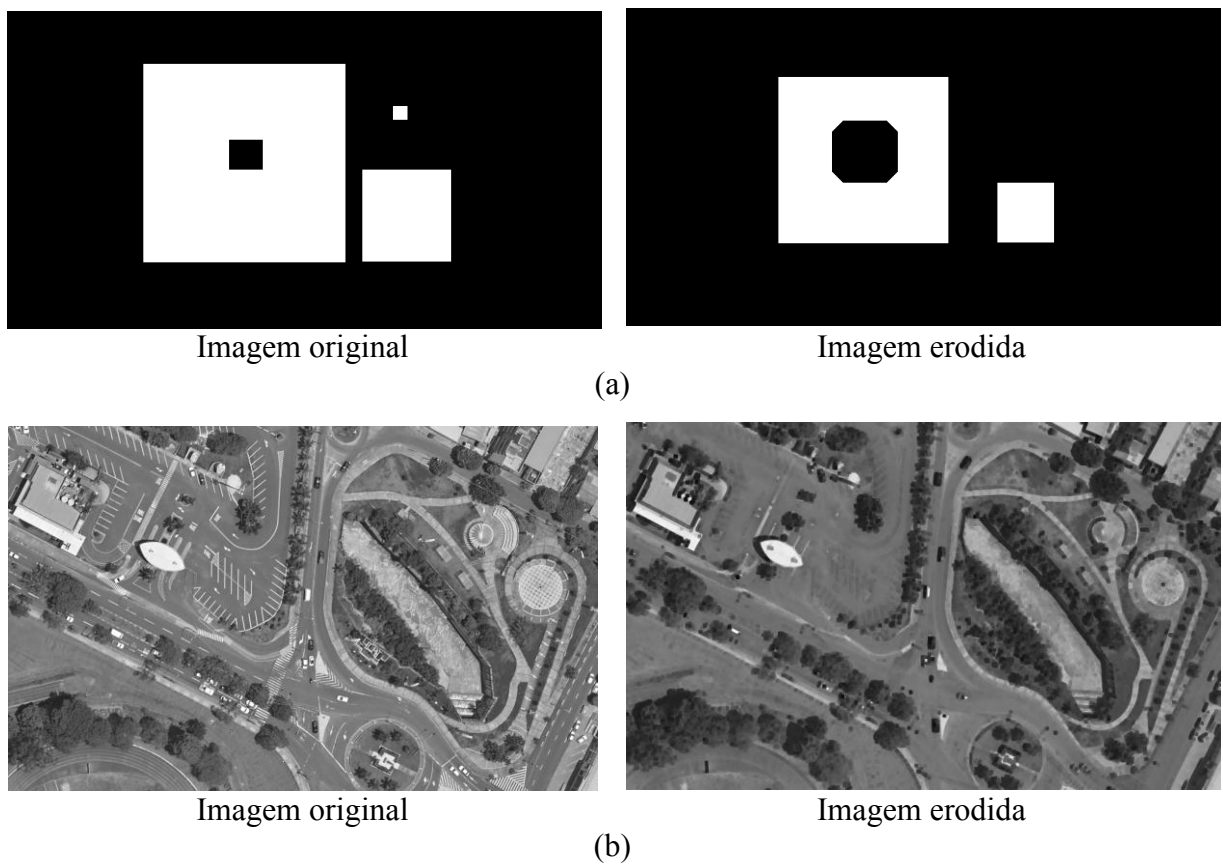
$$[\varepsilon_B(h)](x) = \min_{b \in B} h(x + b) \quad (2.25)$$

Os efeitos da erosão em uma imagem em níveis de cinza são (Figura 20(a)):

- Escurecimento da imagem;
- Conexão de padrões escuros;
- Alargamento e aumento de padrões escuros;
- Redução e eliminação de padrões claros;
- Separação de padrões claros.

Figura 20 – Efeitos da erosão utilizando o EE octógono de tamanho 21 pixels.

(a) Imagem binária. (b) Imagem em níveis de cinza.



O processo de dilatação ( $\delta$ ) do conjunto  $X$  pelo elemento estruturante  $B$  é dado por (SOILLE, 2004):

$$\delta_B(X) = \{x | B_x \cap X \neq \emptyset\} \quad (2.26)$$

onde  $B_x$  é o elemento estruturante centrado no pixel  $x \in X$ , e  $\emptyset$  é um conjunto vazio.

Os efeitos da dilatação em imagens binárias são (Figura 21(a)):

- Preenchimento de pequenos buracos;
- Aumento do tamanho dos objetos;
- Conexão de objetos próximos.

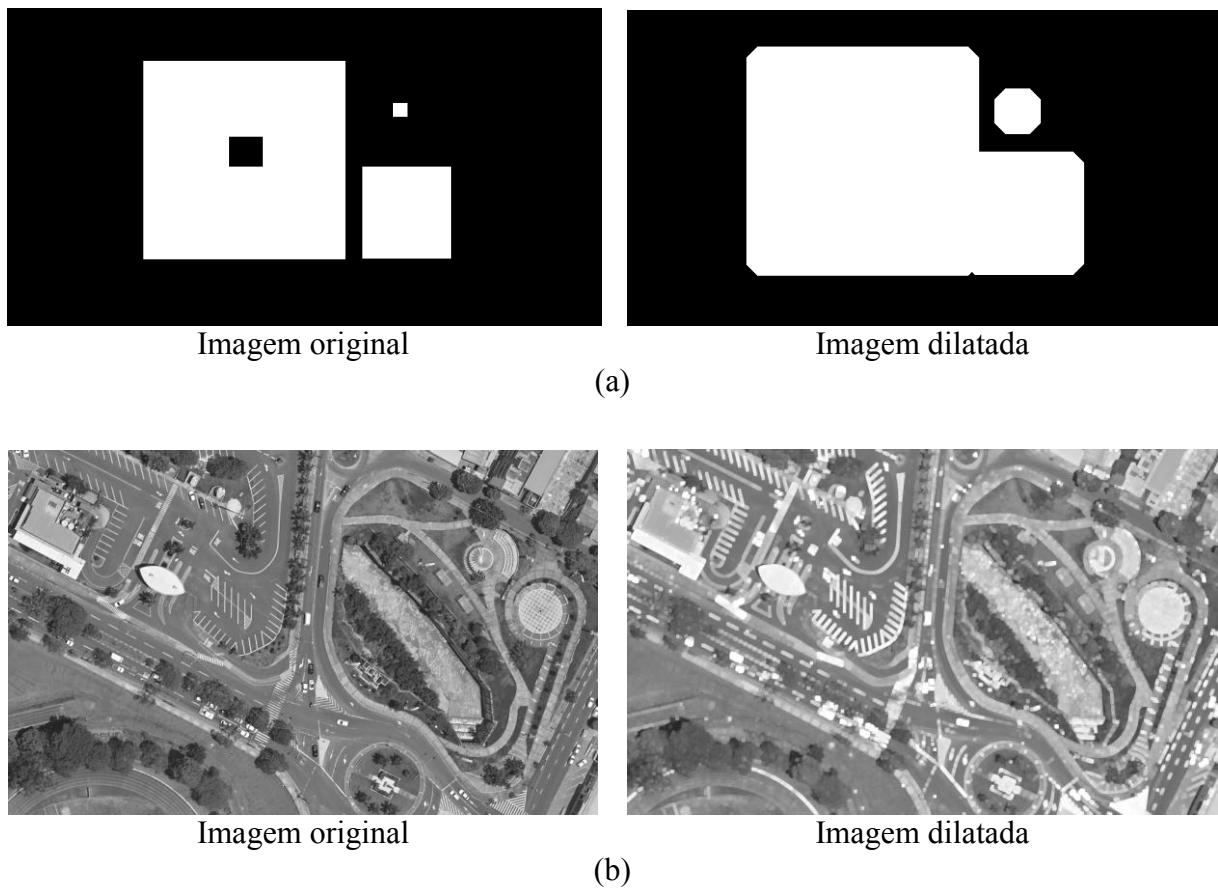
Para imagens em níveis de cinza, a dilatação de uma imagem  $h$  por elemento estruturante  $B$  é dado por (SOILLE, 2004):

$$[\delta_B(h)](x) = \max_{b \in B} h(x + b) \quad (2.27)$$

A dilatação em imagens em níveis de cinza causa os seguintes efeitos (Figura 21(b)):

- Clareamento da imagem;
- Conexão de padrões claros;
- Alargamento e aumento de padrões claros;
- Redução e eliminação de padrões escuros;
- Separação de padrões escuros próximos.

Figura 21 – Efeitos da dilatação utilizando o EE octógono de tamanho 21 pixels.  
(a) Imagem binária. (b) Imagem em níveis de cinza.



### 2.7.2. Abertura e Fechamento

A abertura ( $\gamma$ ) de uma imagem  $h$  por um elemento estruturante  $B$  é definida pela erosão de  $h$  por  $B$ , seguida da dilatação usando o elemento estruturante transposto  $B'$  (SOILLE, 2004):

$$\gamma_B(h) = \delta_{B'}[\varepsilon_B(h)] \quad (2.28)$$

Os efeitos do processo de abertura em imagens em níveis de cinza são:

- Conexão de padrões escuros próximos;
- Conservação de padrões escuros distantes;
- Eliminação de padrões claros menores que o elemento estruturante;
- Separação de padrões claros próximos;
- Diminuição dos detalhes da imagem.

O fechamento ( $\phi$ ) de uma imagem  $h$  por um elemento estruturante  $B$  é definido pela dilatação de  $h$  por  $B$ , seguida da erosão usando o elemento estruturante transposto  $B'$  (SOILLE, 2004):

$$\phi_B(h) = \varepsilon_{B'}[\delta_B(h)] \quad (2.29)$$

Os efeitos do processo de fechamento em imagens em níveis de cinza são:

- Conexão de padrões claros próximos;
- Conservação de padrões claros distantes;
- Eliminação de padrões escuros menores que o elemento estruturante;
- Separação de padrões escuros próximos;
- Diminuição dos detalhes da imagem.

#### 2.7.2.1. *Transformação Top-hat*

A transformação *Top-hat* tem como objetivo detectar os elementos relevantes de uma imagem a partir do realce das regiões claras ou escuras. Para isso ela realiza operações aritméticas entre a imagem original e a imagem resultante do processo de abertura ou fechamento morfológico. O processo de abertura ou fechamento de uma imagem por um EE remove as feições que não se ajustam a esse EE. Dessa forma, as feições removidas podem ser recuperadas através da diferença aritmética com a imagem original. Existem dois tipos de transformação *top-hat*, por abertura ou por fechamento.

##### Top-hat por abertura

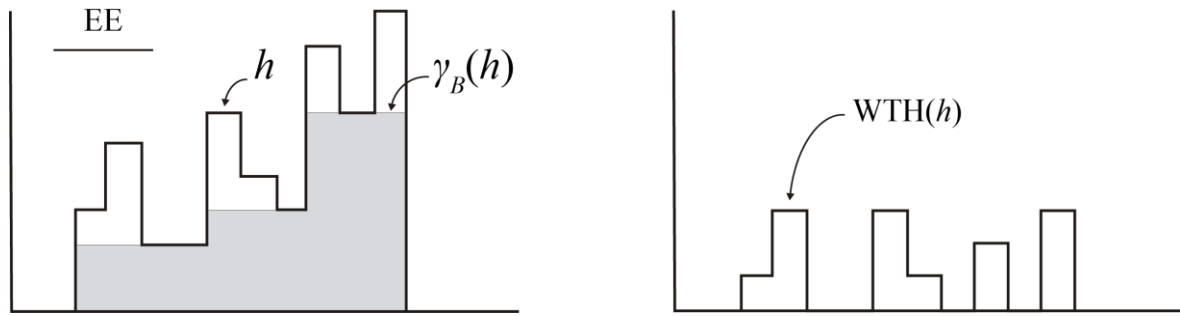
A transformação *top-hat* por abertura, denominada *White Top-hat* (WTH), é obtida a partir da diferença entre uma imagem  $h$  e sua abertura  $\gamma$  (SOILLE, 2004):

$$WTH(f) = h - \gamma(h) \quad (2.30)$$

A Figura 22 apresenta o princípio da transformação *top-hat* por abertura. É possível notar que a diferença apresentada na equação acima permite ressaltar as informações dos maiores valores de intensidade da imagem, ou seja, permite ressaltar os picos das imagens.



Figura 22 – Exemplo da transformação *top-hat* por abertura em um sinal unidimensional.



Fonte: Adaptado de Soille (2004)

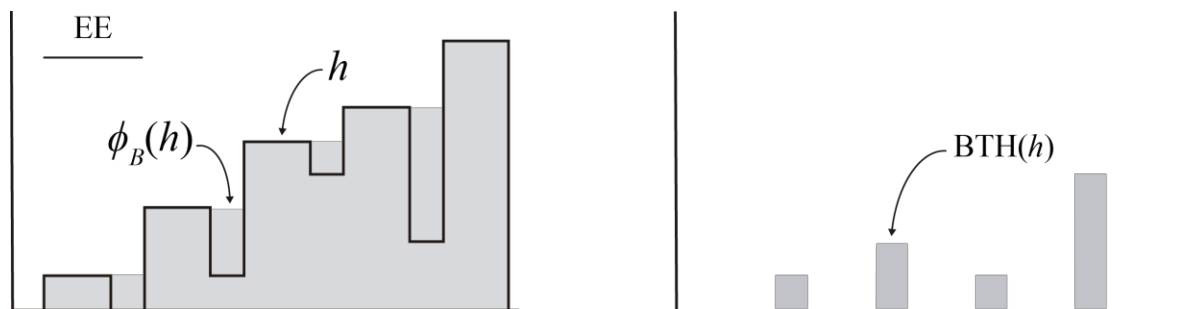
### Top-hat por fechamento

A transformação *top-hat* por fechamento, denominada *Black Top-hat* (BTH), consiste na diferença entre o fechamento  $\phi$  de uma imagem e a imagem original  $h$  (SOILLE, 2004):

$$BTH(f) = \phi(h) - h \quad (2.31)$$

A diferença apresentada na equação acima permite detectar os vales de uma imagem, ou seja, resalta as informações dos menores valores de intensidade da imagem. A Figura 23 ilustra o princípio da transformação *top-hat* por fechamento, onde todas as estruturas escuras relevantes do fundo foram suprimidas e após a diferença com a imagem original por meio do BTH, elas foram recuperadas.

Figura 23 – Exemplo da transformação *top-hat* por fechamento em um sinal unidimensional.



Fonte: Adaptado de Soille (2004)

## 2.8. Modelo de contorno ativo: *Snakes*

### 2.8.1. Modelos deformáveis generalizados

Um modelo deformável generalizado pode ser considerado como sendo uma superfície  $n$ -dimensional constituída fisicamente por um material elástico que se deforma sob a ação de forças internas e externas atuando sobre ele (LI, 1997). Um modelo desse tipo com  $d$  graus de liberdade de deformação, cada qual possuindo dimensionalidade  $p$ , pode ser representado por uma função vetorial definida por (LI, 1997):

$$v(s) = [v_1(s) \ v_2(s) \ \cdots \ v_d(s)] \quad s \in \Omega \quad (2.32)$$

onde:

- $\Omega$  é um subconjunto do espaço real de dimensionalidade  $p$ ;
- $s = (s_1, s_2, \dots, s_p)$  é um ponto em  $\Omega$ ;
- $v(s)$  é uma função vetorial que mapeia cada ponto  $s \in \Omega \subset \mathbb{R}^p$  para o espaço  $\mathbb{R}^d$ .

O modelo deformável representado pela função  $v(s)$  se estabiliza (atinge o estado de equilíbrio) quando sua energia de deformação total, dada pela Equação 2.33, atinge um valor mínimo ou máximo global.

$$E_q(v) = \lambda E_{int}(v) + (1 - \lambda) E_{ext}(v) \quad (2.33)$$

onde  $E_{int}(v)$  representa a energia interna de deformação,  $E_{ext}(v)$  representa a energia externa de deformação, e  $0 \leq \lambda \leq 1$  é o parâmetro que controla a influência relativa dos dois tipos de energia (interna e externa) atuando sobre o modelo.

A função de energia de um modelo deformável generalizado de ordem  $q$  é dada por (LI, 1997):

$$E_q(v) = \sum_{m=1}^q \sum_{j_1 \dots j_m=1}^p \int_{\Omega} w_m(s) \left| \frac{\partial^m v(s)}{\partial s_{j_1} \partial s_{j_2} \dots \partial s_{j_m}} \right|^2 ds + \int_{\Omega} w(s) |P(v(s))| ds \quad (2.34)$$

onde:

- $q$  é uma constante positiva que denota a ordem das derivadas parciais da função  $v(s)$ , que por sua vez determinam a ordem de continuidade que cada componente em  $v(s)$  possui;
- $(j_1, j_2, \dots, j_m)$  é um índice múltiplo relacionado ao índice  $m$ ;
- $w_m(s)$  e  $w(s)$  são funções de ponderação;
- $P(v(s))$  é uma função potencial generalizada.

Os dois primeiros termos da Equação 2.34 representam, respectivamente, a energia interna do modelo e a energia potencial generalizada associada ao campo de forças externas aplicado sobre ele. Tais forças são derivadas dos dados e de propriedades inerentes a eles.

Como dito anteriormente, os modelos deformáveis generalizados possuem  $d$  graus de liberdade, que representam o número de direções nas quais o objeto pode se deformar. E cada grau de liberdade possui dimensão paramétrica de ordem  $p$ . Assim, para um problema de reconstrução de superfícies, por exemplo, o modelo deformável teria  $d = 1$  grau de liberdade, representando o deslocamento  $Z$  em profundidade, com  $p = 2$  parâmetros referentes às coordenadas  $(X, Y)$ , correspondendo ao ponto de coordenadas  $(X, Y, Z)$  no espaço tridimensional ( $n = 3$ ) (LI, 1997).

### 2.8.2. Snakes

A formulação *Snake*, comumente denominada apenas de *Snake*, foi inicialmente proposta por Kass et al. (1988) e, desde então, ela tem sido amplamente aplicada na extração e refinamento de contornos de telhado. Isso se deve a sua capacidade de integrar em um único processo os dados de entrada, uma estimativa inicial e restrições previamente conhecidas do objeto para a realizar a geração do contorno ideal (CHEN et al., 2016b).

A *Snake*, ou modelo de contorno ativo, é o tipo mais conhecido de modelo deformável generalizado. Ela é representada por um conjunto de curvas definidas no  $\mathcal{R}^2$  (espaço euclidiano bidimensional) que se movem ao longo das direções  $x$  e  $y$ . Formalmente, as *Snakes* são instâncias particulares de ordem  $q = 2$  (a energia interna do modelo é constituída por derivadas de primeira e segunda ordem da função  $v(s)$ ) com dois graus de liberdade de deformação ( $d = 2$ ) e espaço paramétrico unitário ( $p = 1$ ) de modelos deformáveis generalizados.

Uma curva *Snake* é representada por uma função vetorial paramétrica bidimensional dada na forma (FAZAN, 2014):

$$v(s) = [x(s) \quad y(s)] \quad (2.35)$$

onde:

- $x(s)$  e  $y(s)$  são funções paramétricas de  $s$ ;
- $s = [s] \in [a \ b] = \Omega \subset \Re$  representa o intervalo no qual a função vetorial está definida e se refere ao comprimento de um arco da curva desde a origem até o ponto considerado.

A função que representa a energia total de uma curva *Snake* é especificada a partir da Equação 2.34 definindo-se  $q = 2$  e  $p = 1$ , resultando na equação (FAZAN, 2014):

$$E_2(v) = \sum_{m=1}^2 \int_a^b w_m(s) \left| \frac{d^m v(s)}{ds^m} \right|^2 ds + E_{ext}(v) \quad (2.36)$$

onde  $E_{ext}(v)$  representa a energia externa da função  $v(s)$ , estimada a partir dos dados da imagem e considerando-se as propriedades conhecidas a respeito das feições de interesse a serem extraídas a partir da imagem, sendo dada por:

$$E_{ext}(v) = \int_a^b w(s)P(v(s))ds \quad (2.37)$$

A expansão da somatória na Equação 2.36 resulta em (FAZAN, 2014):

$$\begin{aligned} E_2(v) &= \int_a^b w_1(s) \left| \frac{dv(s)}{ds} \right|^2 ds + \int_a^b w_2(s) \left| \frac{d^2v(s)}{ds^2} \right|^2 ds + \int_a^b w(s)P(v(s))ds \\ &= \int_a^b \left\{ w_1(s) \left| \frac{dv(s)}{ds} \right|^2 + w_2(s) \left| \frac{d^2v(s)}{ds^2} \right|^2 \right\} ds + \int_a^b w(s)P(v(s))ds \end{aligned} \quad (2.38)$$

Considerando  $x(s) = x$  e  $y(s) = y$ , têm-se (FAZAN, 2014):

$$\frac{dv(s)}{ds} = \left[ \frac{dx(s)}{ds} \quad \frac{dy(s)}{ds} \right] = \left[ \frac{dx}{ds} \quad \frac{dy}{ds} \right] \quad (2.39)$$

$$\frac{d^2v(s)}{ds^2} = \left[ \frac{d^2x(s)}{ds^2} \quad \frac{d^2y(s)}{ds^2} \right] = \left[ \frac{d^2x}{ds^2} \quad \frac{d^2y}{ds^2} \right] \quad (2.40)$$

A partir das Equações 2.39 e 2.40 é possível obter os quadrados das normas das derivadas da função de energia da Equação 2.38:

$$\left| \frac{dv(s)}{ds} \right|^2 = \begin{bmatrix} \frac{dx}{ds} & \frac{dy}{ds} \end{bmatrix} \begin{bmatrix} \frac{dx}{ds} \\ \frac{dy}{ds} \end{bmatrix} = \left( \frac{dx}{ds} \right)^2 + \left( \frac{dy}{ds} \right)^2 = \left[ \sqrt{\left( \frac{dx}{ds} \right)^2 + \left( \frac{dy}{ds} \right)^2} \right]^2 \quad (2.41)$$

$$\left| \frac{d^2v(s)}{ds^2} \right|^2 = \begin{bmatrix} \frac{d^2x}{ds^2} & \frac{d^2y}{ds^2} \end{bmatrix} \begin{bmatrix} \frac{d^2x}{ds^2} \\ \frac{d^2y}{ds^2} \end{bmatrix} = \left( \frac{d^2x}{ds^2} \right)^2 + \left( \frac{d^2y}{ds^2} \right)^2 = \left[ \sqrt{\left( \frac{d^2x}{ds^2} \right)^2 + \left( \frac{d^2y}{ds^2} \right)^2} \right]^2 \quad (2.42)$$

Substituindo as Equações 2.41 e 2.42 na Equação 2.38, têm-se:

$$\begin{aligned} E_2(v) &= \int_a^b \left\{ w_1(s) \left[ \sqrt{\left( \frac{dx}{ds} \right)^2 + \left( \frac{dy}{ds} \right)^2} \right]^2 + w_2(s) \left[ \sqrt{\left( \frac{d^2x}{ds^2} \right)^2 + \left( \frac{d^2y}{ds^2} \right)^2} \right]^2 \right\} ds \\ &\quad + \int_a^b w(s) P(v(s)) ds \\ &= \lambda E_{int}(v) + (1 - \lambda) E_{ext}(v) \end{aligned} \quad (2.43)$$

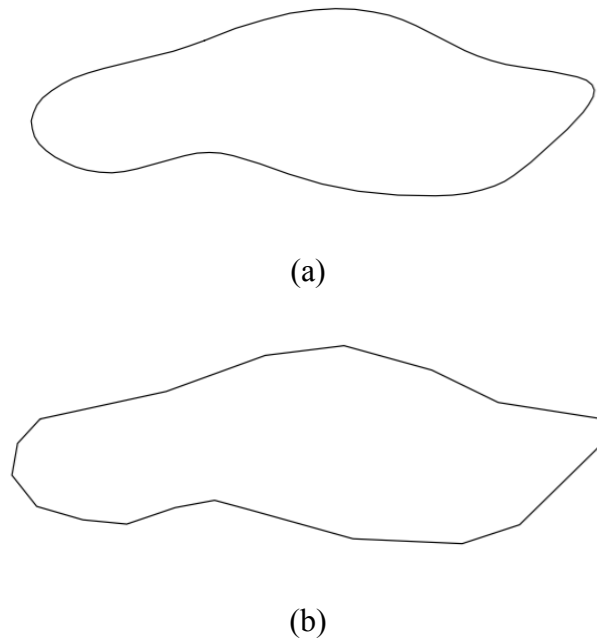
onde  $\lambda$  é o parâmetro derivado das funções de ponderação  $w_1(s)$ ,  $w_2(s)$  e  $w(s)$ , utilizado para controlar a influência relativa dos termos de energia interna e externa da curva. Quando  $\lambda \rightarrow 1$ , a solução fornecida pela Equação 2.43 torna-se genérica e robusta a ruídos. No outro extremo, quando  $\lambda \rightarrow 0$ , o modelo resultante é capaz de detectar detalhes da feição de interesse a ser extraída, mas torna-se sensível a ruídos presentes na imagem.

A Equação 2.44 apresenta uma notação mais usual para a Equação 2.43 (FAZAN, 2014):

$$E_2(v) = \int_a^b [w_1(s)|v_s|^2 + w_2(s)|v_{ss}|^2] ds + E_{ext}(v) \quad (2.44)$$

No caso da *Snake*,  $w_1(s)$  controla a flexibilidade e  $w_2(s)$  controla a rigidez da curva no ponto parametrizado por  $s$ . Quando,  $w_1(s)/w_2(s) \rightarrow 0$  a curva *Snake* apresenta um comportamento suave (Figura 24(a)), entretanto, se  $w_1(s)/w_2(s) \rightarrow \infty$  a curva *Snake* irá se comportar da forma apresentada na Figura 24(b). Em um caso extremo, como o apresentado na Figura 24(b), quando  $w_2(s)$  se anula em todos os pontos  $s$  da curva, a *Snake* perde a rigidez nesses pontos e se degenera em uma linha poligonal.

Figura 24 – Representação geométrica de uma curva *Snake*.  
 (a) Representação suave. (b) Representação não suave.



Fonte: Fazan (2014).

O cálculo da energia externa da curva *Snake* é realizado a partir dos dados da imagem, baseando-se em propriedades bem definidas das feições de interesse. A função da energia externa é mover o contorno em direção à feição a ser extraída a partir da imagem. Sendo assim, uma curva *Snake* é capaz de incorporar diferentes informações a respeito da forma dos objetos de interesse, variando de acordo com a aplicação desejada.

A condição necessária para solucionar a equação de energia *Snake* e, desta forma, extrair a feição de interesse na imagem, é dada pelas equações diferenciais parciais de Euler-Lagrange de 4ª ordem, cuja forma vetorial é representada por (FAZAN, 2014):

$$F_v - \frac{\partial}{\partial s} F_{v_s} + \frac{\partial^2}{\partial s^2} F_{v_{ss}} = 0 \quad (2.45)$$

sendo  $F(s, v, v_s, v_{ss})$  o integrando da função de energia dada pela Equação 2.44, ou seja,

$$F(s, v, v_s, v_{ss}) = w_1(s)|v_s(s)|^2 + w_2(s)|v_{ss}(s)|^2 + w(s)P(v(s)) \quad (2.46)$$

Considerando que as funções de ponderação  $w_1(s)$ ,  $w_2(s)$  e  $w(s)$  sejam dadas, respectivamente, pelas constantes  $\alpha$ ,  $\beta$  e  $\kappa$ , a Equação 2.46 pode ser reescrita da seguinte forma:

$$F(s, v, v_s, v_{ss}) = \alpha[x_s^2(s) + y_s^2(s)] + \beta[x_{ss}^2(s) + y_{ss}^2(s)] + \kappa P(x(s), y(s)) \quad (2.47)$$

Os componentes  $F_v$ ,  $F_{v_s}$  e  $F_{v_{ss}}$ , omitindo-se  $s$  na representação, são dados por (FAZAN, 2014):

$$F_v = [F_x \ F_y] = \left[ \kappa \frac{\partial}{\partial x} P(x, y) \quad \kappa \frac{\partial}{\partial y} P(x, y) \right] \quad (2.48)$$

$$F_{v_s} = [F_{x_s} \ F_{y_s}] = \left[ \alpha \frac{\partial}{\partial x_s} x_s^2 \quad \alpha \frac{\partial}{\partial y_s} y_s^2 \right] = [2\alpha x_s \quad 2\alpha y_s] \quad (2.49)$$

$$F_{v_{ss}} = [F_{x_{ss}} \ F_{y_{ss}}] = \left[ \beta \frac{\partial}{\partial x_{ss}} x_{ss}^2 \quad \beta \frac{\partial}{\partial y_{ss}} y_{ss}^2 \right] = [2\beta x_{ss} \quad 2\beta y_{ss}] \quad (2.50)$$

Substituindo as Equações 2.48, 2.49 e 2.50 na equação diferencial parcial de Euler-Lagrange, tem-se uma equação diferencial vetorial de 4ª ordem:

$$\begin{bmatrix} \kappa \frac{\partial}{\partial x} P(x, y) \\ \kappa \frac{\partial}{\partial y} P(x, y) \end{bmatrix} - \frac{\partial}{\partial s} \begin{bmatrix} 2\alpha x_s \\ 2\alpha y_s \end{bmatrix} + \frac{\partial^2}{\partial s^2} \begin{bmatrix} 2\beta x_{ss} \\ 2\beta y_{ss} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad (2.51)$$

correspondente ao seguinte sistema de equações:

$$\begin{aligned} -2\alpha x_{ss} + 2\beta x_{ssss} + \kappa \frac{\partial}{\partial x} P(x, y) &= 0 \\ -2\alpha y_{ss} + 2\beta y_{ssss} + \kappa \frac{\partial}{\partial y} P(x, y) &= 0 \end{aligned} \quad (2.52)$$

Quando se trata de uma implementação computacional, é conveniente aproximar a curva *Snake* por uma linha poligonal composta por  $n$  vértices, de modo que:

$$v_i = v(s_i) = [x(s_i) \ y(s_i)] \quad i = 1, \dots, n \quad (2.53)$$

Com isso, a solução da função de energia consiste em determinar os  $n$  vértices que descrevem a curva na imagem. Uma maneira conveniente de discretizar os termos da função

de energia (para a implementação computacional) é aproximá-los através de diferenças finitas. Assim, a Equação 2.44 se transforma em (FAZAN, 2014):

$$E_2(v) = \sum_{i=1}^n [\alpha(s_i)|v_{i+1} - v_i|^2 + \beta(s_i)|v_{i-1} - 2v_i + v_{i+1}|^2 + E_{ext}(i)] \quad (2.54)$$

Fazendo  $P_x(x, y) = \frac{\partial}{\partial x} E_{ext}(x(s_i), y(s_i))$  e  $P_y(x, y) = \frac{\partial}{\partial y} E_{ext}(x(s_i), y(s_i))$  e utilizando diferenças finitas, a equação diferencial parcial de 4ª ordem de Euler-Lagrange se torna (FAZAN, 2014):

$$\begin{aligned} & -2[\alpha(s_{i+1})(v_{i+1} - v_i) - \alpha(s_i)(v_i - v_{i-1})] \\ & + 2 \left[ \begin{array}{c} \beta(s_{i-1})(v_{i-2} - 2v_{i-1} + v_i) - \beta(s_i)(v_{i-1} - 2v_i + v_{i+1}) \\ \beta(s_{i+1})(v_i - 2v_{i+1} + v_{i+2}) \end{array} + \right] \\ & + w(s_i) [P_x(x(s_i), y(s_i)) \quad P_y(x(s_i), y(s_i))]^T = 0 \end{aligned} \quad (2.55)$$

A expansão da equação vetorial (Equação 2.55) resulta nas seguintes equações:

$$\begin{aligned} & -2[\alpha(s_{i+1})(x_{i+1} - x_i) - \alpha(s_i)(x_i - x_{i-1})] \\ & + 2 \left[ \begin{array}{c} \beta(s_{i-1})(x_{i-2} - 2x_{i-1} + x_i) - \beta(s_i)(x_{i-1} - 2x_i + x_{i+1}) \\ \beta(s_{i+1})(x_i - 2x_{i+1} + x_{i+2}) \end{array} + \right] \\ & + w(s_i) [P_x(x(s_i), y(s_i)) \quad P_y(x(s_i), y(s_i))]^T = 0 \end{aligned} \quad (2.56)$$

$$\begin{aligned} & -2[\alpha(s_{i+1})(y_{i+1} - y_i) - \alpha(s_i)(y_i - y_{i-1})] \\ & + 2 \left[ \begin{array}{c} \beta(s_{i-1})(y_{i-2} - 2y_{i-1} + y_i) - \beta(s_i)(y_{i-1} - 2y_i + y_{i+1}) \\ \beta(s_{i+1})(y_i - 2y_{i+1} + y_{i+2}) \end{array} + \right] \\ & + w(s_i) [P_x(x(s_i), y(s_i)) \quad P_y(x(s_i), y(s_i))]^T = 0 \end{aligned} \quad (2.57)$$

A condição dada pelas equações diferenciais parciais de Euler-Lagrange deve ser satisfeita pelos  $n$  vértices da linha poligonal. As Equações 2.56 e 2.57 podem ser escritas para todos os pontos na forma matricial:

$$\begin{aligned} Ax + WP_x(x, y) &= 0 \\ Ay + WP_y(x, y) &= 0 \end{aligned} \quad (2.58)$$



onde:

- $W$  é uma matriz diagonal formada pelas funções de ponderação  $w(s_i)$ ;
- $x = [x_1 \ x_2 \ \dots \ x_n]^T$ ;
- $y = [y_1 \ y_2 \ \dots \ y_n]^T$ ;
- $P_x(x, y) = [P_x(x_1, y_1) \ P_x(x_2, y_2) \ \dots \ P_x(x_n, y_n)]^T$ ;
- $P_y(x, y) = [P_y(x_1, y_1) \ P_y(x_2, y_2) \ \dots \ P_y(x_n, y_n)]^T$ ;
- $A$  é uma matriz combinada (formada por submatrizes), dada pela equação  $A = -2AD^2 + 2BD^4$ , na qual:

$$A = \begin{bmatrix} \alpha(s_1) & 0 & \dots & 0 \\ 0 & \alpha(s_2) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \alpha(s_n) \end{bmatrix}; B = \begin{bmatrix} \beta(s_1) & 0 & \dots & 0 \\ 0 & \beta(s_2) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \beta(s_n) \end{bmatrix}; D = \begin{bmatrix} -1 & 1 & 0 & \dots & 0 \\ 0 & -1 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 0 & 0 & \dots & -1 \end{bmatrix}$$

sendo  $D^i = \begin{cases} -D^{i-1}D^t, \forall i \text{ par} \\ D^{i-1}D, \forall i \text{ ímpar} \end{cases}$  o  $i$ -ésimo operador diferencial matricial.

A solução do sistema expresso na Equação 2.58 pode ser calculada iterativamente através das equações (LI, 1997):

$$\begin{aligned} x_{i+1} &= (A + \gamma I)^{-1}[\gamma x_i - WP_x(x_t, y_t)] \\ y_{i+1} &= (A + \gamma I)^{-1}[\gamma y_i - WP_y(x_t, y_t)] \end{aligned} \quad (2.59)$$

onde  $t$  é o número de iterações realizadas e  $\gamma$  é uma constante positiva usada para controlar a velocidade de convergência e estabilizar a solução.

A convergência do sistema apresentado na Equação 2.59 ocorre após  $t+1$  iterações se  $|v_{t+1} - v_t| < T$ , onde  $T$  é um limiar preestabelecido pelo operador no início do processo de extração.

## **2.9. Programação Dinâmica**

### **2.9.1. Conceitos gerais de Programação Dinâmica**

A Programação Dinâmica (PD) é considerada um método de otimização que resolve um problema por meio da combinação das soluções de subproblemas. Os algoritmos trabalham particionando os problemas em subproblemas independentes, em seguida resolvem os subproblemas recursivamente, e por fim, combinam as soluções para resolver o problema original.

Segundo Ziviani (2011), a programação dinâmica calcula a solução para todos os subproblemas partindo dos subproblemas menores para os maiores, armazenando os resultados em uma tabela. A grande vantagem está no fato de que uma vez que o subproblema é resolvido, a resposta é armazenada e o subproblema não precisa ser recalculado.

Um indicativo sobre a possibilidade de usar a programação dinâmica para a solução de um problema é quando o princípio da otimalidade pode ser aplicado (BRASSARD e BRADLEY, 1996). Esse princípio diz que em uma sequência ótima de decisões, cada subsequência também deve ser ótima.

Aho et al. (1974) apresentam a aplicação da programação dinâmica na avaliação do produto de  $n$  matrizes ( $M = M_1 \times M_2 \times \dots \times M_n$ ), onde a ordem na qual as matrizes são multiplicadas tem efeito na quantidade total de operações de adição e multiplicação.

Considerando o produto:

$$M = M_1[10,20] \times M_2[20,50] \times M_3[50,1] \times M_4[1,100]$$

onde as dimensões de cada matriz são mostradas entre colchetes. A solução de  $M$  na ordem:

$$M = M_1 \times (M_2 \times (M_3 \times M_4))$$

requer 125000 operações, enquanto a solução de  $M$  na ordem:

$$M = (M_1 \times (M_2 \times M_3)) \times M_4$$

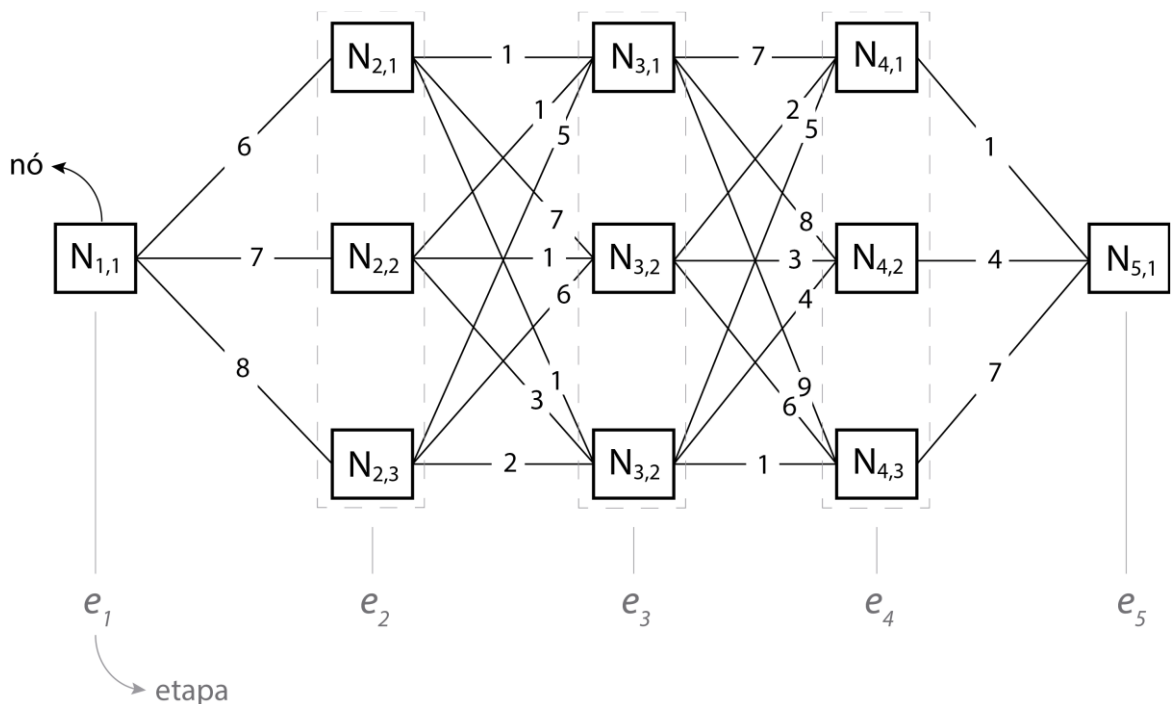
requer apenas 2200 operações.

### **2.9.2. Princípio do algoritmo de Programação Dinâmica**

A Figura 25 apresenta um dos problemas clássicos da PD, denominado problema da viagem, o qual consiste em encontrar o caminho ótimo para ir de uma cidade representada pelo nó  $N_{1,1}$  para outra cidade representada pelo nó  $N_{5,1}$ , passando necessariamente pelas

idades intermediárias. A ideia básica neste exemplo é encontrar o caminho de custo máximo em um grafo, entre os nós  $N_{1,1}$  e  $N_{5,1}$ , cuja solução pode ser encontrada sequencialmente. Na primeira decisão, é necessário escolher um dos caminhos entre o nó  $N_{1,1}$  e os nós  $N_{2,1}$ ,  $N_{2,2}$  ou  $N_{2,3}$ . Supondo que o nó  $N_{2,3}$  é a melhor escolha, a próxima decisão deve ser tomada entre os nós  $N_{3,1}$ ,  $N_{3,2}$  ou  $N_{3,3}$ . Esta estratégia é repetida até que o nó  $N_{5,1}$  seja alcançado. O caminho definido pelos nós  $N_{5,1} \rightarrow N_{4,3} \rightarrow N_{3,1} \rightarrow N_{2,3} \rightarrow N_{1,1}$  é o caminho ótimo procurado, correspondendo à trajetória de custo máximo (DAL POZ e VALE, 2003).

Figura 25 – Problema clássico de Programação Dinâmica.



Fonte: Adaptado de Gallis (2006).

Como no caso do exemplo da Figura 25 o objetivo é percorrer o caminho de custo máximo, a primeira ideia para resolver este problema é escolher sempre a viagem de maior distância. No entanto, este critério pode conduzir a um resultado que não corresponde ao caminho máximo global. O problema pode então ser resolvido de duas formas: técnica da enumeração exaustiva ou PD. A técnica da enumeração exaustiva, ou também conhecida como força bruta, é realizada a partir da combinação de todos os caminhos possíveis através do teste de todas as opções disponíveis, o que torna esta técnica pouco atrativa devido ao seu alto nível de esforço computacional. Dessa forma, a solução por PD do exemplo apresentado na Figura 25 é descrita a seguir.

O exemplo da Figura 25 é formado por 5 etapas ( $e_1, e_2, e_3, e_4, e_5$ ), sendo que as etapas  $e_1$  e  $e_5$  são compostas por 1 nó cada uma, e as etapas  $e_2, e_3$  e  $e_4$  são compostas por 3 nós cada uma. O propósito da PD é selecionar os custos máximos ou mínimos a cada par de etapas ( $e_i$  e  $e_{i+1}$ ), e então, adicionar esses custos aos custos parciais subsequentes. A seguir são descritas os passos para a solução via PD do exemplo da Figura 25 (GALLIS, 2006).

#### Etapa 1: Mapeamento da função objetivo

A função objetivo para resolver o problema apresentado na Figura 25 depende apenas de 2 variáveis (dois estados subsequentes  $e_i$  e  $e_{i+1}$ ). Portanto, a função objetivo pode ser transformada em um somatório de subfunções que dependam de 2 variáveis, conforme a equação a seguir (GALLIS, 2006):

$$g(e_1, e_2, e_3, \dots, e_n) = g_1(e_1, e_2) + g_2(e_2, e_3) + \dots + g_{n-1}(e_{n-1}, e_n) \quad (2.60)$$

O mapeamento (ou valoração) da função objetivo consiste no cálculo dos custos entre os nós de 2 etapas consecutivas. Estes custos são armazenados em tabelas denominadas “custos parciais”. A Tabela 1 apresenta um exemplo de tabela de “custos parciais” onde são considerados os 3 nós das etapas  $e_2$  e  $e_3$ , totalizando 9 possíveis custos, que ficam armazenados em uma matriz quadrada de dimensões  $3 \times 3$ .

Tabela 1 – Tabela da matriz de “custos parciais”.

$e_2, e_3$	$N_{3,1}$	$N_{3,2}$	$N_{3,3}$
$N_{2,1}$	1	7	1
$N_{2,2}$	1	1	3
$N_{2,3}$	5	6	2

#### Etapa 2: Maximização da função objetivo

No exemplo apresentado, a PD será utilizada para a maximização da função objetivo, no entanto, tratamento análogo pode ser dado para realizar a minimização da função objetivo (BALLARD E BROWN, 1982). A maximização da função objetivo é feita através da

verificação dos valores máximos que chegam na etapa seguinte. O processo pode ser dividido em duas partes:

- (a) Obtenção da matriz quadrada de “custos acumulados”: a matriz de “custos acumulados” (Tabela 4) é obtida a partir da soma de cada linha da matriz de “custos parciais” (Tabela 2) ao elemento correspondente da matriz coluna de “custos máximos” (Tabela 3).

Tabela 2 – Tabela da matriz de "custos parciais"  $g_2(e_2, e_3)$ .

<b>Nó posterior</b>	<b>N<sub>3,1</sub></b>	<b>N<sub>3,2</sub></b>	<b>N<sub>3,3</sub></b>
<b>Nó anterior</b>			
<b>N<sub>2,1</sub></b>	1	7	1
<b>N<sub>2,2</sub></b>	1	1	3
<b>N<sub>2,3</sub></b>	5	6	2

Tabela 3 – Tabela da matriz de "custos máximos"  $f_1(e_2)$ .

<b>Nó anterior</b>	<b>Custo</b>	<b>Nó posterior</b>
<b>N<sub>1,1</sub></b>	6	<b>N<sub>2,1</sub></b>
<b>N<sub>1,1</sub></b>	7	<b>N<sub>2,2</sub></b>
<b>N<sub>1,1</sub></b>	8	<b>N<sub>2,3</sub></b>

Tabela 4 – Tabela da matriz de "custos acumulados"  $f_1(e_2) + g_2(e_2, e_3)$ .

<b>Nó posterior</b>	<b>N<sub>3,1</sub></b>	<b>N<sub>3,2</sub></b>	<b>N<sub>3,3</sub></b>
<b>Nó anterior</b>			
<b>N<sub>2,1</sub></b>	7	13	7
<b>N<sub>2,2</sub></b>	8	8	10
<b>N<sub>2,3</sub></b>	13	14	10

- (b) Obtenção da matriz coluna de “custos máximos”: a matriz coluna de “custos máximos” (Tabela 5) é obtida, verificando-se dentro de cada coluna da matriz de “custos acumulados” qual é o maior elemento. Na Tabela 5, o elemento máximo que chega em  $N_{3,2}$  vale 14, e parte de  $N_{2,3}$ .

Tabela 5 – Tabela da matriz de "custos máximos".

Nó anterior	Custo	Nó posterior
N <sub>2,3</sub>	13	N <sub>3,1</sub>
N <sub>2,3</sub>	14	N <sub>3,2</sub>
N <sub>2,3</sub>	10	N <sub>3,3</sub>

Etapa 3: Caminho inverso (tracking back)

O caminho inverso consiste em fazer a busca dos custos máximos e seus respectivos nós de trás para frente. Sendo assim, verifica-se na última tabela de “custos máximos” (Tabela 6) qual o custo máximo e os seus respectivos nós. O que se deseja realmente é o custo máximo dentre todos os custos máximos.

Tabela 6 – Última tabela da matriz de "custos máximos".

Nó anterior	Custo	Nó posterior
N <sub>4,1</sub>	21	N <sub>5,1</sub>
N <sub>4,2</sub>	26	N <sub>5,1</sub>
N <sub>4,3</sub>	28	N <sub>5,1</sub>

Pela Tabela 6 verifica-se que o custo máximo dentre todos os máximos é 28, e que os nós envolvidos são N<sub>5,1</sub> e N<sub>4,3</sub>. A partir daí, pesquisa-se na tabela de “custos máximos” anterior o nó N<sub>4,3</sub>, e assim constrói-se o caminho inverso até a primeira tabela. Isto pode ser feito porque existe a superposição de um nó entre as tabelas de custos máximos.

A Tabela 7 apresenta todo o caminho inverso para a solução do problema apresentado na Figura 25. Resultado da sequência de nós no caminho inverso: N<sub>5,1</sub> → N<sub>4,3</sub> → N<sub>3,1</sub> → N<sub>2,3</sub> → N<sub>1,1</sub>.

Tabela 7 – Caminho inverso com a superposição de 1 nó.

Etapa posterior	Estado posterior	Custo máximo	Etapa anterior	Estado anterior
5	1	28	4	3
4	3	21	3	1
3	1	13	2	3
2	3	8	1	1

### **3. MATERIAL E MÉTODOS**

#### **3.1. Material**

##### **3.1.1. Dados**

Os dados utilizados na realização desse trabalho foram:

- Nuvens de pontos LASER (malha irregular) de uma região urbana de Presidente Prudente – SP (densidade média de 8 pontos/m<sup>2</sup>), parte do UNESP *Photogrammetric Dataset* (TOMMASELLI et al., 2018)<sup>1</sup>. O sistema de varredura a LASER utilizado foi o RIEGL LMS-Q680i, o qual possui um ângulo de cobertura de 60°. O tipo de espelho de varredura utilizado por esse sistema é o polígono de rotação. A nuvem de pontos LASER foi adquirida no ano de 2014 com altura média de voo de 900 m;
- Imagens aéreas digitais de alta resolução com parâmetros de orientação interior e exterior, parte do UNESP *Photogrammetric Dataset* (TOMMASELLI et al., 2018)<sup>1</sup>. A câmara utilizada para a aquisição das imagens foi a Phase One modelo iXA 180, com a tecnologia CCD (*Charge Couple Device*) e lentes Schneider-Kreuznach. A distância focal nominal é 55 mm, o tamanho da imagem é 10328×7760 pixels e o tamanho do pixel é 5,2×5,2 µm. A aquisição das imagens foi realizada juntamente com os dados LASER, no ano de 2014 com uma altura média de voo de 900 m, resultando em um GSD médio de 10 cm;
- Modelos 3D de telhado gerados a partir de dados de VLA, contendo as coordenadas no espaço-objeto dos pontos de quina e cume do telhado, e os coeficientes dos planos que formam esse telhado;
- Mapas de visibilidades (GSD 0,10 m) gerados com base nas imagens aéreas e dados de VLA da área de interesse.

##### **3.1.2. Recursos de hardware e software**

Os principais recursos utilizados e disponíveis no Laboratório de Fotogrametria da FCT/UNESP foram:

---

<sup>1</sup> DOI: 10.1109/LGRS.2017.2779559

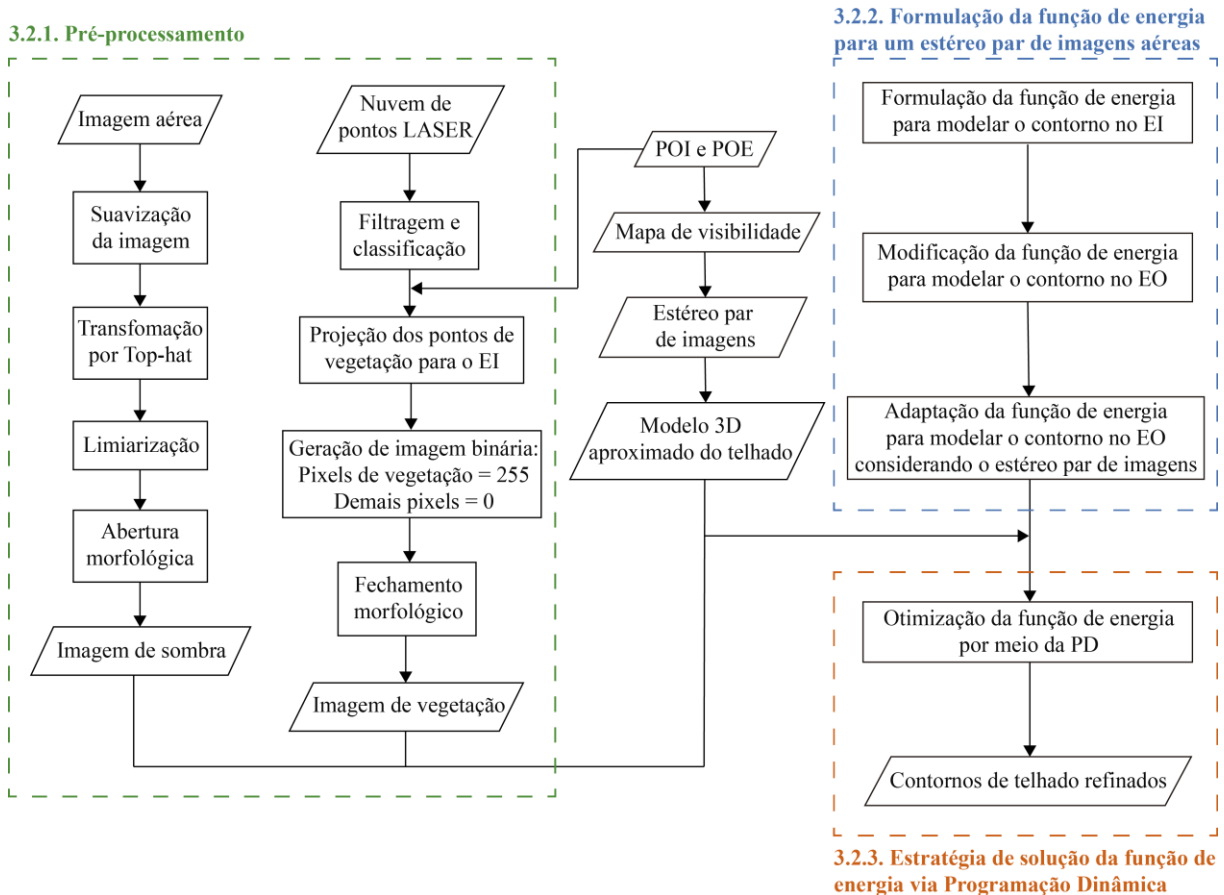
- Microcomputador e seus componentes:  
Processador: Intel Core i5 2,5 GHz  
Placa de vídeo: Intel HD Graphics 4000  
Memória: 8 GB 1600 MHz DDR3
- Software *LAStools* (versão livre) para o processamento dos dados de VLA;
- Biblioteca de classes Halcon/MVTec de processamento de imagem e visão computacional;
- Software Matlab para a implementação da etapa de pré-processamento;
- Software Visual Studio para a implementação do método de otimização via PD;
- Software LPS (*Leica Photogrammetry Suite*) para a obtenção dos contornos de referência.

### **3.2. Métodos**

Esta seção apresenta o método proposto para a extração dos contornos de telhados de edifícios no espaço-objeto. O método pode ser dividido em três etapas principais: 1) realização de um pré-processamento para a geração de duas imagens auxiliares (vegetação e sombra) que foram utilizadas durante o processo de otimização; 2) formulação de um modelo matemático para representar os contornos de telhados no espaço-objeto utilizando um estéreo par de imagens; e 3) solução do modelo matemático através da otimização por PD. As subseções a seguir apresentam os processos envolvidos em cada uma dessas etapas. A Figura 26 apresenta o fluxograma do método proposto.



Figura 26 – Fluxograma do método proposto.



### 3.2.1. Pré-processamento: geração das imagens auxiliares de vegetação e sombra

O processo de extração de contornos de telhados a partir de imagens aéreas é prejudicado por diversas feições presentes nas cenas que podem atrapalhar ou confundir o algoritmo de extração. Em ambientes urbano as principais feições que causam esses problemas são:

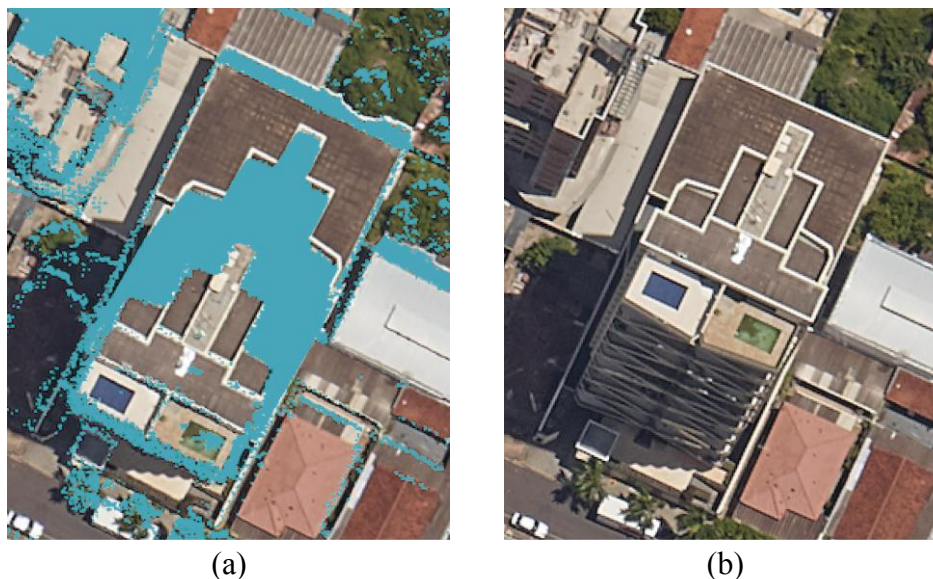
- As vegetações que se elevam acima do telhado, as quais causam oclusões impedindo a obtenção correta da resposta radiométrica das bordas de telhados nas partes oclusas;
- As sombras adjacentes às bordas dos telhados, causadas pelo próprio telhado. Nesse caso as bordas das sombras possuem respostas de gradiente mais altas, se comparadas às respostas das bordas de telhado, e assim, podem ser confundidas com o contorno do telhado;

- As obstruções perspectivas causadas por edifícios elevados e adjacentes. Em ambiente urbano é muito comum a presença de edifícios com vários pavimentos, o que acaba causando a obstrução de telhados vizinhos dependendo do ponto de vista da tomada da fotografia. Assim como no caso das vegetações elevadas, isso impede a obtenção correta da resposta radiométrica das partes obstruídas.

Com base nisso, a etapa de pré-processamento do método proposto tem como objetivo minimizar esses problemas a partir da criação de duas imagens auxiliares contendo: 1) as regiões de vegetação elevada; e 2) as regiões de sombras adjacentes ao telhado. Essas imagens são geradas com a mesma dimensão e resolução das imagens ópticas. Assim, dado um pixel da imagem óptica, é possível saber se o mesmo está em alguma dessas regiões de interesse, e então, adaptar os parâmetros da função de energia *Snake* para tratar cada caso, como será descrito na Subseção 3.2.3.4. A seguir são apresentados os processos para a criação das imagens auxiliares.

Vale ressaltar que além das imagens auxiliares de vegetação e sombra, o mapa de visibilidade também é utilizado durante o processo de solução da função de energia *Snake*, tendo como objetivo indicar as regiões contendo obstruções perspectivas. Ao contrário das imagens de vegetação e sombra, os mapas de visibilidade utilizados não foram gerados neste trabalho, portanto foram considerados como dado de entrada nos experimentos. A Figura 27(a) apresenta uma amostra do mapa de visibilidade utilizado neste trabalho, onde a região em azul corresponde à área oclusa, e a Figura 27(b) apresenta a imagem aérea original.

Figura 27 – (a) Amostra do mapa de visibilidade, em azul a região oclusa. (b) Imagem aérea correspondente.



#### 3.2.1.1. Imagem de vegetação

O primeiro passo para a geração da imagem de vegetação é determinar as regiões de vegetação elevada a partir dos dados de VLA utilizando o software *LAStools*. Primeiro, os pontos elevados são determinados através da ferramenta *LASground*. Essa ferramenta gera uma malha irregular a partir da nuvem de pontos de entrada e, a partir disso, classifica os pontos em “terreno” e “não-terreno” (Figura 28(a)). Essa ferramenta possui quatro parâmetros que podem ser alterados conforme o tipo de região:

- *Step*: parâmetro medido em metros e deve ser alterado conforme o tipo de região. Este parâmetro se refere à dimensão do menor objeto presente na área em análise. Quanto maior for o valor do *step* (passo), menor a chance de um edifício ou árvore ser considerado como ponto do terreno. Para a escolha do valor deste parâmetro o *LASground* fornece quatro opções de tipos de região, podendo ser: “floresta ou colinas”, “cidades pequenas ou planícies”, “cidades grandes ou galpões” ou “metrópoles”; além disso fornece também a opção “customizado”.

- *Spike*: parâmetro medido em metros, responsável por remover picos acima e abaixo do valor estabelecido.

- *Offset*: parâmetro medido em metros, responsável por aceitar como pontos do terreno aqueles que possuírem altura até o valor estabelecido.

- *Stddev*: parâmetro medido em centímetros, que irá considerar um desvio padrão de determinado valor como ponto de terreno para áreas planas.

Na próxima etapa, a ferramenta LASclassify é utilizada para classificar os pontos de “não-terreno” em “edifício” ou “vegetação” (Figura 28(b)). Essencialmente, a ferramenta procura classificar as regiões altas em edifício ou vegetação, através da análise da altura, planaridade, e rugosidade de uma vizinhança. Para isso, a ferramenta possibilita que sejam estabelecido valores para os seguintes parâmetros:

- *Ground offset*: referente à altura mínima que um ponto deve ter para ser considerado “não terreno”.

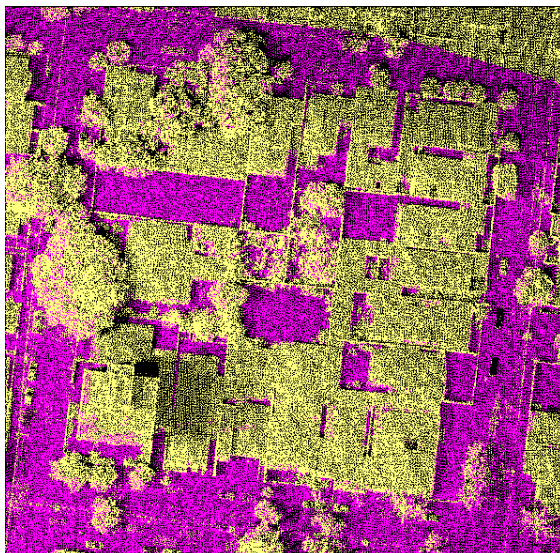
- *Building planarity*: limiar de planaridade para verificar se a vizinhança em análise é plana. Vizinhanças com valor de planaridade abaixo desse limiar serão consideradas planas.

- *Forest ruggedness*: limiar de rugosidade para verificar se a vizinhança em análise é vegetação. Vizinhanças com valor de rugosidade acima desse limiar serão consideradas vegetação.

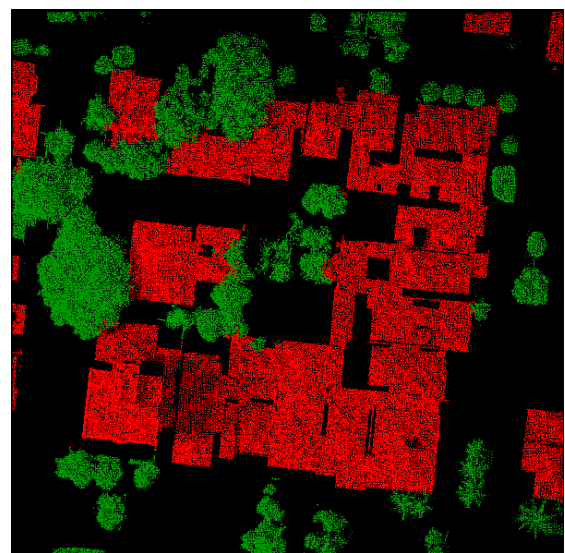
Por fim, é possível se obter somente os pontos classificados como “vegetação”. A Figura 28 apresenta um exemplo dos resultados de cada uma dessas etapas.

Figura 28 – Etapas da classificação da nuvem de pontos LASER.

(a) Pontos de "terreno" (magenta) e "não-terreno" (amarelo); (b) Pontos de "edifício" (vermelho) e "vegetação" (verde).



(a)



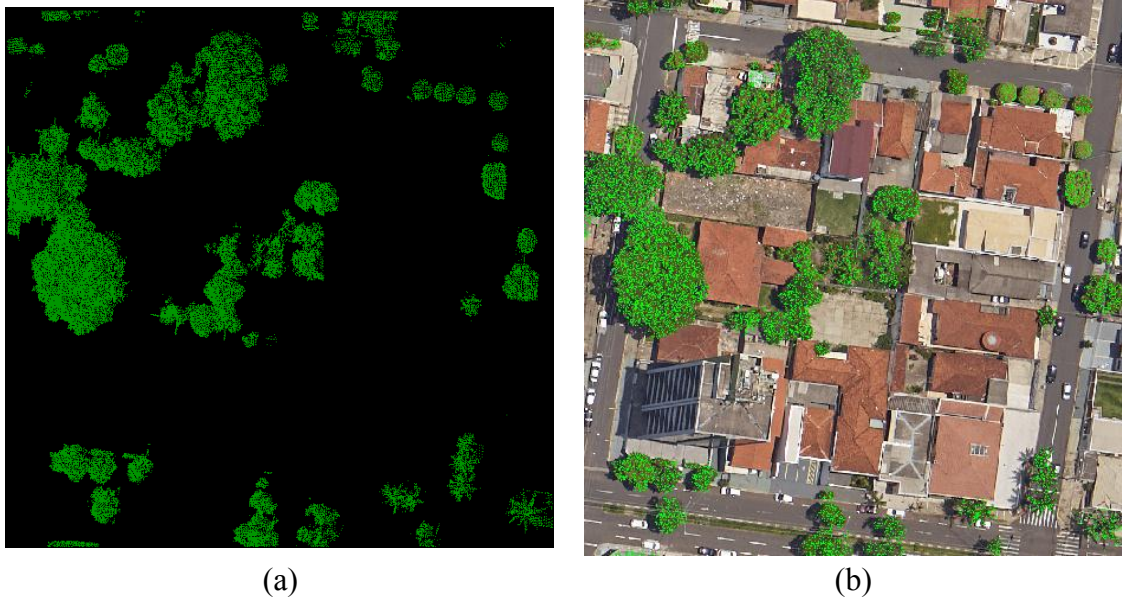
(b)

O próximo passo para a geração da imagem de vegetação consiste em projetar os pontos de vegetação, que estão no sistema de coordenadas dos dados de VLA ( $E, N, H$ ), para

o sistema de coordenadas da imagem ( $l, c$ ). Para isso foi implementado um programa em Linguagem C que realiza as transformações entre espaço-objeto e espaço-imagem através das equações de colinearidade e introdução dos erros sistemáticos, como descrito na Seção 2.1.3. A Figura 29(a) e (b) apresenta a nuvem de pontos de vegetação obtida na etapa anterior, e os pontos de vegetação projetados na imagem aérea, respectivamente.

Figura 29 – Pontos de vegetação.

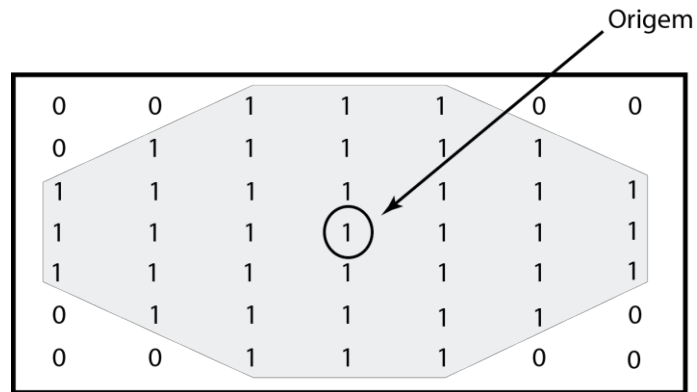
(a) Nuvem de pontos da vegetação; (b) Pontos de vegetação projetados na imagem.



Após a projeção dos pontos de vegetação para o espaço-imagem, é gerada uma imagem onde os pixels de vegetação recebem o valor 255 e os demais pixels recebem o valor 0 (Figura 31(a)).

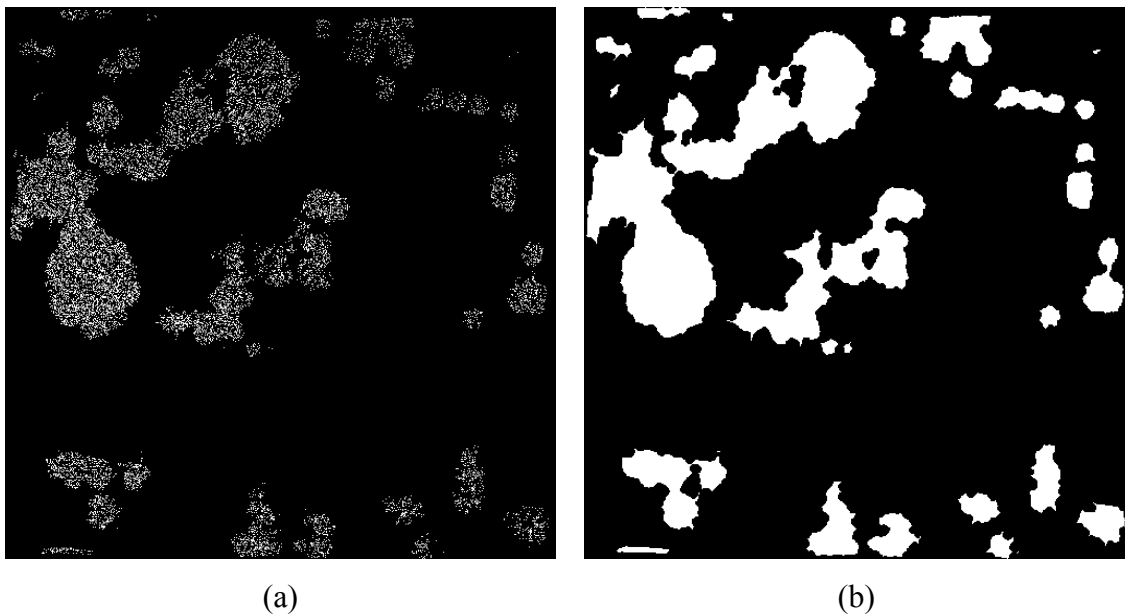
No entanto, é possível notar pela Figura 31(a) que os pontos de vegetação resultantes da nuvem de pontos LASER são espaçados e não formam regiões sólidas e homogêneas na imagem. Dessa forma, é necessário a realização de um processamento para preencher os espaços entre os pixels de vegetação de uma determinada região. Para isso, é realizado o processo de fechamento morfológico na imagem binária de vegetação utilizando o elemento estruturante “octógono” de tamanho 9. A Figura 30 apresenta uma ilustração do elemento estruturante “octógono” de tamanho 3.

Figura 30 – Ilustração do elemento estruturante octógono de tamanho 3.



O fechamento morfológico consiste na realização da operação de dilatação seguida da operação de erosão, e tem como característica a conservação de padrões claros distantes e conexão de padrões claros próximos. A Figura 31(b) apresenta o resultado da imagem de vegetação após passar por este processamento.

Figura 31 – Exemplo de imagem de vegetação gerada pelo método proposto.  
(a) Imagem de vegetação sem processamento; (b) Imagem de vegetação após o processo de fechamento.



### 3.2.1.2. Imagem de sombra adjacente ao telhado

A imagem com as regiões de sombras adjacentes aos telhados é gerada através da detecção dessas regiões na imagem aérea. Essa detecção é realizada a partir do método proposto por Azevedo et al. (2015), com algumas adaptações.

Antes de aplicar o método proposto por Azevedo et al. (2015), é necessário efetuar um processo de suavização da imagem original. Isso ocorre pois como as imagens utilizadas são de alta resolução, o método de detecção de sombra acaba detectando muitas sombras desnecessárias causadas por pequenas feições, como por exemplo, cada telha do telhado. A Figura 32(a) apresenta um exemplo dos detalhes de uma imagem aérea de alta resolução, e a Figura 32(b) traz em branco as áreas de sombra detectadas. Com base nisso, o filtro Gaussiano com  $\sigma = 2$  foi escolhido para realizar a suavização, uma vez que ele é amplamente utilizado para reduzir ruídos e detalhes em imagens.

Figura 32 – Exemplo de imagem de sombra gerada a partir de uma imagem sem suavização. (a) Detalhes de uma imagem aérea de alta resolução. (b) Sombras detectadas (em branco).



(a)



(b)

Aplicada a suavização na imagem, o primeiro passo para a detecção das sombras é realizar a transformação *top-hat* por fechamento na imagem suavizada. Essa transformação consiste em evidenciar padrões escuros da imagem através da diferença aritmética entre a imagem original e a imagem resultante do processo de fechamento. Como as sombras em imagens aéreas não obedecem um padrão específico, o trabalho original de Azevedo et al. (2015) utiliza a operação de fechamento por área, para que não haja a necessidade de estabelecer uma forma particular para o elemento estruturante. No entanto, o

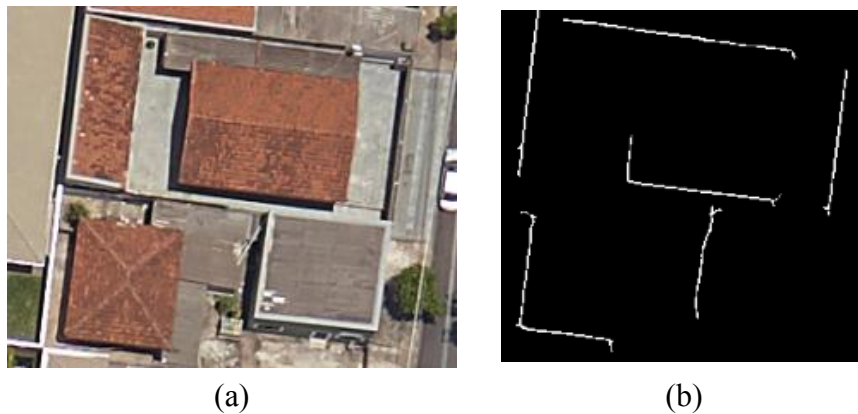
tipo de sombra de interesse do presente trabalho são aquelas formadas pelos próprios edifícios, possuindo formas retangulares e alongadas. Com base nisso, a operação de fechamento utilizada na transformação *top-hat* do método proposto é realizada utilizando-se o elemento estruturante “quadrado” de tamanho 3. Como resultado, tem-se uma imagem onde são realçadas somente as sombras que obedecem o padrão desejado.

O próximo passo consiste em realizar uma limiarização na imagem resultante da transformação *top-hat*, a fim de detectar as sombras realçadas. O histograma da imagem resultante da transformação *top-hat* possui um comportamento bimodal, ou seja, os alvos de interesse se destacam dos demais alvos da imagem. Dessa forma, o método de Otsu (1979) foi escolhido para realizar a limiarização pois ele consegue determinar automaticamente o limiar mais adequado de acordo com uma medida de separação de classes. O resultado é uma imagem binária onde os pixels de sombra possuem o valor 255 e os demais pixels possuem o valor 0.

Por fim, a operação de abertura morfológica por área é aplicada na imagem binária para eliminar regiões de sombra pequenas, as quais podem ser consideradas ruídos. A Figura 33 apresenta a imagem de sombra resultante do processo descrito.

Figura 33 – Imagem de sombra gerada pelo método proposto.

- (a) Amostra de uma imagem aérea. (b) Imagem de sombra resultante, onde as regiões em branco são as sombras detectadas.



### **3.2.2. Formulação da função de energia para um estéreo par de imagens aéreas**

A formulação da função de energia que modela os contornos de telhados no espaço-objeto considerando um estéreo par de imagens é realizada em três etapas: formulação da função de energia para modelar o contorno de telhado no espaço-imagem; modificação dessa



função de energia para modelar o contorno no espaço-objeto; e desenvolvimento da função de energia que modela o contorno no espaço-objeto considerando um estéreo par de imagens. Vale ressaltar que a solução ótima da função de energia é encontrada a partir da minimização da função proposta. A seguir é descrito o desenvolvimento de cada um desses passos.

### 3.2.2.1. *Formulação da função de energia para modelar o contorno no espaço-imagem*

O contorno do telhado é modelado no espaço-imagem a partir da introdução de termos de função de energia externa na curva *Snake* que consideram as propriedades radiométricas e geométricas dos telhados. Uma dessas propriedades refere-se ao fato das bordas de edifícios geralmente apresentarem forma contínua e limitarem regiões bem contrastadas em imagens aéreas, formando bordas com padrão de degrau (*step edges*). Essa propriedade é normalmente modelada pela Equação 3.1 (Kass et al., 1987):

$$E_{borda}(p) = -\sum_{i=1}^n \gamma_i |\nabla G(p_i)|^2 \quad (3.1)$$

onde  $n$  é o número total de pontos considerados ao longo de uma dada borda,  $\gamma_i$  é uma constante positiva associada à cada ponto  $p_i$  de borda, e  $|\nabla G(p_i)|$  é a magnitude do gradiente no ponto de borda  $p_i$  calculada através das máscaras de Prewitt (GONZALEZ e WOODS, 2000). O sinal negativo na Equação 3.1 é adotado para que o respectivo termo contribua com a minimização da função de energia quando a *Snake* coincidir exatamente sobre um contorno de telhado.

Neste trabalho é proposta uma modificação inédita na Equação 3.1, visando atenuar a energia de borda resultante da sombra projetada pelo próprio edifício, cujo contorno de telhado é alvo de extração. A borda de sombra que pode causar problema é aquela que se dispõe paralela e proximamente ao edifício, podendo ser confundida com a borda do telhado desse edifício.

Como as regiões de sombra numa imagem em tom de cinza serão bem mais escuras que as regiões adjacentes, o vetor do gradiente em pontos de contorno de telhados apontará para o interior do telhado e em pontos de contorno de sombra o vetor gradiente apontará no sentido aproximadamente oposto (Figura 34(a)). Portanto, para um lado de telhado com sombra aproximadamente paralela e próxima, os vetores do gradiente tomados nos pontos de

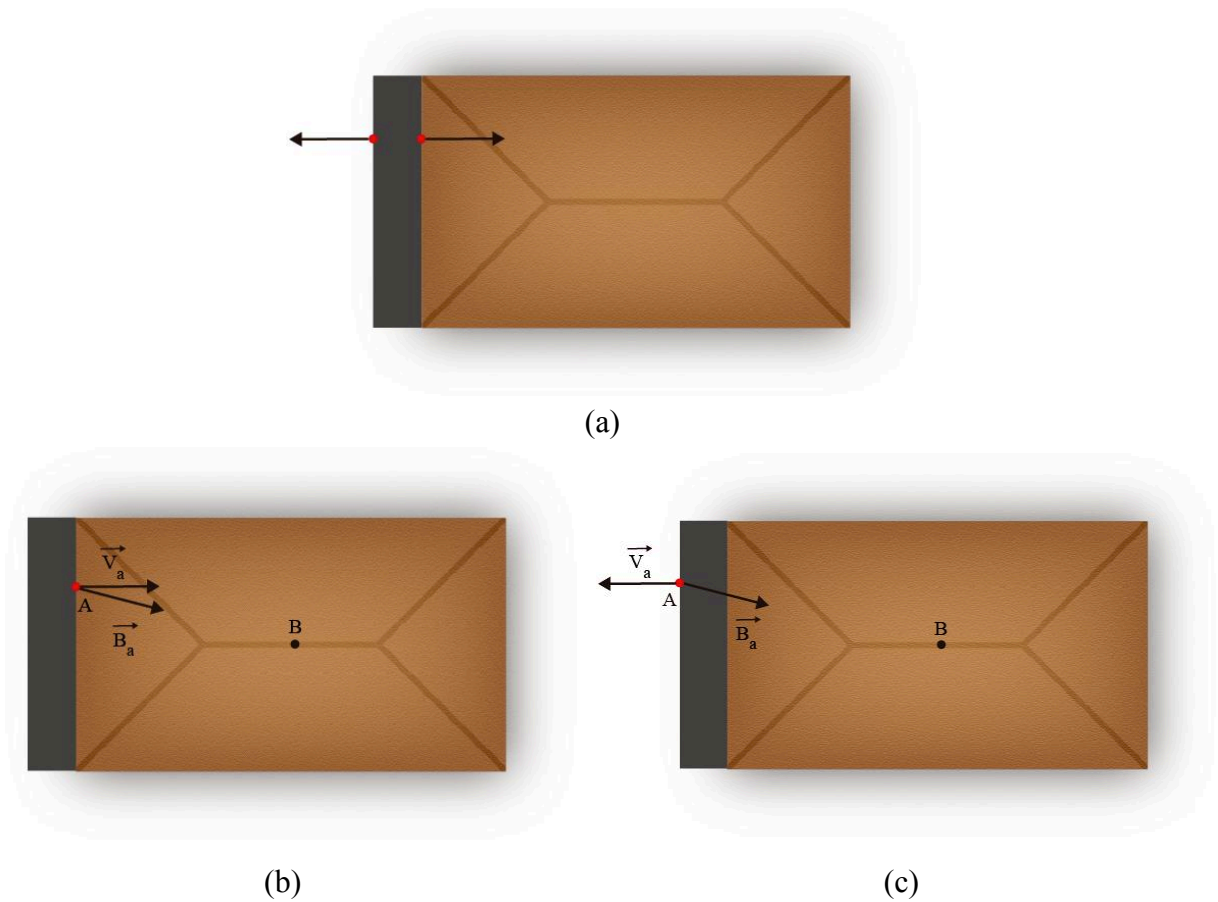
borda de telhado e de borda da sombra serão aproximadamente antiparalelos. Nessas condições, sendo B o baricentro 3D do telhado projetado na imagem, o vetor gradiente num ponto A ( $\vec{V}_a$ ) do contorno de telhado formará um ângulo agudo com o vetor ( $B_a$ ) que aponta do ponto A para B (Figura 34(b)). Consequentemente,  $\vec{V}_a \cdot \vec{B}_a > 0$  (“.” é o operador produto escalar). No entanto, se o ponto a for um ponto de borda de sombra, ambos os vetores formarão um ângulo obtuso e, consequentemente,  $\vec{V}_a \cdot \vec{B}_a < 0$  (Figura 34(c)). Com base nessa discussão, pode-se reescrever a Equação 3.1 de forma a beneficiar pontos de borda de contorno de edifício através da análise do produto escalar  $\vec{V}_a \cdot \vec{B}_a$ :

$$E_{borda}(p) = -\sum_{i=1}^n \gamma_i f(\vec{V}_a \cdot \vec{B}_a) |\nabla G(p_i)|^2 \quad (3.2)$$

onde: 
$$\begin{cases} f(\vec{V}_a \cdot \vec{B}_a) = 1, \text{ se } \vec{V}_a \cdot \vec{B}_a > 0 \\ f(\vec{V}_a \cdot \vec{B}_a) = -1, \text{ se } \vec{V}_a \cdot \vec{B}_a < 0 \end{cases}$$

Figura 34 – Fator de compensação de sombra.

- (a) Vetor gradiente nos pontos de borda de telhado e nos pontos de borda de sombra;
- (b) Vetor gradiente no ponto de borda de telhado e vetor que aponta para o baricentro do telhado;
- (c) Vetor gradiente no ponto de borda de sombra e vetor que aponta para o baricentro do telhado.



A outra propriedade considerada está ligada ao fato de que os telhados de edifícios, em sua maioria, são formados por estruturas retilíneas com quinas descritas por ângulos retos. Neste caso, o ângulo de deflexão nas quinas dos telhados pode ser usado para construir uma função de energia externa que represente esta propriedade (FAZAN e DAL POZ, 2013):

$$E_{quina}(p) = -\sum_{i=1}^n \eta_i [(1 - \cos(\delta_i)) \cdot CS(p_i)]^2 \quad (3.3)$$

onde  $n$  é o número total de quinas,  $\eta_i$  é uma constante positiva,  $\delta_i$  é o ângulo de deflexão no vértice de quina  $p_i$  e  $CS(p_i)$  é a resposta de quina calculada a partir de uma nova abordagem proposta neste trabalho. O sinal negativo na Equação 3.3 tem função idêntica do sinal negativo da Equação 3.1.

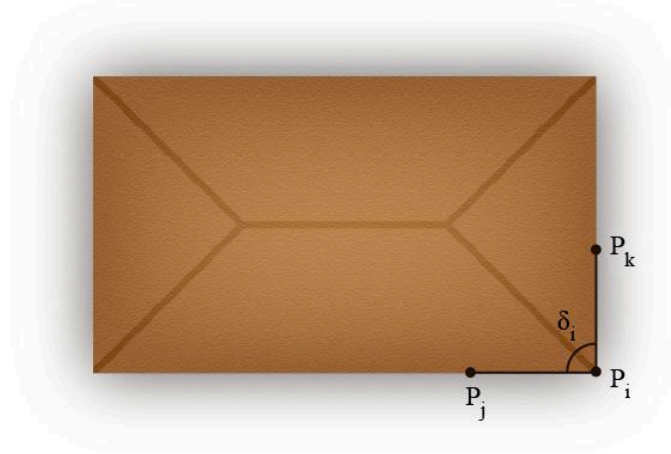
Um operador de detecção de quinas, como por exemplo o operador de Moravec (1977), realiza a detecção das quinas tendo como base somente as respostas radiométricas de uma pequena região em volta do ponto candidato à quina. Contudo, em imagens de alta resolução, como é o caso deste trabalho, os telhados de edifícios apresentam muitos detalhes presentes em sua superfície, como manchas e diferenças de tonalidade das telhas. Isso prejudica a detecção correta das quinas, pois uma região pode conter falsas quinas formadas por detalhes aleatórios do telhado. Com base nisso, uma nova abordagem para o cálculo da resposta de quina foi desenvolvida neste trabalho, combinando as informações radiométricas e as informações geométricas dos pontos candidatos à quina, a partir da ideia a seguir.

Considerando que as quinas de telhados são formadas por duas retas (que representam as bordas do telhado) que formam um ângulo de  $90^\circ$  entre si, duas retas suportes são criadas a partir do ponto candidato à quina ( $P_i$ ) e dos pontos vizinhos candidatos à borda ( $P_j$  e  $P_k$ ), como mostra a Figura 35. Se o ponto de quina e os pontos de borda forem os corretos, a soma do gradiente dos pontos sobre as retas suportes terá um valor alto. Esse atributo é então usado para calcular a resposta de quina  $CS(p_i)$ :

$$CS(p_i) = f(\vec{V}_a, \vec{B}_a) \sum_{j=1}^m |\nabla G(\overline{P_i P_j})|^2 + f(\vec{V}_a, \vec{B}_a) \sum_{k=1}^m |\nabla G(\overline{P_i P_k})|^2 \quad (3.4)$$

onde  $m$  é o número de pontos amostrados sobre cada reta suporte,  $P_i$  é o ponto candidato à quina,  $P_j$  e  $P_k$  são os pontos candidatos à borda, vizinhos de  $P_i$ ,  $|\nabla G(\overline{P_i P_j})|$  e  $|\nabla G(\overline{P_i P_k})|$  são a magnitude do gradiente em cada ponto da reta suporte  $\overline{P_i P_j}$  e  $\overline{P_i P_k}$ , respectivamente. Vale ressaltar que o fator de compensação de sombra  $f(\vec{V}_a, \vec{B}_a)$  também é utilizado nessa equação, de forma semelhante à Equação 3.2.

Figura 35 – Ideia utilizada na criação do termo  $CS$  para o cálculo da resposta de quina.



O outro termo da Equação 3.3  $(1 - \cos(\delta_i))$  é uma função de ponderação que favorece quinas em ângulos retos. Neste trabalho o ângulo de deflexão  $\delta_i$  no ponto candidato à quina também é calculado utilizando as retas suportes  $\overline{P_i P_j}$  e  $\overline{P_i P_k}$  (Figura 35), a partir da equação clássica da geometria analítica.

A função objetivo resultante, que modela o contorno do telhado no espaço-imagem, pode ser escrita incorporando na função de energia *Snake* as funções de energia externa dadas pelas Equações (3.2) e (3.3):

$$E(p) = \sum_{i=1}^n [\alpha_i |p_{i+1} - p_i|^2 + \beta_i |p_{i-1} - 2p_i + p_{i+1}|^2 - \gamma_i f(\vec{V}_a, \vec{B}_a) |\nabla G(p_i)|^2 - \eta_i [(1 - \cos(\delta_i)) CS(p_i)]^2] \quad (3.5)$$

onde  $p$  é um polígono candidato a representar um contorno de telhado no espaço-imagem.

A solução da função de energia dada pela Equação 3.5 pode ser obtida ao se encontrar um valor mínimo global de energia para essa equação, o que corresponde a um conjunto de pontos de borda e quina que representam o contorno de telhado de um edifício.

A Equação 3.5 apresenta uma estrutura em que apenas três vértices sucessivos  $(p_{i-1}(C_{i-1}, L_{i-1}), p_i(C_i, L_i), p_{i+1}(C_{i+1}, L_{i+1}))$  do polígono que representa o contorno do telhado estão simultaneamente inter-relacionados. Sendo assim, pode-se compactamente reescrever a Equação 3.5 da seguinte forma geral:

$$E(p) = \sum_{i=1}^n E_i(p_{i-1}(C_{i-1}, L_{i-1}), p_i(C_i, L_i), p_{i+1}(C_{i+1}, L_{i+1})) \quad (3.6)$$

onde  $p_{i-1}$ ,  $p_i$  e  $p_{i+1}$  são vértices consecutivos do polígono que estão simultaneamente inter-relacionados.

### 3.2.2.2. *Formulação da função de energia para modelar o contorno no espaço-objeto*

No próximo passo, a função objetivo obtida no espaço-imagem (Equação 3.6) é modificada para descrever os contornos de telhado no espaço-objeto. A relação matemática entre um ponto  $P(E, N, H)$  no espaço-objeto e o correspondente ponto  $p(c, l)$  no espaço-imagem é bem conhecida. Ela envolve as transformações projetivas através das equações de colinearidade, a modelagem de erros sistemáticos e as transformações entre referenciais no espaço-imagem. Maiores informações sobre estas transformações podem ser encontradas na literatura específica de Fotogrametria (ANDRADE, 1998; WOLF e DEWITT, 2000; MIKHAIL et al., 2001; KRAUS, 1993).

Os seguintes parâmetros estabelecem a relação matemática entre um ponto no espaço-imagem e o correspondente ponto no espaço-objeto:

- $\kappa, \varphi, \omega, X_0, Y_0, Z_0$ : parâmetros de orientação exterior da imagem aérea;
- $f$ : distância focal calibrada da câmara;
- $x_0, y_0$ : coordenadas do ponto principal no referencial fiducial da imagem aérea;
- $k_1, k_2, k_3$ : parâmetros de distorção radial das lentes da câmara;
- $P_1, P_2$ : parâmetros de distorção descentrada das lentes da câmara;

Sendo assim, considerando que as coordenadas de pontos no espaço-imagem podem ser colocadas em função das coordenadas de pontos no espaço-objeto, tem-se (FAZAN, 2014):

$$\begin{aligned} C_i &= f_1 \left( \vec{Q}, (E_i, N_i, H_i) \right) \\ L_i &= f_2 \left( \vec{Q}, (E_i, N_i, H_i) \right) \end{aligned} \tag{3.7}$$

onde  $\vec{Q} = (\kappa, \varphi, \omega, X_0, Y_0, Z_0, f, x_0, y_0, k_1, k_2, k_3, P_1, P_2)$  e  $E_i, N_i, H_i = P_i$ .

Sendo  $\vec{Q}$  um vetor de parâmetros conhecidos, um ponto no espaço-imagem pode ser expresso somente em função de  $P_i$ :

$$p_i(C_i, L_i) = p_i(f_1(P_i), f_2(P_i)) = p_i(f_1(E_i, N_i, H_i), f_2(E_i, N_i, H_i, )) \quad (3.8)$$

Substituindo a Equação 3.8 na Equação 3.6, obtém-se a seguinte função objetivo:

$$E(p) = \sum_{i=1}^n E_i \left[ \begin{array}{c} p_{i-1}(f_1(E_{i-1}, N_{i-1}, H_{i-1}), f_2(E_{i-1}, N_{i-1}, H_{i-1}, )), \\ p_i(f_1(E_i, N_i, H_i), f_2(E_i, N_i, H_i, )), \\ p_{i+1}(f_1(E_{i+1}, N_{i+1}, H_{i+1}), f_2(E_{i+1}, N_{i+1}, H_{i+1}, )) \end{array} \right] \quad (3.9)$$

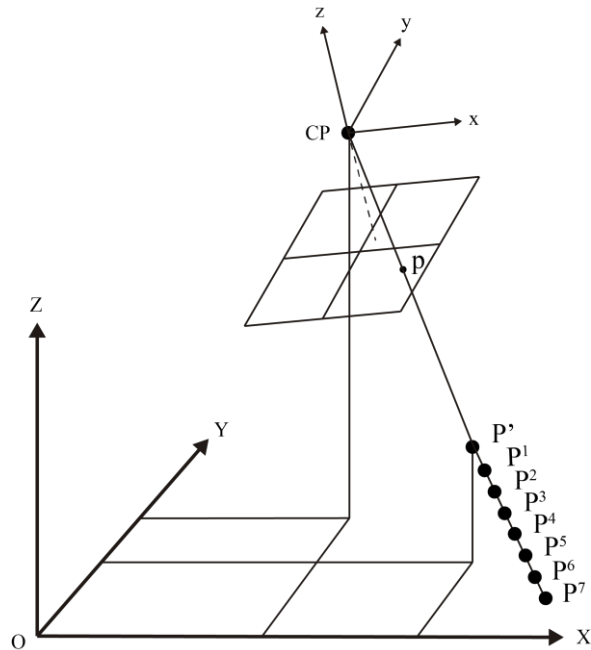
É possível observar que a Equação 3.9 depende apenas das coordenadas de três pontos consecutivos do polígono que descreve o contorno do telhado no espaço-objeto. Com isso, a equação pode ser reescrita da seguinte forma simplificada:

$$E(P) = \sum_{i=1}^n E_i(P_{i-1}(E_{i-1}, N_{i-1}, H_{i-1}), P_i(E_i, N_i, H_i), P_{i+1}(E_{i+1}, N_{i+1}, H_{i+1})) \quad (3.10)$$

onde  $P$  é um polígono 3D no espaço-objeto correspondente ao polígono  $p$  no espaço-imagem.

A função de energia dada pela Equação 3.10 é ambígua e, como tal, não pode ser a base para extração de contornos de telhados, mesmo que usando apenas uma imagem. Esse problema decorre do fato de que, matematicamente, infinitos pontos no espaço-objeto, desde que pertencentes à reta definida por  $p$  e  $P'$  (além do CP da imagem, conforme o princípio fundamental de colinearidade da Fotogrametria), formarão imagem num único ponto:  $p$ , como é ilustrado na Figura 36. Da mesma forma, têm-se infinitos polígonos 3D que formarão imagem num único polígono:  $p_i$ . Portanto, a função de energia dada pela Equação 3.10 possui infinitas soluções, visto que infinitos polígonos 3D no espaço-objeto satisfarão o critério de mínimo de  $E(P)$  num processo de otimização. A solução desta ambiguidade é obtida a partir do uso do estéreo par de imagens, como será apresentado na próxima seção.

Figura 36 – Princípio da ambiguidade.



3.2.2.3. *Formulação da função de energia que modela o contorno no espaço-objeto considerando um estéreo par de imagens*

Finalmente, o modelo matemático do contorno do telhado no espaço-objeto é desenvolvido para um estéreo par de imagens aéreas, tendo por base a Equação 3.10. Entretanto, a questão central é como desenvolver uma função de energia a partir de uma função de energia ambígua e que foi desenvolvida para uma única imagem. A fim de remover esta ambiguidade e possibilitar a obtenção de uma solução única, definiu-se uma função de energia para um estéreo par baseada na soma de funções de energia definidas para as imagens esquerda e direita. Pode-se então demonstrar que esta equação de energia pode ser compactamente dada na forma:

$$E_T(P) = E^e(P) + E^d(P) = \sum_{i=1}^n E_i^T(P_{i-1}(E_{i-1}, N_{i-1}, H_{i-1}), P_i(E_i, N_i, H_i), P_{i+1}(E_{i+1}, N_{i+1}, H_{i+1})) \quad (3.11)$$

onde  $E^e$  e  $E^d$  são as funções de energia (Equação 3.10) definidas para as imagens esquerda e direita, respectivamente, e  $E_i^T$  resulta do agrupamento dos termos semelhantes de  $E^e$  e  $E^d$ .

Para teoricamente justificar a Equação 3.11, supor que  $P^l, \dots, P', \dots, P^n$  são polígonos 3D que produzem o mesmo valor mínimo  $E^e$ . Supor também que  $Q^l, \dots, P', \dots, Q^m$  são



polígonos 3D que agora produzem o mesmo valor mínimo  $E^d$ . Notar que  $P'$  é o polígono 3D pertencente a ambas as sequências de polígonos 3D, sendo ele a representação correta do contorno de um dado telhado. A classe  $P^i$  de polígonos 3D, excluindo  $P'$ , produz valores de energia  $E^d$  maiores que  $E^e$ . Analogamente, a classe  $Q^i$  de polígonos 3D, excluindo também  $P'$ , produz valores de energia  $E^e$  maiores que  $E^d$ . Nas condições estabelecidas, o polígono 3D  $P'$  é o único a minimizar simultaneamente as funções de energia  $E^e$  e  $E^d$  e, conseqüentemente,  $E^e(P) + E^d(P)$ . Desse resultado se conclui que  $P'$  é a única solução da função de energia dada pela Equação 3.11.

### **3.2.3. Estratégia de solução da função de energia via Programação Dinâmica**

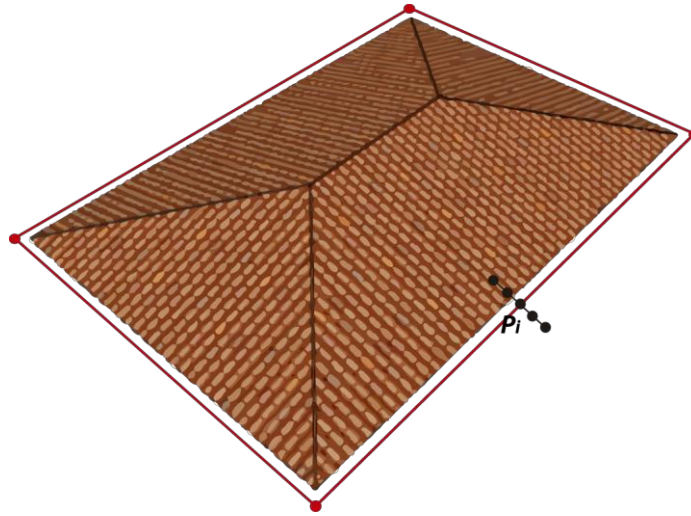
Embora os lados de telhados sejam retos, é previsto que inúmeros pontos dariam melhor suporte para definição dos lados na presença de ruídos e irregularidades geométricas. Conseqüentemente, o número de pontos no polígono 3D candidato a representar um dado contorno de telhado no espaço-objeto poderá ser bem maior que o mínimo necessário em situação ideal.

Com base nisso, dois tipos de espaços de busca são utilizados para criar o espaço poligonal de solução. No caso dos lados de contorno, que são assumidos como retos, é razoável buscar os melhores pontos na direção transversal aos respectivos lados. Com isso, são utilizados espaços de busca 1D que permitem amostrar polígonos 3D numa faixa ampla em torno do polígono 3D inicial, sem usar um número grande de pontos. No entanto, no caso das regiões de quina, é necessário utilizar espaços de busca 2D uma vez que não se sabe quais as direções que se deve seguir para encontrar as melhores quinas. A forma de definição desses dois espaços de busca é apresentada a seguir.

#### *3.2.3.1. Espaços de busca 1D*

Os espaços de busca 1D são criados nos pontos pertencentes aos lados do contorno do telhado, e são estabelecidos via amostragem de pontos ao longo de segmentos de retas simultaneamente transversais ao lado do contorno e ao vetor normal ao plano da face de telhado que contém o ponto em análise, gerando segmentos de retas que acompanham os planos do telhado (Figura 37).

Figura 37 – Criação do espaço de busca 1D no espaço 3D.



A criação dos pontos do espaço de busca 1D é realizada a partir da equação da reta suporte na forma paramétrica (Equação 3.12). Para tanto, basta tomar: 1) o vértice ( $P_i$ ) do polígono inicial onde a reta de suporte o interceptará; e 2) o vetor ( $\vec{v}$ ) transversal ao polígono inicial em  $P_i$  e que também seja ortogonal ao vetor normal do plano de face de telhado que contém esse mesmo vértice ( $P_i$ ).

$$\begin{cases} X = X_{P_i} \pm tv_1 \\ Y = Y_{P_i} \pm tv_2 \\ Z = Z_{P_i} \pm tv_3 \end{cases} \quad (3.12)$$

onde  $(X_{P_i}, Y_{P_i}, Z_{P_i})$  são as coordenadas do ponto  $P_i$  onde a reta intercepta o polígono inicial e  $(v_1, v_2, v_3)$  é o vetor simultaneamente transversal ao polígono inicial e ortogonal ao vetor normal do plano da face de telhado que contém  $P_i$ .

Assim, as coordenadas dos pontos do espaço de busca 1D passam a depender apenas do parâmetro  $t$  da reta que permitirá amostrar os pontos a distâncias preestabelecidas. Considerando o GSD da imagem, adotou-se convenientemente o valor de 1GSD para calcular as distâncias entre pontos amostrados nos espaços de busca 1D. O sinal  $\pm$  da Equação 3.12 permite amostrar pontos sobre o telhado e ao longo de seu prolongamento, tendo o polígono inicial como eixo de simetria.

### 3.2.3.2. Espaços de busca 2D

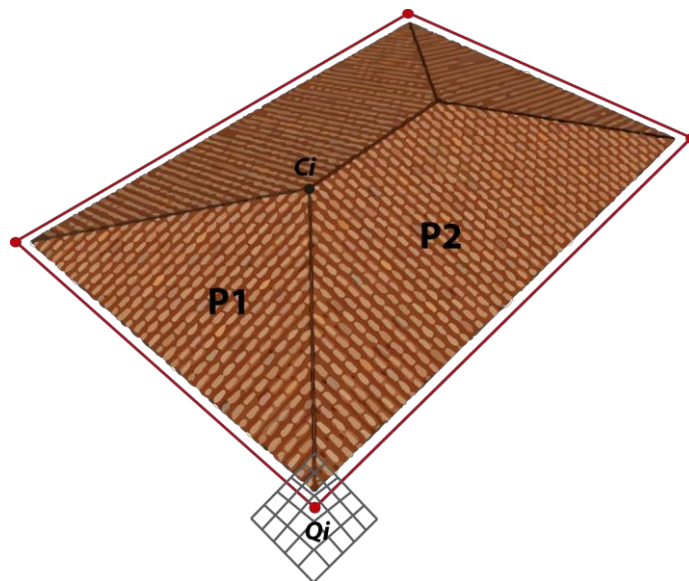
Os espaços de busca 2D são criados nos pontos de quina e são estabelecidos por uma malha regular onde os pontos são amostrados sobre os segmentos de planos que formam a quina em análise (Figura 38).

A Figura 38 ilustra como a malha regular é criada. Primeiro, as coordenadas  $X$  e  $Y$  dos pontos da malha são obtidas considerando: o ponto de quina  $Q_i$  do polígono inicial como o ponto central da malha, a resolução de 1GSD e o tamanho pré-estabelecido da malha. Em seguida é necessário verificar a qual face do telhado cada ponto da malha pertence, a fim de estabelecer a coordenada  $Z$  dos pontos. Para isso, os coeficientes da reta ( $a, b, c$ ) que passa pelo ponto de quina  $Q_i$  e seu respectivo ponto de cume  $C_i$  são calculados, e assim, a seguinte condição é utilizada para verificar se um dado ponto da malha pertence ao plano da face  $P1$  ou  $P2$  do telhado.

$$\begin{cases} \text{Se } aX + bY + c < 0 \rightarrow \text{ponto pertence ao plano } P1 \\ \text{Se } aX + bY + c > 0 \rightarrow \text{ponto pertence ao plano } P2 \end{cases}$$

Por fim, considerando que os coeficientes dos planos do telhado já foram previamente determinados e sabendo a qual plano o ponto em análise pertence, basta utilizar a equação do plano ( $Z = AX + BY + C$ ) para determinar a coordenada  $Z$  desse ponto.

Figura 38 – Criação do espaço de busca 2D no espaço 3D.



### 3.2.3.3. Adensamento dos espaços de busca

Pelo que foi apresentado nas Subseções 3.2.3.1 e 3.2.3.2, a Equação 3.11 poderia assumir a seguinte forma, considerando por simplicidade apenas dois lados sucessivos de contorno e uma quina na intersecção de ambos os lados (Figura 39):

$$\begin{aligned}
 E_T(P) = & \dots + E_i^T(P_{i-1}(X_{i-1}, Y_{i-1}, Z_{i-1}), P_i(X_i, Y_i, Z_i), P_{i+1}(X_{i+1}, Y_{i+1}, Z_{i+1})) + \dots \\
 & + E_q^T(P_f(X_f, Y_f, Z_f), P_q(X_q, Y_q, Z_q), P_{in}(X_{in}, Y_{in}, Z_{in})) + \dots + \\
 & + E_j^T(P_{j-1}(X_{j-1}, Y_{j-1}, Z_{j-1}), P_j(X_j, Y_j, Z_j), P_{j+1}(X_{j+1}, Y_{j+1}, Z_{j+1})) + \dots \quad (3.13)
 \end{aligned}$$

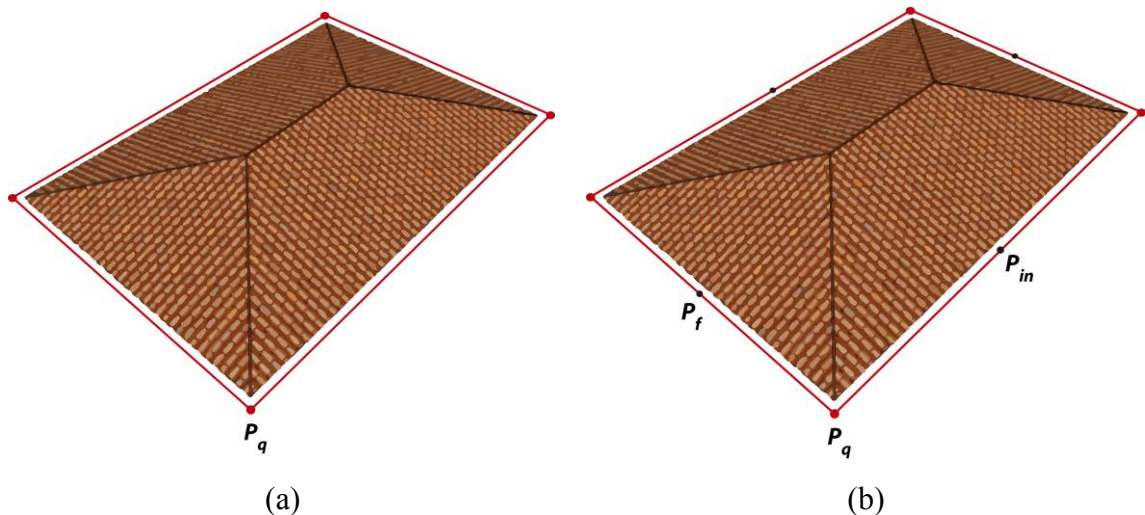
onde:  $E_i^T$  é o termo de energia para a  $i$ -ésima sequência de pontos  $P_{i-1}(X_{i-1}, Y_{i-1}, Z_{i-1}), P_i(X_i, Y_i, Z_i), P_{i+1}(X_{i+1}, Y_{i+1}, Z_{i+1})$  inter-relacionados simultaneamente;  $E_q^T$  é o termo de energia para a  $q$ -ésima quina, definida para o último ponto ( $P_f(X_f, Y_f, Z_f)$ ) do lado anterior à quina  $q$ , para o próprio ponto de quina ( $P_q(X_q, Y_q, Z_q)$ ) e para o ponto inicial ( $P_{in}(X_{in}, Y_{in}, Z_{in})$ ) do lado posterior à quina  $q$ ; e  $E_j^T$  é definido de forma análoga a  $E_i^T$ .

A Equação 3.13 será o alvo do processo de otimização por PD para encontrar o polígono 3D ótimo, entre os polígonos 3D candidatos que compõem o espaço de solução, formado pelos espaços de busca 1D e 2D. O espaço de solução é gerado a partir do polígono 3D inicial projetado na imagem. Entretanto, esta aproximação inicial para o contorno de telhado é refinada iterativamente, conforme segue:

- 1) O contorno aproximado do telhado, formado pelos pontos de quina obtidos através do modelo 3D do telhado, é utilizado como polígono 3D inicial (Figura 39(a)).
- 2) O polígono 3D é densificado via simples inserção de pontos médios entre seus pontos, mantendo os pontos já existentes (Figura 39(b)). Um limiar de separação entre os pontos inseridos é utilizado a fim de evitar uma densidade de pontos excessiva em lados de contornos pequenos e, assim, manter um padrão na densidade de pontos para lados pequenos e grandes. O valor estabelecido para esse limiar foi determinado empiricamente, e resultou em 1,5 m.
- 3) Esses pontos são então utilizados como referência para o estabelecimento dos espaços de busca 1D e 2D, conforme o procedimento apresentado nas Subseções 3.2.3.1 e 3.2.3.2.

- 4) O algoritmo de PD é aplicado considerando o espaço de solução criado, a fim de determinar o polígono 3D que melhor representa o contorno do telhado. Nessa etapa as imagens auxiliares de vegetação, sombra e obstrução perspectiva são utilizadas para verificar se um dado ponto está em alguma dessas regiões e, com isso, estabelecer o valor apropriado para cada parâmetro da função *Snake*. A forma de utilização de cada um dos parâmetros da função *Snake* é apresentada na Subseção 3.2.3.4.
- 5) Verificar a condição:  $\|P_i - P_{i-1}\| < L$ , sendo  $P_i$  e  $P_{i-1}$  os dois últimos polígonos 3D refinados e  $L$  um limiar preestabelecido. Se essa condição não for satisfeita, retornar para a etapa 2, caso contrário, o polígono ótimo  $P_i$  foi encontrado.

Figura 39 – Criação do espaço de solução no espaço 3D.  
(a) Contorno inicial. (b) Contorno densificado na primeira iteração.



#### 3.2.3.4. Determinação adaptativa dos parâmetros da função de energia

A função de energia (Equação 3.5) possui oito parâmetros para cada ponto  $i$ , sendo quatro ( $\alpha_i, \beta_i, \gamma_i$  e  $\eta_i$ ) para uma imagem e outros quatro ( $\alpha'_i, \beta'_i, \gamma'_i$  e  $\eta'_i$ ) para a outra imagem. Entretanto, em condições normais, sem a interferência de nenhuma anomalia, que pode afetar diferentemente ambas as imagens (normalmente causada por obstrução perspectiva), mesmo que pese o fato de ambas terem sido obtidas de perspectiva diferentes, as características geométricas e radiométricas não serão afetadas significativamente. Logo é razoável se ter  $\alpha_i = \alpha'_i, \beta_i = \beta'_i, \gamma_i = \gamma'_i$  e  $\eta_i = \eta'_i$ , sendo todos positivos ou eventualmente nulos. A seguir é apresentada a forma de adaptação que é adotada para cada um desses parâmetros durante o processo de otimização.

Parâmetro  $\alpha_i$ : como os lados de contorno são retos, adota-se um valor único e grande para  $\alpha_i$ , a fim de forçar que os segmentos de contorno para lados sejam pouco flexíveis para diminuir a influência de pontos de borda mais distantes de uma reta média, ou seja, forçando que os segmentos de contornos se aproximem de retas.

Parâmetro  $\beta_i$ : deverá possuir valor único, a menos em pontos de quina, onde deverá ser nula para possibilitar a formação de uma quina (conhecida como descontinuidade de segunda ordem). Vale destacar que os pontos de quina e de borda são previamente rotulados durante a etapa de criação dos espaços de busca.

Parâmetro  $\gamma_i$ : também deverá possuir valor único para pontos de borda e nulo para pontos de quina, uma vez que esse parâmetro está envolvido no termo da função de energia que modela a propriedade dos lados de contorno (Equação 3.2).

Parâmetro  $\eta_i$ : deve possuir valor único nas quinas e nulo nos pontos de borda, pois ele está presente no termo da função de energia que modela as propriedades das quinas (Equação 3.3).

Como apresentado na Seção 3.2.1, as imagens de vegetação e obstrução perspectiva permitem que o método identifique se um dado ponto está em alguma dessas regiões, e com isso, é possível adaptar os parâmetros da função *Snake* ( $\alpha, \beta, \gamma, \eta$ ) para tratar cada caso. As obstruções perspectivas sobre telhados podem ocorrer numa imagem e/ou em outra. Isso não afeta as forças internas da *Snake*, mas sim somente as forças externas. Assim  $\gamma$  e  $\eta$  devem ser anulados nos segmentos de telhado obstruídos, em uma e/ou em outra imagem, deixando apenas as forças internas ( $\alpha$  e  $\beta$ ) agirem. Portanto, se em uma imagem um determinado segmento do telhado não estiver obstruído, as forças externas associadas com essa imagem poderá ser usada para auxiliar o re-traçamento preciso do contorno.

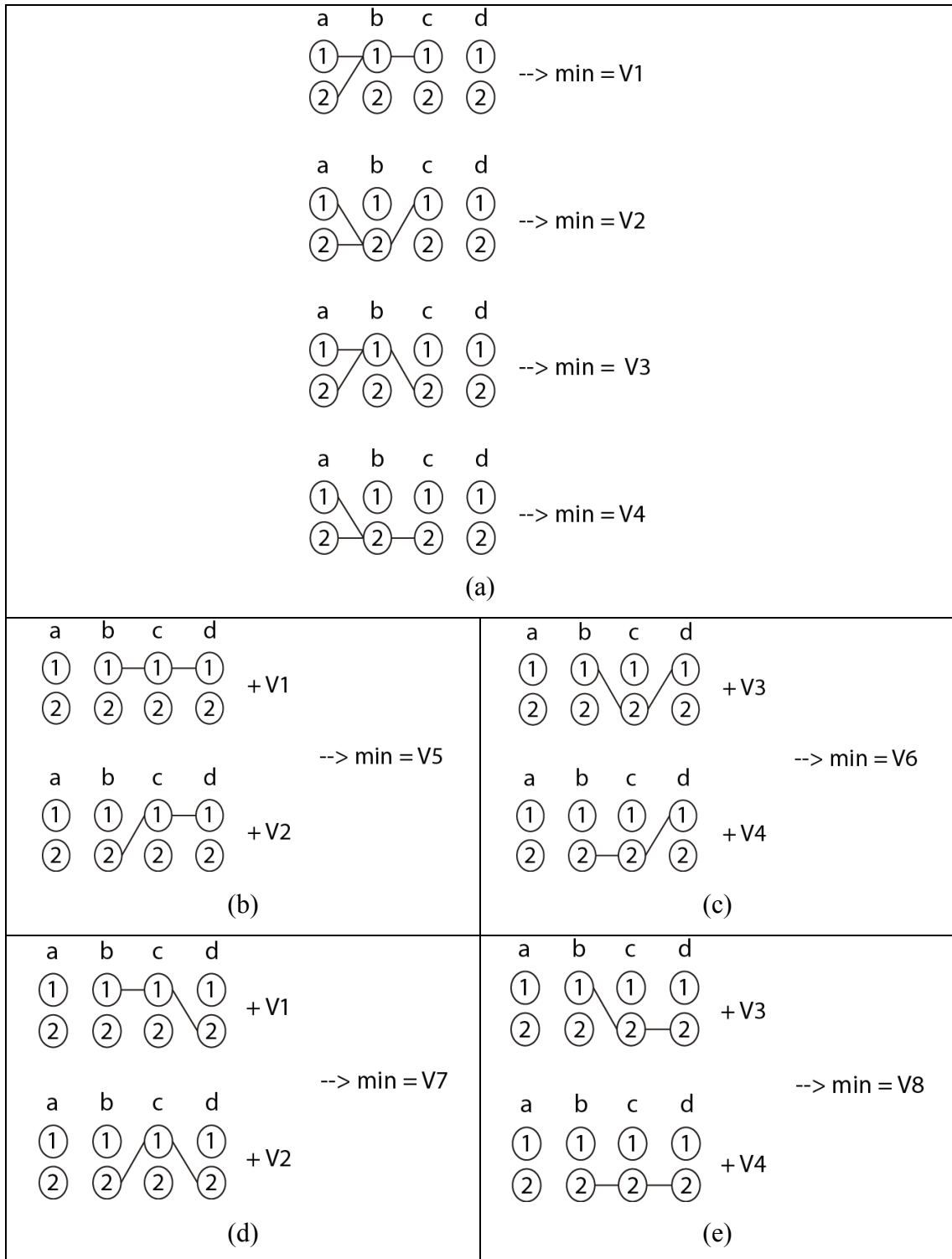
Em se tratando de vegetação obstruindo diretamente os telhados, estes aparecerão obstruídos em ambas as imagens. Consequentemente,  $\gamma$  e  $\eta$  devem ser anulados em ambos os segmentos de imagem obstruídos, deixando o controle apenas para forças internas. Finalizando, ambos casos de obstrução (perspectiva ou direta) podem ser neutralizados para evitar o incorreto traçamento do contorno 3D de telhado.

### 3.2.3.5. Aplicação da Programação Dinâmica

A Figura 40 ilustra um exemplo da aplicação da PD na obtenção da solução função de energia para determinação do polígono representando o contorno do telhado. Considerando 4 espaços de busca ( $a, b, c, d$ ) contendo 2 pontos cada um, os seguintes passos são realizados:

- 1) Calcular os custos da função de energia gerados pela combinação dos pontos  $a1-b1-c1$  e  $a2-b1-c1$ , determinar qual combinação fornece o menor custo e armazenar no vetor  $V1$ . Repetir essa etapa para as demais combinações conforme a Figura 40(a).
- 2) Calcular os custos da função de energia gerados pela combinação dos pontos  $b1-c1-d1$  e somar com o custo  $V1$ ; calcular os custos gerados pela combinação  $b2-c1-d1$  e somar com o custo  $V2$ . Determinar qual dessas combinações fornecem o menor custo e armazenar no vetor  $V5$  (Figura 40(b)). Repetir essa etapa para as demais combinações conforme a Figura 40(c), (d) e (e).
- 3) Por fim, determinar qual vetor ( $V5, V6, V7$  ou  $V8$ ) contém o menor custo e percorrer o caminho inverso para determinar a sequência de pontos que forneceram esses resultados, e conseqüentemente serão o polígono ótimo para representar o contorno do telhado.

Figura 40 – Aplicação da PD.



### 3.3. Forma de análise dos resultados

Os resultados obtidos são analisados visual e numericamente, por meio de comparações entre os contornos extraídos pelo método e os correspondentes contornos de



referência obtidos por restituição fotogramétrica utilizando o software LPS. A análise visual é realizada com base na visualização do contorno extraído e do correspondente contorno de referência sobrepostos na imagem aérea. A análise numérica é feita a partir da comparação numérica entre o contorno extraído e o correspondente contorno de referência, ambos no espaço-objeto, tendo por base os parâmetros de completeza, correção e a raiz do erro médio quadrático (RMSE).

O parâmetro de completeza está relacionado ao erro de omissão, e se refere à porcentagem da área de referência que foi extraída pelo método. Já o parâmetro de correção está relacionado ao erro de comissão, e se refere à porcentagem da área que foi corretamente extraída. A seguir são apresentadas as equações utilizadas para o cálculo da completeza e correção (FAZAN, 2014):

$$\begin{aligned} \text{completeza} &= \frac{A_{E \cap R}}{A_R} 100\% \\ \text{correção} &= \frac{A_{E \cap R}}{A_E} 100\% \end{aligned} \tag{3.14}$$

onde:

- $A_{E \cap R}$  é a área do polígono resultante da intersecção entre os polígonos que descrevem os contornos extraído e de referência;
- $A_R$  é a área do polígono correspondente ao contorno de referência;
- $A_E$  é a área do polígono correspondente ao contorno extraído.

O RMSE é calculado a partir da discrepância entre os vértices do polígono extraído e os correspondentes vértices do polígono de referência, através da seguinte equação:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n d_i^2} \tag{3.15}$$

onde  $d_i$  é a distância entre o vértice  $i$  do polígono de referência e o correspondente vértice no polígono extraído, e  $n$  é o número de vértices em ambos polígonos.

## **4. RESULTADOS**

### **4.1. Introdução**

Nesse capítulo é apresentada a avaliação experimental do método de extração de contornos de telhados no espaço-objeto desenvolvido neste trabalho. A Seção 4.2 apresenta as especificações dos parâmetros e limiares empregados na implementação do método. E a Seção 4.4 apresenta os resultados experimentais obtidos a partir do método proposto.

### **4.2. Parâmetros e limiares utilizados**

Esta seção apresenta os valores estabelecidos para os parâmetros requisitados pelo aplicativo *LAStools* utilizado na etapa de pré-processamento, bem como os parâmetros e limiares empregados na implementação do método proposto para a extração de contornos de telhados. Vale ressaltar que todos os valores foram estabelecidos empiricamente. A Tabela 8 apresenta os valores dos parâmetros e limiares utilizados no desenvolvimento deste trabalho.

Tabela 8 – Parâmetros e limiares utilizados.

<b>Etapa</b>	<b>Parâmetro/Limiar</b>	<b>Valor</b>
Pré-processamento	<i>step</i>	idades pequenas ou planícies
	<i>spike</i>	0,5 m
	<i>stddev</i>	10 cm
	<i>offset</i>	0,05 m
	<i>Ground offset</i>	2 m
	<i>Building planarity</i>	0,10 m
	<i>Forest ruggedness</i>	0,40 m
Extração dos contornos	$\alpha$	500000
	$\beta$	500000
	$\gamma$	1
	$\eta$	500000
	Dimensão do Espaço de Busca 1D	9 pontos
	Dimensão do Espaço de Busca 2D	9×9 pontos
	Limiar de separação mínima entre espaços de busca	1,5 m

### 4.3. Resultados e análise

Essa seção apresenta os resultados obtidos a partir do método desenvolvido neste trabalho. Vale ressaltar que os tipos de telhados considerados no desenvolvimento desse trabalho são aqueles com vértices conectados e quinas formadas por ângulo de 90°. Todas as áreas utilizadas nos experimentos são áreas residenciais da cidade de Presidente Prudente/SP.

Antes de realizar os experimentos para testar o desempenho do método proposto de extração de contornos de telhado, alguns testes envolvendo o algoritmo de PD foram realizados a fim de se testar a eficiência da aplicação do algoritmo.

Os experimentos para testar o desempenho do método proposto para a extração de contornos de telhado foram divididos em 3 grupos, os quais abrangem as situações previstas pelo método:

- Experimento 1: extração dos contornos de telhados que possuem partes oclusas causadas por vegetação que se eleva acima do telhado;

- Experimento 2: extração dos contornos de telhados onde ocorrem casos de sombras adjacentes causadas pelo próprio telhado;
- Experimento 3: extração dos contornos de telhados que possuem casos de obstrução perspectiva causada por edifícios vizinhos.

#### **4.3.1. Aplicação do algoritmo de PD**

Alguns testes para a obtenção da solução da função de energia (Equação 3.5) foram realizados com o objetivo de testar preliminarmente a eficiência do algoritmo de PD. A seguir são apresentados os detalhes de cada teste.

##### Teste 1

Nesse primeiro teste o algoritmo utilizado para a solução da função de energia foi o algoritmo de força bruta, no qual todas as combinações possíveis dos pontos dos espaços de busca eram calculadas e comparadas a fim de determinar qual a combinação fornecia o melhor resultado. Nesse teste as análises das informações contextuais da cena (obstruções e sombras) não foram consideradas e a função de energia dada pela Equação 3.5 não foi implementada de forma completa, sendo utilizado apenas:

$$E(p) = \sum_{i=1}^n [\alpha_i |p_{i+1} - p_i|^2 + \beta_i |p_{i-1} - 2p_i + p_{i+1}|^2 - \gamma_i |\nabla G(p_i)|^2 - \eta_i CS(p_i)^2] \quad (4.1)$$

onde  $\nabla G(p_i)$  é a magnitude do gradiente no ponto de borda e  $CS(p_i)$  é a resposta de quina calculada pelo detector de Moravec (1977).

##### Teste 2

No segundo teste também foi utilizado o algoritmo de força bruta, tendo por base a função de energia dada pela Equação 4.1; e não foram consideradas as análises das informações contextuais da cena. No entanto, com o intuito de otimizar a implementação da força bruta, utilizou-se 4 *threads*<sup>2</sup> para processar o algoritmo.

---

<sup>2</sup> Um processo pode ser dividido em duas ou mais tarefas, denominadas *threads*, que podem ser executadas concorrentemente a fim de melhorar a performance do processamento.

### Teste 3

No terceiro e último teste as análises das informações contextuais da cena passaram a ser consideradas, sendo que a função de energia dada pela Equação 3.5 foi implementada por completo e utilizou-se o algoritmo de PD para encontrar a solução da função de energia.

A Tabela 9 apresenta um resumo das informações de cada teste, e o resultado de tempo de processamento e número de combinações resultante de cada um.

Tabela 9 – Testes de comparação dos algoritmos.

	Teste 1	Teste 2	Teste 3
<b>Algoritmo</b>	Força bruta	Força bruta + <i>thread</i>	PD
<b>Quantidade de pontos no espaço de solução</b>	120	120	120
<b>Tempo de processamento</b>	40 min	15 min	6 seg
<b>Número de combinações</b>	$P^{EB}$	–	$P^3(EB - 2)$

onde  $EB$  é a quantidade de espaços de busca criados e  $P$  é a quantidade de pontos por espaço de busca.

A Tabela 9 mostra que o tempo de processamento melhorou de 40 min, no algoritmo de força bruta, para 6 seg com o algoritmo de PD. Essa melhoria pode ser explicada pela quantidade de combinações que são geradas por cada algoritmo.

A partir das equações apresentadas na Tabela 9 é possível observar que o número de combinações geradas pelo algoritmo de força bruta possui crescimento exponencial ( $P^{EB}$ ), já o número de combinações geradas pela PD apresenta crescimento linear ( $P^3(EB - 2)$ ), uma vez que o termo  $P^3$  pode ser considerado uma constante. Para ilustrar esse fato, a Figura 41 apresenta os gráficos do número total de combinações obtidas em cada estratégia a partir de diferentes quantidades de pontos no espaço de solução. De fato, é possível notar que o gráfico do algoritmo de força bruta apresenta comportamento exponencial, enquanto o gráfico do algoritmo de PD apresenta comportamento linear. A Tabela 10 traz os resultados numéricos desse exemplo.

Por meio desses testes foi possível concluir que a utilização do algoritmo de PD foi essencial para garantir a eficiência do programa implementado neste trabalho.

Figura 41 – Gráficos do número total de combinações obtidas em cada estratégia a partir de diferentes quantidades de pontos no espaço de solução.

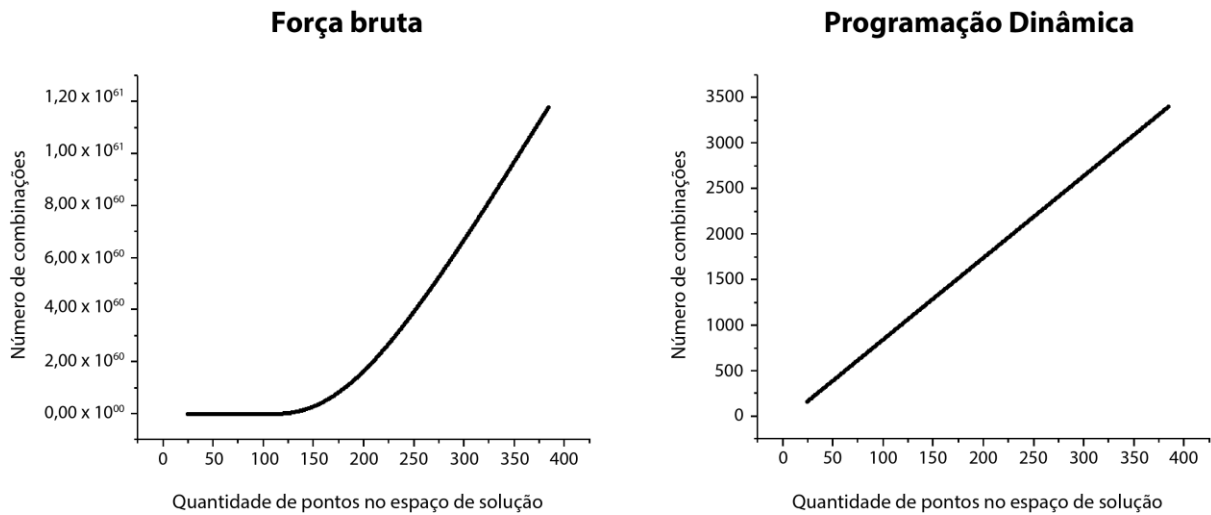


Tabela 10 – Comparação entre a quantidade de pontos no espaço de solução e número de combinações resultantes em cada estratégia.

EB	P	Pontos no espaço de solução	Número de combinações	
			Força bruta $P^{EB}$	PD $P^3(EB - 2)$
8	3	24	6561	162
16	3	48	43046721	378
32	3	96	$1,85302 \times 10^{15}$	810
64	3	192	$3,43368 \times 10^{30}$	1674
128	3	348	$1,17902 \times 10^{61}$	3402

#### 4.3.2. Experimento 1 – Vegetação

O primeiro experimento foi realizado com o objetivo de testar o desempenho do método na extração dos contornos de telhados em casos que ocorrem oclusões diretas causadas por árvores. Para isso foram selecionadas quatro áreas compostas por um total de sete telhados. A Figura 42(a) apresenta a imagem aérea desses telhados, bem como sua rotulação. A Figura 42(b) mostra os contornos de referência (vermelho) utilizados na análise de qualidade dos resultados.

Nos Telhados 1 e 2 a oclusão causada pela vegetação ocorre nas partes laterais dos telhados. Nos Telhados 4 e 5 a vegetação cobre regiões próximas à quina, no entanto apenas um dos lados do telhado que forma a quina é obstruído. Nos Telhados 3, 6 e 7 a oclusão também ocorre em regiões de quina, mas nesses casos a vegetação cobre tanto as quinas quanto os lados que formam essa quina.

Figura 42 – Áreas utilizadas no Experimento 1 contendo telhados cobertos por vegetação.

(a) Imagem aérea das áreas utilizadas; (b) Contornos de referência (vermelho).



(a)



(b)

A Figura 43(a) e (b) apresenta os contornos iniciais (azul) e os contornos extraídos pelo método (verde), respectivamente. É possível notar que os contornos das partes oclusas pela vegetação foram extraídos com sucesso. Isso ocorreu pois mesmo não havendo boas respostas radiométricas para modelar o contorno nessas áreas, as características contidas na função de energia desenvolvida conseguiram definir o contorno correto.

No caso dos Telhados 1 e 2 cuja oclusão aparece em uma parte dos lados dos telhados, a extração correta do contorno foi favorecida pelo termo 1 da Equação 3.5, uma vez que atribuindo um valor alto para o parâmetro  $\alpha$  é possível forçar que o segmento de contorno para aquele lado seja pouco flexível, se aproximando ao máximo a uma reta.

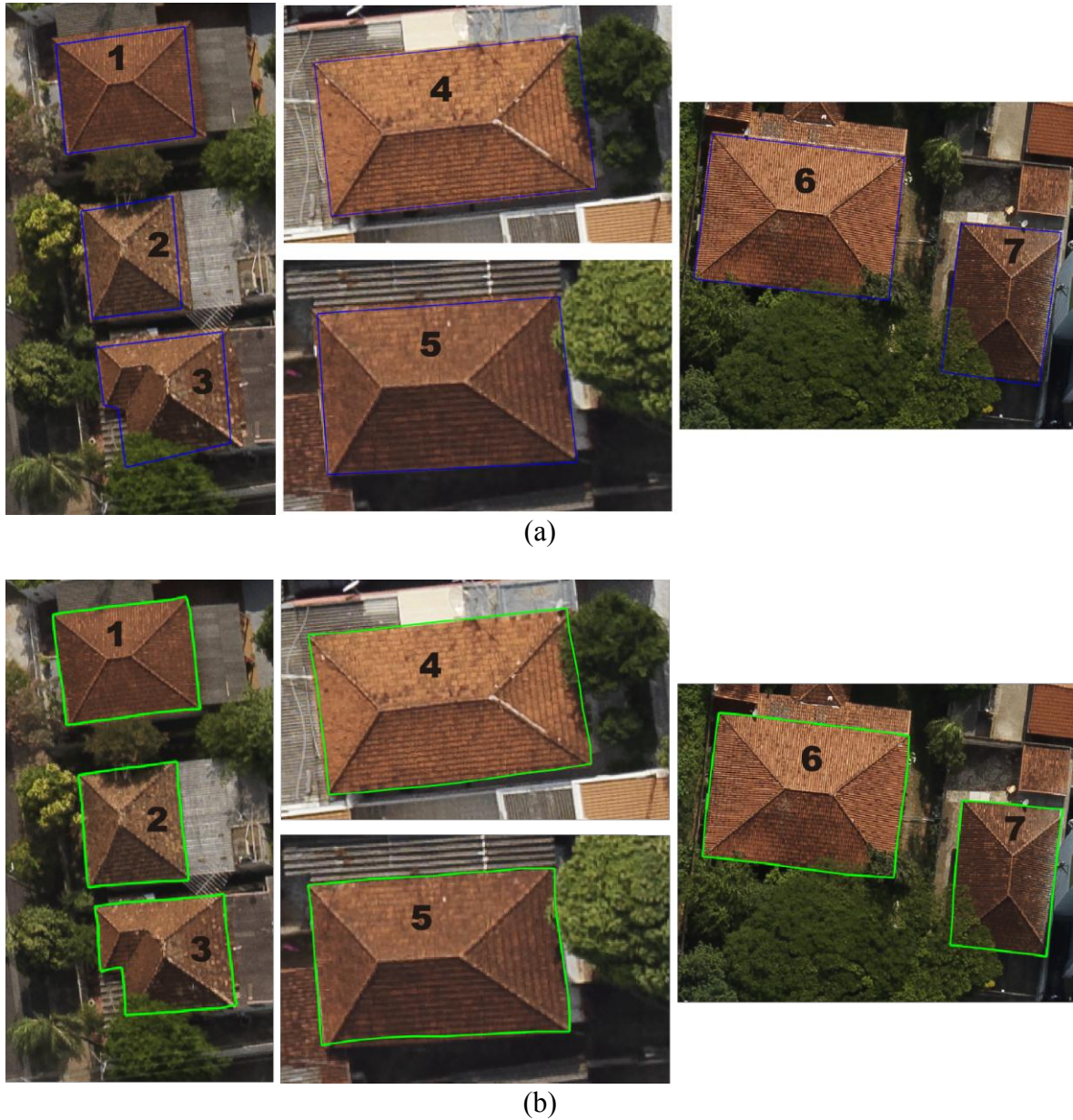
Nos Telhados 4 e 5 a oclusão ocorre próximo à quina, no entanto apenas um dos lados do telhado que forma a quina está ocluso. Com isso, a resposta radiométrica da reta suporte formada no lado não ocluso pode ser usada no termo 4 da Equação 3.5 para auxiliar na extração correta do contorno naquela área.

Nos Telhados 3, 6 e 7 a oclusão ocorre em regiões de quina, obstruindo tanto as quinas quanto os lados que formam essa quina. Nesse caso, a componente angular do termo 4 da Equação 3.5 favoreceu para que fossem escolhidas as quinas que formassem ângulos retos.



Figura 43 – Resultado visual do Experimento 1.

(a) Contornos iniciais (azul); e (b) Contornos extraídos pelo método proposto (verde).



A Tabela 11 e a Tabela 12 apresentam os resultados dos parâmetros de completude, correção e RMSE para os contornos iniciais e extraídos, respectivamente. A partir da análise das médias apresentadas nas duas tabelas é possível observar que o parâmetro de completude dos contornos extraídos (96,88%) foi maior que o parâmetro de completude dos contornos iniciais (90,87%), indicando que houve uma melhora dos contornos ao se aplicar o método proposto.

Já o parâmetro de correção teve um resultado pior para os contornos extraídos. O parâmetro de correção avalia o quanto da área analisada está correta em relação à área de

referência. Como em quase todos os telhados a área do polígono inicial era menor do que a área do polígono de referência, maior foi a chance do polígono inicial estar contido dentro do polígono de referência e, assim, maior foi a chance do polígono inicial estar correto e fornecer resultados melhores de correção.

Os contornos extraídos apresentaram valores médios de  $RMSE_{E,N}$  em torno de 0,20 m, já os contornos iniciais tiveram valores médios em torno de 0,30 m. O  $RMSE_H$  também melhorou de 0,80 m para 0,70 m. Isso mostra que o método refinou com sucesso os contornos iniciais dos telhados.

Tabela 11 – Parâmetros de qualidade dos contornos iniciais do Experimento 1.

<b>Telhado</b>	<b>Completeza (%)</b>	<b>Correção (%)</b>	<b><math>RMSE_E</math> (m)</b>	<b><math>RMSE_N</math> (m)</b>	<b><math>RMSE_H</math> (m)</b>
1	88,76	99,67	0,27	0,35	0,75
2	83,65	100,00	0,29	0,39	0,69
3	89,69	97,02	0,41	0,27	0,60
4	92,35	96,99	0,33	0,46	1,25
5	89,14	98,95	0,36	0,28	1,13
6	96,25	98,86	0,21	0,21	0,38
7	96,25	93,31	0,35	0,30	0,95
Média	90,87	97,83	0,32	0,32	0,82

Tabela 12 – Parâmetros de qualidade dos contornos extraídos pelo Experimento 1.

<b>Telhado</b>	<b>Completeza (%)</b>	<b>Correção (%)</b>	<b><math>RMSE_E</math> (m)</b>	<b><math>RMSE_N</math> (m)</b>	<b><math>RMSE_H</math> (m)</b>
1	99,23	92,18	0,30	0,22	0,61
2	94,51	96,49	0,23	0,16	0,54
3	95,86	95,68	0,22	0,28	0,52
4	97,42	94,69	0,27	0,25	1,14
5	96,66	93,24	0,34	0,21	0,94
6	98,29	98,19	0,11	0,23	0,34
7	96,20	98,39	0,22	0,18	0,91
Média	96,88	95,55	0,24	0,22	0,71

### 4.3.3. Experimento 2 – Sombra

O segundo experimento teve como objetivo avaliar o método proposto aplicado na extração de contornos de telhados contendo sombras adjacentes e alongadas, as quais poderiam ser confundidas com as bordas dos telhados. Para esse experimento foi selecionada uma área contendo 5 telhados com sombras adjacentes. A Figura 44(a) apresenta a imagem aérea dos telhados utilizados nesse experimento, e sua rotulação. A Figura 44(b) apresenta os contornos de referência (vermelho) utilizados na análise de qualidade dos resultados.

Figura 44 – Área utilizada no Experimento 2 contendo sombras adjacente aos telhados.  
(a) Imagem aérea das áreas utilizadas; (b) Contornos de referência (vermelho).



Os contornos iniciais (azul) e os contornos extraídos (verde) são apresentados na Figura 45(a) e (b), respectivamente. Pode-se observar que os contornos dos telhados foram extraídos com sucesso mesmo nas regiões com a presença de sombras adjacentes.

Teoricamente as bordas de sombras produziriam energias maiores pois as respostas do gradiente nesses locais são mais altas, se comparadas às respostas das bordas dos telhados. No entanto, o fator de compensação de sombra introduzido no termo 3 da Equação 3.4 conseguiu analisar quais pontos pertenciam a borda de telhado e quais pertenciam a borda de sombra, e assim, favorecer os pontos de borda de telhado resultando na escolha correta do contorno. O Telhado 5, além da sombra adjacente, também apresenta obstrução causada por vegetação, e como já apresentado no Experimento 1, o método trabalha corretamente na presença dessas feições.

Em alguns casos, como por exemplo no Telhado 1, é possível notar que há um pequeno desvio na detecção de algumas partes da borda. A Figura 45(c) apresenta em detalhe essa falha. Isso ocorreu devido a feições presentes no entorno dos telhados, o que acabou confundindo e prejudicando a detecção correta da borda. Esse resultado mostra que a realização de um pré-processamento da imagem, como por exemplo a segmentação, poderia minimizar esse tipo de problema.

Figura 45 – Resultado visual do Experimento 2.

- (a) Contornos iniciais (azul); (b) Contornos extraídos pelo método proposto (verde); e  
(c) Falha na detecção da borda devido às feições presentes no entorno do telhado.



A Tabela 13 e a Tabela 14 trazem os resultados dos parâmetros de completeza, correção e RMSE dos contornos iniciais e extraídos, respectivamente. Os parâmetros de completeza dos contornos extraídos (99,32%) resultaram em valores médios melhores que dos contornos iniciais (90,05%), mostrando que o método refinou corretamente os contornos dos telhados. Já a média dos parâmetros de correção teve valor pior para os contornos extraídos, pelo mesmo motivo apresentado no Experimento 1.

Em relação ao RMSE, os valores médios das componentes E e N foram aproximadamente 0,30 m para os contornos iniciais e 0,10 m para os contornos extraídos. Para a componente H o valor médio do RMSE foi aproximadamente 0,50 m para os contornos iniciais e 0,35 m para os contornos extraídos. Isso mostra, mais uma vez, que o método conseguiu refinar com sucesso os contornos iniciais dos telhados.

Tabela 13 – Parâmetros de qualidade dos contornos iniciais do Experimento 2.

<b>Telhado</b>	<b>Completeza (%)</b>	<b>Correção (%)</b>	<b>RMSE<sub>E</sub> (m)</b>	<b>RMSE<sub>N</sub> (m)</b>	<b>RMSE<sub>H</sub> (m)</b>
1	92,50	99,97	0,23	0,29	0,51
2	89,73	100,00	0,34	0,30	0,35
3	93,28	100,00	0,22	0,26	0,64
4	87,08	100,00	0,37	0,35	0,40
5	87,67	100,00	0,28	0,35	0,48
Média	90,05	99,99	0,29	0,31	0,48

Tabela 14 – Parâmetros de qualidade dos contornos extraídos pelo Experimento 2.

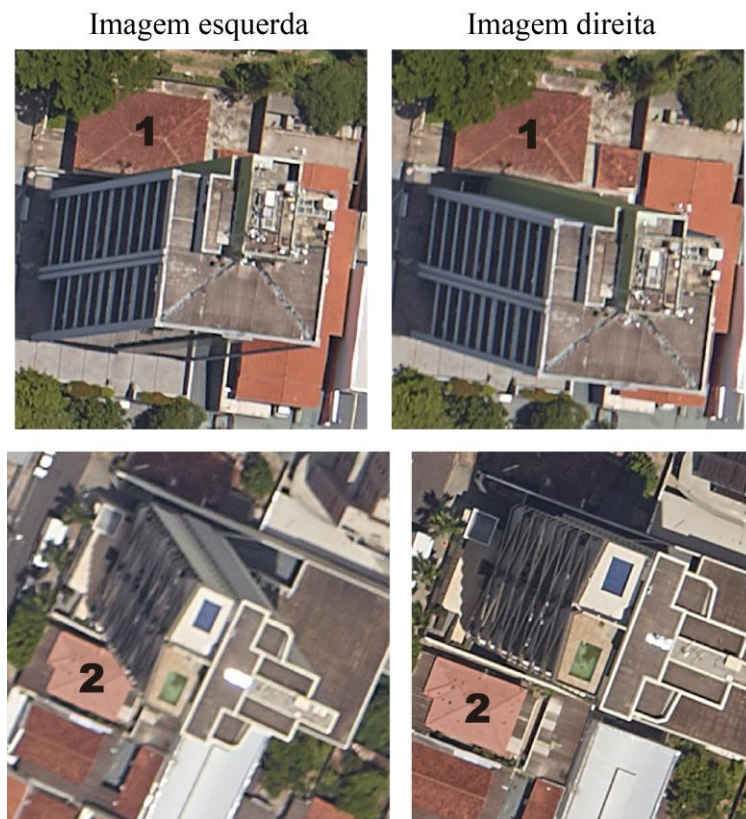
<b>Telhado</b>	<b>Completeza (%)</b>	<b>Correção (%)</b>	<b>RMSE<sub>E</sub> (m)</b>	<b>RMSE<sub>N</sub> (m)</b>	<b>RMSE<sub>H</sub> (m)</b>
1	99,91	97,37	0,13	0,09	0,43
2	99,75	98,75	0,04	0,12	0,22
3	99,72	97,04	0,15	0,12	0,53
4	98,13	97,70	0,12	0,10	0,28
5	99,09	96,92	0,16	0,14	0,36
Média	99,32	97,56	0,12	0,11	0,36

#### **4.3.4. Experimento 3 – Obstrução perspectiva**

O terceiro experimento teve como objetivo avaliar o desempenho do método proposto em casos onde ocorrem obstruções perspectivas. Para explorar a vantagem do uso do estéreo par de imagens, visou-se escolher situações onde a obstrução causada por edifícios elevados aparecesse em apenas uma das imagens. Embora a presença de edifícios elevados em ambientes urbanos seja muito comum, a situação considerada pelo experimento não é tão fácil de ser encontrada, pois ela depende de diversos fatores como altura e direção do voo, altura da edificação que causa a oclusão, posição do CP, relevo, etc.

Com isso, apenas duas áreas puderam ser utilizadas para a realização do terceiro experimento. A Figura 46 apresenta as imagens do estéreo par dos telhados utilizados nesse experimento e a sua rotulação.

Figura 46 – Áreas utilizadas no Experimento 3 contendo obstruções perspectivas.



A Figura 47(a), (b) e (c) apresenta os contornos de referência (vermelho), os contornos iniciais (azul) e os contornos extraídos pelo método proposto (verde), respectivamente. Pode-se notar que o método conseguiu delinear corretamente as bordas nas

regiões obstruídas. Isso foi possível pois nessas regiões a resposta da imagem não obstruída pôde ser utilizada no cálculo da função de energia, e assim, fornecer um resultado correto.

No Telhado 1, além da obstrução perspectiva, há também a presença de obstrução causada por vegetação e, como já apresentado no Experimento 1, o método é eficiente na presença dessa situação. No entanto, como é possível notar pela Figura 48, há uma mancha branca próximo à borda que faz divisa com o edifício elevado, o que atrapalhou o delineamento correto da borda fazendo com que o método a confundisse com a borda verdadeira (linha rosa tracejada). Esse resultado mostra que o método é sensível a ruídos presentes nas proximidades das bordas, tornando interessante a realização de um pré-processamento da imagem para a eliminação desses ruídos antes da aplicação do método de extração de contornos.



Figura 47 – Resultado visual do Experimento 3.

(a) Contornos de referência (vermelho); (b) Contornos iniciais (azul); e (c) Contornos extraídos pelo método proposto (verde).

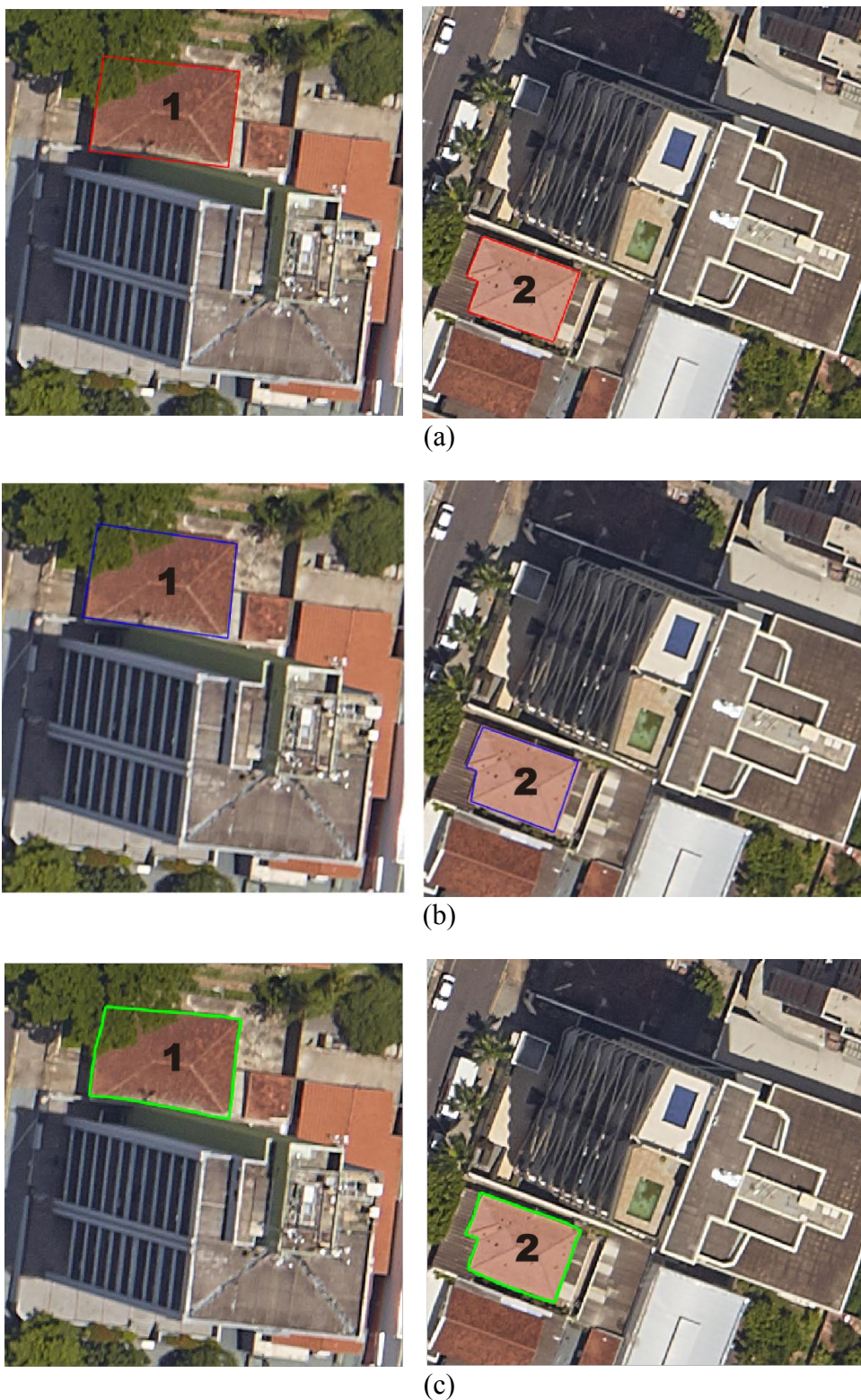


Figura 48 – Detalhe da mancha branca no telhado que atrapalhou o delineamento correto da borda, contorno detectado (verde) e contorno correto (rosa tracejado).



A Tabela 15 e a Tabela 16 apresentam os resultados quantitativos dos contornos iniciais e extraídos, respectivamente. Analisando os valores médios obtidos para cada parâmetro é possível notar que, novamente, o parâmetro de completeza dos contornos extraídos (96,16%) teve um valor melhor que dos contornos iniciais (93,26%), e o parâmetro de correção teve um valor pior para os contornos extraídos. No entanto, ambos ficaram com valores acima de 95%.

Os valores de RMSE dos contornos extraídos foram melhores para as três componentes. Para as componentes E e N o valor do RMSE ficou abaixo de 0,50 m, no entanto, a componente H resultou em um RMSE acima de 1 m.

Tabela 15 – Parâmetros de qualidade dos contornos iniciais do Experimento 3.

<b>Telhado</b>	<b>Completeza (%)</b>	<b>Correção (%)</b>	<b>RMSE<sub>E</sub> (m)</b>	<b>RMSE<sub>N</sub> (m)</b>	<b>RMSE<sub>H</sub> (m)</b>
1	93,50	95,50	0,42	0,33	0,93
2	93,01	97,73	0,28	0,29	1,46
Média	93,26	96,62	0,35	0,31	1,20

Tabela 16 – Parâmetros de qualidade dos contornos extraídos pelo Experimento 3.

<b>Telhado</b>	<b>Completeza (%)</b>	<b>Correção (%)</b>	<b>RMSE<sub>E</sub> (m)</b>	<b>RMSE<sub>N</sub> (m)</b>	<b>RMSE<sub>H</sub> (m)</b>
1	94,46	97,96	0,31	0,28	0,91
2	97,86	94,83	0,29	0,27	1,34
Média	96,16	96,40	0,30	0,28	1,13

## **5. CONCLUSÕES E RECOMENDAÇÕES**

Neste trabalho foi desenvolvida uma metodologia para a extração de contornos de telhados de edifícios diretamente no espaço-objeto, integrando um estéreo par de imagens aéreas e modelos 3D aproximados de telhados extraídos a partir de dados de varredura a LASER. As bordas de telhados não são bem definidas em dados de VLA com baixa densidade, o que torna interessante a integração com os dados de imagem para se ter uma melhoria posicional dessas bordas, uma vez que em imagens ópticas as bordas de telhados são bem definidas. Por outro lado, a utilização dos modelos 3D auxiliam no fornecimento das informações altimétricas sobre os pontos das bordas de telhado. Assim, a integração desses dois tipos de dados proposta por este trabalho mostrou ser uma solução para se obter os contornos de telhado diretamente no espaço-objeto.

Esse trabalho também propôs solucionar problemas comuns quando se trata de extração de contornos de telhados em ambientes urbanos, sendo eles: obstrução causada por vegetação que se eleva acima dos telhados; sombras adjacentes aos telhados, as quais podem ser confundidas com as bordas dos telhados; e obstrução perspectiva dos telhados causadas por edifícios vizinhos elevados. Para contornar essas situações, a função de energia *Snake* (Equação 3.5) foi criada incorporando termos que possibilitam a detecção correta do contorno de telhado, mesmo na presença dos problemas citados. Além disso, a função de energia *Snake* criada apresenta uma nova abordagem para o cálculo da resposta de quina, combinando as informações radiométricas e as informações geométricas dos pontos candidatos à quina.

No método proposto, o espaço de solução criado a partir dos espaços de busca 1D e 2D é densificado iterativamente, o que cria uma grande quantidade de possíveis combinações de pontos para a solução da função de energia *Snake*. Dessa forma, a realização da otimização da função de energia via PD foi fundamental para garantir a eficiência do algoritmo implementado, uma vez que o número de combinações geradas pelo algoritmo possui crescimento linear.

A avaliação experimental foi realizada de modo a testar o desempenho do método nas três situações descritas anteriormente. O Experimento 1 teve como objetivo avaliar o método na presença de obstruções causadas por vegetação elevada. Em todos os casos apresentados nesse experimento o método conseguiu delinear o contorno corretamente, mesmo nas regiões obstruídas. Isso ocorreu pois, na presença de obstrução, os termos da função de energia relacionados às informações radiométricas não são considerados, já que não

existem informações radiométricas do telhado nessas regiões. Assim, a resposta da função de energia passa a depender apenas dos termos relacionados às informações geométricas do telhado, as quais conseguem assegurar a resposta correta do contorno.

No Experimento 2, foram testadas as áreas que possuíam telhados com sombras adjacentes a eles. Nesses casos o método também conseguiu identificar corretamente os contornos dos telhados, mesmo as bordas de sombra tendo respostas de gradiente maiores que as bordas de telhado. Esse resultado mostra que o fator de compensação de sombra (Equação 3.2) adicionado na função de energia *Snake* trabalha da maneira esperada, favorecendo as bordas de telhado.

O Experimento 3 foi realizado a fim de testar o método em casos de obstruções perspectivas. Para explorar a vantagem do uso do estéreo par, foram selecionadas apenas áreas com telhados que apareciam obstruídos em apenas uma das imagens do estéreo par. Novamente, os resultados obtidos mostraram que o método conseguiu delinear corretamente as bordas dos telhados. Isso foi possível pois a função de energia era desligada nas regiões obstruídas, ficando a cargo apenas da imagem não obstruída fornecer os dados para o cálculo da resposta.

Os experimentos mostraram que, em alguns casos, o método é sensível a ruídos. A detecção da borda foi afetada por objetos e feições presentes no entorno dos telhados. Isso mostra que um pré-processamento da imagem a partir de um bom método de segmentação auxiliaria na melhoria dos resultados.

Por fim, considerando a média dos valores de RMSE calculados nos três experimentos, os polígonos refinados pelo método proposto apresentou melhores resultados se comparados com os polígonos iniciais obtidos dos modelos 3D. O parâmetro de completude também apresentou melhores resultados para os polígonos refinados pelo método, no entanto, o parâmetro de correção apresentou resultados inferiores.

No intuito de melhorar o desempenho do método proposto, algumas recomendações são apresentadas a seguir:

- Adaptar o algoritmo para tratar outros tipos de telhados, não somente os estilos de telhados utilizados neste trabalho;
- Realizar um pré-processamento da imagem, a fim de minimizar as influências de feições presentes ao redor dos telhados durante o cálculo da função de energia;
- Gerar as imagens de vegetação utilizando as próprias imagens ópticas, pois ao se utilizar dados de VLA deve-se tomar o cuidado de escolher dados de VLA com

data próxima à data das imagens. Caso contrário, é provável que muitas feições presentes nas imagens não sejam as mesmas presentes nos dados de VLA;

- Realizar experimentos utilizando imagens e dados de VLA com resoluções/densidades variadas, a fim de verificar de forma mais ampla os limites da aplicação do método proposto;
- Estender o método proposto para tratar dados de imagem multisensor, envolvendo, por exemplo, imagens de plataforma VANT e aérea convencional.

## REFERÊNCIAS

- AHO, A. V.; HOPCROFT, J. E.; ULLMAN, J. D. **The design and analysis of computer algorithms.** 1974.
- AMHAR, F.; JANSÁ, J.; RIES, C., **The generation of true orthophotos using a 3D building model in conjunction with a conventional DTM.** In: The International Archives of Photogrammetry and Remote Sensing, v. 32, p.16-22. 1998.
- ANDRADE, J.B. **Fotogrametria.** Curitiba: SBEE, 1998.
- AXELSSON, P. **Processing of laser scanner data: algorithms and applications.** ISPRS Journal of Photogrammetry and Remote Sensing, v. 54 (2-3), p. 138-147, 1999.
- AZEVEDO, S. C.; SILVA, E. A.; PEDROSA, M. M. **Shadow detection improvement using spectral indices and morphological operators in urban areas in high resolution images.** The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, v. XL-7/W3, p. 587-592, 2015.
- BALLARD, D.; BROWN, C. M. **Computer vision.** New Jersey: Prentice Hall, Inc., Englewood Cliffs, 1982.
- BALTSAVIAS, E. P. **Airborne LASER Scanning: basic relation and formulas.** In: ISPRS Journal of Photogrammetry and Remote Sensing, v. 54, p. 199-214, 1999.
- BRASSARD, G; BRADLEY, P. **Fundamentals of algorithmics.** New Jersey: Prentice Hall, 1996.
- BENCIOLINI, B.; RUGGIERO, V.; VITTI, A.; ZANETTI, M. **Roof planes detection via a second-order variational model.** ISPRS Journal of Photogrammetry and Remote Sensing, v. 138, p. 101-120, 2018.
- BROVELLI, M. A.; CANNATA, M.; LONGONI, U. M. **Managing and processing LIDAR data within GRASS.** In: GRASS Users Conference, Proceedings... Trento, Italia, 2002.
- CHEN, D.; et al. **A methodology for automated segmentation and reconstruction of urban 3-D buildings from ALS point clouds .** IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, v. 7, n. 10, p. 4199-4217, 2014.
- CHEN, J. et al. **Information from imagery: ISPRS scientific vision and research agenda.** ISPRS Journal of Photogrammetry and Remote Sensing, v. 115, p. 3-21, 2016a.
- CHEN, Q.; WANG, S.; LIU, X. **An improved Snake model for refinement of LiDAR-derived building roof contours using aerial images.** The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, v. XLI-B3, p. 583-589, 2016b.

COTE, M.; SAEEDI, P. **Automatic rooftop extraction in nadir aerial imagery of suburban regions using corners and variational level set evolution.** IEEE Transactions on Geoscience and Remote Sensing, v. 51, n. 1, p. 313-328, 2013.

DAL POZ, A. P., VALE, G. M. **Dynamic programming approach for semi-automated road extraction from medium- and high-resolution images.** In: Ebner, H.; Heipke, C.; Mayer, H.; Pakzad, K. (Orgs.). Hannover, Germany, 34: 87-92, 2003.

DAL POZ, A. P.; FAZAN, A. J.; CORREIA, L. S.; MARCATO, V. J. **Automated building extraction: comparison of paradigms and examples.** VI International Symposium on Mobile Mapping Technology, Presidente Prudente, 2009.

DAL POZ, A. P. **Filtragem de dados de varredura a LASER: princípios e exemplos.** XVI Simpósio Brasileiro de Sensoriamento Remoto, p. 6057-6064, Foz do Iguaçu, 2013.

ELMQVIST, M. **Ground estimation of LASER RADAR data using active shape models.** In: OEEPE workshop on Airborne Laser scanning and Interferometric SAR for Detailed Digital Elevation Models, Proceedings... 2001.

EL-SHEIMY, N.; NIU, X. **The promise of MEMS to the navigation community.** Inside GNSS, v. 2, p. 46-56, 2007.

FACON, J. **Morfologia Matemática: teoria e exemplos.** Curitiba: PUC, 1996.

FAZAN, A. J.; DAL POZ, A. P. **Predição de sombras de edifícios sobre vias urbanas com base em modelos digitais de elevação e dados de imagem aérea de alta resolução.** Boletim de Ciências Geodésicas, v. 14, p. 350-369, 2008.

FAZAN, A. J.; DAL POZ, A. P. **Rectilinear building roof contour extraction based on snakes and dynamic programming.** International Journal of Applied Earth Observation and Geoinformation, v. 25, p. 1-10, 2013.

FAZAN, A. J. **Extração de contornos de telhados de edifícios no espaço-objeto baseada em snakes e programação dinâmica a partir de TIN/LASER e imagens aéreas de alta resolução.** Tese (doutorado em Ciências Cartográficas) – Faculdade de Ciências e Tecnologia, Universidade Estadual Paulista, Presidente Prudente, 2014.

GALLIS, R. B. A. **Extração semi-automática de malha viária em imagens aéreas digitais de áreas rurais utilizando otimização por programação dinâmica no espaço objeto.** Tese (doutorado em Ciências Cartográficas) – Faculdade de Ciências e Tecnologia, Universidade Estadual Paulista, Presidente Prudente, 2006.

GALO, M.; TOZZI, C. L. **Extração de pontos com acurácia sub-pixel em imagens digitais.** Série em Ciências Geodésicas – Pesquisa em Ciências Geodésicas, v. 2, p. 289-313, 2002.

GERKE, M.; XIAO, J. **Fusion of airborne laserscanning point clouds and images for supervised and unsupervised scene classification.** ISPRS Journal of Photogrammetry and Remote Sensing, v. 87, p. 78-92, 2014.

GILANI, S. A. N.; AWRANGJEB, M.; LU, G. **Fusion of LiDAR data and multispectral imagery for effective building detection based on graph and connected component analysis.** The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Munich, Germany, Vol. XL, Part 3/W2, pp. 65-72, 2015.

GONZALEZ, R. C.; WOODS, R. E. **Processamento de imagens digitais.** 3ª ed. São Paulo: Edgard Blücher, 2000.

GRIFFITHS, D.; BOEHM, J. **Improving public data for building segmentation from Convolutional Neural Networks (CNNs) for fused airborne LIDAR and image data using active contours.** ISPRS Journal of Photogrammetry and Remote Sensing, v. 154, p. 70–83, 2019.

HABIB, A. F.; KIM, E.; KIM, C. **New Methodologies for True Orthophoto Generation.** Photogrammetric Engineering & Remote Sensing. v.73, n. 1, p. 25-36. 2007.

HABIB, A. F. **Integration of lidar and photogrammetric data: triangulation and orthorectification.** In: SHAN, J.; TOTH, C. K. Topographic laser ranging and scanning - principles and processing. CRC Press. 2008.

HUANG, J.; ZHANG, X.; XIN, Q.; SUN, Y.; ZHANG, P. **Automatic building extraction from high-resolution aerial images and LiDAR data using gated residual refinement network.** ISPRS Journal of Photogrammetry and Remote Sensing, v. 151, p. 91-105, 2019.

JAIN, R.; KASTURI, R.; SCHUNCK, B. G. **Machine vision.** McGraw-Hill, 1995.

JAZAYERI, I.; FRASER, C. S. **Interest operator for feature-based matching in close range photogrammetry.** The Photogrammetric Record, v. 25(129), p. 24-41, 2010.

JENSEN, J. R. **Sensoriamento Remoto do Ambiente: Uma Perspectiva em Recursos Terrestres.** Editora Parêntese, ed. 2, 2009.

KASS, M.; WITKIN, A.; TERZOPOULOUS, D. **Snakes: Active contour models.** International Journal of Computer Vision, v. 1, n. 4, p. 321-331, 1988.

KRAUS, K. **Photogrammetry: fundamentals and standard processes.** Bonn: Dümmler, 1993.

LARI, Z.; EL-SHEIMY, N.; HABIB, A. **A new approach for realistic 3D reconstruction of planar surfaces from laser scanning data and imagery collected onboard modern low-cost aerial mapping systems.** Remote Sensing, 2017.

LI, H. **Semi-automatic road extraction from satellite and aerial images.** Tese (Technical Sciences), Swiss Federal Institute of Technology Zurich, 1997.

LI, Y.; SASAGAWA, T.; GONG, P. **A system of the shadow detection and shadow removal for high resolution city aerial photo.** International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, v. 35, p. 802-807, Istambul, 2004.



LUGNANI, J. B. **Introdução a Fototriangulação**. Curitiba: Editora UFPR, 1987.

MADHAVAN, B. B.; TACHIBANA, K.; SASAGAWA, T.; OKADA, H.; SHIMOZUMA, Y. **Automatic extraction of shadow regions in high-resolution ADS40 images - By robust approach of feature spaces analysis**. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, v. 35, p. 808-810, Istanbul, 2004.

MASSALABI, A.; HE, D.; BÉNIÉ, G. B.; BEAUDRY, E. **Restitution of information under shadow in remote sensing high space resolution images: Application to IKONOS data of Sherbrooke City**. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, v. 35, p. 173-178, Istanbul, 2004.

MEHROTRA, R.; NICHANI, S.; RANGANATHAN, N. **Corner detection**. Pattern Recognition, v. 23, n. 11, p. 1223-1233, 1990.

MIKHAIL, E. M.; BETHEL, J. S.; McGLONE, J. C. **Introduction to modern Photogrammetry**. John Wiley & Sons, 2001.

MORAVEC, H. P. **Towards automatic visual obstacle avoidance**. Proc. of the 5<sup>th</sup> Int. Joint Conference on Artificial Intelligence, p. 584, 1977.

NIELSEN, M. O. **True Orthophoto Generation**. Dissertação de Mestrado – Universidade Técnica da Dinamarca, Informatics and Mathematical Modelling, Kgs, Lyngby, 2004.

PETRIE, G.; TOTH, C. K. **Airborne and Spaceborne LASER Profilers and Scanners**. In: SHAN, J.; TOTH, C. K. (Eds.) Topographic Laser Ranging and Scanning - Principles and Processing. Boca Raton: CRC Press. 2008.

PFEIFER, N.; KOSTLI, A.; KRAUS, K. **Interpolation and filtering of laser scanner data: implementation and first results**. In: International Archives of Photogrammetry and Remote Sensing, Proceedings... 1998.

OLIVEIRA, H. C. **Detecção de áreas de oclusão para geração de ortoimagem verdadeira utilizando dados LASER**. Dissertação (mestrado em Ciências Cartográficas) – Faculdade de Ciências e Tecnologia, Universidade Estadual Paulista, Presidente Prudente, 2013.

OLIVEIRA, H. C. **Detecção de oclusão via análise de gradientes de superfície sobre um poliedro para geração de ortoimagem verdadeira**. Tese (doutorado em Ciências Cartográficas) – Faculdade de Ciências e Tecnologia, Universidade Estadual Paulista, Presidente Prudente, 2016.

OTSU, N. **A threshold Selection Method from Gray-Level Histograms**. IEEE Transactions on Systems, Man and Cybernetics, v. 9, n. 1, p. 62-69, 1979.

ROGGERO, M. **Airborne LASER scanning: clustering in raw data**. In: International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Proceedings... 2001.

ROTTENSTEINER, F.; SOHN, G.; GERKE, M.; WEGNER, J. D.; BREITKOPF, U.; JUNG, J. **Results of the ISPRS benchmark on urban object detection and 3D building**

**reconstruction.** ISPRS Journal of Photogrammetry and Remote Sensing, v. 93, p. 256-271, 2014.

SEO, S.; LEE, J.; KIM, Y. **Extraction of boundaries of rooftop fenced buildings from airborne laser scanning data using rectangle models.** IEEE Geoscience and Remote Sensing Letters, v. 11, n. 2, p. 404-408, 2014.

SITHOLE, G. **Filtering of LASER altimetry data using a slope adaptive filter.** In: International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Proceedings... 2001.

SITHOLE, G.; VOSSSELMAN, G. **Experimental comparison of filter algorithms for bare-Earth extraction from airborne laser scanning point clouds.** ISPRS Journal of Photogrammetry and Remote Sensing, v. 59, p. 85-101, 2004.

SOHN, G.; DOWMAN, I. **Terrain surface reconstruction by the use of tetrahedron model with the MDL Criterion.** In: XXXIV International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences (Pt. 3A), p. 336–344, 2002.

SOILLE, P. **Morphological Image Analysis.** Berlin. Springer-Verlag, 2004.

SÜLI, E.; MAYERS, D. **An Introduction to Numerical Analysis.** Cambridge University Press, 2003.

TOMMASELLI, A. M. G.; GALO, M.; DOS REIS, T. T.; RUY, R. S.; DE MORAES, M. V. A.; MATRICARDI, W. V. **Development and assessment of a dataset containing frame images and dense airborne laser scanning point clouds.** IEEE Geoscience and Remote Sensing Letters, v. 15(2), p. 192–196, 2018.

VOLOTÃO, C. F. S. **Geração de Ortoimagens para aplicações urbanas: desenvolvimento das equações e protótipo experimental.** Dissertação de Mestrado – Instituto Nacional de Pesquisas Espaciais (INPE), São José dos Campos, 2001.

WACK, R.; WIMMER, A. **Digital terrain models from airborne laser scanner data: a grid based approach.** In: International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Proceedings..., 2002.

WOLF, P. R.; DEWITT, B. A. **Elements of Photogrammetry - with Applications in GIS.** ed. 3, New York: McGraw-Hill, 2000.

WOLF, P. R.; DEWITT, B. A. WILKINSON, B. E. **Elements of Photogrammetry – with applications in GIS.** ed. 4, USA: McGraw-Hill Education, 2014.

WU, B.; TANG, S.; ZHU, Q.; TONG, K.; HU, H.; LI, G. **Geometric integration of high-resolution satellite imagery and airborne LiDAR data for improved geopositioning accuracy in metropolitan areas.** ISPRS Journal of Photogrammetry and Remote Sensing, v. 109, p. 139–151, 2015.

YAN, J.; ZHANG, K.; ZHANG, C. **Automatic construction of 3-D building model from airborne LiDAR data through 2-D Snake algorithm.** IEEE Transactions on Geoscience and Remote Sensing, v. 53, n. 1, 2015.

ZHOU, G.; ZHOU, X. **Seamless Fusion of LiDAR and Aerial Imagery for Building Extraction.** IEEE Transactions on Geoscience and Remote Sensing, v. 52, p. 7393-7407, 2014.

ZIVIANI, N. **Projeto de algoritmos: com implementação em Java e C++.** São Paulo: Cengage Learning, 2011.