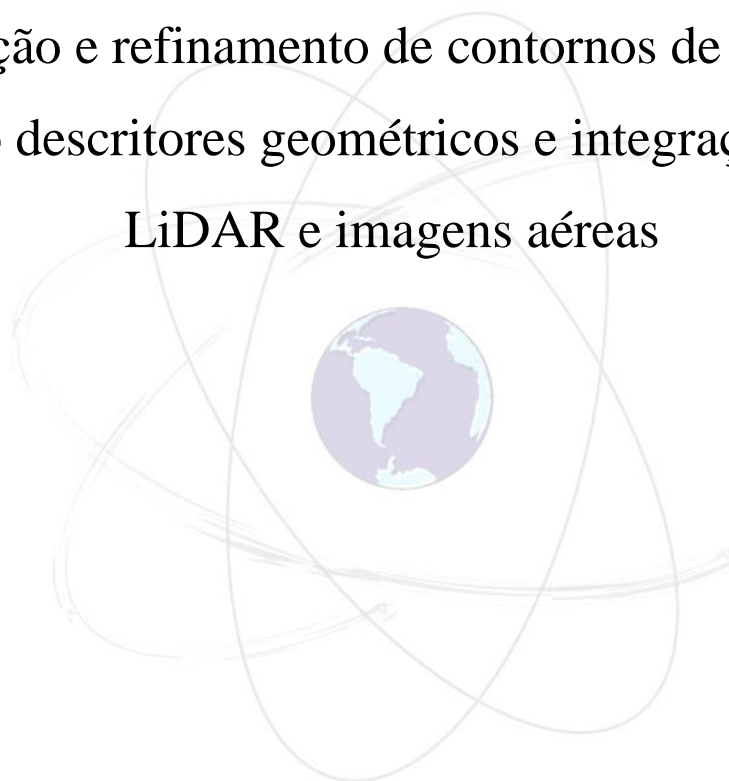




UNIVERSIDADE ESTADUAL PAULISTA
“JÚLIO DE MESQUITA FILHO”
Faculdade de Ciências e Tecnologia – Presidente Prudente
Programa de Pós-Graduação em Ciências Cartográficas

ANDRÉ CACERES CARRILHO

Extração e refinamento de contornos de telhados
utilizando descritores geométricos e integração de dados
LiDAR e imagens aéreas



PRESIDENTE PRUDENTE

2021

UNIVERSIDADE ESTADUAL PAULISTA
“JÚLIO DE MESQUITA FILHO”
Faculdade de Ciências e Tecnologia – Presidente Prudente
Programa de Pós-Graduação em Ciências Cartográficas

ANDRÉ CACERES CARRILHO

Extração e refinamento de contornos de telhados
utilizando descritores geométricos e integração de dados
LiDAR e imagens aéreas

Tese de doutorado apresentada ao Programa de Pós-Graduação em Ciências Cartográficas (PPGCC), da Faculdade de Ciências e Tecnologia (FCT), Universidade Estadual Paulista (UNESP), campus de Presidente Prudente/SP, como parte dos requisitos para a obtenção do título de Doutor em Ciências Cartográficas.

Área de concentração: Aquisição, Análise e Representação de Informações Espaciais.

Orientador: Prof. Dr. Mauricio Galo

PRESIDENTE PRUDENTE

2021

C317e

Carrilho, André Caceres

Extração e refinamento de contornos de telhados utilizando descritores geométricos e integração de dados LiDAR e imagens aéreas / André Caceres Carrilho. -- Presidente Prudente, 2021
132 f.

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

1. Dados LiDAR aerotransportado. 2. Imagens aéreas de alta
resolução. 3. Descritores geométricos. 4. Assinatura do contorno. 5.
Redes Neurais. 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 e refinamento de contornos de telhados utilizando descritores geométricos e integração de dados LIDAR e imagens aéreas

AUTOR: ANDRÉ CACERES CARRILHO

ORIENTADOR: MAURICIO GALO

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

Prof. Dr. MAURICIO GALO (Participação Virtual) **VIDEOCONFERÊNCIA**
Departamento de Cartografia / Faculdade de Ciências e Tecnologia de Presidente Prudente

Prof. Dr. DANIEL RODRIGUES DOS SANTOS (Participação Virtual) **VIDEOCONFERÊNCIA**
Departamento de Geomática / Universidade Federal do Paraná

Prof. Dr. IRINEU DA SILVA (Participação Virtual) **VIDEOCONFERÊNCIA**
Escola de Engenharia de São Carlos / Universidade de São Paulo

Prof. Dr. ANTONIO MARIA GARCIA TOMMASELLI (Participação Virtual) **VIDEOCONFERÊNCIA**
Departamento de Cartografia / Faculdade de Ciências e Tecnologia de Presidente Prudente

Prof. Dr. ALUIR PORFIRIO DAL POZ (Participação Virtual) **VIDEOCONFERÊNCIA**
Departamento de Cartografia / Faculdade de Ciências e Tecnologia de Presidente Prudente

Presidente Prudente, 08 de janeiro de 2021

DEDICATÓRIA

À minha mãe Maria Isabel.

À minha tia Angela.

Ao meu irmão Bruno.

Ao Nino e a Lina.

AGRADECIMENTOS

Agradeço ao professor Dr. Mauricio Galo, pela diligência durante a orientação, atenção, pelos ensinamentos passados e pela paciência.

Aos professores Dr. Irineu da Silva, Dr. Daniel Rodrigues dos Santos, Dr. Antonio Maria Garcia Tommaselli e Dr. Aluir Porfírio Dal Poz pelas sugestões feitas na defesa.

Aos professores e funcionários do Departamento de Cartografia da FCT/UNESP, pelo apoio e subsídio para o desenvolvimento deste trabalho.

Ao Programa de Pós-Graduação em Ciências Cartográficas (PPGCC), à Universidade Estadual Paulista (UNESP) pela infraestrutura fornecida.

A todos os colegas do PPGCC, especialmente ao Renato e ao Guilherme (Alemão) pelas conversas e sugestões.

À empresa Sensormap Geotecnologia pela disponibilização das imagens aéreas e dos dados de varredura a LASER utilizados nos experimentos.

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

RESUMO

Este trabalho trata da extração automática de telhados de edificações a partir de dados de Sistema de Varredura LASER Aerotransportado (SVLA) e imagens aéreas de alta resolução. O foco se refere à regularização de telhados cujo contorno tenha sido obstruído por algum outro objeto, como copas de árvore, por exemplo. Existem diversos problemas envolvidos nesse processo, tais como a preparação dos dados (remoção de ruídos, segmentação de objetos e classificação), extração dos contornos, modelagem dos contornos, detecção das oclusões e a reconstrução do contorno. Dentre as contribuições desta Tese de doutorado, a principal é a utilização de descritor geométrico (assinatura do contorno) como uma maneira de modelar o polígono e detectar possíveis obstruções. Embora não seja possível resolver todos os casos, como oclusões severas, a abordagem proposta deve auxiliar na automação de processos por meio da diminuição da intervenção do operador. Além da contribuição mencionada, outras soluções foram propostas para etapas como a remoção de pontos espúrios, filtragem do modelo digital de terreno (MDT) utilizando um método adaptativo de ajuste de superfície polinomial bivariada, integração da informação radiométrica na nuvem de pontos, e a extração de contornos de telhados utilizando redes neurais convolucionais (RNC). Os experimentos foram realizados com dados derivados do *Unesp Photogrammetry Data Set*, adquiridos sobre o município de Presidente Prudente, São Paulo, Brasil. As análises quantitativas dos experimentos realizados demonstram que o método proposto consegue reconstruir obstruções em telhados de edificações que não sejam muito severas, isto é, não sejam maiores que 1/3 do segmento linear da borda do telhado e que não sejam em uma quina (vértice) do polígono. Além disso, todas as etapas precedentes apresentaram resultados satisfatórios, como a acurácia global da classificação acima de 93%.

Palavras-chave: Extração de contornos de telhados. Dados SVLA. Imagens aéreas.

ABSTRACT

This thesis discusses the automatic building roofs extraction from airborne laser scanning (ALS) data and high-resolution aerial images. The principal concern is the refinement of roofs whose contour has been obstructed by other objects, such as treetops. There are several problems involved in generating this result, such as data preparation (outlier detection, object segmentation and classification), contour extraction, contour modeling, occlusion detection and contour reconstruction. Among the contributions of this doctoral thesis, the main one is the use of geometric descriptor (signature) to model the polygon and detect possible obstructions. Although it is not possible to resolve all cases, such as severe occlusions, the proposed approach should aid in process automation by reducing operator intervention. In addition to the main contribution, other innovative solutions were proposed for steps such as the removal of outliers, digital terrain model (MDT) filtering using an adaptive bivariate polynomial surface fitting, integration of radiometric information in the point cloud, and the extraction of roof contours using convolutional neural networks (CNN). The experiments were carried out with data derived from the Unesp Photogrammetry Data Set, acquired over the municipality of Presidente Prudente, São Paulo, Brazil. The quantitative analysis of the experiments demonstrate that the proposed method can reconstruct obstructions in roofs of buildings that are not severe, that is, not larger than $1/3$ of the linear segment of the roof edge and that is not in a corner (vertex) of the polygon. In addition, all the preceding steps showed adequate results, such as the overall accuracy of the classification above 93%, for example.

Keywords: Building boundary extraction. ALS data. Aerial imagery.

LISTA DE FIGURAS

| | |
|--|----|
| FIGURA 1 - PROBLEMAS DEVIDO A SOBREPOSIÇÃO DE VEGETAÇÃO EM EDIFICAÇÕES..... | 8 |
| FIGURA 2 - COBERTURA DE UMA FAIXA COM UM SVLA. | 10 |
| FIGURA 3 - SISTEMAS DE COORDENADAS E PARÂMETROS ENVOLVIDOS NA AQUISIÇÃO DE DADOS SVLA..... | 12 |
| FIGURA 4 - INTERAÇÃO DO PULSO LASER COM A VEGETAÇÃO CAUSANDO MÚLTIPLOS RETORNOS. | 13 |
| FIGURA 5 - EXEMPLOS DE ESTRUTURAS COMUMENTE EMPREGADAS NA REPRESENTAÇÃO DE SUPERFÍCIES. | 15 |
| FIGURA 6 - REPRESENTAÇÃO EM NOTAÇÃO CIENTÍFICA DA COORDENADA N DA ESTAÇÃO GEODÉSICA SAT 92579/1995..... | 19 |
| FIGURA 7 - SOBREPOSIÇÃO ENTRE DUAS FUNÇÕES DE DISTRIBUIÇÃO DE PROBABILIDADE NORMAIS COM MÉDIAS E VARIÂNCIAS DIFERENTES. OS PONTOS EM QUE ELAS SE INTERCEPTAM ESTÃO INDICADOS PELAS LINHAS PONTILHADAS..... | 21 |
| FIGURA 8 - CLASSES LINEARMENTE SEPARÁVEIS. ILUSTRAÇÃO DE DUAS POSSÍVEIS SOLUÇÕES PARA O HIPERPLANO. | 24 |
| FIGURA 9 - TOPOLOGIA <i>FEED-FORWARD</i> PARA UM PERCEPTRON DE MÚLTIPLAS CAMADAS. | 25 |
| FIGURA 10 - APROXIMAÇÃO DA FUNÇÃO DEGRAU A PARTIR DA FUNÇÃO LOGÍSTICA SIGMÓIDE. | 27 |
| FIGURA 11 – EXEMPLO DE REDE NEURAL CONVOLUTIONAL PARA CLASSIFICAÇÃO DE IMAGENS COLORIDAS. | 29 |
| FIGURA 12 - SELEÇÃO DO CONJUNTO DE PONTOS PARA O CÁLCULO DA MATRIZ DE COVARIÂNCIAS. | 33 |
| FIGURA 13 – CLASSIFICAÇÃO SUPERVISIONADA DE NUVEM DE PONTOS POR ENGENHARIA DE ATRIBUTOS. | 36 |
| FIGURA 14 – FLUXOGRAMA DA CLASSIFICAÇÃO DE NUENS DE PONTOS COM ENGENHARIA DE ATRIBUTOS. | 37 |
| FIGURA 15 - TRÊS FIGURAS GEOMÉTRICAS SIMPLES E AS RESPECTIVAS ASSINATURAS. | 39 |
| FIGURA 16 - PROBLEMA COM O DESCRITOR DE CONTORNO $r(\theta)$. (A) OBJETO CIRCULAR COM PEQUENA PERFURAÇÃO NA BORDA. (B) MESMO OBJETO, PORÉM PARCIALMENTE SUJEITO À OCLUSÕES. | 40 |
| FIGURA 17 - DIFERENÇA ENTRE OS GRÁFICOS DOS DESCRITORES $r(\theta)$ E $k(s)$ | 40 |
| FIGURA 18 – FLUXOGRAMA DO MÉTODO PROPOSTO. EM AZUL SÃO DESTACADAS AS CONTRIBUIÇÕES..... | 41 |
| FIGURA 19 – PROBLEMAS AO REGISTRAR INFORMAÇÃO RADIOMÉTRICA À INFORMAÇÃO POSICIONAL..... | 47 |
| FIGURA 20 - ANÁLISE DE VISIBILIDADE DA NUVEM DE PONTOS CONSIDERANDO SEIS IMAGENS AÉREAS..... | 48 |
| FIGURA 21 - IMAGENS DA MESMA REGIÃO TOMADAS EM HORÁRIOS DIFERENTES. ALTERAÇÕES NA ILUMINAÇÃO (POSIÇÃO DO SOL) E NA COMPOSIÇÃO DE OBJETOS NA CENA (VEÍCULOS, PRINCIPALMENTE)..... | 49 |
| FIGURA 22 - DETERMINAÇÃO DO VETOR NORMAL MÉDIO PARA O VÉRTICE CENTRAL DE UMA PEQUENA MALHA TRIANGULAR..... | 50 |
| FIGURA 23 – PRINCÍPIO DO AJUSTE PROGRESSIVO POR MEIO DO AJUSTE DE SUPERFÍCIE POLINOMIAL AOS PONTOS DO MDT. | 51 |
| FIGURA 24 - SUBDIVISÃO DA NUVEM DE PONTOS EM UMA MALHA REGULAR DE CÉLULAS RETANGULARES. | 52 |
| FIGURA 25 – DETERMINAÇÃO DE SEGMENTO DE BORDA COM O ALGORITMO EM (A) DOIS PONTOS DE BORDA E EM (B) DOIS PONTOS INTERNOS..... | 56 |
| FIGURA 26 - RESULTADOS OBTIDOS CONSIDERANDO (A) $A = 2$, (B) $A = \infty$ E (C) $A = 0$ | 56 |
| FIGURA 27 - CORREÇÃO DO CONTORNO DO TELHADO AFETADO POR SILHUETA DE VEGETAÇÃO (COPA DE ÁRVORE) SOBREPOSTA A EDIFICAÇÃO. | 59 |
| FIGURA 28 – CONTORNO DE TELHADO DE EDIFICAÇÃO EM FORMA DE L SIMULADO, NO QUAL O CENTRO DE MASSA (PONTO NA COR VERMELHA) ENCONTRA-SE EXTERNO AO POLÍGONO. | 60 |

| | |
|---|----|
| FIGURA 29 – CADA UM DOS SEGMENTOS DE TELHADO MOSTRADO NA FIGURA 28 E RESPECTIVAS CURVAS AJUSTADAS. OBSERVA-SE QUE EM ALGUMAS SITUAÇÕES A PARÁBOLA NÃO ATENDE (RESÍDUO ALTO), SENDO NECESSÁRIO A UTILIZAÇÃO DE UMA OUTRA CÔNICA PARA MODELAR A CURVA DADA POR $(\theta, r\theta)$ | 61 |
| FIGURA 30 – GERAÇÃO DO DESCRITOR $r(\theta)$ PARA UM SEGMENTO DE RETA ij E DO CENTRO DE MASSA (C) DE UM POLÍGONO REFERENTE AO CONTORNO DO TELHADO DE UMA EDIFICAÇÃO..... | 63 |
| FIGURA 31 - TERMINOLOGIA DOS ELEMENTOS DERIVADOS DA MATRIZ DE ERRO. EM (C): VERDADEIRO POSITIVO (VP); VERDADEIRO NEGATIVO (VN); FALSO POSITIVO (FP); E FALSO NEGATIVO (FN). | 67 |
| FIGURA 32 – NUVEM DE PONTOS DA REGIÃO 1: MORADIA ESTUDANTIL DO CAMPUS DA UNESP..... | 71 |
| FIGURA 33 – NUVEM DE PONTOS DA REGIÃO 2: DIRETORIA E ALGUNS DOS DEPARTAMENTOS E LABORATÓRIOS DO CAMPUS DA UNESP. | 71 |
| FIGURA 34 – REGIÃO 3: BAIRRO MORUMBI, ÁREA PARCIALMENTE URBANIZADA. | 72 |
| FIGURA 35 – CONTORNOS DOS TELHADOS OBTIDOS A PARTIR DE RESTITUIÇÃO DAS IMAGENS AÉREAS. | 72 |
| FIGURA 36 – ÁREA DE ESTUDOS: PROXIMIDADES DA REGIÃO 3 (BAIRRO MORUMBI). COORDENADAS UTM FUSO 22S. | 75 |
| FIGURA 37 – EFEITO DA QUANTIDADE DE PONTOS VIZINHOS (K) NO FILTRO SOR, E DO LIMIAR DE FREQUÊNCIA NO FE. CONSIDEROU-SE CÉLULAS DE 2 M PARA O FE. | 75 |
| FIGURA 38 – EFEITO DO LIMIAR DE FREQUÊNCIA NO MÉTODO DE FILTRAGEM PROPOSTO (HC). CONSIDEROU-SE CÉLULAS QUADRADAS DE 50 M E HISTOGRAMA COM INTERVALOS DE 15 CM. | 76 |
| FIGURA 39 – OCORRÊNCIAS DE FALSO NEGATIVO NAS PAREDES DE EDIFICAÇÕES. | 77 |
| FIGURA 40 – CONJUNTO DE DADOS PARA TREINAMENTO DE RNC CONTENDO DEZ AMOSTRAS PARA CADA UMA DAS SETE CLASSES DE OBJETOS PRESENTES EM CENAS URBANAS. | 78 |
| FIGURA 41 – ABORDAGEM ADOTADA PARA EXTRAÇÃO DE OBJETOS UTILIZANDO RNC. | 78 |
| FIGURA 42 – ILUSTRAÇÃO DA ARQUITETURA DA RNC ADOTADA..... | 79 |
| FIGURA 43 – ABORDAGEM SLIC MULTI-ESCALA. (A) IMAGEM ORIGINAL, SUPERPIXELS GERADOS COM (B) $k = 50$, (C) $k = 80$ E (D) $k = 150$ | 79 |
| FIGURA 44 – DETECÇÃO DE OBJETOS EM DUAS REGIÕES URBANAS. RETÂNGULOS VERMELHOS: EDIFICAÇÕES, VERDES: VEGETAÇÃO, AZUL: PISCINAS..... | 80 |
| FIGURA 45 – EXTRAÇÃO DE OBJETOS. IMAGENS AÉREAS EM (A) E (B), RESULTADOS EM (C) E (D), RESPECTIVAMENTE. LEGENDA: POLÍGONOS VERMELHOS: EDIFICAÇÕES, VERDES: VEGETAÇÃO, AZUL: PISCINAS. | 81 |
| FIGURA 46 – RESULTADO DA AMOSTRAGEM EM UMA REGIÃO DE SOBREPOSIÇÃO DE FAIXAS. COMPARAÇÃO ENTRE DOIS MECANISMOS DE VARREDURA DIFERENTES. (A) ESPELHO OSCILANTE, E (B) POLÍGONO DE ROTAÇÃO. | 82 |
| FIGURA 47 – CONTORNOS EXTRAÍDOS CONSIDERANDO DIFERENTES VALORES PARA O PARÂMETRO α | 83 |
| FIGURA 48 – INSPEÇÃO VISUAL DOS RESULTADOS DA INTEGRAÇÃO RADIOMÉTRICA AOS DADOS SVLA E A UTILIZAÇÃO DO CRITÉRIO DA MODA DOS VALORES RGB NO MAPEAMENTO $1 \rightarrow n$ | 85 |
| FIGURA 49 – RESULTADO DA CLASSIFICAÇÃO DA METADE SUPERIOR DA REGIÃO 3. | 88 |
| FIGURA 50 – TELHADOS SELECIONADOS COM OBSTRUÇÕES CAUSADAS POR VEGETAÇÃO. | 90 |
| FIGURA 51 – RESULTADOS PARA O TELHADO 1 (REGIÃO 1 - MORADIA ESTUDANTIL) | 91 |
| FIGURA 52 – RESULTADOS PARA O TELHADO 2 (REGIÃO 2 – LABORATÓRIO) | 92 |

| | |
|--|----|
| FIGURA 53 – RESULTADOS PARA O TELHADO 3 (REGIÃO 2 – LABORATÓRIO) | 93 |
| FIGURA 54 – RESULTADOS PARA O TELHADO 4 (REGIÃO 2 – LABORATÓRIO) | 94 |
| FIGURA 55 – SITUAÇÕES REAIS ONDE AS OCLUSÕES SÃO DE GRANDE MAGNITUDE, DIFICULTANDO A IDENTIFICAÇÃO DE ALGUMAS DAS BORDAS DAS EDIFICAÇÕES. | 96 |
| FIGURA 56 – UTILIZAÇÃO DO CRITÉRIO DE CONECTIVIDADE COM CRESCIMENTO DE REGIÕES PARA MITIGAR PROBLEMAS DE SEGMENTAÇÃO DE PLANOS COM O ALGORITMO RANSAC. | 97 |

SUMÁRIO

| | |
|---|-----------|
| CAPÍTULO 1 DETECÇÃO AUTOMÁTICA DE EDIFICAÇÕES..... | 1 |
| 1.1 Introdução | 1 |
| 1.2 Caracterização do problema a partir do estado da arte | 3 |
| 1.3 Objetivos..... | 8 |
| 1.4 Contribuições científicas..... | 8 |
| CAPÍTULO 2 AMOSTRAGEM E REPRESENTAÇÃO DE SUPERFÍCIES..... | 10 |
| 2.1 Sistemas de varredura a LASER aerotransportados | 10 |
| 2.2 Interação do pulso emitido com a vegetação | 13 |
| 2.3 Representação de superfícies | 14 |
| 2.4 Considerações sobre o armazenamento dos dados | 16 |
| CAPÍTULO 3 RECONHECIMENTO DE PADRÕES..... | 20 |
| 3.1 Coeficiente de sobreposição entre duas curvas Gaussianas..... | 20 |
| 3.2 Discriminantes lineares | 22 |
| 3.3 Redes neurais artificiais | 24 |
| 3.3.1 Funções de ativação | 25 |
| 3.3.2 Redes Neurais Convolucionais..... | 28 |
| 3.3.3 Normalização dos dados de entrada | 31 |
| 3.4 Descritores geométricos..... | 32 |
| 3.4.1 Informação de variância local | 32 |
| 3.4.2 Atributos geométricos invariantes à densidade de pontos | 36 |
| 3.4.3 Assinatura do contorno..... | 38 |
| CAPÍTULO 4 EXTRAÇÃO DE CONTORNOS DE TELHADOS..... | 41 |
| 4.1 Remoção de pontos espúrios..... | 42 |
| 4.2 Integração da informação radiométrica | 45 |
| 4.3 Cálculo dos vetores normais | 49 |
| 4.4 Filtragem por ajuste progressivo de superfícies polinomiais..... | 50 |
| 4.5 Detecção de regiões planas | 53 |
| 4.6 Algoritmo RANSAC..... | 54 |
| 4.7 Algoritmo α -shape..... | 55 |
| 4.8 Algoritmo SLIC | 56 |
| 4.9 Refinamento do contorno dos telhados..... | 58 |
| 4.10 Seleção do modelo matemático para o ajuste de segmentos de contorno..... | 59 |
| 4.11 Avaliação quantitativa dos resultados..... | 66 |
| CAPÍTULO 5 EXPERIMENTOS E RESULTADOS..... | 70 |
| 5.1 Materiais utilizados..... | 70 |
| 5.2 Experimentos e discussões..... | 74 |

| | |
|---|------------|
| 5.2.1 Remoção de pontos espúrios..... | 74 |
| 5.2.2 Extração do contorno de objetos utilizando RNC e SLIC | 77 |
| 5.2.3 Extração do contorno dos telhados de edificações com o algoritmo α -shape..... | 82 |
| 5.2.4 Integração da informação radiométrica aos dados SVLA | 84 |
| 5.2.5 Classificação das nuvens de pontos | 86 |
| 5.2.6 Extração e regularização dos contornos utilizando a assinatura..... | 89 |
| 5.2.7 Segmentação dos planos de telhados de edificações | 96 |
| CAPÍTULO 6 CONCLUSÕES E RECOMENDAÇÕES | 98 |
| REFERÊNCIAS | 101 |
| APÊNDICE A | 114 |

Capítulo 1

DETECÇÃO AUTOMÁTICA DE EDIFICAÇÕES

O primeiro capítulo apresenta o estado da arte referente à detecção de edificações em dados de Sistema de Varredura LASER Aerotransportado (SVLA) e expõe situações problemáticas em que os algoritmos atuais são incapazes de tratar corretamente. As questões científicas associadas aos problemas mencionados, hipótese, objetivos, e a estrutura do documento também são detalhados.

1.1 Introdução

Com o decorrer do tempo a superfície terrestre sofre alterações decorrentes de ações antrópicas e/ou naturais, o que torna necessária a atualização da base cartográfica que a representa. Países emergentes apresentam mudanças significativas em seus centros urbanos, mudanças essas resultantes de atividades antrópicas, as quais englobam a ampliação da malha viária, a construção de novas edificações, reformas ou demolições das existentes para destinar o espaço para outras finalidades e/ou novos empreendimentos.

A detecção automática de edificações a partir de dados obtidos remotamente é importante para uma série de tarefas, tais como o planejamento urbano, identificação de construções ilegais, monitoramento do crescimento urbano e avaliações de desastres naturais. O processo de atualização de bases cartográficas por técnicas fotogramétricas, embora tenha um elevado nível de automação em algumas etapas, ainda é dependente do operador em algumas tarefas como a restituição e vetorização, o que acaba por ser um procedimento oneroso. Por outro lado, a manutenção dessas bases cartográficas necessita de métodos confiáveis, com procedimentos sistemáticos para a detecção de edificações (HERMOSILLA *et al.*, 2011).

Métodos automáticos para a detecção de feições têm recebido atenção da comunidade científica nas últimas décadas, devido ao interesse prático envolvido, isto é, a redução dos custos do processo e do tempo de processamento, sem comprometer a qualidade do resultado obtido, quando comparado ao obtido por uma referência gerada manualmente por um operador treinado. Atualmente, no Grupo de Trabalho 3 (*Point Cloud Processing*) da Comissão II da *International Society for Photogrammetry and Remote Sensing (ISPRS)*, aparecem os seguintes termos de referência: desenvolvimento de novos métodos e algoritmos para processamento de

nuvens de pontos; extração de informação de nuvens de pontos (segmentação e classificação); computação de alta performance para grandes conjuntos de dados; e geração de modelos de edificações (*building information modelling* - BIM) por meio de nuvem de pontos. Tais termos atestam tanto a atualidade do tema quanto o empenho da comunidade científica no desenvolvimento de novos métodos, algoritmos e aplicações para o processamento de dados SVLA.

Segundo Pang *et al.* (2014) e Awrangjeb *et al.* (2015), os métodos automáticos para a detecção de edificações podem ser categorizados de acordo com a origem dos dados utilizados: imagens multiespectrais de alta resolução (aéreas e orbitais); dados de elevação obtidos tanto por métodos de correspondência de imagens quanto por sensores ativos, como SVLA ou SAR (*Synthetic Aperture Radar*); e a combinação de dados provenientes de múltiplas fontes. Historicamente as imagens obtidas por plataformas aéreas e orbitais, assim como os produtos derivados, isto é, ortoimagens e ortomosaicos, são conjuntos de dados frequentemente empregados para a geração e atualização de bases cartográficas. Entretanto, nas últimas décadas houve um avanço na utilização de nuvens de pontos providas de perfilamento a LASER (*Light Amplification by Stimulated Emission of Radiation*) para a obtenção de informação geométrica da superfície terrestre (RÖNNHOLM *et al.*, 2007; HERMOSILLA *et al.*, 2011; BUJÁN *et al.*, 2012; GILANI *et al.*, 2017; DAL POZ e YAWATA, 2019; SÁNCHEZ-APARICIO *et al.*, 2020). De acordo com Buján *et al.* (2012), os dados SVLA possuem algumas vantagens com relação às nuvens de pontos obtidas por outras técnicas: independência de condições atmosféricas, uma vez que o pulso LASER não é afetado por sombras e ângulo solar (sensor ativo); registro de informação de intensidade do retorno, o que permite inferir algumas características da superfície dos alvos; penetração do pulso LASER na vegetação e o registro dos múltiplos retornos; e a coleta de pontos de coordenadas 3D com alta densidade. No entanto, a detecção automática de edificações pode apresentar resultados incompletos quando os dados utilizados são obtidos por sistemas de aquisição diferentes (SVLA e imageamento por sensor óptico) caso eles sejam utilizados isoladamente, o que motivou a integração dessas informações para aprimorar a qualidade dos resultados (HERMOSILLA *et al.*, 2011; BUJÁN *et al.*, 2012; MALPICA *et al.*, 2013).

Com o intuito de apresentar o estado da arte, realizou-se um levantamento de trabalhos publicados em periódicos e em eventos arbitrados. O panorama obtido permitiu caracterizar os problemas relacionados aos métodos de detecção de edificações utilizando imagens aéreas e dados SVLA.

1.2 Caracterização do problema a partir do estado da arte

Encontram-se na literatura diversas abordagens para a extração de feições em regiões urbanas, dentre elas Murakami *et al.* (1999) apresentam um método para identificar mudanças em edificações detectadas com dados SVLA, utilizando imagens digitais para a validação visual dos resultados. Os autores comparam os modelos digitais de superfície (MDS) obtidos em épocas distintas, negligenciando o coregistro entre os dados, o que resulta em problemas nos contornos das edificações. Os autores salientam a sinergia entre os dados SVLA e imagens aéreas, citando o potencial de utilização conjunta.

Vosselman *et al.* (2004) propõem um método de atualização de bases cartográficas por meio da filtragem de dados SVLA aliado a imagens digitais. O MDS é filtrado para a obtenção do modelo digital do terreno (MDT), e os pontos filtrados (que não pertencem ao terreno) são classificados em vegetação ou edificações, segundo os seguintes critérios: rugosidade da superfície; área e altura do segmento; diferença entre o primeiro e último retorno; radiometria (se disponível), onde o valor da componente matiz (cor) pode ser utilizado para distinguir a vegetação dos telhados de edificações. Cita-se no artigo alguns problemas relacionados com a classificação errônea da vegetação próxima a edificações.

Rottensteiner (2007) apresenta um método automático para a atualização de bases cadastrais urbanas utilizando MDSs obtidos por sistemas de varredura LASER aerotransportado (SVLA) e correspondência de imagens, integrados à informação radiométrica de imagens aéreas multiespectrais. O método utiliza os valores de *normalized difference vegetation index* (NDVI) para a detecção da vegetação. A integração entre os dados SVLA e as imagens multiespectrais é feita utilizando a teoria de Dempster-Shafer (KLEIN, 2012). Concluiu-se na avaliação que a integração dos dados possibilitou detectar mudanças que não eram triviais ao utilizar cada conjunto de dados isoladamente. Entretanto, o autor relata problemas para a detecção de edificações pequenas (área inferior a 50 m²).

No trabalho de Hermosilla *et al.* (2011) são apresentadas duas abordagens para a extração de edificações por meio da integração de dados SVLA e imagens multiespectrais de alta resolução. A primeira abordagem utiliza dois limiares, um referente a altitude mínima do ponto no modelo digital de superfície normalizado (MDSn) para que seja considerado edificação e outro relacionado ao valor mínimo de NDVI para ser rotulado como vegetação. A segunda abordagem é baseada em objetos e requer uma etapa de segmentação do MDSn, na qual os segmentos resultantes são caracterizados pela resposta espectral, textura, altitude e forma. O

artigo evidencia a utilização do contexto da edificação (como sombras, por exemplo) para a sua detecção. Os autores concluem que a abordagem baseada em limiares é mais vantajosa (por ser mais simples) e apresenta melhores resultados. Assim como apontado anteriormente por Vosselman *et al.* (2004), o trabalho refere-se às regiões com árvores muito próximas a edificações (bairros residenciais) como mais problemáticas para a classificação que bairros industriais ou densamente edificados.

Fazan (2011) propõe um método de extração de contornos de telhado de edificações por meio da integração de malha triangular (gerada a partir de dados SVLA) e imagens aéreas de alta resolução. A abordagem baseia-se na solução de um problema de otimização com modelagem via *snakes* e programação dinâmica (PD). O autor menciona a sensibilidade do método quanto à complexidade da cena urbana e cita como exemplo a obstrução de informação causada por objetos.

Buján *et al.* (2012) apresentam um método hierárquico de classificação baseado em objetos, utilizando dados SVLA e ortoimagens. O método consiste em uma árvore de decisões que leva em conta o MDSn, valor de intensidade do retorno, diferença entre o primeiro e último retorno, e o *ratio vegetation index* (RVI). A classe pavimento é obtida a partir dos pontos de menor elevação e com intensidades do retorno baixas. Para os pontos mais elevados e com valores de intensidade mais altos, a separação entre as classes vegetação e edificações foi feita por meio da avaliação do valor RVI. Embora o método tenha permitido obter bons resultados na detecção de edificações, os autores alertam sobre a utilização apenas em regiões com baixa densidade de construções (áreas rurais).

Malpica *et al.* (2013) apresentam um método de detecção de mudanças em regiões urbanas baseado em classificações supervisionadas com base no uso de *support vector machines* (SVM), utilizando imagens orbitais (SPOT5) e dados SVLA. Os autores empregaram imagens aéreas e informações da base vetorial da agência de mapeamento espanhola - Instituto Geográfico Nacional (IGN) - como referência para a atualização e verificação dos resultados. Algumas limitações são explicitadas devido a resolução das imagens orbitais (2,50 m), o que inviabiliza a utilização do método para escalas grandes. No trabalho são relatadas melhorias quanto a utilização de imagens multiespectrais (com as bandas do infravermelho) perante o uso de imagens aéreas obtidas apenas nas bandas do espectro visível, ainda que apresentem maior resolução espacial (0,50 m).

Xu *et al.* (2013) propõem uma abordagem para a detecção de mudanças em edificações por meio da comparação de dados SVLA classificados, obtidos em épocas diferentes. O método

de classificação da nuvem de pontos consiste no cálculo de descritores locais baseados nas coordenadas dos pontos vizinhos ao ponto estudado. Esses descritores permitem distinguir o tipo de entidade geométrica que o ponto pertence. Assim como observado por Murakami *et al.* (1999), os autores concluem que apesar de robusta, a abordagem depende da qualidade do registro entre as nuvens de pontos.

Dal Poz (2014) reforça a sinergia entre dados SVLA e imagens aéreas ao apresentar um método híbrido para o refinamento do contorno de telhados de edificações. No artigo é evidenciado o potencial dos dados SVLA para a extração de faces planas, enquanto a delimitação do contorno do telhado é melhor resolvida quando são utilizadas imagens aéreas. Fernandes e Dal Poz (2014, 2016), e Dal Poz e Fernandes (2017) apresentam um método para realizar o refinamento geométrico dos telhados de edificações, os quais são extraídos dos dados SVLA. A abordagem consiste na aplicação dos campos aleatórios de Markov (MRF – *Markov Random Field*) utilizando as relações de comprimento, proximidade e orientação de retas extraídas em imagens aéreas de alta resolução como dados de entrada. A minimização da função de energia foi realizada com o uso de algoritmo genético. Os autores mencionam as vantagens da integração dos dados para a descrição topológica da estrutura do telhado.

No trabalho de Tian *et al.* (2014) é apresentado um método de detecção de mudanças não-supervisionado utilizando imagens orbitais estereoscópicas providas dos satélites Ikonos e WorldView-2, e MDSs obtidos com o método de correspondência de imagens *semi-global matching* (SGM). Os indicadores de mudança obtidos por meio do MDS e por imagens orbitais foram integrados a partir da teoria de Dempster-Shafer (KLEIN, 2012). Além do método proposto, os autores comparam MDSs obtidos por correspondência de imagens e por SVLA, e sugerem o uso dos primeiros para situações práticas em que a aquisição de dados SVLA seja inviável ou onerosa.

Dresch e Santos (2015) propõem um método automático para o controle de qualidade relativo de dados SVLA aerotransportado. Os autores empregaram uma variante do método *iterative closest point* (ICP) para estimar os parâmetros de transformação entre faixas com sobreposição. O método baseia-se na extração e correspondência de feições planas, retas (segmentos) e pontuais, extraídas dos dados com o uso do algoritmo RANSAC e crescimento de regiões.

Oliveira (2016) apresenta um método de extração de contornos de edificações por meio da integração de dados SVLA e imagens aéreas de alta resolução. A abordagem consiste na determinação dos pontos do contorno das edificações nos dados SVLA os quais são projetados

nas imagens aéreas. O método emprega uma etapa de refinamento baseada no ajuste de retas nas bordas detectadas próximas às regiões dos contornos projetados. O autor menciona a presença de sombras ou objetos próximos às edificações como fatores adversos.

Fernandes *et al.* (2017) propõem uma abordagem para a filtragem de nuvem de pontos utilizando o método de interpolação por krigagem. A identificação de pontos semente (superfície aproximada do terreno) é realizada por meio de pontos amostrados sobre vias urbanas, as quais são detectadas pelo operador de Steger¹. A hipótese formulada considera que o interior das quadras possui comportamento suave e altitude similar à das vias.

Nos últimos anos o avanço científico e tecnológico proporcionou melhorias consideráveis tanto na capacidade de processamento quanto na acurácia dos algoritmos de reconhecimento de padrões. Neste contexto, Ševo e Avramović (2016) apresentaram um método para detecção de objetos em imagens aéreas e orbitais de alta resolução. A abordagem consiste em uma classificação supervisionada com redes neurais convolucionais (RNC, *convolutional neural networks* – CNN – ou *ConvNets*). Uma ideia similar foi apresentada por Alidoost e Arefi (2016), na qual a informação altimétrica (SVLA) foi integrada a cada pixel das imagens de entrada, as quais foram processadas com RNC para a geração de modelos tridimensionais das edificações.

Similarmente, Ghamisi *et al.* (2016) apresentaram um método para a integração de dados SVLA com imagens hiperespectrais utilizando *extinction profiles* (EP) e aprendizagem profunda (*deep-learning*), RNC no caso. O termo aprendizado profundo refere-se ao uso de técnicas de aprendizado de máquina baseadas em Redes Neurais Artificiais (RNA) com topologia *feed-forward* e com no mínimo três camadas ocultas, podendo ser convolucionais ou não. Neste artigo os autores compararam a aplicação de RNC profundas em conjunto com EP para a classificação dos dados. Segundo os autores, a principal desvantagem dessas abordagens baseadas em RNC é a necessidade de reamostrar a nuvem de pontos em uma malha regular para a aplicação das convoluções, processo este que pode introduzir erros nos dados.

Zarea and Mohammadzadeh (2016) propuseram uma abordagem para detecção de edificações e indivíduos arbóreos baseados na integração de dados SVLA e imagens aéreas. Os autores empregaram SVM para a classificação dos dados, e aplicaram diferentes técnicas de pós-processamento visando a melhoria da detecção, sendo que uma delas consiste na separação

¹ STEGER, C. Extracting lines using differential geometry and Gaussian smoothing. In: International Archives of Photogrammetry and Remote Sensing, v.31, 1996, Vienna, **Anais...**, ISPRS, 1996, p. 821-826.

de edificações muito próximas, as quais aparecem como um único segmento. Este último processamento utiliza as diferenças de altitudes dos telhados no agrupamento por k-médias.

Existem ainda outras abordagens como a proposta por Kang *et al.* (2017), na qual os autores aplicaram um método de classificação baseado em redes Bayesianas visando a integração de dados SVLA e imagens ópticas. Neste caso os resultados foram promissores, sendo que o método resultou em acurácia superior à 90% na classificação, entretanto, o potencial do classificador está voltado para a separação de pontos do terreno e vegetação rasteira, sendo que edificações e vegetação média/alta ainda são problemáticas.

Em Dal Poz e Ywata (2019) os autores propõem um método adaptativo de segmentação de planos de telhados amostrados por sistemas de varredura a laser aerotransportado baseado no algoritmo RANSAC. O método é composto de três etapas principais: delimitação da área de interesse (individualização dos telhados); segmentação dos planos de cada telhado; e verificação de consistência por meio de distância, e ângulos do vetor gradiente e vetor normal. Esta última etapa serve para mitigar os problemas de planos fragmentados durante a segmentação pelo algoritmo RANSAC.

A influência da densidade da nuvem de pontos obtida por SVLA na caracterização geométrica dos telhados foi estudada em Sánchez-Aparicio *et al.* (2020). Os autores fizeram uma comparação da modelagem de telhados utilizando dados com 0.5, 1.0, 2.0 e 14.0 pontos/m² disponibilizados pelo Instituto de Geografia da Espanha. O estudo conclui que quando são utilizadas nuvens de pontos com densidades baixas, maiores são os erros na estimativa das áreas, entretanto a melhoria em acurácia não é proporcional ao aumento na densidade.

A partir dos trabalhos descritos verificam-se problemas quanto à presença de vegetação próxima aos telhados de edificações. Embora essas situações ocorram com mais frequência em regiões específicas (residenciais, no caso), os métodos atuais não são capazes de lidar com elas. Observou-se também limitações referentes a edificações cujos telhados sejam pequenos e que podem ser confundidos com outras feições. Na Figura 1 são ilustrados exemplos de edificações com vegetação ao redor.

Com base no que foi discutido, pretende-se desenvolver um método para a atualização da base cartográfica urbana com foco na extração e refinamento dos contornos de edificações sujeitas a oclusão por vegetação no seu entorno. A hipótese formulada pressupõe a utilização de descritores geométricos (assinatura do contorno dos telhados das edificações) em conjunto com técnicas de classificação supervisionada para a extração de telhados de edificações.

Figura 1 - Problemas devido a sobreposição de vegetação em edificações.



(a) Árvore próxima ao telhado (b) Árvore em calçada (c) Vegetação em quintal

Fonte: Imagens aéreas obtidas pela empresa Sensormap Geotecnologia em 2014.

1.3 Objetivos

O propósito geral deste trabalho consiste no desenvolvimento de um método híbrido para a detecção de edificações por meio da integração de dados SVLA e imagens aéreas de alta resolução. O foco, no entanto, consiste na extração e a regularização, quando necessária, dos contornos dos telhados detectados utilizando descritores geométricos. Os objetivos específicos são:

- Extrair o MDT utilizando um filtro adaptativo baseado no ajuste progressivo de superfícies polinomiais bivariadas;
- Associar aos dados SVLA a informação radiométrica provinda de imagens aéreas de alta resolução obtidas por sensor de quadro;
- Identificar planos de telhados com o método supervisionado baseado em RNA, utilizando diferentes métricas como indicadores geométricos;
- Realizar a extração automática e o refinamento dos contornos dos telhados de edificações por meio de descritores como a assinatura geométrica; e
- Avaliar quantitativamente o resultado da classificação com dados de referência gerados manualmente e por outros métodos (comparação relativa).

1.4 Contribuições científicas

Apesar dos esforços da comunidade científica no desenvolvimento de novos métodos de detecção de edificações, estes ainda apresentam problemas devido à complexidade da cena, como por exemplo, quando se tem a presença de vegetação próxima aos telhados de edificações. Almejando mitigar esses problemas, esta Tese investiga e propõe um método para a extração

dos contornos de telhados de edificações, por meio da integração de dados SVLA e imagens de alta resolução. Além de assimilar dados complementares, outro aspecto relevante consiste no emprego da assinatura do contorno dos telhados, com o intuito de detectar as situações problemáticas devido a obstruções causadas por outras feições, como a vegetação, por exemplo.

Em comparação às abordagens presentes no estado da arte, observa-se uma tendência no uso de RNC, entretanto, estes métodos requerem a interpolação dos dados em malhas regulares. De modo alternativo, o método proposto neste projeto prevê a aplicação de descritores geométricos para possibilitar a utilização de RNA ao invés das RNC. A relevância científica e tecnológica deste trabalho é evidenciada pela *International Society for Photogrammetry and Remote Sensing* (ISPRS), na qual os tópicos tratados neste projeto compõem em diversos termos de referência de alguns Grupos de Trabalho da Comissão Técnica II. O Quadro 1 apresenta alguns destes termos de referência vigentes durante o período de 2016 a 2020.

Quadro 1 – Alguns termos de referência de Grupos de Trabalho da ISPRS.

| Grupo de Trabalho | Termos de referência |
|--|--|
| II/2 - Geração de nuvens de pontos | Filtragem e integração de nuvens de pontos geradas a partir de diferentes fontes, visando a reconstrução de superfícies; |
| II/3 - Processamento de nuvens de pontos | Desenvolvimento de novas metodologias, algoritmos e aplicações para processamento de nuvens de pontos; |
| | Extração de informação a partir de nuvens de pontos, segmentação e classificação; |
| II/4 - Reconstrução 3D e análise da cena | Classificação e segmentação semântica de nuvens de pontos e superfícies com ou sem informação radiométrica; |
| | Geração e atualização de modelos 3D de cidades de alta resolução e bancos de dados de vias [...] |

Fonte: Adaptado de ISPRS².

² Disponível em: <http://www2.isprs.org/commissions/comm2/wgs.html>. Acessado em: 18 ago. 2017.

Capítulo 2

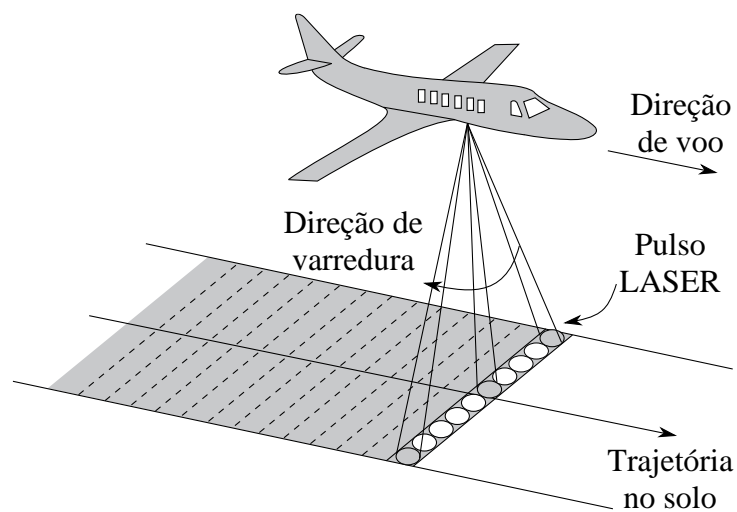
AMOSTRAGEM E REPRESENTAÇÃO DE SUPERFÍCIES

Este capítulo apresenta uma breve revisão sobre Sistemas de Varredura LASER Aerotransportado, representação de superfícies e finaliza com algumas considerações relativas ao armazenamento de dados.

2.1 Sistemas de varredura a LASER aerotransportados

O princípio de funcionamento de um Sistema de Varredura LASER consiste na medição de distâncias a partir da emissão de pulsos por um sistema óptico-mecânico e a recepção de seus retornos, após terem sido refletidos pela superfície amostrada (Figura 2). Pelo fato de o sistema emitir os pulsos, ou seja, ser independente de fontes de radiação externas (como o Sol, por exemplo), ele é classificado como um sistema ativo. De acordo com Wehr e Lohr (1999), um SVLA pode ser subdividido nas seguintes unidades principais: sistema de varredura a LASER; sistema de posicionamento e orientação; e unidade de controle e processamento. Quando detectados pelo sistema, os pulsos refletidos pela superfície resultam em medidas de distância, cuja orientação é controlada por sistemas óptico-mecânicos.

Figura 2 - Cobertura de uma faixa com um SVLA.



Fonte: Adaptado de Petrie e Toth (2008).

O sistema de varredura a LASER aerotransportado é composto pelo emissor de feixes LASER, sistema de varredura (geralmente composto por espelhos), receptor eletro-óptico e o detector/contador de tempo (WEHR e LOHR, 1999). O posicionamento e a orientação são feitos utilizando DGPS (*Differential Global Positioning System* – Sistema Diferencial de Posicionamento Global) ou RTK (*Real Time Kinematic* – Cinemático em Tempo Real) que opera em conjunto e de modo sincronizado com um INS (*Inertial Navigation System* – Sistema de Navegação Inercial). As observações do sistema GNSS (*Global Navigation Satellite System* – Sistema Global de Navegação por Satélite) contribuem para minimizar os erros sistemáticos do INS e vice-versa.

Existem diversas configurações para o sistema de varredura, o qual é responsável pela orientação do pulso emitido, tradicionalmente em direção transversal³ à trajetória do voo. Os dispositivos comumente empregam soluções baseadas em espelhos oscilantes, polígono de rotação, espelho de nutação (ou Palmer), dentre outros. Cada solução proporciona diferentes padrões de amostragem na superfície, sendo que a escolha apropriada do sistema depende de diversos fatores, como a densidade de pontos desejada, por exemplo.

De acordo com Baltsavias (1999), a medida de distância com LASER pode ser realizada por meio de dois princípios fundamentais: pelo tempo de propagação entre a emissão e a recepção do pulso LASER; e pela diferença de fase. A distância pode ser obtida trivialmente quando o tempo de propagação é conhecido. A medição a partir da diferença de fase demanda desenvolvimentos mais rigorosos.

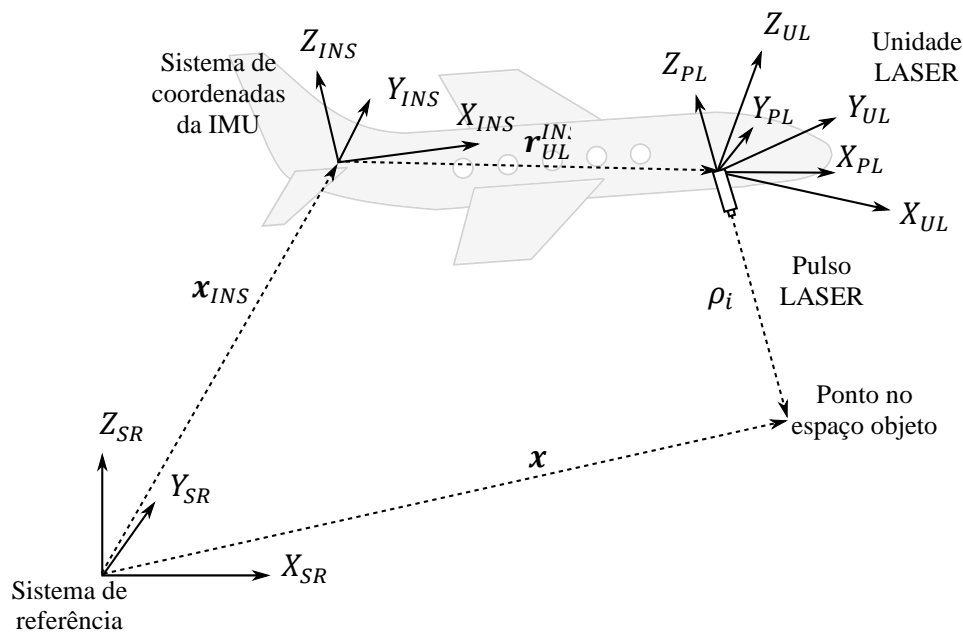
As coordenadas dos pontos são obtidas indiretamente por meio das medidas realizadas pelas diferentes unidades que compõem o sistema (Figura 3). A acurácia da nuvem de pontos resultante dependerá da qualidade com a qual essas medidas foram realizadas. Tornam-se necessários sistemas rigorosos e robustos para a determinação da trajetória da aeronave (posição e orientação), uma vez que as coordenadas finais, isto é, referenciadas a um sistema geodésico, dependem destes valores. Outro fator importante que impacta diretamente na qualidade do produto gerado é a acurácia da medida da distância. Caso a distância esteja sendo obtida pelo princípio do tempo de propagação do pulso, é necessário que o sistema tenha oscilador com resolução adequada⁴.

³ Na prática alguns dispositivos realizam a varredura em direções parcialmente transversais, isto é, não possuem um ângulo reto com a direção da trajetória de voo, por isso é necessário o cuidado com a utilização do termo.

⁴ Um oscilador com resolução de 1 ns proporcionará uma medida de distância com incerteza da ordem de 30 cm.

Genericamente, o modelo matemático (Eq. 1) envolvido na obtenção das coordenadas (\mathbf{x}) do ponto no sistema de referência adotado requer que algumas quantidades sejam registradas no instante da coleta: a medida da distância realizada pela unidade LASER (ρ_i); a posição (\mathbf{x}_{INS}) e atitude (\mathbf{M}_{INS}^{SR}) da plataforma registrada pelo INS; e a orientação a qual o pulso LASER foi emitido (\mathbf{M}_{PL}^{UL}). As demais quantidades devem ser conhecidas *a priori* e obtidas a partir de um procedimento de calibração, são elas: a separação (\mathbf{r}_{UL}^{INS}) e orientação⁵ (\mathbf{M}_{UL}^{INS}) entre o sistema da unidade LASER e o INS. No caso da plataforma contar com receptores GNSS acoplados (não está ilustrado na Figura 3), o processo de calibração deve considerar também a separação entre o centro de fase das antenas e o INS (*lever-arm*).

Figura 3 - Sistemas de coordenadas e parâmetros envolvidos na aquisição de dados SVLA.



Fonte: Adaptado de El-Sheimy *et al.* (2005).

$$\mathbf{x}(t) = \mathbf{x}_{SI}(t) + \mathbf{M}_{INS}^{SR}(t) \cdot \mathbf{M}_{UL}^{INS} (\mathbf{r}_{UL}^{INS} + \mathbf{M}_{PL}^{UL}(t) \cdot \mathbf{r}(t)) \quad (1)$$

sendo $\mathbf{r}^T = [0 \quad 0 \quad \rho_i]$, e \mathbf{M}_A^B a matriz de rotação⁶ que transforma de A para B .

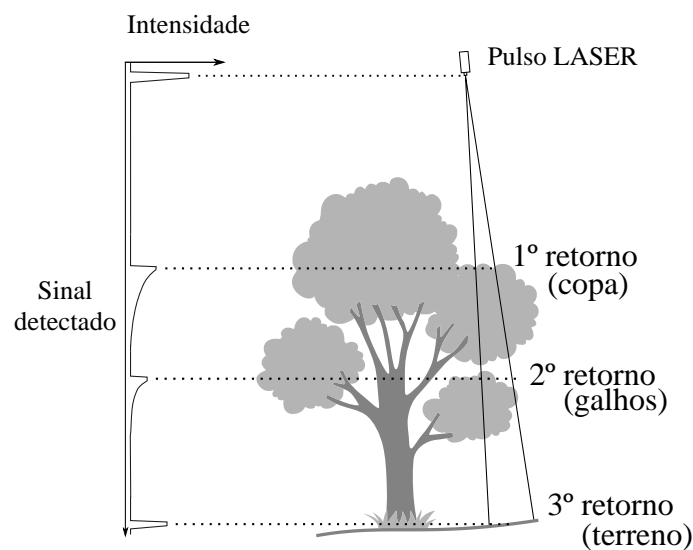
⁵ Os ângulos que descrevem o desalinhamento entre esses sistemas são conhecidos como *boresight*.

⁶ A formulação original presente na literatura pode considerar tanto os ângulos de rotação (*yaw, pitch, roll*) quanto (ω, φ, κ).

2.2 Interação do pulso emitido com a vegetação

A interação do pulso LASER com o dossel da vegetação (Figura 4) pode ser complexa em virtude de dois principais fatores: a vegetação não constitui uma superfície contínua, isto é, existem vazios em sua cobertura, dependendo do espaçamento dos indivíduos arbóreos, suas dimensões, e espécies, por exemplo; e, apesar do rigor empregado pela fabricante no processo de construção do conjunto de lentes dos sistemas o pulso emitido não é perfeitamente colimado, ou seja, atinge uma área (*footprint*) da superfície devido sua divergência⁷. Essas características dão origem a certos fenômenos como a penetração parcial ou total do pulso no dossel, o qual pode resultar em múltiplos retornos para uma única emissão.

Figura 4 - Interação do pulso LASER com a vegetação causando múltiplos retornos.



Fonte: Adaptado de Hyypä *et al.* (2008).

Caso o SVLA utilizado faça o registro dos múltiplos retornos refletidos, os dados coletados podem fornecer informações úteis e detalhadas para aplicações em inventário florestal, por exemplo. O interesse dessa tecnologia para a aplicação mencionada origina-se na

⁷ Alguns instrumentos incluem a possibilidade de configurar a divergência do pulso LASER.

possibilidade de inferir sobre algumas variáveis dendrométricas, o que é menos vantajoso quando obtidos por métodos aerofotogramétricos tradicionais⁸.

2.3 Representação de superfícies

Na literatura existe uma certa divergência com relação aos conceitos associados à terminologia usual para a modelagem digital da superfície terrestre (El-Sheimy *et al.*, 2005). Assim como observado por alguns autores, o conceito de terreno é subjetivo, podendo acomodar diferentes nuances. Neste trabalho, os termos “terreno”, “chão” e “solo” serão entendidos como sinônimos, definidos conforme o conceito proposto por Sithole e Vosselman (2004): superfície contínua delimitadora entre o material sólido da crosta terrestre e o ar. A partir desta definição algumas camadas finas de pavimento rígido são entendidas como terreno, tais como asfalto e concreto, por exemplo, entretanto, viadutos são considerados como objetos.

De acordo com Li *et al.* (2005), pode-se entender o MDS como a superfície do terreno acrescida de quaisquer feições existentes sobre ela, como vegetação e edificações (feições antrópicas), por exemplo. Uma definição genérica para o MDT é dada por Pfeifer e Mandlbürger (2008), a qual é independente da estrutura de dados: o MDT é uma função contínua que mapeia a posição bidimensional (planimétrica) para a elevação do terreno, ou seja: $z=f(x,y)$. De acordo com os autores, a utilização ou não de linhas de quebra (*breaklines*) é arbitrária, dependendo do nível de detalhamento almejado. Outra definição é apresentada por Peckham e Jordan (2007), na qual os autores entendem o MDT como um arranjo ordenado de números que representam a distribuição espacial dos atributos do terreno.

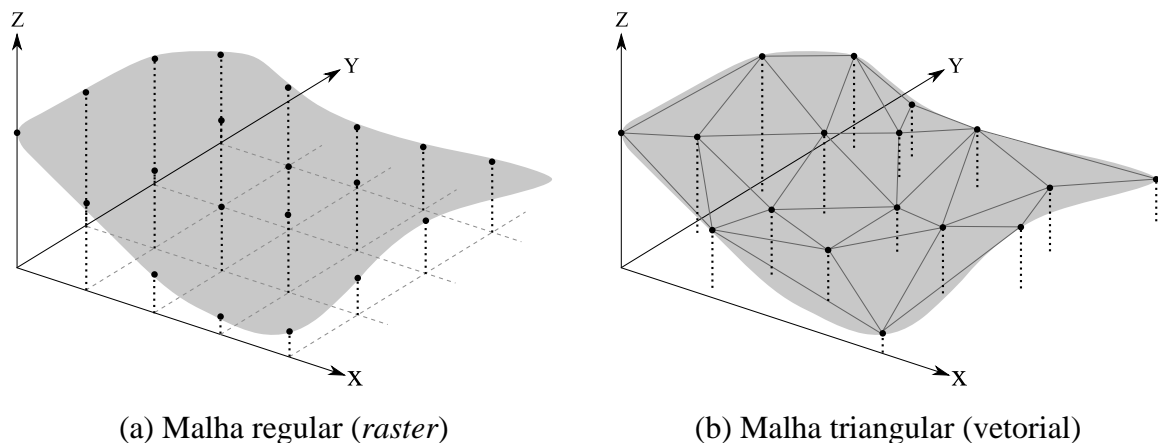
O termo “elevação” expressa uma medida de altitude relativa a um datum vertical (Li *et al.*, 2005), podendo ser tanto de um ponto no terreno quanto em um objeto sobre ele⁹. Entretanto, alguns autores (El-Sheimy *et al.* (2005); Li *et al.* (2005); e, Pfeifer e Mandlbürger (2008)) tomam os modelos digitais de elevação (MDE) como sinônimos de MDT, ou seja, limitando esta representação aos pontos do terreno. Do ponto de vista semântico, esta associação é incorreta, uma vez que é incompleta por excluir a representação dos objetos sobre o terreno. Outro conceito discutível origina-se ao admitir MDE como arranjos regulares

⁸ Cita-se especificamente a Aerofotogrametria pois, a Fotogrametria Terrestre pode ser utilizada para a geração de inventário florestal, entretanto, com menor produtividade quando comparada ao SVLA, em vista da área recoberta.

⁹ De acordo com o significado apresentado nos dicionários Aurélio e Cambridge (neste considera-se o equivalente em inglês – *elevation*), o termo “elevação” não é restrito ao terreno.

(padrões retangulares ou hexagonais) de elevações, assim como apresentado por El-Sheimy *et al.* (2005). Nesta situação limita-se o conceito pela estrutura de dados utilizada (Figura 5). Dos pontos de vista teórico e conceitual é preferível que as definições para MDE, MDT e MDS evitem as possíveis incoerências semânticas, e que não estejam limitadas às estruturas de dados utilizadas, assim como ocorre nas concepções apresentadas anteriormente. Este trabalho adota as definições citadas, com a exceção das restrições discutidas.

Figura 5 - Exemplos de estruturas comumente empregadas na representação de superfícies.



Fonte: Adaptado de Namikawa *et al.* (2003).

Existem ainda outras terminologias encontradas principalmente em estudos florestais, tais como o modelo de altura do dossel (MAD, do inglês *canopy height model* - CHM) e modelo digital do dossel (MDD, do inglês *digital canopy model* - DCM). Pfeifer e Mandlburger (2008) citam o uso do MDD e MDS como sinônimos, e embora este último tenha um âmbito mais genérico, ambos podem ser obtidos a partir do primeiro retorno do pulso LASER. Os autores ainda definem o MDSn como a diferença entre o MDS e o MDT, isto é: $MDSn_{(x,y)} = MDS_{(x,y)} - MDT_{(x,y)}$, neste caso as altitudes registradas passam a ser as alturas dos objetos com relação ao terreno, sendo esta representação designada por alguns autores como MDS normalizado.

O processo de filtragem consiste basicamente em uma classificação binária entre pontos de terreno e não-terreno (objetos), ou sucintamente, a obtenção do MDT a partir do MDS. O problema é formalmente apresentado na sequência: considere um conjunto \mathbf{P} de pontos $\mathbf{p}_i = [x_i \ y_i \ z_i \ \mathbf{a}_i]^T, i = 1 \dots n$, sendo \mathbf{a}_i o vetor com os demais atributos disponíveis (intensidade do retorno, cor obtida a partir de uma imagem digital, por exemplo), podendo este

ser inexistente. Assuma também os conjuntos de pontos de terreno (\mathbf{P}_T) e não-terreno (\mathbf{P}_{NT}) tal que a união dos dois subconjuntos resulte no conjunto original ($\mathbf{P}_T \cup \mathbf{P}_{NT} = \mathbf{P}$) e a intersecção entre eles resulte no conjunto vazio ($\mathbf{P}_T \cap \mathbf{P}_{NT} = \emptyset$). Nestas condições deve existir ao menos um critério de decisão $C(\mathbf{p}_i)$, que permita rotular o conjunto total de pontos:

$$C(\mathbf{p}_i) = \begin{cases} 1 & \text{se } \mathbf{p}_i \in \mathbf{P}_T \\ 0 & \text{se } \mathbf{p}_i \in \mathbf{P}_{NT} \end{cases} \quad i = 1 \dots n \quad (2)$$

As estruturas de dados comumente utilizadas na representação de superfícies são o formato matricial (malha regular, grade ou *raster*) e as malhas triangulares. As malhas regulares compõem estruturas discretas, cujos elementos estão alinhados e regularmente espaçados. Tais características permitem a aplicação de algoritmos do campo de processamento digital de imagens, cujas operações de vizinhança podem ser executadas rapidamente. Entretanto, a geração de malhas regulares requer a interpolação dos dados originais, acarretando perda de precisão e de informação. Algumas possibilidades incluem a interpolação por vizinho mais próximo, inverso da potência da distância, krigagem, entre outras.

De acordo com Floriani e Puppo (1989), as malhas triangulares exercem funções importantes na representação de superfícies. Estas estruturas apresentam algumas características vantajosas como a adaptação à densidade dos dados e a facilidade de atualização quando a inserção ou remoção de alguns vértices se fazem necessárias. Dentre as diversas possibilidades, a triangulação de Delaunay é considerada como uma das mais apropriadas, tendo em vista que os triângulos satisfazem as condições de equiangularidade, uma propriedade importante para tarefas de interpolação, por exemplo. O problema envolvido na construção da malha triangular com o critério de Delaunay é conhecido como *constrained Delaunay triangulation* (CDT), o qual pode ser resolvido de diferentes maneiras, como por exemplo os algoritmos de divisão e conquista, *sweep*, e *flip* (CHENG, S.; DEY, T. K.; SHEWCHUK, J. R., 2012)

2.4 Considerações sobre o armazenamento dos dados

Com o objetivo de facilitar as operações de transferência de dados SVLA (nuvens de pontos) a *American Society for Photogrammetry and Remote Sensing* (ASPRS) criou um formato público de arquivos binários (LAS) como alternativa aos arquivos de texto (ASCII –

American Standard Code for Information Interchange) e aos formatos proprietários. A adoção do formato pelas empresas do segmento permite que os dados sejam coletados e operados a partir de diferentes equipamentos e sistemas computacionais. O padrão comporta o registro de diversos metadados para cada retorno dos pulsos LASER, assim como pode ser visto no Quadro 2. Além do formato LAS, a ASPRS padronizou códigos utilizados para rotular os pontos, como por exemplo: terreno (2), edificações (6), entre outros¹⁰.

Quadro 2 – Exemplos de metadados armazenadas para cada retorno no formato LAS.

| Item | Tamanho | Tipo |
|-----------------------------------|-----------------------------|---------|
| Coordenadas (X,Y,Z) | 12 bytes | Inteiro |
| Intensidade do retorno | 2 bytes | Inteiro |
| Número do retorno | De 3 a 4 bits ¹¹ | Inteiro |
| Quantidade de retornos | De 3 a 4 bits | Inteiro |
| Rótulo (classificação) | 1 byte | Inteiro |
| Tempo GPS | 8 bytes | Real |
| Coloração RGB + NIR ¹² | 8 bytes | Inteiro |

Fonte: Adaptado de ASPRS (2013).

O cabeçalho do arquivo contém diversos metadados, dentre eles a quantidade de registros (pontos) e os parâmetros de transformação (três fatores de escala e três translações), conforme Eq. 3.

$$\begin{bmatrix} X_i \\ Y_i \\ Z_i \end{bmatrix} = \begin{bmatrix} S_X & 0 & 0 \\ 0 & S_Y & 0 \\ 0 & 0 & S_Z \end{bmatrix} \begin{bmatrix} X_i^L \\ Y_i^L \\ Z_i^L \end{bmatrix} + \begin{bmatrix} T_X \\ T_Y \\ T_Z \end{bmatrix} \quad (3)$$

Os fatores de escala (S_X , S_Y , S_Z) são responsáveis pela conversão das coordenadas durante as operações de leitura e escrita no arquivo. Como exemplo, o fator de escala deverá ser igual a 0,001 para a leitura de coordenadas armazenadas com três casas decimais. As translações (T_X , T_Y , T_Z) não são mandatórias, entretanto devem ser sempre consideradas (mesmo que sejam nulas), possibilitando transformar as coordenadas para um sistema local.

¹⁰ A vegetação é dividida em: (3) baixa (até 0,3 m); (4) média (0,3 à 2,0 m); e (5) alta (acima de 2,0 m).

¹¹ Alguns formatos utilizam 3 bits enquanto os mais novos permitem 4 bits.

¹² A componente *near infra-red* (NIR) foi introduzida na versão 1.4 do formato LAS.

A ordem de grandeza das coordenadas em certos sistemas de projeção, como na Projeção Universal Transversa de Mercator¹³, por exemplo, associada a representação aproximada dos números reais feita pelos computadores pode causar mau condicionamento na solução numérica de alguns problemas matemáticos (como sistemas de equações). Nestes casos, isto é, aplicações práticas que envolvem a implementação de rotinas para processamento de coordenadas, a adoção de sistemas de coordenadas locais proporciona alguns benefícios.

Existem diversas possibilidades para a representação de números reais em dispositivos digitais, dentre elas a de ponto fixo e a de ponto flutuante. Em 1985 o Instituto de Engenheiros Eletricistas e Eletrônicos (IEEE) definiu o padrão IEEE 754 (revisões em 2008 e 2019)¹⁴ o qual dita as regras de representação de números reais a partir de ponto flutuante em circuitos eletrônicos. A representação de números reais em ponto flutuante é derivada da notação científica, onde tem-se um sinal (positivo ou negativo), um coeficiente e um expoente. O padrão IEEE 754-1985 define as representações de precisão simples (*float*, 32 bits) e precisão dupla (*double*, 64 bits) que são adotados por diversas arquiteturas de computadores e linguagens de programação.

Tomando como exemplo a representação da coordenada N da estação geodésica SAT 92579/1995¹⁵ em notação científica, uma opção seria: $+0,7553951875 \times 10^7$ m, correspondente a 7553951,875 m. Por conta da ordem de grandeza da coordenada (milhões de metros) são necessários no mínimo 10 algarismos (7553951875) para poder representar o valor do coeficiente com precisão de milímetros. Além disso, é mandatória a utilização do expoente (potência de dez, neste caso) e o sinal do valor, este último requer apenas um bit.

Assim, como pode ser visto na Figura 6, só é possível representar corretamente o valor da coordenada utilizando ponto flutuante de precisão dupla. Isso ocorre devido à quantidade limitada de bits reservados para representar o coeficiente (23 bits no caso de precisão simples contra 53 na precisão dupla). Entretanto, a utilização de precisão dupla implica no dobro de memória alocada para armazenar esta coordenada.

Este simples exemplo permite perceber um problema computacional, ao tentar responder a seguinte questão: a representação em precisão dupla seria suficiente para computar N ao cubo? Como se percebe, a resposta é não. De modo geral, observa-se que quanto maior o valor a ser

¹³ Utilizada no mapeamento sistemático do Brasil.

¹⁴ IEEE 754-2019 *Standard for Floating-Point Arithmetic (Revision of 754-2008)*. pp. 1-84. Julho, 2019. DOI: [10.1109/IEEESTD.2019.8766229](https://doi.org/10.1109/IEEESTD.2019.8766229)

¹⁵ Disponível em: <http://www.bdg.ibge.gov.br/bdg/pdf/relatorio.asp?L1=92579>. Acessado em: 10 nov 2020.

armazenado, menor será a precisão obtida pela representação em ponto flutuante, o que de certa maneira reitera a importância da utilização de sistemas de coordenadas locais. A principal desvantagem referente à adoção destes sistemas resume-se aos processamentos adicionais necessários para a transformação (direta e inversa). Embora seja necessário arbitrar uma origem para o sistema local, esta origem pode ser definida de diferentes modos, como por exemplo pela adoção do centro de massa das coordenadas a serem processadas.

Figura 6 - Representação em notação científica da coordenada N da estação geodésica SAT 92579/1995.

| | | | |
|------------------------------|---------------|---------------|---------|
| $+7,553952000 \times 10^6$ m | 7553952,000 m | <i>float</i> | 32 bits |
| $+7,553951875 \times 10^6$ m | 7553951,875 m | <i>double</i> | 64 bits |

Valores armazenados em ponto flutuante utilizando computador com arquitetura x86_64.

Fonte: Elaborado pelo autor.

Capítulo 3

RECONHECIMENTO DE PADRÕES

Este capítulo apresenta algumas das técnicas de reconhecimentos de padrões e métricas existentes na literatura que, no escopo deste trabalho, podem ser utilizadas para inferir sobre determinado contexto geométrico ou topológico das estruturas e objetos amostrados na superfície terrestre.

De acordo com Novo (2008), as técnicas utilizadas para atribuir significado a um indivíduo em função de suas propriedades são conhecidas como classificação. A associação ou agrupamento de indivíduos a partir de suas propriedades é realizada a partir de técnicas matemáticas conhecidas como reconhecimento de padrões e aprendizado de máquina (HAYKIN, 2009). Os processos de classificação podem ser supervisionados ou não, de acordo com o grau de intervenção do operador. A distinção entre classificação supervisionada e não-supervisionada consiste na utilização ou não de dados amostrais de treinamento (RICHARDS, 2013). Os dados de treinamento são necessários para a determinação de descritores (vetores de atributos) para as respectivas classes, os quais serão comparados aos descritores dos objetos, a partir de diferentes medidas de distância¹⁶, almejando a associação de cada um dos objetos a uma classe. Bishop (1995) distingue os problemas de reconhecimento de padrões em: classificação, quando deseja-se associar os dados de entrada à uma classe, dentre um conjunto discreto (finito) de possibilidades; e regressão, para as situações em que a associação dos dados de entrada é feita com uma ou mais variáveis contínuas. Ainda de acordo com este autor, ambos os problemas mencionados compõem casos de aproximação de funções.

3.1 Coeficiente de sobreposição entre duas curvas Gaussianas

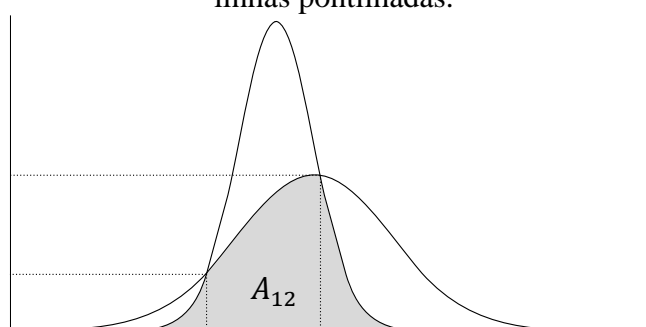
Os métodos de reconhecimento de padrões supervisionados almejam a obtenção de informações - a partir de dados amostrais - que permitam caracterizar as classes e inferir univocamente a qual delas os indivíduos pertencem. A seleção de quais atributos devem ser empregados na classificação é uma tarefa complexa, entretanto, o conhecimento sobre o potencial de separabilidade que cada um deles proporcionará no resultado pode contribuir na

¹⁶ As técnicas de classificação empregam diferentes medidas de similaridade para a tomada de decisões (rotulação dos objetos nas diferentes classes).

escolha. Neste sentido, o coeficiente de sobreposição (Figura 7) entre duas distribuições normais, e suas respectivas curvas Gaussianas, pode ser utilizado como um indicador do potencial de separabilidade entre dois atributos.

As classes podem ser caracterizadas numericamente a partir do vetor média \boldsymbol{x} e da matriz de covariâncias \boldsymbol{S} das k variáveis selecionadas, obtidos a partir das amostras de treinamento. Assumindo que as variáveis possuam comportamento normal¹⁷ é possível avaliar o potencial de separabilidade (para cada uma delas, configurando uma análise univariada) entre duas classes C_1 e C_2 . Segundo Dhaker *et al.* (2017), uma das maneiras¹⁸ de avaliar a separabilidade entre duas distribuições normais é a partir do coeficiente de sobreposição (a sigla utilizada na literatura internacional é: OVL – *overlapping coefficient*). Diversos estudos foram realizados nas últimas décadas nos quais os autores propuserem diferentes medidas.

Figura 7 - Sobreposição entre duas funções de distribuição de probabilidade normais com médias e variâncias diferentes. Os pontos em que elas se interceptam estão indicados pelas linhas pontilhadas.



Fonte: Adaptado de Inman e Bradley (1988).

Os pontos de intersecção entre duas curvas Gaussianas podem ser obtidos a partir da solução de uma simples equação quadrática. O coeficiente de sobreposição entre duas curvas Gaussianas (Δ) proposto por Weitzman (1970¹⁹ *apud* INMAN e BRADLEY, 1988) para o caso univariado é igual ao dobro da área em comum entre as duas curvas (A_{12}) dividido pela soma das áreas destas curvas ($A_1 + A_2$):

¹⁷ Existem diversos testes estatísticos para verificar a suposição de normalidade dos dados, dentre eles o de Shapiro-Wilk é um dos mais usuais.

¹⁸ Outras possíveis medidas são a distância de Bhattacharyya e a distância de Mahalanobis, por exemplo.

¹⁹ WEITZMAN, M. S. *Measures of overlap of income distributions of white and Negro families in the United States*. Relatório Técnico, Department of Commerce, Washington, D. C. 1970.

$$\Delta = \frac{2 A_{12}}{A_1 + A_2} \in [0,1] \quad (4)$$

Entretanto, sabe-se que: $\lim_{x \rightarrow \infty} \Phi(x, \mu, \sigma) = 1$, isto é, a função de distribuição de probabilidade normal acumulada tende a um (1) quando x tende ao infinito, ou seja, a área total de qualquer curva normal é unitária. Deste modo pode-se escrever:

$$A_1 = A_2 = 1 = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \sigma^{-1} \exp\left\{\frac{-(x - \mu)^2}{2\sigma^2}\right\} dx \quad \forall \mu, \sigma \in \mathbb{R} \quad (5)$$

logo, a soma das áreas totais das distribuições é igual a dois ($A_1 + A_2 = 2$), e, portanto:

$$\Delta(\mu_1, \sigma_1, \mu_2, \sigma_2) = A_{12} = \int_{-\infty}^{\infty} \min\{P(x, \mu_1, \sigma_1), P(x, \mu_2, \sigma_2)\} dx \quad (6)$$

Em termos práticos o resultado numérico pode ser obtido por diversas abordagens. As soluções mais robustas fazem uso de integração numérica. Algumas propriedades particulares do coeficiente de sobreposição são apresentadas por Inman e Bradley (1988), sendo que uma delas trata da invariância do resultado quando uma única transformação $g(x)$ é aplicada nas duas distribuições, isto é:

$$\Delta(\mu_1, \sigma_1, \mu_2, \sigma_2) = \int_{-\infty}^{\infty} \min\{P(g(x), \mu_1, \sigma_1), P(g(x), \mu_2, \sigma_2)\} dx \quad (7)$$

Esta propriedade é útil em situações em que as variáveis em questão apresentem mal condicionamento. Dessa maneira, o uso dessa propriedade pode reduzir o erro na integração numérica da solução.

3.2 Discriminantes lineares

A tarefa de distinguir indivíduos entre duas categorias constitui uma operação básica, quando se trata de reconhecimento de padrões. Em sua instância mais simples, quando as classes são linearmente separáveis, é possível aplicar funções discriminantes lineares $y(\mathbf{x})$ de modo que o indivíduo caracterizado pelo vetor \mathbf{x} (em um espaço k-dimensional) seja

classificado em C_1 se $y(\mathbf{x}) > 0$ e em C_2 se $y(\mathbf{x}) < 0$. A função mais simples²⁰ que permite essa separação pode ser escrita como (BISHOP, 1995):

$$y(\mathbf{x}) = \mathbf{w}^T \mathbf{x} + w_0 \quad (8)$$

em que o vetor \mathbf{w} contém pesos aplicados na transformação de mapeamento (projeção) entre os espaços $\mathbb{R}^k \rightarrow \mathbb{R}$, e w_0 é uma translação realizada no sistema resultante de modo que a origem (máxima separação entre as classes) seja no zero (origem).

Esta função discriminante linear corresponde ao hiperplano definido no espaço (k-1)-dimensional²¹. Assim como pode ser visualizado na Figura 8, existem infinitas soluções para os coeficientes (valores dos pesos e da translação), o que torna necessário adotar um ou mais critérios (restrições) durante o treinamento do classificador, de modo que seja possível resolver numericamente o problema. Uma das abordagens mais comuns consiste na aplicação da técnica dos multiplicadores de Lagrange²², a qual pode ser utilizada para encontrar o mínimo de uma função sujeito a uma ou mais restrições. Outras abordagens incluem a solução pelo método dos mínimos quadrados (MMQ) ou por gradiente descendente (*gradient descent*, do inglês).

Fisher (1936) apresentou uma possível solução para a determinação dos coeficientes baseada na técnica dos multiplicadores de Lagrange. A função discriminante resultante consiste em uma projeção linear que maximiza a separação entre as classes²³. Considerando que as classes C_1 e C_2 apresentem vetores médios $\bar{\mathbf{x}}_1$ e $\bar{\mathbf{x}}_2$, e matrizes de covariâncias \mathbf{S}_1 e \mathbf{S}_2 , calculados a partir de n_1 e n_2 indivíduos de treinamento, respectivamente, os coeficientes da função discriminante linear podem ser obtidos fazendo:

$$\mathbf{w} = \mathbf{S}_{ag}^{-1} (\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2) \quad (9)$$

$$w_0 = -\frac{1}{2} (\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2)^T \mathbf{S}_{ag}^{-1} (\bar{\mathbf{x}}_1 + \bar{\mathbf{x}}_2) - \ln \left[\frac{c\{1|2\}}{c\{2|1\}} \frac{p(2)}{p(1)} \right]$$

²⁰ Também conhecido por *threshold logic unit* (TLU) por realizar *hard thresholding*.

²¹ Na classificação de indivíduos caracterizados por vetores bidimensionais (k=2), isto é, $\mathbf{x}^T = [x_1 \ x_2]$, o hiperplano resultante corresponde a uma reta.

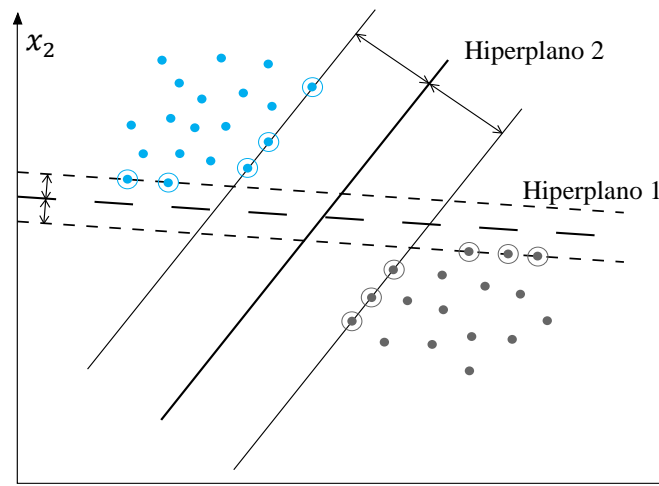
²² Também conhecidos como multiplicadores indeterminados de Lagrange.

²³ Para a classificação de indivíduos entre duas classes apenas, a solução utilizando os critérios propostos por Fisher (1936) pode ser obtida como um caso especial da solução pelo MMQ.

em que $c\{a|b\}$ representa o custo de classificar erroneamente o indivíduo na classe a , quando sabe-se que ele pertence à b ; e $p(a)$ é a probabilidade a priori do indivíduo ser da classe a . Nesta situação assume-se que as matrizes de covariâncias são estatisticamente equivalentes, e uma matriz de covariâncias agrupada pode ser obtida fazendo:

$$\mathbf{S}_{ag} = \frac{(n_1 - 1)\mathbf{S}_1 + (n_2 - 1)\mathbf{S}_2}{n_1 + n_2 - 2} \quad (10)$$

Figura 8 - Classes linearmente separáveis. Ilustração de duas possíveis soluções para o hiperplano.



Fonte: Adaptado de Theodoridis e Koutroumbas (2003).

3.3 Redes neurais artificiais

As RNA²⁴ compreendem modelos estudados em uma das vertentes do reconhecimento de padrões. Tais modelos buscam simular o funcionamento básico do sistema nervoso central (cérebro) presentes nos animais e humanos, e podem ser aplicados em diversas tarefas complexas de classificação supervisionada. De acordo com Haykin (1999), as RNA podem ser entendidas como processadores²⁵ paralelos os quais são compostos por unidades de processamento menores (neurônios) que são capazes de armazenar conhecimento e torná-lo disponível para uso. Ainda de acordo com o autor, elas se assemelham com o cérebro em dois

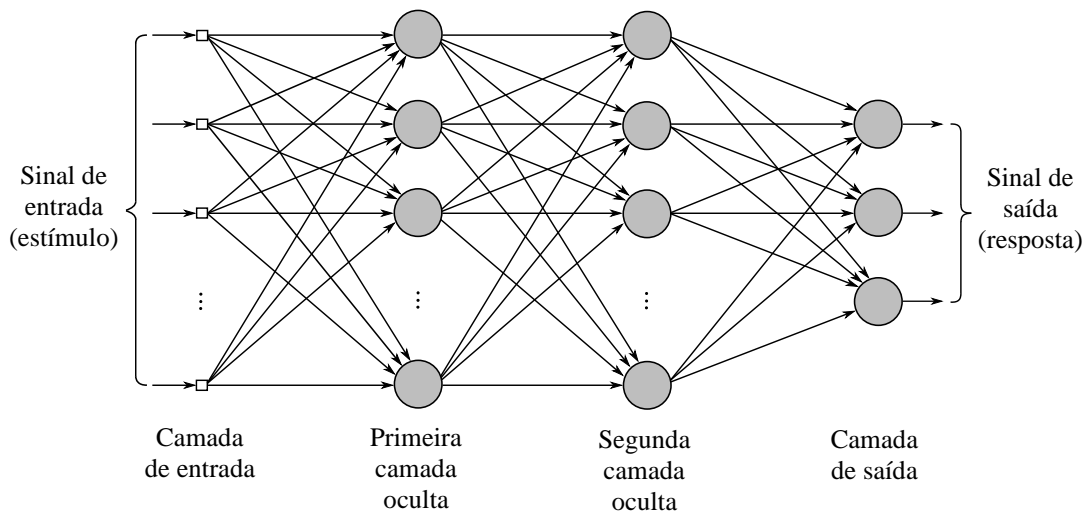
²⁴ Também conhecidas como *multi-layer perceptrons* (MLP), isto é, perceptron de múltiplas camadas.

²⁵ No sentido de que fornece uma resposta única a partir de um determinado estímulo.

aspectos principais: a rede adquire conhecimento a partir de um processo de aprendizado (treinamento); e este conhecimento é armazenado pelas conexões entre os neurônios (pesos).

Existem diversos tipos de RNA, tais como perceptron de camada única, utilizado como classificador linear, e os perceptrons de múltiplas camadas, mais comuns e que são empregados em tarefas de predição e classificações não lineares (SAMARASINGHE, 2007). A topologia empregada na configuração de tais redes (Figura 9) usualmente considera o formato *feed-forward*, isto é, adota-se a restrição de que a transferência dos sinais ocorre apenas em uma direção: da entrada para a saída (BISHOP, 1995).

Figura 9 - Topologia *feed-forward* para um perceptron de múltiplas camadas.



Fonte: Adaptado de Haykin (2009).

3.3.1 Funções de ativação

As funções discriminantes lineares resultam em uma classificação binária, entretanto, a utilização em conjunto com as funções de ativação $g(\cdot)$ generaliza o modelo matemático e proporciona um melhor tratamento para o problema. Embora o limite de decisão ainda seja linear, a aplicação destas funções permite que o resultado seja interpretado como uma probabilidade a posteriori (BISHOP, 1995). De acordo com Richards (2013) a utilização de funções de ativação permite que a RNA consiga lidar com classes que não sejam linearmente separáveis, as quais não seriam processadas corretamente apenas com limiares. Desta forma, pode-se reescrever a Eq. 8 considerando uma função de ativação:

$$y(\mathbf{x}) = g(\mathbf{w}^T \mathbf{x} + w_0) \quad (11)$$

Segundo Theodoridis e Koutroumbas (2003), o principal problema envolvido na formulação das RNA consiste na descontinuidade da função de ativação implícita no modelo matemático discriminante mais básico. Esta formulação é decorrente da função degrau (também conhecida por função de Heaviside²⁶) não permitir a operação de diferenciação, a qual é necessária para aplicar o algoritmo de retropropagação (*backpropagation*).

$$g(a) = \begin{cases} 1 & \text{se } a \geq 0 \\ 0 & \text{se } a < 0 \end{cases} \quad (12)$$

Alternativamente é possível adotar funções contínuas (e, portanto, diferenciáveis) que se aproximam da função degrau. Uma possibilidade é a função discriminante logística sigmóide²⁷ como função de ativação, assim como ilustra a Figura 10. Este modelo matemático mapeia os valores de entrada para o intervalo $[0,1]$. Outra função usual é a tangente hiperbólica, a qual apresenta convergência mais rápida no treinamento da rede (BISHOP, 1995), uma vez que esta função mapeia os valores de entrada para o intervalo $[-1,1]$.

$$g(x) = \frac{1}{1 + e^{-ax}} \quad (13)$$

ou:

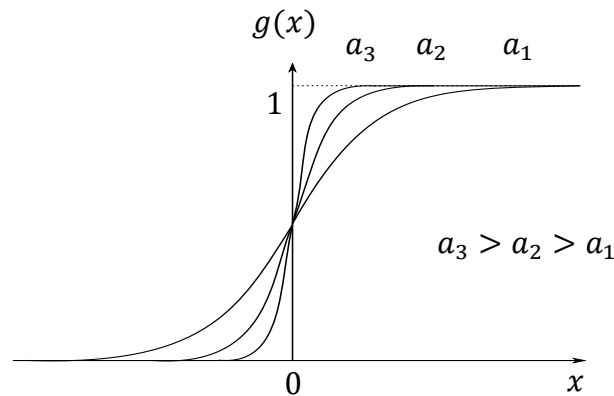
$$g(x) = \tanh(ax) = \frac{e^{ax} - e^{-ax}}{e^{ax} + e^{-ax}} \quad (14)$$

Assim como os discriminantes lineares, o processo de treinamento de uma RNA requer a determinação de todos os pesos das camadas ocultas e da camada de saída simultaneamente. Esta tarefa requer estímulos e respostas conhecidos, de modo que permita avaliar o erro médio quadrático (EMQ) obtido pelo modelo (GONZALEZ e WOODS, 2007). Após este procedimento é possível utilizar o modelo para obter respostas de problemas que forneçam outros estímulos análogos aos de treinamento, desde que estes estejam formatados e padronizados nas mesmas condições.

²⁶ Denominação em homenagem ao pesquisador Oliver Heaviside (1850-1925).

²⁷ Denominação sigmóide decorrente do formato de “S” da curva do gráfico desta função.

Figura 10 - Aproximação da função degrau a partir da função logística sigmóide.



Fonte: Adaptado de Theodoridis e Koutroumbas (2003).

Considerando RNAs com topologias *feed-forward* o sinal de saída calculado para o k -ésimo neurônio da j -ésima camada é dado pela soma ponderada de todos os seus sinais de entrada, ou seja:

$$y_k = \sum_{i=1}^m w_{ij} g(y_i) \quad (15)$$

em que os m neurônios da camada anterior com sinais y_i estão conectados ao neurônio atual e a tais conexões estão associados aos pesos w_{ij} . Na prática este processo consiste em um cálculo progressivo, uma vez que são necessários os valores das camadas anteriores. Para a determinação desses pesos é possível aplicar o algoritmo *backpropagation*, o qual inicializa os valores a partir de uma distribuição pseudoaleatória. Estes pesos são então utilizados no processo de inferência²⁸ com os dados de treinamento, o que permite o cálculo do erro médio quadrático (GONZALEZ e WOODS, 2007) por:

$$\varepsilon = \frac{1}{2} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (16)$$

sendo n a quantidade de neurônios da camada de saída, y_i e \hat{y}_i os sinais calculado e esperado, respectivamente, para o i -ésimo neurônio da camada de saída. Em termos práticos essa métrica

²⁸ O processo de inferência em uma RNA é também conhecido por *forward propagation* devido ao fluxo em que os cálculos são realizados, isto é, da primeira à última camada.

serve como indicativo da qualidade da inferência realizada pela RNA (uma função de custo que deve ser minimizada): caso o EMQ calculado tenha sido grande, significa que são necessárias severas mudanças nos pesos associados às conexões. Por outro lado, se a inferência resultou em valores próximos aos esperados, pouca ou nenhuma mudança deve ser realizada. Assim, os pesos podem ser atualizados a partir da seguinte função:

$$w_{ij}^{(t)} = w_{ij}^{(t-1)} + \Delta w_{ij} \quad (17)$$

As correções para os valores dos pesos (Δw_{ij}) são calculadas a partir dos gradientes das funções de custo e de ativação. A velocidade com que o método atinge convergência depende da constante η , a qual depende do problema em questão e do tipo da função de ativação (BISHOP, 1995).

$$\Delta w_{ij} = \eta \delta_j c_i \quad (18)$$

em que δ_j é o gradiente calculado a partir da função de custo, função de ativação, e dos dados de treinamento. Alguns problemas podem ocorrer nesta solução devido à seleção equivocada da topologia da RNA, valores iniciais dos pesos, e dados de treinamento insuficientes ou problemáticos. Estas situações podem ocasionar em não convergência do algoritmo.

3.3.2 Redes Neurais Convolucionais

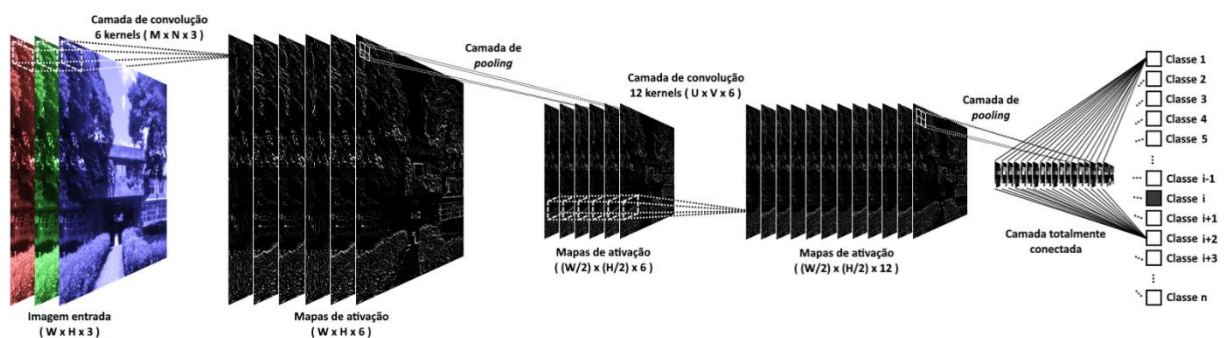
Dentre outros motivos, a evolução no poder de processamento dos computadores²⁹ e o contínuo desenvolvimento das áreas de visão computacional, processamento de imagens e estatística, culminaram no surgimento das RNC. De acordo com Goodfellow *et al.* (2016) as RNC surgiram com o objetivo de aproveitar a estrutura matricial presente em certos dados de entrada, tais como imagens e malhas regulares de pontos (obtidas a partir de nuvens de pontos reamostrada, por exemplo). De forma geral, este tipo de rede neural apresenta o formato *feed-forward* e pode conter uma ou mais camadas de convolução seguidas por camadas completamente conectadas, tais como as encontradas em RNA convencionais, assim como

²⁹ Principalmente a evolução das *graphics processing units* (GPU) a partir dos anos 2000. Atualmente utilizadas como processadores de propósito geral.

ilustrado na Figura 11. O objetivo por trás do desenvolvimento desta categoria de redes neurais é a redução da quantidade de parâmetros de treinamento, uma vez que a camada de entrada irá conter tantos neurônios quanto são os elementos do conjunto de dados.

O conceito de redes neural convolucional foi introduzido originalmente por Fukushima (1980) com o objetivo de identificar automaticamente padrões (estímulos visuais) a partir de similaridade geométrica em imagens matriciais. Este modelo apresentou um lento progresso nas primeiras duas décadas, mas tem despertado a atenção e sido amplamente disseminado, devido a dois fatores: (i) avanços importantes nos algoritmos, tais como os propostos nos trabalhos notáveis de LeCun *et al.* (1998) e Krizhevsky *et al.* (2012), por exemplo; e (ii) hardware, software, e conjuntos de dados necessários para o processamento estar acessível para mais usuários e instituições de pesquisa.

Figura 11 – Exemplo de rede neural convolucional para classificação de imagens coloridas.



Fonte: Elaborado pelo autor.

Uma das vantagens das RNC em relação às RNA é a habilidade dessas redes “aprenderem” e “extraírem” os próprios atributos, o que é conhecido como *feature learning* ou *representation learning*. A idéia principal, como apontada por Fukushima (1980), é simular o córtex visual encontrado no cérebro. Esta característica é importante visto que dispensa o processo manual de definição dos atributos pelo analista. No entanto, conforme a topologia da rede torna-se mais complexa, como no caso das RNC, o número de parâmetros tende a aumentar consideravelmente, necessitando de mais processamento e conjuntos de dados maiores (AMIRKOLAEI e AREFI, 2019).

De acordo com Gonzalez e Woods (2007), a operação de convolução pode ser entendida como um processo similar ao de filtragem espacial linear, que depende da definição prévia de uma matriz de pesos. Esta matriz de pesos é geralmente chamada de filtro, máscara, *kernel*, ou

template, dependendo da literatura (GONZALEZ e WOODS, 2007). Considerando uma imagem f de dimensões $M \times N$ a ser filtrada por uma máscara w de dimensão $m \times n$, a resposta obtida para o pixel (x, y) é obtida por:

$$g(x, y) = \sum_{s=-a}^a \sum_{t=-b}^b w(s, t) f(x + s, y + t) \quad (19)$$

sendo $a = (m - 1)/2$, e $b = (n - 1)/2$. A imagem resultante (filtrada) g é obtida ao realizar a operação em todos os elementos da imagem f , ou seja, para $x = 0, 1, 2, \dots, M - 1$ e $y = 0, 1, 2, \dots, N - 1$. Esta operação básica de convolução permite gerar diversos resultados importantes, tais como extração de feições de interesse, suavização, remoção de ruído, entre outros. Alguns operadores comuns são o de Prewitt e o de Sobel, os quais são utilizados para a extração de bordas, e até mesmo em algumas etapas de métodos como o de Canny, Harris e Förstner, por exemplo.

No contexto de redes neurais convolucionais, a imagem resultante da operação de convolução é chamada de mapa de ativação. Entretanto, as RNC utilizam-se de uma forma mais generalizada da convolução, a qual opera em tensores³⁰. É importante entender o conceito de tensor neste contexto para evitar ambiguidade com o conceito utilizado na física. Tensor, no contexto de aprendizado de máquina, é entendido como uma generalização dos conceitos de escalar, vetor e matriz. De forma simplificada, escalar, vetor e matriz são tensores de dimensões (grau) 0, 1, e 2, respectivamente. Ao trabalhar com imagens multiespectrais ou hiperespectrais, cada pixel é composto por mais de um valor, neste caso, é necessário representar a imagem como um tensor de dimensão 3, isto é, uma “lista de matrizes”. Por exemplo, uma imagem colorida RGB com $M \times N$ pixels é representada por um tensor de dimensão 3, ou seja, uma lista com 3 matrizes ($M \times N \times 3$). A operação de convolução é então generalizada para levar em conta os tensores de maiores dimensões:

$$g(x, y) = \sum_{s=-a}^a \sum_{t=-b}^b \sum_{k=0}^P w(s, t, k) f(x + s, y + t, k) \quad (20)$$

³⁰ Daí o nome da biblioteca TensorFlow, uma biblioteca de código aberto desenvolvida pela Google Brain para aprendizado de máquina. <https://www.tensorflow.org/guide/tensor> Acesso: 10 nov 2020.

É importante notar que tanto a imagem quanto a máscara devem ser tensores de mesmo grau, isto é, conter a mesma quantidade de matrizes: a imagem $f: M \times N \times P$ e a máscara $w: m \times n \times P$. Outra observação importante é que a operação resulta em um tensor de grau inferior, ou seja, um único mapa de ativação para cada máscara, como pode ser visto na Figura 11. Na RNC exemplificada na Figura 11, a primeira camada de convolução considera 6 máscaras as quais resultam em 6 mapas de ativação, ou seja, um tensor de dimensão 3, porém cada matriz é o resultado de uma convolução individual realizada no tensor de entrada (imagem RGB). Desta forma, se em uma camada de convolução forem utilizadas K máscaras (*kernel*), serão gerados K mapas de ativação.

3.3.3 Normalização dos dados de entrada

Para problemas que utilizam variáveis de entrada em diferentes escalas, isto é, grandeza, unidade ou magnitude diferentes, um pré-processamento pode ser aplicado para transformar os valores para um intervalo comum. Considerando N variáveis de entrada independentes (x_i), é possível determinar uma transformação linear para normalizá-las, fazendo:

$$x_i^{(n)} = \frac{x_i - \bar{x}_i}{s_i} \quad i = 1 \dots N \quad (21)$$

sendo \bar{x}_i e s_i^2 a média e variância amostral, respectivamente, calculados para a i -ésima variável.

Nota-se que após esta transformação as variáveis possuem média zero e desvio padrão unitário, isto é, estarão contidas no intervalo $[-1,1]$. Outra possibilidade é normalizar os dados no intervalo $[0,1]$, o que pode tornar o processo de convergência do treinamento mais rápido. Este último processo de normalização é também conhecido como normalização min-max, e pode ser feito como:

$$x_i^{(n)} = \frac{x_i - \min(x_i)}{\max(x_i) - \min(x_i)} \quad i = 1 \dots N \quad (22)$$

No caso de problemas de regressão é possível aplicar a transformação inversa como pós-processamento, de modo a converter as variáveis para suas escalas originais. De acordo com Bishop (1995) este subterfúgio matemático não tem impacto direto no resultado da

classificação, a não ser pelas vantagens (em termos práticos) do treinamento da RNA, uma vez que os valores utilizados estão em um intervalo similar.

Uma das vantagens de ter os dados normalizados é a possibilidade de armazená-los em variáveis do tipo ponto flutuante de 16 bits (*half precision*) ao invés dos tradicionais 32 e 64 bits (*single e double precision*). Este tipo de variável foi introduzido no padrão IEEE 754-2008, e é particularmente útil quando as redes neurais convolucionais possuem milhões de coeficientes³¹, e quando os conjuntos de dados chegam a ocupar dezenas de *gigabytes*³² de memória. Nestes casos, substituir as variáveis de ponto flutuante de 32 bits para 16 bits permitem reduzir efetivamente a utilização de memória pela metade, e teoricamente reduzir o tempo de processamento pela metade, enquanto mantém níveis aceitáveis de acurácia numérica (HO e WONG, 2017).

3.4 Descritores geométricos

De acordo com West *et al.* (2004), Gross e Thoennessen (2006), Demantké *et al.* (2011), Weinmann *et al.* (2013), Hackel *et al.* (2016) e Becker *et al.* (2017), existem diversas possibilidades para a geração de descritores geométricos, bem como para a aplicação deles. Dentre os descritores mais comum pode-se considerar histogramas, momentos de primeira ou segunda ordem, e assinaturas, por exemplo. Na sequência serão apresentadas as aplicações de momentos de segunda ordem e assinaturas do contorno, com foco no contexto deste trabalho, isto é, classificação da nuvem de pontos e o refinamento da detecção dos telhados de edificações, respectivamente.

3.4.1 Informação de variância local

Dentre outras possibilidades, a classificação de um pixel presente em uma imagem aérea pode ser realizada a partir dos valores radiométricos registrados pelo sensor, que por sua vez compõem um vetor de atributos. Diferentemente, as componentes posicionais (coordenadas

³¹ Por exemplo, a AlexNet, proposta em por Krizhevsky *et al.* (2012), contém aproximadamente 60 milhões de parâmetros, enquanto que outras arquiteturas mais recentes, como a VGG-16 de 2014 contém 138 milhões.

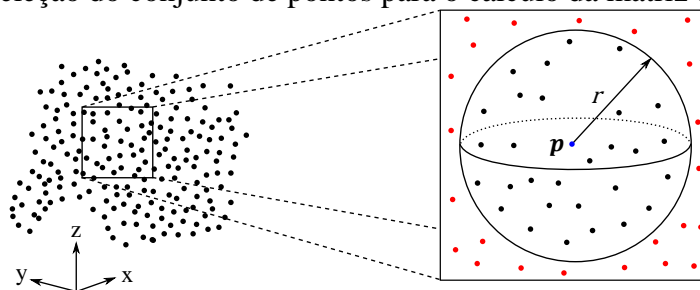
³² Como o conjunto de dados disponibilizados pela Microsoft, o *Common Objects in Context* (COCO), por exemplo, cujos dados de treinamento para a versão de 2017 ocupam 18 GB de memória. Disponível em <http://cocodataset.org>. Acessado em: 03 mar 2019.

X, Y, Z) de um ponto nos dados SVLA não são suficientes para discernir a classe do mesmo quando tratadas isoladamente, pois podem existir infinitas soluções. De modo alternativo, é possível inferir a classe do ponto com base na geometria local ao considerar sua vizinhança. Desta forma, é necessário construir um descritor que, quando avaliado como um vetor de atributos, possibilite diferenciar as possíveis feições presentes na superfície, com base nas suas propriedades geométricas.

A análise de componentes principais (ACP) é uma das técnicas de estatística multivariada mais utilizadas, que segundo Johnson e Wichern (2007), consiste na explicação dos dados a partir de combinações lineares das variáveis originais, visando inicialmente a redução da quantidade destas (volume de dados), assim como sua interpretação (BISHOP, 1995). A ACP consiste em uma transformação utilizada no reconhecimento de padrões não supervisionado, uma vez que o procedimento utiliza apenas os dados de entrada, isto é, nenhuma informação do resultado é conhecida *a priori*. A única restrição imposta nesta transformação é que as componentes (os versores) sejam ortonormais. Entretanto, também pode ser utilizada para extrair atributos que sejam aplicáveis a modelos de classificação supervisionados, assim como veremos a seguir.

No contexto de classificação de dados, a ACP pode ser utilizada como ferramenta para o reconhecimento de padrões a partir das propriedades estatísticas dos dados. Segundo Shi e Zakhor (2011), a ideia básica da classificação da nuvem de pontos utilizando ACP consiste em determinar as relações numéricas dos autovalores ($\lambda_1, \lambda_2, \lambda_3$) da matriz de covariâncias de um subconjunto de pontos selecionado a partir de uma certa vizinhança em torno de um ponto p (como apresentado na Figura 12).

Figura 12 - Seleção do conjunto de pontos para o cálculo da matriz de covariâncias.



Fonte: Adaptado de Santos e Galo (2016).

Com base nesses conceitos, gera-se o descritor geométrico do ponto como um vetor de atributos contendo os autovalores da matriz de covariâncias (momento de segunda ordem) do mesmo. Neste caso, as proporções dos autovalores indicam a dispersão espacial dos pontos vizinhos, o que pode ser correlacionado com primitivas como planos, intersecções de planos, segmentos de retas, quinas, entre outros. Considerando que se deseja classificar um ponto \mathbf{p} , e que se dispõe das coordenadas \mathbf{x}_i de todos os pontos que estejam a uma distância igual ou inferior a r , a matriz de covariâncias para este ponto pode ser calculada como:

$$\Sigma_{\mathbf{p}} = \sum (\mathbf{x}_i - \bar{\mathbf{x}})(\mathbf{x}_i - \bar{\mathbf{x}})^T \quad \forall \mathbf{x}_i \text{ tal que } d(\mathbf{x}_i, \mathbf{p}) \leq r \quad (23)$$

Assumindo uma amostragem regular, isto é, pontos distribuídos uniformemente, uma característica pode ser explorada na implementação deste cálculo, a qual consiste na aproximação do centro de massa para o centro do volume esférico. Esta propriedade pode ser útil para aplicações em tempo real, ou em situações em que a obtenção rápida dos resultados seja crítica, sendo possível obter uma aproximação da matriz de covariâncias fazendo:

$$\lim_{n \rightarrow \infty} \bar{\mathbf{x}} = \mathbf{p} \Rightarrow \Sigma_{\mathbf{p}} \approx \sum_{i=1}^n (\mathbf{x}_i - \mathbf{p})(\mathbf{x}_i - \mathbf{p})^T \quad (24)$$

Qualquer matriz real ou complexa de dimensões m por n pode ser decomposta a partir do produto de outras três matrizes: $\Sigma_{\mathbf{p}} = \mathbf{USV}^T$, o que configura a decomposição em valores singulares (ou *Singular Value Decomposition* – SVD). Para os casos particulares onde a matriz é real, simétrica e positiva-definida, esta decomposição configura a fatorização da matriz para sua forma canônica de Jordan, isto é, sua decomposição espectral (GOLUB e VAN LOAN, 1996):

$$\Sigma_{\mathbf{p}} = \sum_{i=1}^k \lambda_i \mathbf{e}_i \mathbf{e}_i^T = \mathbf{USU}^T = \mathbf{VSV}^T \quad (25)$$

Dentre as propriedades desse caso particular, as mais úteis para este trabalho consistem em: os autovalores e os correspondentes valores singulares são iguais ($\lambda_i = \sigma_i$); o traço da matriz original coincide com o traço da matriz diagonal \mathbf{S} , isto é, $\text{tr}(\Sigma_{\mathbf{p}}) = \text{tr}(\mathbf{S})$. Estas duas

propriedades implicam que a variância total da matriz original é igual a soma dos autovalores (ou soma dos valores singulares).

$$\text{prop}(\lambda_i) = \tilde{\lambda}_i = \frac{\lambda_i}{\text{tr}(\mathbf{\Sigma}_p)} \quad \text{para } i = 1, 2 \dots k. \quad (26)$$

em que k é a dimensão da matriz $\mathbf{\Sigma}_p$.

Por conveniência, algumas implementações da SVD realizam a ordenação decrescente dos autovalores, isto é, $\lambda_1 > \lambda_2 > \dots > \lambda_k$. Assumindo que esta ordenação seja feita, é possível estabelecer algumas relações sobre os autovalores. Em uma situação ideal (hipotética), um ponto pertencente à um plano teria seu terceiro autovalor nulo, entretanto, caso ele pertença à uma feição linear, tanto o segundo quanto terceiro autovalor seriam nulos. Alguns trabalhos (Shi e Zakhor, 2011; Yang e Dong, 2013) associam algumas relações entre autovalores à estruturas geométricas, enquanto que outros (WEST *et al.*, 2004; GROSS e THOENNESSEN, 2006; DEMANTKÉ *et al.*, 2011; WEINMANN *et al.*, 2013; HACKEL *et al.*, 2016; e BECKER *et al.*, 2017) propuseram e adotaram métricas que ajudam a distinguir qual tipo de estrutura o ponto avaliado está inserido, isto é, pertence. Alguns desses atributos estão listados no Quadro 3.

Quadro 3 – Exemplos de atributos geométricos utilizados para a classificação de nuvem de pontos, derivados dos autovalores da MVC das coordenadas.

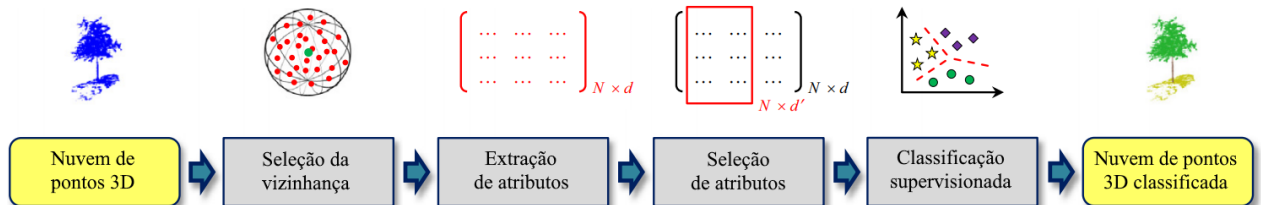
| | |
|---------------|---|
| Omnivariância | $(\lambda_1 \lambda_2 \lambda_3)^{1/3}$ |
| Anisotropia | $(\lambda_1 - \lambda_3)/\lambda_1$ |
| Planaridade | $(\lambda_2 - \lambda_3)/\lambda_1$ |
| Linearidade | $(\lambda_1 - \lambda_2)/\lambda_1$ |
| Dispersão | λ_3/λ_1 |
| Entropia | $-\sum_{i=1}^3 \lambda_i \ln \lambda_i$ |

Fonte: Adaptado de Weinmann *et al.* (2015).

3.4.2 Atributos geométricos invariantes à densidade de pontos

De acordo com Hackel *et al.* (2016), um dos principais problemas envolvidos na classificação de nuvens de pontos, consiste na mudança da densidade de pontos. Esta dificuldade afeta tanto dados provindos de sistemas de varredura LASER, quanto nuvens de pontos obtidas por correspondência de imagens. Dados de varredura LASER aéreos ou terrestres geralmente são obtidos por múltiplas faixas (ou estações no caso terrestre), as quais apresentam sobreposição entre si. Nessas situações, espera-se que a densidade de pontos amostrados na região de sobreposição seja aproximadamente o dobro das demais, amostradas apenas uma vez. Além disso, outros fatores podem acometer a almejada constância e regularidade na amostragem da superfície, tais como variações na velocidade, aceleração e rotações da plataforma, assim como observado em Santos *et al.* (2019). Tais fatores podem causar problemas na determinação dos atributos geométricos, e até mesmo na definição da vizinhança. A Figura 13 apresenta o processo comumente empregado para a classificação de nuvens de pontos utilizando métodos de aprendizado de máquina e engenharia de atributos.

Figura 13 – Classificação supervisionada de nuvem de pontos por engenharia de atributos.

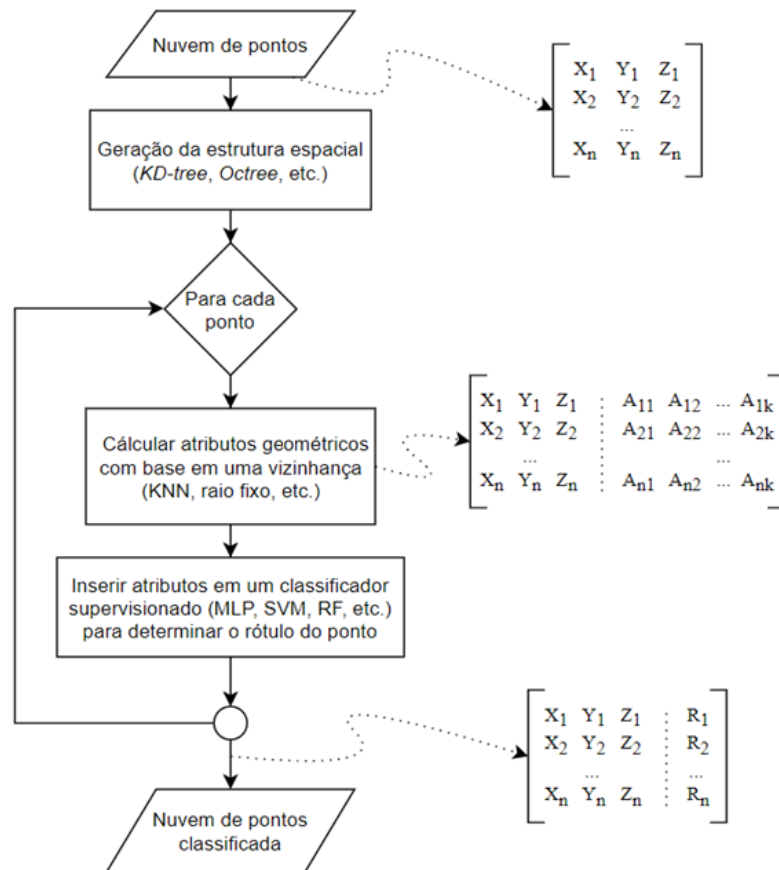


Fonte: Adaptado de Weinmann *et al.* (2015).

Visando minimizar os efeitos da variação de densidade, alguns estudos propuseram soluções multi-escala. Weinmann *et al.* (2015) estudou a definição da vizinhança do ponto e o impacto da quantidade de pontos vizinhos na acurácia da classificação de nuvens de pontos utilizando diversos modelos supervisionados, tais como SVM, *Random Forest*, e Redes Neurais, por exemplo. Os autores concluíram que a melhor maneira de definir a vizinhança é a partir de uma quantidade k de pontos próximos (k NN – *nearest neighbors*), ao contrário de utilizar um raio fixo. A quantidade de pontos vizinhos, no entanto, deve ser definida a partir de um intervalo de busca ideal (no qual os autores propuseram $k \in [10,100]$) e uma métrica, bem como algum critério de seleção. Para o critério de seleção, os autores adaptaram a estratégia proposta por Demantké *et al.* (2011), a qual basea-se em adotar a vizinhança que minimiza a

entropia dos autovalores. Em suma, essa abordagem sugere que a entropia dos autovalores deve ser calculada para $k = 1 \dots 100$ vizinhos mais próximos, adotando-se a vizinhança k cujo valor da entropia seja mínimo. O processo deve ser realizado para cada ponto da nuvem, e após selecionada a escala, isto é, vizinhança, os demais atributos geométricos são calculados e utilizados na classificação (Figura 14).

Figura 14 – Fluxograma da classificação de nuvens de pontos com engenharia de atributos.



Fonte: Elaborada pelo autor.

A questão da variação de densidade deve ser abordada na seleção da vizinhança (primeira etapa). Uma possibilidade para a seleção dos atributos é a utilização da Análise de Componentes Principais, a qual busca identificar atributos que possuam correlação entre si. Desta maneira, apenas os atributos que possam auxiliar na classificação são levados em consideração. Tendo os atributos calculados, basta treinar e aplicar algum modelo matemático de classificação supervisionado. Como pode ser visto em Weinmann *et al.* (2015), as Redes Neurais podem atingir bons níveis de acurácia quando os dados de treinamento e a topologia escolhida são adequados.

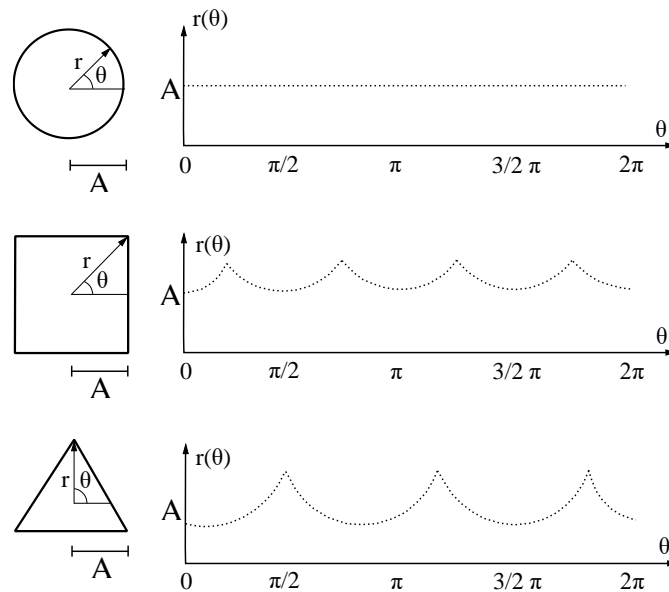
Uma das maneiras de reduzir o custo computacional no cálculo dos atributos consiste na geração de alguma estrutura espacial sobre os dados de entrada. Algumas dessas estruturas, tais como a KD-tree, Octree, e Quad-Tree, por exemplo, podem reduzir o espaço de busca para a vizinhança de um dado ponto. Desta forma, ao invés de percorrer todos os pontos da nuvem, apenas os das células (ou também chamados voxels) mais próximos são visitados. O tipo da estrutura espacial vai depender da geometria escolhida para a vizinhança.

3.4.3 Assinatura do contorno

Na literatura encontram-se diversos esquemas de representação e descrição de objetos, ou da forma de objetos, baseados na obtenção de características de regiões segmentadas nas imagens. A escolha correta da estrutura de representação facilita o armazenamento e manipulação dos segmentos, além de simplificarem o cálculo de certos descritores de forma (PEDRINI e SCHWARTZ, 2007). Assinaturas são representações unidimensionais das bordas dos objetos, cuja maneira convencional mais simples é a partir da construção do gráfico da distância entre o centro de massa do segmento e sua borda (descritor $r(\theta)$), em função do ângulo de varredura (GONZALEZ e WOODS, 2007). São várias as possibilidades de obter a assinatura, uma alternativa recorrente consiste em percorrer as bordas do objeto e representar o ângulo formado entre a tangente do ponto atual e o segmento de reta que une tal ponto e o centro de massa (descritor $\kappa(s)$). A Figura 15 apresenta três objetos simples e suas respectivas assinaturas (gráfico da distância pelo ângulo de varredura, ou seja, $\theta \times r(\theta)$).

Como pode ser visto em Gonzalez e Woods (2007), em sua abordagem mais simples, isto é, descritor $r(\theta)$, a assinatura é independente da translação e dependente da orientação e escala. Uma maneira de solucionar a dependência da orientação é a adoção de critérios de seleção para o ponto de partida da assinatura. Uma das alternativas é adotar o centro de massa do objeto. Tais critérios podem ser definidos a partir de características intrínsecas ao segmento, como o ponto no autovetor principal que apresente a máxima distância ao centro de massa (PEDRINI e SCHWARTZ, 2007). Em algumas situações, no entanto, a adoção deste critério pode acarretar em ambiguidades, como nos casos de objetos circulares ou polígonos que apresentem múltiplos vértices de máxima distância ao centro de massa.

Figura 15 - Três figuras geométricas simples e as respectivas assinaturas.

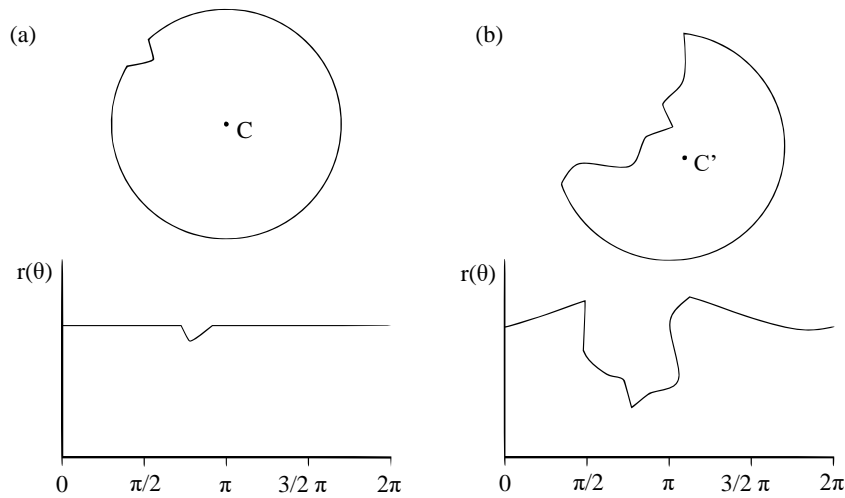


Fonte: Adaptado de Pedrini e Schwartz (2007).

A dependência de escala pode ser minimizada a partir da normalização das distâncias no intervalo $[0,1]$, assumindo que o sistema de coordenadas do objeto seja ortogonal e apresente escala uniforme nos dois eixos. A normalização das distâncias é feita em função de seus valores de máximo e mínimo, porém, em casos práticos, os objetos podem apresentar ruídos (contornos serrilhados ou sinuosos). Nessas situações a solução pode ser obtida pela divisão de cada distância pela variância da assinatura, quando esta for não nula (GONZALEZ e WOODS, 2007). Uma alternativa seria a simplificação do contorno do objeto utilizando o algoritmo de simplificação de curvas proposto por Douglas e Peucker (1973), no qual os vértices que não contribuem significativamente para a delimitação do contorno são removidos.

Existem ainda outras situações problemáticas com o descritor $r(\theta)$, as quais não possuem solução trivial. Assim como pode ser visualizado na Figura 16, oclusões nas bordas dos objetos podem causar a translação do centroide de sua verdadeira posição e, desta forma os processos de correspondência serão prejudicados. A solução mais direta para este caso é a substituição do centroide por algum outro ponto de referência do objeto. Além disso, alguns objetos podem causar ambiguidades na geração do descritor (Figura 17), isto é, múltiplos valores de r para um único valor de θ . Neste caso o processo de correspondência passa a ser bidimensional, o que acarreta em processamentos adicionais. Uma possível solução para este problema consiste na adoção de algum critério, como o registro apenas da menor distância, por exemplo.

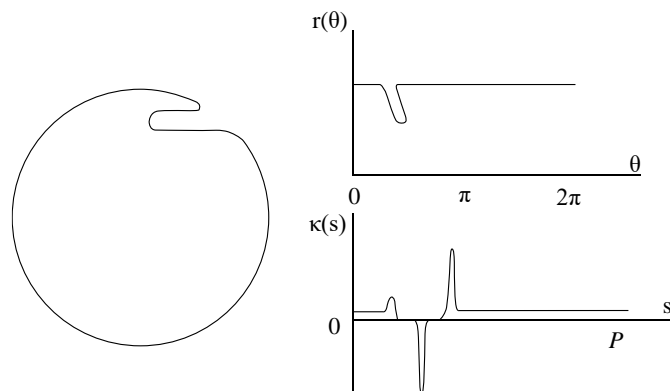
Figura 16 - Problema com o descritor de contorno $r(\theta)$. (a) Objeto circular com pequena perfuração na borda. (b) Mesmo objeto, porém parcialmente sujeito à oclusões.



Fonte: Adaptado de Davies (2012).

Em virtude das situações problemáticas mencionadas para o descritor $r(\theta)$, foram desenvolvidas outras soluções, tais como o descritor $\kappa(s)$, o qual é mais robusto. A primeira diferença surge por conta do gráfico resultante não possuir periodicidade, como na representação anterior, o que dificulta o processo de correspondência. No entanto, algumas considerações podem ser feitas para a adoção do ponto de partida para o descritor $r(\theta)$, assim como a normalização do valor do perímetro do objeto, o que soluciona a dependência de orientação e escala. Desta maneira, este descritor mais simples pode ser utilizado de maneira similar para objetos mais complexos, como os telhados das edificações.

Figura 17 - Diferença entre os gráficos dos descritores $r(\theta)$ e $\kappa(s)$.



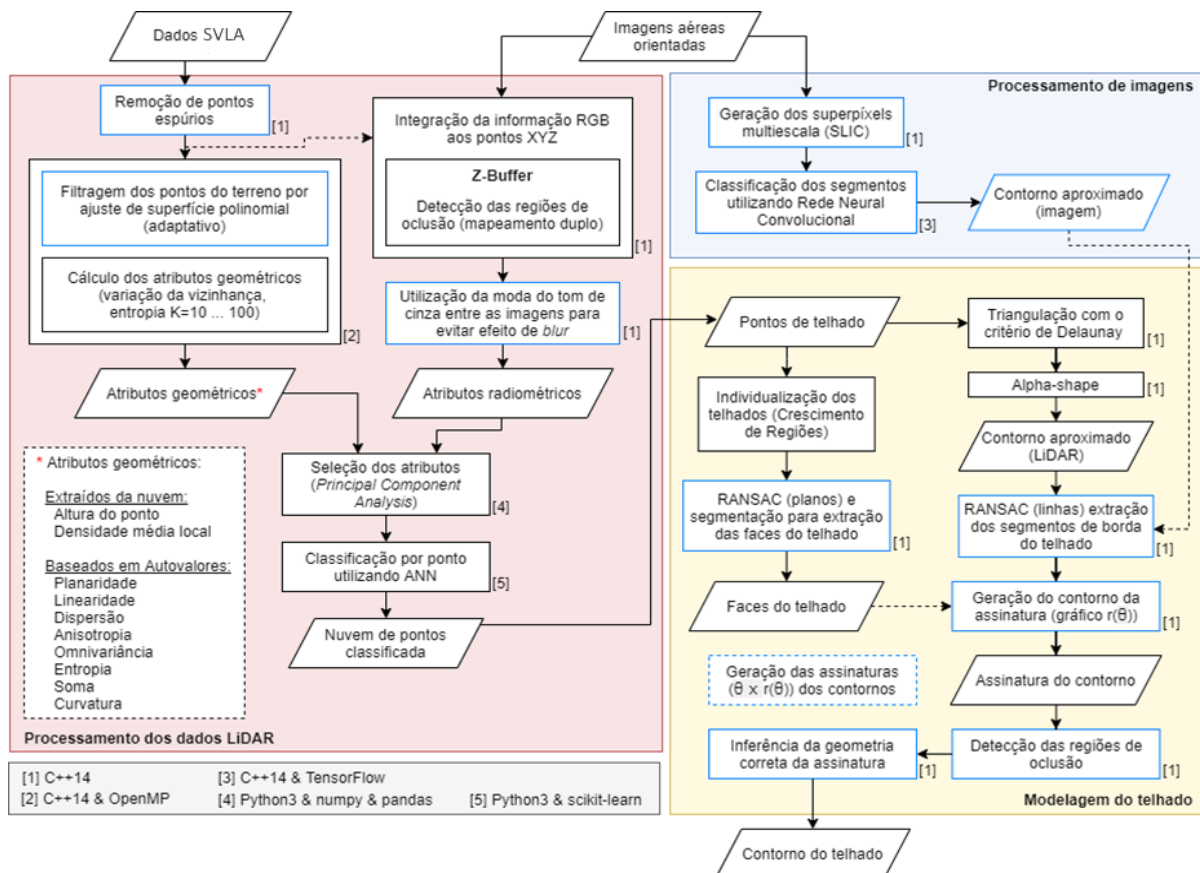
Fonte: Adaptado de Davies (2012).

Capítulo 4

EXTRAÇÃO DE CONTORNOS DE TELHADOS

Este capítulo tem o propósito de apresentar o método para a extração de contornos de telhados. O método proposto é dividido em três etapas: preparação dos dados; classificação da nuvem de pontos; e a extração e regularização nos contornos de telhados delimitados (Figura 18). A fase de preparação dos dados compreende a remoção de pontos espúrios presentes na nuvem de pontos, a integração da informação radiométrica às coordenadas dos pontos da nuvem de pontos, a geração de uma malha de triângulos irregular (TIN – *Triangular Irregular Network*) seguido do cálculo de vetores normais para cada vértice. Além dessas quantidades, são calculados os autovalores da matriz de covariâncias das coordenadas dos pontos vizinhos, com o objetivo de obter um indicador da geometria local a partir da informação de variância das coordenadas.

Figura 18 – Fluxograma do método proposto. Em azul são destacadas as contribuições.



Fonte: Elaborado pelo autor.

A segunda fase visa a identificação de quais pontos da nuvem pertencem à telhados de edificações. Nesta etapa é aplicado um processo de filtragem progressiva a partir do ajuste de superfícies polinomiais, desta forma os pontos do terreno e de objetos são identificados. Emprega-se então o classificador supervisionado (RNA) apenas nos pontos de objetos com o intuito de separar pontos de telhados de edificações dos de vegetação. Esta classificação emprega métricas (indicadores) de planos, e a classe vegetação é obtida por exclusão, ou seja, o que não pertencer à planos ou ao terreno.

A última etapa consiste na extração e regularização dos contornos dos telhados. Esta fase tem como entrada o conjunto de pontos de telhados obtidos na classificação, nos quais aplica-se uma segmentação por crescimento de regiões visando a individualização de cada telhado. Finalmente, os contornos dos segmentos que compõem as edificações são extraídos por assinaturas (descritores geométricos), e caso identificadas, as regiões inconsistentes são regularizadas.

4.1 Remoção de pontos espúrios

Apesar da qualidade e robustez dos equipamentos atuais, e ainda que a aquisição dos dados durante os voos seja feita com diligência, os dados SVLA podem conter medidas indesejadas decorrentes de fatores externos presentes na cena. Sotoodeh (2006) diferencia as possíveis causas da presença de pontos espúrios em dados SVLA entre: oclusões em bordas de objetos, reflectância da superfície e multicaminhamento. Diversas abordagens podem ser utilizadas para a remoção de pontos espúrios em dados SVLA aerotransportado, sendo as mais simples baseadas em informações externas como a altitude média do terreno da região estudada, ou a diferença máxima entre as altitudes do primeiro e último retorno do pulso LASER (MATKAN *et al.*, 2014).

Matkan *et al.* (2014) classificam os pontos espúrios em positivos e negativos: os positivos são resultantes de retornos dos pulsos LASER refletirem objetos próximos ao sistema de varredura, tais como pássaros ou pequenas aeronaves não tripuladas, por exemplo. Além das causas mencionadas, Leslar *et al.* (2010) listam partículas suspensas na atmosfera (neve e poeira, por exemplo) como possíveis fontes de pontos espúrios positivos. Já aos pontos espúrios negativos são atribuídos ao efeito de multicaminhamento do pulso emitido, de forma análoga ao que ocorre com a tecnologia GNSS. Neste caso, o pulso percorre um trajeto mais longo até retornar ao sistema, gerando um ponto em uma elevação mais baixa do que a real. Pontos

espúrios negativos acarretam problemas nas operações de filtragem (geração do MDT), pois os métodos comumente aplicados para essa finalidade assumem os pontos de menor altitude como pertencentes ao terreno (VOSSELMAN, 2000).

Existem na literatura diversas abordagens para a remoção dos pontos espúrios presentes nos dados SVLA, as quais usualmente baseam-se no fato de que eles normalmente compareçam isolados (sozinhos ou em pequenos agrupamentos) das demais feições. Nesta etapa do método proposto emprega-se uma versão adaptada do método proposto por Carrilho e Galo (2017), cujos critérios de remoção dos pontos espúrios fundamentam-se na análise estatística das altitudes, considerando dois critérios. Inicialmente é feita a leitura dos dados SVLA e o cálculo da média (\bar{z}) e desvio-padrão (σ_z) das altitudes. Estes valores, juntamente com o valor de n que multiplica o desvio-padrão, são utilizados para a determinação do primeiro intervalo de aceitação, conforme Eq. 24.

$$[\bar{z}-n\sigma_z, \bar{z}+n\sigma_z] \quad (27)$$

O segundo intervalo de aceitação é obtido a partir do histograma das altitudes dos pontos. Para isso o algoritmo percorre as classes, até encontrar a primeira e a última ocorrência (h_{\min} , h_{\max}) nas quais a frequência (quantidade de pontos) seja superior a um limiar estipulado pelo usuário. Assim, os pontos que simultaneamente não estiverem na região de aceitação dos pontos, pelos dois critérios, serão considerados espúrios. A região de aceitação poderá ser obtida a partir da intersecção dos dois intervalos computados anteriormente, isto é:

$$[\min(\bar{z}-n\sigma_z, h_{\min}), \max(\bar{z}+n\sigma_z, h_{\max})] \quad (28)$$

No entanto, assim como observado por Carrilho e Galo (2017), a utilização desses dois critérios possibilita a remoção apenas dos pontos espúrios que estejam grosseiramente destoantes da superfície, isto é, realiza-se uma filtragem parcial desses erros grosseiros. Desta forma o presente trabalho propõe uma adaptação por meio da inclusão de um terceiro critério, cujo intuito é detectar pontos espúrios cujas altitudes sejam semelhantes às dos objetos presentes na superfície ou da ordem da diferença de elevação do terreno. Com a inclusão deste critério adicional espera-se que esta abordagem seja particularmente útil para terrenos com relevo acidentado.

O terceiro critério é concebido a partir da frequência espacial dos pontos: percorre-se todos os pontos do conjunto de dados, e para cada um deles é contabilizada a quantidade de pontos cuja distância Euclidiana até este seja inferior a um dado limiar. Caso a quantidade de pontos vizinhos seja pequena (ou nula), há indícios para considerar o ponto como isolado, e, portanto, espúrio. O critério pode ser estabelecido matematicamente da seguinte forma: considerando um conjunto de m pontos \mathbf{P} , um dado ponto $\mathbf{p}_i \in \mathbf{P}$ é considerado como espúrio caso a condição seja verdadeira:

$$\sum T(\mathbf{p}_i, \mathbf{p}_j, r) \leq k \quad i = 1 \dots m, \quad j = 1 \dots m, \quad i \neq j \quad (29)$$

sendo k o limiar de frequência de pontos, r é o limiar de distância máxima para considerar o ponto como vizinho³³, e $T(\mathbf{p}_i, \mathbf{p}_j, r)$ a função degrau responsável por contabilizar a frequência de pontos vizinhos, isto é:

$$T(\mathbf{p}_i, \mathbf{p}_j, r) = \begin{cases} 1 & \text{se } d(\mathbf{p}_i, \mathbf{p}_j) \leq r \\ 0 & \text{caso contrário} \end{cases} \quad (30)$$

Desta maneira, os possíveis pontos espúrios que não foram removidos pelos dois critérios anteriores podem ser identificados por esta terceira condição. Resumindo, esta abordagem permite a remoção de pontos espúrios sem a necessidade de informações externas aos dados. Contudo, cinco valores devem ser informados *a priori* pelo operador: o multiplicador dos sigmas (n), o intervalo para a criação das classes do histograma, o limiar de frequência mínima para a geração do segundo intervalo, o limiar de frequência de pontos (k), e o limiar de distância máxima para definir a vizinhança (r).

Na sequência é apresentado o Algoritmo 1 baseado no uso dos três critérios mencionados.

³³ O algoritmo para consultar pontos vizinhos possui complexidade $O(n^2)$, entretanto, as implementações podem considerar tipos abstratos de dados (TAD) como *quadtrees*, *octrees*, *KD-trees*, por exemplo, visando facilitar a operação de busca com base na distribuição espacial dos pontos (topologia).

Algoritmo 1 – Remoção dos pontos espúrios a partir de análises estatísticas (intervalos e frequência).

Entrada: dados LiDAR sem pré-processamento

Saída: dados LiDAR após a remoção dos pontos espúrios

- 1: calcular a média (\bar{z}) e desvio-padrão (σ_z) das altitudes;
 - 2: definir o intervalo 1: $[\bar{z}-n\sigma_z, \bar{z}+n\sigma_z]$;
 - 3: calcular o histograma de altitudes;
 - 4: definir o intervalo 2: $[h_{\min}, h_{\max}]$;
 - 5: definir o intervalo 3 como a intersecção entre 1 e 2: $[\min(\bar{z}-n\sigma_z, h_{\min}), \max(\bar{z}+n\sigma_z, h_{\max})]$;
 - 6: **para cada** ponto da nuvem **faça:**
 - 7: **caso** a altitude não esteja contida no intervalo 3, **então:**
 - 8: remover o ponto;
 - 9: **para cada** ponto restante **faça:**
 - 10: calcula da quantidade de pontos vizinhos considerando um raio de busca r ;
 - 11: **caso** o resultado seja menor que o limiar (k), **então:**
 - 12: remover o ponto.
-

4.2 Integração da informação radiométrica

Assim como mencionado por Habib (2008), a natureza complementar entre as informações posicional e radiométrica, obtidas por sensores embarcados em plataformas aéreas, motivou a integração desses conjuntos de dados com o objetivo de proporcionar uma melhor descrição da superfície amostrada. Entretanto, tal integração só é possível quando os dados coletados podem ser referenciados em um único sistema. Geralmente, a informação posicional pode ser obtida tanto por técnicas fotogramétricas convencionais, como a intersecção de imagens, por exemplo, quanto por SVLA. Já a informação radiométrica é usualmente adquirida em larga escala³⁴ a partir de sensores de quadro e/ou lineares (*pushbroom*), estes últimos principalmente para o caso de sensores orbitais. Este trabalho limita-se a tratar de dados posicionais obtidos por SVLA e dados radiométricos obtidos por sensores digitais de quadro. Outras fontes de informações radiométricas são as ortoimagens, entretanto a utilização desses produtos não se enquadra no escopo desta tese.

³⁴ A especificidade da escala nesse contexto se dá em função da possível confusão referente à aquisição de informações radiométricas com outros instrumentos que operam em áreas menores, como espectrorradiômetros, por exemplo.

Considerando o conceito de imagem digital como uma função contínua $f(x,y)$ apresentado por Gonzalez e Woods (2007), e assumindo como conhecidos os parâmetros de orientação interior e exterior para as imagens aéreas de uma dada região, é possível associar a informação radiométrica às coordenadas posicionais dos dados SVLA desta região a partir das equações de colinearidade. Este modelo matemático permite relacionar as coordenadas (X_P, Y_P, Z_P) de cada ponto da nuvem com o nível digital (ND) dos pixels registrados em uma ou mais imagens, dependendo da configuração do voo, como a sobreposição, principalmente.

$$\begin{aligned} x &= -c \frac{m_{11}(X_P - X_{CP}) + m_{12}(Y_P - Y_{CP}) + m_{13}(Z_P - Z_{CP})}{m_{31}(X_P - X_{CP}) + m_{32}(Y_P - Y_{CP}) + m_{33}(Z_P - Z_{CP})} \\ y &= -c \frac{m_{21}(X_P - X_{CP}) + m_{22}(Y_P - Y_{CP}) + m_{23}(Z_P - Z_{CP})}{m_{31}(X_P - X_{CP}) + m_{32}(Y_P - Y_{CP}) + m_{33}(Z_P - Z_{CP})} \end{aligned} \quad (31)$$

sendo c a distância focal da câmara, (X_{CP}, Y_{CP}, Z_{CP}) as coordenadas do centro perspectivo no sistema de referência, e m_{ij} $i, j = 1 \dots 3$ os coeficientes da matriz de rotação passiva \mathbf{M} , calculada em função dos ângulos de Euler $(\omega, \varphi, \kappa)$ por:

$$\mathbf{M} = \mathbf{R}_Z(\kappa) \mathbf{R}_Y(\varphi) \mathbf{R}_X(\omega) \quad (32)$$

Em termos práticos, a transformação entre o sistema fotogramétrico (x,y) e o sistema de coordenadas da imagem (i,j) envolve a correção de alguns efeitos, como as distorções do conjunto de lentes e a redução ao ponto principal, por exemplo. O processo completo envolve as seguintes operações: $(X_P, Y_P, Z_P) \rightarrow (x,y) \rightarrow (x',y') \rightarrow (i,j)$:

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} x_0 \\ y_0 \end{bmatrix} + \begin{bmatrix} \delta r_x + \delta d_x \\ \delta r_y + \delta d_y \end{bmatrix} \quad (33)$$

em que $(\delta r_x, \delta r_y)$ representam as componentes da distorção radial simétrica, calculadas a partir dos coeficientes k_1, k_2 e k_3 :

$$\begin{bmatrix} \delta r_x \\ \delta r_y \end{bmatrix} = \begin{bmatrix} (x' - x_0)(k_1 r^2 + k_2 r^4 + k_3 r^6) \\ (y' - y_0)(k_1 r^2 + k_2 r^4 + k_3 r^6) \end{bmatrix} \quad (34)$$

e $(\delta d_x, \delta d_y)$ representam as componentes da distorção descentrada, calculadas a partir dos coeficientes P_1 e P_2 :

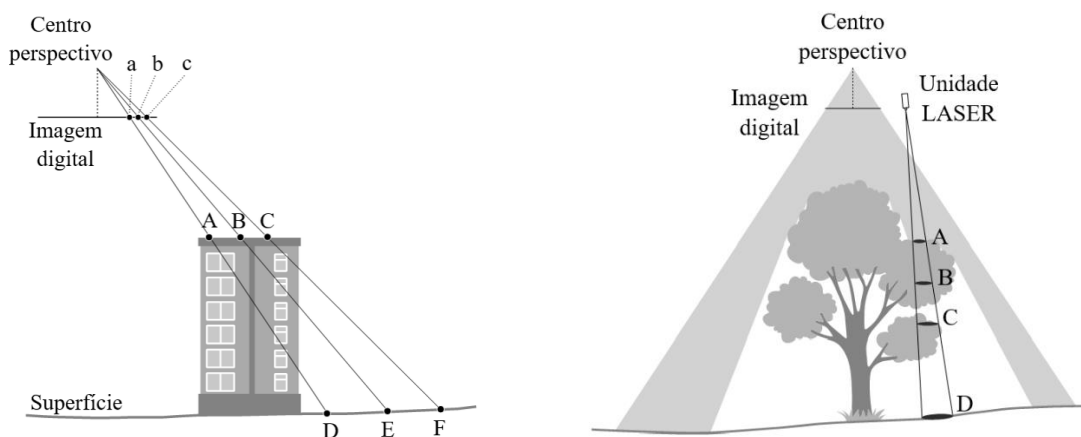
$$\begin{bmatrix} \delta d_x \\ \delta d_y \end{bmatrix} = \begin{bmatrix} P_1[r^2 + 2(x' - x_0)^2] + 2P_2(x' - x_0)(y' - y_0) \\ P_2[r^2 + 2(y' - y_0)^2] + 2P_1(x' - x_0)(y' - y_0) \end{bmatrix} \quad (35)$$

sendo $r = [(x' - x_0)^2 + (y' - y_0)^2]^{1/2}$ a distância plana entre o ponto medido no sensor e o ponto principal.

$$\begin{bmatrix} j \\ i \end{bmatrix} = \begin{bmatrix} s_x^{-1} & 0 \\ 0 & -s_y^{-1} \end{bmatrix} \begin{bmatrix} x' \\ y' \end{bmatrix} + \begin{bmatrix} (W-1)/2 \\ (H-1)/2 \end{bmatrix} \quad (36)$$

No entanto, um problema se origina devido ao cálculo das distorções, que é realizado a partir das coordenadas intermediárias (x', y') , as quais não são conhecidas, neste caso. Desta forma, é necessário utilizar um processo iterativo, tomando (x, y) como aproximações iniciais. Uma vez obtida a convergência após este processo iterativo, basta associar o tom de cinza do pixel (i, j) ao ponto no espaço objeto situado na posição (X_P, Y_P, Z_P) . Esse processo é similar ao da geração de ortofotos, compartilhando do mesmo problema: o mapeamento duplo dos pixels, assim como ilustrado na Figura 19.

Figura 19 – Problemas ao registrar informação radiométrica à informação posicional.



(a) Duplo mapeamento dos pixels (a,b,c) nos pontos (A,B,C) corretamente, e (D,E,F) incorretamente.

Fonte: Adaptado de Habib *et al.* (2007).

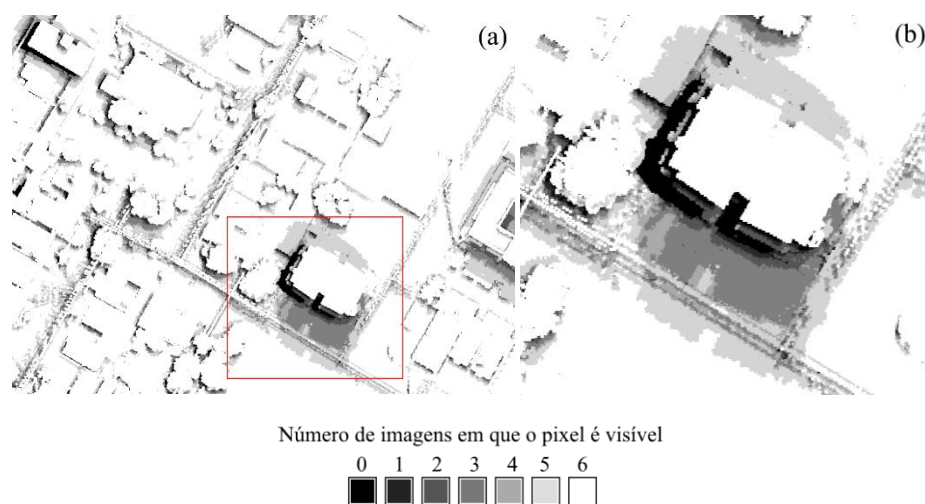
(b) Retornos (B,C,D) do pulso LASER no interior da copa não visíveis na imagem, apenas o primeiro (A).

Fonte: Adaptado de: Hyypä *et al.* (2008).

Existem diversos métodos para tratar o problema do mapeamento duplo. Neste trabalho optou-se por utilizar o método Z-Buffer (HABIB *et al.*, 2007), o qual detecta as regiões de oclusão nas imagens, com o intuito de descartar tais pixels. As regiões oclusas podem ser então, obtidas a partir de outras imagens tomadas em pontos de vistas diferentes. Dependendo da configuração do voo, isto é, da sobreposição, é possível que um mesmo ponto seja imageado várias vezes. A Figura 20 apresenta uma análise visual da oclusão causada por objetos elevados. Em alguns locais os pontos são visíveis por apenas uma imagem.

Além do que foi mencionado, outros dois problemas podem ocorrer: (i) alteração da cena durante a aquisição das imagens (Figura 21), e (ii) quando um ponto não ocluso aparece em múltiplas imagens: qual valor RGB (de qual imagem) deve ser associado à sua coordenada. Dependendo da dimensão do bloco, a mesma região pode ser imageada por duas ou mais faixas de voo, as quais podem demorar alguns minutos. Nesse intervalo de tempo é possível (e até esperado) que alguns objetos presentes na cena urbana se movimentem. Essa característica pode resultar em um problema na associação da informação radiométrica caso a média dos valores RGB sejam utilizados, visto que irá resultar em um “borrão” nos casos em que o objeto se moveu.

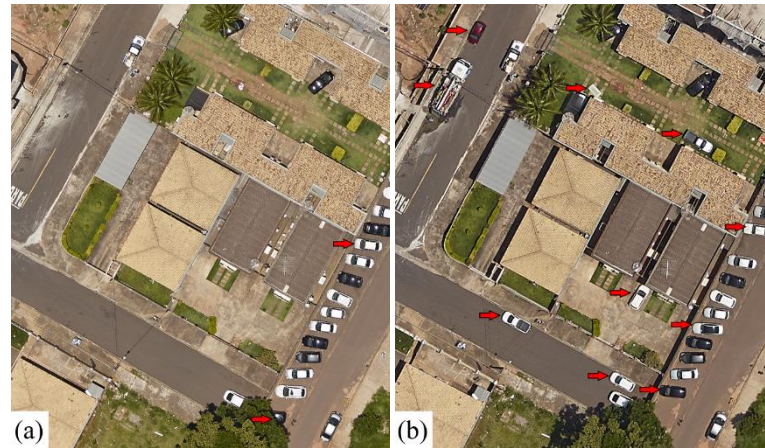
Figura 20 - Análise de visibilidade da nuvem de pontos considerando seis imagens aéreas.



Fonte: Elaborado pelo autor.

Na Figura 21 é possível observar que, alguns veículos se moveram, e até mesmo a posição das sombras dos objetos foi alterada entre uma faixa e outra. A Figura 21a foi tomada mais próxima do meio-dia, visto que quase não aparecem as sombras dos objetos. Já na Figura 21b, tomada momentos mais tarde, a sombra das edificações começa a aparecer.

Figura 21 - Imagens da mesma região tomadas em horários diferentes. Alterações na iluminação (posição do Sol) e na composição de objetos na cena (veículos, principalmente).



Fonte: Elaborado pelo autor.

4.3 Cálculo dos vetores normais

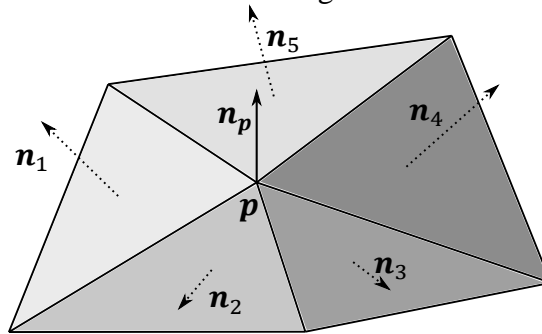
O último processo da etapa de preparação dos dados consiste na geração do TIN e cálculo dos vetores normais para cada vértice das faces triangulares (Figura 22). Dentre as possíveis técnicas para a triangulação cita-se o critério de Delaunay, o qual prioriza a geração de faces com geometria interna equilibrada, isto é, maximiza o menor ângulo do triângulo. Após gerada a malha triangular é possível calcular o vetor normal de cada vértice a partir de algum estimador, tal como a média aritmética dos vetores normais de cada face que contenham este vértice (ZHAO *et al.*, 2019). Supondo m vetores normais unitários \mathbf{n}_q e suas correspondentes faces triangulares que contenham um dado vértice \mathbf{p} , o vetor normal deste vértice (\mathbf{n}_p) pode ser estimado como:

$$\mathbf{n}_p = \frac{1}{m} \sum \mathbf{n}_q \quad (37)$$

Os vetores normais calculados para cada vértice são particularmente úteis para a detecção de tendências de orientação, o que torna possível inferir sobre possíveis entidades geométricas presentes nos dados. Neste sentido, é proposta uma medida de colimação, isto é, paralelismo entre vetores normais vizinhos, a qual serve como um indicador do tipo de superfície em que o

vértice se encontra. A motivação para o desenvolvimento desse indicador origina-se a partir do ângulo formado entre vetores no espaço \mathbb{R}^3 .

Figura 22 - Determinação do vetor normal médio para o vértice central de uma pequena malha triangular.



Fonte: Elaborada pelo autor.

$$\Omega(\mathbf{n}_p) = \sum \arccos \frac{\mathbf{n}_p \cdot \mathbf{n}_q}{|\mathbf{n}_p| |\mathbf{n}_q|} \quad \forall q \mid d(\mathbf{p}, \mathbf{q}) < \varepsilon \quad (38)$$

sendo ε o raio delimitador da vizinhança considerada no cálculo.

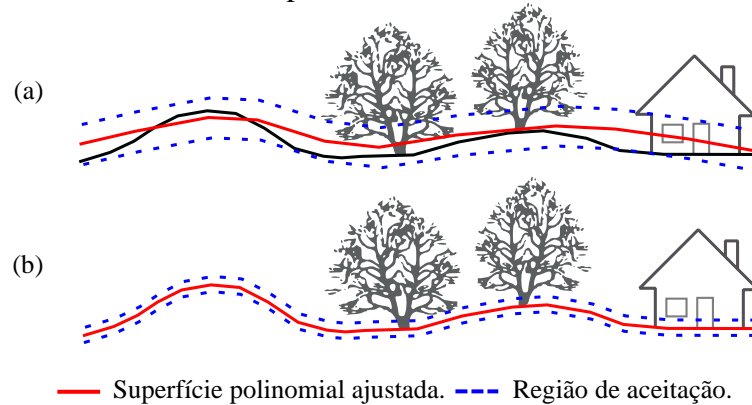
Esta medida é particularmente útil na detecção de regiões planas, nas quais o resultado esperado é um valor pequeno ou nulo, idealmente. Assim sendo, este indicador pode ser utilizado conjuntamente com a informação de variância local como um dos atributos de entrada para o classificador supervisionado baseado em RNA. Considerando que os pontos já tenham sido rotulados entre as classes terreno e objetos, é possível associar as regiões planas detectadas nos pontos de objetos a telhados de edificações. Em outras palavras, esta medida pode ser entendida como um indicador de edificações em dados SVLA.

4.4 Filtragem por ajuste progressivo de superfícies polinomiais

Após a etapa de preparação, o conjunto de dados (MDS) é filtrado com o intuito de distinguir os pontos entre terreno (MDT) e objetos, isto é, separar retornos de pulsos LASER refletidos pelo terreno, dos que foram refletidos por outras feições, tais como edificações e a vegetação, por exemplo. Na literatura encontram-se diversas possíveis abordagens para a realização desta filtragem (MENG *et al.*, 2010). Neste projeto pretende adaptar o método

proposto por Qin *et al.* (2017), cuja abordagem baseia-se no ajuste progressivo de superfícies polinomiais bivariadas (Figura 23) pelo MMQ.

Figura 23 – Princípio do ajuste progressivo por meio do ajuste de superfície polinomial aos pontos do MDT.



Fonte: Elaborado pelo autor.

A operação da filtragem consiste em aproximações sucessivas do MDT a partir de um processo iterativo de ajuste de superfícies matemáticas. Quando a seleção dos pontos semente é realizada de forma adequada espera-se que o método obtenha convergência, removendo os pontos de objetos ao mesmo tempo que preserva as características do relevo. Este processamento resulta em um conjunto com os pontos do terreno, que por comparação com o MDS original permite separar os pontos pertencentes a outros objetos.

Alguns autores (XIONG *et al.*, 2012; QIN *et al.*, 2017) utilizam superfícies polinomiais bivariadas quadráticas³⁵, entretanto uma abordagem mais abrangente pode ser descrita a partir da generalização da ordem dos polinômios considerados. A formulação proposta na Eq. 36 permite a definição do tipo de superfície representada em função do grau do polinômio.

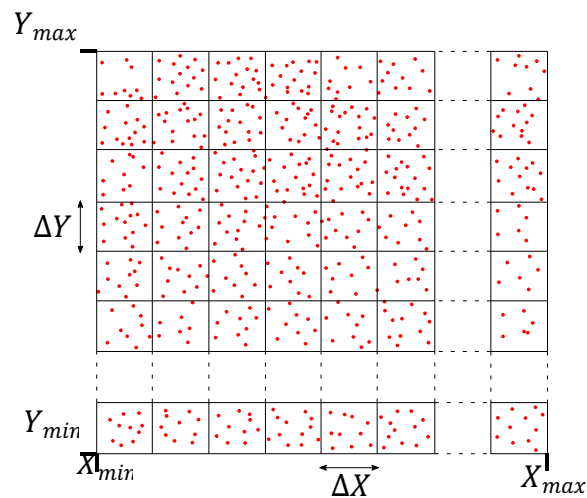
$$z=f(x,y)=\sum_{i,j=0}^n a_{ij}x^i y^j \quad n \in \mathbb{N}^* \quad (39)$$

Como o objetivo é aproximar o relevo do terreno de uma pequena região, a utilização de polinômios de primeiro grau pode acarretar em uma representação grosseira da superfície, uma vez que será ajustado um plano. Já a escolha de polinômios com graus elevados (quarto ou

³⁵ Expressas por: $z=f(x,y)=a_0x^2+a_1y^2+a_2xy+a_3x+a_4y+a_5$.

superiores) pode resultar na superparametrização da superfície, tornando o processo computacional mais oneroso. Opta-se, portanto, pelo uso de polinômios de segundo ou terceiro grau, dependendo da extensão da área de interesse. O próprio algoritmo determina a ordem do polinômio (2 ou 3) de acordo com o resíduo do ajuste. O processo de filtragem inicia-se com a subdivisão da nuvem de pontos em uma malha regular de células retangulares, assim como ilustrado na Figura 24. Em cada uma das células são escolhidos alguns pontos de menor altitude como semente para o ajuste da superfície polinomial.

Figura 24 - Subdivisão da nuvem de pontos em uma malha regular de células retangulares.



Fonte: Adaptado de Xiong *et al.* (2012) e Santos (2015).

Todos os pontos contidos na célula são classificados com relação à superfície ajustada a partir da altitude do ponto (z) e a prevista pelo polinômio utilizado, ou seja $f(x,y)$. Caso o valor absoluto da diferença seja inferior a um dado limiar (ϵ) o ponto é classificado como terreno e, portanto, será considerado como semente na próxima iteração. A cada nova iteração o limiar é reduzido por um fator $1/\sqrt{2}$ (determinado empiricamente a partir dos experimentos realizados), e o processo se repete até que não ocorram alterações na classificação dos pontos. Na sequência é mostrado o pseudo-código (Algoritmo 2) que permite o ajuste progressivo de superfícies polinomiais, conforme descrito nesta seção.

Algoritmo 2 – Filtragem da nuvem de pontos por ajuste progressivo de superfícies polinomiais.

Entrada: nuvem de pontos após a remoção dos pontos espúrios.

Saída: nuvem de pontos cujos pontos estão rotulados como terreno ou objeto.

- 1: ler a nuvem de pontos;
 - 2: calcular o retângulo envolvente: $(X_{\min}, Y_{\min}), (X_{\max}, Y_{\max})$;
 - 3: dividir a nuvem de pontos em um conjunto regular de células;
 - 4: **para cada** célula, **faça**:
 - 5: buscar por pontos semente;
 - 6: **caso** seja a primeira iteração, **então**:
 - 7: utilizar as menores altitudes;
 - 8: **caso contrário**:
 - 9: utilizar os pontos classificados como terreno na iteração anterior;
 - 10: ajustar a superfície polinomial (grau 2) nos pontos semente (aproximação do terreno);
 - 11: **caso** o resíduo do ajuste seja maior que um limiar, **então**:
 - 12: ajustar novamente a superfície utilizando um polinômio de grau 3;
 - 13: **para cada** ponto da célula, **faça**:
 - 14: calcular $\Delta z = z - f(x, y)$;
 - 15: **caso** $|\Delta z|$ seja menor que o limiar (ϵ), **então**:
 - 16: rotular o ponto como terreno;
 - 17: **caso contrário**:
 - 18: rotular o ponto como objeto;
 - 19: **caso** tenham ocorrido alterações na classificação dos pontos de terreno, **então**:
 - 20: atualizar o valor do limiar ($\epsilon_{i+1} = \epsilon_i / \sqrt{2}$);
 - 21: retornar ao passo 4;
 - 22: **caso contrário**:
 - 23: finalizar.
-

4.5 Detecção de regiões planas

Após a filtragem da nuvem de pontos, o classificador supervisionado baseado em RNA é aplicado apenas nos pontos de objetos com o intuito de discriminá-los entre telhados de edificações e vegetação³⁶. No entanto, este processamento empregará os atributos integrados à nuvem de pontos durante a etapa de preparação dos dados, isto é, informação radiométrica,

³⁶ Neste caso a classe vegetação é obtida por exclusão, ou seja, a partir dos pontos que não foram identificados como pertencentes ao terreno ou a algum plano.

autovalores e os vetores normais, os quais servem como sinal de entrada (estímulo) na primeira camada da rede neural.

A informação radiométrica é particularmente útil para evitar que pontos de vegetação sejam classificados como telhado, entretanto não é necessário utilizar as três componentes coloridas (RGB) como atributos na RNA. Neste sentido, transforma-se estas informações para o espaço de cor HSV (*Hue, Saturation and Value*) e utiliza-se apenas a componente matiz (H). Portanto, definiu-se a topologia da RNA com três neurônios na camada de entrada para os atributos descritos (ver Eq. 40), e dois neurônios na camada de saída (plano ou não). Como componentes de \mathbf{x}_p , que irá compor o vetor dos atributos, são considerados os seguintes elementos: a métrica $\Lambda(\mathbf{p})$, estimada em função dos autovalores calculados para a vizinhança de \mathbf{p} (Eq. 23), o valor do ângulo $\Omega(\mathbf{n}_p)$, e finalmente o valor do matiz (*hue*) em \mathbf{p} (H_p), ou seja:

$$\mathbf{x}_p = [\Lambda(\mathbf{p}) \quad \Omega(\mathbf{n}_p) \quad H_p]^T \quad (40)$$

Por se tratar de um problema em que as classes não são linearmente separáveis, optou-se por uma RNA com duas camadas ocultas. O treinamento do classificador é realizado com dados de referência obtidos por processamentos e edições manuais no programa Trimble SCOP++. Desta forma, uma vez determinados os pesos das conexões e os limiares das funções de ativação, a RNA poderá ser utilizada para detectar planos em outros conjuntos de dados que possuam as mesmas características.

4.6 Algoritmo RANSAC

O *Random Sample Consensus* (RANSAC) é um algoritmo que gera soluções para um modelo matemático usando a quantidade mínima de observações necessárias para estimar os parâmetros desse modelo. No método dos mínimos quadrados (MMQ) assume-se que os erros grosseiros e sistemáticos foram removidos, restando apenas os de origem aleatória. Quando essa suposição não é válida, é necessário utilizar outras abordagens, como o algoritmo RANSAC, que usa o menor conjunto de dados, na tentativa de acrescentar iterativamente as observações corretas a um conjunto consistente (consenso). O resultado esperado é um conjunto de dados composto por *inliers*, ou seja, este algoritmo pode ser utilizado para obter os dados de entrada para o MMQ.

De acordo com Fischler e Bolles (1981), assumindo um modelo matemático ϕ com n parâmetros, são necessárias $m \geq n$ observações para realizar o ajustamento. Considerando um conjunto de pontos p que contenha no mínimo n *inliers*, o algoritmo RANSAC pode ser sintetizado nas seguintes etapas:

1. Selecionar aleatoriamente n observações do conjunto de dados;
2. Resolver o modelo matemático ϕ com as observações selecionadas;
3. Quantificar quantas das observações do conjunto p atendem a algum critério de convergência (contar os *inliers*);
4. Verificar se a quantidade de *inliers* é maior do que a melhor solução até a iteração atual. Se for, atualizar a solução atual como sendo a melhor solução (solução consenso);
5. Verificar se o limite de iterações foi atingido. Se não, voltar ao passo 1. Caso contrário, solucionar o modelo matemático ϕ pelo MMQ utilizando o conjunto consenso obtido.

A quantidade de iterações pode ser determinada de diversas maneiras, entretanto, quando há informações sobre o conjunto de dados é possível fazer uma estimativa mais acurada com base em probabilidades. Considerando z como a probabilidade de selecionar apenas *inliers* em ao menos uma das k iterações, e w como a proporção estimada entre *inliers* para *outliers* no conjunto de dados, o número de iterações pode ser calculado como:

$$k = \frac{\log(1 - z)}{\log(1 - w^n)}. \quad (41)$$

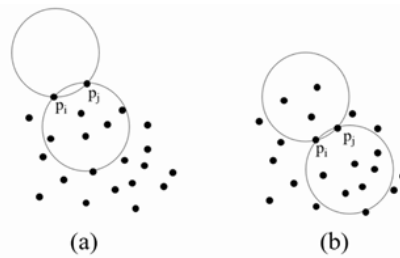
4.7 Algoritmo α -shape

O algoritmo α -shape consiste na seleção dos segmentos de reta que compõem o contorno de um dado polígono a partir de um conjunto de pontos no espaço 2D e o parâmetro α (EDELSBRUNNER, KIRKPATRICK e SEIDEL, 1983). De acordo com Shen (2008) este algoritmo foi inicialmente desenvolvido pela comunidade de Computação Gráfica e posteriormente aplicado no processamento de dados SVLA devido a algumas vantagens, tais como a capacidade de detectar polígonos côncavos e convexos, bem como contornos internos e externos. O desafio no uso desse algoritmo está relacionado à seleção de um valor ótimo para

o parâmetro α , uma vez que esse depende do espaçamento médio dos pontos e do nível de detalhamento desejado para os contornos (SANTOS, GALO e CARRILHO, 2019).

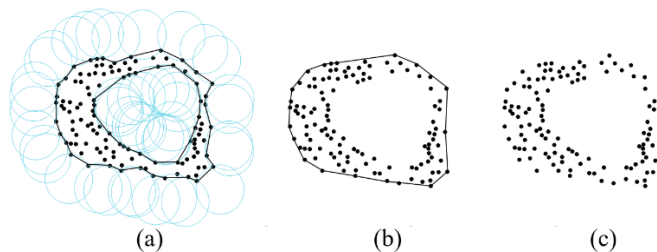
Considerando quaisquer dois pontos p_i e p_j não coincidentes, podem ser definidos dois círculos com raio α que contenham esses dois pontos, assim como ilustra a Figura 25. O critério estabelecido no algoritmo α -shape consiste em rotular os pontos p_i e p_j como pontos de contorno, caso não exista nenhum outro ponto no interior de um dos dois círculos. No caso em que o valor de α tende ao infinito ($\alpha \rightarrow \infty$), o contorno coincide com o fecho convexo. Quando o valor α tende a zero ($\alpha \rightarrow 0$), todos os pontos são considerados como pontos de contorno, ou seja, o resultado coincide com o conjunto de pontos original (Figura 26).

Figura 25 – Determinação de segmento de borda com o algoritmo em (a) dois pontos de borda e em (b) dois pontos internos.



Fonte: Adaptado de Shen *et al.* (2011).

Figura 26 - Resultados obtidos considerando (a) $\alpha = 2$, (b) $\alpha = \infty$ e (c) $\alpha = 0$.



Fonte: Adaptado de Carrilho *et al.* (2019).

4.8 Algoritmo SLIC

De acordo com Achanta *et al.* (2012), dentre as diversas abordagens possíveis para realizar a segmentação de imagens, o algoritmo Agrupamento Iterativo Linear Simples - *Simple Linear Iterative Clustering* (SLIC) é um dos mais indicados quando o objetivo é a interpretação

de imagens devido às características dos resultados. Além de ser simples, o algoritmo consegue agrupar pixels em segmentos de tamanho e forma similares.

Assumindo uma imagem com $W \times H = N$ pixels, a primeira etapa do algoritmo é a seleção da quantidade de pontos sementes (k) que devem ser gerados. Os pontos sementes são posicionados em uma grade regularmente espaçada sobre a imagem, cuja distância é de $S = [N/k]^{1/2}$. Após isso, são calculados os gradientes em uma janela 3×3 centrada em cada ponto semente, cuja posição é movida para o pixel com o menor gradiente nesta janela. Esta etapa é importante para evitar que os pontos sementes sejam posicionados sobre bordas ou em locais ruidosos. Cada ponto semente é definido como:

$$C_i = [l_i \ a_i \ b_i \ x_i \ y_i]^T \quad i = 1 \dots k \quad (42)$$

Neste vetor, as primeiras três componentes correspondem ao espaço de cor CIELAB, enquanto as outras duas são a posição do ponto semente. O algoritmo funciona de maneira similar ao k -médias, agregando pixels vizinhos a cada ponto semente. Esta etapa iterativa é repetida até que não ocorra mais nenhuma alteração nos agrupamentos. Assim como pode ser visto em Achanta *et al.* (2012), a métrica utilizada para identificar a qual agrupamento um determinado pixel pertence leva em conta tanto o espaço de cores quanto o espaço geométrico. A distância é definida como:

$$D = \sqrt{d_c^2 + \left(\frac{d_s}{S}\right)^2 m^2} \quad (43)$$

em que:

$$d_c = \sqrt{(l_j - l_i)^2 + (a_j - a_i)^2 + (b_j - b_i)^2} \quad (44)$$

corresponde à distância no espaço de cor CIELAB, e:

$$d_s = \sqrt{(x_j - x_i)^2 + (y_j - y_i)^2} \quad (45)$$

à distância euclidiana entre dois pixels i e j . Na Eq. 40, $m \in [1,40]$ (Achanta *et al.*, 2012) é uma constante que pondera a importância entre a distância no espaço de cores CIELAB (d_c) e a distância espacial (d_s) entre os pixels. O algoritmo SLIC é interessante justamente pelo fato de utilizar desse espaço de cores e desse fator de ponderação, permitindo a geração de segmentos grandes, sem que a distância espacial interfira demais no resultado, isto é, o contorno dos segmentos é mais detalhado de acordo com a tonalidade dos pixels.

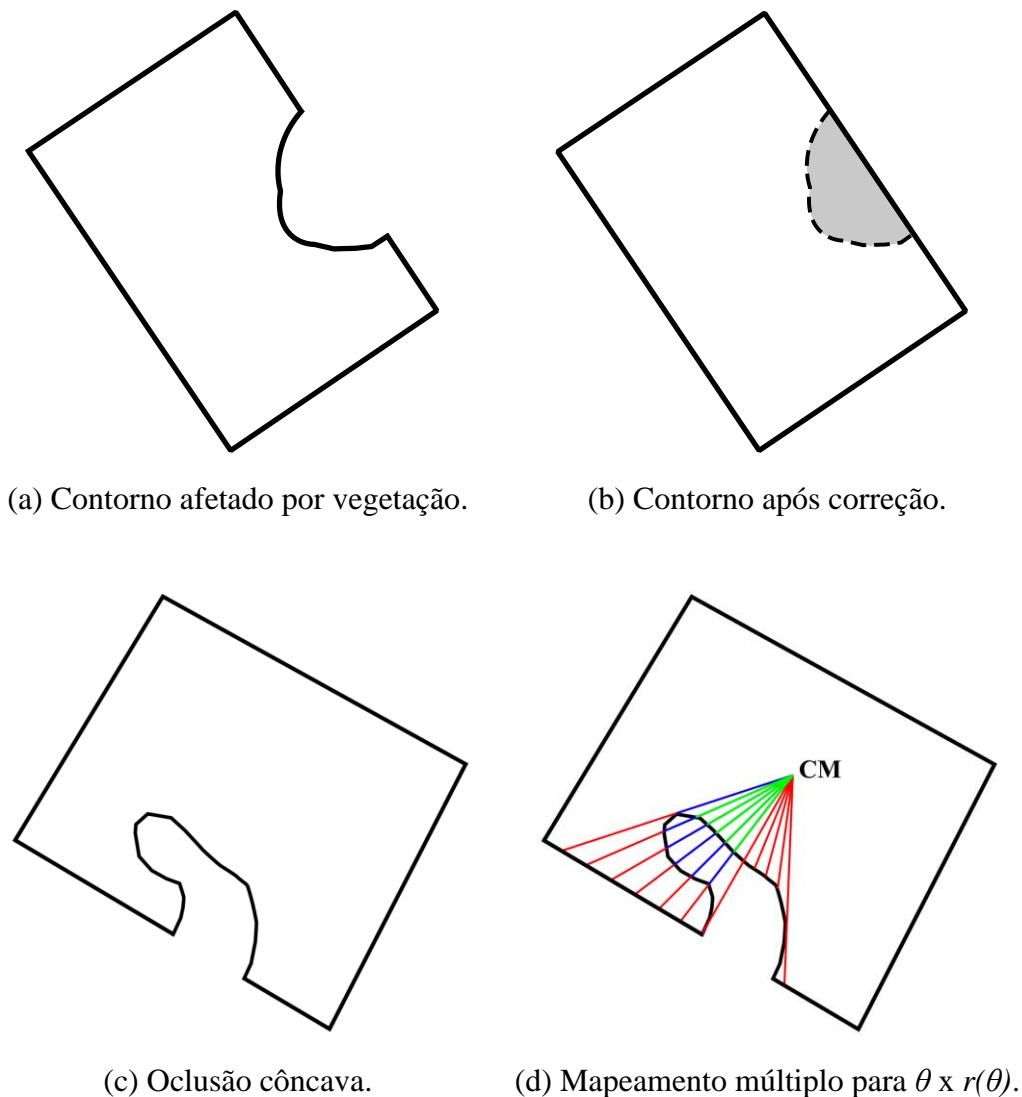
4.9 Refinamento do contorno dos telhados

Embora seja esperado e desejável a robustez do procedimento proposto, é inevitável que ocorram situações problemáticas ocasionadas pela complexidade da cena, as quais necessitam de processamentos adicionais. Com base nisso, é proposta a última etapa do método, a qual consiste na geração e análise da assinatura dos contornos das edificações, cujo intuito é detectar as situações problemáticas, como por exemplo a presença de silhuetas de copas de árvores sobrepostas às edificações. A análise da assinatura pode auxiliar também na validação de pequenos telhados a partir da geometria de seu contorno. As silhuetas vistas na Figura 27 são exemplos de telhados cujas assinaturas dos contornos se apresentam com anomalia. Tais anomalias poderão ser detectadas a partir da correspondência com contornos de referência. A parte obstruída (Figura 27a) pode ser substituída pela correspondente na referência, de modo a corrigir o contorno (Figura 27b).

Outras situações mais críticas podem ocorrer, como a oclusão em uma quina, ou ter um formato côncavo, assim como apresentado na Figura 27c. Para oclusões em quinas torna-se difícil até para operadores humanos restituírem o contorno corretamente. Entretanto, para oclusões côncavas, é possível adotar um critério durante a geração da assinatura, como visto na Figura 27d. Embora ocorram mais de uma intersecção entre o centro de massa (CM) do telhado e o polígono do contorno (registradas em verde e azul), o método proposto adota apenas a maior distância ($r(\theta)$) para cada valor de θ , isto é, os segmentos em vermelho.

O princípio da adoção deste critério é de que é mais provável que as intersecções com vértices mais distantes do CM sejam válidas, isto é, embora alguma oclusão tenha sido detectada, o contorno naquele segmento não foi afetado. Desta forma, apenas a parte convexa da oclusão é registrada na assinatura, simplificando o processo de reconstrução.

Figura 27 - Correção do contorno do telhado afetado por silhueta de vegetação (copa de árvore) sobreposta a edificação.



Fonte: Elaborado pelo autor.

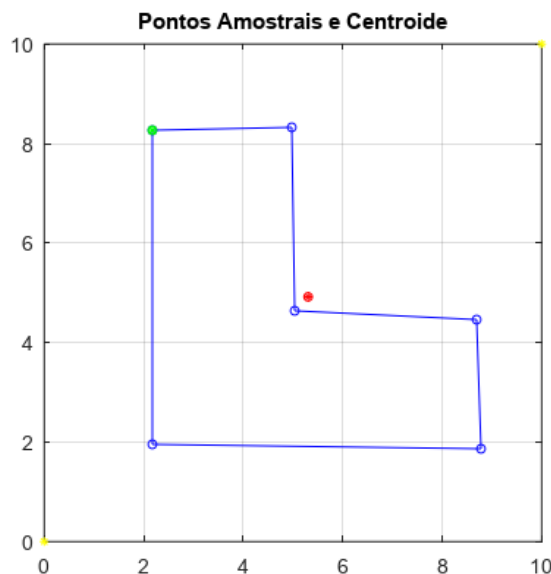
4.10 Seleção do modelo matemático para o ajuste de segmentos de contorno

O método proposto depende do conhecimento da assinatura do contorno do telhado pelo descritor $r(\theta)$, calculado a partir do centro de massa do polígono. Neste método pressupõe-se que os segmentos de curva compreendidos no intervalo $[0, 2\pi]$ sejam contínuos. Caso isso não ocorra em algum dos segmentos, é possível interpretar que exista alguma deformidade no contorno devido a oclusões, como visto nas Figura 27 e 28. A primeira dificuldade surge na seleção do modelo matemático adequado para o descritor $r(\theta)$. Como pode-se ver na Figura 15,

para um objeto circular tem-se uma assinatura linear, uma vez que para qualquer valor de θ o valor de $r(\theta)$ será o mesmo. Como pode-se notar o mesmo não acontece para objetos com outra forma, como na Figura 15, onde percebe-se que as curvas correspondentes às assinaturas de objetos quadrados e triangulares não são lineares, podendo-se notar visualmente que elas podem corresponder a segmentos de circunferência, parábolas, elipses ou outras curvas planas. Neste sentido, foram realizados alguns testes com o propósito inicial de permitir um melhor entendimento, e conseqüentemente obter de forma algébrica um modelo matemático que permita modelar as assinaturas geradas para diferentes edificações.

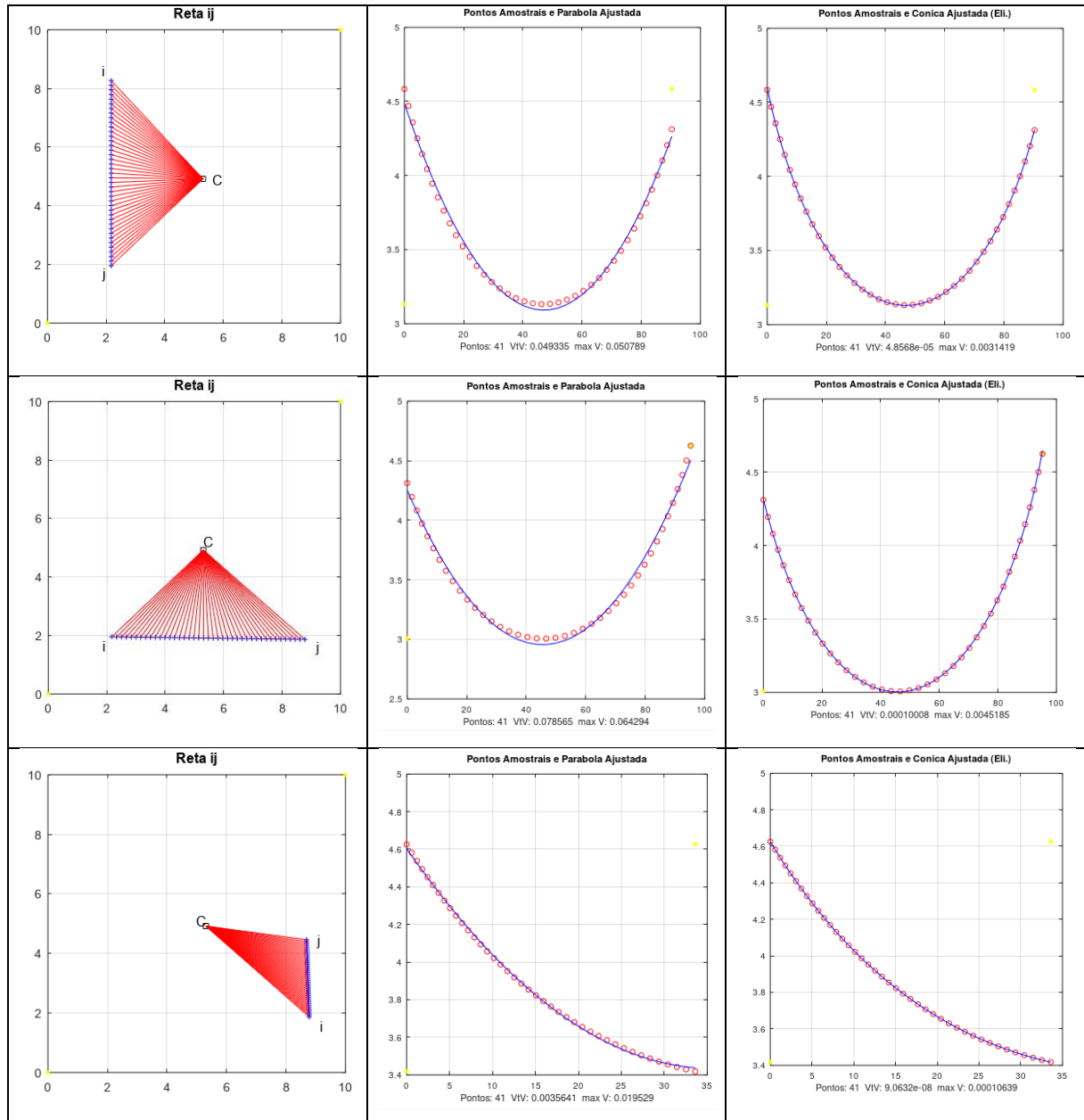
Na Figura 28 apresenta-se um objeto na forma de L, composto por 6 vértices. Para cada um dos segmentos deste objeto são mostradas as assinaturas na forma de um gráfico na forma $\theta \times r(\theta)$, bem como o ajuste de uma parábola e de uma cônica.

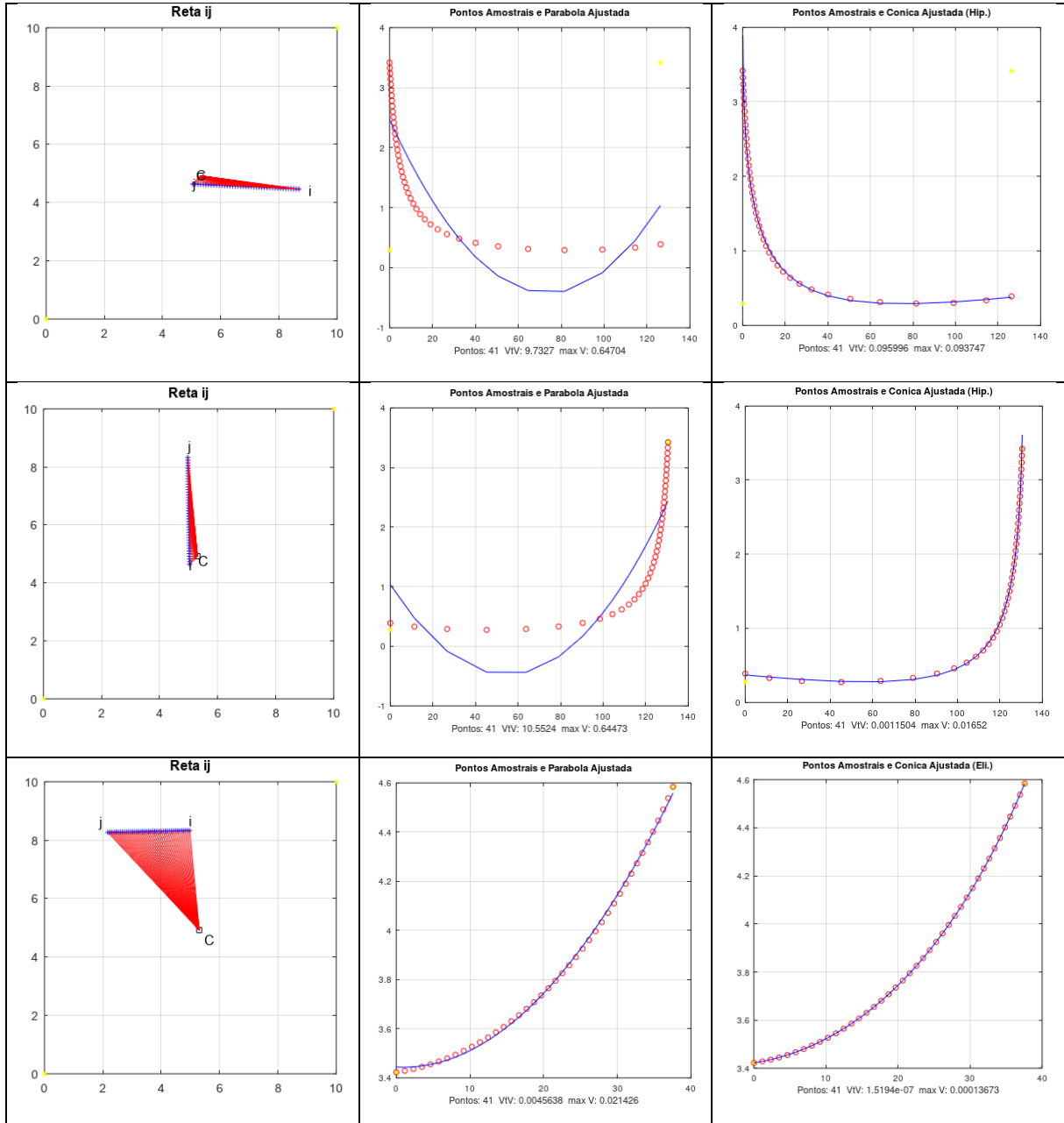
Figura 28 – Contorno de telhado de edificação em forma de L simulado, no qual o centro de massa (ponto na cor vermelha) encontra-se externo ao polígono.



Como pode-se observar na Figura 29, em algumas situações o ajuste de uma curva parabólica se mostrou adequada, mas em algumas situações curvas cônicas como elipses ou até hipérbolas se ajustam melhor aos pontos amostrais da função $\theta \times r(\theta)$, o que induz a considerarmos que curvas cônicas podem ser utilizadas para a modelagem das assinaturas mostradas.

Figura 29 – Cada um dos segmentos de telhado mostrado na Figura 28 e respectivas curvas ajustadas. Observa-se que em algumas situações a parábola não atende (resíduo alto), sendo necessário a utilização de uma outra cônica para modelar a curva dada por $(\theta, r(\theta))$.

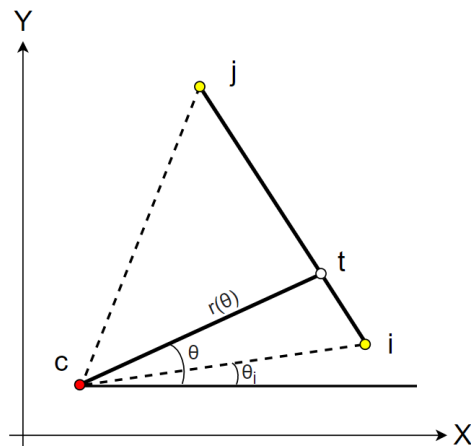




O intuito do desenvolvimento apresentado na sequência, feito a partir do que é mostrado na Figura 30, é obter uma equação em termos de (x, y) que descreva algebricamente o comportamento do descritor geométrico $r(\theta)$. Considerando esta figura, bem como o segmento de reta \bar{ij} que liga os pontos extremos desse segmento, ou seja, $p_i = (x_i, y_i)$ e $p_j = (x_j, y_j)$, é possível obter as coordenadas do ponto $t = (x(t), y(t))$ a partir da equação da reta na forma paramétrica:

$$\begin{cases} x(t) = x_i + t(x_j - x_i) \\ y(t) = y_i + t(y_j - y_i) \end{cases} \quad (46)$$

Figura 30 – Geração do descritor $r(\theta)$ para um segmento de reta \bar{ij} e do centro de massa (C) de um polígono referente ao contorno do telhado de uma edificação.



A partir das Equações 46 pode-se observar que para um dado valor de t tem-se um ponto do segmento \bar{ij} que pode ser determinado a partir das coordenadas polares $(t, r(\theta))$ ou $(\theta, r(\theta))$. Deste modo, é importante obter uma relação que envolva as variáveis t e θ . Assim sendo, assumindo que o centro de massa C possua coordenadas $C = (x_c, y_c)$, a seguinte relação trigonométrica pode ser obtida a partir da Figura 30:

$$\tan \theta = \frac{y_{(t)} - y_c}{x_{(t)} - x_c} \quad (47)$$

sendo possível determinar a variável t fazendo:

$$x_{(t)} \tan \theta - x_c \tan \theta = y_{(t)} - y_c \quad \Rightarrow \quad x_{(t)} \tan \theta - y_{(t)} = x_c \tan \theta - y_c.$$

Substituindo na equação acima as Eq(s). 46 e desenvolvendo de modo a isolar t , tem-se:

$$\begin{aligned} [x_i + t(x_j - x_i)] \tan \theta - [y_i + t(y_j - y_i)] &= x_c \tan \theta - y_c \\ x_i \tan \theta + t(x_j - x_i) \tan \theta - y_i - t(y_j - y_i) &= x_c \tan \theta - y_c \\ t[(x_j - x_i) \tan \theta - (y_j - y_i)] + x_i \tan \theta - y_i &= x_c \tan \theta - y_c \\ t[(x_j - x_i) \tan \theta + y_i - y_j] &= (x_c - x_i) \tan \theta + y_i - y_c \end{aligned}$$

$$t = \frac{(x_c - x_i) \tan \theta + y_i - y_c}{(x_j - x_i) \tan \theta + y_i - y_j} \quad (48)$$

Deste modo, dado o valor de θ pode-se obter o valor de t e, inversamente, pode-se obter a equação que permite determinar θ a partir de t , conforme segue:

$$\theta = \arctan\left(\frac{t(y_j - y_i) + y_i - y_c}{t(x_j - x_i) + x_i - x_c}\right). \quad (49)$$

Considerando que o interesse é obter o valor de θ para um ponto t qualquer do segmento que liga os pontos i e j , e que para o ponto i tem-se o ângulo θ_i , pode-se verificar que a mudança do valor de θ_i implicará apenas na translação da assinatura, não afetando a sua forma. Aplicando a lei dos cossenos ao triângulo com vértices em i , C e t , tem-se:

$$d_{it}^2 = d_{ci}^2 + r_{(\theta)}^2 - 2d_{ci}r_{(\theta)}\cos(\theta - \theta_i) \quad (50)$$

A distância d_{it} poderá ser obtida por:

$$d_{it}^2 = (x_t - x_i)^2 + (y_t - y_i)^2,$$

onde cada uma dessas componentes poderá ser escrita em função da Equação 46, obtendo-se:

$$d_{it}^2 = (x_t - x_i)^2 + (y_t - y_i)^2 = t^2(x_j - x_i)^2 + t^2(y_j - y_i)^2 = t^2d_{ij}^2. \quad (51)$$

Substituindo a Eq. 51 na 50 chega-se a:

$$t^2d_{ij}^2 = d_{ci}^2 + r_{(\theta)}^2 - 2d_{ci}r_{(\theta)}\cos(\theta - \theta_i) \quad (52)$$

Desenvolvendo o termo do cosseno da Equação 52, considerando que $\cos(\theta_i) = a$ e $\sin(\theta_i) = b$, sendo a e b constantes, e ainda usando a Equação 46 obtém-se:

$$\begin{aligned} t^2d_{ij}^2 &= d_{ci}^2 + r_{(\theta)}^2 - 2d_{ci}r_{(\theta)}[\cos\theta \cdot \cos\theta_i + \sin\theta \cdot \sin\theta_i] \\ t^2d_{ij}^2 &= d_{ci}^2 + r_{(\theta)}^2 - 2d_{ci}r_{(\theta)}[a \cdot \cos\theta + b \cdot \sin\theta] \\ t^2d_{ij}^2 &= d_{ci}^2 + r_{(\theta)}^2 - 2d_{ci} \cdot a \cdot r_{(\theta)}\cos\theta - 2d_{ci} \cdot b \cdot r_{(\theta)}\sin\theta \end{aligned} \quad (53)$$

Por outro lado, os termos $r_{(\theta)}\cos\theta$ e $r_{(\theta)}\sin\theta$ podem ser escritos, a partir da Figura 30, por:

$$\begin{aligned} r_{(\theta)}\cos\theta &= x_t - x_c = x_i + t(x_j - x_i) - x_c = t(x_j - x_i) + x_i - x_c \\ r_{(\theta)}\sin\theta &= y_t - y_c = y_i + t(y_j - y_i) - y_c = t(y_j - y_i) + y_i - y_c \end{aligned}$$

Aplicando este resultado na Equação 53 chega-se a:

$$t^2 d_{ij}^2 = d_{ci}^2 + r_{(\theta)}^2 - 2d_{ci} \cdot a(t(x_j - x_i) + x_i - x_c) - 2d_{ci} \cdot b.(t(y_j - y_i) + y_i - y_c)$$

$$t^2 d_{ij}^2 = d_{ci}^2 + r_{(\theta)}^2 + t[2d_{ci} \cdot a(x_i - x_j) + 2d_{ci} \cdot b(y_i - y_j)]$$

$$+ 2d_{ci} \cdot [a(x_c - x_i) + b(y_c - y_i)]$$

Como pode-se notar, considerando que os pontos i, j e C são conhecidos *a priori*, pode-se perceber que vários termos desta equação são constantes. Rearranjando os termos em uma outra ordem pode-se escrever:

$$d_{ij}^2 t^2 - r_{(\theta)}^2 + t[2d_{ci} \cdot a(x_j - x_i) + 2d_{ci} \cdot b(y_j - y_i)] - d_{ci}^2 + 2d_{ci}[a(x_i - x_c) + b(y_i - y_c)] = 0$$

Considerando as seguintes trocas de variáveis:

$$A = d_{ij}^2,$$

$$C = -1,$$

$$D = 2d_{ci} \cdot a(x_j - x_i) + 2d_{ci} \cdot b(y_j - y_i),$$

$$F = 2d_{ci} \cdot [a(x_i - x_c) + b(y_i - y_c)] - d_{ci}^2,$$

onde todos estes termos são constantes, pode-se escrever a equação anterior por:

$$F(t, r_{(\theta)}) = At^2 + Cr_{(\theta)}^2 + Dt + F = 0 \quad (54)$$

Como pode-se notar nesta equação, escrita na forma implícita, ela relaciona as variáveis $(t, r_{(\theta)})$ ou $(t(\theta), r(\theta))$, uma vez que t é função de θ . De acordo com Boldrini (1984) a equação geral das cônicas no plano (\mathbb{R}^2) é definida como:

$$Ax^2 + Bxy + Cy^2 + Dx + Ey + F = 0 \quad (55)$$

A partir dos coeficientes da Equação 55 é possível classificar uma determinada curva cônica a partir do cálculo de $\Delta = B^2 - 4AC$, do seguinte modo:

$$\begin{cases} \Delta < 0 & \text{Elipse} \\ \Delta = 0 & \text{Parábola} \\ \Delta > 0 & \text{Hipérbole} \end{cases}$$

Fazendo a comparação da equação $F(t, r_{(\theta)})$ com a equação geral das curvas cônicas pode-se notar que a curva $F(t, r_{(\theta)})$ corresponde ao caso particular de uma cônica na qual alguns termos se anulam. Pode-se notar ainda que, neste caso particular da curva $F(t, r_{(\theta)})$, a classificação da cônica será obtida a partir da análise do sinal de $\Delta = -4AC$. Assim sendo, considerando o modelo obtido, dado por $F(t, r_{(\theta)})$, uma vez que ele correspondente a um caso particular de uma curva cônica, cuja equação geral é dada pela Eq. 55, é possível modelar as curvas obtidas pela assinatura das bordas lineares das edificações a partir tanto da função $F(t, r_{(\theta)})$ deduzida quanto da equação geral da cônica, que foi o modelo escolhido para realizar a modelagem.

É importante destacar que o desenvolvimento realizado se refere ao caso de assinaturas de bordas lineares de edificações, que estão sujeitas tanto ao efeito da amostragem de pontos por um sistema a LASER aerotransportado quanto ao efeito da oclusão de pontos provocados por vegetação, por exemplo, como ilustrado na Figura 27, o que modifica as respectivas assinaturas. No entanto, dependendo da magnitude dos efeitos da amostragem e da oclusão, algumas dessas bordas podem ser reconstruídas.

4.11 Avaliação quantitativa dos resultados

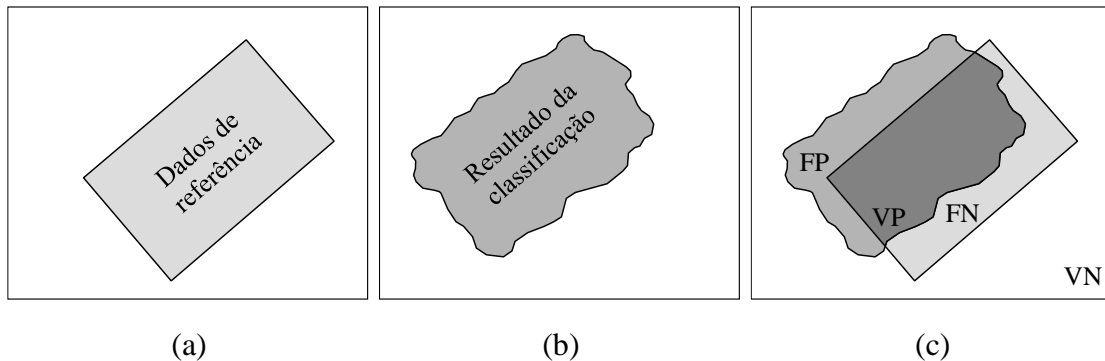
A validação de qualquer método de processamento de dados espaciais requer uma avaliação estatística da conformidade dos resultados com relação aos dados de referência³⁷ (RICHARDS, 2013). Na literatura existem diferentes abordagens utilizadas para estimar a qualidade dos resultados, acarretando em falta de consenso e dificuldades para comparar a performance dos métodos a partir dos resultados publicados. A falta de padronização está relacionada aos critérios utilizados durante a seleção das áreas para realizar a análise e quais métricas devem ser consideradas para esta avaliação. Observa-se que no geral os autores adotam elementos derivados da matriz de erro (Figura 31), também conhecida como matriz de confusão ou contingência.

Em Viera e Mather (2005) é apresentada uma revisão dos métodos tradicionais utilizados no controle de qualidade de produtos cartográficos gerados a partir de dados de sensoriamento

³⁷ De acordo com Richards (2013), a expressão verdade terrestre (*ground truth*) foi utilizada por muito tempo como sinônimo para dados de referência. Atualmente prefere-se o termo referência por ser menos dogmático.

remoto. Para a acurácia posicional os autores sugerem a correspondência entre entidades geométricas (pontos ou retas, principalmente) observadas nos dados e seus homólogos na referência, enquanto a acurácia temática foi baseada nas estatísticas obtidas pela matriz de erro, isto é, abordagens de controle de qualidade por pixel.

Figura 31 - Terminologia dos elementos derivados da matriz de erro. Em (c): verdadeiro positivo (VP); verdadeiro negativo (VN); falso positivo (FP); e falso negativo (FN).



Fonte: Adaptado de Buján *et al.* (2012).

Para o caso específico de uma classificação binária, como no problema da filtragem de nuvens de pontos, por exemplo, Sithole e Vosselman (2004) avaliaram a performance de oito algoritmos a partir dos erros tipo I e II. Os autores geraram dados de referência manualmente para diferentes conjuntos de dados, os quais foram processados por todos os algoritmos considerados. Na revisão apresentada por Meng *et al.* (2010), os autores discutem os desafios para a avaliação quantitativa dos métodos de filtragem. Os trabalhos mencionados no artigo utilizam diferentes procedimentos para gerar os dados de referência, e oito das publicações citadas realizaram apenas uma avaliação qualitativa a partir de inspeções visuais.

Para classificações com mais de duas classes, Murakami *et al.* (1999) utilizaram os índices de omissão e comissão, enquanto Rottensteiner (2007), Hermosilla *et al.* (2011), Malpica *et al.* (2013), Pang *et al.* (2014) e Awrangjeb *et al.* (2015) adotaram os índices completude e nível de acerto. Buján *et al.* (2012) quantificaram os quatro índices citados, além da qualidade e F-score, que podem ser obtidos em função dos anteriores. Em Richards (2013) é apresentada uma formulação para determinar a quantidade de amostras (n) necessárias para realizar o controle de qualidade de produtos temáticos, isto é, resultantes de processos de classificação. Este modelo matemático assume que a quantidade de amostras é muito menor do que a quantidade de elementos classificados, e proporciona um nível de confiança de 95%.

$$n = \frac{4P(1 - P)}{\varepsilon^2} \quad (56)$$

sendo ε a margem de erro admitida, e P a propoção de pixels corretamente identificados.

Alguns trabalhos consideram diferentes metodologias para a seleção das áreas utilizadas na análise. Li *et al.* (2013) apresentam uma abordagem para avaliar a acurácia da segmentação de planos de telhados de edificações, a qual consiste em realizar amostras aleatórias ao invés do conjunto de dados por inteiro. Encontra-se na literatura outras possibilidades semelhantes, como a amostragem aleatória estratificada (McCOY, 2005; CONGALTON e GREEN, 2009), entretanto, é necessário delimitar os estratos antes que a análise possa ser conduzida.

Na área de reconhecimento de padrões é comum a utilização das métricas: precisão (p) e revocação (r), ou *recall* em inglês, para avaliar tanto problemas de classificação quanto de regressão (BISHOP, 1995). Estes valores podem ser estimados em função da matriz de erro, ou seja, a partir de VP, FP e FN, podendo ser obtidos por:

$$p = \frac{VP}{VP+FP} \quad r = \frac{VP}{VP+FN} \quad (57)$$

A comunidade científica das áreas de Fotogrametria e Sensoriamento Remoto adota a terminologia completudeza e nível de acerto como sinônimos de precisão e revocação, respectivamente. Existem ainda outras denominações para esses indicadores, como sensibilidade e especificidade, assim como apresentado por Sokolova *et al.* (2006) e discutido em Carrilho *et al.* (2017). Além disso, alguns autores (SOKOLOVA *et al.*, 2006; LU *et al.*, 2014; VEGA *et al.*, 2014, por exemplo) adotam um valor único conhecido por F-score, calculado como a média harmônica da precisão e revocação. O F-score é particularmente útil uma vez que leva em conta as três quantidades (VP, FP e FN). Se os pesos para p e r forem iguais, o cálculo do F-score pode ser simplificado para:

$$F\text{-score} = \frac{2}{1/p + 1/r} = \frac{2VP}{2VP+FP+FN} \quad (58)$$

Além das métricas apresentadas, existem algumas medidas que podem ser consideradas no contexto de comparação de polígonos, como a distância de Hausdorff e de Chamfer, por

exemplo. Para a comparação de resultados de extração de polígonos de telhados, Avbelj *et al.* (2015) propuseram a métrica *Polygons and Line Segments* (PoLiS) a qual é simples e robusta, tendo sido idealizada para a realização de análises quantitativas de similaridade entre polígonos. Dados os pontos pertencentes ao contorno de dois polígonos (A e B), a métrica PoLiS pode ser determinada a partir da seguinte equação:

$$p(A, B) = \frac{1}{2q} \sum_{a_j \in A} \min_{b \in \partial B} \|a_j - b\| + \frac{1}{2r} \sum_{b_k \in B} \min_{a \in \partial A} \|b_k - a\| \quad (59)$$

sendo q e r o número de vértices dos polígonos A e B , respectivamente; ∂A e ∂B correspondem aos polígonos A e B , e $\|a - b\|$ é a distância Euclidiana entre dois pontos a e b .

Capítulo 5

EXPERIMENTOS E RESULTADOS

Este capítulo apresenta uma breve descrição dos materiais utilizados e os experimentos realizados com o intuito de validar o método proposto e apresentado no capítulo anterior. Os experimentos foram feitos para todas as etapas do método, com foco na parte final que é a reconstrução das oclusões utilizando o descritor geométrico.

5.1 Materiais utilizados

Os experimentos foram realizados com dados do município de Presidente Prudente/SP, sendo composto por nuvens de pontos (ver Figuras 32, 33 e 34), os quais fazem parte do *Unesp Photogrammetric Data Set*³⁸ (TOMMASELLI *et al.*, 2018), tendo sido adquiridos e disponibilizados pela empresa Sensormap Geotecnologia. Durante os voos foram obtidas imagens aéreas de alta resolução, parâmetros de orientação exterior (POE) e dados SVLA. De acordo com Tommaselli *et al.* (2018), o SVLA utilizado foi o RIEGL modelo LMS Q680i. O artigo reporta que, após a fototriangulação, na qual obteve-se os POE refinados, o RMSE determinado a partir de doze pontos de verificação foi da ordem de 4 cm (planimétrico) e 7 cm (altimétrico).

A partir das nuvens de pontos das três regiões mostradas nas Figuras 32, 33 e 34 foram obtidos os dados de referência utilizados para realizar as análises quantitativas visando avaliar o método proposto. Inicialmente, foram removidos pontos espúrios. Utilizando as imagens aéreas e os dados de orientação, a informação RGB foi associada à nuvem utilizando a metodologia descrita anteriormente. Também com as imagens, os contornos de alguns dos telhados foram restituídos no software ERDAS IMAGINE no módulo LPS (Figura 35), exportados no formato DXF e depois convertidos para SHP (*Shapefile*). Por fim, as três nuvens de pontos foram classificadas manualmente com o software CloudCompare. Nesta etapa foram consideradas algumas classes de interesse para área urbana, e principalmente que representem a cobertura urbana.

³⁸ Disponível em: <http://www.fct.unesp.br/unespdataset>. Acessado em: 10 ago. 2017.

Figura 32 – Nuvem de pontos da Região 1: Moradia estudantil do campus da Unesp.

(a) Nuvem de pontos com coloração RGB.

(b) Nuvem classificada.

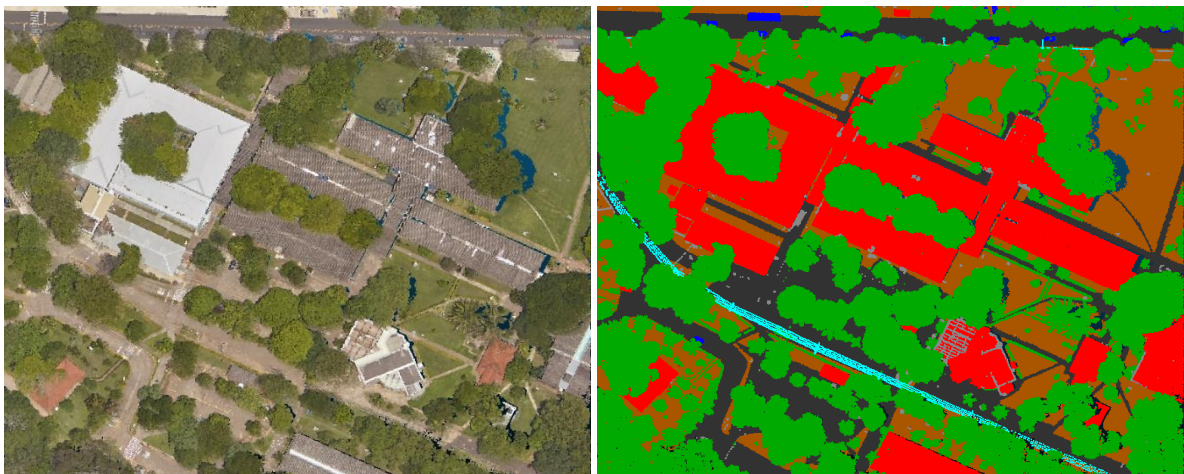


(●) Vegetação (●) Edificações (●) Linha de transmissão (●) Veículos
 (●) Solo não-pavimentado (●) Solo pavimentado (●) Não classificado

Figura 33 – Nuvem de pontos da Região 2: Diretoria e alguns dos departamentos e laboratórios do campus da Unesp.

(a) Coloração RGB.

(b) Classificada.



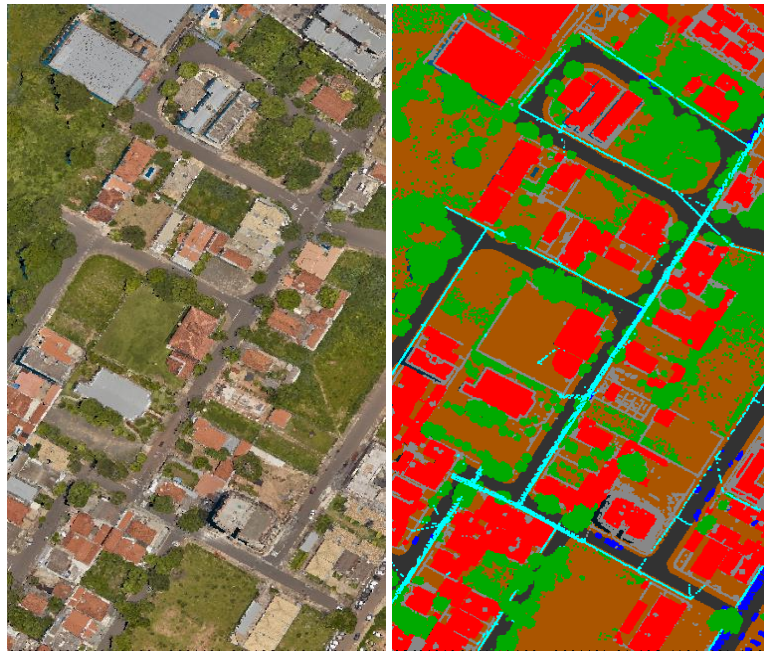
(●) Vegetação (●) Edificações (●) Linha de transmissão (●) Veículos
 (●) Solo não-pavimentado (●) Solo pavimentado (●) Não classificado

A classificação levou em conta as especificações do formato LAS proposta pela ASPRS, versão 1.4, que estabelece os códigos padrões para cada classe. Embora o interesse principal neste trabalho sejam as edificações, foram classificadas também a vegetação, linhas de transmissão, veículos, e solo pavimentado e não-pavimentado. Os pontos que não se enquadraram em nenhuma dessas classes foram rotulados como não classificados.

Figura 34 – Região 3: Bairro Morumbi, área parcialmente urbanizada.

(a) Coloração RGB.

(b) Classificada.

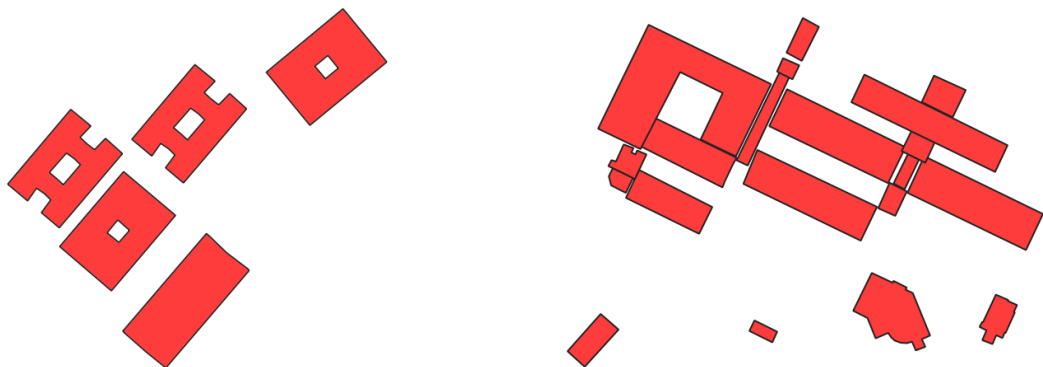


(●) Vegetação (●) Edificações (●) Linha de transmissão (●) Veículos
(●) Solo não-pavimentado (●) Solo pavimentado (●) Não classificado

Figura 35 – Contornos dos telhados obtidos a partir de restituição das imagens aéreas.

(a) Região 1 – Moradia estudantil.

(b) Região 2 – Diretoria e departamentos.



Fonte: Elaborado pelo autor.

As três regiões selecionadas correspondem a duas porções do campus da Unesp (moradia estudantil e a parte central, isto é, diretoria, laboratórios e departamentos), e a uma parte do bairro Morumbi. Nas três nuvens de pontos é possível observar telhados com obstruções causadas por vegetação, em alguns casos sendo difícil até para operadores humanos

identificarem os contornos dos telhados. Algumas informações adicionais são apresentadas na Tabela 1.

Tabela 1 – Dados das regiões selecionadas para a realização os experimentos.

| <i>Região</i> | <i>Número de pontos</i> | <i>Dimensões da região (Largura e Altura)</i> | <i>Densidade média (pontos/m²)</i> | <i>Coordenadas do centro de cada região (UTM Fuso 22S)</i> |
|---------------|-------------------------|---|---|--|
| 1 | 399.502 | L = 206,3 m A = 144,1 m | 13,4 | E = 457.768 m N = 7.553.958 m |
| 2 | 741.151 | L = 211,8 m A = 181,0 m | 19,3 | E = 457.952 m N = 7.553.712 m |
| 3 | 725.222 | L = 184,7 m A = 351,5 m | 11,2 | E = 457.762 m N = 7.552.489 m |

No decorrer do desenvolvimento do trabalho, diferentes ferramentas foram utilizadas, como indicado no fluxograma do método proposto (Capítulo 4). Para a implementação dos algoritmos utilizou-se a linguagem C++ (padrão C++14), inicialmente no ambiente de desenvolvimento integrado Code::Blocks versões 16.01 e 17.12, e posteriormente no Microsoft Visual Studio 2019. Para alguns algoritmos de álgebra linear optou-se pelo uso da biblioteca Eigen, a qual faz uso de Programação Genérica e disponibiliza diversas decomposições de matrizes com robustez numérica. As etapas que envolvem aprendizado de máquina foram majoritariamente implementadas em *scripts* Python 3.7 utilizando os pacotes *numpy*, *scikit-learn*, *Keras*, *OpenCV*, e *TensorFlow*. Para o gerenciamento dos ambientes Python optou-se pelo Anaconda3. Os softwares GNU Octave 5.2.0 e GNU Plot 5.2.8 foram utilizados para a geração de gráficos e auxílio no desenvolvimento de algumas implementações.

Algumas das etapas do método proposto demandam bastante processamento quando o conjunto de dados é grande (milhões de pontos ou mais). Desta forma, alguns recursos de paralelismo foram implementados utilizando o conceito de *Single Instruction Multiple Data* (SIMD) com o conjunto de instruções *Streaming SIMD Extensions* (SSE) da arquitetura x86 e x86_64, e *multithreading* com a *Application Program Interface* (API) *OpenMP* (*Open Multi-Processing*). Para mais detalhes sobre a implementação, verificar o Apêndice A.

Para as etapas que envolvem o pré-processamento dos arquivos de dados SVLA, foi utilizado o conjunto de ferramentas (aplicativos) rapidlasso LAStools, versões 160329 até 200304, em conjunto com o CloudCompare versões 2.8.1 e 2.11 alpha. A implementação do

método e os processamentos foram realizados em um computador AMD Ryzen7 3800X 4.5GHz, 32GB DDR4, nVIDIA GeForce RTX 2080 8GB, NVMe M.2 de 256GB e dois HDD SATA de 4TB, com sistema operacional Microsoft Windows 10 Pro x64 *build* 2004.

5.2 Experimentos e discussões

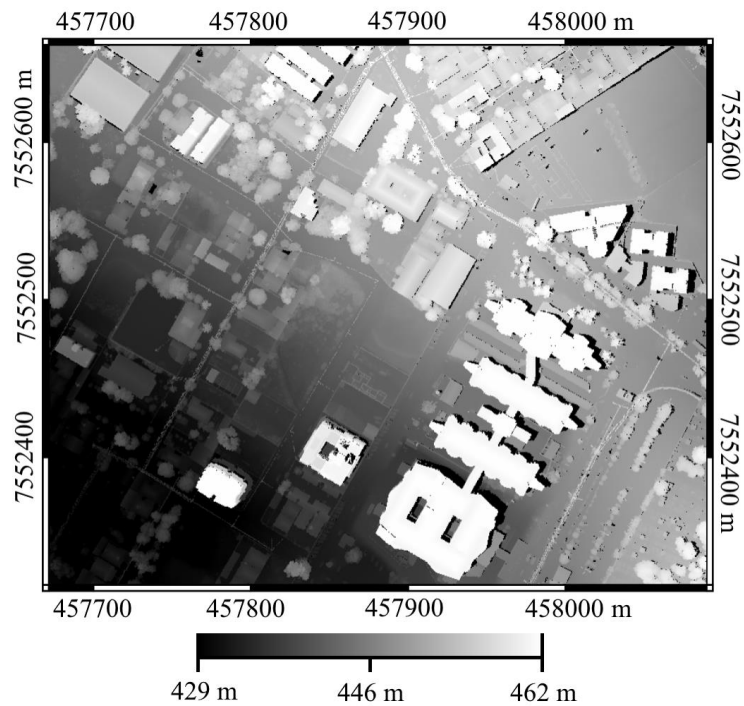
Neste trabalho, três experimentos foram realizados com o intuito de avaliar o método proposto. Embora o método contenha diversas etapas, optou-se aqui por analisar as mais importantes com relação a Tese, isto é: (i) remoção de pontos espúrios (*outliers*); (ii) extração de contorno de objetos a partir de imagens de alta resolução; (iii) extração de contorno de telhados de edificações utilizando o algoritmo α -shape; (iv) associação da informação radiométrica na nuvem de pontos (dados SVLA); (v) classificação da nuvem de pontos; e (vi) extração e reconstrução dos contornos dos telhados de edificações a partir da assinatura.

5.2.1 Remoção de pontos espúrios

Assim como pode ser visto em Carrilho e Galo (2017), a estratégia originalmente proposta para a remoção de pontos espúrios apresentava algumas limitações com relação ao relevo da região trabalhada, restringindo o uso para regiões aproximadamente planas (relevo pouco acidentado). Neste sentido, foi realizado um estudo mais aprofundado e uma revisão dos métodos já existentes (presentes em softwares como LAStools, PointCloudLibrary - PCL e CloudCompare, por exemplo), o que culminou em um novo método (filtro baseado no histograma das células - HC) apresentado em Carrilho *et al.* (2018). A área de estudos apresentada na Figura 36 foi processada manualmente, contabilizando 848 *outliers* no total, sendo 650 negativos e 198 positivos (normalmente pontos isolados).

Ao avaliar quantitativamente os métodos implementados em softwares de processamento de nuvens de pontos (LAStools e CloudCompare), foi possível observar o efeito da modificação do número de vizinhos (K) e limiar de frequência, no número de Verdadeiros Positivos, Falsos Positivos e Falsos Negativos (Figura 37 e 38). Este tipo de análise é relevante pois permite ao usuário configurar os parâmetros para realizar uma filtragem mais rigorosa visando ter mais Verdadeiros Positivos, ao mesmo tempo que reduz Falsos Positivos, por exemplo.

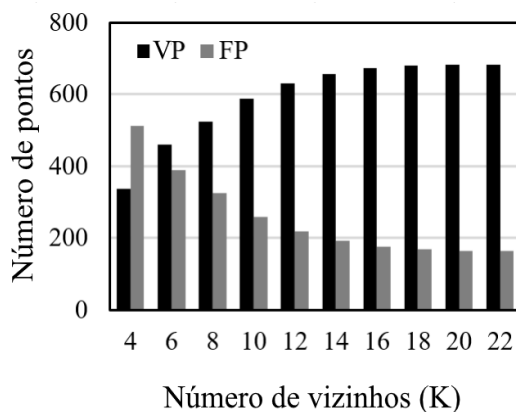
Figura 36 – Área de estudos: Proximidades da Região 3 (bairro Morumbi).
Coordenadas UTM fuso 22S.



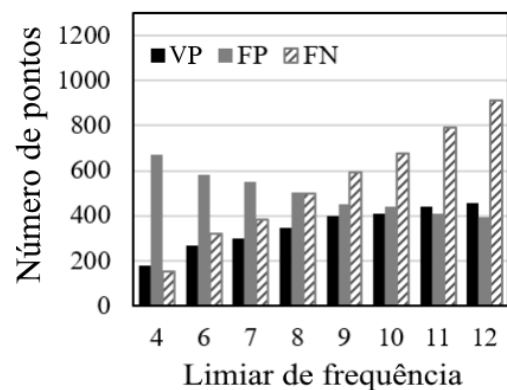
Fonte: Adaptado de Carrilho *et al.* (2018).

Figura 37 – Efeito da quantidade de pontos vizinhos (K) no filtro SOR, e do limiar de frequência no FE. Considerou-se células de 2 m para o FE.

(a) Filtro SOR – PCL/CloudCompare.



(b) LAStools (filtro espacial – FE).



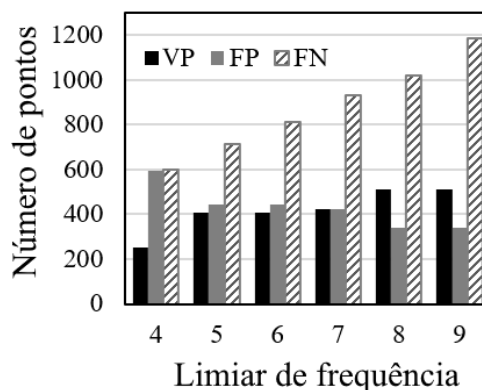
Fonte: Adaptado de Carrilho *et al.* (2018).

O método proposto consegue minimizar um pouco esse efeito, por remover apenas os pontos que estiverem fora do intervalo central do histograma de alturas. Dessa forma, embora pontos amostrados no interior das copas das árvores em paredes de edificações (Figura 39) aparentem ser *outliers* (por estarem isolados), eles não são removidos. Esse aspecto torna o

método proposto (HC) útil para estudos florestais, por exemplo, uma vez que é menos invasivo e preserva mais a geometria original da vegetação.

Figura 38 – Efeito do limiar de frequência no método de filtragem proposto (HC).

Considerou-se células quadradas de 50 m e histograma com intervalos de 15 cm.

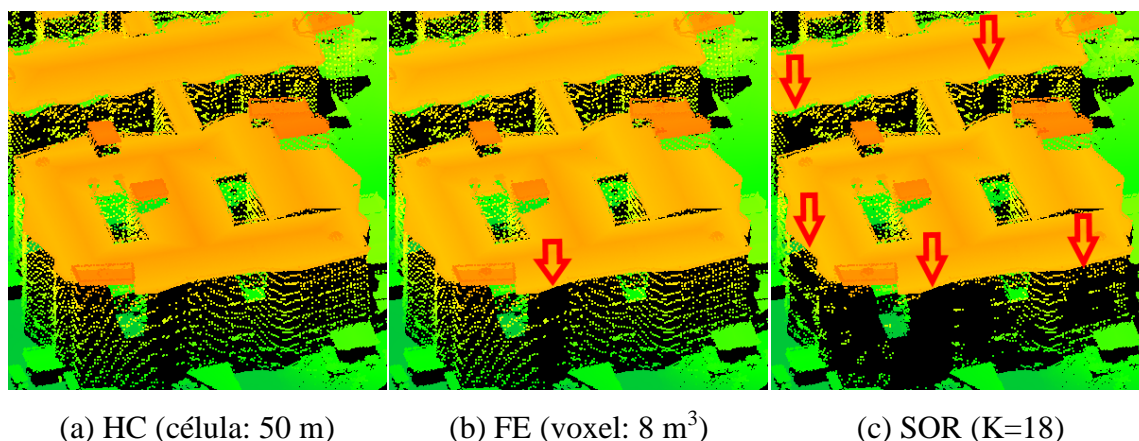


Fonte: Adaptado de Carrilho *et al.* (2018).

Uma das situações mencionadas anteriormente está ilustrada na Figura 39, na qual os pontos amostrados nas paredes das edificações são removidos quando se utiliza dos métodos FE e SOR (LAStools e PCL/CloudCompare, respectivamente). Os três métodos utilizam a mesma premissa para detectar pontos espúrios: identificar pontos isolados. A vantagem da abordagem proposta é assumir que se trata de dados da superfície terrestre obtidos por SVLA. Os demais (FE e SOR) foram concebidos para lidar com qualquer nuvem de pontos, isto é, sem qualquer conhecimento adicional sobre os dados, já que os softwares podem ser utilizados para processar nuvens de pontos terrestres.

Sumarizando, embora o método proposto tenha algumas vantagens, a conclusão dos experimentos é de que este assunto ainda requer esforço da comunidade, pois as abordagens existentes resolvem apenas parcialmente o problema, sendo ainda necessários ajustes e correções por operadores humanos. Uma das principais dificuldades é a variação de densidade local na nuvem de pontos, visto que regiões com densidades baixas estão sujeitas a serem confundidas com *outliers* por conterem pontos aparentemente isolados.

Figura 39 – Ocorrências de Falso Negativo nas paredes de edificações.



Fonte: Adaptado de Carrilho *et al.* (2018).

5.2.2 Extração do contorno de objetos utilizando RNC e SLIC

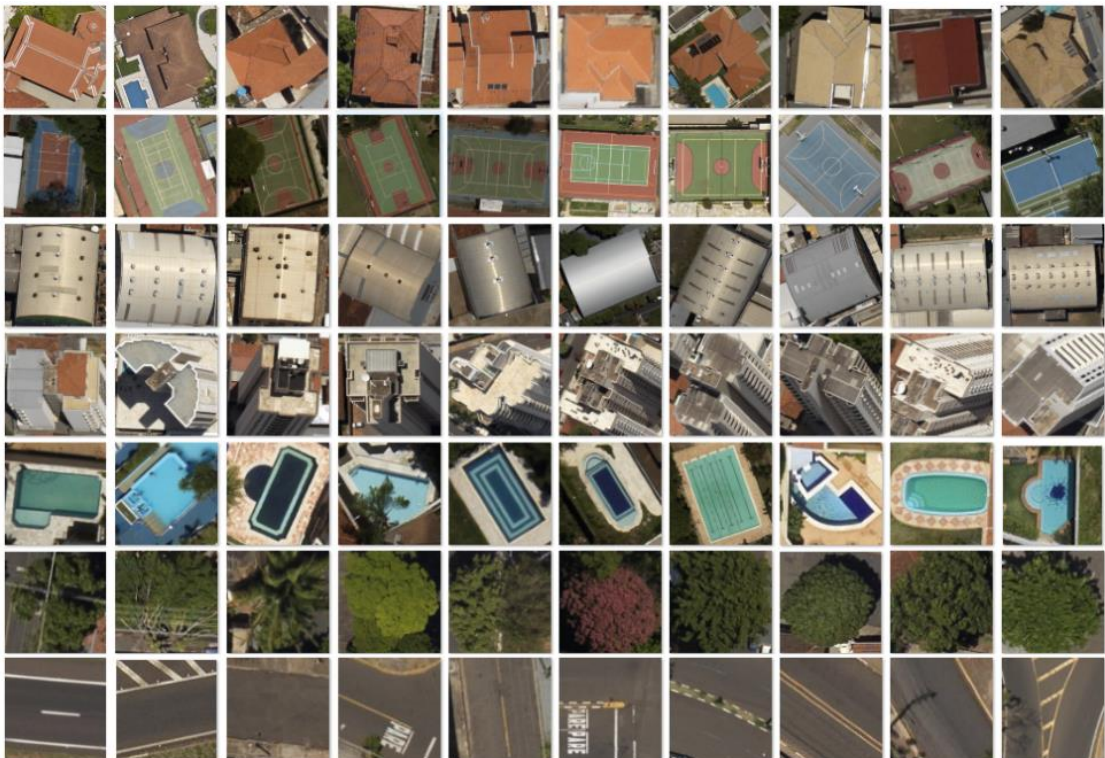
No âmbito de processamento e análise de imagens existem muitas maneiras distintas para a extração de informação. Os desenvolvimentos nas áreas de visão computacional e aprendizado de máquina culminaram em novas metodologias para realizar essas tarefas. De acordo com Carrilho e Galo (2019), o problema de extração do contorno de objetos pode ser dividido em tarefas mais específicas: (i) segmentação dos objetos³⁹; (ii) detecção e localização dos objetos; e (iii) classificação dos objetos. As RNC se mostraram muito eficazes na tarefa de classificação de imagens, e muitos trabalhos foram desenvolvidos nos últimos anos com o intuito de expandir a capacidade de informação extraída a partir dessa metodologia. Uma das maneiras é utilizar de técnicas de segmentação aliadas às RNC, como é o caso da R-CNN e suas variantes, Fast R-CNN e Faster R-CNN, por exemplo. No entanto, essas abordagens empregam uma segunda RNC para a segmentação, isto é, gerar as regiões de interesse. Em Carrilho e Galo (2019) optou-se por utilizar o *Simple Linear Iterative Clustering* (SLIC) para extrair superpixels, utilizando a RNC apenas para classificá-los.

Para a realização dos experimentos gerou-se um conjunto de dados de treinamento (Figura 40) com aproximadamente 100 a 200 amostras de 64 por 64 pixels para 7 classes extraídas do *Unesp Photogrammetric Data Set*. As classes selecionadas, foram: edificação, quadra esportiva, galpão, prédio, piscina, vegetação, e via pública. A seleção da topologia da

³⁹ Ressalta-se que é uma tarefa mais complexa do que a segmentação da imagem tradicional, utilizando crescimento de regiões, por exemplo.

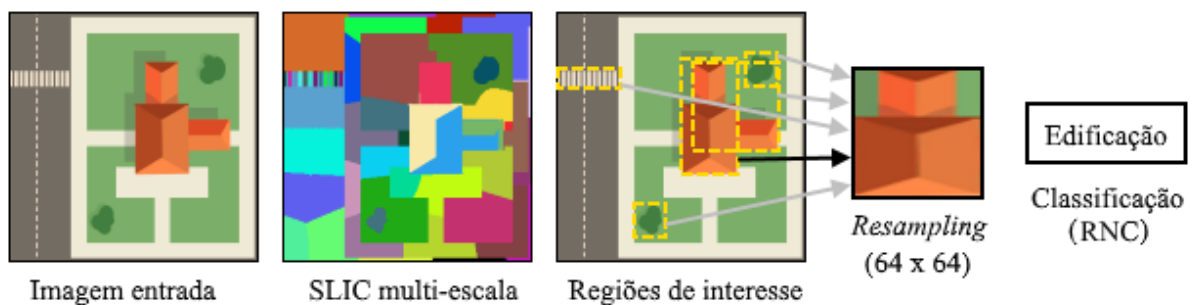
rede, dos parâmetros de treinamento e do algoritmo SLIC não são triviais. Os detalhes da implementação (como o emprego da técnica de *data augmentation*) e os parâmetros utilizados podem ser vistos em Carrilho e Galo (2019). Nas Figuras 41 e 42 são apresentadas a abordagem adotada para a extração de objetos e a arquitetura da RNC adotada, respectivamente.

Figura 40 – Conjunto de dados para treinamento de RNC contendo dez amostras para cada uma das sete classes de objetos presentes em cenas urbanas.



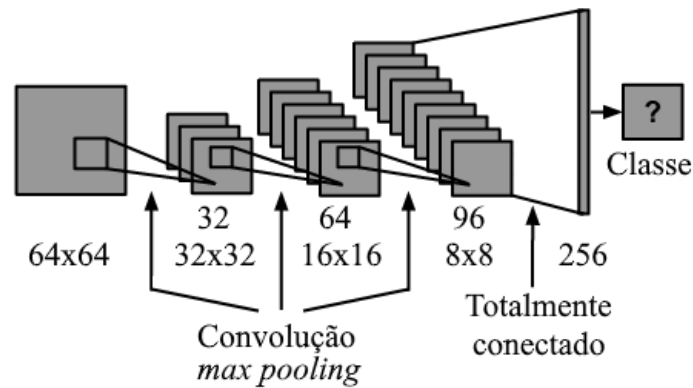
Fonte: Adaptado de Carrilho e Galo (2019).

Figura 41 – Abordagem adotada para extração de objetos utilizando RNC.



Fonte: Adaptado de Carrilho e Galo (2019).

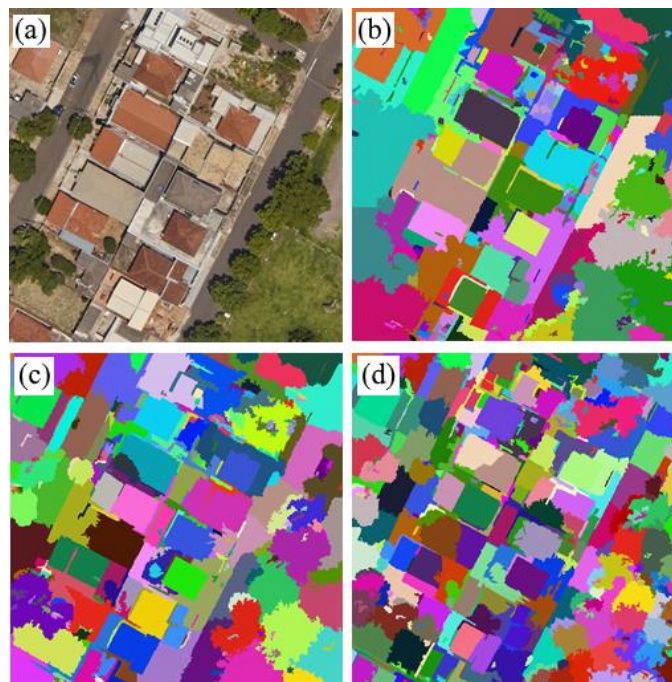
Figura 42 – Ilustração da arquitetura da RNC adotada.



Fonte: Adaptado de Carrilho e Galo (2019).

Como pode-se ver na Seção 4.8, uma variável que deve ser considerada no algoritmo SLIE se refere à quantidade (k) de pontos sementes. Quanto maior o valor de k , mais superpixels são gerados. Neste sentido, é importante definir um valor de k que resulte em superpixels pequenos, sendo a situação ideal aquela em que cada objeto de interesse seja segmentado em um superpixel. Na Figura 43 é mostrado o resultado do uso de diferentes parâmetros (k) no algoritmo SLIC visando extrair os superpixels.

Figura 43 – Abordagem SLIC multi-escala. (a) Imagem original, superpixels gerados com (b) $k = 50$, (c) $k = 80$ e (d) $k = 150$.



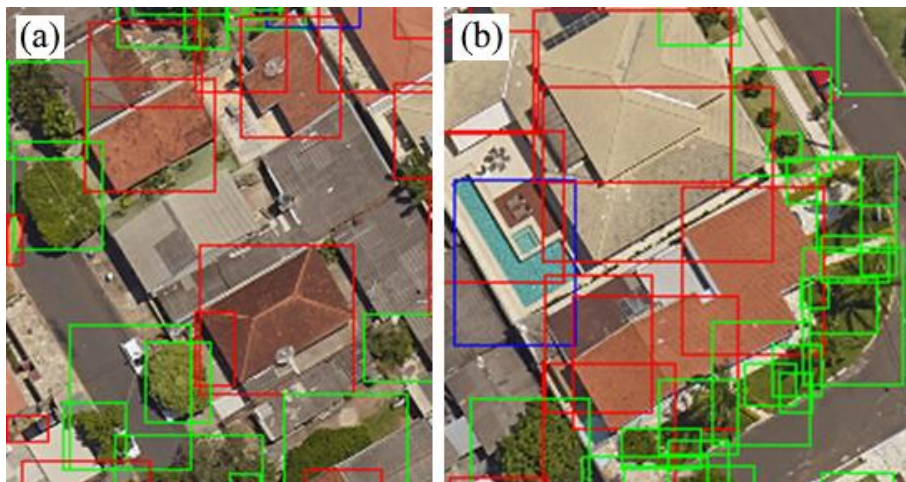
Fonte: Adaptado de Carrilho e Galo (2019).

Entretanto, é complicado obter uma solução que comporte objetos de diferentes tamanhos. Portanto, o parâmetro m (na Equação 40) deve ser escolhido de forma que dê um peso maior para a distância no espaço de cores CIELAB, a fim de mitigar esse problema. Neste trabalho foi utilizado $k = 80$ e $m = 10$. Dessa forma, são geradas regiões de interesse de tamanho adequado, correspondentes às amostras de treinamento da RNC. Para a extração de telhados de edificações com imagens aéreas de alta resolução é interessante selecionar os valores que proporcionem a geração de superpixels no intervalo de $30 - 500 \text{ m}^2$, o que deve acomodar boa parte das edificações. É possível selecionar os parâmetros quando o GSD das imagens é conhecido, como normalmente é o caso. Chen *et al.* (2019) mencionam que essa abordagem proporciona bons resultados.

Cada uma das regiões de interesse (superpixel) resulta em uma subimagem delimitada pelo retângulo envolvente, que são reamostrados para a dimensão 64×64 pixels e classificados com a RNC. Na Figura 44 são apresentadas as delimitações dos objetos edificações, vegetação e piscinas, extraídos após este procedimento. Ou seja, a RNC associa um rótulo a cada região de interesse (superpixel). Alguns alvos, como a vegetação por exemplo, podem conter diversas regiões de interesse em diferentes dimensões (Figura 44b). O ideal é que o SLIC gere regiões de interesse cujo formato resulte em retângulos envoltentes com a altura similar da largura, isto é, o mais quadrado possível, a fim de minimizar as distorções durante a reamostragem da subimagem para 64×64 pixels.

Figura 44 – Detecção de objetos em duas regiões urbanas.

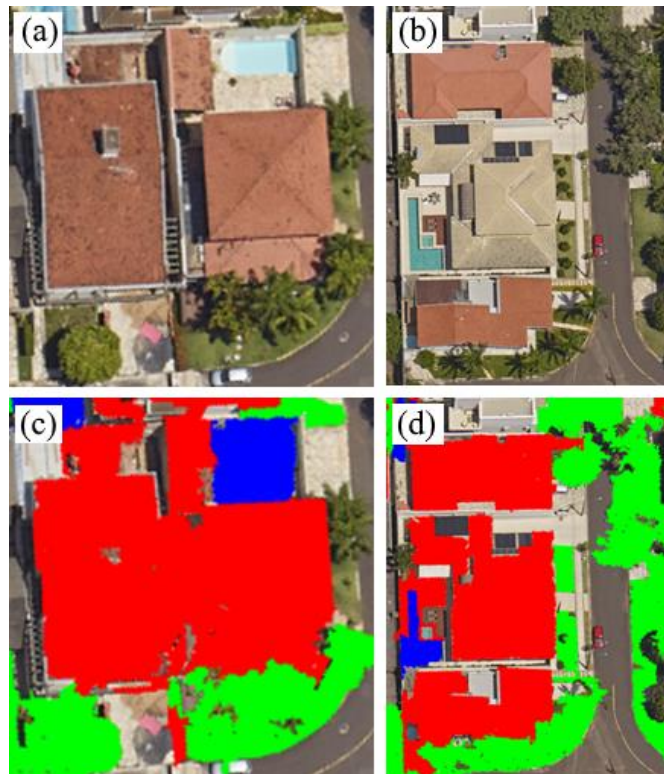
Retângulos vermelhos: edificações, verdes: vegetação, azul: piscinas.



Fonte: Adaptado de Carrilho e Galo (2019).

Por fim, os superpixels rotulados são considerados como os objetos extraídos, assim como ilustra a Figura 45. Nesse caso, como um telhado pode ter sido segmentado em mais de um superpixel, realizou-se um agrupamento por proximidade, isto é, todos os superpixels de mesmo rótulo que estejam conectados são agrupados em um único objeto. O contorno dos objetos vai depender então da qualidade das imagens (GSD, principalmente) e dos parâmetros k e m utilizados.

Figura 45 – Extração de objetos. Imagens aéreas em (a) e (b), resultados em (c) e (d), respectivamente. Legenda: Polígonos vermelhos: edificações, verdes: vegetação, azul: piscinas.



Fonte: Adaptado de Carrilho e Galo (2019).

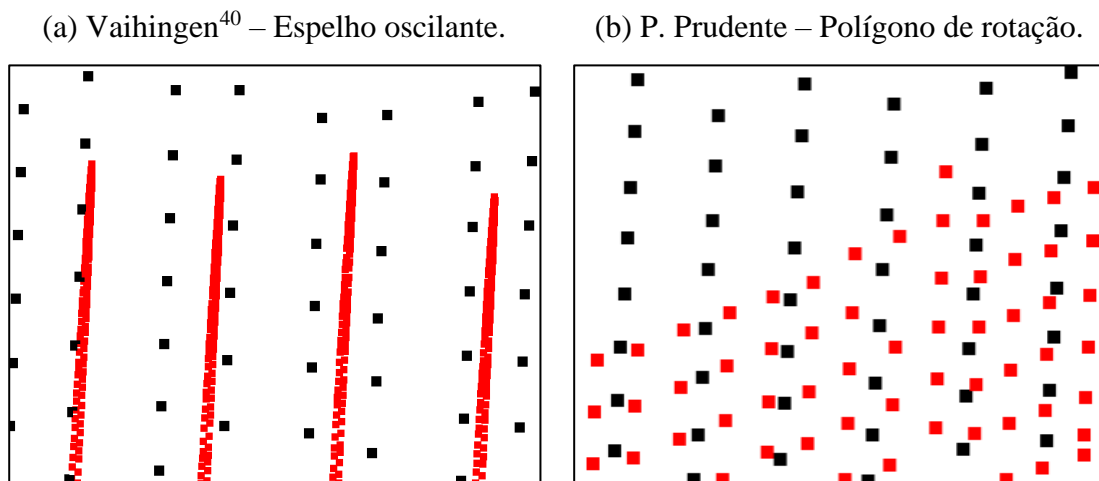
Em Carrilho e Galo (2019) é possível verificar que a abordagem proposta é promissora e resultou em resultados satisfatórios (89% de acurácia, considerando 80% dos dados para treinamento e 20% para validação). O trabalho também utilizou a mesma abordagem para classificar imagens do satélite QuickBird e imagens extraídas do Google Earth, obtendo uma acurácia de 95,2%, o que demonstra sua robustez, uma vez que a rede não foi treinada para lidar com esses dados originalmente.

5.2.3 Extração do contorno dos telhados de edificações com o algoritmo α -shape

A maior dificuldade na utilização do algoritmo α -shape consiste justamente na seleção do parâmetro α . O resultado do contorno obtido é sensível a este parâmetro, sendo necessário utilizar um valor ótimo para extrair o contorno desejado. Com relação aos dados SVLA aerotransportados, observa-se uma correlação entre a densidade média e o valor do parâmetro α , assim como mostra Santos *et al.* (2019). Durante o desenvolvimento deste trabalho, o grupo de pesquisa se empenhou em desenvolver soluções adaptativas e robustas à variação de densidade dos dados.

Os Sistemas de Varredura a LASER Aerotransportado podem gerar dados muitas vezes problemáticos para o algoritmo α -shape. Os motivos principais são referentes a variações de densidades na nuvem de pontos resultante. Essas variações de densidade são decorrentes de diversos fatores, tais como áreas de sobreposição entre faixas (na qual a densidade de pontos pode duplicar), alterações na velocidade e na orientação da plataforma durante a aquisição dos dados, e até mesmo o próprio mecanismo do sistema de varredura (Figura 46).

Figura 46 – Resultado da amostragem em uma região de sobreposição de faixas. Comparação entre dois mecanismos de varredura diferentes. (a) Espelho oscilante, e (b) Polígono de rotação.



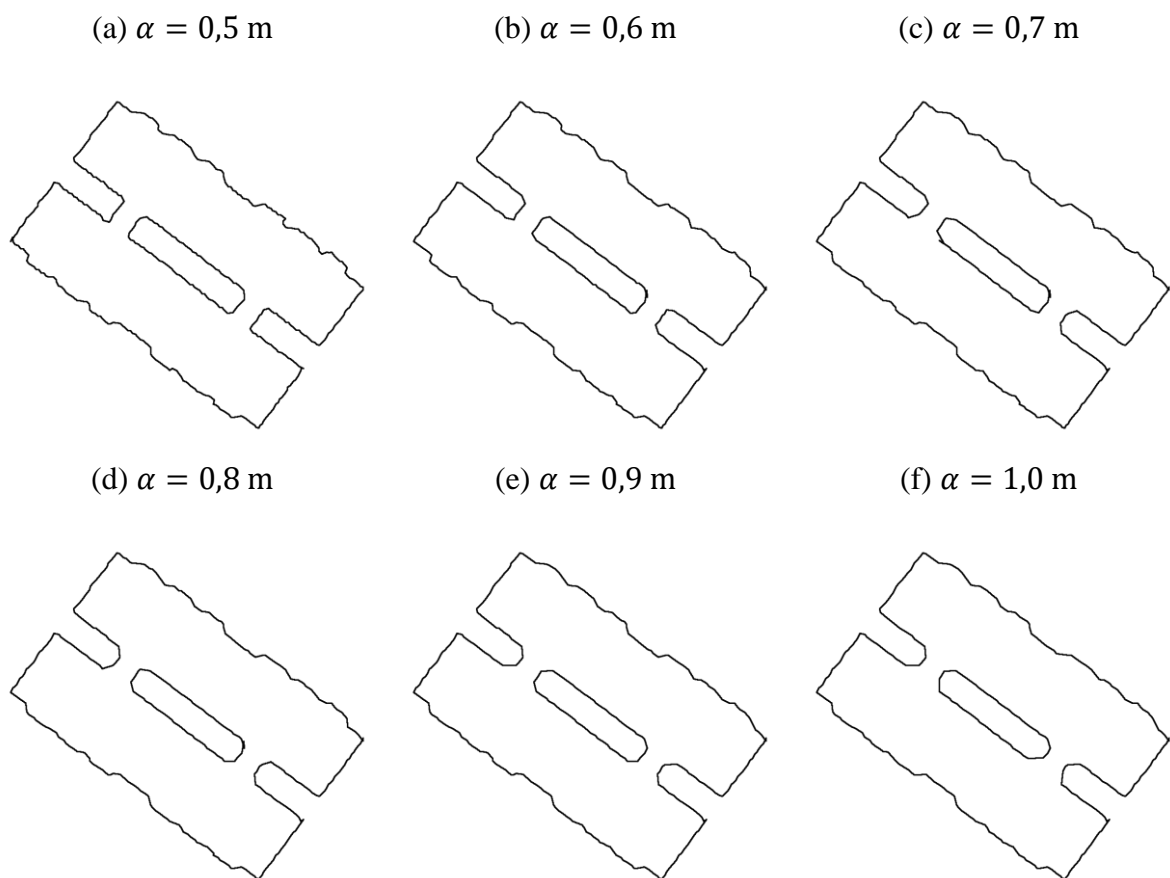
Legenda: (●) Faixa 1 (●) Faixa 2.

Fonte: Elaborada pelo autor.

⁴⁰ Disponível em: <https://www2.isprs.org/commissions/comm2/wg4/benchmark/2d-sem-label-vaihingen/>

Alguns modelos de SVLA utilizam espelhos oscilantes para direcionar o pulso LASER, cujo mecanismo requer desaceleração, parada e aceleração no sentido oposto, fazendo com que a borda das faixas tenha uma maior densidade de pontos do que a parte central. De acordo com Santos *et al.* (2019), existem diversas estratégias para selecionar o valor do parâmetro α , tais como: utilizar um valor predefinido empiricamente; determinar o espaçamento médio dos pontos para a nuvem de pontos; determinar o espaçamento médio dos pontos para cada objeto (como ilustrado na Figura 47); e determinar o espaçamento médio dos pontos localmente, utilizando uma certa vizinhança. Santos *et al.* (2020) apresentam um método adaptativo baseado na seleção do parâmetro α a partir da vizinhança do ponto. Esta abordagem se mostrou robusta para variações severas de densidade, principalmente em situações de telhados obstruídos parcialmente por vegetação, onde apenas alguns pulsos penetram a copa da árvore.

Figura 47 – Contornos extraídos considerando diferentes valores para o parâmetro α .



Fonte: Elaborada pelo autor.

Assim como pode ser observado na Figura 47, o valor do parâmetro α é de suma importância na extração do contorno do objeto. Nota-se que a partir do valor $\alpha = 0,6$ m os vértices do polígono interno passam a ficar mais suavizados (chanfrados), o que não condiz com a realidade. Entretanto, ao reduzir demais o valor do parâmetro, o contorno passa a conter ruídos (aparência de serrilhamento).

5.2.4 Integração da informação radiométrica aos dados SVLA

Embora a integração da informação radiométrica aos dados SVLA forneça mais detalhes sobre a cena, é necessário tomar certos cuidados para que certos problemas não prejudiquem a etapa de classificação. A diferença na geometria de aquisição dos sensores (câmara e SVLA) utilizados na coleta dos dados acarreta em imperfeições no registro, as quais são causadas principalmente por oclusões. Vale ressaltar também, que a inserção das informações radiométricas requer todo o rigor envolvido em sua interpretação, tal como as correções de efeitos atmosféricos, e remoção de sombras, por exemplo.

Considerando a abordagem apresentada na Seção 4.2, foram realizados alguns experimentos para avaliar a qualidade da associação da informação radiométrica aos dados SVLA. O principal problema é a oclusão causada pela geometria da aquisição da imagem com relação à superfície. Entretanto, utilizando diversas imagens do bloco que recobrem a mesma região é possível minimizar esse efeito, uma vez que pontos oclusos em uma imagem podem ter sua informação radiométrica associada a partir de outra(s) imagens.

Uma condição surge a partir dessa solução: cada ponto (XYZ) do terreno terá um ou mais pixel(s) associados no espaço imagem, isto é, teremos um mapeamento $1 \rightarrow n$, onde n é o número de imagens em que o ponto não está ocluso. Considerando que partes das cenas não são estáticas, um segundo problema surge: qual dos n pixels associar à coordenada do ponto? Existem diversas possíveis soluções, como: adotar o mais claro, o mais escuro, a média, entre outros. Intuitivamente, a média parece ser uma boa solução, entretanto, como pode ser visto na Figura 48, alguns artefatos (destacados em vermelho) podem surgir caso objetos da cena tenham movido (como veículos, por exemplo) durante a aquisição das n imagens. Tais artefatos podem causar problemas durante a classificação, uma vez que os valores RGB tem um peso grande como atributos.

Figura 48 – Inspeção visual dos resultados da integração radiométrica aos dados SVLA e a utilização do critério da moda dos valores RGB no mapeamento $1 \rightarrow n$.



Fonte: Elaborada pelo autor.

Para mitigar esse problema, um critério pode ser estabelecido: adotar a moda dos valores RGB entre os n possíveis. Dessa forma, há uma probabilidade maior de que o valor RGB associado ao ponto seja consistente aos dados, uma vez que corresponde a um intervalo de tempo maior (apareceu em mais imagens). A Figura 48 apresenta os resultados da integração da informação radiométrica para a Região 3 utilizando 6 imagens aéreas em duas situações: um dos resultados sem considerar o critério da moda dos valores RGB e o outro considerando este critério. É possível observar que o uso deste critério removeu algumas ocorrências de objetos

que apareciam em menos de 3 imagens, o que causava um efeito de “*blur*” (os veículos aparentam estar esmaecidos com relação aos demais, que apareceram nas 6 imagens).

5.2.5 Classificação das nuvens de pontos

Para a classificação da nuvem de pontos, quatro etapas implementadas em códigos independentes foram consideradas. Na primeira etapa é feita a leitura da nuvem de pontos (matriz $n \times 7$ ou $n \times 8$ se forem dados de treinamento, sendo n o número de pontos). As 7 colunas correspondem a: 3 coordenadas (X,Y,Z), intensidade do pulso LASER, e os dados radiométricos (R,G,B). Já quando se trata de dados de treinamento há também o rótulo do ponto, totalizando 8 colunas. O algoritmo implementado extrai os atributos de cada ponto da nuvem e salva em um arquivo de texto, resultando em uma matriz $n \times m$, onde m é o número de atributos calculados. Na segunda etapa é feita a leitura dos atributos e o treinamento do modelo matemático de classificação, como apresentado na Seção 3.4.2. Na terceira etapa é feita leitura do arquivo de atributos e os parâmetros do classificador treinado anteriormente visando a aplicação em uma nova nuvem de pontos. Na última etapa é feita a comparação do resultado da classificação com os rótulos de referência, sendo geradas as métricas quantitativas.

Considerando duas nuvens de pontos NP1 e NP2, sendo que NP1 contém os dados de treinamento (rótulos) e NP2 a nuvem a ser classificada, o seguinte procedimento é empregado:

- Etapa 1 – Lê a nuvem de pontos e gera a matriz de atributos para NP1 e NP2;
- Etapa 2 – Lê a matriz de atributos de NP1 e treina o classificador (C);
- Etapa 3 – Lê a matriz de atributos de NP2 e os parâmetros do classificador (C) e gera NP3, contendo o resultado da classificação;
- Etapa 4 – Lê o resultado da classificação e compara com os rótulos de referência, obtendo a acurácia da classificação.

Para a realização da classificação apenas as Etapas 1 até 3 são suficientes. No entanto, em função da avaliação da classificação proposta a Etapa 4 deve ser executada. A sequência apresentada acima serve apenas para a escolha do classificador, seleção dos hiperparâmetros e validação. Isto se dá pois, em um ambiente real é pouco provável que existam rótulos de referência para a quantificação dos resultados.

Considerando as três regiões selecionadas para os experimentos, todas elas foram divididas ao meio, isto é, foram usadas 50% para treinamento e 50% para validação. Assim, como pode ser visto na Tabela 2, os atributos utilizados para a classificação foram: bandas R,

G e B, intensidade do pulso LASER, omnivariância, anisotropia, curvatura, dispersão, linearidade, planaridade, entropia dos autovalores, desvio padrão da altitude, e a altura do ponto.

Tabela 2 – Análise de Componentes Principais calculada para a matriz dos atributos.

| Atributo | ACP 1 | ACP 2 | ACP 3 | ACP 4 |
|------------------------|----------------|----------------|----------------|----------------|
| Banda R | 38,75% | -- | -- | -- |
| Banda G | 27,57% | -- | -- | -- |
| Banda B | 33,67% | -- | -- | -- |
| Intensidade do pulso | 0,00% | -- | -- | -- |
| Altura do ponto | 0,01% | 99,99% | -- | -- |
| Desvio padrão altitude | 0,00% | 0,01% | 93,43% | -- |
| Omnivariância | 0,00% | 0,00% | 0,91% | 4,89% |
| Anisotropia | 0,00% | 0,00% | 0,20% | 1,29% |
| Curvatura | 0,00% | 0,00% | 0,11% | 0,57% |
| Dispersão | 0,00% | 0,00% | 0,20% | 1,29% |
| Linearidade | 0,00% | 0,00% | 0,71% | 21,82% |
| Planaridade | 0,00% | 0,00% | 0,16% | 12,48% |
| Entropia de autovalor | 0,00% | 0,00% | 4,30% | 57,65% |
| Total | 100,00% | 100,00% | 100,00% | 100,00% |

Como pode ser visto na Tabela 2, a classificação se beneficia consideravelmente quando a informação radiométrica é associada às coordenadas dos pontos. Este resultado ressalta a importância de realizar a integração da informação radiométrica da forma mais rigorosa possível, uma vez que a etapa de classificação passa a depender completamente nesses dados.

Ao desconsiderar totalmente a informação radiométrica, o atributo mais importante passa a ser a altura do ponto, e por sua vez quando este atributo também é removido, a classificação pode ser resumida com os atributos desvio padrão das altitudes e entropia dos autovalores. Quando apenas os atributos derivados dos autovalores são considerados, a entropia, linearidade e planaridade são os atributos que possuem maior influência na solução.

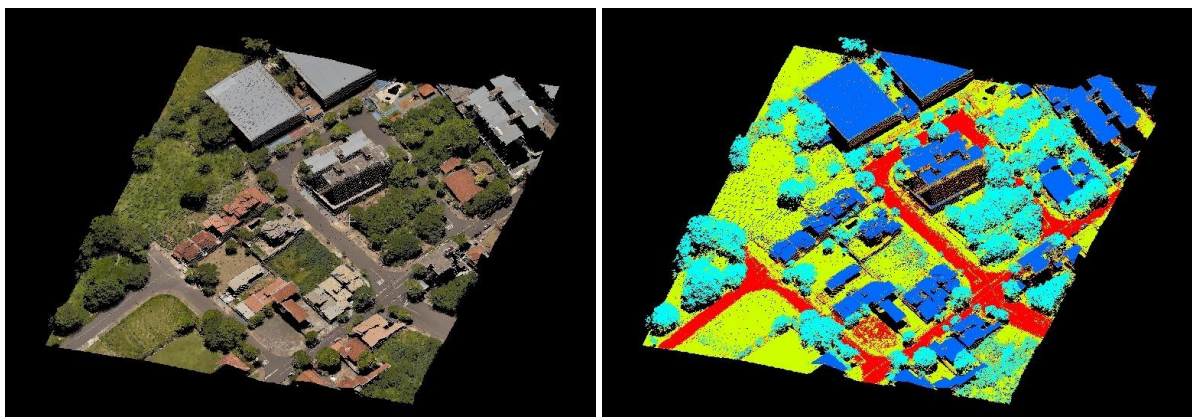
A partir dessa análise, é possível adaptar a implementação e desconsiderar completamente os atributos omnivariância, anisotropia, curvatura e dispersão, por não contribuírem de modo significativo na classificação, uma vez que possuem alta correlação com os demais atributos, e por sua vez, são redundantes. Desta maneira, é possível reduzir o tempo de processamento e minimizar os parâmetros do modelo de classificação uma vez que passa a lidar com menos atributos.

Partindo dessa análise⁴¹, é possível assumir os principais atributos para a classificação como: RGB, altura do ponto, desvio padrão das altitudes, entropia dos autovalores, linearidade e planaridade. As três regiões selecionadas foram classificadas utilizando os códigos implementados e considerando os atributos selecionado, sendo os resultados para a Região 3 apresentados na Figura 49 e Tabela 3. O classificador selecionado foi a Rede Neural Artificial (*Multi-Layer Perceptron - MLP*) com duas camadas ocultas de 10 neurônios em cada.

Figura 49 – Resultado da classificação da metade superior⁴² da Região 3.

(a) Nuvem de pontos com RGB associado.

(b) Resultado da classificação.



Legenda: (●) Vegetação (●) Edificações (●) Linha de transmissão
(●) Solo não-pavimentado (●) Solo pavimentado (●) Não classificado

A classe veículos foi desconsiderada pois os atributos e classificador selecionados não foram capazes de identificar os pontos efetivamente. Além disso, essa classe não contribui para as etapas posteriores deste trabalho, cujo objetivo restringe-se às edificações.

Tabela 3 – Análise quantitativa da classificação utilizando MLP.

| Classe | Pontos | Acurácia |
|----------------------|---------|----------|
| Não classificado | 16.454 | 89,07% |
| Solo não pavimentado | 163.689 | 98,24% |
| Solo pavimentado | 32.022 | 84,74% |
| Edificações | 81.239 | 98,21% |
| Vegetação | 56.398 | 95,96% |
| Linha de transmissão | 2.740 | 92,37% |
| Total | 352.542 | 93,10% |

⁴¹ De forma simplória, essa análise pode ser considerada como engenharia de atributos.

⁴² Corresponde aos 50% dos dados utilizados para validação da classificação.

Como pode ser visto na Figura 49, apesar da alta acurácia numérica (93,1%), a classificação falha em algumas situações. A mais notável (visivelmente) é a confusão feita entre “solo pavimentado” e “não-pavimentado” quando o “solo pavimentado” não está em boas condições (pequenos arbustos e poeira/terra acumulada). As demais classes tiveram um resultado satisfatório na classificação.

As duas classes com melhores resultados foram “solo não pavimentado” (98,24%) e “edificações” (98,21%), seguido pela classe “vegetação” (95,96%). Observando a Figura 49 é notável a baixa acurácia da classe “solo pavimentado”, visto que ocorreu confusão no estacionamento próximo da parte inferior-central da imagem 5.16b. Apesar desses erros também afetarem a classe “solo não pavimentado”, é necessário lembrar que essa é a classe com maior número de pontos e, em proporção, este erro é menor quando comparado a classe “solo pavimentado”.

5.2.6 Extração e regularização dos contornos utilizando a assinatura

Após a classificação das nuvens de pontos é possível separar os telhados individualmente, aplicando uma segmentação na classe edificações. Para cada telhado é aplicado o mesmo procedimento: geração da malha triangular utilizando o critério de Delaunay; extração do contorno aproximado com o algoritmo α -shape; cálculo do centro de massa e geração da assinatura do contorno; aplicação do RANSAC com o modelo matemático da equação geral da cônica para cada segmento do telhado; e reconstrução da borda do contorno utilizando a cônica ajustada.

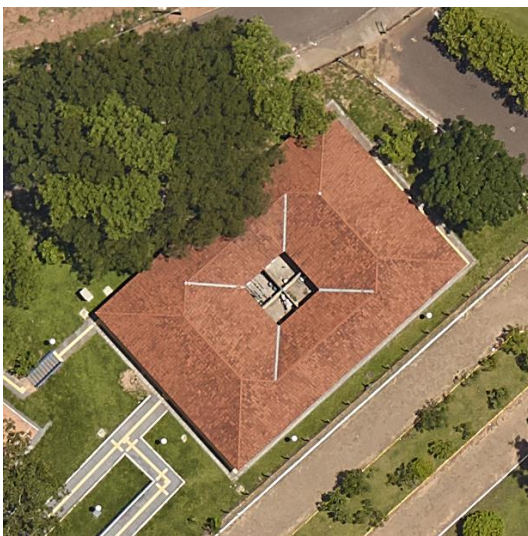
Quatro edificações com telhados parcialmente obstruídos por vegetação foram selecionadas para avaliar o método proposto de reconstrução a partir do descritor geométrico baseado em assinatura. As edificações (Figura 50) apresentam obstruções não severas, isto é, menos que 50% da dimensão de um segmento lateral, sem que a oclusão seja em alguma das quinas. Esses casos selecionados são as situações ideais esperadas para o método, assim como previsto anteriormente.

Nas Figuras 51 e 52 são apresentados os resultados obtidos para o Telhado 1 e Telhado 2, respectivamente. Para os dois telhados o método proposto foi capaz de reconstituir a obstrução causada pela vegetação. O primeiro telhado apresenta uma obstrução maior que o segundo, entretanto, a oclusão no segundo telhado possui um formato peculiar devido à copa da árvore que causou a obstrução estar parcialmente preenchida com folhas.

Já nas Figuras 53 e 54 são apresentados os resultados obtidos para o Telhado 3 e Telhado 4, respectivamente. Os dois telhados possuem orientação e oclusões similares. No Telhado 3, a penetração do pulso LASER auxiliou bastante a execução do método, pois não obstruiu a quina superior esquerda. Entretanto a área total ocluída é a maior entre os quatro telhados selecionados. O método proposto foi capaz de reconstruir todos os telhados, apesar de visualmente na imagem aérea o Telhado 3 estar bem ocluído.

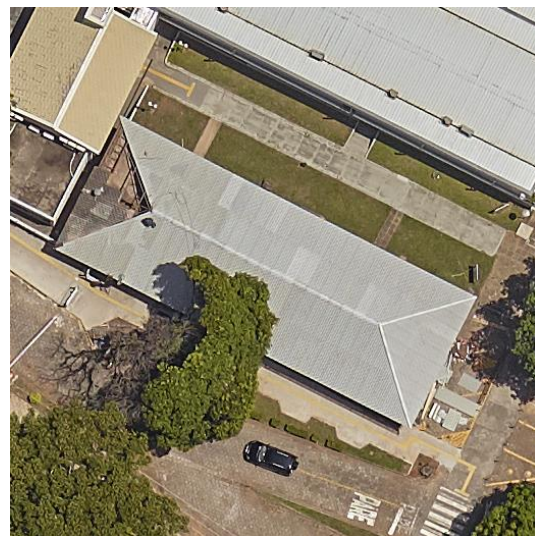
Figura 50 – Telhados selecionados com obstruções causadas por vegetação.

(a) Telhado 1 – Moradia estudantil.



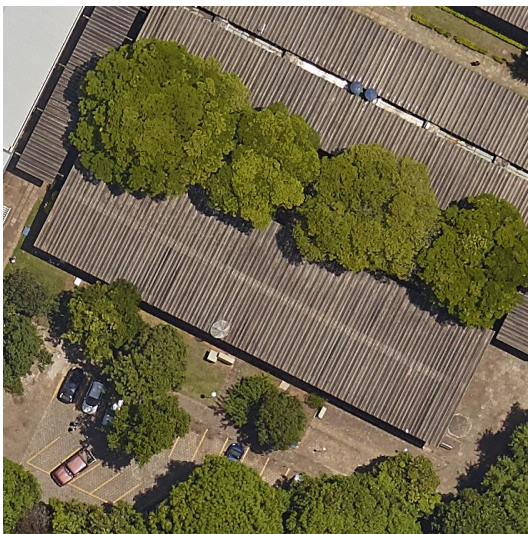
Segmento linear: 30,0 m
Oclusão: 10,0 m (33,3%)

(b) Telhado 2 – Laboratório.



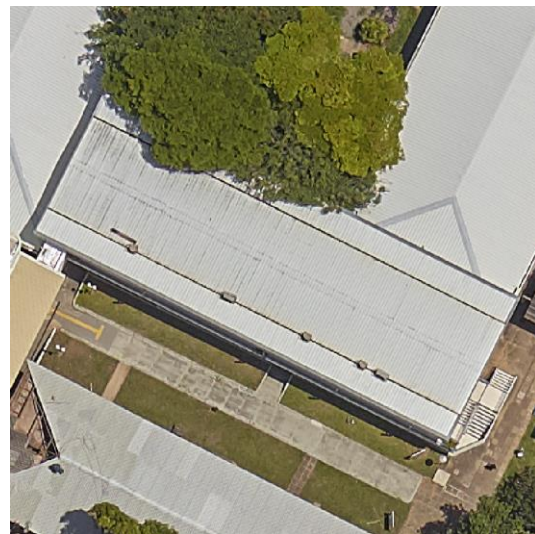
Segmento linear: 32,0 m
Oclusão: 5,5 m (17%)

(c) Telhado 3 – Laboratório.



Segmento linear: 51,5 m
Oclusões: várias ~20,0 m

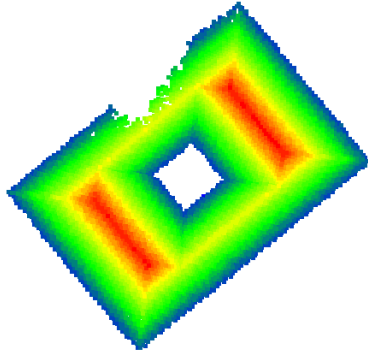
(d) Telhado 4 – Laboratório.



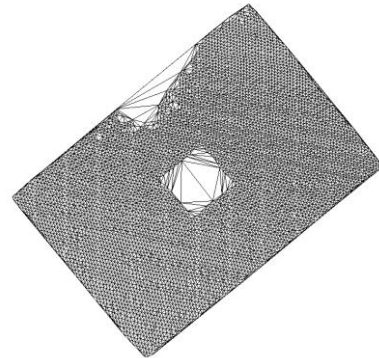
Segmento linear: 36,0 m
Oclusão: 13,0 m (36%)

Figura 51 – Resultados para o Telhado 1 (Região 1 - moradia estudantil)

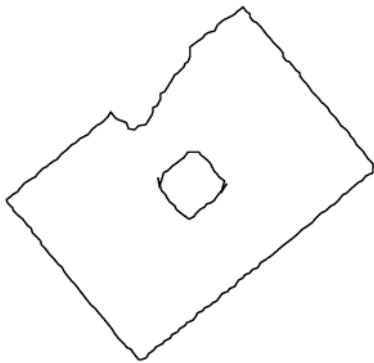
(a) Nuvem de pontos do telhado.



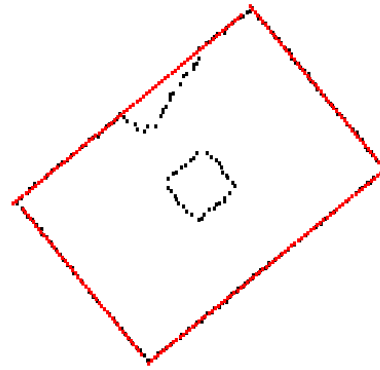
(b) Malha triangular (Delaunay).



(c) Contorno α -shape ($\alpha = 0,60$ m).

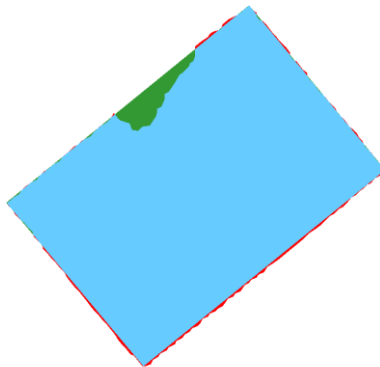


(d) Contorno reconstruído pelo método.

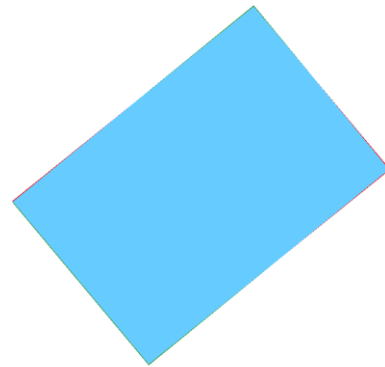


Legenda: Contorno α -shape (●), contorno reconstruído a partir da assinatura (●).

(e) Erro cometido com o α -shape.



(f) Erro cometido com o método proposto.



Legenda (e) e (f): VP (●), VN (), FN (●) e FP (●).

(g) Assinatura normalizada e com origem na quina esquerda (no sentido anti-horário).

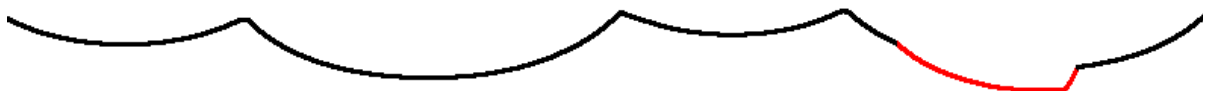
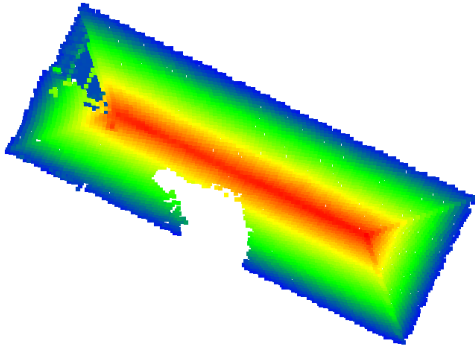
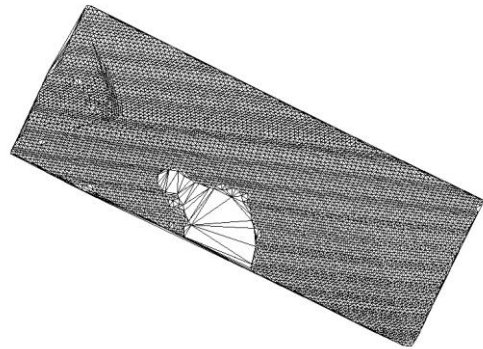


Figura 52 – Resultados para o Telhado 2 (Região 2 – laboratório)

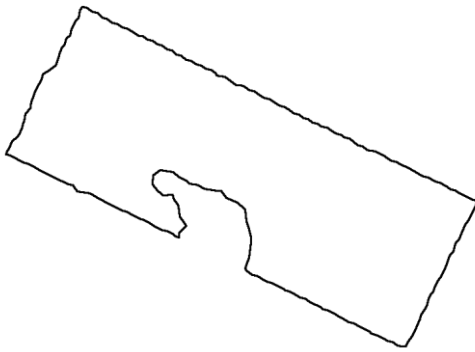
(a) Nuvem de pontos do telhado.



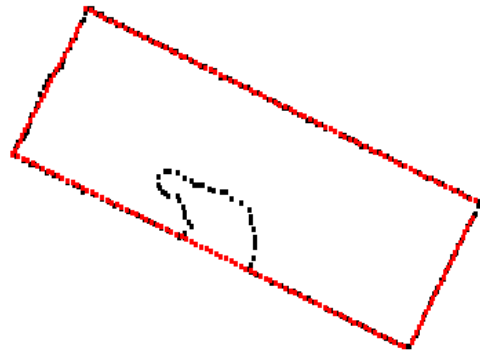
(b) Malha triangular (Delaunay).



(c) Contorno α -shape ($\alpha = 0,50$ m).

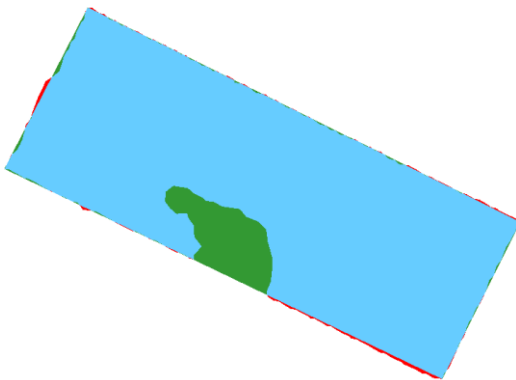


(d) Contorno reconstruído pelo método.

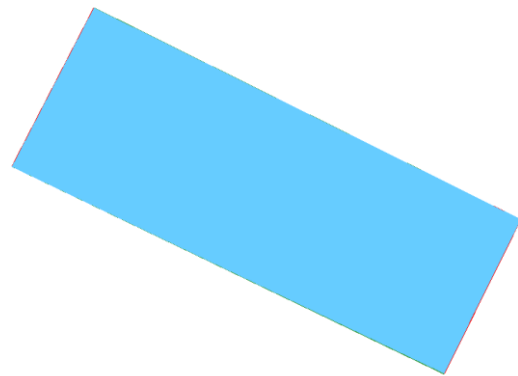


Legenda: Contorno α -shape (●), contorno reconstruído a partir da assinatura (●).

(e) Erro cometido com o α -shape.



(f) Erro cometido com o método proposto.



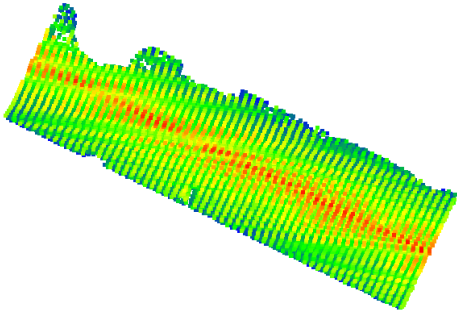
Legenda (e) e (f): VP (●), VN (), FN (●) e FP (●).

(g) Assinatura normalizada e com origem na quina superior (no sentido anti-horário).

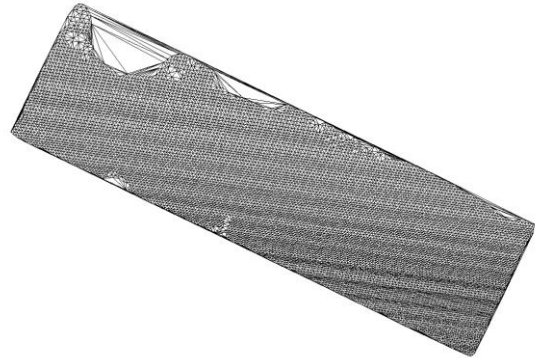


Figura 53 – Resultados para o Telhado 3 (Região 2 – laboratório)

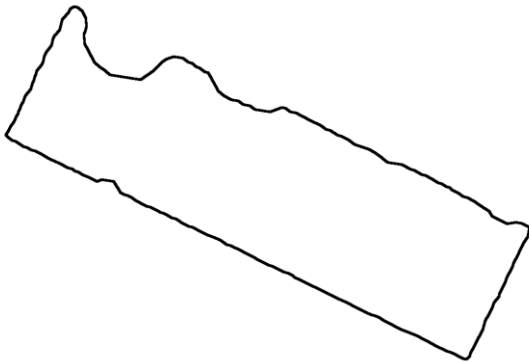
(a) Nuvem de pontos do telhado.



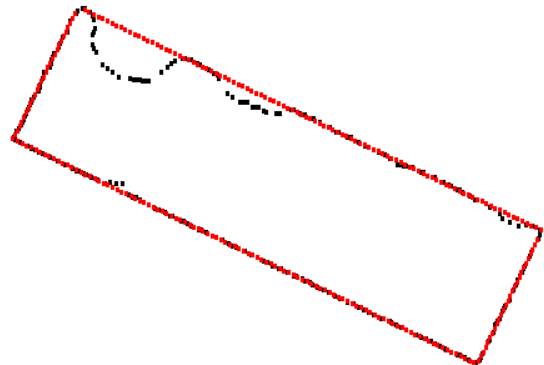
(b) Malha triangular (Delaunay).



(c) Contorno α -shape ($\alpha = 0,50$ m).

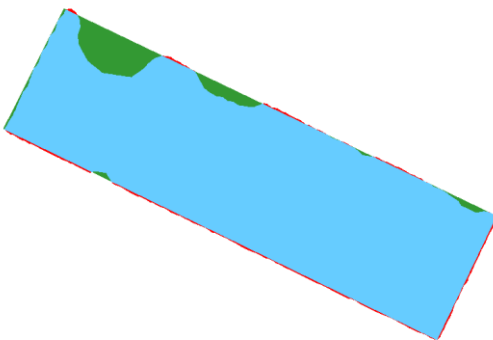


(d) Contorno reconstruído pelo método.

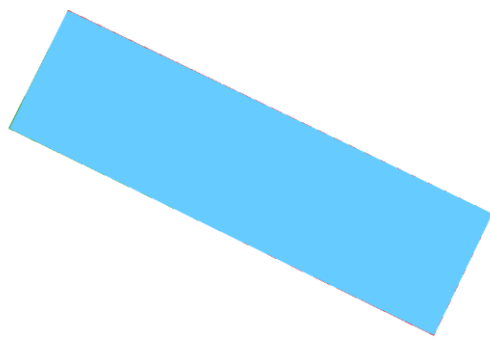


Legenda: Contorno α -shape (●), contorno reconstruído a partir da assinatura (●).

(e) Erro cometido com o α -shape.



(f) Erro cometido com o método proposto.



Legenda (e) e (f): VP (●), VN (), FN (●) e FP (●).

(g) Assinatura normalizada e com origem na quina esquerda (no sentido anti-horário).

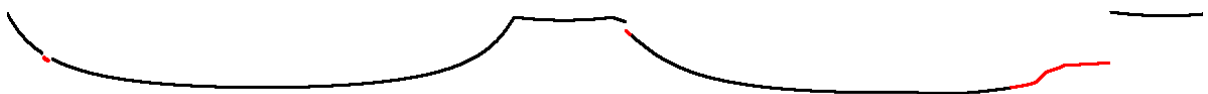
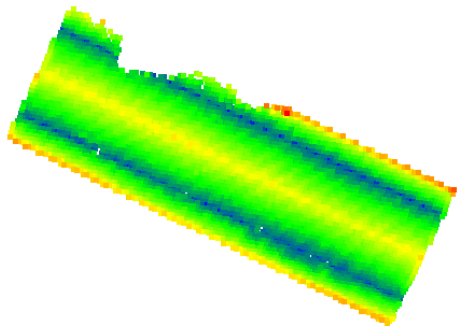
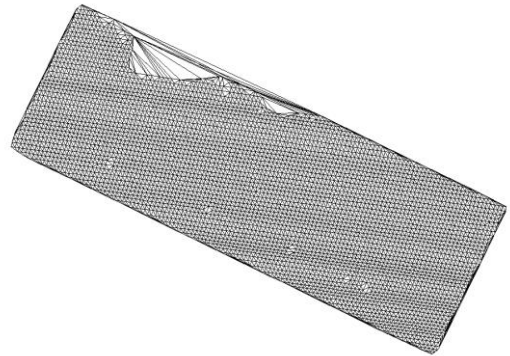


Figura 54 – Resultados para o Telhado 4 (Região 2 – laboratório)

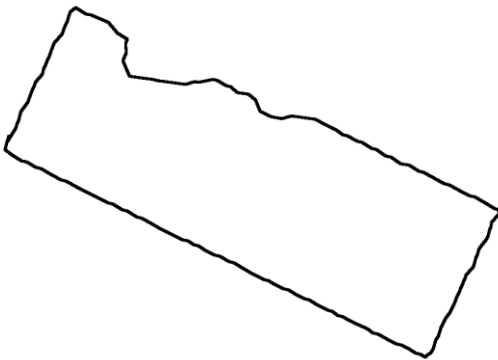
(a) Nuvem de pontos do telhado.



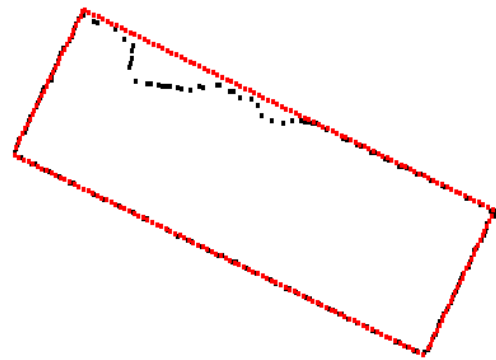
(b) Malha triangular (Delaunay).



(c) Contorno α -shape ($\alpha = 0,50$ m).

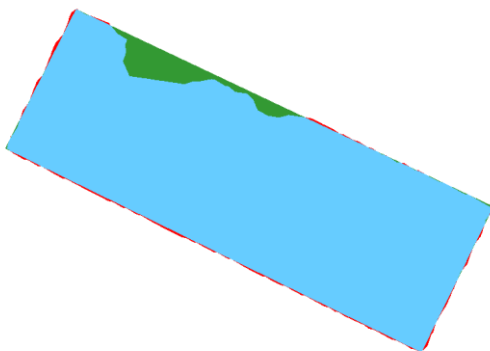


(d) Contorno reconstruído pelo método.

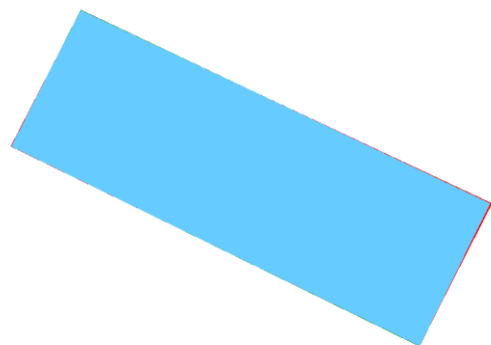


Legenda: Contorno α -shape (●), contorno reconstruído a partir da assinatura (●).

(e) Erro cometido com o α -shape.



(f) Erro cometido com o método proposto.



Legenda (e) e (f): VP (●), VN (), FN (●) e FP (●).

(g) Assinatura normalizada e com origem na quina esquerda (no sentido anti-horário).



É importante observar que apesar da Figura 50 ilustrar oclusões grandes (principalmente no Telhado 3), a real dimensão é um pouco menor devido à penetração do pulso LASER na copa das árvores. Os valores informados na Figura 50 foram medidos a partir das malhas triangulares mostradas nas Figuras 51 – 54.

Uma síntese dos resultados quantitativos da aplicação do método proposto é apresentada na Tabela 4. Observando os resultados apresentados na Tabela 4 os percentuais de *Precision*, *Recall*, F-score e da métrica PoLiS podem dar a impressão de que houve pouca mudança. Entretanto, quando a diferença em área é levada em consideração torna-se mais nítida a contribuição da reconstrução da obstrução por meio da assinatura. Considerando apenas o resultado do contorno obtido pelo algoritmo α -shape, os telhados apresentam uma grande diferença em área com relação ao polígono de referência.

Tabela 4 – Resultados da extração dos contornos de telhados.

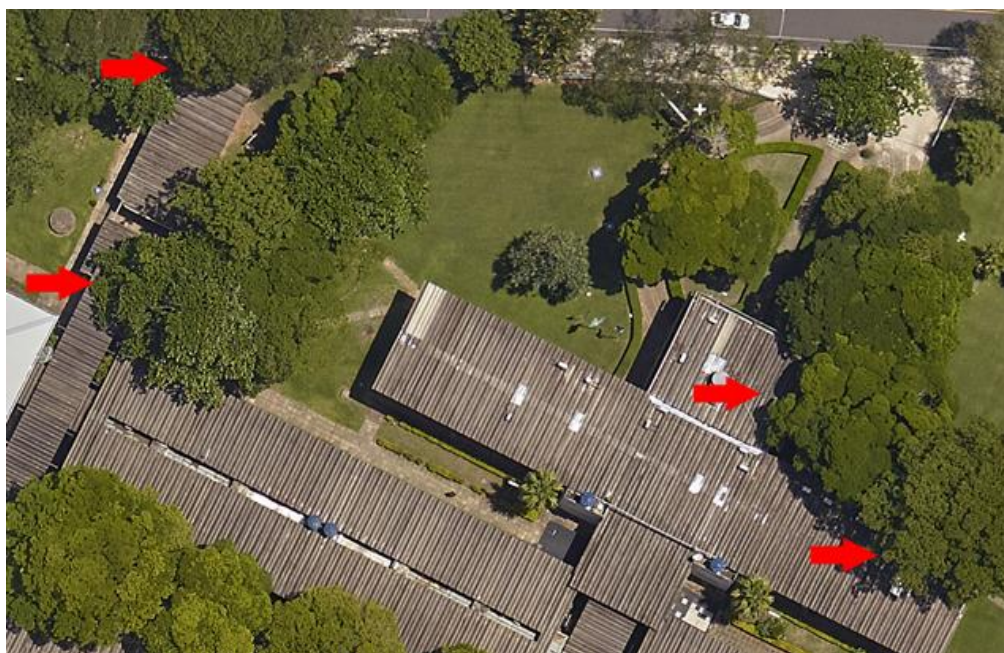
| | Telhado 1 | | Telhado 2 | | Telhado 3 | | Telhado 4 | |
|-----------------------------|-----------------|--------|-----------------|--------|-----------------|--------|-----------------|--------|
| | α -shape | Método | α -shape | Método | α -shape | Método | α -shape | Método |
| <i>Precision</i> (%) | 98,59 | 99,76 | 99,13 | 99,60 | 98,54 | 99,74 | 98,80 | 99,61 |
| <i>Recall</i> (%) | 97,12 | 99,53 | 92,72 | 99,82 | 94,13 | 99,95 | 95,59 | 99,78 |
| F-score (%) | 97,85 | 99,65 | 95,83 | 99,71 | 96,28 | 99,80 | 97,17 | 99,69 |
| Métrica PoLiS | 0,14 | 0,09 | 0,26 | 0,06 | 0,26 | 0,03 | 0,17 | 0,04 |
| Dif. área (m ²) | 8,93 | 1,41 | 24,33 | 0,83 | 33,73 | 0,87 | 14,41 | 0,75 |

Apesar da menor dimensão linear no segmento da borda do Telhado 2, a obstrução é maior do que a do Telhado 1 em termos de área, uma vez que adentra mais ao centro do telhado. Após a reconstrução do contorno com o método proposto, a diferença em área com relação ao polígono de referência é minimizada consideravelmente. Sem corrigir as oclusões, o erro em área é significativo, chegando a afetar quase 34 m² para o Telhado 3. Para os quatro telhados selecionados os valores de acurácia foram maiores que 99,5% quando o método proposto é utilizado. É importante notar que uma diferença de 2% no valor do F-score, pode representar uma área significativa, como nos casos em que apenas o α -shape é utilizado.

Apesar dessas situações de sucesso, o método proposto não será capaz de reconstruir todas as obstruções em telhados de edificações causadas por vegetação. A rigor, dependendo da magnitude da oclusão um operador humano também poderá ter dúvidas e em alguns casos uma inspeção *in loco* é necessária para ter uma garantia integral da qualidade. A Figura 55 ilustra diversos casos problemáticos onde até mesmo operadores humanos teriam dificuldade na interpretação e restituição correta dos vértices dos polígonos de algumas das bordas. Em

algumas das situações, um ou mais vértices estão ocultos, em outras, um segmento completo do telhado está sob a copa das árvores. Para essas ocorrências severas de obstrução outras soluções devem ser adotadas, como a coleta de dados terrestres por meios topográficos ou por sistemas de varredura a LASER terrestres; ou na época de voo, dando-se preferência para a coleta de dados no inverno, onde o número de folhas é menor.

Figura 55 – Situações reais onde as oclusões são de grande magnitude, dificultando a identificação de algumas das bordas das edificações.



Fonte: Elaborada pelo autor.

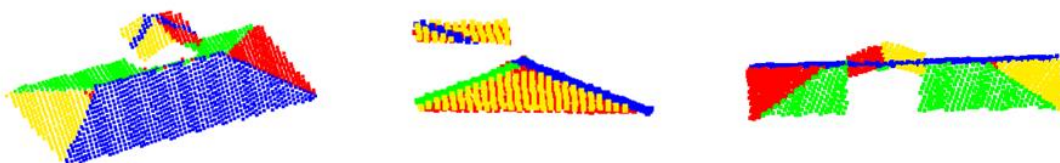
5.2.7 Segmentação dos planos de telhados de edificações

Além do contorno dos telhados das edificações, outro resultado que pode ser útil para diversos trabalhos urbanísticos e de planejamento é a segmentação dos planos. A modelagem geométrica dos telhados serve de base para estudos de incidência de luz solar por exemplo, o que auxilia no planejamento e implantação de sistemas de energia renováveis. Além disso, existem outras aplicações como na parte de Engenharia Civil, isto é, geração do *Building Information Modeling* (BIM). Em Carrilho (2016) e Carrilho e Galo (2018), algumas abordagens iniciais foram propostas utilizando o algoritmo RANSAC. Dal Poz e Ywata (2019) também abordaram o problema de maneira similar, adotando uma etapa de refinamento da solução após a segmentação com o RANSAC.

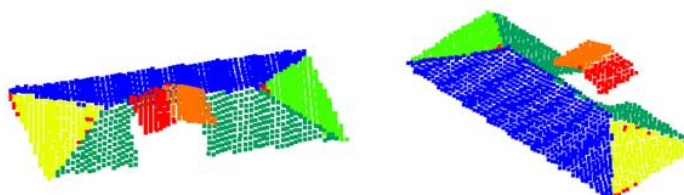
Um dos problemas com a abordagem tradicional com o RANSAC para a segmentação dos planos dos telhados é ilustrado na Figura 56a, onde o plano azul acaba seccionando outros planos do telhado, e os pontos pertencentes a estas secções são inclusos como *inliers* desse plano. Uma das possíveis soluções para este problema foi implementado na forma de um critério de conectividade. Esse critério é necessário pois, a equação geral do plano não delimita o domínio do plano, isto é, ele tem dimensões infinitas. Ao utilizar a distância do ponto ao plano como métrica para calcular os *inliers*, pode não ser suficiente, causando alguns problemas como visto na Figura 56a. Nesta figura pode-se observar que pontos pertencentes a segmentos de telhados diferentes foram agrupados como sendo de um mesmo plano (nas cores vermelha, azul e amarelo), indicando que critérios adicionais são necessários.

Figura 56 – Utilização do critério de conectividade com crescimento de regiões para mitigar problemas de segmentação de planos com o algoritmo RANSAC.

(a) Segmentação dos planos com RANSAC padrão.



(b) Segmentação dos planos com RANSAC e critério de conectividade.



Fonte: Elaborado pelo autor.

A cada iteração do RANSAC, o conjunto de pontos considerado como consenso é submetido à uma segmentação por crescimento de regiões, utilizando um limiar de distância máxima, visando agrupar pontos de acordo com o espaçamento médio da amostragem da nuvem de pontos. Desta forma, o segmento com a maior quantidade de pontos é considerado como o novo consenso, e os demais pontos a *outliers* da atual iteração do RANSAC. O resultado da aplicação desse critério é ilustrado na Figura 56b.

Capítulo 6

CONCLUSÕES E RECOMENDAÇÕES

Este trabalho apresentou uma abordagem para a extração de contornos de telhados de edificações a partir de dados SVLA aerotransportado e imagens aéreas de alta resolução. Apesar do tema ser abordado em diversos trabalhos na literatura, nem todos os métodos encontrados como estado da arte utilizam-se de procedimentos para lidar com telhados obstruídos por outros alvos presentes na cena, como a vegetação, por exemplo. Foram revisados todos os assuntos utilizados na implementação, desde a revisão do sistema de aquisição dos dados até os modelos estatísticos de aprendizagem de máquina e visão computacional.

Além disso, o presente trabalho abordou todos os problemas que compõem as etapas preliminares, como a remoção de pontos espúrios, filtragem do MDS para a geração do MDT, integração da informação radiométrica à nuvem de pontos, extração de atributos, classificação da nuvem de pontos, extração e regularização do polígono do contorno do telhado, e a segmentação dos planos. Desta forma, pode-se dizer que o método proposto compõe uma sequência de operações completa, executadas a partir dos dados iniciais adquiridos por um sistema de varredura a LASER aerotransportado (SVLA), sistema de imageamento e de posicionamento/orientação (para a obtenção da posição/orientação da plataforma), resultando no contorno dos telhados extraídos.

Diante dos problemas mencionados, propôs-se um novo método para extração de contornos de telhados de edificações, visando validar a seguinte hipótese:

“Dados SVLA associados às imagens aéreas de alta resolução adquiridos em regiões urbanas, e a utilização de algoritmos de aprendizado de máquina supervisionados juntamente com descritores geométricos (como a assinatura do contorno) podem minimizar os erros na extração automática de telhados de edificações mesmo que estejam parcialmente oclusos por vegetação.”

Múltiplas soluções foram propostas para etapas como a remoção de pontos espúrios, filtragem do modelo digital de terreno (MDT) utilizando um método adaptativo de ajuste de superfície polinomial bivariada, integração da informação radiométrica na nuvem de pontos, e a extração de contornos de telhados utilizando redes neurais convolucionais (RNC). Além disso, o principal processo do trabalho foi implementado com base na solução da equação geral da cônica para segmentos da assinatura utilizando o algoritmo RANSAC.

Foram realizados experimentos com o intuito de validar todas as etapas do método. Para isso utilizou-se do conjunto de dados de Presidente Prudente/Brasil, disponibilizados no *Unesp Photogrammetry Data Set*. Os dados de referência foram obtidos manualmente para fim de gerar as métricas quantitativas. As análises quantitativas dos experimentos realizados demonstram que:

- Considerando apenas telhados compostos por bordas de segmentos lineares, o método proposto consegue reconstruir obstruções não severas, isto é, que não sejam maiores que $1/3$ do segmento linear da borda do telhado. Além disso é importante que estas obstruções não ocorram em quinas (vértices) do polígono. Considerando estas situações, os experimentos realizados resultaram em parâmetros de *precision*, *recall* e F-score acima de 99% nos telhados estudados;
- A reconstrução das obstruções presentes em contornos de telhado de edificações a partir da assinatura geométrica pode reduzir consideravelmente o erro cometido em área, que em algumas situações chega a dezenas de metros quadrados;
- A integração da informação radiométrica na nuvem de pontos auxilia nas etapas de classificação e segmentação dos objetos quando aliada aos atributos geométricos. Entretanto, são necessários cuidados para resolver os problemas de oclusão (mapeamento duplo), similares ao da geração de ortofotos;
- A extração dos atributos geométricos em múltiplas escalas, juntamente com a utilização da informação radiométrica e a seleção correta da topologia do classificador RNA gerou resultados satisfatórios com acurácia acima de 93%;
- O nível de detalhe do contorno extraído a partir do algoritmo α -shape depende principalmente do valor do parâmetro α , sendo sensível à mudança de densidade da nuvem de pontos;
- As RNC, aliadas a algoritmos de segmentação como o SLIC, são capazes de extrair objetos a partir de imagens aéreas de alta resolução, entretanto, requerem um conjunto de treinamento de grande dimensão (com milhares de amostras); e
- Mesmo com os avanços realizados com relação aos demais métodos presentes na literatura, ainda são necessários mais estudos e desenvolvimentos para melhorar a detecção e remoção de pontos espúrios.

Assim como mencionado anteriormente, apesar das melhorias trazidas pela abordagem proposta, nem todas as situações são solucionadas. Na sequência, são destacadas algumas das limitações existentes:

- Apesar da utilização da equação geral da cônica, o método proposto pressupõe telhados com bordas compostas por segmentos lineares e cujo formato seja composto por um polígono convexo, de preferência; e
- Nem todas as etapas puderam ser automatizadas completamente, como a remoção de pontos espúrios, que ainda requer intervenção do operador.

Como recomendações para trabalhos futuros, indica-se:

- Adaptar o método para tratar de telhados que contenham segmentos curvos, mesmo que estes não sejam frequentes;
- Estudar e desenvolver melhorias no algoritmo de detecção de pontos espúrios de modo que minimize a ocorrência de falsos positivos; e
- Experimentar e realizar as devidas alterações necessárias para utilizar a metodologia proposta com nuvens de pontos fotogramétricas.

REFERÊNCIAS

- ACHANTA, R.; SHAJI, A.; SMITH, K.; LUCCHI, A.; FUA, P.; SUSSTRUNK, S. SLIC Superpixels compared to state-of-the-art superpixel methods. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, v. 34, n. 11, p. 2274-2282, 2012.
- ALIDOOST, F.; AREFI, H. Knowledge based 3D building model recognition using convolutional neural networks from LiDAR and aerial images. In: International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, v. 41, 2016, Prague, **Anais...**, ISPRS, 2016, p. 833-840.
- AMIRKOLAEI, H. A.; AREFI, H. Height estimation from single aerial images using a deep convolutional encoder-decoder network. **ISPRS Journal of Photogrammetry and Remote Sensing**, 149, 2019. p. 50-66.
- AVBELJ, J.; MÜLLER, R.; BAMLER, R. A metric for polygon comparison and building extraction evaluation. **IEEE Geoscience and Remote Sensing Letters**, v. 12, n. 1, 2015. p. 170-174.
- AXELSSON, P., 2000. DEM generation from laser scanner data using adaptive TIN models. In International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, v. 33, 2000, Amsterdam, **Anais...**, ISPRS, p. 110-117.
- AWRANGJEB, M.; FRASER, C. S.; LU, G. Building change detection from LiDAR point cloud data based on connected component analysis. In: Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences, v. 2, 2015, La Grande Motte, **Anais...**, ISPRS, 2015, p. 393-400.
- BALTSAVIAS, E. Airborne laser scanning: basic relations and formulas. **ISPRS Journal of Photogrammetry and Remote Sensing**, v. 54, n. 2-3, 1999, p. 199-214.
- BECKER, C.; HANI, N.; ROSINSKAYA, E.; D'ANGELO, E.; STRECHA, C. Classification of aerial photogrammetric 3D point clouds. In: **ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences**, v. IV-1, Hannover, Alemanha, 2017.

BISHOP, C. M. **Neural networks for pattern recognition**. Oxford: Clarendon Press, 1995. 482 p.

BOYER, V.; EL BAZ, D. Recent Advances on GPU Computing in Operations Research. In: **IEEE Conference on Parallel and Distributed Processing**. p. 1778-1787, 2013.

BUJÁN, S.; GONZÁLEZ-FERREIRO, E.; REYES-BUENO, F.; BARREIRO-FERNÁNDEZ, L.; CRECENTE, R.; MIRANDA, D. Land use classification from LiDAR data and ortho-images in a rural area. **The Photogrammetric Record**, v. 27, n. 140, p. 401–422, 2012.

CARRILHO, A. C.; IVÁNOVÁ, I.; GALO, M. Quality assessment for automatic LiDAR data classification methods. In: *XVIII SBSR - Simpósio Brasileiro de Sensoriamento Remoto*, p. 6772-6779, 28-31 de maio, Santos - SP, 2017. ISBN: 987-85-17-00088-1.

CARRILHO, A. C.; GALO, M. Extraction of building roof planes with stratified random sample consensus. **The Photogrammetric Record**. v. 33, n. 163, 2018, p. 363-380.

CARRILHO, A. C.; GALO, M. Automatic object extraction from high resolution aerial imagery with simple linear iterative clustering and convolutional neural networks. In: ISPRS Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 2019, Stuttgart, **Anais...**, ISPRS, 2019, p. 61-66.

CARRILHO, A. C.; GALO, M. Remoção de pontos espúrios em dados LiDAR aerotransportado a partir da análise estatística das altitudes. In: IV Simpósio Brasileiro de Geomática - SBG2017 e II Jornadas Lusófonas - Ciências e Tecnologias de Informação Geográfica - CTIG2017, 2017, **Anais...**, UNESP, 2017, p. 323-327.

CARRILHO, A. C.; GALO, M.; SANTOS, R. C. Statistical outlier detection method for airborne LiDAR data. In: ISPRS Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 2018, Karlsruhe, **Anais...**, ISPRS, 2018, p. 87-92.

CARRILHO, A. C.; GALO, M.; SANTOS, R. C.; PESSOA, G. G. Extração automática de edificações a partir de dados LiDAR aerotransportado aplicado ao cadastro territorial multifinalitário urbano. In: 13º Congresso Brasileiro de Cadastro Técnico Multifinalitário e Gestão Territorial, 2018, Florianópolis. **Anais do COBRAC 2018**, 2018.

CARRILHO, A. C.; SANTOS, R. C.; PESSOA, G. G.; GALO, M. Extração automática de contornos de edificações a partir de dados LiDAR aerotransportado. **Revista Brasileira de Cartografia**, v. 71, p. 832-855, 2019.

CARRILHO, A. C.; SANTOS, R. C.; PESSOA, G. G.; GALO, M. Comparison between airborne LiDAR data and digital image matching for surface reconstruction. In: ISPRS Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 2020, Santiago, **Anais...**, ISPRS, p. 509-514, 2020.

CHEN, Y.; MING, D.; Lv, X. Superpixel based land cover classification of VHR satellite image combining multi-scale CNN and scale parameter estimation. **Earth Science Informatics**, p. 1-23, 2019.

CHENG, S.; DEY, T. K.; SHEWCHUK, J. **Delaunay Mesh Generation**. Boca Raton: CRC Press, 2012. 410 p.

CONGALTON, R. G.; GREEN, K., **Assessing the accuracy of remotely sensed data: principles and practices**. Florida: CRC Press, 2009. 183 p.

DAL POZ, A. P. Synergy between LiDAR and image data in context of building extraction. In: International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, v. 40, 2014, Denver, **Anais...**, ISPRS, 2014, p. 89-93.

DAL POZ, A. P.; FERNANDES, V. J. M. Building roof boundary extraction from LiDAR and image data based on Markov random field. In: International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, v. 42, 2017, Hannover, **Anais...**, ISPRS, 2017, p. 339-344.

DAL POZ, A. P.; YWATA, M. S. Y. Adaptive random sample consensus approach for segmentation of building roof in airborne laser scanning point cloud. **International Journal of Remote Sensing**, v. 41, p. 2047-2061, 2020.

DAVIES, E. R. **Computer and machine vision: theory, algorithms, practicalities**. San Diego: Academic Press, 2012, 912 p.

DEMANTKÉ, J.; MALLET, C.; DAVID, N., VALLET, B. Dimensionality based scale selection in 3D lidar point clouds. **2011 ISPRS Workshop on Laser Scanning**, 2011, Calgary, Canadá, 2011.

DHAKER, H.; NGOM, P.; MBODJ, M. Overlap coefficients based on Kullback-Leibler divergence: exponential populations case. **International Journal of Applied Mathematical Research**, v. 6, n. 4, p. 135-140, 2017.

DOUGLAS, D. H.; PEUCKER, T. K. Algorithms for the reduction of the number of points required to represent a digitized line or its caricature. **Cartographica: The International Journal for Geographic Information and Geovisualization**, v. 10, n. 2, p. 112-122, 1973.

DRESCH, G.; SANTOS, D. R. Avaliação automática da acurácia relativa de dados LiDAR aerotransportado. **Boletim de Ciências Geodésicas**, v. 21, n. 4, 2015, p. 730-749.

DONGARA, J.; FOSTER, I.; FOX, G.; GROPP, W.; KENNEDY, K.; TORCZON, L.; WHITE, A. **Sourcebook of parallel computing**. San Francisco: Morgan Kaufmann, 2002. 864 p.

EIJKHOUT, V.; VAN DE GEIJN, R.; CHOW, E. **Introduction to High Performance Computing**. Zenodo. 2016.

EL-SHEIMY, N.; VALEO, C.; HABIB, A. **Digital terrain modeling: Acquisition, manipulation, and applications**. Londres: Artech House, 2005. 257 p.

FAZAN, A. J. **Extração de contornos de telhado de edifícios no espaço-objeto baseada em snakes e programação dinâmica a partir de TIN/LASER e imagens aéreas de alta resolução**. 2011. 113 f. Tese (Doutorado em Ciências Cartográficas) – Faculdade de Ciências e Tecnologia, Universidade Estadual Paulista, Presidente Prudente, 2011.

NAMIKAWA, L.M.; FELGUEIRAS, C.A.; MURA, J.C.; ROSIM, S.; LOPES, E.S.S. **Modelagem numérica de terreno e aplicações**. São José dos Campos: INPE. Disponível em: <<http://mtc-m12.sid.inpe.br/col/sid.inpe.br/marciana/2003/03.10.11.36/doc/publicacao.pdf>>. Acesso em: 21 dez 2020.

FERNANDES, V. J. M.; DAL POZ, A. P. Refinamento geométrico de contornos e cumeeiras de telhados de edifícios extraídos de dados LASER com uso de imagem aérea. **Boletim de Ciências Geodésicas**, v. 20, n. 3, 2014. p. 647-668.

FERNANDES, V. J. M.; DAL POZ, A. P. A Markov-random-field approach for extracting straight-line segments of roofs from high-resolution aerial images, **IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing**, v. 9, n. 12, 2016, p. 5493-5505.

FERNANDES, V. J. M.; MARTINS, E. F. O.; DAL POZ, A. P.; IMAI, N. N. Filtragem de nuvem LASER para geração de MDT por krigagem. **Boletim de Ciências Geodésicas**, v. 23, n. 1, 2017. p. 196-212.

FISCHLER, M. A.; BOLLES, R. C. Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography. **Commun. ACM**, v. 24, n. 6, p. 381–395, 1981.

FISHER, R. A. The use of multiple measurements in taxonomic problems. **Annals of Eugenics**, v. 7, 1936, p. 179-188.

FLORIANI, L.; PUPPO, E. A survey of constrained Delaunay triangulation algorithms for surface representation. In: PIERONI, G. G. (Ed.) **Issues on machine vision**. Berlin: Springer-Verlag, 1989. 339 p.

FUKUSHIMA, K. Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position. **Biological cybernetics**, v. 36, n. 4, 1980. p. 193-202.

GHAMISI, P.; HÖFLE, B.; ZHU, X. X. Hyperspectral and LiDAR data fusion using extinction profiles and deep convolutional neural network. **IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing**. v. 10, n. 6, p. 3011-3024, 2016.

GILANI, S. A. N.; AWRANGJEB, M.; LU, G. Segmentation of airborne point cloud data for automatic building roof extraction. **GIScience and remote sensing**, v. 55, p. 63-89, 2017.

GOODFELLOW, I.; BENGIO, Y.; COURVILLE, A. **Deep Learning**. The MIT Press, 2016. 800 p.

GONZALEZ, R. C.; WOODS, R. E. **Digital image processing**. Upper Saddle River: Pearson Education, 2007. 954 p.

GOLUB, G. H.; VAN LOAN, C. F. **Matrix Computations**. Baltimore: The Johns Hopkins Press, 1996. 694 p.

GROSS, H.; THOENNESSEN, U. Extraction of lines from laser point clouds. **International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences**, v. 36, p. 86-91, 2006.

HABIB, A. F. Integration of LiDAR and photogrammetric data: triangulation and orthorectification. In: SHAN, J.; TOTH, C. (Ed.) **Topographic laser ranging and scanning: principles and processing**. Londres: CRC Press, 2008. 590 p.

HABIB, A. F.; KIM, E. M.; KIM, C. J. New methodologies for true orthophoto generation. **Photogrammetric Engineering and Remote Sensing**, v. 73, n. 1, p. 25-36, 2007.

HACKEL, T.; WEGNER, J. D.; SCHINDLER, K. Fast semantic segmentation of 3D point clouds with strongly varying density. In: **ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences**, v. III-3, Praga, p. 177-184, 2016.

HAYKIN, S. **Neural networks and learning machines**. Londres: Pearson Education, 2009. 906 p.

HENNESSY, J. L.; PATTERSON, D. A. **Computer architecture: a quantitative approach**. Morgan Kaufmann, 2011. 856 p.

HERMOSILLA, T.; RUIZ, L. A.; RECIO, J. A.; ESTORNELL, J. Evaluation of automatic building detection approaches combining high resolution images and LiDAR data. **Remote Sensing**, v. 3, n. 6, p. 1188-1210, 2011.

HYYPPÄ, J.; HYYPPÄ, H.; YU, X.; KAARTINEN, H.; KUKKO, A.; HOLOPAINEN, M. Forest inventory using small-footprint airborne LiDAR. In: SHAN, J.; TOTH, C. (Ed.) **Topographic laser ranging and scanning: principles and processing**. Londres: CRC Press, 2008. 590 p.

HO, N.; WONG, W. Exploiting half precision arithmetic in NVIDIA GPUs. In: **IEEE Conference High Performance Extreme Computing (HPEC)**. Waltham, USA, 2017.

INMAN, H. F.; BRADLEY, E. L. The overlapping coefficient as a measure of agreement between probability distributions and point estimation of the overlap of two normal densities. **Communications in Statistics – Theory and Methods**, v. 18, n. 10, p. 3851-3874, 1988.

JOHNSON, R. A.; WICHERN, D. W. **Applied multivariate statistical analysis**. Delhi: Pearson Education, 2007. 773 p.

KANG, Z.; YANG, J.; ZHONG, R. A Bayesian-Network-Based classification method integrating airborne LiDAR data with optical images. **IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing**, v. 10, n. 4, p. 1651-1661, 2017.

KLEIN, L. A. **Sensor and data fusion: a tool for information assessment and decision-making**, 2a edição. SPIE Press, 512 p.

KRAUS, K. **Photogrammetry: geometry from images and laser scans**. Berlim: Walter de Gruyter, 2007. 459 p.

KRIZHEVSKY, A., SUTSKEVER, I.; HINTON, G. ImageNet Classification with Deep Convolutional Neural Networks. In: **Proceedings of the 26th International Conference on Neural Information Processing Systems (NIPS)**. 2012

LECUN, Y.; BOTTOU, L.; BENGIO, Y.; HAFFNER, P. Gradient-Based Learning Applied to Document Recognition. In: **Proc. of the IEEE**, vol. 86, no. 11, 1998. p. 2278–2324.

LESLAR, M.; WANG, J. G.; HU, B. A comparison of two new methods of outlier detection for mobile terrestrial LiDAR data. In: International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, v. 38, 2010, Calgary, **Anais...**, ISPRS, 2010, p. 78-84.

LI, J.; XIAO, Y.; WANG, C. Quality assessment of building roof segmentation from airborne LiDAR data. In: International Conference on Geoinformatics, 2013, Kaifeng, **Anais...**, IEEE, 2013, p. 1-4.

LI, Z.; ZHU, C.; GOLD, C. **Digital terrain modeling: Principles and methodology**. Londres: CRC Press, 2004. 318 p.

LU, X.; GUO, Q.; LI, W.; FLANAGAN, J. A bottom-up approach to segment individual deciduous trees using leaf-off LiDAR point cloud data. **ISPRS Journal of Photogrammetry and Remote Sensing**, v. 94, 2014, 12 p.

MALPICA, J. A.; ALONSO, M. C.; PAPI, F.; AROZARENA, A.; AGIRRE, A. M. Change detection of buildings from satellite imagery and LiDAR data. **International Journal of Remote Sensing**, v. 34, n. 5, p. 1652-1675, 2013.

MATKAN, A. A.; HAJEB, M.; MIRBAGHERI, B.; SADEGHIAN, S.; AHMADI, M. Spatial analysis for outlier removal from LiDAR data. In: International Archives of Photogrammetry and Remote Sensing, v. 40, 2014, Tehran, **Anais...**, ISPRS, 2014, p. 187-190.

MCCOY, R. M. **Field methods in remote sensing**. New York: The Guilford Press. 159 p.

MENG, X.; CURRIT, N.; ZHAO, K. Ground filtering algorithms for airborne lidar data: A review of critical issues. **Remote Sensing**, v. 2, n. 3, 2010, p. 833-860.

MIKHAIL, E. M.; BETHEL, J. S.; MCGLONE, J. C. **Introduction to modern photogrammetry**. New York: John Wiley and Sons, 2001. 479 p.

MURAKAMI, H.; NAKAGAWA, K.; HASEGAWA, H.; SHIBATA, T.; IWANAMI, E. Change detection of buildings using an airborne laser scanner. **ISPRS Journal of Photogrammetry and Remote Sensing**, v. 54, 1999, p. 148-152.

MURDOCCA, M. J.; HEURING, V. P. **Principles of computer architecture**. Prentice Hall, 1999. 654 p.

NOVO, E. M. L. M. **Sensoriamento remoto: princípios e aplicações**. São Paulo: Edgard Blucher, 2008. 363 p.

OLIVEIRA, G. R. K. **Uso integrado de dados lidar e imagens aéreas aplicado na extração de contornos de telhados de edificações**. 2016. 105 f. Dissertação (Mestrado em Ciências Cartográficas) – Faculdade de Ciências e Tecnologia, Universidade Estadual Paulista, Presidente Prudente, 2016.

PANG, S.; HU, X.; WANG, Z.; LU, Y. Object-based analysis of airborne LiDAR data for building change detection. **Remote Sensing**, v. 6, n. 11, p. 10733-10749, 2014.

PECKHAM, R.; JORDAN, G. **Digital terrain modelling: Development and applications in a policy support environment**. Berlin: Springer-Verlag, 2007. 313 p.

PEDRINI, H.; SCHWARTZ, W. R. **Análise de imagens digitais: Princípios, algoritmos e aplicações**. São Paulo: Thomson, 2007. 528 p.

PESSOA, G. G.; SANTOS, R. C.; CARRILHO, A. C.; GALO, M.; AMORIM, A. Urban scene classification using features extracted from photogrammetric point clouds acquired by UAV. In: ISPRS Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 2019, Enschede, **Anais...**, ISPRS, 2019, p. 511-518.

PETRIE, G.; TOTH, C. Introduction to laser ranging, profiling, and scanning. In: SHAN, J.; TOTH, C. (Ed.) **Topographic laser ranging and scanning: principles and processing**. Londres: CRC Press, 2008. 590 p.

PFEIFER, N.; MANDLBURGER, G. LiDAR data filtering and DTM generation. In: SHAN, J.; TOTH, C. (Ed.) **Topographic laser ranging and scanning: principles and processing**. Londres: CRC Press, 2008. 590 p.

PFEIFER, N.; STADLER, P.; BRIESE, C. Derivation of digital terrain models in the SCOP++ environment. In: OEEPE Workshop on Airborne Laserscanning and Interferometric SAR for Digital Elevation Models, 2001, Estocolmo, **Anais...**, 2001, 14 p.

QIN, L.; WU, W.; TIAN, Y.; XU, W. LiDAR filtering of urban areas with region growing based on moving-window weighted iterative least-squares fitting. **IEEE Geoscience and Remote Sensing Letters**, v. 14, n. 6, p. 841-845, 2017.

RICHARDS, J. A. **Remote sensing digital image analysis: an introduction**. Londres: Springer, 2013. 494 p.

RÖNNHOLM, P.; HONKAVAARA, E.; LITKEY, P.; HYYPÄ, H.; HYYPPÄ, J. Integration of laser scanning and photogrammetry. In: International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, v. 36, 2007, Espoo, **Anais...**, ISPRS, 2007, p. 355-362.

ROTTENSTEINER, F. Automated updating of building data bases from digital surface models and multi-spectral images: potentials and limitations. In: International Archives of

Photogrammetry, Remote Sensing and Spatial Information Sciences, v. 37, 2007, Beijing, **Anais...**, ISPRS, 2007, p. 265-270.

THEODORIDIS, S.; KOUTROUMBAS, K. **Pattern recognition**. San Diego: Academic Press, 2003. 689 p.

TIAN, J.; CUI, S.; REINARTZ, P. Building change detection based on satellite stereo imagery and digital surface models. **IEEE Transactions on Geoscience and Remote Sensing**, v. 52, n. 1, p. 406-417, 2014.

SAINI, S.; JIN, H.; HOOD, R.; BARKER, D.; MEHROTRA, P.; BISWAS, R. The impact of Hyper-Threading on processor resource utilization in production applications. **18^a International Conference on High Performance Computing**, Bangalore, 2011. p. 1-10, 2011.

SAMARASINGHE, S. **Neural networks for applied sciences and engineering: from fundamentals to complex pattern recognition**. Boca Raton: Taylor and Francis, 2007. 570 p.

SÁNCHEZ-APARICIO, M.; DEL POZO, S.; MARTÍN-JIMÉNEZ, J. A.; GONZÁLEZ-GONZÁLEZ, E.; ANDRÉS-ANAYA, P.; LAGUELA, S. Influence of LiDAR Point Cloud Density in the Geometric Characterization of Rooftops for Solar Photovoltaic Studies in Cities. **Remote Sensing**, 12, 2020.

SANTOS, R. C.; GALO, M. Classificação de nuvem de pontos laser utilizando o conceito de análise de componentes principais e o fator de não ambiguidade. **Boletim de Ciências Geodésicas**. V. 22, n. 2, 2016. p. 196-216.

SANTOS, R. C.; GALO, M.; CARRILHO, A. C. Extraction of building roof boundaries from LiDAR data using an adaptive alpha-shape algorithm. **IEEE Geoscience and Remote Sensing Letters**, v. 16, 2019.

SANTOS, R. C.; GALO, M.; CARRILHO, A. C. Building boundary extraction from LiDAR data using a local estimated parameter for alpha shape algorithm. In: ISPRS Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 2018, Karlsruhe, **Anais...**, ISPRS, 2018, p. 127-132.

SANTOS, R. C.; PESSOA, G. G.; CARRILHO, A. C.; GALO, M. Building detection from LiDAR data using entropy and the k-means concept. In: ISPRS Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 2019, Enschede, **Anais...**, ISPRS, 2019, p. 969-974.

SANTOS, R, C. **Extração de feições retas e cálculo de entidades pontuais a partir de dados LASER para o ajustamento relativo de faixas**. 2015. Dissertação (Mestrado em Ciências Cartográficas) – Faculdade de Ciências e Tecnologia, Universidade Estadual Paulista, Presidente Prudente.

ŠEVO, I.; AVRAMOVIĆ, A. Convolutional Neural Networks based automatic object detection on aerial images. **IEEE Geoscience and Remote Sensing Letters**, v. 13, n. 5, 2016, p. 740-744.

SITHOLE, G.; VOSELMAN, 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.

SHI, X.; ZAKHOR, A. Fast approximation for geometric classification of lidar returns. In: IEEE International Conference on Image Processing (ICIP), v. 28, 2011, Bruxelas, **Anais...**, IEEE, 2011, p. 2925–2928.

SOTOODEH, S. Outlier detection in LASER scanner point clouds. In: International Archives of Photogrammetry and Remote Sensing, v. 36, 2006, Dresden, **Anais...**, ISPRS, 2006, p. 297-302.

ST-LAURENT, S. **The complete effect and HLSL guide**. Paradoxal Press, 2005. 324 p.

TOMMASELLI, A. M. G.; GALO, M.; TIEDKE, T.; RUY, R. S.; MORAES, M. V. A.; MATRICARDI, W. V. Development and assessment of a data set containing frame images and dense airborne LASER scanning point clouds. **IEEE Geoscience and Remote Sensing Letters**, v. 15, n. 2, 2018, p. 192-196.

VEGA, C.; HAMROUNI, A.; MOKHTARI, S. E.; MOREL, J.; BOCK, J. RENAUD, J. P.; BOUVIER, M.; DURRIEU, S. Ptrees: A point-based approach to forest tree extraction from

lidar data. **International Journal of Applied Earth Observation and Geoinformation**, v. 33, 2014. p. 98–108

VIEIRA, C. A. O.; MATHER, P. M. Techniques for estimating the positional and thematic accuracy of remotely sensed products. In: Anais do XII Simpósio Brasileiro de Sensoriamento Remoto, 2005, Goiânia, **Anais...**, INPE, 2005, p. 4351–4359.

VOSSelman, G. Slope based filtering of laser altimetry data. In: International Archives of Photogrammetry and Remote Sensing, v. 33, 2000, Amsterdam, **Anais...**, ISPRS, 2000, p. 935–942.

VOSSelman, G.; GORTE, B.; SITHOLE, G. Change detection for updating medium scale maps using laser altimetry. In: Proceedings of the XX ISPRS Congress, v. 35, 2004, Istanbul, **Anais...**, ISPRS, 2004, p. 207-212.

XIONG, J.; FANG, Y.; JIN, B.; ZHAO, Z. Automated DTM generation in urban areas with airborne LiDAR data. In: International Conference on Intelligent Human-Machine System and Cybernetics, 2012, Nanchang, **Anais...**, IEEE, 2012, p. 181-184.

XU, S.; VOSSelman, G.; ELBERINK, S. O. Detection and classification of changes in buildings from airborne laser scanning data. In: International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, v. II-5/W2, 2013, Antalya, **Anais...**, ISPRS, 2013, p. 343-348.

WEHR, A.; LOHR, U. Airborne laser scanning: an introduction and overview. **ISPRS Journal of Photogrammetry and Remote Sensing**, v. 54, n. 2-3, 1999, p. 68-82.

WEINMANN, M.; JUTZI, B.; HINZ, S.; MALLETT, C. Semantic point cloud interpretation based on optimal neighborhoods, relevant features and efficient classifiers. **ISPRS Journal of Photogrammetry and Remote Sensing**, v. 105, p. 286-304, 2015.

WEST, K.F.; WEBB, B.N.; LERSCH, J.R.; POTHIER, S.; TRISCARI, J.M.; IVERSON, A.E. Context-driven automated target detection in 3-D data. **Proceedings of SPIE**, Vol. 5426, p. 133–143. 2004.

WOLF, P. R.; DEWITT, B. A. **Elements of photogrammetry: with applications in GIS**. Boston: McGraw-Hill, 2000. 608 p.

YANG, B.; DONG, Z. A shape-based segmentation method for mobile laser scanning point clouds. **ISPRS Journal of Photogrammetry and Remote Sensing**, v. 81, 2013, p. 19-30.

ZAREA, A.; MOHAMMADZADEH, A. A novel building and tree detection method from LiDAR data and aerial images. **IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing**, v. 9, n. 5, p. 1864-1875, 2016.

ZHAO, R.; PANG, M.; LIU, C.; ZHANG, Y. Robust normal estimation for 3D LiDAR point clouds in urban environments. **Sensors**, v. 19, n. 5, 2019.

APÊNDICE A

DETALHES DA IMPLEMENTAÇÃO

Embora atualmente existam diferentes possibilidades de plataformas de desenvolvimento, tais como dispositivos móveis, por exemplo, que utilizam processadores com arquiteturas *Reduced Instruction Set Computer* (RISC) adaptadas da *Advanced RISC Machines* (ARM), o método apresentado neste trabalho foi implementado em computadores pessoais, que majoritariamente adotam a arquitetura x86_64⁴³, isto é, uma arquitetura *Complex Instruction Set Computer* (CISC) (MURDOCCA e HEURING, 1999). De acordo com Murdocca e Heuring (1999), independente do conjunto de instruções (*instruction set architecture* – ISA) considerado, os processadores funcionam da mesma forma: o ciclo “busca”, “decodifica” e “executa”, também conhecido como ciclo de von Neumann⁴⁴. Existem circuitos específicos para realizar cada uma dessas tarefas no processador. A primeira etapa “busca” faz a consulta na memória do computador pela próxima instrução a ser executada, após isso, o endereço de memória lido é então “decodificado” para identificar qual circuito lógico deve ser utilizado na próxima etapa, isto é, a “execução”. Dentre as diferenças entre modelos de processadores os mais importantes são: o conjunto de instruções, o quão rápido as instruções são executadas (frequência), e quantos núcleos de execução o circuito possui.

Em uma arquitetura RISC por exemplo, cada etapa do ciclo de von Neumann é executada em cada ciclo do processador, desta forma seriam necessários 3 ciclos para a execução de uma única instrução. Neste caso, apenas uma parte do circuito é utilizada em cada ciclo. De forma alternativa é possível empregar a técnica de “*pipelining*”, na qual a próxima instrução a ser executada é buscada na memória imediatamente após o término da busca da instrução atual, assim como apresenta a Figura A.1. Esta técnica permite uma maior taxa de execução de instruções por ciclo e uma melhor utilização dos circuitos do processador. Ao considerar a técnica de *pipelining* seriam necessários $n + 2$ ciclos para executar n instruções (em uma situação ideal), ao invés dos $3n$ ciclos sem *pipelining*. No entanto, durante o ciclo de execução de determinadas instruções, uma situação especial denominada de *branch* pode ocorrer

⁴³ Também denominada por x64 ou AMD64. Este trabalho adotou o termo x86_64 para remeter ao fato de que mantém a compatibilidade com a ISA x86 original.

⁴⁴ Por ter sido proposto por John von Neumann.

(Figura A.2). Neste caso ocorre um *jump* (desvio de fluxo)⁴⁵ para uma outra posição de memória na qual outra sequência de instruções será executada (MURDOCCA e HEURING, 1999). Assim, a instrução “agendada para execução”, isto é, que havia sido buscada na memória e estava sendo decodificada é desconsiderada (*flush*), e o ciclo é reiniciado no novo endereço de memória.

Figura A.1 – Ilustração da técnica de *pipelining* utilizada em algumas arquiteturas para aumentar a quantidade de instruções executadas em um determinado número de ciclos.

Legenda: B - busca, D - decodifica, E - executa a instrução.

| (a) Sem <i>pipelining</i> de instruções | | | | (b) Com <i>pipelining</i> de instruções | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Instrução 1 | B | D | E | | | | | | |
| Instrução 2 | | | | B | D | E | | | |
| Instrução 3 | | | | | | | B | D | E |
| Ciclo | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |

| | | | | | | | |
|-------------|---|---|---|---|---|--|--|
| Instrução 1 | B | D | E | | | | |
| Instrução 2 | | B | D | E | | | |
| Instrução 3 | | | B | D | E | | |
| Ciclo | 1 | 2 | 3 | 4 | 5 | | |

Fonte: Adaptado de Murdocca e Heuring (1999); Dongara *et al.* (2002).

Figura A.2 – Problema da técnica de *pipelining*: desvio (*branch*) do fluxo para uma outra instrução a qual não estava agendada. Nos ciclos 5 e 6 a unidade de execução ficou inoperante.

| | | | | | | | | |
|----------------------|---|---|---|---|---|---|---|---|
| Instrução <i>n</i> | B | D | E | | | | | |
| Instrução <i>n+1</i> | | B | D | E | | | | |
| Instrução <i>n+2</i> | | | B | D | | | | |
| Instrução <i>m</i> | | | | | B | D | E | |
| Instrução <i>m+1</i> | | | | | | B | D | E |
| Ciclo | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |

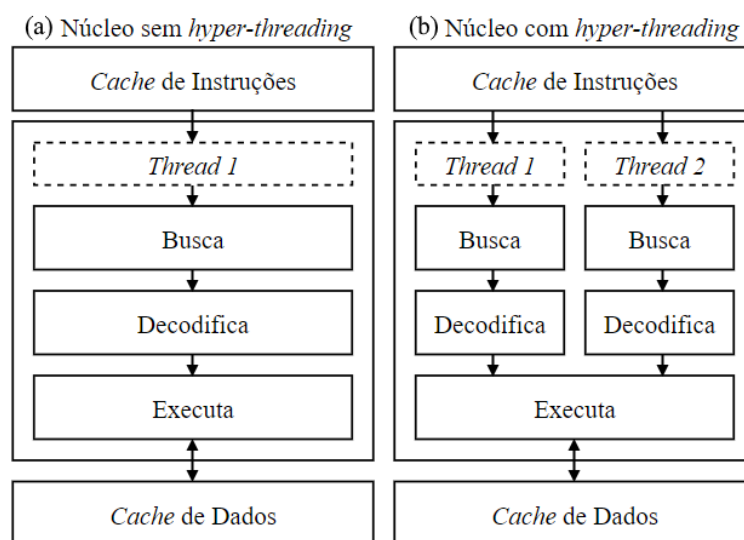
Fonte: Adaptado de Murdocca e Heuring (1999).

A ocorrência de desvio de fluxos é a principal desvantagem para a técnica de *pipelining*. Nessas situações algumas unidades do processador não são utilizadas durante alguns ciclos. Uma segunda técnica surgiu para minimizar esse problema, o *hyper-threading* (HT), assim como ilustra a Figura A.3. Introduzido pela Intel CorporationTM, o princípio consiste em incluir

⁴⁵ Um exemplo simples é a declaração de uma condição *if {...} else {...}* para controlar o fluxo de execução em um programa. Uma sequência diferente de instruções é executada de acordo com o predicado.

duas unidades de busca e decodificação e apenas uma unidade de execução em um núcleo, pois os circuitos de busca e decodificação correspondem a uma pequena porção da área do chip (EIJKHOUT *et al.*, 2016). Desta forma a unidade de execução não fica inoperante uma vez que pode executar instruções decodificadas por duas *threads* diferentes, mesmo que ocorra um desvio de fluxo em uma delas. O sistema operacional (SO) interpreta como dois núcleos, embora seja apenas uma unidade de execução. Assim, o SO reconhece $2n$ núcleos a partir um processador com n núcleos e HT. Vale ressaltar que os ganhos em performance são modestos, uma vez que só haverá vantagem quando ocorrer desvio de fluxo em uma das *threads*. Essa é uma das maneiras de processamento paralelo que existem (SAINI *et al.*, 2011), o *simultaneous multi-threading* (SMT), que pode ser decorrente de: (i) um núcleo com HT; (ii) múltiplos núcleos; e (iii) múltiplos núcleos com HT.

Figura A.3 – Diagrama simplificado das unidades que compõem a lógica principal de um processador.

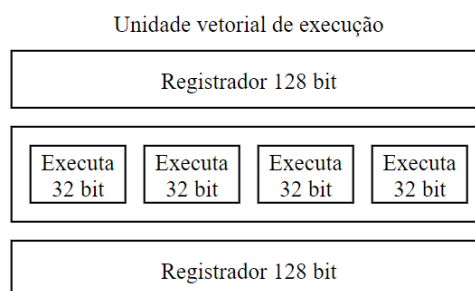


Fonte: Adaptado de Eijkhout *et al* (2016).

De acordo com Dongara *et al.* (2002), uma segunda maneira de processamento paralelo ocorre em nível de instrução com o objetivo de realizar múltiplas operações com uma única instrução. Essa estratégia é bastante útil, particularmente em processamento de dados matriciais (imagens digitais e álgebra linear, por exemplo), uma vez que as operações devem ser repetidas sobre uma grande quantidade de elementos. Neste caso, quando implementado no hardware, os registradores passam a comportar múltiplos valores ao invés de apenas um (Figura A.4). O interior da unidade de execução é alterado para comportar operações nesses registradores,

contendo mais unidades de lógica e aritmética, por exemplo. Na arquitetura x86 os registradores possuem originalmente 32 bits, os quais foram estendidos para 128 bits com as instruções SSE, e posteriormente para 256 bits e 512 bits com as instruções *Advanced Vector Extensions* (AVX) e AVX-512, respectivamente. Nota-se que o endereçamento de memória continua sendo os mesmos 32 bits para x86 e 64 bits para x86_64.

Figura A.4 – Exemplo simplificado da técnica de paralelismo em nível de instrução.



Fonte: Elaborada pelo autor.

Um exemplo de aplicação de paralelismo em nível de instrução seria a soma de quatérnios. Assumindo um quatérnio $q = (x, y, z, w)$ cujos elementos sejam representados por quatro variáveis de tipo ponto flutuante de precisão simples (32 bits no padrão IEEE 754), este quatérnio pode ocupar um dos oito registradores SSE (XMM0, ..., XMM7) de 128 bits. A operação de soma entre dois quatérnios se resume a uma única instrução (*addps*) ao invés de 4 operações individuais (*fadd*).

Ao contrário da técnica de *pipelining*, a qual o próprio processador se encarrega de agendar as próximas instruções, os meios mencionados de processamento paralelo (SMT e paralelismo em nível de instrução) devem ser explicitamente codificados para que se extraia maior performance. Apesar dos compiladores modernos realizarem muitas otimizações durante a geração do código de máquina (sequência de instruções), eles não empregam paralelismo em nível de instrução. Para contornar essa característica, a forma mais tradicional é substituir trechos específicos do código utilizando *assembly inline* nas linguagens C e C++. Por ser um trabalho moroso, é usual que sejam utilizados métodos de diagnóstico (*profiling*) no código para identificar trechos críticos que necessitem serem otimizados, deixando todo o resto do código para o compilador.

Para a implementação com SMT existem diferentes alternativas, entretanto a complexidade é maior devido ao problema de concorrência, isto é, múltiplas *threads* tentando

modificar o mesmo endereço de memória. Isso faz com que o programador tenha que gerenciar cuidadosamente o acesso aos recursos. Existem diversos mecanismos que auxiliam na sincronização dos processadores e as tarefas executadas por eles. Um mecanismo elegante é o *thread pool*, o qual começa com a inicialização de *threads* que ficam em repouso aguardando tarefas independentes⁴⁶ para executar. Dessa forma o programador implementa rotinas *thread-safe* (isto é, sem que haja conflitos de acesso aos endereços de memória) as quais são despachadas para o gerenciador que se encarrega de coordenar o trabalho.

Por serem processadores que precisam servir a diversos propósitos, os modelos modernos de CPU x86_64 empregam mais de mil instruções⁴⁷. No entanto, existem tarefas específicas que necessitam de um alto poder de processamento, tais como a renderização de cenas virtuais em tempo real. As técnicas de rasterização e traçado de raios (as duas mais comuns) necessitam que cada pixel da imagem tenha seu valor atualizado várias vezes por segundo para uma experiência fluída do usuário. Rasterizar um polígono como um triângulo envolve iterar sobre diversos pixels vizinhos, realizando operações similares neles. Nesse sentido, surgiram circuitos mais simples, específicos para essa aplicação: as unidades de processamento gráfico (*graphic processing unit* - GPU). Os primeiros modelos implementavam em hardware (fixo) os algoritmos de renderização, limitando os efeitos que poderiam ser aplicados. A partir de 2001, com a introdução da GeForce 3 e o padrão *Open Graphics Library* - OpenGL 1.3, as GPU passaram a contar com processadores de *shader*, os quais flexibilizaram os recursos e permitiram que os desenvolvedores tenham mais liberdade para implementar efeitos especiais. Os processadores de *shader* são simples, contendo poucas instruções (menos de 100 na versão 2.0), como descrito por St-Laurent (2005), com isso, os circuitos ocupam uma área menor e as fabricantes conseguem acoplar muitos deles em um único chip. A partir desse momento, os desenvolvedores passaram a implementar a parte de renderização com o uso de linguagens de *shader* (*OpenGL Shading Language* - GLSL e *High Level Shading Language* - HLSL⁴⁸, por exemplo), as quais são compiladas e submetidas especificamente para a GPU.

De acordo com Boyer e El Baz (2013), as GPU causaram impacto no segmento científico nas últimas duas décadas devido a alguns fatores: (i) existe uma correlação alta entre o conjunto de instruções utilizado para a geração de gráficos 3D e para processamentos de álgebra linear;

⁴⁶ No sentido de ser um trecho de código conhecido por *thread-safe*, isto é, não irá concorrer com recursos utilizados por outras *threads*.

⁴⁷ Uma descrição detalhada pode ser encontrada em: <http://ref.x86asm.net/coder32.html>. Acesso em: 05 maio 2020.

⁴⁸ Linguagem de *shader* para o Microsoft DirectX.

(ii) as GPU apresentam maior desempenho em ponto flutuante comparado aos CPU; (iii) a relação custo-benefício, no qual os pesquisadores conseguem ter acesso a um poder computacional maior por um preço menor. Inicialmente a comunidade científica se utilizava de *hacks*⁴⁹, isto é, ao invés de sintetizar uma imagem a partir de modelos 3D e texturas, os pesquisadores submetiam matrizes e os algoritmos programados na linguagem *shader* para a GPU executar. No entanto, as fabricantes flexibilizaram⁵⁰ a utilização das GPU para outros propósitos ao perceber o potencial nicho de novos consumidores: os cientistas. Nos últimos anos as fabricantes passaram a disponibilizar linhas de produtos específicos para fins científicos, tais como o chip A100 da nVIDIA, o qual conta com 6912 núcleos CUDA (*Compute Unified Device Architecture*, uma espécie de processador *shader*) que correspondem a 19,5 TeraFLOPS (para ponto flutuante de 32 bits).

⁴⁹ No sentido de utilizar algo para outra finalidade além da qual foi concebido originalmente.

⁵⁰ Principalmente com a introdução do *Open Compute Language* (OpenCL) e do CUDA.