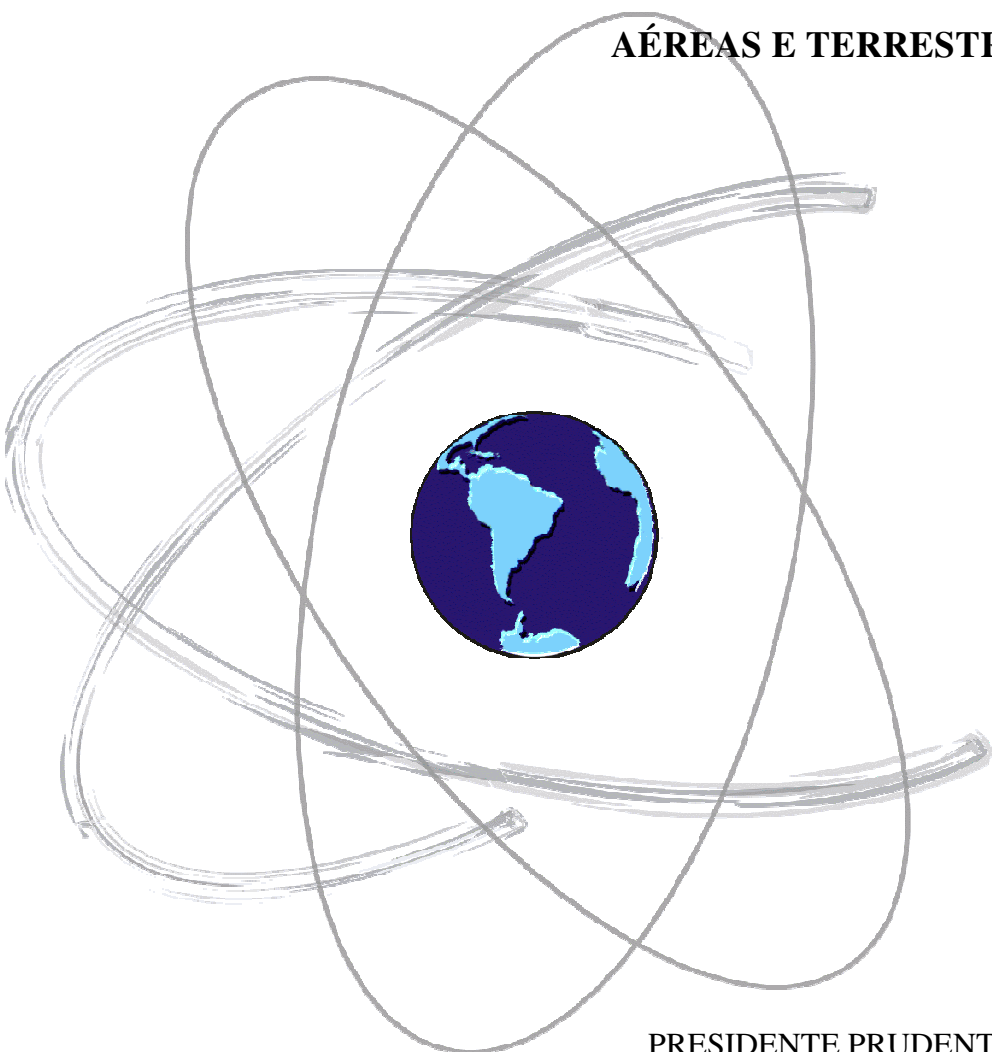




UNIVERSIDADE ESTADUAL PAULISTA
"JÚLIO DE MESQUITA FILHO"
Faculdade de Ciências e Tecnologia
Pós-Graduação em Ciências Cartográficas

JÉSSYCA MARIA DA SILVA PESTANA

**ESTABELECIMENTO DA CORRESPONDÊNCIA ENTRE IMAGENS
AÉREAS E TERRESTRES**



PRESIDENTE PRUDENTE
2017

JÉSSYCA MARIA DA SILVA PESTANA

**ESTABELECIMENTO DA CORRESPONDÊNCIA ENTRE IMAGENS
AÉREAS E TERRESTRES**

Dissertação apresentada ao Programa de Pós-Graduação em Ciências Cartográficas da Faculdade de Ciências e Tecnologia – UNESP, *campus* de Presidente Prudente.

Orientador: Prof. Dr. Antonio Maria Garcia Tommaselli

PRESIDENTE PRUDENTE
2017

FICHA CATALOGRÁFICA

P571e Pestana, Jéssyca Maria da Silva.
Estabelecimento da correspondência entre imagens aéreas e terrestres /
Jéssyca Maria da Silva Pestana. - Presidente Prudente : [s.n], 2017
114 f. :il.

Orientador: Antonio Maria Garcia Tommaselli
Dissertação (mestrado) - Universidade Estadual Paulista, Faculdade de
Ciências e Tecnologia
Inclui bibliografia

1. Correspondência de imagens. 2. Correspondência entre imagens
aéreas e terrestres. 3. Operadores de interesse. I. Pestana, Jéssyca Maria da
Silva. II. Tommaselli, Antonio Maria Garcia. III. Universidade Estadual
Paulista. Faculdade de Ciências e Tecnologia. IIII. Título.

CERTIFICADO DE APROVAÇÃO

TÍTULO DA DISSERTAÇÃO: Estabelecimento da correspondência entre imagens aéreas e terrestres

AUTORA: JÉSSYCA MARIA DA SILVA PESTANA

ORIENTADOR: ANTONIO MARIA GARCIA TOMMASELLI

Aprovada como parte das exigências para obtenção do Título de Mestre em CIÊNCIAS CARTOGRÁFICAS, área: AQUISIÇÃO, ANÁLISE E REPRESENTAÇÃO DE INFORMAÇÕES ESPACIAIS pela Comissão Examinadora:


Prof. Dr. ANTONIO MARIA GARCIA TOMMASELLI
Departamento de Cartografia / Faculdade de Ciências e Tecnologia de Presidente Prudente


Dr. ROBERTO DA SILVA RUY
Sensormap / Sensormap Soluções em Geotecnologia e Aerolevantamentos Ltda EPP


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

Presidente Prudente, 18 de setembro de 2017

À minha mãe, por me mostrar como é gratificante dar o melhor de si em tudo, e como pode ser duro não o fazer, e por me mostrar que a dedicação, por si só, já alcança conquistas incríveis.

AGRADECIMENTOS

À força espiritual que eu chamo de Deus, por me permitir seguir, sempre, e me fornecer a estrutura emocional necessária para finalizar esta etapa.

À minha família, em especial meus pais, irmã, e madrinhas, por me fornecerem todo o apoio psicológico, financeiro, emocional e espiritual de que eu precisei. Por estarem sempre tão disponíveis para mim e para minhas necessidades. Pelos exemplos de perseverança, compromisso e luta, que indiretamente tiveram tanta influência neste processo.

Aos meus grandes amigos, Gabriel e Camila, que estiveram ao meu lado nos momentos mais difíceis que enfrentei neste período, sendo o apoio que eu precisava e fazendo uma diferença tão positiva na minha vida. A eles, e aos meus amigos Jéssica e Pablo, que, juntos, compõem minhas definições de parceria e amizade.

Às famílias que fiz em Prudente: Iara e Pedro, Creuza e Roberto, e Dulce, por serem pais tão maravilhosos quando os meus biológicos estavam distantes, e pessoas que vou levar comigo para onde for.

Ao meu querido orientador: Prof^o Tommaselli. Pelo exemplo impecável de profissional, por toda a paciência, orientação, incentivo e amizade. Pela enorme contribuição na minha formação como profissional e como pessoa.

À todas as pessoas maravilhosas que eu conheci em Prudente, pela ajuda e amizade, por todas as experiências boas e ruins que tanto contribuíram para o meu crescimento pessoal. De forma especial: Hanson, Vanessa, Felipe, Ju, Alemão e Raquel.

À UNESP, por toda a infraestrutura fornecida. Aos colegas do PPGCC, em especial Guilherme, André e Marcos, pelas contribuições sem as quais não seria possível a realização deste trabalho. A todos os professores do Departamento de Cartografia e do PPGCC, em especial Prof^o Paulo, por ser meu primeiro contato com a pesquisa, pela amizade e contribuições durante este período, e ao Prof^o Galo, bem como ao Eng. Roberto Ruy, membros da banca, por todas as contribuições fornecidas no decorrer deste trabalho.

À empresa Sensormap e ao Grupo Engemap, pelos dados cedidos e por todo o suporte fornecido durante a realização do trabalho.

À CAPES, pelo suporte financeiro durante o desenvolvimento da pesquisa.

“Quando aceitamos nossos limites,
consequimos ir além deles”

Albert Einstein

RESUMO

A integração de plataformas aéreas e terrestres representa uma solução promissora para diversas aplicações de mapeamento, como projetos rodoviários e ferroviários. Estes sistemas possuem características complementares que permitem a visualização de um objeto na superfície a partir de diferentes perspectivas, além de, quando combinados, minimizarem limitações presentes na utilização individual dos mesmos, como o alcance da região imageada (caso terrestre) e o nível de detalhe dos dados adquiridos (caso aéreo). O problema ao integrar estes sistemas está no estabelecimento de correspondência entre as imagens. A aquisição a partir de diferentes perspectivas gera imagens com geometrias muito diferentes, dificultando o processo de correspondência, de modo que os métodos existentes não consigam solucionar o problema de maneira satisfatória. Com esta motivação, o presente trabalho propõe uma metodologia para o estabelecimento de correspondência entre imagens aéreas e terrestres, baseada na alteração prévia da geometria destas imagens, sendo necessário o conhecimento da orientação das mesmas e de um Modelo Digital do Terreno (MDT). A metodologia parte de pontos bem definidos, identificados na imagem aérea, que são projetados para o terreno e então para a imagem terrestre, definindo origens para os recortes que serão projetados e, posteriormente, submetidos à busca por correspondência. O trabalho apresenta uma revisão do estado da arte no que diz respeito a soluções para o problema de correspondência entre imagens aéreas e terrestres, bem como verifica o desempenho de métodos já existentes na solução do problema. Os experimentos foram baseados em imagens reais adquiridas sobre rodovias. Os resultados obtidos validam a metodologia proposta, mostrando que, ainda que existam limitações, é possível o estabelecimento de correspondência entre imagens aéreas e terrestres, e que este pode ser realizado tanto por métodos baseados em características como por métodos baseados em área.

Palavras-chave: Correspondência de imagens. Correspondência entre imagens aéreas e terrestres. Operadores de interesse. Mapeamento multi-plataformas.

ABSTRACT

The integration of aerial and terrestrial platforms represents a promising solution for several mapping applications. These systems have complementary features that allow the visualization of an object on the surface from several perspectives, and, when combined, minimize existing limitations in each individual system, such as the range of the imaged region (terrestrial case) and the level of detail of the acquired data (aerial case). When integrating these systems, the main problem that arises is related to the establishment of corresponding images. The acquisition from different perspectives generates images with very different geometries, hampering the matching process, so that the existing methods cannot solve the problem satisfactorily. With this motivation, the present work proposes a methodology for the establishment of correspondences between aerial and terrestrial images, based on the previous alteration of the geometry of these images, being necessary the knowledge of their orientation and of a Digital Terrain Model (DTM). The methodology starts with well-defined points, identified in the aerial image, that are projected to the terrain and then to the terrestrial image, defining origins for the cutouts that will be projected and later submitted to the correspondence search. The work presents a review of the state of the art regarding solutions of the correspondence between aerial and terrestrial images problem, as well as verifies the performance of existing methods in the solution of the problem. The experiments were based on actual images acquires on highways. The results obtained validate the proposed methodology, showing that, although there are limitation, it is possible to establish correspondence between aerial and terrestrial images, and that this can be done either by feature based methods or by area based methods.

Keywords: Image matching. Aerial and terrestrial images matching. Interest operators. Multi-platform mapping.

LISTA DE FIGURAS

Figura 1. Formação da imagem em um sistema óptico - caso ideal.....	23
Figura 2. Geometria da distorção radial.	27
Figura 3. Geometria da distorção descentrada.	28
Figura 4. Princípio da monorrestituição.	32
Figura 5. Retificação de uma imagem digital aérea.	34
Figura 6. Método de interpolação bilinear.	38
Figura 7. Diagrama do processo de busca por correspondência em imagens.	41
Figura 8. Imagens e subimagens em destaque de referência (a) e pesquisa (b).	42
Figura 9. Linha do tempo com operadores de interesse desenvolvidos até 2016.....	45
Figura 10. Resumo da etapa de extração dos locais candidatos no método SIFT: versões modificadas da imagem de entrada são obtidas e também diferenças dessas versões (DoGs), comparação entre pixels para determinação de pontos candidatos é realizada nas DoGs.	47
Figura 11. Magnitude e orientação dos gradientes são calculadas para os pixels no entorno do ponto-chave. Histograma de orientação é construído para cada subregião.....	49
Figura 12. Síntese do processo de construção do descritor local no operador SIFT.....	50
Figura 13. Visão geral do operador ASIFT: várias versões simuladas a partir das imagens originais (destacadas na figura) e comparadas entre si na busca por correspondência.	51
Figura 14. Representação dos conceitos de latitude (θ) e longitude (ϕ) empregados no ASIFT.	52
Figura 15. Amostragem dos parâmetros latitude (θ) e longitude (ϕ) no método ASIFT. Apenas as inclinações $t = 2, 2\sqrt{2}, 4$ foram representadas. Intervalos de variação $\Delta t = \sqrt{2}$ e $\Delta\phi = 72^\circ/b$	53
Figura 16. Princípio da identificação de quinas por meio da interseção de segmentos de retas.	55
Figura 17. Problema de oclusão no processo de correspondência: feições só aparecem em uma das imagens devido à variação de perspectiva (destaque).....	57
Figura 18. Problema de ambiguidade no processo de correspondência causado pela presença de padrões repetitivos (em destaque) na imagem.....	57
Figura 19. Sistemas aéreo (a) e terrestre (b) utilizados na aquisição dos dados.....	60
Figura 20. Localização da área de estudo, destaque para os extremos do trecho rodoviário contemplado no mapeamento.	61

Figura 21. Etapas da metodologia proposta.....	62
Figura 22. a) Conceito de retificação para o caso terrestre, onde c representa a distância focal da imagem e c' a distância focal da imagem retificada. b) Projeção da imagem terrestre realizada neste trabalho por meio do conceito de retificação.....	69
Figura 23. Recortes terrestres projetados, originados a partir de pontos de partida distintos (a, b).....	71
Figura 24. a) e b) Recortes terrestres retificados, gerados com um percentual muito pequeno de informação representada. c) Imagem terrestre que originou o recorte a).	72
Figura 25. Exemplo de retas que ligam pontos homólogos - mesma inclinação.....	74
Figura 26. Exemplos de casos em que não é possível aplicar o método de verificação baseado na inclinação das retas.	75
Figura 27. Processo de verificação visual das correspondências: pontos homólogos representados com mesmo símbolo e cor em ambas as imagens.	77
Figura 28. Conjunto de imagens utilizado nos experimentos.....	80
Figura 29. Conjunto de imagens utilizado nos experimentos (continuação).....	81
Figura 30. Localização das imagens terrestres em relação à aérea.	82
Figura 31. Resultado da extração de pontos de partida com o operador Harris, com destaque para a região considerada neste trabalho (área útil).	83
Figura 32. Resultado da extração de pontos de partida com o operador Förstner, com destaque para a região considerada neste trabalho (área útil).	84
Figura 33. Resultado da busca por correspondência com o operador SIFT nos recortes projetados (par 2). Correspondências incorretas ainda não eliminadas.....	85
Figura 34. Resultado da busca por correspondência com o operador SIFT nos recortes projetados (par 3). Correspondências incorretas ainda não eliminadas.....	86
Figura 35. Resultado da busca por correspondência com o operador SIFT nos recortes projetados (par 4). Correspondências incorretas ainda não eliminadas.....	87
Figura 36. Resultado da busca por correspondência com o operador SIFT nos recortes projetados (par 5). Correspondências incorretas ainda não eliminadas.....	88
Figura 37. Resultado da busca por correspondência com o operador SIFT nos recortes projetados (par 5 - continuação). Correspondências incorretas ainda não eliminadas.....	89
Figura 38. Resultado da busca por correspondência com o operador SIFT nos recortes projetados (par 6). Correspondências incorretas ainda não eliminadas.....	90
Figura 39. Resultado da busca por correspondência com o operador SIFT nos recortes projetados (par 6 - continuação). Correspondências incorretas ainda não eliminadas.....	91

Figura 40. Resultado da busca por correspondência com o operador SIFT nos recortes projetados (par 7). Correspondências incorretas ainda não eliminadas.....	92
Figura 41. Resultado da busca por correspondência com o operador SIFT nos recortes projetados (par 8). Correspondências incorretas ainda não eliminadas.....	93
Figura 42. Resultado final da classificação de correspondências (par 2).....	94
Figura 43. Resultado final da classificação de correspondências (par 3).....	94
Figura 44. Resultado final da classificação de correspondências (par 4).....	94
Figura 45. Resultado final da classificação de correspondências (par 5).....	95
Figura 46. Resultado final da classificação de correspondências (par 6).....	95
Figura 47. Resultado final da classificação de correspondências (par 7).....	96
Figura 48. Resultado final da classificação de correspondências (par 8).....	96
Figura 49. Resultado final da busca por correspondência na imagem aérea, representando apenas correspondências corretas, com <i>zoom</i> para região em destaque.....	97
Figura 50. Regiões correspondentes em um par de recortes retificados, selecionadas em: a) recorte aéreo e b) recorte terrestre.	98
Figura 51. Pares de recortes selecionados para o experimento, com destaque para a região definida como janela de referência no recorte aéreo.	98
Figura 52. Resultado da correspondência baseada em área para o recorte 1202: a) janela de referência, b) janela de pesquisa, com destaque para a região que apresentou maior correlação, e c) superfície gerada a partir dos valores de correlação calculados para a imagem de pesquisa, com destaque para o ponto que representa a correspondência verdadeira.	99
Figura 53. Resultado da correspondência baseada em área para o recorte 1109: a) janela de referência, b) janela de pesquisa, com destaque para a região que apresentou maior correlação, c) resultado esperado pelo programa, e d) superfície gerada a partir dos valores de correlação calculados para a imagem de pesquisa, com destaque para o ponto que representa a correspondência verdadeira.	100
Figura 54. Resultado da busca por correspondência em um caso ideal, realizada com os operadores SIFT e ASIFT.....	102
Figura 55. Resultado da busca por correspondência nas imagens inteiras, realizada com os operadores SIFT e ASIFT.....	102
Figura 56. Resultado da busca por correspondência no recorte 1202: a) correspondência baseada em área, com destaque para a região que apresentou maior correlação; e b) busca realizada com o SIFT, com a cor verde representando as correspondências corretas.....	105

Figura 57. Resultado da busca por correspondência no recorte 1109: a) correspondência baseada em área, com destaque para a região que apresentou maior correlação; e b) busca realizada com o SIFT, com a cor verde representando as correspondências corretas e vermelho as incorretas. 105

LISTA DE SIGLAS

AGAST: *Adaptive and Generic Accelerated Segment Test*

ASIFT: *Affine Scale Invariant Feature Transform*

CP: Centro Perspectivo

DoG: *Difference of Gaussians*

DTM: *Digital Terrain Model*

FAST: *Features from Accelerated Segment Test*

GNSS: *Global Navigation Satellite Systems*

GSA: *European GNSS Agency*

GSD: *Ground Sample Distance*

ICOMOS: *International Council of Monuments and Sites*

IMU: *Inertial Measurement Unit*

INS: *Inertial Navigation System*

IPOL: *Image Processing On Line*

MDT: Modelo Digital de Terreno

MM: Mapeamento Móvel

MMQ: Método dos Mínimos Quadrados

MVC: Matriz Variância-Covariância

OE: Orientação Exterior

OI: Orientação Interior

ORB: *Oriented BRIEF - Binary Robust Independent Elementary Features*

PGM: Portable Gray Map

POE: Parâmetros de Orientação Exterior

POI: Parâmetros de Orientação Interior

PP: Ponto Principal

RANSAC: RANdom SAmple Consensus

SIFT: *Scale Invariant Feature Transform*

SMM: Sensormap Mapeamento Móvel

SURF: *Speeded Up Robust Features*

SUSAN: *Smallest Univalve Segment Assimilating Nucleus*

UNESCO: *United Nations Educational, Scientific and Cultural Organization*

VANT: Veículo Aéreo Não Tripulado

SUMÁRIO

CAPÍTULO I INTRODUÇÃO

1.1	CONSIDERAÇÕES INICIAIS	16
1.2	OBJETIVOS	18
1.2.1	Geral	18
1.2.2	Específicos.....	18
1.3	REVISÃO DO ESTADO DA ARTE	18
1.4	ORGANIZAÇÃO DO TRABALHO.....	21

CAPÍTULO II OPERAÇÕES FOTOGRAMÉTRICAS

2.1	ORIENTAÇÃO DE IMAGENS	23
2.1.1	Orientação interior.....	23
2.1.2	Orientação exterior	29
2.2	MONORRESTITUIÇÃO	31
2.3	RETIFICAÇÃO DE IMAGENS DIGITAIS	33
2.3.1	Método direto	34
2.3.2	Método inverso ou indireto	36
2.4	REAMOSTRAGEM DE IMAGENS	37
2.4.1	Interpolação bilinear	38

CAPÍTULO III CORRESPONDÊNCIA DE IMAGENS

3.1	DEFINIÇÃO DO CONCEITO	40
3.2	CORRESPONDÊNCIA BASEADA EM ÁREA (<i>AREA-BASED MATCHING</i>).....	42
3.3	CORRESPONDÊNCIA BASEADA EM ATRIBUTOS OU CARACTERÍSTICAS (<i>FEATURE-BASED MATCHING</i>)	44
3.3.1	Operadores de interesse	44
3.4	PROBLEMAS NA CORRESPONDÊNCIA DE IMAGENS	56

CAPÍTULO IV
METODOLOGIA PARA CORRESPONDÊNCIA DE IMAGENS AÉREAS E
TERRESTRES

4.1	SÍNTESE DA PROPOSTA	59
4.2	RECURSOS MATERIAIS	59
4.3	ÁREA DE ESTUDO	60
4.4	METODOLOGIA	61
4.4.1	ETAPA I – Extração dos pontos de partida.....	62
4.4.2	ETAPA II – Projeção dos recortes	64
4.4.3	ETAPA III – Busca por correspondência.....	73
4.4.4	ETAPA IV – Projeção inversa para as imagens originais.....	75

CAPÍTULO V
EXPERIMENTOS, RESULTADOS E ANÁLISES

5.1	DESCRIÇÃO DOS EXPERIMENTOS	79
5.2	RESULTADOS	82
5.2.1	Etapa I.....	82
5.2.2	Etapa II	84
5.2.3	Etapa III	85
5.2.4	Etapa IV	93
5.2.5	Outros resultados	97
5.3	ANÁLISES	103

CAPÍTULO VI

CONCLUSÕES E RECOMENDAÇÕES	107
REFERÊNCIAS BIBLIOGRÁFICAS	110

CAPÍTULO I

INTRODUÇÃO

1.1 CONSIDERAÇÕES INICIAIS

A integração de imagens e dados adquiridos por plataformas aéreas e terrestres representa uma solução promissora para agências de cartografia e mapeamento (NEX *et al.*, 2015). O uso combinado desses dados multi-plataformas vem se mostrando interessante a diversas aplicações, uma vez que permite a aquisição de informações complementares da cena: enquanto a perspectiva aérea permite visualizar partes como os telhados das edificações, que são inacessíveis às plataformas terrestres, a perspectiva terrestre completa com as vistas laterais, gerando assim um mapeamento mais completo da cena. Um trabalho nesse sentido foi realizado por Córdoba (2012), no qual imagens adquiridas com um Veículo Aéreo Não Tripulado (VANT) foram combinadas a imagens terrestres do Palácio Itamaraty, em Brasília - DF, com a finalidade de restauração arquitetônica. De acordo com Córdoba (2012), o Comitê Internacional de Conservação e Restauração de Monumentos e Sítios (ICOMOS/UNESCO) “recomenda a Fotogrametria como uma poderosa técnica para a obtenção [...] de modelos tridimensionais, permitindo a restituição de diversos tipos de objetos [...]”, o que deve contribuir ainda mais para a aplicação da técnica, e consequentemente dos dados integrados, na restauração de patrimônios. Outras aplicações do uso combinado desses dados podem ser encontradas em Remondino *et al.* (2011), Molina *et al.* (2015), Nex *et al.* (2015) e Molina *et al.* (2016).

Além da potencial ferramenta de mapeamento que representam as tecnologias multi-plataformas, elas permitem minimizar limitações apresentadas pelos sistemas quando operados individualmente. Para o caso terrestre tem-se a limitação quanto ao alcance das imagens. No mapeamento rodoviário, por exemplo, não é possível extrair muita informação sobre o entorno da rodovia partindo somente de imagens terrestres. Esta informação é útil para o estudo de ocupações na faixa de domínio mantida no entorno das rodovias, sendo necessário, portanto, utilizar também imagens aéreas, que possibilitam uma vista mais ampla da cena. Em relação aos sistemas aéreos pode-se mencionar as áreas de oclusão: regiões cobertas por pontes e viadutos que são inacessíveis às imagens aéreas. Outra limitação deste sensor é quanto ao nível de detalhe alcançado, que é bastante inferior ao obtido no caso

terrestre. O mapeamento de sinalização vertical em rodovias, por exemplo, é bastante limitado quando se utilizam apenas imagens aéreas.

Como pode ser visto, a integração das plataformas terrestres e aéreas apresenta diversas vantagens em aplicações de mapeamento. O problema ao se trabalhar com esses dados de forma integrada está no estabelecimento de correspondência entre estas imagens. As técnicas de correspondência convencionais, baseadas em funções de correlação, permitem a identificação em imagens com alto grau de semelhança radiométrica e geométrica. No caso do problema tratado, as imagens são obtidas a partir de diferentes pontos de vista. Essa mudança de perspectiva faz com que as imagens tenham geometrias muito diferentes, dificultando ou, na maioria das vezes, impossibilitando o estabelecimento de correspondência a partir dos métodos convencionais.

Os chamados operadores de interesse são algoritmos capazes de identificar pontos estáveis nas imagens, ou seja, pontos cujos descritores (ou características) sofrem pequena variação com as mudanças de perspectiva. Estes pontos são chamados pontos de interesse ou pontos-chave (*keypoints*) e podem, após a extração, ser utilizados para finalidade de correspondência. O estado da arte destes operadores busca a invariância aos seis parâmetros de uma transformação afim, podendo então representar o que se tem de mais próximo de uma solução para o problema de correspondência entre imagens de diferentes plataformas. No entanto, alguns autores (ALCANTARILLA *et al.*, 2011; RUBLEE *et al.*, 2011; LEVI e HASSNER, 2016) afirmam que há limites para as variações de iluminação e perspectiva para as quais estes operadores conseguem obter resultados satisfatórios.

Gerke *et al.* (2016) apresentam um estudo sobre a aplicabilidade de operadores de interesse e outras estratégias na solução do problema de correspondência entre imagens com variações de perspectiva. Os resultados obtidos mostraram que em situações mais críticas, como é o caso de imagens obtidas a partir de diferentes plataformas e em dias diferentes, a maioria das abordagens não apresentou solução. Jende *et al.* (2016) afirmam que para grandes variações, como acontece entre imagens aéreas e terrestres, mesmo o estado da arte dos operadores de interesse não é capaz de solucionar o problema de maneira satisfatória.

A correspondência automática, ou semiautomática, entre imagens aéreas e terrestres pode então ser considerada um desafio que ainda não foi solucionado. Nesse contexto, destaca-se a importância de se estudar uma metodologia que permita estabelecer correspondência entre essas imagens, que é justamente a proposta deste trabalho.

O presente trabalho utiliza dados cedidos pela empresa Sensormap Geotecnologia, pertencente ao Grupo Engemap e integrante do projeto internacional mapKITE (MOLINA *et*

al., 2016), financiado pela Agência Europeia de GNSS (*European GNSS Agency – GSA da European Commission – processo 641518, EDITAL HORIZON 2020*). O projeto mapKITE, sob responsabilidade da empresa espanhola GeoNumerics S. L., portadora da patente, visa o desenvolvimento de um protótipo inédito baseado na integração das tecnologias VANT e veículo de mapeamento móvel terrestre, ambas operando simultaneamente. Desta forma, o projeto MapKITE representa uma aplicação concreta da integração das plataformas terrestre e aérea e, conseqüentemente, da necessidade de se estabelecer correspondência entre as imagens provenientes destes sensores. Este projeto foi um dos motivadores deste trabalho.

1.2 OBJETIVOS

1.2.1 Geral

Desenvolver uma metodologia capaz de estabelecer correspondência entre pontos em imagens aéreas e terrestres.

1.2.2 Específicos

- Implementar um programa para aplicar a metodologia proposta de forma automática ou semiautomática;
- Testar o programa com dados reais;
- Realizar análise da busca por correspondência observando fatores como quantidade de correspondências identificadas, quantidade de correspondências corretas e incorretas;
- Comparar resultados obtidos com métodos já existentes.

1.3 REVISÃO DO ESTADO DA ARTE

Apesar do interesse na correspondência e integração de dados terrestres e aéreos crescer constantemente, ainda é limitada a quantidade de trabalhos desenvolvidos a respeito do assunto (NEX *et al.*, 2015). Nesta seção são apresentados dois trabalhos encontrados na literatura envolvendo combinação de imagens ópticas: um deles analisa a aplicabilidade de operadores de interesse no contexto da correspondência de imagens, envolvendo imagens

adquiridas por diferentes plataformas (GERKE *et al.*, 2016), e o outro propõe uma nova abordagem para correspondência de imagens aéreas e terrestres (JENDE *et al.*, 2016). Pode-se mencionar ainda os trabalhos relacionados ao desenvolvimento do sistema mapKITE que, embora não dependa da correspondência das imagens, trata da integração dos sistemas e, assim, representa uma potencial aplicação da metodologia proposta neste trabalho, bem como das demais propostas com a mesma finalidade.

O primeiro estudo, realizado por Gerke *et al.* (2016), analisa o estado da arte dos descritores/detectores de características quanto ao seu desempenho no processo de correspondência de imagens e a influência de técnicas de pré-processamento neste contexto. Os algoritmos testados foram: SIFT (LOWE, 1999), ASIFT (MOREL e YU, 2009), SURF (BAY *et al.*, 2008), ORB (RUBLEE *et al.*, 2011), KAZE (ALCANTARILLA *et al.*, 2012), e AKAZE (ALCANTARILLA *et al.*, 2011). A influência da filtragem Wallis, bem como de outras técnicas de pré-processamento, também foi analisada.

Os testes foram realizados com várias combinações de imagens, considerando imagens adquiridas pela mesma e por diferentes plataformas. Os experimentos mostraram que todos os algoritmos analisados apresentaram solução satisfatória apenas em condições ideais, isto é, mesmo sensor e plataforma, condições de iluminação semelhantes e pouca variação de perspectiva. Embora todos os algoritmos tenham solucionado este caso sem grandes problemas, não identificando nenhuma correspondência incorreta, os algoritmos SURF e ASIFT se destacaram, obtendo cerca de 500 ~ 600 correspondências, aproximadamente o dobro do total identificado pelos demais algoritmos.

Ao se comparar imagens com grande variação de perspectiva, ainda que obtidas pela mesma plataforma, o desempenho dos algoritmos decaiu bruscamente. Para este teste, o ASIFT superou os demais, obtendo mais que o dobro de correspondências corretas identificadas pelo segundo colocado, SIFT. Ainda assim, esse número representa menos de 10% da quantidade identificada ao se analisar uma situação ideal, ou seja, menos de 60 correspondências, sendo 90% delas verdadeiras. Ao se analisar imagens adquiridas por diferentes plataformas, mas com perspectivas semelhantes, o SURF foi a única abordagem que apresentou solução satisfatória, mas em quantidade bastante reduzida: cerca de 20 correspondências corretas (90% do total identificado). Ainda para este teste, a combinação do SIFT com a filtragem Wallis apresentou a melhor solução: cerca de 100 correspondências corretas, não sendo fornecido o total identificado. Nos demais algoritmos a filtragem Wallis não apresentou melhora significativa do resultado.

O autor analisou ainda um “par errado” de imagens, ou seja, imagens de regiões que não eram correspondentes. Neste experimento o ASIFT foi a única abordagem que não encontrou nenhuma correspondência para o caso aéreo, resultado que era esperado para o teste. Por fim, nas situações mais críticas em que há uma mudança significativa na perspectiva das imagens, combinado à presença de variação de iluminação na cena, nenhuma abordagem apresentou solução.

Jende *et al.* (2016) apresentam um procedimento avançado de correspondência que permite estabelecer relação entre ortoimagens aéreas e imagens terrestres. O procedimento é um pré-requisito para a determinação da orientação de imagens panorâmicas terrestres, baseado na utilização de parâmetros de orientação acurados, provenientes de imagens aéreas. O problema que se pretende tratar é o de posicionamento em cenários de “*canyon urbano*”, onde a grande quantidade de arranha-céus impede o estabelecimento da linha de comunicação com os satélites de navegação e, conseqüentemente, a determinação da posição da plataforma terrestre com alta acurácia.

Dados provenientes de diferentes sensores geram dificuldades com respeito a mudanças de iluminação, radiométricas e às diferenças de perspectivas originais. Para lidar com estas diferenças na detecção de feições, a extração de pontos com base em características é realizada apenas na imagem aérea, utilizando uma versão modificada do operador Förstner (FÖRSTNER e GULCH, 1987) proposta por Köthe (2003). Em seguida é realizada uma busca na vizinhança por áreas com pontos esparsos, utilizando uma estrutura *kd-tree*, de modo a evitar regiões com marcações rodoviárias repetitivas que podem influenciar de forma negativa o processo de correspondência. Por fim, os pontos são projetados para as imagens terrestres e o estabelecimento de correspondência é realizado por um método baseado em área, que utiliza como métrica a correlação cruzada.

Experimentos realizados por Jende *et al.* consideram 14 recortes de uma ortofoto aérea e imagens panorâmicas terrestres, provenientes de Mapeamento Móvel (MM). Para avaliar o desempenho do método proposto, os resultados foram comparados a outros métodos baseados em características: o detector de cantos AGAST (MAIR *et al.*, 2010) em conjunto com o descritor SURF (BAY *et al.*, 2008); e o detector Förstner, mesma versão utilizada no método proposto, em conjunto com o descritor SURF. Dos 14 recortes analisados, 9 deles foram solucionados pelo método proposto, o que indica uma taxa de sucesso de 64%. Os testes iniciais mostraram uma melhora considerável da abordagem proposta em comparação com as demais analisadas nesse cenário particular de correspondência. Todas as abordagens foram capazes de estabelecer correspondência na maior parte dos casos analisados, sendo o

grande diferencial do método proposto a quantidade de correspondências corretas encontradas. Enquanto os outros métodos identificaram até 30 correspondências nos pares analisados, a abordagem proposta identificou, nos melhores casos, cerca de 10 correspondências, sendo identificado em boa parte dos recortes 5 ou menos correspondências. Embora apresente uma quantidade inferior de correspondências em alguns casos, a nova abordagem garantiu resultados quase sempre corretos: cerca de 90% de correspondências corretas. Algumas limitações ainda não tratadas adequadamente nesta nova abordagem referem-se ao problema de padrões repetitivos e ao fato de só trabalhar com feições planas.

Neste contexto de integração de sistemas terrestres e aéreos, foi proposto o sistema mapKITE, que é um protótipo baseado na integração das tecnologias VANT e veículo de mapeamento móvel terrestre, representando um novo paradigma na aquisição simultânea de dados terrestres e aéreos para finalidade de mapeamento. Por meio de um sistema de navegação localizado sobre o veículo terrestre, coordenadas são enviadas para o VANT em tempo real, estabelecendo assim uma ligação virtual entre a aeronave e o veículo terrestre e agindo como uma “pipa eletrônica de mapeamento”, ideia que deu origem ao nome adotado para o sistema – mapKITE. Além disso, no teto do veículo existe um alvo sinalizado, cujas coordenadas em movimento são utilizadas como apoio cinemático para o sistema aéreo. Alguns testes já foram realizados e os resultados preliminares mostraram que as estimativas de orientação exterior, interior, e localização de *tie points*, utilizando o controle cinemático do sistema, são tão boas quanto as que utilizam pontos de controle estáticos no terreno, o que resulta em um levantamento de menor custo e menor tempo de operação. Mais detalhes sobre o sistema mapKITE, aplicações e testes realizados podem ser encontrados em Molina *et al.* (2015, 2016). Este sistema, como mencionado anteriormente, representa uma aplicação concreta da integração das plataformas terrestre e aérea e, assim, uma potencial aplicação da metodologia proposta neste trabalho.

1.4 ORGANIZAÇÃO DO TRABALHO

O presente trabalho é dividido em seis capítulos. No primeiro, são apresentadas algumas considerações iniciais, onde se contextualiza o problema de correspondência que se pretende tratar, definem-se os objetivos deste trabalho e apresenta-se uma breve revisão do estado da arte em correspondência de imagens multi-plataformas. No segundo capítulo se inicia a parte de fundamentação teórica, apresentando-se algumas operações fotogramétricas que compõem a metodologia de correspondência que está sendo proposta neste trabalho. O

terceiro capítulo ainda trata da fundamentação teórica, apresentando o conceito de correspondência de imagens, métodos e problemas existentes, bem como o conceito de operadores de interesse, que se insere no contexto deste trabalho, sendo apresentado com maiores detalhes os operadores utilizados na metodologia e/ou nos experimentos: SIFT, ASIFT, Harris e Förstner. O quarto capítulo apresenta em detalhes a proposta metodológica deste trabalho, recursos materiais utilizados e área de estudo. O quinto apresenta os experimentos realizados e o conjunto de dados analisado, bem como os resultados obtidos e as análises. No sexto e último capítulo são apresentadas as conclusões do trabalho e algumas sugestões para trabalhos futuros.

CAPÍTULO II

OPERAÇÕES FOTOGRAMÉTRICAS

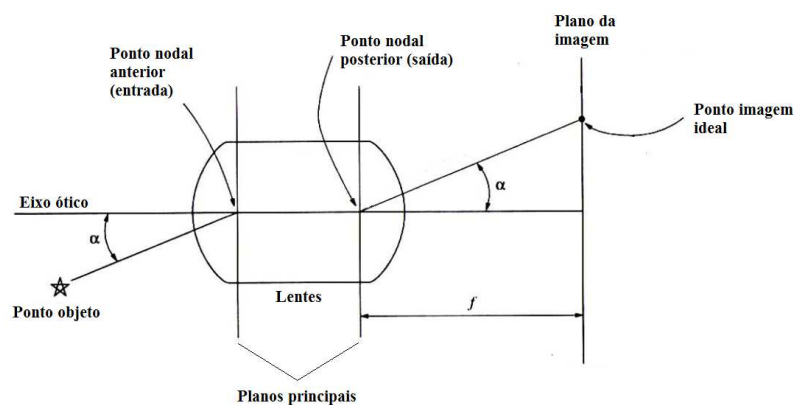
2.1 ORIENTAÇÃO DE IMAGENS

2.1.1 Orientação interior

Orientação interior (OI) é o processo de reconstrução da geometria interna da câmara, por meio do qual é possível reproduzir os feixes de raios formadores da imagem. Os parâmetros de orientação interior permitem a correção dos erros sistemáticos de modo que o feixe de raios possa ser reconstruído corretamente (MUGNIER *et al.*, 2013). Esta reconstrução é necessária pois, na prática, o sistema óptico de uma câmara nunca segue o modelo ideal.

O modelo ideal de um sistema óptico consiste em dois planos principais, que representam a entrada e saída de luz na lente. O eixo óptico é considerado perpendicular a estes planos e o feixe de luz atravessa as lentes sem sofrer nenhum desvio. A interseção desses planos com o eixo óptico da câmara define os pontos nodais anterior (entrada de luz) e posterior (saída de luz), como mostra a Figura 1 (MIKHAIL *et al.*, 2001; KRAUS, 2007).

Figura 1. Formação da imagem em um sistema óptico - caso ideal.



Fonte: Adaptado de Mikhail *et al.* (2001).

Na prática, o sistema óptico da câmara é diferente desse modelo ideal. Imperfeições na construção e montagem das lentes causam desvios na trajetória do feixe de luz ao atravessar as lentes, fazendo com que os ângulos de entrada e saída do sistema sejam diferentes (KRAUS, 2007). Uma vez que existe este desvio de trajetória, é necessária a

correção dos erros envolvidos para que seja possível reproduzir perfeitamente a geometria do feixe de raios. Assim, a orientação interior é o processo que define todos os parâmetros necessários à reconstrução do feixe de raios no espaço-objeto a partir de seu ponto correspondente no espaço-imagem. Estes parâmetros são intrínsecos e também chamados de Parâmetros de Orientação Interior (POI) (MUGNIER *et al.*, 2013). Mikhail *et al.* (2001) lista alguns dos principais POI:

- *distância focal (f)*: distância do centro perspectivo (CP) ao plano da imagem;
- *localização do ponto principal (PP)*: ponto definido pela projeção do CP no plano da imagem;
- *coeficientes das distorções que afetam o sistema de lentes*: distorção radial simétrica (k_1, k_2, k_3) e distorção descentrada (P_1, P_2, P_3).

Estes parâmetros intrínsecos adotados no procedimento de OI são determinados no processo de calibração geométrica da câmara e os valores são informados em um certificado de calibração (MUGNIER *et al.*, 2013; KRAUS, 2007). Em algumas situações usa-se, ainda, o modelo de afinidade (HABIB *et al.*, 2002). Para a câmara utilizada neste trabalho este modelo não é considerado devido ao efeito apresentado não ser significativo, isto é, ser muito menor que o tamanho do pixel.

A OI consiste principalmente na correção dos erros sistemáticos envolvidos, sendo necessário, primeiro, relacionar o valor medido na imagem original, que no nosso caso é digital, com o sistema considerado no sensor – sistema fotogramétrico (MUGNIER *et al.*, 2013). O procedimento que permite esta relação está descrito no item 2.1.1.1 e o refinamento das coordenadas, ou seja, a correção dos erros sistemáticos envolvidos, está apresentado no item 2.1.1.2.

2.1.1.1 Relação entre sistemas do espaço imagem

Uma imagem digital é formada por uma matriz de pixels organizados em colunas e linhas. Esta matriz define o sistema considerado na imagem: origem no centro do pixel superior esquerdo, eixo horizontal definido pelas colunas e eixo vertical pelas linhas. A unidade de medida desse sistema é o pixel e as coordenadas são dadas em coluna e linha. Assim, o primeiro pixel da imagem, localizado no canto superior esquerdo, possui coordenada (0,0) e a imagem tem dimensão $(N_{col}-1) \times (N_{lin}-1)$, onde N_{col} e N_{lin} representam respectivamente o número total de colunas e linhas da imagem. A primeira operação consiste

em transformar este sistema em um sistema com mesmo sentido do sistema fotogramétrico (dextrogiro) – consequentemente mesmo sentido do sistema de terreno – com origem no centro da imagem. As Equações 1 permitem esta transformação.

$$\begin{aligned}x &= (C - C_x) * \mu_x \\y &= -(L - C_y) * \mu_y\end{aligned}\tag{1}$$

onde,

x, y = coordenadas do ponto no sistema com origem no centro da imagem;

C, L = coordenadas do ponto no sistema imagem (coluna, linha);

μ_x, μ_y = dimensão do pixel em x e y , respectivamente;

C_x, C_y = coordenadas do centro geométrico da imagem, obtidas como apresentado nas Equações 2:

$$\begin{aligned}C_x &= (N_{col} - 1)/2 \\C_y &= (N_{lin} - 1)/2\end{aligned}\tag{2}$$

O centro geométrico da imagem, calculado pelas equações acima, não coincide com a projeção do centro perspectivo no plano da imagem (ponto principal), sendo necessária a redução das coordenadas a este ponto. As Equações 3 permitem essa redução, resultando nas coordenadas planimétricas do ponto no sistema fotogramétrico.

$$\begin{aligned}x_f &= x - x_0 \\y_f &= y - y_0\end{aligned}\tag{3}$$

onde,

x_f, y_f = coordenadas x, y do ponto no sistema fotogramétrico, não corrigidas dos erros sistemáticos;

x_0, y_0 = coordenadas do ponto principal.

Obtidas as coordenadas x_f, y_f do ponto no sistema fotogramétrico realiza-se o processo de refinamento das mesmas. A distorção radial simétrica, a distorção descentrada e a refração atmosférica são correções comumente consideradas no refinamento de coordenadas. No entanto, a refração aumenta conforme a altitude da câmara (ANDRADE, 2003). Imagens obtidas em baixa altitude sofrem pouquíssima influência deste efeito – muito menor que o tamanho do pixel – e, por isso, ele é desconsiderado no caso deste trabalho. Informações a respeito da correção e efeito da refração atmosférica podem ser encontrados em Mikhail *et al.* (2001) e Wolf (1983).

2.1.1.2 Refinamento de coordenadas

Os modelos matemáticos adotados nos procedimentos fotogramétricos foram desenvolvidos para uma geometria de imagem ideal, desconsiderando possíveis distorções, a fim de simplificar as equações e assumindo que esses efeitos seriam corrigidos em uma etapa de pré-processamento. O objetivo do refinamento de coordenadas é eliminar erros sistemáticos do processo, garantindo, assim, o pressuposto adotado na modelagem matemática de que os erros presentes são estritamente randômicos (MIKHAIL *et al.*, 2001).

Os erros sistemáticos são corrigidos na ordem oposta à que ocorrem (MIKHAIL *et al.*, 2001). O primeiro erro a afetar o sistema de lentes é a refração atmosférica, seguida da distorção descentrada e posteriormente da distorção radial. Logo, corrige-se primeiramente a distorção radial, seguida da distorção descentrada e por último a refração atmosférica. As coordenadas no sistema fotogramétrico (x_{fot}, y_{fot}) corrigidas dos erros sistemáticos podem ser obtidas como apresentado nas Equações 4.

$$\begin{aligned}x_{fot} &= x_f - \delta x_r - \delta x_d - \Delta r_x \\y_{fot} &= y_f - \delta y_r - \delta y_d - \Delta r_y\end{aligned}\tag{4}$$

onde,

x_f, y_f = coordenadas do ponto no sistema fotogramétrico, não corrigidas dos erros sistemáticos;

$\delta x_r, \delta y_r$ = efeito da distorção radial simétrica nas coordenadas x e y, respectivamente;

$\delta x_d, \delta y_d$ = efeito da distorção descentrada nas coordenadas x e y, respectivamente;

$\Delta r_x, \Delta r_y$ = efeito da refração atmosférica nas coordenadas x e y, respectivamente.

Estes valores ($\delta x_r, \delta y_r, \delta x_d, \delta y_d, \Delta r_x, \Delta r_y$) podem ser somados ou subtraídos das coordenadas no sistema fotogramétrico, dependendo de como estão apresentados no certificado de calibração (correção ou distorção) (OLIVEIRA, 2013). No presente trabalho, o procedimento de orientação interior é realizado no código implementado, não sendo necessária a utilização de imagens já corrigidas destes efeitos. Como já mencionado, a refração atmosférica não apresenta efeito significativo neste trabalho, sendo, portanto, desconsiderada. Os efeitos das distorções das lentes foram corrigidos e são apresentadas na sequência.

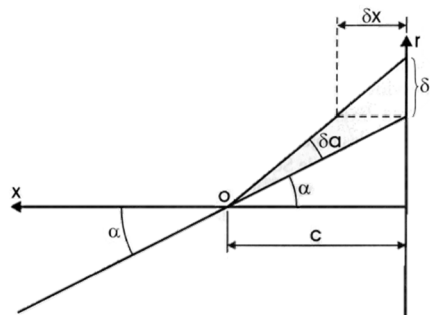
2.1.1.2.1 Distorções das lentes

As distorções, diferente de outras aberrações, têm efeito direto na medida realizada na imagem e não na qualidade da mesma (FIETE, 2013). Mikhail *et al.* (2001) considera as distorções das lentes a aberração mais relevante para a prática fotogramétrica. Estas distorções podem ser divididas em radial simétrica e descentrada, e devem ser corrigidas depois de as imagens serem reduzidas ao ponto principal (WOLF, 1983).

a) Distorção radial simétrica

Segundo Andrade (2003) a distorção radial pode ser encarada como a parcela não desejável da refração sofrida por um raio de luz ao atravessar um sistema de lentes. Assim, um raio de luz ao entrar no sistema de lentes forma um ângulo α com o eixo óptico, mas sai formando um ângulo de $(\alpha + \delta\alpha)$ devido à refração sofrida ao atravessar o sistema, causando um deslocamento de δr na posição da imagem no plano do negativo (ANDRADE, 2003), como mostra a Figura 2.

Figura 2. Geometria da distorção radial.



Fonte: Andrade (2003).

Na Figura 2, c representa a distância focal. Este tipo de distorção pode assumir valores tanto positivos quanto negativos. O efeito da distorção radial positiva é conhecido por efeito de almofada e ocorre quando um ponto é projetado radialmente para fora da imagem. Quando negativa, a distorção radial causa um efeito na imagem conhecido por barril, que é uma espécie de encolhimento. Mais detalhes podem ser vistos em Mikhail *et al.* (2001).

De acordo com Mikhail *et al.* (2001), o modelo polinomial, apresentado na Equação 5, permite o cálculo da distorção radial simétrica δr .

$$\delta r = k_1 r^3 + k_2 r^5 + k_3 r^7 + \dots \quad (5)$$

onde k_1, k_2, k_3 são os coeficientes da distorção radial, determinados no processo de calibração, e r é o raio calculado por:

$$r = \sqrt{(x - x_0)^2 + (y - y_0)^2} \quad (6)$$

Decompondo δr em componentes x e y tem-se seu efeito nas coordenadas:

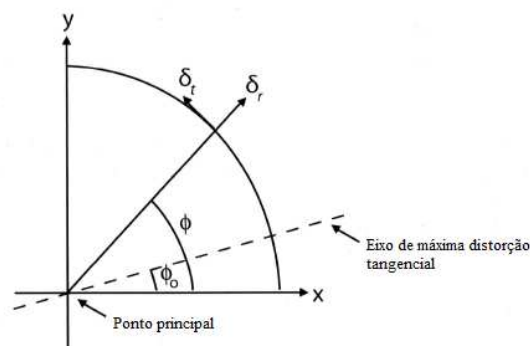
$$\begin{aligned} \delta x_r &= (x - x_0)(k_1 r^2 + k_2 r^4 + k_3 r^6 + \dots) \\ \delta y_r &= (y - y_0)(k_1 r^2 + k_2 r^4 + k_3 r^6 + \dots) \end{aligned} \quad (7)$$

b) Distorção descentrada

A distorção descentrada é o deslocamento no sentido perpendicular às linhas radiais, resultado de imperfeições no alinhamento das lentes. Lentes sofisticadas, utilizadas atualmente na Fotogrametria aérea, não são afetadas de forma significativa pela distorção descentrada. Já as lentes comerciais com zoom variável, comumente utilizadas em aplicações fotogramétricas terrestres, apresentam efeitos significativos desta distorção (MIKHAIL *et al.*, 2001; FIETE, 2013).

O efeito da distorção descentrada pode ser separado em duas componentes: radial (não simétrica) e tangencial; que variam conforme o ângulo $(\phi - \phi_0)$ formado pelo vetor que liga o ponto principal ao ponto de interesse e o eixo de máxima distorção tangencial. Quando este vetor é paralelo ao eixo de máxima distorção tangencial, a componente tangencial é maximizada e a componente radial é nula. Quando o vetor é perpendicular ao eixo, a componente radial é máxima e a tangencial é nula (MIKHAIL *et al.*, 2001; FIETE, 2013). A Figura 3 apresenta a geometria da distorção descentrada.

Figura 3. Geometria da distorção descentrada.



Fonte: Fiete (2013).

O efeito da distorção descentrada pode ser modelado em x (δx_d) e y (δy_d) pela função polinomial apresentada nas Equações 8.

$$\begin{aligned}\delta x_d &= [P_1(r^2 + 2(x - x_0)^2) + 2P_2(x - x_0)(y - y_0)] [1 + P_3r^2 + \dots] \\ \delta y_d &= [P_2(r^2 + 2(y - y_0)^2) + 2P_1(x - x_0)(y - y_0)] [1 + P_3r^2 + \dots]\end{aligned}\quad (8)$$

onde P_1 , P_2 , P_3 são coeficientes da distorção descentrada, sendo P_3 insignificante no caso abordado neste trabalho e, portanto, desconsiderado. Estes parâmetros são determinados no processo de calibração, e r é o raio calculado pela Equação 6.

2.1.2 Orientação exterior

A Orientação Exterior (OE) permite a recuperação da posição e atitude da fotografia em relação a um referencial terrestre (ANDRADE, 2003). A inclinação da plataforma de aquisição é inevitável e faz com que as fotografias sejam obtidas com o eixo da câmara inclinado. Na prática, mesmo com os equipamentos de estabilização, é impossível manter o eixo óptico da câmara perfeitamente vertical (WOLF, 1983), sendo necessária a determinação da OE das imagens.

Três parâmetros são suficientes para representar a inclinação de uma fotografia: *omega* (ω), que representa a rotação em torno do eixo x , *phi* (ϕ), que representa a rotação em torno do eixo y , e *kappa* (κ), rotação em torno do eixo z . Além destes, outros três parâmetros completam os chamados elementos de orientação exterior, que são os responsáveis por fornecer a posição no espaço e a orientação angular de uma fotografia inclinada. A posição espacial é comumente dada por X_0 , Y_0 , Z_0 , coordenadas tridimensionais da estação – ou centro perspectivo da câmara – no referencial do terreno (WOLF, 1983). Desta forma, a orientação exterior pode ser definida como a reprodução da posição e orientação da fotografia no momento em que ela foi obtida, representada então por esses seis parâmetros (X_0 , Y_0 , Z_0 , ω , ϕ , κ) conhecidos por Parâmetros de Orientação Exterior (POE).

Os POE variam conforme o movimento da câmara no espaço, ou podem ser constantes quando a câmara está fixa, como quando se utiliza um tripé para aquisição das imagens, por exemplo (MUGNIER *et al.*, 2013). Estes parâmetros são determinados por meio da relação entre o espaço-objeto e o espaço-imagem, que pode ser estabelecida de maneira direta, indireta ou integrada, como apresentado no item 2.1.2.1. Um dos modelos mais adotados na Fotogrametria para relacionar espaço-objeto e espaço-imagem é o modelo de colinearidade, que estabelece a condição de que o centro perspectivo de uma fotografia, um ponto no terreno e seu correspondente na imagem encontram-se todos localizados sob uma

linha reta (WOLF, 1983). As equações que expressam essa condição são chamadas equações de colinearidade e podem ser encontradas na forma direta (Equações 12), transformando coordenadas do espaço-objeto para o espaço-imagem, ou na forma inversa (Equações 13) que, sendo conhecida a coordenada Z do terreno, permitem transformam coordenadas do espaço-imagem para o espaço objeto. O modelo apresentado considera medidas realizadas no plano do negativo (WOLF, 1983).

$$\begin{aligned} x &= -f \frac{m_{11}(X - X_0) + m_{12}(Y - Y_0) + m_{13}(Z - Z_0)}{m_{31}(X - X_0) + m_{32}(Y - Y_0) + m_{33}(Z - Z_0)} \\ y &= -f \frac{m_{21}(X - X_0) + m_{22}(Y - Y_0) + m_{23}(Z - Z_0)}{m_{31}(X - X_0) + m_{32}(Y - Y_0) + m_{33}(Z - Z_0)} \end{aligned} \quad (12)$$

e

$$\begin{aligned} X &= X_0 + (Z - Z_0) \frac{m_{11}x + m_{21}y - m_{31}f}{m_{13}x + m_{23}y - m_{33}f} \\ Y &= Y_0 + (Z - Z_0) \frac{m_{12}x + m_{22}y - m_{32}f}{m_{13}x + m_{23}y - m_{33}f} \end{aligned} \quad (13)$$

onde,

x, y = coordenadas do ponto no sistema fotogramétrico, corrigidas dos efeitos sistemáticos;

f = distância focal da câmara;

X, Y, Z = coordenadas do ponto no referencial do terreno;

X₀, Y₀, Z₀ = coordenadas do centro perspectivo da imagem no referencial do terreno;

m₁₁, m₁₂, ...m₃₃ = elementos da matriz de rotação M dada por (MIKHAIL *et al.*, 2001):

$$M = \begin{bmatrix} \cos\varphi \cos\kappa & \cos\omega \operatorname{sen}\kappa + \operatorname{sen}\omega \operatorname{sen}\varphi \cos\kappa & \operatorname{sen}\omega \operatorname{sen}\kappa - \cos\omega \operatorname{sen}\varphi \cos\kappa \\ -\cos\varphi \operatorname{sen}\kappa & \cos\omega \cos\kappa - \operatorname{sen}\omega \operatorname{sen}\varphi \operatorname{sen}\kappa & \operatorname{sen}\omega \cos\kappa + \cos\omega \operatorname{sen}\varphi \operatorname{sen}\kappa \\ \operatorname{sen}\varphi & -\operatorname{sen}\omega \cos\varphi & \cos\omega \cos\varphi \end{bmatrix}$$

onde ω , φ , κ são parâmetros de orientação da imagem.

2.1.2.1 Formas de determinação dos Parâmetros de Orientação Exterior

O vetor que liga o centro perspectivo de uma fotografia e as coordenadas de um ponto no espaço imagem definem um feixe de raios no espaço. Os elementos de OE de todos os feixes em um bloco de imagens podem ser calculados simultaneamente pelo processo de fototriangulação por feixe de raios (*bundle adjustment*) (KRAUS, 2007). Quando são conhecidos pelo menos três pontos de apoio em cada imagem, o processo pode ser realizado também por resseção espacial, onde os POE são determinados para cada foto individualmente – mais detalhes podem ser encontrados em Mikhail *et al.* (2001). O método de ajuste

simultâneo possui a vantagem de não necessitar de apoio de campo em todas as imagens. Em vez disso, pontos de enlace (*tie points*) estabelecem ligações entre as imagens, formando um bloco e, assim, alguns pontos de controle nos extremos e na parte central são suficientes para orientar o bloco todo. A relação matemática entre as coordenadas imagem e o sistema de coordenadas do espaço objeto é feita pela condição de colinearidade, apresentada anteriormente (KRAUS, 2007). Estes são métodos de determinação indireta dos elementos de OE.

Outra forma de determinação dos POE é o chamado georreferenciamento direto, que permite a medida direta dos parâmetros de posição e orientação do sensor por meio da integração dos sistemas de navegação via satélite e inercial, em inglês *Global Navigation Satellite System* (GNSS) e *Inertial Navigation System* (INS), respectivamente. Quando a atitude da unidade de medida inercial, em inglês *Inertial Measurement Unit* (IMU), e a posição da antena GNSS são medidas com acurácia suficiente para a aplicação, o método dispensa o uso de pontos de apoio e não há necessidade de se realizar o procedimento de fototriangulação com as imagens (ABDULLAH *et al.*, 2013). Para detalhes sobre a aplicação do georreferenciamento direto pode-se consultar Kraus (2007), Cramer *et al.* (2000) e Abdullah *et al.* (2013). Os dados provenientes da IMU e GNSS quando combinados ao uso de pontos de apoio podem servir de suporte à fototriangulação. Este último é conhecido como determinação integrada dos elementos de OE e pode ser encontrado em Kraus (2007).

No presente trabalho, as informações de orientação exterior são conhecidas a priori e provenientes, no caso terrestre, de georreferenciamento direto e, no caso aéreo, de uma fototriangulação. O método utilizado na determinação destas informações não é relevante, desde que a qualidade necessária para estes dados seja alcançada. A metodologia não realiza a orientação exterior das imagens, mas recebe estas informações como dados de entrada.

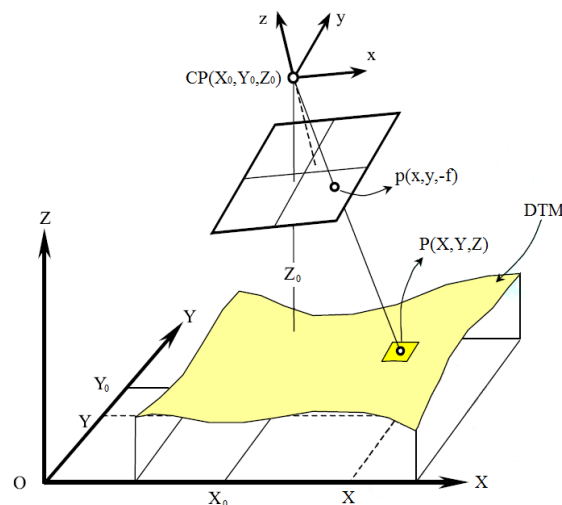
2.2 MONORRESTITUIÇÃO

A monorrestituição é o procedimento fotogramétrico que possibilita a determinação da posição de pontos no terreno a partir de medidas realizadas em uma única imagem. A fundamentação matemática consiste em transformar fotocoordenadas medidas em uma imagem para um referencial de terreno, utilizando as equações de colinearidade na forma inversa e informações da superfície provenientes, por exemplo, de um Modelo Digital de Terreno (MDT), em inglês *Digital Terrain Model* (DTM) (MITISHITA, 1997). De acordo

com Jauregui *et al.* (2002), uma importante aplicação da técnica é a atualização de bases de dados espaciais, onde a monorrestituição, juntamente com o avanço da tecnologia, permitem a realização do procedimento de forma mais barata e mais rápida. A fundamentação teórica apresentada neste item foi elaborada de acordo com o material apresentado por Galo (2014).

O princípio da monorrestituição é ilustrado na Figura 4, onde o raio passante pelos pontos p e CP intercepta o terreno (DTM), determinando o ponto $P(X,Y,Z)$.

Figura 4. Princípio da monorrestituição.



Fonte: Adaptado de Galo (2014).

Desta forma, pelo princípio da monorrestituição, um ponto P no espaço objeto, cujas coordenadas fotogramétricas são conhecidas bem como os parâmetros f , ω , ϕ , κ , X_0 , Y_0 e Z_0 da fotografia na qual o ponto foi medido, tem suas coordenadas planimétricas determinadas no terreno por meio das equações de colinearidade na forma inversa (Equações 13), sendo necessário também conhecer a coordenada altimétrica (Z) deste ponto no terreno. Segundo Galo (2014), de todas as informações necessárias à realização da monorrestituição, a determinação do valor de Z é a mais difícil de se obter. Uma solução iterativa para determinação deste valor pode ser realizada a partir das seguintes etapas:

- 1) Definição de um limiar ε como tolerância para o valor de Z ;
- 2) Estimativa inicial da componente Z (Z_i) para um dado ponto (x,y) na imagem;
- 3) Cálculo da posição planimétrica do ponto no terreno (X,Y) , utilizando as coordenadas fotogramétricas (x,y) e a componente Z_i , por meio das equações de colinearidade na forma inversa;

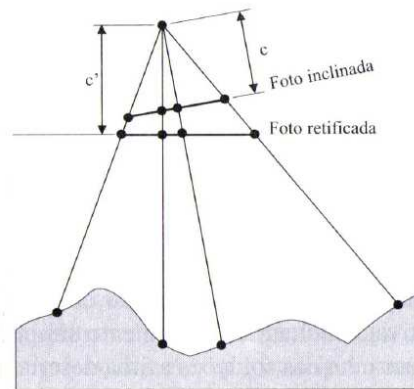
- 4) Obtida a posição planimétrica do ponto (X,Y), extrai-se o valor de Z correspondente à esta posição no MDT (Z_{MDT});
- 5) Verifica-se a condição $|Z_i - Z_{MDT}| \leq \varepsilon$
 - a. Se condição é verdadeira, realiza etapa 6;
 - b. Se condição é falsa, Z_i recebe o valor de Z_{MDT} ($Z_i=Z_{MDT}$) e volta para etapa 3;
- 6) Z_i recebe o valor de Z_{MDT} e calcula-se posição corrigida do ponto no terreno, ou seja, cálculo da posição planimétrica do ponto no terreno (X,Y), utilizando as coordenadas fotogramétricas (x,y) e a componente Z_{MDT} , por colinearidade inversa.

O presente trabalho utiliza a solução iterativa apresentada acima em uma das primeiras etapas realizadas na metodologia proposta: a determinação da posição de pontos no terreno que foram, inicialmente, identificados na imagem aérea. Outros métodos de determinação do valor de Z podem ser encontrados em Galo (2014).

2.3 RETIFICAÇÃO DE IMAGENS DIGITAIS

O processo de retificação no caso aéreo consiste na projeção de uma imagem segundo seu próprio feixe perspectivo para um plano horizontal. O procedimento permite corrigir a inclinação da câmara no momento da exposição fotográfica, de modo que uma imagem retificada possui orientação em ω e φ sempre nula (ANDRADE, 2003), mas ainda contém deslocamentos devido ao relevo do terreno. Este último pode ser corrigido por um processo denominado retificação diferencial ou ortoretificação, cujo produto gerado é uma ortofoto ou ortoimagem. Embora a ortofoto seja um produto mais acurado do que uma imagem retificada, as imagens retificadas são muito utilizadas por serem boas substitutas em casos nos quais a variação de relevo do terreno é moderada (WOLF, 1983). Também é possível, a partir do processo de retificação, alterar a distância focal, o que significa alterar a escala da imagem. Este procedimento permite a geração de várias imagens em um mesmo plano de referência, o que é útil, por exemplo, para a geração de estereomodelos ou mosaicos de imagens digitais (WOLF, 1983; ANDRADE, 2003). Para o caso terrestre, a retificação é feita através da projeção da imagem para um plano vertical (WOLF, 1983). A Figura 5 ilustra a ideia da geração de uma imagem retificada para o caso aéreo com alteração da distância focal, apresentada na imagem como c e c' .

Figura 5. Retificação de uma imagem digital aérea.



Fonte: Andrade (2003).

As imagens retificadas podem ser produzidas a partir de duas operações matemáticas principais: as transformações (afim, projetiva, polinomial, etc.) e as equações de colinearidade. As transformações podem ser realizadas conhecendo-se os parâmetros que permitem a relação entre a imagem original e a retificada. Estes parâmetros são normalmente determinados por um processo de ajustamento com o auxílio de alguns pontos de apoio, cuja quantidade pode variar de acordo com o modelo de transformação utilizado. A relação entre o espaço imagem e o sistema da imagem retificada também pode ser estabelecida com as equações de colinearidade. A vantagem ao utilizar o modelo de colinearidade é que, se houver conhecimento prévio da orientação da imagem original não há necessidade de apoio de campo. O processo de retificação pode ser realizado de duas formas: o chamado método direto, que parte da imagem original e obtém, para cada pixel, um correspondente na imagem retificada; e o chamado método inverso, que parte da imagem retificada e encontra, para cada um de seus pixels, um correspondente na imagem original. No presente trabalho, os dois métodos são utilizados: o centro de cada imagem retificada é calculado pelo método direto e o preenchimento da mesma é realizado pelo método inverso. Ambos são apresentados na sequência.

2.3.1 Método direto

O método direto de retificação parte da imagem original, a qual contém distorções, e calcula para cada pixel desta imagem sua posição na imagem retificada; ou seja, cada pixel da imagem original é corrigido e projetado no plano adotado para a retificação. Este método garante que todos os pixels da imagem original sejam retificados. As

coordenadas de cada pixel são então lidas no sistema da imagem (coluna, linha), transformadas para o sistema fotogramétrico e corrigidas dos erros sistemáticos pelo procedimento de OI, apresentado no item 2.1.1.

Realizado o procedimento de OI, o resultado fornece as coordenadas do ponto no sistema fotogramétrico, corrigidas dos erros sistemáticos, dadas por $(x_f, y_f, -f)$. Conhecida então a posição corrigida do ponto no sistema fotogramétrico, calcula-se sua projeção no plano de retificação por meio das equações de colinearidade na forma inversa (Equações 13). Uma vez que não é necessário o conhecimento prévio da posição da câmara ($X_0 = Y_0 = Z_0 = 0$), e que se deseja obter uma imagem retificada, sem alteração da escala, ou seja, Z representa a própria distância focal ($Z = -f_R$), o modelo de colinearidade na forma inversa pode ser reescrito como apresentado nas Equações 14 (LOPES, 2010):

$$\begin{aligned} X &= -f_R \frac{m_{11}x_f + m_{21}y_f - m_{31}f}{m_{13}x_f + m_{23}y_f - m_{33}f} \\ Y &= -f_R \frac{m_{12}x_f + m_{22}y_f - m_{32}f}{m_{13}x_f + m_{23}y_f - m_{33}f} \end{aligned} \quad (14)$$

onde,

$x_f, y_f, -f$ = coordenadas fotogramétricas na imagem original, e f é a distância focal da câmara;
 $X, Y, -f_R$ = coordenadas na imagem retificada, e f_R representa a distância do plano da imagem retificada ao centro perspectivo (plano de projeção), que neste caso é igual à distância focal;
 $m_{11}, m_{12}, \dots, m_{33}$ = elementos da matriz de rotação M em função de ω, ϕ, κ , apresentada no item 2.1.2.

As Equações 14 permitem a projeção dos pontos para o plano de retificação. A principal diferença deste modelo para o que permite a geração de uma ortofoto é o valor de Z , que na retificação representa o plano para o qual a imagem está sendo projetada e no caso da ortofoto assume valores diferentes para cada ponto, com base em um modelo da superfície do terreno, sem considerar árvores e feições antrópicas previamente conhecidas. Por fim, as coordenadas obtidas no plano de retificação são então transformadas para o sistema da imagem digital retificada, como apresentado nas Equações 15.

$$\begin{aligned} C_R &= \frac{X}{\mu_{x_R}} + C_{x_R} \\ L_R &= -\frac{Y}{\mu_{y_R}} + C_{y_R} \end{aligned} \quad (15)$$

onde,

C_R, L_R = coordenadas do ponto no sistema da imagem retificada (coluna, linha);

C_{x_R}, C_{y_R} = coordenadas do centro da imagem retificada, em pixels;

X, Y = coordenadas planimétricas do ponto no plano de retificação;

μ_{x_R}, μ_{y_R} = dimensão do pixel na imagem retificada em x e y , respectivamente.

As Equações 15 representam a forma inversa das Equações 1, considerando valores para a imagem retificada. Os valores de C_R, L_R obtidos não serão necessariamente valores inteiros, sendo necessário um processo de interpolação para a atribuição do tom de cinza ao pixel da imagem retificada.

O método direto de retificação garante que todo pixel da imagem original seja representado na imagem retificada, o que não significa que a imagem retificada será preenchida por completo, caso seja usado o preenchimento pelo vizinho mais próximo. Como os valores de coluna e linha calculados para a imagem retificada podem não ser inteiros, o arredondamento presente no processo de interpolação destes valores pode fazer com que mais de um pixel da imagem original resulte em uma mesma posição na imagem retificada. Desta forma, alguns pixels da imagem retificada podem armazenar informação duplicada e outros não armazenar informação nenhuma, resultando em “buracos”, que são vazios nas imagens.

2.3.2 Método inverso ou indireto

Ao contrário do método direto de retificação, o método inverso parte da imagem retificada, buscando para cada pixel o tom de cinza correspondente na imagem original. Para isto uma imagem vazia é gerada inicialmente, representando a imagem retificada, que neste momento ainda não está formada. As coordenadas de cada pixel nesta imagem ($coluna_R, linha_R$) são transformadas como apresentado nas Equações 1, considerando agora valores para a imagem retificada. Assim, a relação entre o sistema imagem (c_R, l_R) e o plano de retificação é dada por:

$$\begin{aligned} X &= (C_R - C_{x_R}) * \mu_{x_R} \\ Y &= -(L_R - C_{y_R}) * \mu_{y_R} \end{aligned} \quad (16)$$

onde X, Y são as coordenadas do ponto no plano de retificação. A formulação apresentada neste item envolve, basicamente, as formas inversas de equações apresentadas no item 2.3.1. As Equações 16 representam exatamente a forma inversa das Equações 15, que também consideram os valores para a imagem retificada. Realiza-se então o cálculo das coordenadas fotogramétricas a partir das equações de colinearidade na forma direta em que se desconsidera

a posição inicial da câmara e tem-se $Z = -f_R$, como apresentado nas Equações 17 (LOPES, 2010).

$$\begin{aligned} x_f &= -f \frac{m_{11}X + m_{12}Y - m_{13}f_R}{m_{31}X + m_{32}Y - m_{33}f_R} \\ y_f &= -f \frac{m_{21}X + m_{22}Y - m_{23}f_R}{m_{31}X + m_{32}Y - m_{33}f_R} \end{aligned} \quad (17)$$

As coordenadas (x_f, y_f) no sistema fotogramétrico, resultantes das equações acima, foram obtidas a partir da imagem retificada e, por isso, não estão afetadas pelas distorções do sistema de lentes. Para obter os valores na imagem original somam-se então os efeitos das distorções. Este processo é realizado de forma iterativa, uma vez que as coordenadas que permitirão o cálculo dos efeitos devem ser distorcidas, ao passo que os valores obtidos com as Equações 17 não apresentam distorção. Obtidas as coordenadas no sistema fotogramétrico sob efeito das distorções, soma-se a posição do ponto principal isolando, das Equações 3, as coordenadas do ponto no sistema com origem no centro da imagem:

$$\begin{aligned} x &= x_f + x_0 \\ y &= y_f + y_0 \end{aligned} \quad (18)$$

Realiza-se então a transformação de x e y para coluna e linha por meio das Equações 15, mas considerando agora valores para a imagem original:

$$\begin{aligned} C &= \frac{x}{\mu_x} + C_x \\ L &= -\frac{y}{\mu_y} + C_y \end{aligned} \quad (19)$$

Da mesma forma que no método direto, a posição obtida acima pode não representar um valor inteiro, sendo necessária a aplicação de um método de interpolação. No método inverso, cada pixel da imagem retificada tem seu correspondente encontrado na imagem original, e o tom de cinza é atribuído por meio de uma interpolação dos vizinhos, cujo conceito é apresentado na sequência. Esse procedimento garante que todos os pixels da imagem retificada recebam informação, eliminando a possibilidade de “buracos”, como ocorre no método direto.

2.4 REAMOSTRAGEM DE IMAGENS

Ao final das correções geométricas a que são submetidas as imagens é realizada a atribuição do tom de cinza, onde cada pixel da imagem de saída é preenchido com o nível de

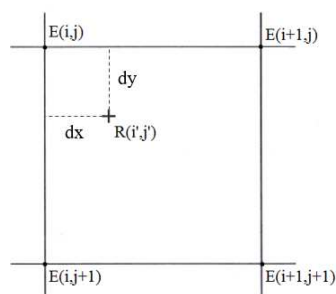
cinza da sua posição correspondente na imagem de entrada. No entanto, essa posição nem sempre representa um valor inteiro, ou seja, o centro de um pixel; podendo ser necessária a aplicação de um método de interpolação que considere a vizinhança da posição calculada (TEMIZ E KÜLÜR, 2008; BISCHOF e LEBERL, 2013).

Um dos métodos mais simples de interpolação é o do vizinho mais próximo, que apenas associa ao pixel o nível de cinza do vizinho mais próximo a ele na imagem. Este método atende de maneira satisfatória os casos em que a mudança de escala e geometria entre as imagens não é muito grande. Para todos os outros casos, a aparência da imagem de saída pode ser muito degradada devido ao efeito de serrilhamento (*aliasing*), um dos principais problemas envolvidos no processo de reamostragem de imagens. Para estes casos, o método de interpolação bilinear pode ser considerado o mais adequado (MIKHAIL *et al.*, 2001). Pode-se mencionar ainda os métodos de interpolação com base na utilização de *splines* bicúbicas e polinômios de Lagrange. Parker *et al.* (1983) realiza uma comparação entre estes e outros métodos de interpolação. Nos experimentos realizados neste estudo foi aplicado o método de interpolação bilinear, apresentado na sequência.

2.4.1 Interpolação bilinear

O método de interpolação bilinear realiza a atribuição do tom de cinza considerando os quatro pixels vizinhos à posição calculada. Desta forma, os pixels mais próximos do ponto são os que assumem maior influência no resultado (MIKHAIL *et al.*, 2001). Assim, um pixel $R(i',j')$ na imagem retificada tem seu valor de intensidade calculado com base na interpolação linear nas direções x e y , dadas por dx e dy , respectivamente, como apresentado na Figura 6.

Figura 6. Método de interpolação bilinear.



Fonte: Adaptado de Galo (2015) e Mikhail *et al.* (2001).

Para determinar o nível de cinza de $R(i',j')$ na imagem retificada, que se encontra entre os pixels $E(i,j)$, $E(i+1,j)$, $E(i,j+1)$ e $E(i+1,j+1)$ da imagem original, calcula-se primeiro valores auxiliares em x (dx) e y (dy):

$$\begin{aligned} dx &= i' - i \\ dy &= j' - j \end{aligned} \tag{20}$$

onde i,j representam a posição coluna e linha, respectivamente, do pixel na imagem original e i',j' analogamente para a imagem retificada. A partir destes valores, o tom de cinza final atribuído a $R(i',j')$ pode ser obtido pela Equação 21 (GALO, 2015).

$$\begin{aligned} R(i',j') &= (1 - dx)(1 - dy)E(i,j) + dx(1 - dy)E(i + 1,j) + \\ &dy(1 - dx)E(i,j + 1) + dxdyE(i + 1,j + 1) \end{aligned} \tag{21}$$

CAPÍTULO III

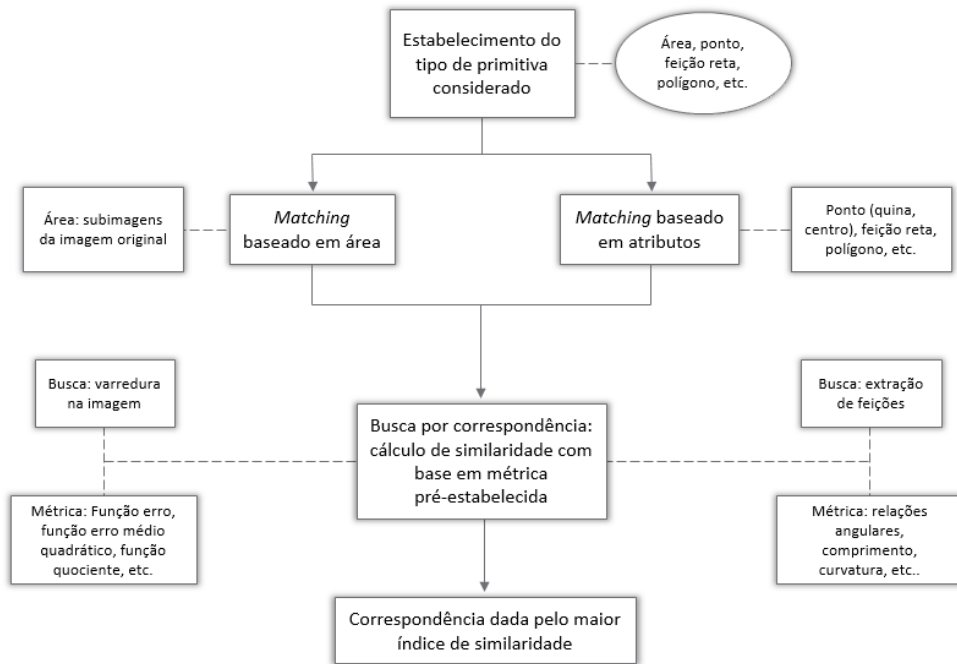
CORRESPONDÊNCIA DE IMAGENS

3.1 DEFINIÇÃO DO CONCEITO

O problema de correspondência na Fotogrametria pode ser definido como sendo a operação de busca por partes ou entidades (pontos, linhas, regiões) correspondentes em duas ou mais imagens de um conjunto, que permitem realizar uma relação de ligação entre estas imagens (GALO, 2003; MAYER *et al.*, 2013). A correspondência representa uma componente chave em diversas aplicações, uma vez que serve de base a operações na Fotogrametria, visão computacional e análise de imagens, e é utilizada em diversas outras como navegação, orientação, supervisão automática, visão robótica, análise de imagens médicas e ciências de modelagem e mapeamento (GRUEN, 2012).

De maneira geral, a busca por correspondência pode ser realizada seguindo algumas etapas. Primeiro, deve-se estabelecer o tipo de primitiva que será utilizada na correspondência das imagens. Em seguida, a busca na imagem pelas primitivas correspondentes é feita através de uma medida de similaridade. Esta similaridade é calculada a partir de uma métrica, que deve ser definida previamente. Por último, as medidas de similaridade são comparadas buscando o menor ou maior valor, dependendo da métrica utilizada. O diagrama apresentado na Figura 7 ilustra o processo de busca por correspondência nas imagens.

Figura 7. Diagrama do processo de busca por correspondência em imagens.



Não se pode esperar que um único método de correspondência satisfaça todas as aplicações. Existem diversas influências externas tanto em relação à cena imageada quanto em relação ao sensor usado na aquisição dos dados e até à própria imagem adquirida: condições físicas do ambiente (condições climáticas), variedade de sensores que podem ser utilizados na aquisição, características específicas das feições imageadas, geometria da cena, entre outras (FONSECA e MANJUNATH, 1996). Pelo procedimento descrito a respeito da realização de busca por correspondência, e também pelo diagrama apresentado, é possível notar alguns fatores variantes para os quais se pode atribuir diversas soluções. Devido à essa diversidade de soluções e também à diversidade de influências a que está submetida a situação de correspondência, é importante a classificação dos métodos de correspondência (HEIPKE, 1997).

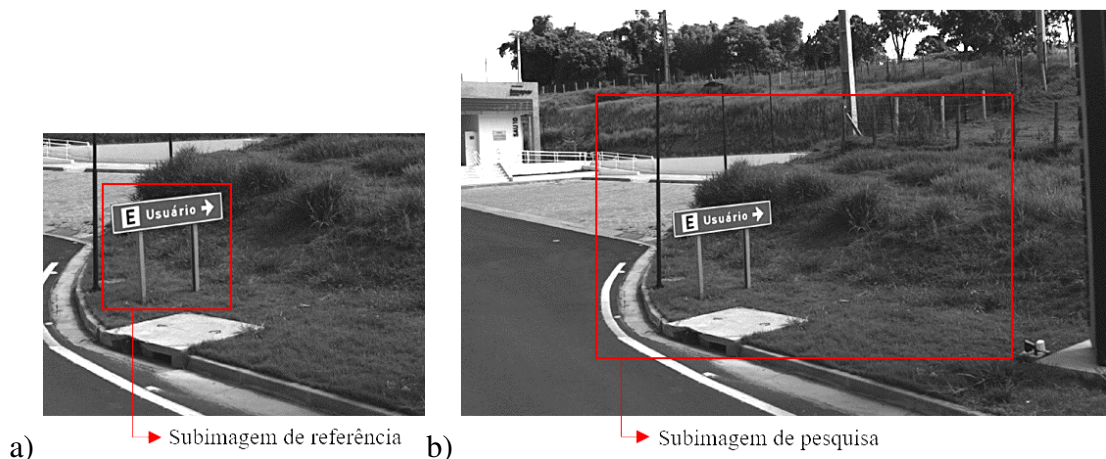
As duas categorias encontradas com maior frequência entre os algoritmos de correspondência são: correspondência baseada em área e correspondência baseada em características ou atributos. Os algoritmos de correspondência baseada em área comparam os níveis de cinza de regiões consideradas em duas imagens, e atribui a correspondência com base na similaridade entre os valores de cinza comparados. Os algoritmos de correspondência baseada em características extraem primeiro características como pontos, linhas ou regiões das imagens de forma independente, e posteriormente compara atributos destas características (MAYER *et al.*, 2013). A metodologia proposta neste trabalho utiliza um método baseado em

características, mas também verifica a possibilidade de aplicação de um método baseado em área. Os dois métodos de correspondência são apresentados na sequência.

3.2 CORRESPONDÊNCIA BASEADA EM ÁREA (*AREA-BASED MATCHING*)

O método de correspondência baseada em área identifica entidades conjugadas buscando por locais com padrões de tons de cinza similares nas imagens (MAYER *et al.*, 2013). Para isso, considera-se um par de imagens: uma na qual foi selecionado o alvo ou região de interesse e outra na qual se deseja identificar a área correspondente. A primeira imagem, onde foi selecionada a região de interesse, recebe o nome de imagem de referência. A segunda, onde será realizada a busca por correspondentes, recebe o nome de imagem de pesquisa ou imagem de busca. A busca por correspondência neste método é normalmente realizada varrendo-se uma região da imagem de pesquisa e comparando os tons de cinza com os da região (janela) de referência (GALO, 2003). Essas regiões são definidas por recortes ou subimagens das imagens originais, a fim de reduzir o custo computacional do método e também evitar resultados ambíguos (BERVEGLIERI, 2014). A Figura 8 ilustra o conceito apresentado.

Figura 8. Imagens e subimagens em destaque de referência (a) e pesquisa (b).



A comparação entre os tons de cinza é realizada com base em uma medida de similaridade, que normalmente considera funções de correlação: função erro, função erro quadrático, função quociente, função correlação cruzada, função covariância cruzada e coeficiente de correlação. Para detalhes sobre funções de correlação pode-se consultar Galo (2003). Realizado o cálculo de similaridade para a primeira posição, a janela de busca se

desloca mudando para a posição imediatamente ao lado da anterior (deslocamento de uma coluna na imagem), e realiza-se uma nova medida de similaridade. Ao término de uma linha, a janela de busca se desloca para a posição imediatamente abaixo da anterior (deslocamento de uma linha na imagem) e repete o procedimento até que se tenha percorrido toda a área de busca. Ao final, a posição assumida pela janela de busca quando foi obtido o maior nível de similaridade indica a posição do alvo correspondente.

De acordo com o modelo apresentado por Yanniris (1974)¹ *apud* Galo (2003), para todo (α, β) com $\alpha \in \{1, 2, \dots, L_p - L_r + 1\}$ e $\beta \in \{1, 2, \dots, C_p - C_r + 1\}$, a função correlação cruzada (F_{CorC}) pode ser calculada por:

$$F_{CorC}(\alpha, \beta) = \frac{1}{L_r C_r} \sum \sum g_r(i, j) g_p(i + \alpha, j + \beta) \quad (22)$$

onde,

C_r, L_r = total de colunas e linhas, respectivamente, da subimagem de referência;

$g_r(i, j), g_p(i, j)$ = valores dos níveis de cinza nas subimagens de referência e pesquisa, respectivamente;

A correspondência é dada pelo máximo valor obtido na Equação 22. Vale ressaltar ainda que $(L_p - L_r + 1)$, definido como valor limite para α , representa o número de subimagens de L_r linhas contidas na subimagem de pesquisa, e $(C_p - C_r + 1)$ analogamente para colunas, definido como valor limite para β . Estes valores permitem calcular a quantidade de combinações que serão realizadas na busca por correspondência. O pequeno teste realizado com correspondência baseada em área neste trabalho utiliza uma função que considera a covariância cruzada normalizada (*normalized cross-correlation*), isto é, uma variante da correlação cruzada que modifica a constante $L_r C_r$ presente na Equação 22. Mais informações podem ser vistas em Fonseca e Manjunath (1996).

Os resultados deste tipo de correspondência são fornecidos com precisão de um pixel, já que este é o valor do deslocamento da janela de busca. Existem diferentes maneiras de refinar estes resultados a fim de se obter uma solução com qualidade subpixel. Nenhuma solução foi aplicada aos resultados obtidos neste trabalho. Informações sobre métodos de refinamento a nível subpixel podem ser encontradas em Kraus (2007) e Mayer *et al.* (2013).

¹ YANNIRIS, Y. Improvements to the off-line epipolar correlation. 1974. 104 p. Thesis (Master of Science) - The University of New Brunswick, Fredericton, Canada, 1974.

3.3 CORRESPONDÊNCIA BASEADA EM ATRIBUTOS OU CARACTERÍSTICAS (*FEATURE-BASED MATCHING*)

Neste tipo de correspondência a comparação é realizada com características das feições selecionadas, que podem ser caracterizadas como locais e globais. Feições locais são as do tipo ponto, linha, quina, etc. E as feições globais são as estabelecidas a partir de feições locais, ou seja, polígonos (estabelecidos a partir de linhas) ou outras estruturas mais complexas (HEIPKE, 1996² *apud* OLIVEIRA, 2013). Gruen (2012) e Mayer *et al.* (2013) apresentam diferentes técnicas empregadas na solução da correspondência baseada em atributos.

Segundo Gruen (2012), a realização deste tipo de correspondência pode acontecer em duas etapas. Primeiro, extrai-se características básicas das imagens, como retas, bordas, cantos, etc, que serão utilizadas posteriormente na correspondência. Em seguida, compara-se a similaridade entre atributos destas feições, como por exemplo comprimento, curvatura, coordenadas, etc. O autor menciona ainda que as características, às vezes, podem ser mais estáveis quanto às variações radiométricas nas imagens.

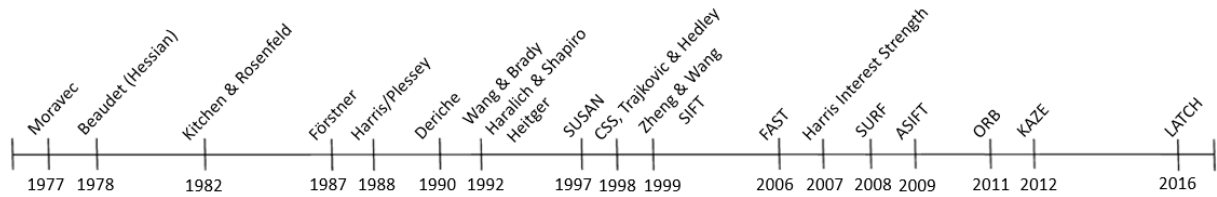
Muitos métodos de correspondência baseada em características têm sido propostos nos últimos anos (MAYER *et al.*, 2013), especialmente os chamados operadores de interesse, cujo conceito será apresentado na sequência.

3.3.1 Operadores de interesse

Detectores e descritores de características, ou simplesmente operadores de interesse, são algoritmos capazes de identificar cantos, bordas, regiões e outras feições de interesse (JAZAYERI e FRASER, 2010) que sofrem pouca variação com a mudanças de perspectiva existentes nas imagens. De acordo com Rodehorst e Koschan (2006), o primeiro operador de interesse foi desenvolvido por Hans P. Moravec, em 1977, quando foi introduzido o conceito de pontos de interesse e se concluiu que estes pontos poderiam ser utilizados na correspondência de imagens consecutivas. A linha do tempo apresentada na Figura 9 traz alguns dos operadores desenvolvidos até 2016.

² HEIPKE, C. Overview of image matching techniques. In: Workshop on the Application of digital photogrammetric workstations, 1996, Lausanne. Proceedings... Lausanne: OEEPE, 1996.

Figura 9. Linha do tempo com operadores de interesse desenvolvidos até 2016.



Fonte: Adaptado de Jazayeri e Fraser, 2010.

Além destes, existem diversos outros operadores desenvolvidos, cada um com características específicas de invariância, sendo as principais: rotação, escala, e translação. Pode se destacar na Figura 9 os operadores Förstner (FÖRSTNER e GULCH, 1987), Harris/Plessey (HARRIS e STEPHENS, 1988), SUSAN (SMITH e BRADY, 1997), SIFT (LOWE, 1999), FAST (ROSTEN e DRUMMOND, 2006), SURF (BAY *et al.*, 2008), e ASIFT (MOREL e YU, 2009), que são operadores bastante utilizados.

O estado da arte dos operadores de interesse compreende a invariância não só a rotação, escala, e translação, mas também ao fator de cisalhamento, resultado da não ortogonalidade entre os eixos da câmara (MOREL e YU, 2009); sendo assim, invariantes aos parâmetros de uma transformação afim. A respeito destes operadores pode-se mencionar as abordagens propostas por Mikolajczyk e Schmid (2004), e Morel e Yu (2009), desenvolvidas a partir dos já conhecidos operadores Harris e SIFT, respectivamente. O operador SIFT, amplamente utilizado, é apresentado na sequência em conjunto com sua versão afim-invariante, denominada ASIFT. Estes dois operadores foram utilizados na busca por correspondência no conjunto de dados analisado neste trabalho. Além deles, os operadores Harris e Förstner também são apresentados, sendo estes utilizados na extração dos pontos de partida, Etapa I da metodologia. Mais informações sobre o conceito, aplicações e exemplos de operadores de interesse podem ser encontradas em Remondino (2006), Kraus (2007), Jazayeri e Fraser (2010), Mayer *et al.* (2013).

3.3.1.1 O operador SIFT

O método proposto por Lowe (1999) trata de um operador invariante a escala e rotação, e parcialmente invariante a mudanças na iluminação da cena e mudanças de ponto de vista. De maneira geral, o algoritmo denominado SIFT, acrônimo de *Scale Invariant Feature Transform*, detecta pontos-chave usando uma abordagem multi-resolução, que permite a identificação de locais candidatos que serão examinados e, no caso de representarem um

ponto-chave, discretizados posteriormente. O procedimento realizado pelo algoritmo pode ser sintetizado em quatro etapas, que serão apresentadas na sequência. Todo o conteúdo apresentado neste item foi elaborado com base nos trabalhos de Lowe (1999, 2004) e Berveglieri (2014).

1) Detecção de pontos-extremos no espaço-escala

O espaço escala de uma imagem é definido por $L(x, y, \sigma)$, e formado pela convolução de um filtro gaussiano $G(x, y, \sigma)$ e uma imagem de entrada $I(x, y)$. A Equação 23 define o conceito.

$$L(x, y, \sigma) = G(x, y, \sigma) * I(x, y) \quad (23)$$

onde * representa a operação de convolução em x e y e $G(x, y, \sigma)$ o filtro gaussiano definido pela Equação 24.

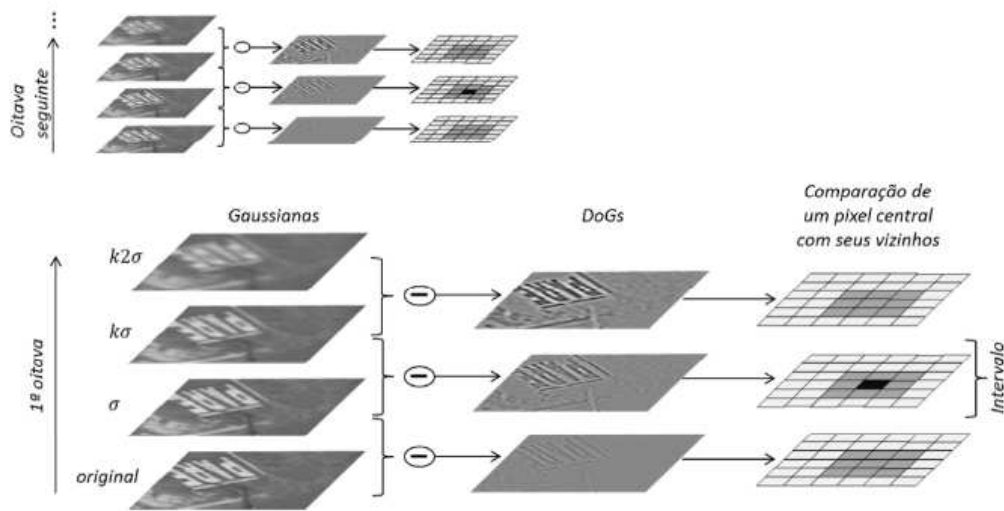
$$G(x, y, \sigma) = \frac{1}{2\pi\sigma^2} e^{-(x^2+y^2)/2\sigma^2} \quad (24)$$

Esse filtro gaussiano é aplicado à imagem com desvio-padrão inicial igual a σ e depois variando em função de um escalar k da seguinte forma: $\sigma, k\sigma, k^2\sigma, \dots$. Desta forma, o procedimento gera versões modificadas da imagem original, onde $k = 2^{1/s}$, e s é o número de intervalos inteiros entre as DoGs, *Difference of Gaussian*. As DoGs ($D(x, y, \sigma)$) são imagens obtidas pela diferença das imagens geradas pela filtragem, incluindo a original, que podem ser obtidas pela Equação 25.

$$D(x, y, \sigma) = L(x, y, k\sigma) - L(x, y, \sigma) \quad (25)$$

Esta etapa tem como finalidade a extração de pontos máximos e mínimos nas imagens. A Figura 10 ilustra o procedimento de determinação das DoGs.

Figura 10. Resumo da etapa de extração dos locais candidatos no método SIFT: versões modificadas da imagem de entrada são obtidas e também diferenças dessas versões (DoGs), comparação entre pixels para determinação de pontos candidatos é realizada nas DoGs.



Fonte: Berveglieri (2014).

Este primeiro conjunto de imagens geradas recebe o nome de primeira oitava. A segunda oitava é gerada a partir da subamostragem para a metade da imagem filtrada com desvio-padrão σ . Para cada nova oitava é considerada uma imagem com metade da resolução da oitava anterior, e o número de oitavas geradas varia com a finalidade da aplicação. Ao término da geração de oitavas, cada pixel das DoGs é comparado a seus vizinhos na própria imagem e nas imagens adjacentes, o que significa realizar 26 comparações para cada pixel. Na Figura 10 há somente uma DoG entre outras duas, o que indica um intervalo $s = 1$. Ao término das comparações, os pontos que representam máximos e mínimos locais são selecionados e classificados como pontos candidatos. Esse processo de filtragem em cascata reduz o custo computacional do método, uma vez que permite que as operações mais trabalhosas sejam realizadas apenas para pontos que já passaram por um teste inicial.

2) Localização dos pontos-chave

A segunda etapa consiste em selecionar os pontos, definidos na etapa anterior como candidatos, que realmente representam pontos-chave. O procedimento realizado permite eliminar pontos com baixo contraste, sensíveis a ruídos, e localizados ao longo de bordas. Para isto, uma superfície quadrática é ajustada aos pontos de amostragem local, permitindo a interpolação dos extremos da função. Esta abordagem usa extensão de Taylor da função do espaço-escala, como apresentado na Equação 26.

$$D(x) = D + \frac{\partial D^T}{\partial x} x + \frac{1}{2} x^T \frac{\partial^2 D}{\partial x^2} \quad (26)$$

Os extremos da função ocorrem nos pontos em que a derivada é nula. A equação anterior é então verificada nos pontos da amostra e $x = (x, y, \sigma)^T$ representa o deslocamento do ponto. Se esse deslocamento for superior a 0,5 significa que o extremo da função está mais próximo de outro ponto, e não do analisado. Neste caso, troca-se o ponto e repete-se o procedimento. Quando o resultado for inferior a 0,5, este deslocamento é acrescido à localização do ponto, resultando na posição do extremo da função. Lowe (2004) recomenda rejeitar resultados inferiores a um limiar pré-estabelecido, a fim de eliminar extremos instáveis, com baixo contraste. Realiza-se ainda um procedimento baseado no conceito da matriz *Hessian* – uma matriz de derivadas parciais – que garante a estabilidade do método. O procedimento não será abordado neste trabalho. Mais detalhes podem ser encontrados em Lowe (2004).

3) Atribuição da orientação

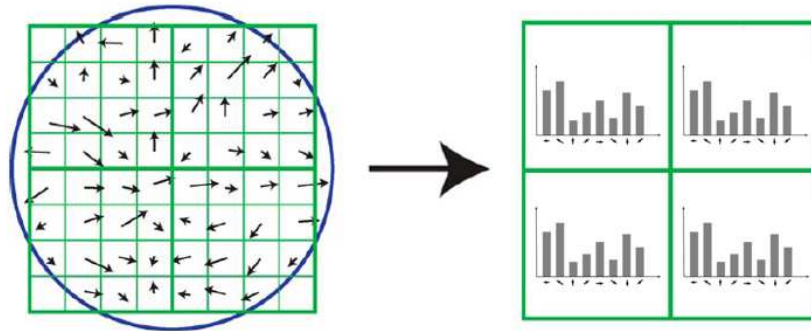
Os procedimentos descritos até o momento permitem a detecção e extração dos pontos-chave. Esta etapa inicia o processo de construção do descritor, com o pré-cálculo dos gradientes das imagens. Os procedimentos realizados neste momento permitem a invariância a rotação e escala: a atribuição da orientação realizada com base em informações locais dos pontos-chave permite a invariância parcial a rotação; e a seleção das imagens filtradas que é realizada a partir do desvio-padrão σ de cada ponto-chave permite a invariância a mudanças de escala. Para cada imagem selecionada, a magnitude do gradiente, $m(x, y)$, e a orientação, $\theta(x, y)$, são pré-calculadas. As Equações 27 e 28 permitem esses cálculos.

$$m(x, y) = \sqrt{(L(x + 1, y) - L(x - 1, y))^2 + (L(x, y + 1) - L(x, y - 1))^2} \quad (27)$$

$$\theta(x, y) = \text{tg}^{-1} \left(\frac{L(x, y + 1) - L(x, y - 1)}{L(x + 1, y) - L(x - 1, y)} \right) \quad (28)$$

A partir das informações de orientação dos gradientes, um histograma de orientação é criado para os pixels da região no entorno do ponto-chave. Cada amostra adicionada ao histograma recebe diferentes ponderações, que permitem dar maior ênfase aos gradientes de maior magnitude e evitar mudanças bruscas no descritor, além de outras vantagens apresentadas em Lowe (2004). A Figura 11 ilustra o procedimento de criação dos histogramas.

Figura 11. Magnitude e orientação dos gradientes são calculadas para os pixels no entorno do ponto-chave. Histograma de orientação é construído para cada subregião.



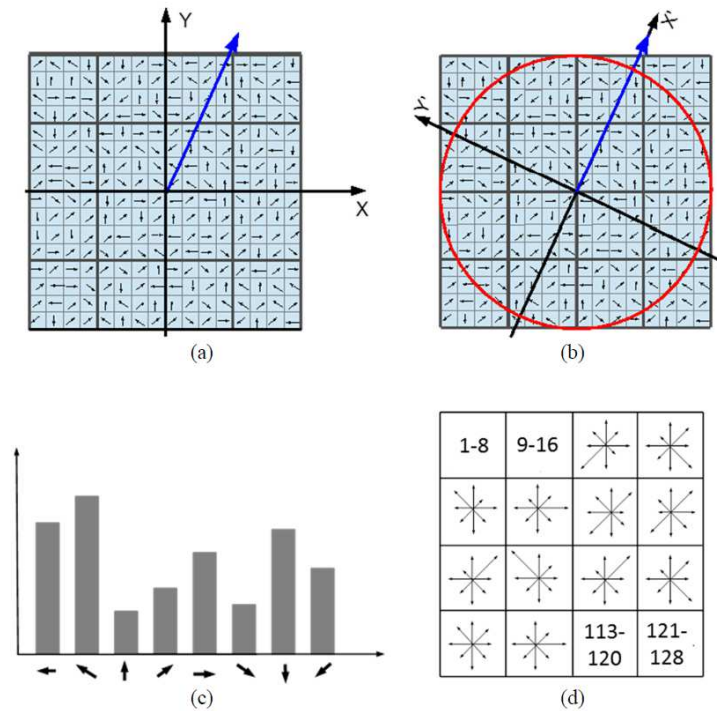
Fonte: Adaptado de Lowe (2004).

Desta forma, os picos destes histogramas representam a direção dominante dos gradientes locais. Os pontos próximos 80% ou mais dos picos também são usados para criação de novos pontos-chave com tal orientação. Por fim, uma parábola é ajustada aos três valores mais próximos de cada pico a fim de interpolar a posição do máximo com melhor acurácia.

4) Construção do descritor local

A primeira parte da construção do descritor é realizada na etapa 3, onde as orientações e magnitudes dos gradientes são amostradas ao redor da localização do ponto-chave. Esse pré-cálculo dos gradientes, para todos os níveis da pirâmide, é uma espécie de otimização do método. Depois disso, as coordenadas do descritor e as orientações dos gradientes são rotacionadas em relação à orientação do ponto-chave, o que permite que o método seja invariante a orientação. A partir disto, os histogramas de orientação são construídos atribuindo diferentes ponderações a cada amostra adicionada. O descritor representa o agrupamento das informações contidas nos histogramas, que por sua vez, representam o agrupamento das informações de gradientes de cada sub-região no entorno do ponto-chave. A Figura 12 ilustra o processo de construção do descritor local.

Figura 12. Síntese do processo de construção do descritor local no operador SIFT.



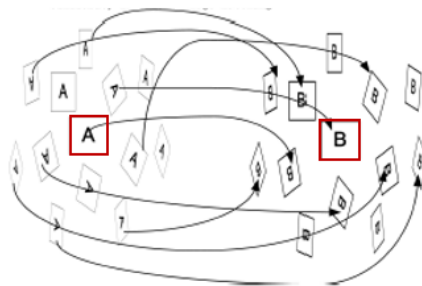
Fonte: Berveglieri (2014).

A Figura 12 apresenta uma síntese do processo de construção do descritor local: em (a) é apresentado os gradientes calculados para a região no entorno do ponto-chave e a orientação extraída para ele, representada em azul; em (b) ocorre a rotação do descritor em relação à orientação do ponto-chave; em (c) ocorre a construção dos histogramas e; em (d) ocorre o armazenamento das informações de orientação dos histogramas no vetor descritor, de modo que os elementos de 1 a 8 deste vetor armazenem informações do histograma referente à primeira sub-região, os elementos de 9 a 16 armazenem informações do histograma referente à segunda sub-região, e assim por diante. Lowe (1999, 2004) afirma que os melhores resultados foram obtidos com uma matriz de histogramas de dimensão 4×4 com 8 direções de orientação representadas em cada histograma, como apresentado no exemplo da Figura 12. Para esse caso, o descritor tem dimensão $4 \times 4 \times 8 = 128$ elementos. O vetor descritor sofre ainda modificações a fim de reduzir os efeitos da mudança de iluminação. No processo de correspondência, são estes descritores que são utilizados para fins de comparação e estabelecimento da correspondência.

3.3.1.2 O operador ASIFT

Além dos parâmetros de similaridade já modelados pelo descritor SIFT, o ASIFT modela também o fator de cisalhamento, tornando o método invariante aos parâmetros de uma transformação afim. Acrônimo de *Affine Scale Invariant Feature Transform*, o ASIFT simula as possíveis distorções da imagem causadas por alterações na orientação dos eixos da câmara e analisa a correspondência entre essas imagens por meio da aplicação do já conhecido método SIFT. O conteúdo apresentado neste item foi elaborado com base em trabalhos apresentados por Morel e Yu (2009) e Yu e Morel (2009). Basicamente, o operador gera várias versões das imagens, submetendo-as a diferentes inclinações e rotações, ou seja, diferentes perspectivas. Em seguida, estas versões, bem como as imagens originais, são comparadas entre si na busca por correspondência. A Figura 13 ilustra este conceito, com destaque para as imagens originais.

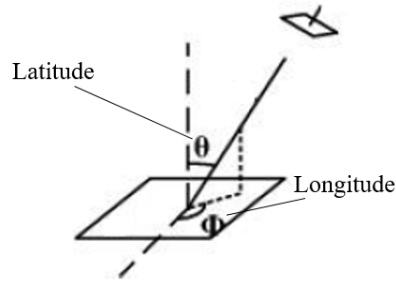
Figura 13. Visão geral do operador ASIFT: várias versões simuladas a partir das imagens originais (destacadas na figura) e comparadas entre si na busca por correspondência.



Fonte: Adaptada de Morel e Yu (2009).

O procedimento realizado pelo algoritmo se inicia com a transformação das imagens, onde cada uma delas é transformada pela simulação de todas as possíveis distorções afins causadas pela mudança de orientação nos eixos da câmara, a partir de uma vista frontal. Estas simulações acontecem em posições específicas, dadas por latitude (θ) e longitude (ϕ), sendo latitude (θ) o ângulo entre o eixo óptico da câmara e a normal ao plano da imagem, e longitude (ϕ), neste caso, sendo o ângulo que o plano que contém a normal e o eixo óptico da câmara faz com o plano vertical. A Figura 14 ilustra estes conceitos.

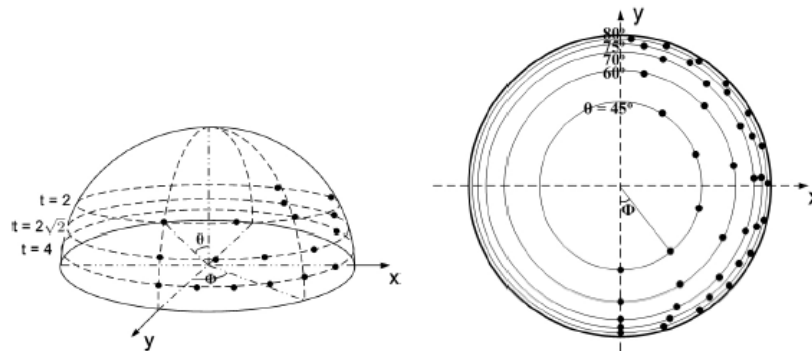
Figura 14. Representação dos conceitos de latitude (θ) e longitude (ϕ) empregados no ASIFT.



Fonte: Adaptado de Yu e Morel (2009).

Em cada uma das simulações realizadas a imagem é submetida a inclinações e rotações, considerando intervalos de amostragem para os parâmetros posicionais (θ , ϕ) em função do parâmetro de inclinação (t). Cada inclinação t representa uma subamostragem da imagem na direção x em t vezes. De acordo com os autores, o intervalo $[0, \pi]$ é suficiente para cobrir todas as possíveis distorções afins da imagem. As latitudes simuladas (θ) correspondem então a inclinações $t = 1, a, a^2, \dots, a^n$, com $a > 1$. Os autores indicam $a = \sqrt{2}$ como uma boa aproximação para o método. Testes realizados pelos autores mostram que a inclinação t é dada por $t = 1/\cos\theta$. Sabendo disto, o intervalo de amostragem considerado para a latitude mostra que todas as imagens simuladas na latitude $\theta = 0^\circ$ serão submetidas a inclinação $t = 1$; depois, todas as imagens simuladas na latitude $\theta = 45^\circ$ serão submetidas a inclinação $t = \sqrt{2}$; e assim sucessivamente até um valor de t máximo (t_{\max}). As longitudes simuladas (ϕ) seguem para cada inclinação t uma série aritmética dada por $0, b/t, \dots, kb/t$, onde k é o último inteiro tal que $kb/t < 180^\circ$; e $b = 72^\circ$ de acordo com os autores. Isto significa que as imagens simuladas na latitude $\theta = 0^\circ$, que receberam inclinação $t = 1$, serão amostradas nas longitudes $\phi = 0^\circ, 72^\circ, 144^\circ$; as imagens simuladas na latitude $\theta = 45^\circ$, que receberam inclinação $t = \sqrt{2}$, serão amostradas nas longitudes $\phi = 0^\circ, 51^\circ, 102^\circ, 153^\circ$; e assim sucessivamente até que todo o intervalo $[0, \pi]$ seja coberto. A Figura 15 ilustra a ideia de amostragem dos parâmetros posicionais.

Figura 15. Amostragem dos parâmetros latitude (θ) e longitude (ϕ) no método ASIFT. Apenas as inclinações $t = 2, 2\sqrt{2}, 4$ foram representadas. Intervalos de variação $\Delta t = \sqrt{2}$ e $\Delta\phi = 72^\circ/b$.



Fonte: Morel e Yu (2009).

A partir disto, a busca por correspondência é realizada nas imagens simuladas, utilizando o método SIFT.

3.3.1.3 O operador de Harris

O método de Harris utiliza uma combinação de detectores de canto e borda, desenvolvida a partir do operador de Moravec (1977) e baseada na função de autocorrelação local. Todo o conteúdo apresentado neste item foi elaborado a partir do trabalho de Harris e Stephens (1988), autores do método. O detector de Moravec se baseia na medida de variância direcional da imagem, considerando, para isto, uma janela local e determinando as mudanças médias de intensidade na imagem a partir do deslocamento desta janela em, inicialmente, quatro diferentes direções. A respeito da análise da mudança mencionada anteriormente, três situações podem ser consideradas:

- a) Se o recorte da imagem na janela representa uma região homogênea, nenhuma mudança será muito significativa, isto é, todas as direções representarão apenas uma pequena alteração;
- b) Se o recorte da imagem na janela representar uma borda, a mudança no sentido da borda será pequena, mas esta mesma mudança no sentido transversal à borda será grande;
- c) Se o recorte da imagem na janela representar um ponto ou canto, todas as direções representarão grandes mudanças.

Assim, o detector procura um máximo local dentro de uma mudança (E) mínima que atenda a um limiar. Sendo I a imagem de intensidades e w a janela considerada nesta imagem, a mudança E produzida por um deslocamento (x,y) é dada por:

$$E_{x,y} = \sum_{u,v} w_{u,v} [I_{x+u,y+v} - I_{u,v}]^2 \quad (29)$$

Sendo este cálculo realizado para quatro direções, intercaladas em 45°, a Equação 29 determina para cada ponto da imagem, o valor mínimo do operador nestas quatro direções. A partir destes valores, os máximos locais de $E_{x,y}$ são encontrados e os pixels com valores superiores a um limiar pré-estabelecido são classificados como cantos. Para considerar as pequenas mudanças existentes entre os intervalos de 45°, uma expansão pode ser realizada na Equação 29, de modo que:

$$E_{x,y} = \sum_{u,v} w_{u,v} [xX + yY + O(x^2, y^2)]^2$$

onde os primeiros gradientes são aproximados por:

$$X = I \otimes (-1,0,1) \approx \partial I / \partial x$$

$$Y = I \otimes (-1,0,1)^T \approx \partial I / \partial y$$

Assim, para pequenos deslocamentos, a mudança E pode ser escrita por:

$$E_{x,y} = Ax^2 + 2Cxy + By^2$$

onde,

$$A = X^2 \otimes w$$

$$B = Y^2 \otimes w$$

$$C = (XY) \otimes w$$

Os autores afirmam ainda que a resposta do operador é ruidosa, devido à janela considerada no método ser binária e retangular. Isto pode ser solucionado, por exemplo, com a utilização de uma janela circular de suavização gaussiana:

$$w_{u,v} = \exp -(u^2 + v^2) / 2\sigma^2 \quad (30)$$

Pela consideração b) feita anteriormente, pode-se perceber que o operador permite a identificação de bordas, sendo necessário apenas alterar a medida de canto para utilizar a variação de E na direção do deslocamento. Assim, a mudança E para um pequeno deslocamento (x, y) pode ser escrita por:

$$E(x, y) = (x, y)M(x, y)^T \quad (31)$$

onde M é uma matriz simétrica 2x2:

$$M = \begin{bmatrix} A & C \\ C & B \end{bmatrix}$$

A partir disto, a mudança E pode ser relacionada à função de autocorrelação local, com M descrevendo sua forma na origem. Uma resposta R (Equação 32) é então calculada para cada janela, com os autovalores λ_1 e λ_2 e a constante do fator de sensibilidade de Harris k :

$$R = \text{Det}(M) - k \text{Tr}(M)^2 \quad (32)$$

onde,

$$\text{Tr}(M) = \alpha + \beta = A + B$$

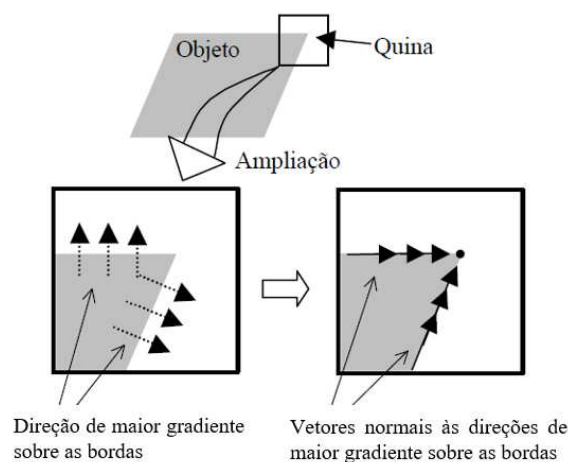
$$\text{Det}(M) = \alpha\beta = AB - C^2$$

A classificação é realizada sabendo-se que R é positivo em regiões de canto, negativo em regiões de borda e pequeno em regiões planas, sendo estas últimas especificadas pelo $\text{Tr}(M)$, que fica abaixo de um limite selecionado. Mais considerações a respeito da classificação das janelas devem ser vistas em Harris e Stephens (1988).

3.3.1.4 O operador de Förstner

A identificação de cantos, ou quinas, pode ser feita a partir do princípio de que uma quina pode ser formada pela interseção de dois segmentos lineares, como na interseção dos vetores normais às direções de maior gradiente das bordas (GALO e TOZZI, 2002). A Figura 16 ilustra esta afirmação.

Figura 16. Princípio da identificação de quinas por meio da interseção de segmentos de retas.



Adaptado de Galo e Tozzi (2002).

Partindo deste conceito, Galo e Tozzi (2002) mostram que a posição de um canto dentro de uma janela pode ser estimada por um sistema linear em função dos componentes do

gradiente ao longo das linhas e colunas, resultando na Equação 33, utilizada no operador de Förstner, e cujo desenvolvimento também é apresentado em Förstner e Gulch (1987):

$$\begin{bmatrix} L_0 \\ C_0 \end{bmatrix} = \begin{bmatrix} \sum g_{Li}^2 & \sum g_{Li}g_{Ci} \\ \sum g_{Li}g_{Ci} & \sum g_{Ci}^2 \end{bmatrix}^{-1} \begin{bmatrix} \sum (L_i g_{Li}^2 + C_i g_{Li}g_{Ci}) \\ \sum (C_i g_{Ci}^2 + L_i g_{Li}g_{Ci}) \end{bmatrix} \quad (33)$$

onde,

(C_0, L_0) = posição coluna e linha do canto/quina estimada pelo operador com qualidade subpixel;

(C_i, L_i) = posição coluna e linha de cada um dos pixels da janela;

(g_{Ci}, g_{Li}) = componentes do gradiente ao longo das colunas e linhas, respectivamente, da janela.

A matriz 2x2 presente na Equação 33 pode ser vista como a matriz variância-covariância do ajustamento, podendo ser utilizada na obtenção da elipse dos erros associada à quina (C_0, L_0) . Análises sobre esta elipse permite a discriminação de pontos com diferentes características (FÖRSTNER, 1986; GALO e TOZZI, 2002).

Para mais detalhes sobre este operador recomenda-se os trabalhos de Förstner (1986), Förstner e Gulch (1987) e Galo e Tozzi (2002).

3.4 PROBLEMAS NA CORRESPONDÊNCIA DE IMAGENS

Diversos fatores podem dificultar ou até mesmo impedir a correspondência de feições homólogas, como colocado por Heipke (1997). Mayer *et al.* (2013) expõe os principais fatores a serem considerados:

Diferenças de perspectivas entre as imagens: As fotografias tomadas a partir de diferentes pontos de vista geram imagens sob diferentes perspectivas, com geometrias diferentes. Estas diferenças aumentam de acordo com o ângulo entre as direções de visada e também com a irregularidade da superfície imageada. A fim de simplificar os cálculos, diversos algoritmos assumem um modelo de superfície suave para o processo de correspondência. Quando esta associação é violada, como acontece em paisagens urbanas com relevo muito acidentado, a busca por pontos homólogos pode falhar ou gerar resíduos nas posições correspondentes estimadas.

Oclusão: Em casos extremos, a diferença de perspectiva entre as imagens pode levar à oclusão de partes da cena em uma ou mais imagens, tornando a correspondência nestas partes impossível de se estabelecer. A Figura 17 ilustra o problema de oclusão, destacando partes da imagem que não aparecem devido à variação de perspectiva.

Figura 17. Problema de oclusão no processo de correspondência: feições só aparecem em uma das imagens devido à variação de perspectiva (destaque).



Fonte: Adaptado de Nielsen (2004).

Reflectâncias diferentes: A maioria das superfícies refletoras são anisotrópicas, o que significa que a reflexão da luz vai variar de acordo com a direção em que ela é observada. Nestes casos, perspectivas diferentes irão resultar em variações de tonalidade nas imagens, fazendo com que os níveis de cinza de áreas correspondentes possam diferir. Algoritmos de correspondência que não modelam essas diferenças poderão rejeitar correspondências corretas e encontrar correspondências incorretas.

Iluminação diferente: Imagens que foram tomadas sob condições de iluminação diferentes irão, com certeza, resultar no mesmo problema com níveis de cinza diferentes. Em particular, esse problema pode ocorrer na correspondência de imagens aéreas adquiridas em datas ou até mesmo horários diferentes.

Padrões repetitivos: Algumas cenas contêm padrões repetitivos, de modo que um certo padrão em uma imagem irá corresponder a vários padrões similares na outra imagem, gerando confusão no processo de correspondência (correspondências ambíguas). A Figura 18 ilustra este problema, destacando um dos padrões repetitivos da imagem.

Figura 18. Problema de ambiguidade no processo de correspondência causado pela presença de padrões repetitivos (em destaque) na imagem.



Baixa relação sinal-ruído: Algoritmos de correspondência tanto baseados em área como baseados em atributos apresentam melhores resultados quando analisam imagens com alto contraste. Uma baixa relação sinal-ruído irá reduzir a similaridade entre fragmentos correspondentes nas imagens e, conseqüentemente, dificultar a extração de características para o processo de correspondência.

Além dos fatores apresentados, deve-se ressaltar que a busca por correspondência pode ser um processo lento, uma vez que é baseado no princípio de varredura das imagens. Devido à essas limitações, restrições devem ser consideradas como forma de otimização do método. Uma forma de restringir a busca por correspondência é acrescentar características à feição utilizada como primitiva. Quanto mais específica for a entidade selecionada, menos ambigüidade existirá. Esta restrição pode ser realizada por meio dos métodos de correspondência baseados em descritores (vetores de características), como o SIFT e o ASIFT, apresentados no item 3.3.1. Outra restrição que pode ser considerada é a de redução do espaço de busca, que consiste na delimitação de um espaço menor para a busca por correspondência a partir de um conhecimento prévio da localização do ponto homólogo. Para exemplos de estratégias consideradas na redução do espaço de busca pode-se consultar Gruen (2012) e Mayer *et al.* (2013).

CAPÍTULO IV

METODOLOGIA PARA CORRESPONDÊNCIA DE IMAGENS AÉREAS E TERRESTRES

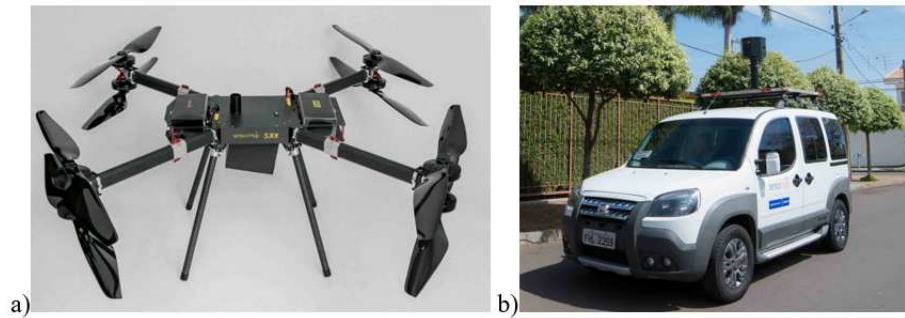
4.1 SÍNTESE DA PROPOSTA

Diante do problema de correspondência entre imagens aéreas e terrestres, ainda não solucionado de maneira satisfatória, o objetivo geral deste trabalho é o desenvolvimento de uma metodologia capaz de realizar esta correspondência de forma automática ou semiautomática. Com base no estudo de Gerke *et al.* (2016), bem como em experimentos realizados pela autora, notou-se que a correspondência a partir das imagens originais pode não representar a melhor solução para o problema devido à grande variação de perspectiva presente nestas imagens. Assim, a ideia neste trabalho é realizar uma mudança de perspectiva prévia nas imagens de modo que apresentem aspectos geométricos semelhantes e, a partir disto, aplicar uma técnica de correspondência. Uma vez que as imagens tenham geometrias semelhantes, a busca por correspondência se torna mais simples e as chances de que seja bem-sucedida são maiores.

4.2 RECURSOS MATERIAIS

As imagens utilizadas nos experimentos foram fornecidas pela empresa Sensormap Geotecnologia, que realizou a aquisição com sistemas aéreo e terrestre de fabricação própria. O sistema aéreo é composto por um VANT, denominado SX8; sistema GNSS de dupla frequência acoplado, modelo NovAtel OEMV-2; e câmara digital Sony NEX-7, com resolução de 24 MP. O sistema terrestre, denominado Sensormap Mapeamento Móvel (SMM), é composto por um veículo terrestre equipado com um sistema omnidirecional de seis câmaras JAI BB-500GE, resolução de 5 MP, e sistema GNSS/INS acoplado. O receptor GNSS, neste caso, é um NovAtel ProPak-V3 e a unidade de medida inercial é uma Honeywell HG1700. A Figura 19 apresenta os sistemas aéreo (a) e terrestre (b) utilizados na aquisição dos dados.

Figura 19. Sistemas aéreo (a) e terrestre (b) utilizados na aquisição dos dados.



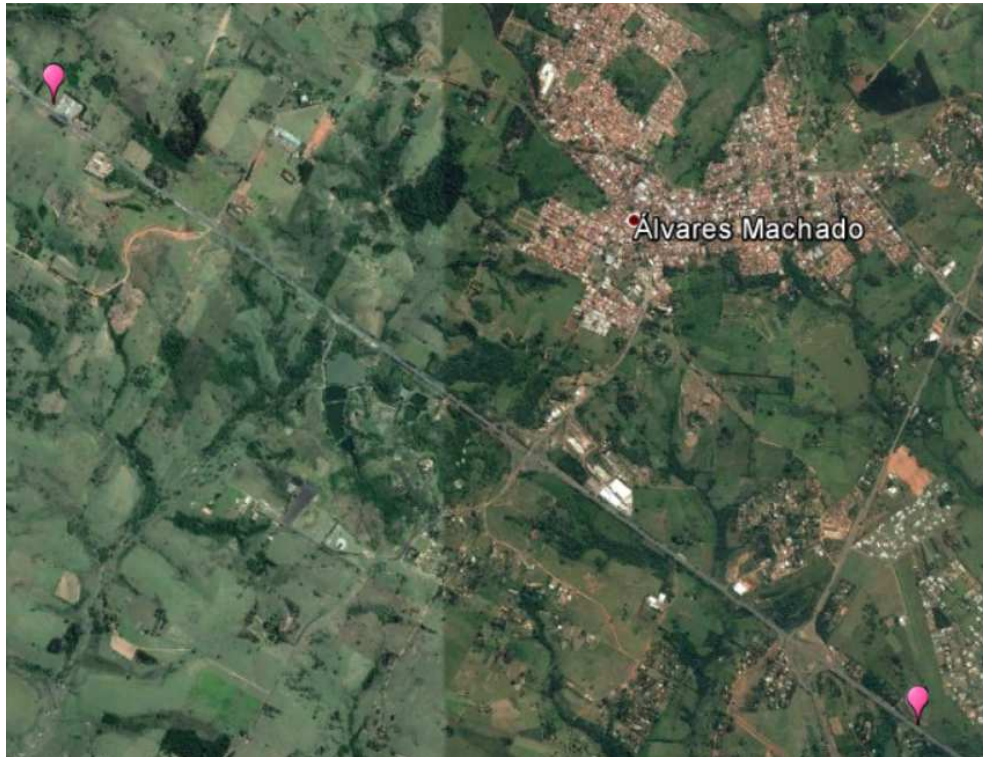
Fonte: Sensormap Geotecnologia (2015).

Juntamente com as imagens a empresa forneceu os arquivos de orientação exterior das mesmas, provenientes no caso aéreo de fototriangulação e no caso terrestre de georreferenciamento direto, bem como o arquivo de MDT gerado a partir das imagens aéreas. A empresa forneceu também os arquivos de orientação interior das câmaras utilizadas. A aquisição dos dados foi realizada em março de 2015, de forma não simultânea – houve um intervalo de um dia entre as aquisições terrestre e aérea.

4.3 ÁREA DE ESTUDO

A área de estudo compreende um trecho da rodovia Raposo Tavares (SP-270), localizado próximo ao município de Alvares Machado - SP. A Figura 20 apresenta a localização da área de estudo, com destaque para os extremos do trecho rodoviário mapeado.

Figura 20. Localização da área de estudo, destaque para os extremos do trecho rodoviário contemplado no mapeamento.



Fonte: Google Earth, 2017.

As coordenadas dos cantos que delimitam a área de estudo estão apresentadas no Quadro 1.

Quadro 1. Coordenadas dos cantos que delimitam a área de estudo.

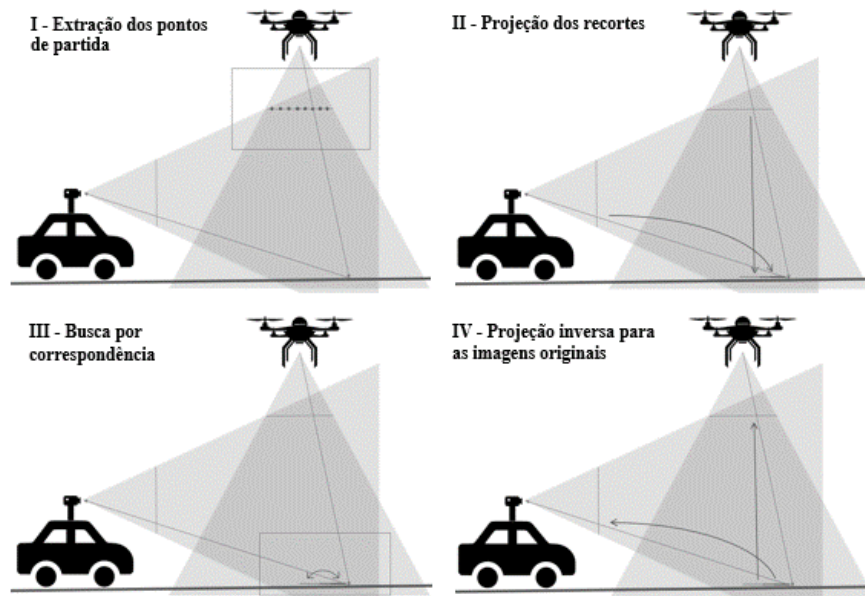
Canto	Coordenadas UTM Fuso 22K		Coordenadas WGS84	
	E (m)	N (m)	Latitude	Longitude
Sup. Esq.	446.583,00	7.559.851,00	22°03'55,20"S	51°31'3,84"O
Inf. dir.	453.510,00	7.554.691,00	22°06'43,73"S	51°27'2,68"O

4.4 METODOLOGIA

De maneira geral, a metodologia proposta parte de pontos de interesse bem definidos, extraídos da imagem aérea, cujas posições são projetadas para a imagem terrestre. Estes pontos são chamados de pontos de partida, pois iniciam o processo realizado na metodologia. Regiões consideradas no entorno destes pontos definem recortes que são projetados para um plano horizontal, tornando a perspectiva aérea em ambos os casos. A busca por correspondência é então realizada nestes recortes projetados, e o resultado é

transformado de volta para as imagens originais. Desta forma, a metodologia pode ser dividida em quatro principais etapas (Figura 21): (I) Extração dos pontos de partida, (II) Projeção dos recortes, (III) Busca por correspondência, e (IV) Projeção inversa para as imagens originais.

Figura 21. Etapas da metodologia proposta.



As etapas II e IV foram implementadas pela autora em linguagem C/C++, utilizando o *software* CodeBlocks, e as demais foram realizadas utilizando códigos disponibilizados por terceiros. Na sequência é apresentado o detalhamento de cada uma das etapas.

4.4.1 ETAPA I – Extração dos pontos de partida

Inicialmente é realizada a extração dos pontos que servirão como partida para a metodologia implementada. Estes pontos – aqui chamados de pontos de partida – são pontos de interesse extraídos na imagem aérea e, posteriormente, projetados para as imagens terrestres, definindo assim as origens dos recortes considerados na etapa de projeção. Existem diversos operadores de interesse na literatura cuja finalidade é a extração de pontos, bem como de cantos, bordas e outras feições distinguíveis. Além das feições extraídas, os operadores podem variar de acordo com sua característica de invariância, podendo extrair

feições invariantes à mudança de escala, rotação e até perspectiva. Para a aplicação neste primeiro estudo foram testados dois detectores de cantos: o operador de Harris (HARRIS e STEPHENS, 1988) e o de Förstner (FÖRSTNER, 1986), referidos no texto apenas pelo nome dos autores. Além de cantos, estes operadores permitem a detecção de pontos, bordas e, no caso do Förstner, de círculos também. Optou-se por utilizar cantos pela restrição que esta feição permite na cena, evitando a identificação de pontos em regiões de mata e favorecendo esta identificação no percurso rodoviário, cenário em que se espera uma melhor identificação das correspondências. As características de invariância dos operadores, neste momento, não serão discutidas.

Para a extração com o operador Harris utilizou-se uma implementação realizada por Kovesi (2002), que realiza, inicialmente, uma operação de realce de bordas na imagem, seguida de uma suavização gaussiana e de uma dilatação morfológica, a partir da qual são extraídos os máximos locais. Por último, são encontrados os cantos na imagem que correspondam aos máximos extraídos na imagem dilatada e sejam maiores que um limiar pré-estabelecido. A implementação se baseia em três parâmetros principais que devem ser fornecidos: *sigma*, *thresh* – que faz menção à palavra *threshold*, em português “limiar”, e *radius* – em português “raio”. O primeiro, *sigma*, representa o desvio-padrão considerado na suavização gaussiana. O segundo, *thresh*, refere-se ao limiar considerado na seleção dos cantos. O terceiro, *radius*, representa o raio da região considerada na supressão não-máxima, ou seja, define o tamanho da máscara utilizada na dilatação morfológica da imagem. Kovesi (2002) sugere os valores de 1 a 3 para *sigma* e *radius*, e para o *thresh* valores próximos de 1000.

Para a extração com o operador Förstner utilizou-se uma implementação realizada por Luxen (2005), a respeito da qual os principais parâmetros considerados estão apresentados no Quadro 2.

Quadro 2. Parâmetros considerados na implementação do operador Förstner, realizada por Luxen (2005), e seus respectivos significados.

SIGMA_N	ruído como desvio padrão padrão: SIGMA_N = 2,5
DERIVATIVE_FILTER	filtro gaussiano utilizado para gradiente: 'gaussian1d' (unidimensional) ou 'gaussian2d' (bidimensional) padrão: 'gaussian2d'
INTEGRATION_FILTER	filtro de convolução para integração: 'gaussian' - filtro gaussiano bidimensional, ou 'box_filt' - filtro de caixa bidimensional padrão: 'box_filt'
SIGMA_DERIVATIVE_FILTER	tamanho do filtro de gradiente (sigma) padrão: SIGMA_DIFF=0.7
SIGMA_INTEGRATION_FILTER	tamanho do filtro de integração (sigma) padrão: SIGMA_INT=1.41*SIGMA_DIFF
VISUALIZATION	exibir ou não os resultados na figura: 'on' ou 'off'

Como os experimentos são realizados combinando uma única imagem aérea a diversas imagens terrestres correspondentes, realizou-se a extração dos pontos de partida em toda a imagem aérea, discriminando uma área de interesse para cada imagem terrestre analisada. Esta discriminação foi feita observando, visualmente, as coordenadas aproximadas de início e fim de cada imagem terrestre na aérea, e inserindo estes valores como uma primeira restrição imposta no programa.

O resultado desta primeira parte da metodologia é a lista de coordenadas dos cantos extraídos na imagem aérea. Esta lista é lida no programa de forma que a cada novo ponto toda a metodologia é realizada. Ao fim do processamento para aquele par de coordenadas, o próximo é lido e todo o processo repetido, até que se chegue ao fim do arquivo de coordenadas.

4.4.2 ETAPA II – Projecção dos recortes

Dados de entrada no programa:

- Par de imagens (aérea e terrestre),
- POE das imagens,
- POI das câmaras,
- MDT da região,
- Coordenadas dos pontos de partida (obtidas na etapa anterior).

Após a leitura dos dados de entrada, cada ponto de interesse na imagem aérea tem sua posição determinada no terreno e então na imagem terrestre. A primeira determinação é realizada utilizando o conceito de monorrestituição, que permite a obtenção da posição de um ponto no terreno de acordo com a modelagem apresentada no item 2.2. Essa posição é então projetada para a imagem terrestre utilizando o modelo de colinearidade na forma direta (Equações 12).

A partir da posição do ponto nas duas imagens, são definidas regiões no entorno desses pontos que serão os recortes a serem utilizados na etapa de projeção. Como os pontos são extraídos inicialmente na imagem aérea, a posição de cada um deles neste momento não considera nenhum desvio-padrão, o que permite que a janela, neste caso, seja definida de maneira arbitrária. Testes mostraram que o valor 1000x1000 pixels, com centro no respectivo ponto de interesse, representa uma boa configuração para o recorte, sendo este valor adotado como fixo para todos os casos aéreos. No caso terrestre, o tamanho do recorte é definido com base no desvio-padrão da posição do ponto calculada na imagem, sendo estes desvios obtidos por propagação de covariâncias, cujo conceito é apresentado na sequência.

Considerando o modelo não linear $Y = F(X)$, a lei de propagação das covariâncias é dada por (GEMAEL *et al.*, 2015):

$$\Sigma_Y = D \Sigma_X D^T \quad (34)$$

onde,

D = matriz das derivadas parciais do modelo $\left(\frac{\partial F}{\partial X}\right)$;

Σ_X = matriz variância-covariância (MVC) dos parâmetros;

Σ_Y = MVC dos parâmetros incógnitos.

Para se obter o desvio-padrão da posição calculada na imagem terrestre, duas relações precisam ser consideradas na propagação de covariâncias: a relação imagem aérea – terreno, que estima a precisão da posição do ponto no terreno, e que, por sua vez, influencia diretamente na posição do ponto projetada para a imagem terrestre, e a relação terreno – imagem terrestre. Os erros envolvidos no processo de OI, neste caso, são desprezíveis. A primeira relação é feita utilizando o conceito de monorrestituição (item 2.2), que utiliza o modelo de colinearidade na forma inversa (Equações 13) e requer o conhecimento de um MDT da região, sendo os POE da imagem aérea os únicos parâmetros envolvidos no modelo. Desta forma, a matriz das derivadas parciais (D) e a matriz variância-covariância (MVC) dos parâmetros (Σ_X) para esta primeira relação são dadas, respectivamente, por:

$$D = \begin{bmatrix} \frac{\partial \omega}{\partial X_p} & \frac{\partial \varphi}{\partial X_p} & \frac{\partial \kappa}{\partial X_p} & \frac{\partial X_0}{\partial X_p} & \frac{\partial Y_0}{\partial X_p} & \frac{\partial Z_0}{\partial X_p} \\ \frac{\partial \omega}{\partial Y_p} & \frac{\partial \varphi}{\partial Y_p} & \frac{\partial \kappa}{\partial Y_p} & \frac{\partial X_0}{\partial Y_p} & \frac{\partial Y_0}{\partial Y_p} & \frac{\partial Z_0}{\partial Y_p} \end{bmatrix}, e$$

$$\Sigma_X = \begin{bmatrix} \sigma_\omega^2 & \sigma_{\omega\varphi} & \sigma_{\omega\kappa} & \sigma_{\omega X_0} & \sigma_{\omega Y_0} & \sigma_{\omega Z_0} \\ \sigma_{\omega\varphi} & \sigma_\varphi^2 & \sigma_{\varphi\kappa} & \sigma_{\varphi X_0} & \sigma_{\varphi Y_0} & \sigma_{\varphi Z_0} \\ \sigma_{\omega\kappa} & \sigma_{\varphi\kappa} & \sigma_\kappa^2 & \sigma_{\kappa X_0} & \sigma_{\kappa Y_0} & \sigma_{\kappa Z_0} \\ \sigma_{\omega X_0} & \sigma_{\varphi X_0} & \sigma_{\kappa X_0} & \sigma_{X_0}^2 & \sigma_{X_0 Y_0} & \sigma_{X_0 Z_0} \\ \sigma_{\omega Y_0} & \sigma_{\varphi Y_0} & \sigma_{\kappa Y_0} & \sigma_{X_0 Y_0} & \sigma_{Y_0}^2 & \sigma_{Y_0 Z_0} \\ \sigma_{\omega Z_0} & \sigma_{\varphi Z_0} & \sigma_{\kappa Z_0} & \sigma_{X_0 Z_0} & \sigma_{Y_0 Z_0} & \sigma_{Z_0}^2 \end{bmatrix},$$

onde,

σ_i^2 = variância do parâmetro i;

σ_{ij} = covariância entre os parâmetros i, j.

As derivadas parciais que compõem a matriz D foram calculadas utilizando o *software* Matlab. Os valores utilizados como desvios-padrão dos POE aéreos foram 20 cm para os desvios posicionais (X_0 , Y_0 , Z_0) – valor extraído do manual do equipamento – e 1' para os desvios angulares (ω , φ , κ) – valor definido empiricamente, considerando que estas imagens foram previamente trianguladas. A partir disto, a MVC dos parâmetros incógnitos (Σ_Y) para esta primeira relação é dada por:

$$\Sigma_Y = \begin{bmatrix} \sigma_{X_p}^2 & \sigma_{X_p Y_p} \\ \sigma_{X_p Y_p} & \sigma_{Y_p}^2 \end{bmatrix}$$

A segunda relação considerada na propagação de covariâncias é a relação terreno – imagem terrestre, que utiliza o modelo de colinearidade na forma direta (Equações 12). Como já mencionado, as coordenadas X e Y do terreno são determinadas pelo modelo de colinearidade na forma inversa, e a coordenada altimétrica Z é determinada por interpolação, a partir do MDT da região, e foi tomada como constante no processo de propagação de covariâncias. Desta forma, a matriz das derivadas parciais (D) e a MVC dos parâmetros (Σ_X) para esta segunda relação são dadas, respectivamente, por:

$$D = \begin{bmatrix} \frac{\partial \omega}{\partial x_p} & \frac{\partial \varphi}{\partial x_p} & \frac{\partial \kappa}{\partial x_p} & \frac{\partial X_0}{\partial x_p} & \frac{\partial Y_0}{\partial x_p} & \frac{\partial Z_0}{\partial x_p} & \frac{\partial X_p}{\partial x_p} & \frac{\partial Y_p}{\partial x_p} & \frac{\partial Z_p}{\partial x_p} \\ \frac{\partial \omega}{\partial y_p} & \frac{\partial \varphi}{\partial y_p} & \frac{\partial \kappa}{\partial y_p} & \frac{\partial X_0}{\partial y_p} & \frac{\partial Y_0}{\partial y_p} & \frac{\partial Z_0}{\partial y_p} & \frac{\partial X_p}{\partial y_p} & \frac{\partial Y_p}{\partial y_p} & \frac{\partial Z_p}{\partial y_p} \end{bmatrix}, e$$

$$\Sigma_X = \begin{bmatrix} \sigma_\omega^2 & \sigma_{\omega\varphi} & \sigma_{\omega\kappa} & \sigma_{\omega X_0} & \sigma_{\omega Y_0} & \sigma_{\omega Z_0} & \sigma_{\omega X_p} & \sigma_{\omega Y_p} & \sigma_{\omega Z_p} \\ \sigma_{\omega\varphi} & \sigma_\varphi^2 & \sigma_{\varphi\kappa} & \sigma_{\varphi X_0} & \sigma_{\varphi Y_0} & \sigma_{\varphi Z_0} & \sigma_{\varphi X_p} & \sigma_{\varphi Y_p} & \sigma_{\varphi Z_p} \\ \sigma_{\omega\kappa} & \sigma_{\varphi\kappa} & \sigma_\kappa^2 & \sigma_{\kappa X_0} & \sigma_{\kappa Y_0} & \sigma_{\kappa Z_0} & \sigma_{\kappa X_p} & \sigma_{\kappa Y_p} & \sigma_{\kappa Z_p} \\ \sigma_{\omega X_0} & \sigma_{\varphi X_0} & \sigma_{\kappa X_0} & \sigma_{X_0}^2 & \sigma_{X_0 Y_0} & \sigma_{X_0 Z_0} & \sigma_{X_0 X_p} & \sigma_{X_0 Y_p} & \sigma_{X_0 Z_p} \\ \sigma_{\omega Y_0} & \sigma_{\varphi Y_0} & \sigma_{\kappa Y_0} & \sigma_{X_0 Y_0} & \sigma_{Y_0}^2 & \sigma_{Y_0 Z_0} & \sigma_{Y_0 X_p} & \sigma_{Y_0 Y_p} & \sigma_{Y_0 Z_p} \\ \sigma_{\omega Z_0} & \sigma_{\varphi Z_0} & \sigma_{\kappa Z_0} & \sigma_{X_0 Z_0} & \sigma_{Y_0 Z_0} & \sigma_{Z_0}^2 & \sigma_{Z_0 X_p} & \sigma_{Z_0 Y_p} & \sigma_{Z_0 Z_p} \\ \sigma_{\omega X_p} & \sigma_{\varphi X_p} & \sigma_{\kappa X_p} & \sigma_{X_0 X_p} & \sigma_{Y_0 X_p} & \sigma_{Z_0 X_p} & \sigma_{X_p}^2 & \sigma_{X_p Y_p} & \sigma_{X_p Z_p} \\ \sigma_{\omega Y_p} & \sigma_{\varphi Y_p} & \sigma_{\kappa Y_p} & \sigma_{X_0 Y_p} & \sigma_{Y_0 Y_p} & \sigma_{Z_0 Y_p} & \sigma_{X_p Y_p} & \sigma_{Y_p}^2 & \sigma_{Y_p Z_p} \\ \sigma_{\omega Z_p} & \sigma_{\varphi Z_p} & \sigma_{\kappa Z_p} & \sigma_{X_0 Z_p} & \sigma_{Y_0 Z_p} & \sigma_{Z_0 Z_p} & \sigma_{X_p Z_p} & \sigma_{Y_p Z_p} & \sigma_{Z_p}^2 \end{bmatrix}$$

As derivadas parciais que compõem a matriz D foram calculadas com auxílio do Matlab. Os valores utilizados para os desvios-padrão dos POE terrestres foram 20 cm para os desvios posicionais (X_0 , Y_0 , Z_0) – valor extraído do manual do equipamento – e 3° para os desvios angulares em ω e φ – valor definido empiricamente. O manual do equipamento especifica valores muito otimistas para os desvios angulares no caso terrestre, por isto optou-se por defini-los empiricamente. O desvio angular em κ não apresentou qualquer efeito no resultado, sendo, portanto, desconsiderado. Os desvios-padrão considerados para as coordenadas planimétricas do ponto no terreno (X_p , Y_p) foram obtidos da primeira parte da propagação de covariâncias, e para a coordenada altimétrica Z foi adotado o valor constante de 0,5 m, definido empiricamente. Deste modo, a MVC dos parâmetros incógnitos (Σ_Y) para esta segunda relação é dada por:

$$\Sigma_Y = \begin{bmatrix} \sigma_{x_p}^2 & \sigma_{x_p y_p} \\ \sigma_{x_p y_p} & \sigma_{y_p}^2 \end{bmatrix}$$

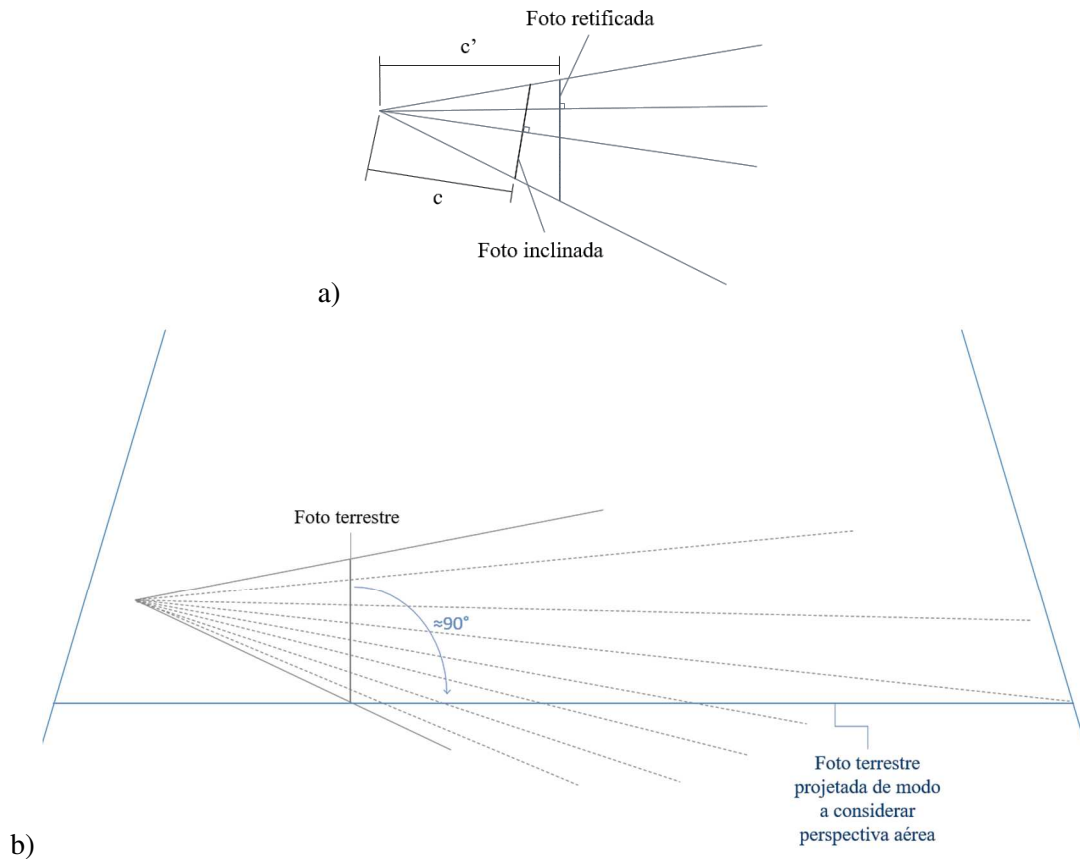
sendo σ_{x_p} e σ_{y_p} o desvio-padrão da posição calculada na imagem terrestre em x e y, respectivamente. Define-se então o recorte terrestre centrado na posição do ponto, com tamanho igual ao dobro do desvio-padrão obtido em x e y.

Os recortes gerados definem a área de cada imagem que será projetada para posterior determinação de semelhança das geometrias. Inicialmente, esta projeção seria realizada para um plano no terreno, onde uma região no entorno do ponto seria considerada no MDT, definindo uma superfície local, e um plano seria ajustado a esta superfície. A orientação deste plano, considerada na etapa de projeção dos recortes, seria fornecida por uma de suas retas normais. No entanto, foram feitos testes com esta metodologia, nos quais o ajuste de planos foi realizado de acordo com Canon (2009), que apresenta um método baseado no Método dos Mínimos Quadrados (MMQ), e os resultados obtidos não foram satisfatórios. A primeira limitação é que a metodologia se restringe a regiões planas. Não realizar um tratamento prévio que garanta que ela seja aplicada apenas neste tipo de cenário já torna o

método inviável, uma vez que não se pode garantir o comportamento da projeção em feições não-planas. Planos muito inclinados geram imagens muito deformadas na etapa de projeção. Além disto, os erros envolvidos no processo, especialmente o do MDT, influenciam na posição calculada do ponto, fazendo com que o mesmo não seja localizado exatamente sobre a superfície esperada e, assim, resulte em um plano que não representa fielmente o terreno. Portanto, devido à dificuldade de se obter um plano que represente a superfície local considerada no método (por exemplo, um trecho de rodovia), optou-se por utilizar outra estratégia de projeção para a determinação de correspondência.

Testes realizados anteriormente pela autora, bem como por Jende *et al.* (2016), mostraram que alterar a geometria da imagem terrestre a fim de coincidi-la com o caso aéreo pode representar a melhor solução para o problema. Optou-se então por projetar ambas as imagens para um plano horizontal, sendo pequeno o efeito causado na imagem aérea, que já é obtida de modo aproximadamente vertical, e realizando uma alteração de aproximadamente 90° em ω no caso terrestre. As projeções são realizadas considerando a orientação das imagens e utilizando o conceito de retificação: os cantos dos recortes são projetados para a imagem retificada utilizando o método de retificação direta (item 2.3.1), e a informação dentro deste intervalo é preenchida utilizando o método de retificação inversa (item 2.3.2). No caso aéreo, é realizada a retificação propriamente dita, na qual a imagem é projetada segundo seu próprio feixe perspectivo para um plano horizontal (ver Figura 5), apenas corrigindo-a da inclinação da câmara, enquanto no caso terrestre o plano de projeção considerado é aproximadamente ortogonal ao da imagem original, ou seja, também horizontal; gerando assim imagens com perspectivas similares. A Figura 22 ilustra o conceito de retificação para o caso terrestre e o que é, de fato, realizado neste trabalho.

Figura 22. a) Conceito de retificação para o caso terrestre, onde c representa a distância focal da imagem e c' a distância focal da imagem retificada. b) Projeção da imagem terrestre realizada neste trabalho por meio do conceito de retificação.



Mesmo não sendo realizada propriamente uma retificação no caso terrestre, no texto a operação é referida como tal. Ao final desta etapa, obtém-se então os pares de recortes “retificados”, juntamente com um conjunto de informações armazenadas a respeito de cada um deles: o tamanho do pixel considerado na imagem retificada, a distância focal retificada, e a posição do início do recorte na imagem original. Estas informações são úteis no cálculo inverso das coordenadas dos pontos correspondentes, Etapa IV da metodologia (item 4.4.4).

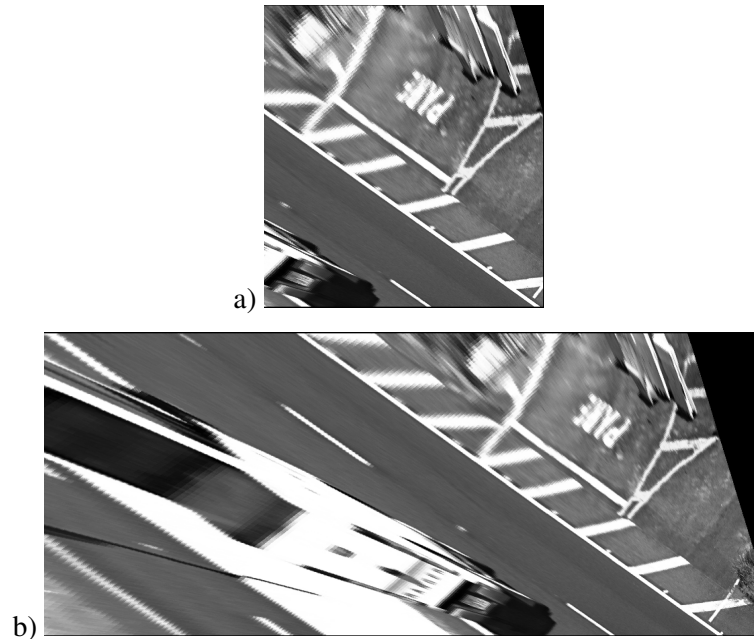
A quantidade de recortes gerados para cada par de imagens é bastante reduzida se comparada à quantidade de cantos extraídos pelos operadores, o que ocorre, principalmente, devido à restrição quanto à área de interesse, mencionada na Etapa I (item 4.4.1). Além desta, outras restrições foram impostas no algoritmo a fim de reduzir a ocorrência de resultados errôneos e, também, o custo computacional envolvido no processamento. Cada uma delas é apresentada na sequência.

(1) *Restrição quanto ao desvio-padrão da janela de busca:* Segundo a metodologia apresentada, cada ponto de interesse extraído na imagem aérea tem sua posição

determinada no terreno e então na imagem terrestre, com as respectivas precisões destas determinações. Esta restrição diz respeito ao tamanho da janela no momento da projeção do ponto para a imagem terrestre, que é determinado de acordo com o desvio-padrão da posição calculada. Foi possível notar, durante a realização dos experimentos, que janelas muito grandes resultam de pontos errados, isto é, pontos que não representam regiões correspondentes. Ainda que haja uma definição prévia da área de interesse, não é possível garantir que todos os pontos extraídos nesta área estejam presentes na imagem terrestre analisada. Testes realizados mostraram que pontos conhecidos eram projetados com erro máximo em torno de 350 pixels em x e 450 pixels em y, aproximadamente. Como estes valores são determinados considerando os erros envolvidos na OE das imagens e o MDT, a presença de um desvio-padrão maior na direção y implica em um maior erro no MDT, uma vez que ele é o responsável pela determinação da altura do ponto neste processo. Considerando então uma margem de segurança, adotou-se o maior valor como limiar para os desvios-padrão nas duas direções. Desta forma, pontos projetados com desvios-padrão superiores a 500 pixels são desconsiderados no processamento.

(2) *Restrição quanto ao tamanho do recorte terrestre:* Como já mencionado, o desvio-padrão da posição calculada na imagem terrestre define o tamanho da janela nesta imagem. Esta janela ou recorte será retificado para posterior busca por correspondência. No entanto, este processo chamado aqui de retificação é uma projeção de aproximadamente 90° no caso terrestre, o que ocasiona bastante deformação na imagem gerada, fazendo com que boa parte dela não possa ser aproveitada. Testes mostraram que a área útil das imagens retificadas, isto é, a área que pode ser utilizada para fins de correspondência, para o caso estudado, é de aproximadamente 2300x3000 pixels, cerca de 30% do total da imagem. Considerar valores superiores a estes apenas incluiria nas imagens uma área com presença de elementos bastante deformados, o que dificilmente resultaria em qualquer correspondência. A Figura 23 ilustra esta afirmativa.

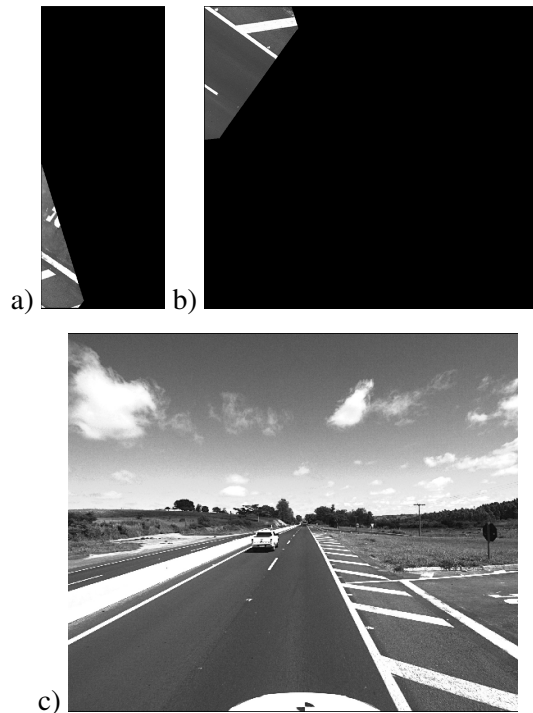
Figura 23. Recortes terrestres projetados, originados a partir de pontos de partida distintos (a, b).



É possível perceber na Figura 23 que a área útil é a mesma nos dois casos. O que acontece então é que se considera um número maior de imagens, mas as correspondências identificadas se repetem, sempre nas regiões de menor deformação, o que não é um resultado interessante considerando o objetivo deste trabalho. A partir disto e considerando-se uma margem de segurança, estabeleceu-se o limiar de 2500x3200 pixels para colunas e linhas, respectivamente. Recortes projetados cujo tamanho seja superior a este limiar são desconsiderados.

(3) *Restrição quanto ao conteúdo do recorte terrestre:* O processo de retificação corrige a imagem da inclinação da câmara, projetando-a em um plano vertical ou horizontal. Este processo acarreta uma “perda” de informação na imagem, ou seja, uma região da imagem original deixa de ser considerada após a projeção. Isto ocorre devido ao plano de projeção ser diferente do original. Quanto maior for a diferença entre esses planos, maior será a informação perdida no processo de projeção. No caso terrestre abordado neste trabalho, a projeção realizada é de aproximadamente 90°, o que acarreta uma perda bastante significativa da imagem. A Figura 24 apresenta dois recortes gerados com um percentual muito pequeno de informação representada, bem como a imagem terrestre que originou um desses recortes.

Figura 24. a) e b) Recortes terrestres retificados, gerados com um percentual muito pequeno de informação representada. c) Imagem terrestre que originou o recorte a).



Nas Figuras 24.a) e 24.b) a pequena porção que representa um trecho da rodovia é a área do recorte que possui informação, e a região em preto é a área sem informação representada. Esta presença de região sem informação pode ocorrer por dois motivos: (1) alguns pontos são projetados com desvio-padrão tal que aparecem no canto do recorte retificado. Quando estes pontos estão localizados na borda da imagem original (Figura 24.c)), não haverá informação a ser representada no entorno dele no recorte gerado, fazendo com que se assemelhe ao apresentado na Figura 24.a); (2) alguns pontos estão contemplados na área de interesse e não aparecem realmente na imagem terrestre, estando apenas próximos da região imageada. Caso estes pontos aparecessem, eles estariam localizados nesta região sem informação do recorte e, por isto, acabam sendo considerados pelo método. Testes mostraram que recortes que se enquadram neste segundo caso possuem menos de 1/5 de informação representada, enquanto os que se enquadram no primeiro caso representam uma porcentagem maior de informação. A fim de evitar os recortes que se enquadram no segundo caso, esta restrição avalia o conteúdo do recorte terrestre retificado. A informação contida em cada um deles deve representar pelo menos 1/5 do total do recorte; caso contrário, o mesmo é desconsiderado.

Ao final desta etapa são obtidos os pares de recortes retificados.

4.4.3 ETAPA III – Busca por correspondência

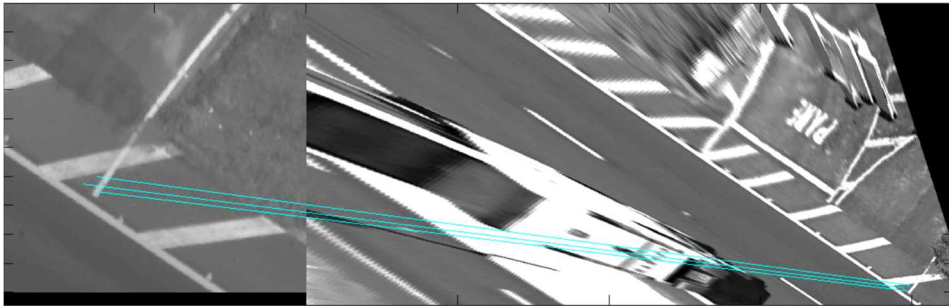
A busca por correspondência é realizada nos recortes projetados, gerados na etapa anterior, utilizando o operador SIFT, cuja implementação foi realizada no Matlab e disponibilizada por Yan (2011). O programa então lê os pares de recortes gerados na Etapa II e realiza o processo de busca por correspondência em cada um deles. Para cada par analisado é gerado um arquivo contendo as coordenadas dos pontos correspondentes, caso existam, e uma figura mostrando o resultado da busca.

A implementação utilizada tem como base um parâmetro – *distRatio* – que funciona como um limiar. Calcula-se o ângulo entre os dois primeiros vizinhos mais próximos do ponto analisado em cada imagem, se a proporção entre estes ângulos for menor que o limiar estabelecido (*distRatio*) os pontos são classificados como correspondentes. O valor assumido para o parâmetro é arbitrário, sendo menor quanto maiores forem as semelhanças entre as imagens. Quanto menor for este valor, mais restrito é o limiar, isto é, menor a quantidade de correspondências identificadas e maior a qualidade da identificação. Neste trabalho foram utilizados dois diferentes valores para os pares de imagens analisados. Testes mostraram que os pares formados na região central da imagem aérea retornavam uma quantidade maior de correspondências que os demais. Por isso, para estes casos – pares 5, 6 e 7 (ver item 5.1) – adotou-se o valor $distRatio = 0,6$ e para os demais $distRatio = 0,7$. Considerar valores mais restritos para este parâmetro permite a identificação de uma quantidade menor de correspondências, desconsiderando, inclusive, correspondências verdadeiras. No entanto, considerando um objetivo futuro deste trabalho, que é a realização de uma fototriangulação em conjunto, a qualidade da identificação é mais importante do que a quantidade de correspondências identificadas, não sendo necessária uma grande quantidade de correspondências, especialmente em uma região tão pequena como as que são consideradas nos recortes. Embora o parâmetro $distRatio = 0,6$ permita um resultado com melhor qualidade, utilizá-lo em todos os casos resultaria em uma quantidade muito pequena ou até insuficiente de correspondências nas áreas de periferia da imagem aérea.

Não existe um padrão de identificação de correspondências, podendo ser identificada uma grande quantidade em um par de recortes e em outro nenhuma. Além disto, a quantidade identificada não representa, necessariamente, a quantidade de correspondências verdadeiras. Com este fim, uma primeira verificação foi realizada levando em conta o ângulo de inclinação das retas que ligam pontos homólogos. Uma simples transformação, como a isogonal ou afim, seria suficiente para modelar as variações entre os recortes, representando

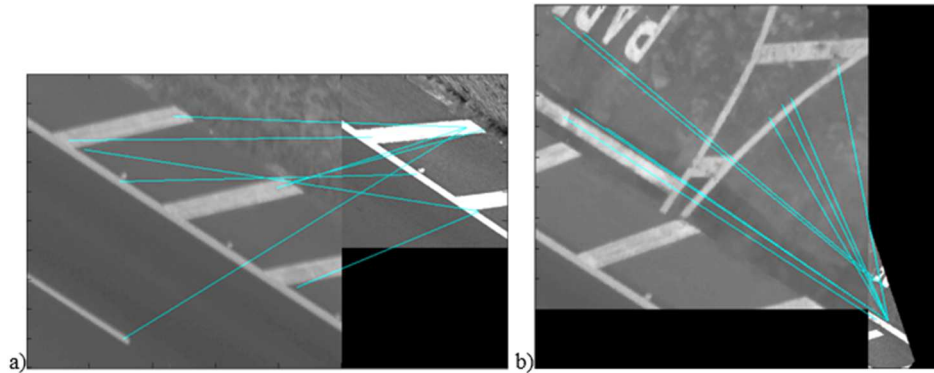
assim um modelo seguro para a verificação das correspondências. No entanto, os parâmetros desta transformação seriam diferentes para cada par analisado, pois os recortes terrestres assumem tamanhos variados, tornando necessária a redefinição dos parâmetros a cada nova análise e, assim, inviabilizando a aplicação do método. A variação na posição de um ponto em duas imagens é influenciada, neste caso, apenas por movimentos de translação e rotação, que são únicos para todo o par. Desta forma, as retas que ligam pontos homólogos devem apresentar aproximadamente a mesma inclinação, como pode ser visto no exemplo da Figura 25.

Figura 25. Exemplo de retas que ligam pontos homólogos - mesma inclinação.



Com base nisto, utilizou-se o ângulo de inclinação das retas como métrica para a verificação das correspondências verdadeiras. Partindo do pressuposto de que a maioria das retas do conjunto representam correspondências verdadeiras, a verificação é realizada considerando a variação entre a inclinação de uma reta e a inclinação média daquele conjunto. Retas que diferem cerca de 10° da média do conjunto tem seus pontos desconsiderados, sendo este valor definido empiricamente. Esta verificação contribui para uma primeira filtragem do resultado, mas não pode ser aplicada a todos os casos. A Figura 26 apresenta dois casos em que o método não pode ser aplicado.

Figura 26. Exemplos de casos em que não é possível aplicar o método de verificação baseado na inclinação das retas.



Na Figura 26.a) não é possível identificar um padrão nas retas apresentadas, não sendo possível, portanto, a aplicação deste método de verificação. Já na Figura 26.b), cinco retas aparecem com a mesma inclinação, mas outras quatro apresentam inclinações variadas. A ocorrência de retas cujas inclinações são diferentes do padrão presente no par é muito alta, fazendo com que o resultado se torne tendencioso e impossibilitando a aplicação do método de verificação. Este problema poderia ser evitado se, ao invés da média, fosse utilizado o conceito de moda ou mediana. Além destas, o algoritmo RANSAC (FISCHLER e BOLES, 1981), acrônimo de *RANdom SAmple Consensus*, e o método de “*majority voting*”, apresentado em Berveglieri e Tommaselli (2015), representam outras possíveis soluções para o problema.

Assim, o método baseado na inclinação das retas realiza uma verificação preliminar das correspondências, mas possui aplicação restrita e não identifica todos os erros envolvidos – como o problema de padrões repetitivos, por exemplo – fazendo com que mesmo após a verificação ainda exista inconsistências nos resultados. Realizou-se então uma verificação visual com os pontos plotados nas imagens originais, operação apresentada na sequência – Etapa IV da metodologia. Recomenda-se para trabalhos futuros estudar este problema, comparando-se as opções de filtragem mencionadas.

4.4.4 ETAPA IV – Projeção inversa para as imagens originais

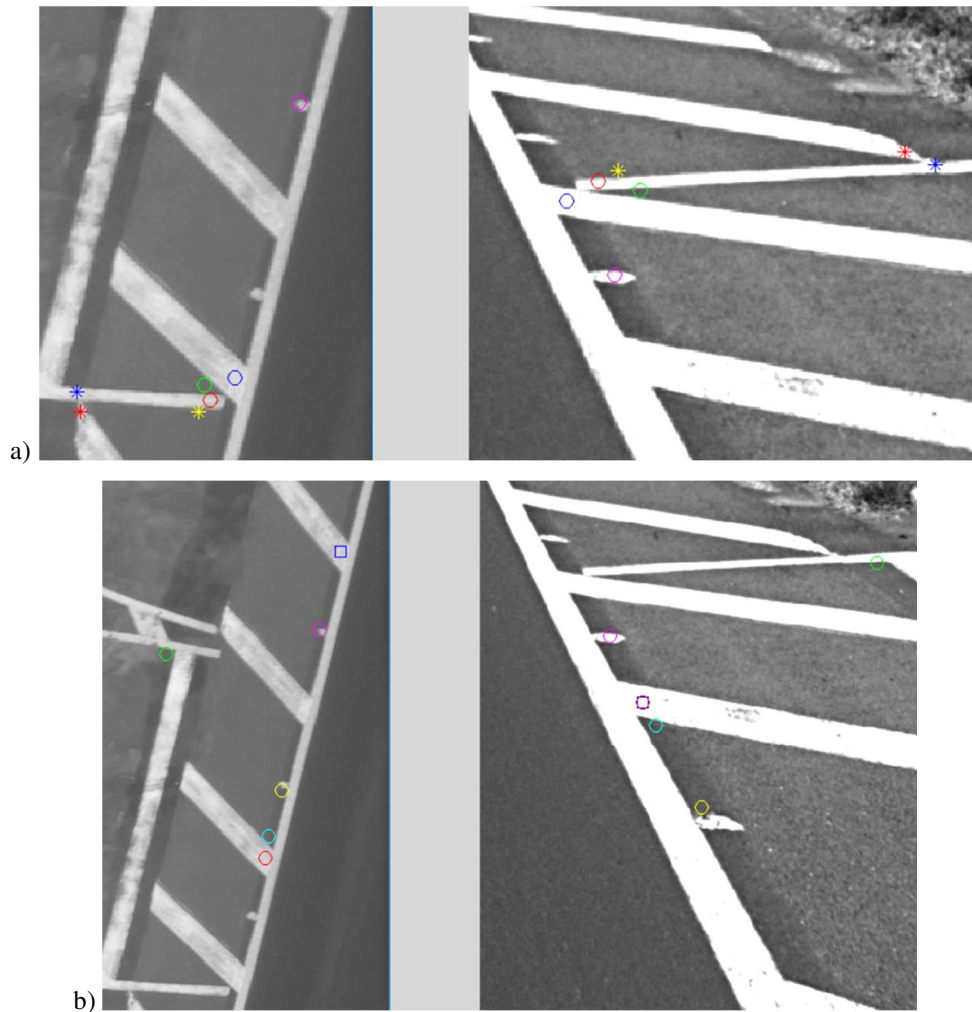
As coordenadas dos pontos tidos como correspondentes neste momento são referentes às posições nos recortes retificados, quando o que se deseja são estas posições identificadas nas imagens originais. Para isto, uma função foi implementada e acrescentada ao

código original. O processo se assemelha ao método inverso da retificação, isto é, busca a posição correspondente na imagem original de pontos identificados na imagem retificada.

A função utiliza como dados de entrada os arquivos gerados na Etapa III (item 4.4.3), contendo as coordenadas dos pontos correspondentes em cada par, e também as informações armazenadas juntamente com cada recorte retificado: o tamanho do pixel na imagem retificada, a distância focal retificada e a origem do recorte na imagem original. Utilizar estes dados garante que as informações utilizadas neste processo inverso sejam as mesmas do processo de obtenção do recorte retificado, minimizando assim os erros ocasionados por aproximações. O resultado desta função é a mesma lista de coordenadas da Etapa III (item 4.4.3), com exceção dos pontos reprovados na primeira verificação, mas agora representando posições nas imagens originais. Estes pontos são então plotados nas imagens originais e uma última verificação é realizada a fim de eliminar todas as correspondências incorretas remanescentes.

A verificação final das correspondências é visual e não automática: os pontos foram plotados utilizando símbolos e cores diferentes, de modo que, em um par de imagens analisado, não houvesse mais de um par de pontos plotados com mesma cor e símbolo, permitindo assim a identificação e verificação de cada um deles. A plotagem foi realizada com auxílio do Matlab. A Figura 27 ilustra o processo de verificação nos pares de imagens, apresentando sempre a ordem imagem aérea (à esquerda) seguida da imagem terrestre (à direita). Embora na figura esteja representado um *zoom* da área onde os pontos estão localizados, a fim de facilitar a visualização do leitor, a análise foi realizada considerando o cenário presente em toda a imagem, permitindo, assim, a identificação de erros ocasionados pela presença de padrões repetitivos. Vale lembrar ainda que foi mantida a orientação original das imagens, o que faz com que elas representem sentidos opostos.

Figura 27. Processo de verificação visual das correspondências: pontos homólogos representados com mesmo símbolo e cor em ambas as imagens.



A Figura 27.a) apresenta a análise de um recorte onde foram identificadas sete correspondências. Pode-se perceber um pequeno deslocamento na posição de alguns pontos – como os representados por um círculo magenta e por um asterisco vermelho – o que é esperado devido à grande mudança de perspectiva entre as imagens e também à característica do método SIFT, utilizado na correspondência dos pontos, que procura pontos invariantes com a escala. Neste caso, apenas o par de pontos representado com o círculo magenta indica uma falsa correspondência, os demais foram classificados como correspondências verdadeiras.

A Figura 27.b) apresenta a análise para um recorte em que foram identificadas seis correspondências. Os pontos representados com o círculo vermelho e o quadrado azul resultaram em uma mesma posição na imagem terrestre e, por isso, estão representados em sobreposição. A classificação final das correspondências para este caso foi:

- Corretas: círculos amarelo, cyan e vermelho;
- Incorretas: círculos verde e magenta, e quadrado azul.

Deste modo, a verificação foi realizada para todos os pontos e recortes. O resultado final é a lista de coordenadas dos pontos homólogos nas imagens originais, bem como uma figura com a representação visual destes pontos. Com a realização de uma fototriangulação conjunta destas imagens, a última etapa de verificação das correspondências seria uma análise dos resíduos da fototriangulação, o que permitiria a eliminação das falsas correspondências.

CAPÍTULO V

EXPERIMENTOS, RESULTADOS E ANÁLISES

5.1 DESCRIÇÃO DOS EXPERIMENTOS

A fim de verificar e validar a metodologia proposta neste trabalho, os experimentos realizados consistem na busca por correspondência em dez pares de imagens. Os pares foram formados por combinações de uma única imagem aérea com dez imagens terrestres, cujas regiões imageadas são comuns às duas plataformas. A seleção destas imagens foi realizada considerando um objetivo futuro deste trabalho: a possibilidade de fototriangulação simultânea de imagens aéreas e terrestres. Assim, as imagens terrestres utilizadas garantem a cobertura da área de rodovia presente no caso aéreo, e esta mesma busca por correspondência, se realizada em outras duas imagens aéreas adjacentes, seria suficiente para a realização de uma fototriangulação. No entanto, este trabalho apenas verifica a possibilidade de se obter correspondência entre estas imagens.

Primeiramente, foi verificada toda a metodologia proposta neste trabalho, apresentando os resultados obtidos em cada etapa, para todos os pares de imagens. Posteriormente, verificou-se os resultados apresentados por métodos disponíveis na literatura para esta finalidade, e também a possibilidade de aplicação de um método de correspondência baseada em área, uma vez que os recortes, após a projeção, tornaram-se muito semelhantes. Neste último experimento realizou-se um teste analisando apenas dois pares de imagens, pois as estratégias foram inicialmente desenvolvidas para a correspondência por características. Mais detalhes são apresentados juntamente com o experimento.

As Figuras 28 e 29 apresentam o conjunto de imagens utilizado nos experimentos.

Figura 28. Conjunto de imagens utilizado nos experimentos.

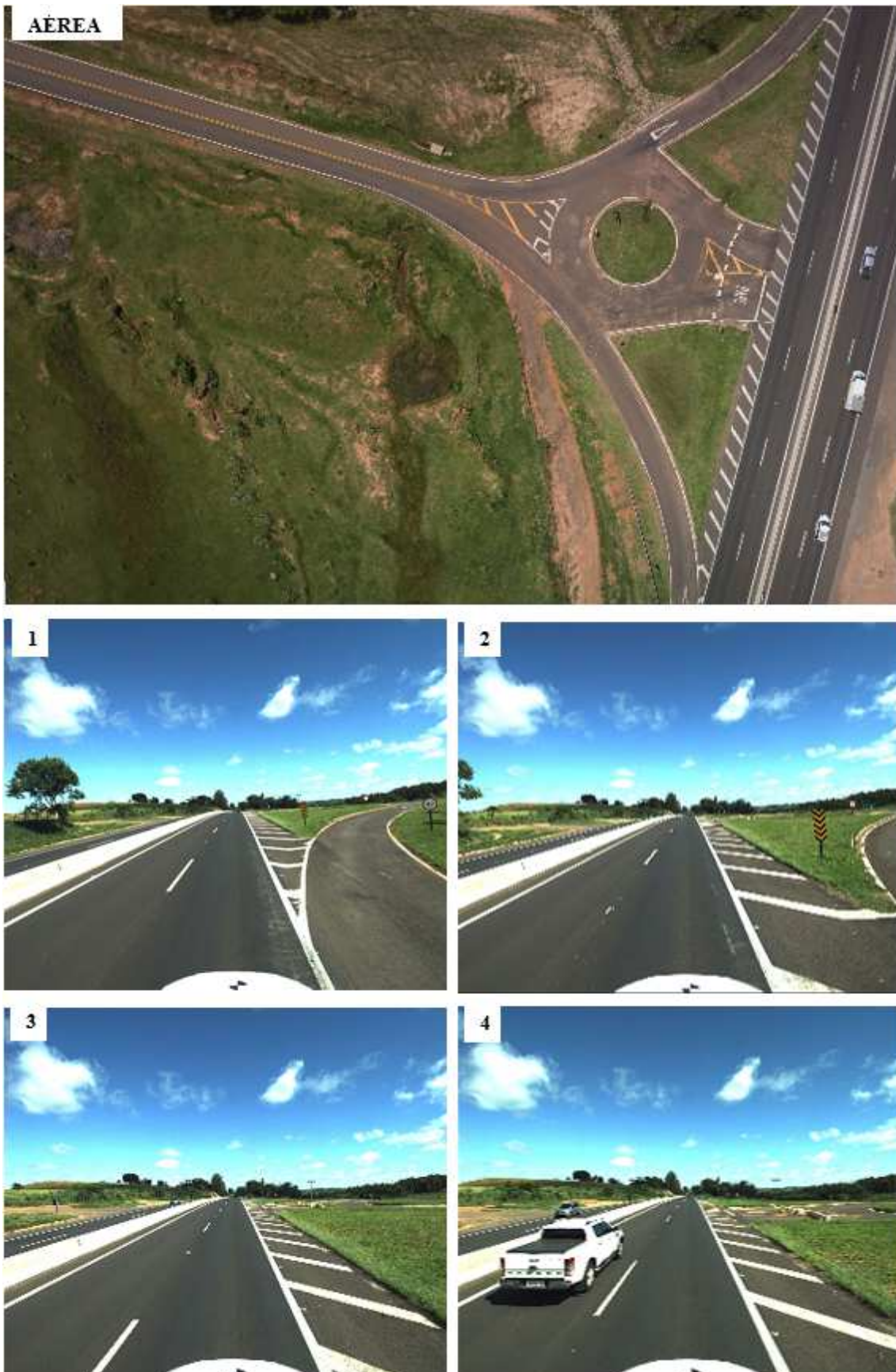
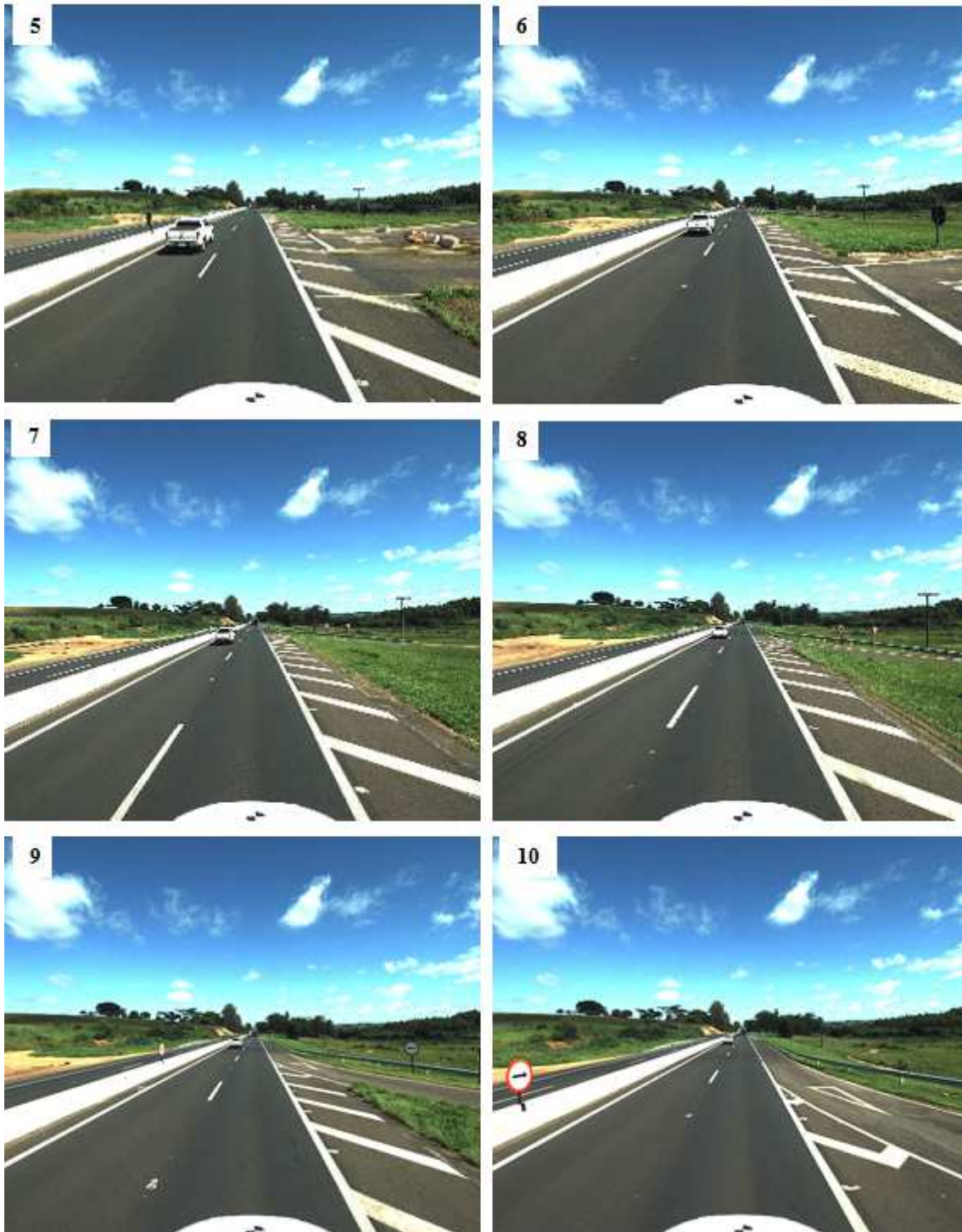
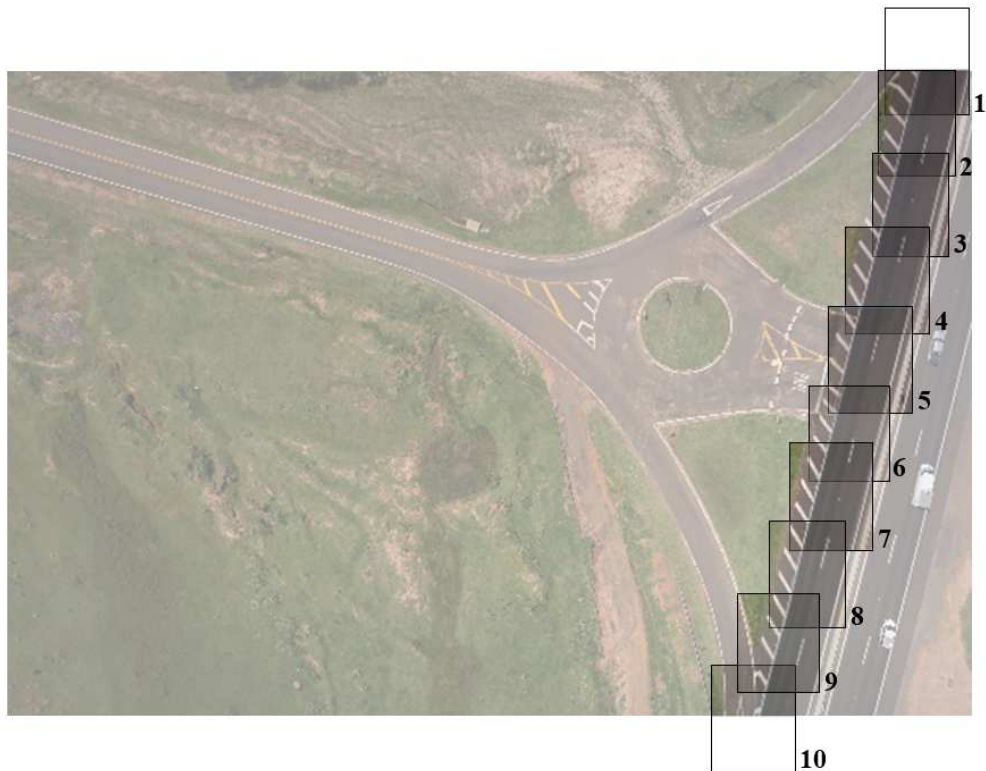


Figura 29. Conjunto de imagens utilizado nos experimentos (continuação).



Os pares foram enumerados de acordo com sua composição, sendo par 1 composto pela imagem aérea e a terrestre 1, par 2 composto pela imagem aérea e a terrestre 2, e, assim, sucessivamente. A Figura 30 apresenta a localização de cada uma das imagens terrestres em relação à aérea.

Figura 30. Localização das imagens terrestres em relação à aérea.



Para os experimentos, as imagens foram convertidas para o formato PGM (tons de cinza), acrônimo de *Portable Gray Map*, com auxílio do *software* Irfanview. Os resultados são apresentados de acordo com a ordem considerada na metodologia e, conseqüentemente, na realização das operações.

5.2 RESULTADOS

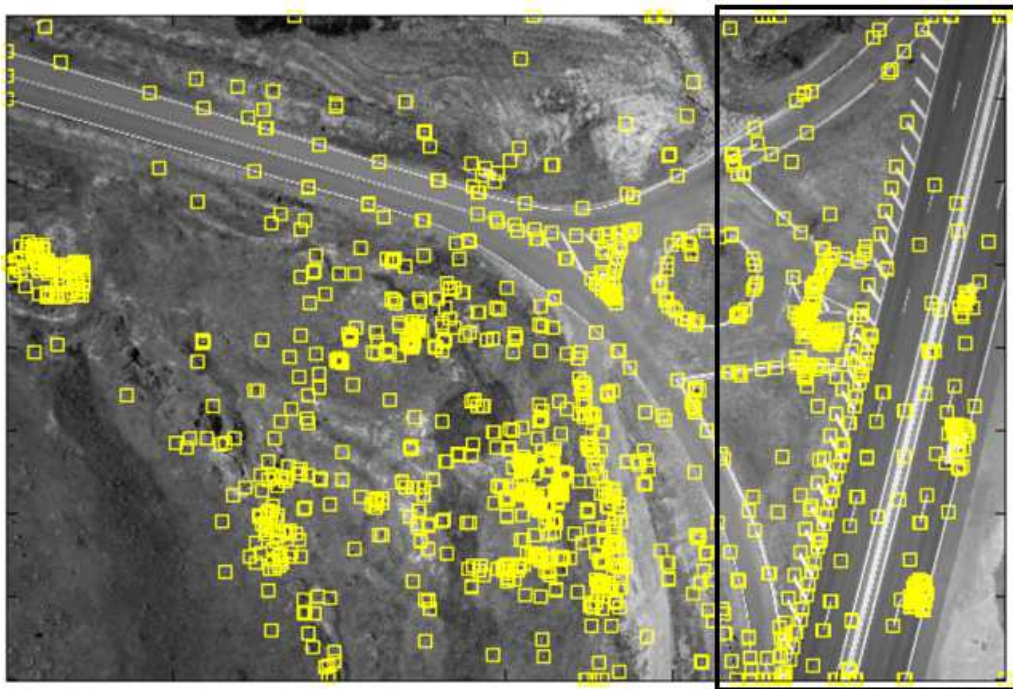
5.2.1 Etapa I

Como a imagem aérea utilizada é a mesma para todos os pares, a extração dos pontos de partida foi realizada uma única vez para toda a imagem e criou-se uma restrição quanto à área de interesse no programa. Desta forma, em cada processamento são considerados apenas os pontos de partida localizados na região correspondente à imagem terrestre (região de interesse).

Entre os testes realizados com o operador Harris, o melhor resultado foi obtido considerando os parâmetros $\sigma = 1$; $thresh = 3500$; e $radius = 3$; que permitiram uma melhor distribuição dos pontos. O valor de limiar ($thresh$) considerado é bastante superior ao

sugerido pelo autor. No entanto, valores menores ocasionavam uma geração exagerada de pontos, isto é, uma grande quantidade de pontos gerados próximos uns dos outros, havendo sobreposição na maioria deles e cuja distribuição era muito concentrada na parte inferior da imagem, não sendo adequada para o caso tratado neste trabalho. O resultado utilizado é apresentado na Figura 31.

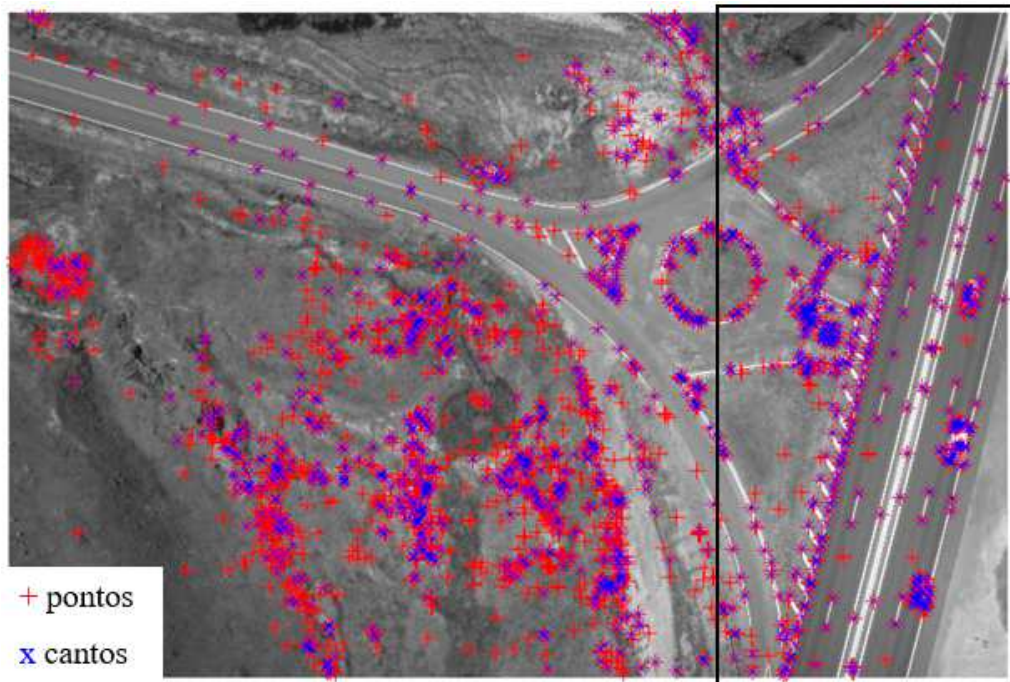
Figura 31. Resultado da extração de pontos de partida com o operador Harris, com destaque para a região considerada neste trabalho (área útil).



Na Figura 31, os quadrados amarelos indicam os cantos detectados pelo operador Harris. Foram extraídos no total 1409 cantos, sendo 616 deles localizados em área útil, ou seja, na área de rodovia e proximidades – área recoberta pelo sensor terrestre (destacada na Figura 31).

Entre os testes realizados com o detector Förstner, o melhor resultado foi obtido utilizando os valores padrões dos parâmetros, exceto para o valor de desvio-padrão. Para este, adotou-se o valor $SIGMA_N = 3$, pois o valor sugerido pelo autor retornava uma grande quantidade de pontos, de modo que começavam a se sobrepôr, não sendo interessante para o objetivo deste trabalho. O resultado da extração de pontos com o operador Förstner é apresentado na Figura 32.

Figura 32. Resultado da extração de pontos de partida com o operador Förstner, com destaque para a região considerada neste trabalho (área útil).



Na Figura 32, as cruzetas em vermelho representam os pontos extraídos pelo operador e os identificados com um 'x' azul representam os cantos classificados pelo operador com qualidade subpixel. Neste trabalho foram utilizados apenas os pontos identificados como cantos. Foram extraídos no total 1296 cantos, sendo 509 deles localizados em área útil (destacada na Figura 32).

Ao analisar as Figuras 31 e 32 é possível perceber uma melhor distribuição dos pontos identificados pelo operador Förstner (Figura 32), garantindo a cobertura quase que completa dos cantos presentes na marcação rodoviária contida na cena, com exceção apenas dos primeiros padrões da rodovia (canto superior direito). O operador Harris, embora tenha identificado uma grande quantidade de pontos na parte mais baixa da figura, considerando a área de rodovia, deixou o início dela bastante vazio, tendo pouquíssimos pontos identificados e não garantindo a cobertura da região. Além da melhor distribuição dos pontos e cobertura da imagem, o operador Förstner possui a vantagem de identificar os cantos com qualidade subpixel, que não é apresentada pelo operador Harris. Portanto, foi adotado para os experimentos o resultado obtido com o operador Förstner.

5.2.2 Etapa II

Na segunda etapa realiza-se a projeção dos recortes por meio da ferramenta implementada pela autora, considerando as restrições e a metodologia descritas anteriormente. Os pares 1, 9 e 10 não apresentaram nenhum recorte considerável na busca por correspondência. Isto porque, embora visualmente as imagens apresentem regiões correspondentes, as restrições impostas pelo algoritmo, bem como a própria distribuição dos pontos de partida impediram e/ou inviabilizaram a utilização de qualquer recorte gerado neste processo. Deste modo, estes pares foram desconsiderados das etapas seguintes. Os pares de recortes gerados neste experimento podem ser vistos na Etapa III (item 5.2.3), juntamente com os resultados da correspondência.

5.2.3 Etapa III

Na terceira etapa realiza-se a busca por correspondência nos recortes projetados. Cada conjunto de recortes obtido na etapa anterior foi analisado com o operador SIFT. Os resultados obtidos estão apresentados nas Figuras 33 a 41, com destaque para as correspondências identificadas como incorretas pelo método de verificação baseado na inclinação das retas. Vale lembrar que o método não é aplicável a qualquer situação, sendo esta discriminação, neste caso, realizada visualmente pela própria autora. Os pares que não apresentam retas em destaques não foram submetidos à verificação ou não tiveram nenhum retorno quando submetidos. Os recortes foram agrupados nas figuras de acordo com o par original a que correspondem, e a numeração que acompanha cada um deles indica o ponto de partida que originou aquele recorte. Em todos os casos a representação traz recorte aéreo e terrestre, respectivamente.

Figura 33. Resultado da busca por correspondência com o operador SIFT nos recortes projetados (par 2). Correspondências incorretas ainda não eliminadas.

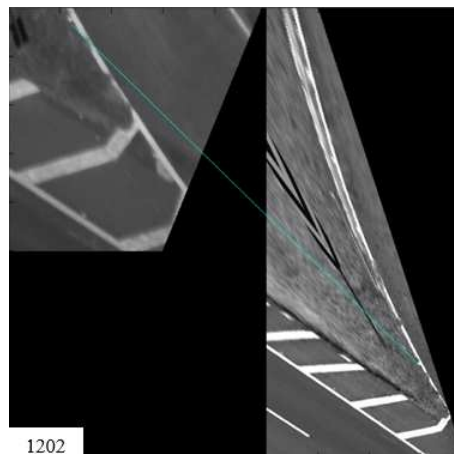


Figura 34. Resultado da busca por correspondência.a com o operador SIFT nos recortes projetados (par 3). Correspondências incorretas ainda não eliminadas.

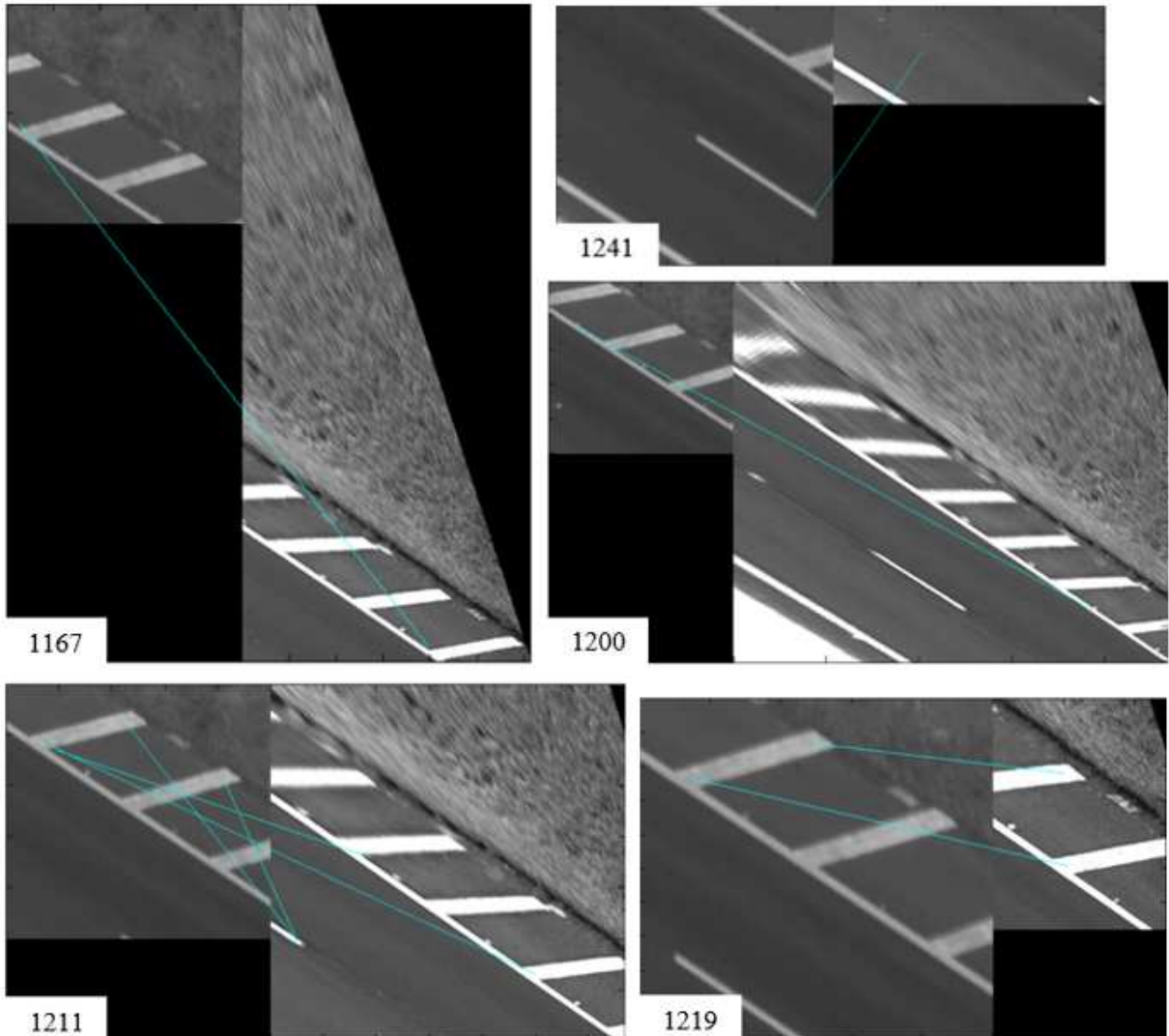


Figura 35. Resultado da busca por correspondência com o operador SIFT nos recortes projetados (par 4). Correspondências incorretas ainda não eliminadas.

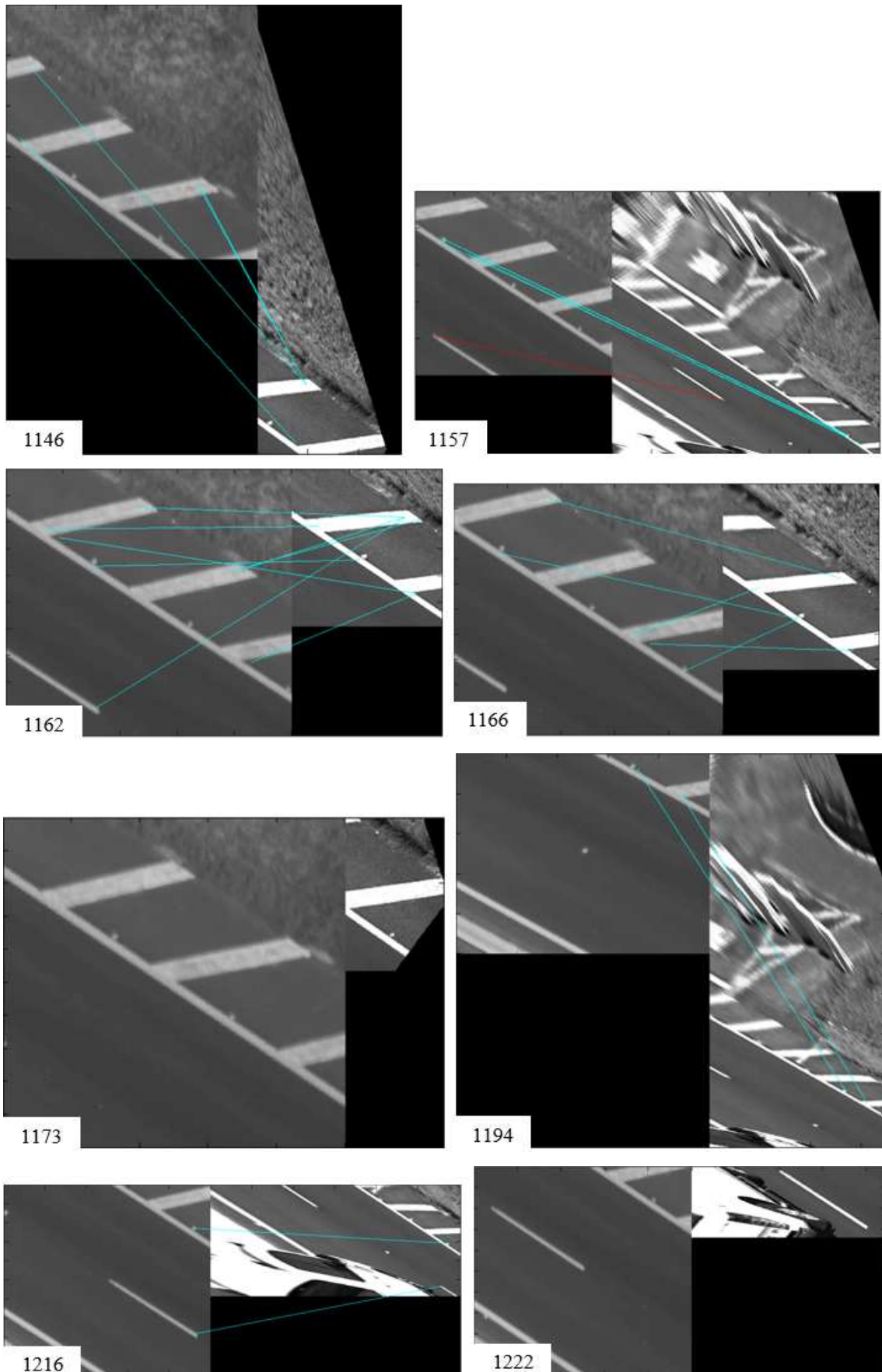


Figura 36. Resultado da busca por correspondência com o operador SIFT nos recortes projetados (par 5). Correspondências incorretas ainda não eliminadas.

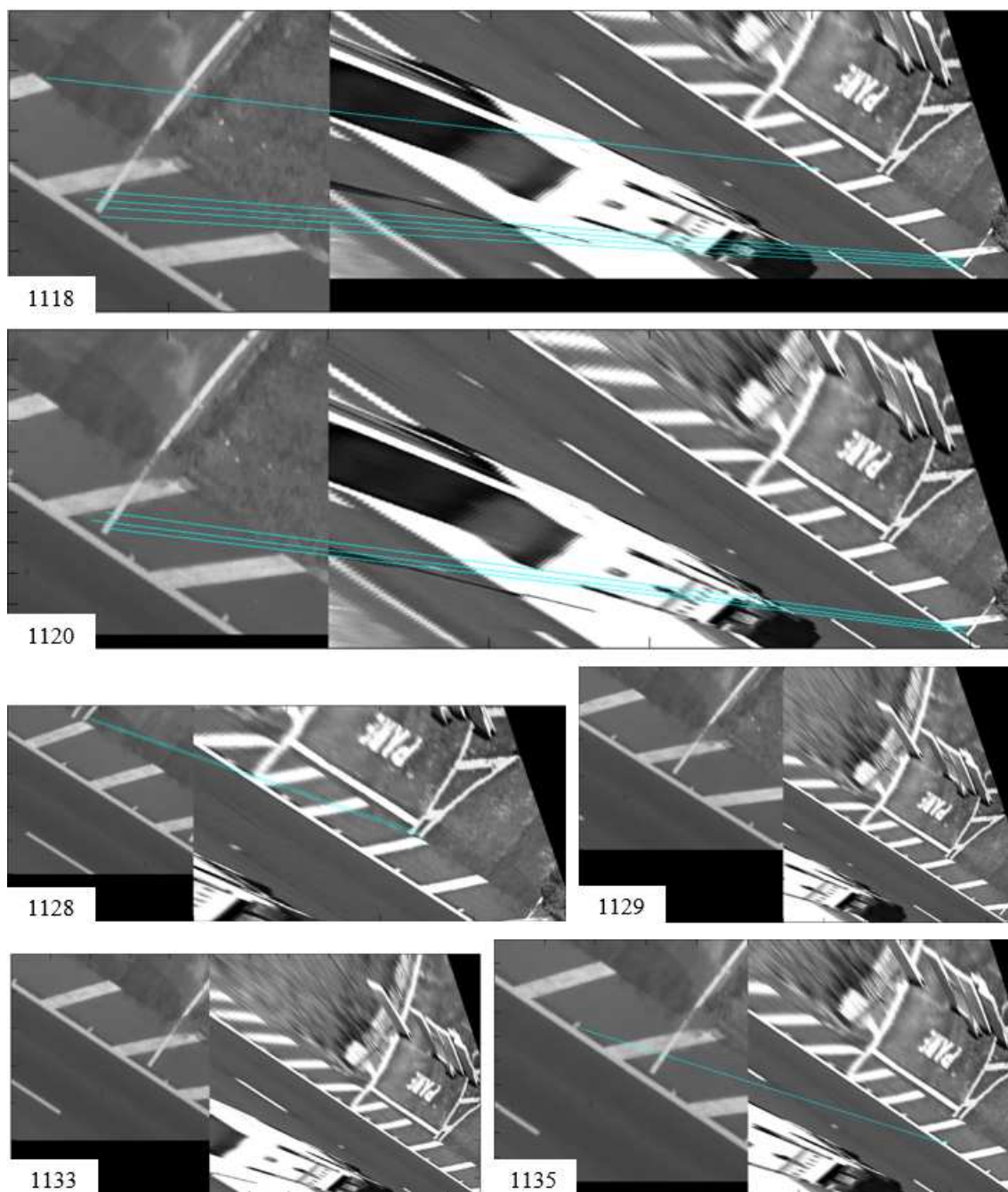


Figura 37. Resultado da busca por correspondência com o operador SIFT nos recortes projetados (par 5 - continuação). Correspondências incorretas ainda não eliminadas.

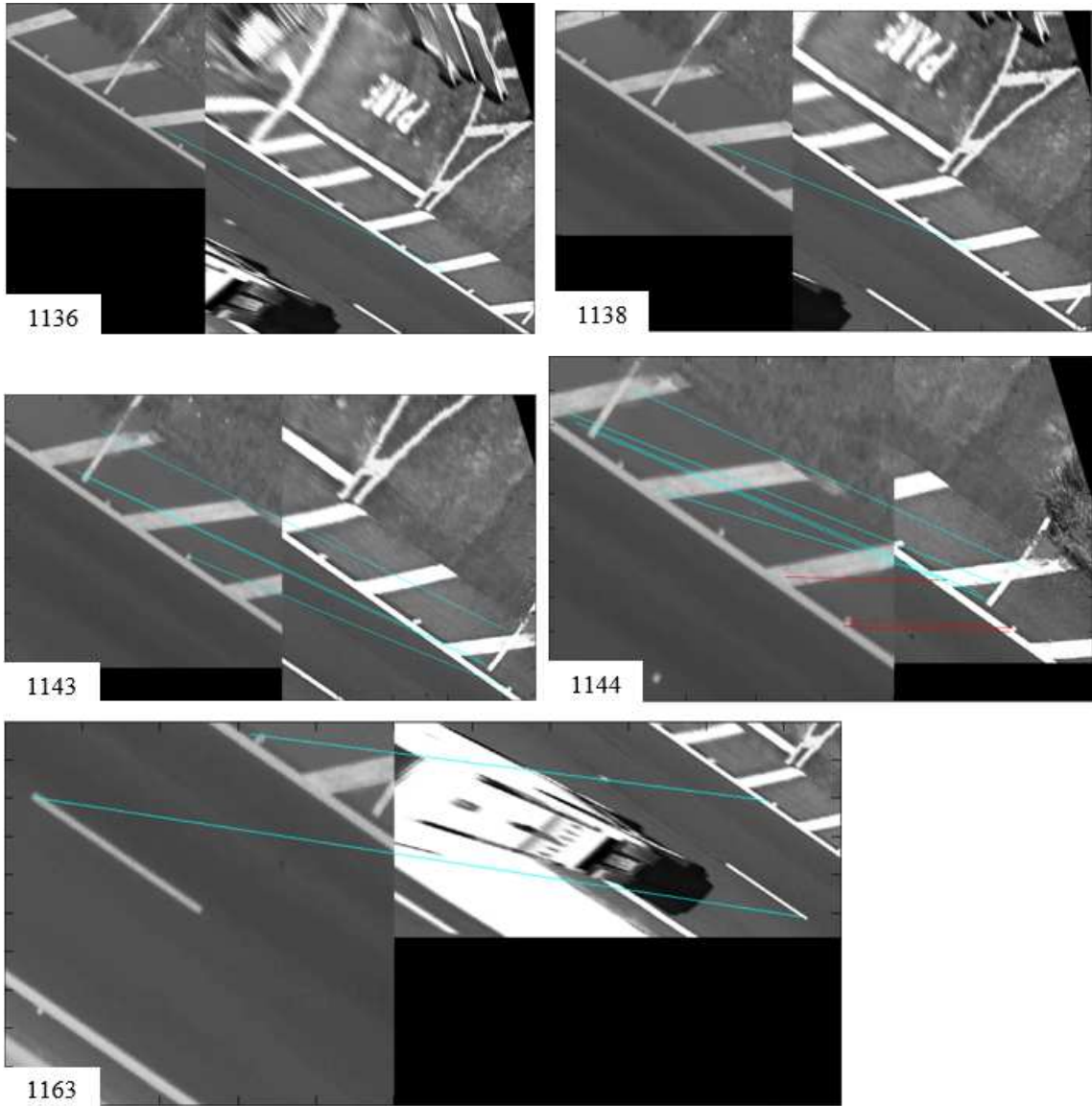


Figura 38. Resultado da busca por correspondência com o operador SIFT nos recortes projetados (par 6). Correspondências incorretas ainda não eliminadas.

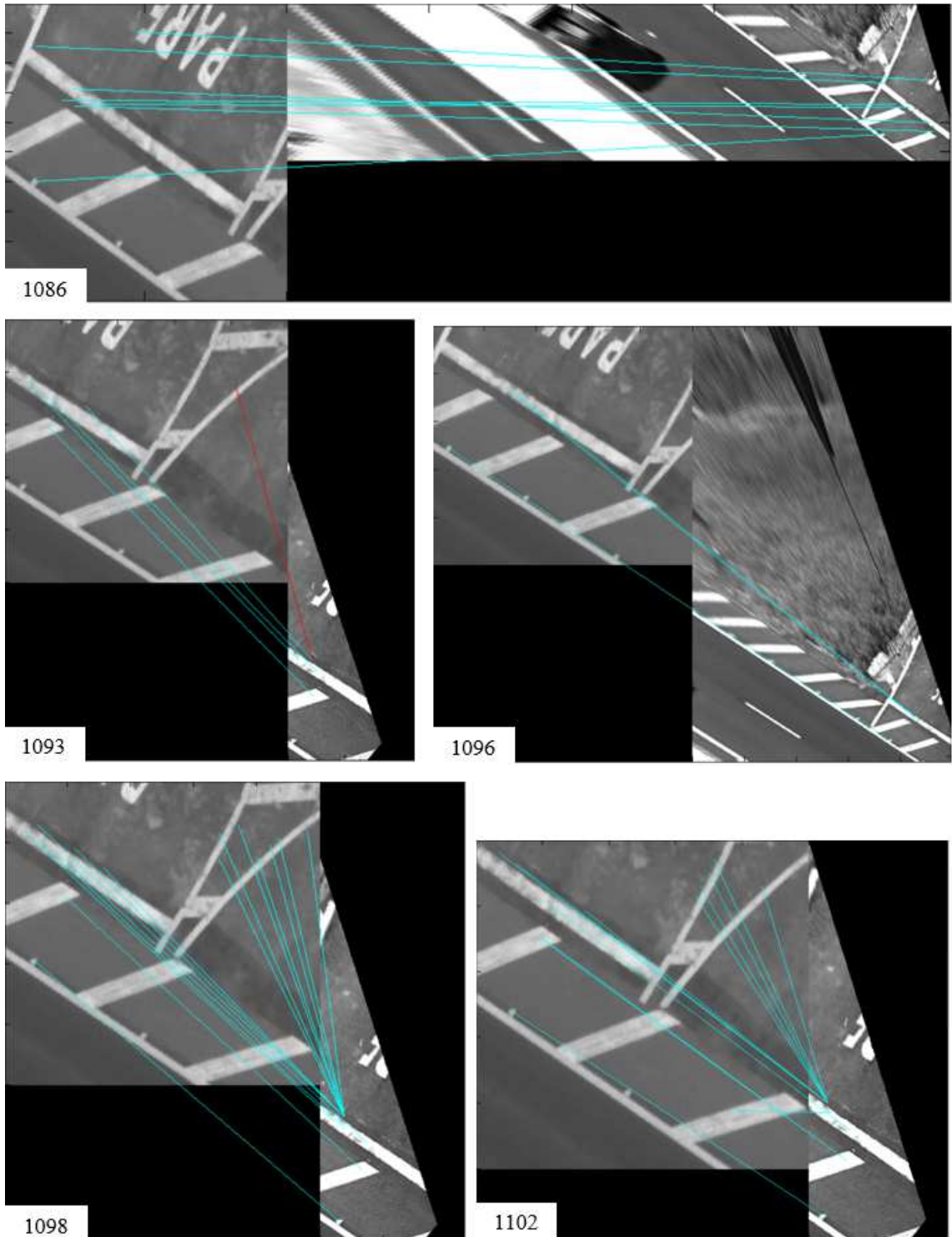


Figura 39. Resultado da busca por correspondência com o operador SIFT nos recortes projetados (par 6 - continuação). Correspondências incorretas ainda não eliminadas.

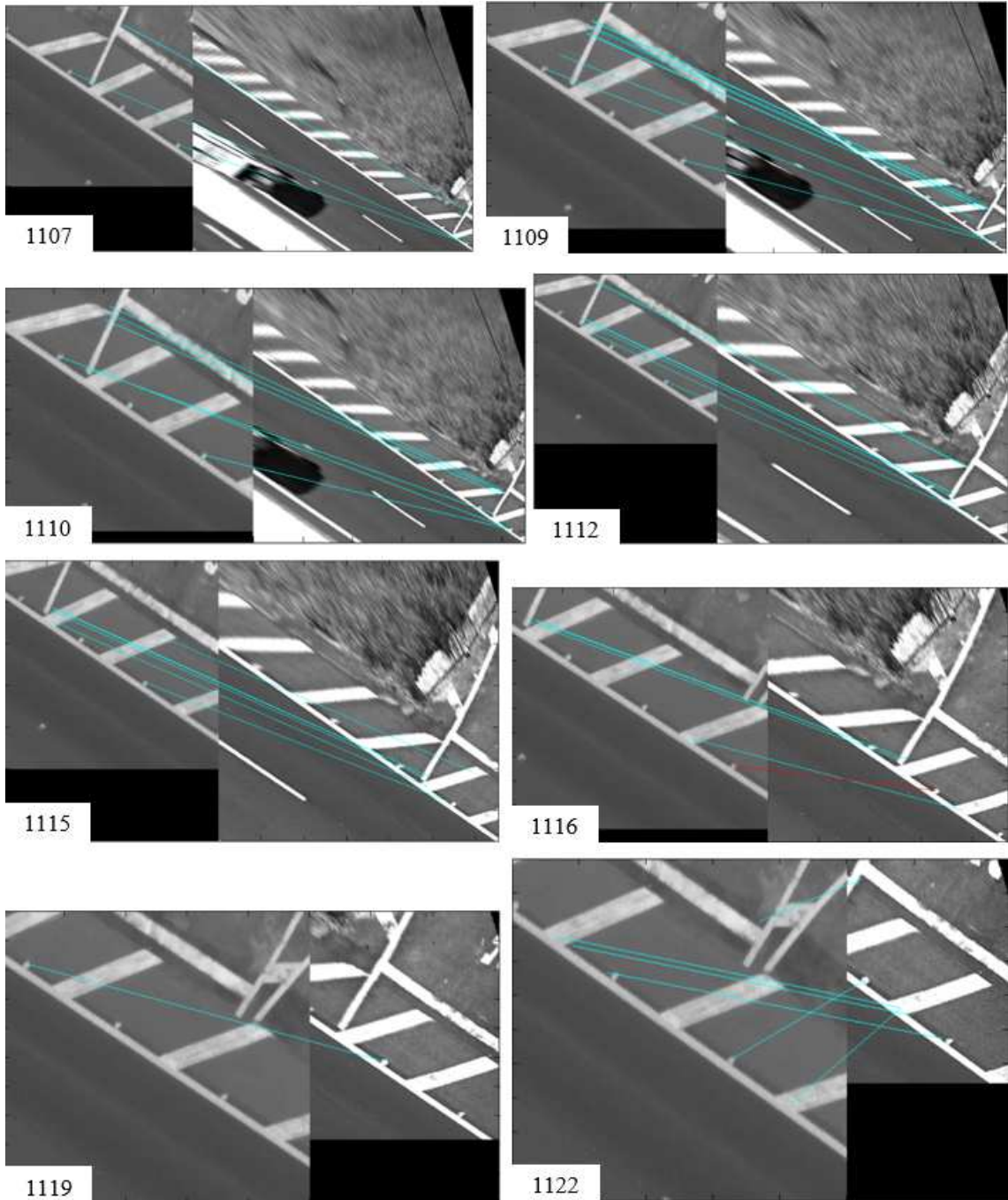


Figura 40. Resultado da busca por correspondência com o operador SIFT nos recortes projetados (par 7). Correspondências incorretas ainda não eliminadas.

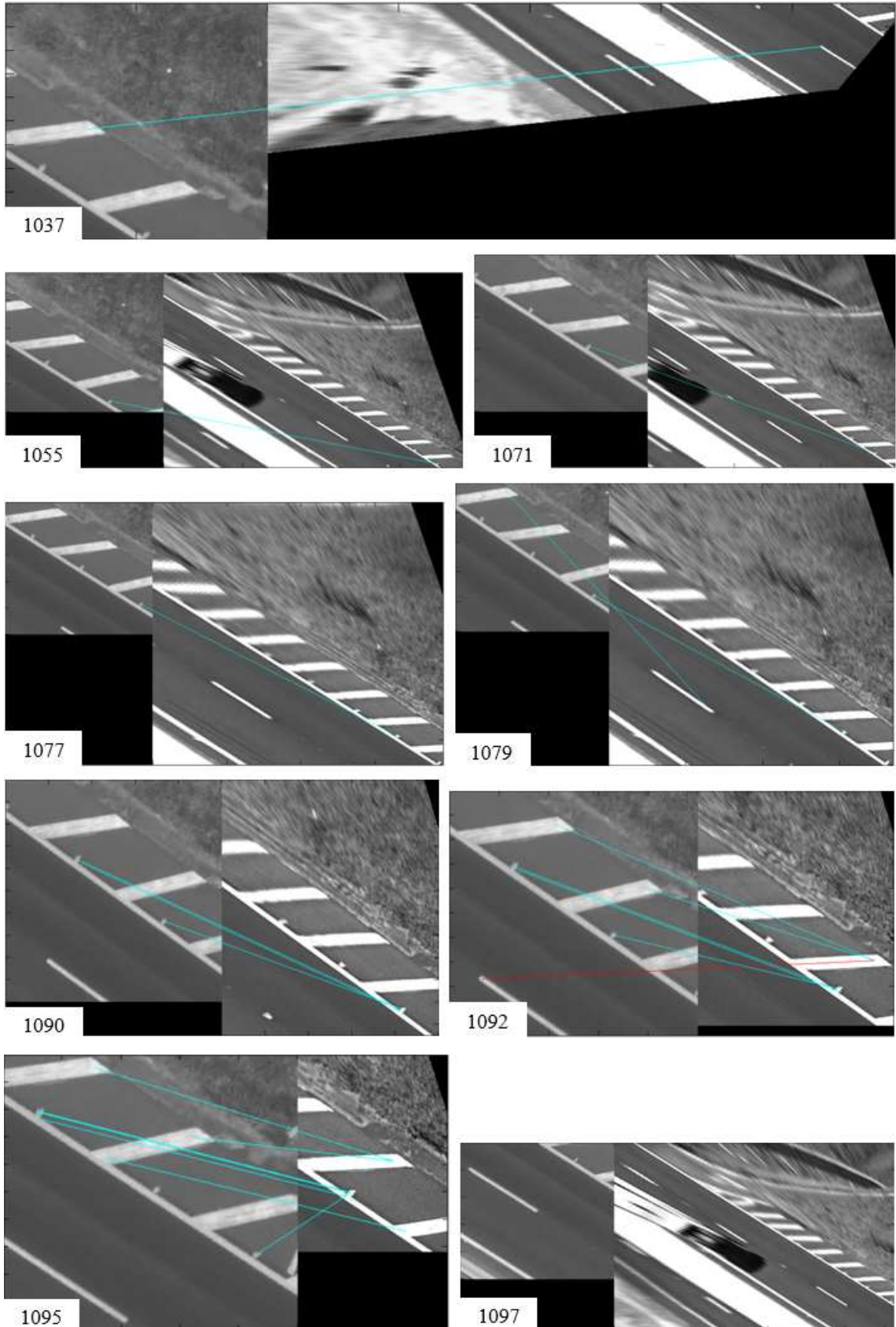
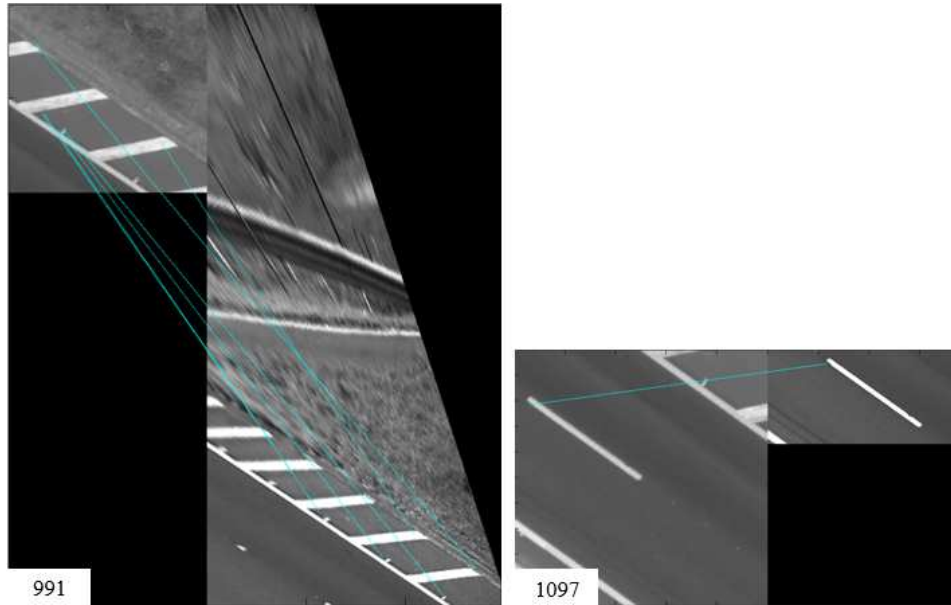


Figura 41. Resultado da busca por correspondência com o operador SIFT nos recortes projetados (par 8). Correspondências incorretas ainda não eliminadas.



Alguns recortes – 1173 e 1222 do par 4 (Figura 35), por exemplo – aparecem sem representação de correspondência, pois não retornaram nenhum resultado quando submetidos à busca com o SIFT. Destaca-se ainda a concentração de correspondências em locais próximos, como acontece nos recortes 1115 (par 6) e 1092 (par 7), entre outros, que não representa um resultado positivo para finalidade de correspondência.

Além da verificação baseada na inclinação das retas, alguns casos foram excluídos devido à possibilidade de se perceber visualmente que todas as correspondências representadas no par estavam incorretas. Os recortes 1211 e 1241 do par 3 (Figura 34) são exemplos destes casos. As demais correspondências incorretas foram excluídas no último processo de verificação, abordado na Etapa IV (item 5.2.4).

5.2.4 Etapa IV

Na quarta e última etapa, os pontos remanescentes tiveram suas posições calculadas nas imagens originais e estas foram plotadas como última forma de verificação das correspondências. As Figuras 42 a 48 apresentam o resultado final da classificação para cada um dos pares analisados. Como as correspondências estão todas localizadas sobre a área de rodovia, os resultados apresentam *zooms* da imagem aérea a fim de facilitar a visualização.

Figura 42. Resultado final da classificação de correspondências (par 2).

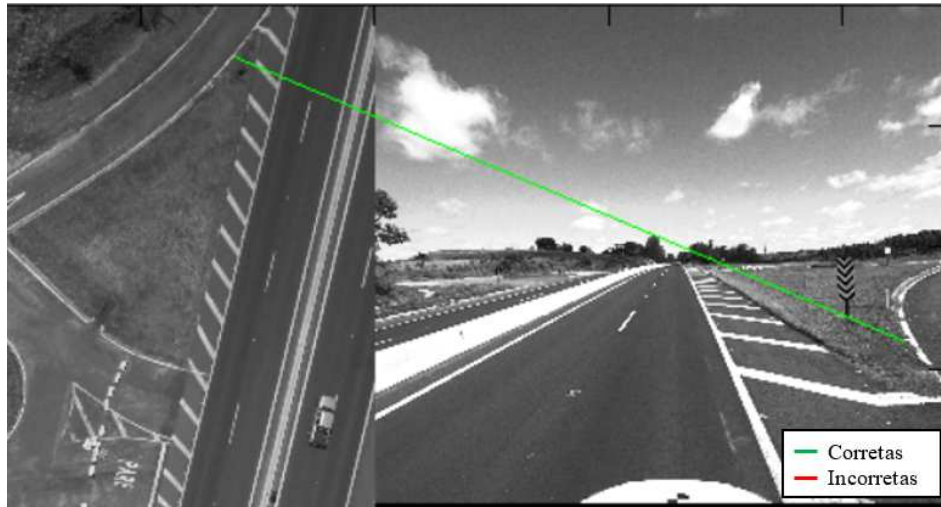


Figura 43. Resultado final da classificação de correspondências (par 3).

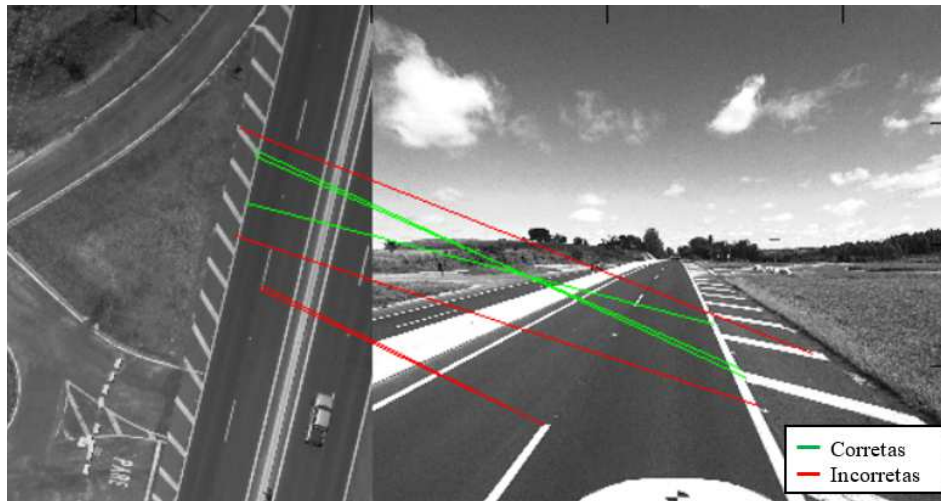


Figura 44. Resultado final da classificação de correspondências (par 4).

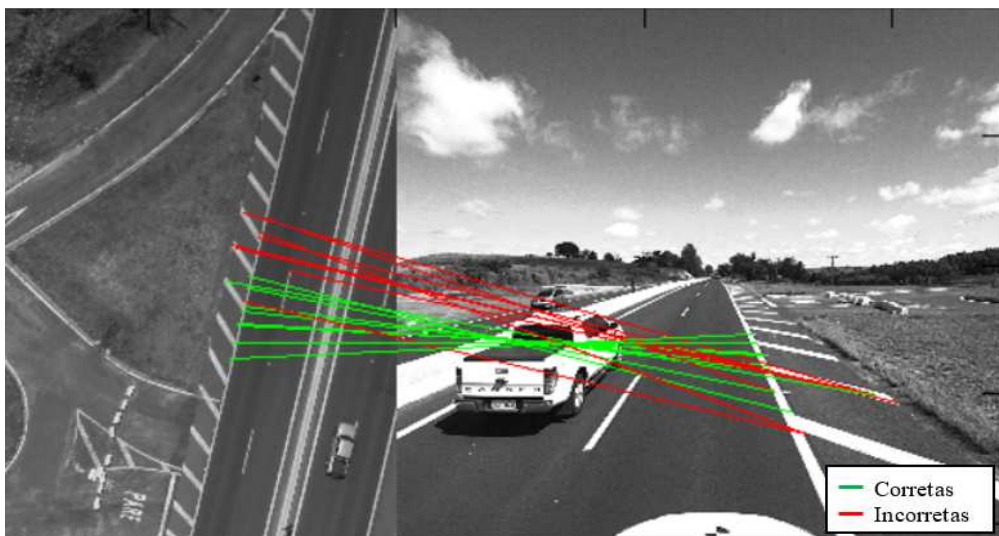


Figura 45. Resultado final da classificação de correspondências (par 5).

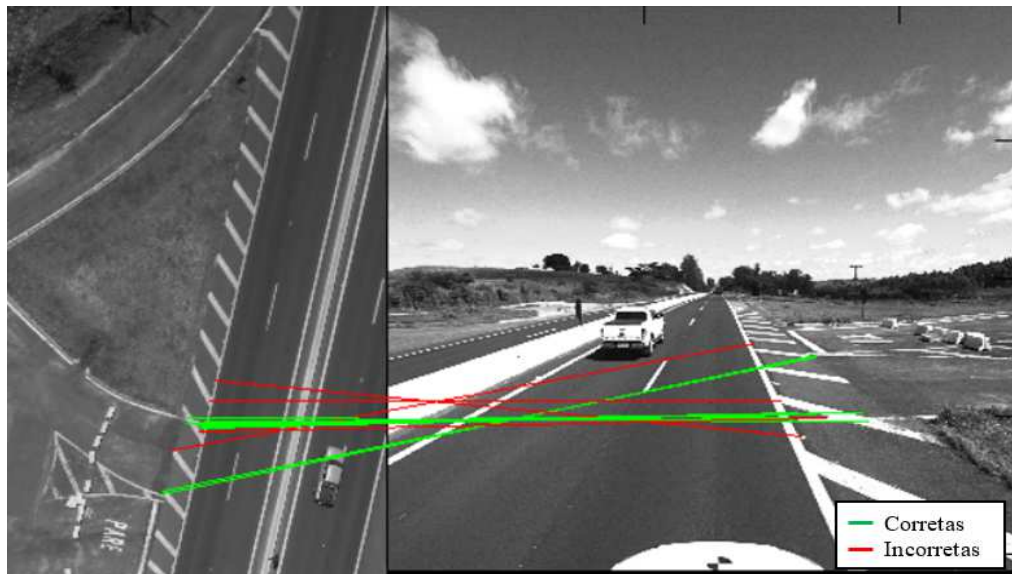


Figura 46. Resultado final da classificação de correspondências (par 6).

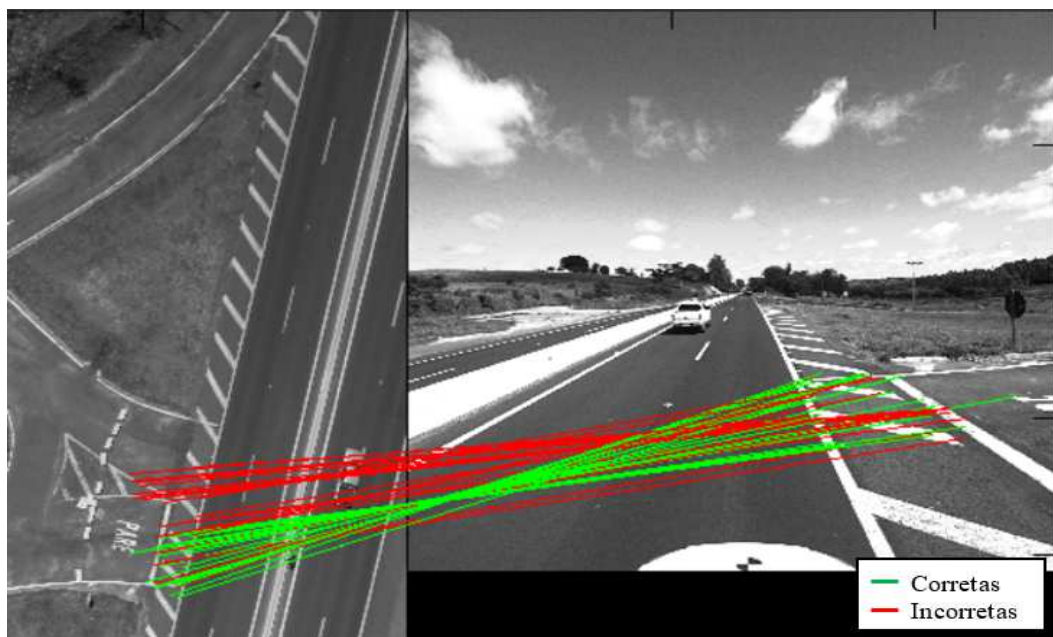


Figura 47. Resultado final da classificação de correspondências (par 7).

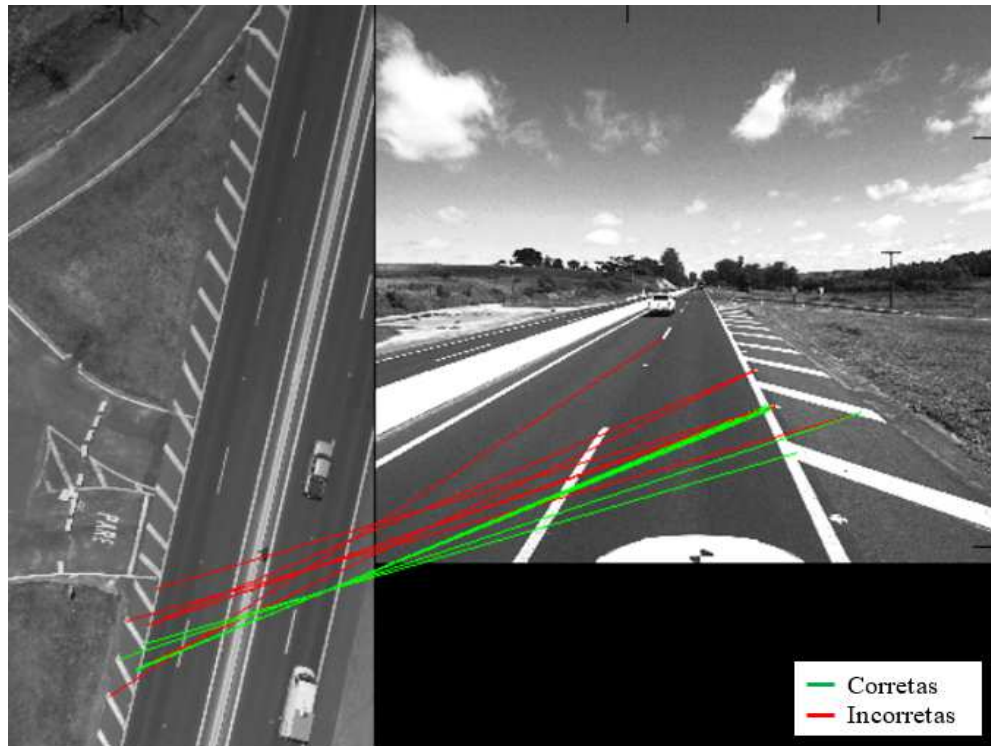
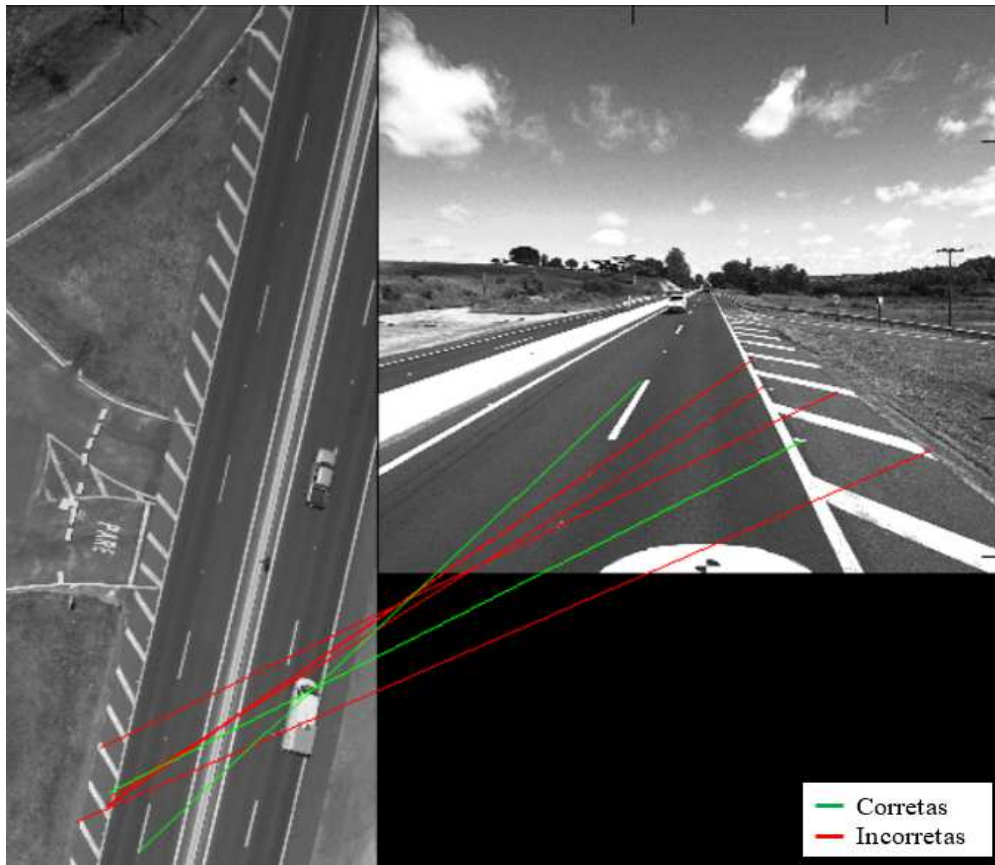
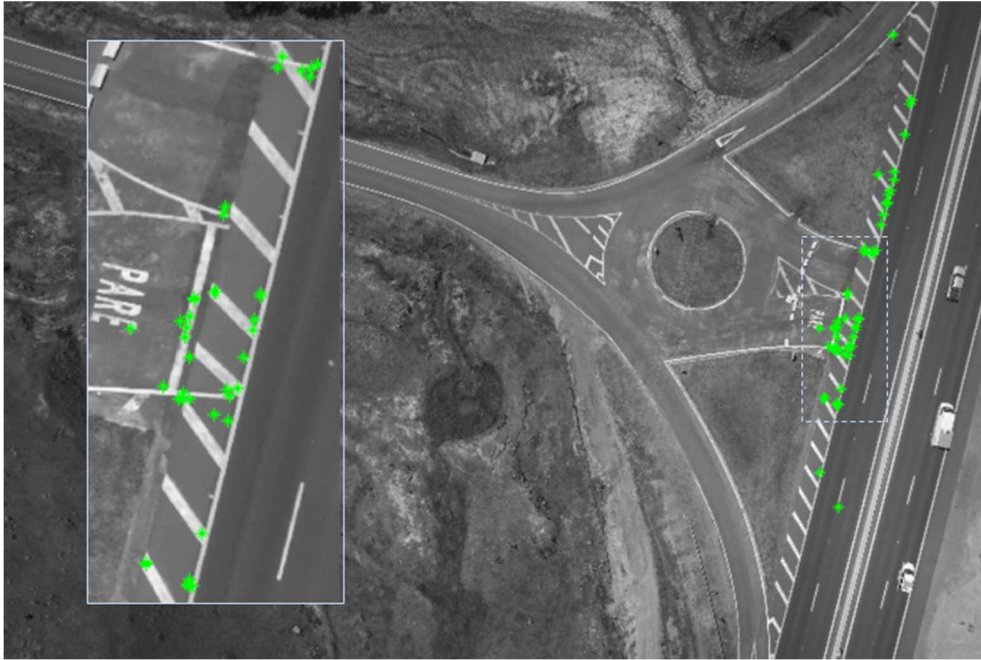


Figura 48. Resultado final da classificação de correspondências (par 8).



A Figura 49 apresenta todas as correspondências verdadeiras identificadas na imagem aérea, considerando todas as terrestres correspondentes, juntamente com um *zoom* da área destacada na imagem, a qual apresenta sobreposição de alguns pontos.

Figura 49. Resultado final da busca por correspondência na imagem aérea, representando apenas correspondências corretas, com *zoom* para região em destaque.

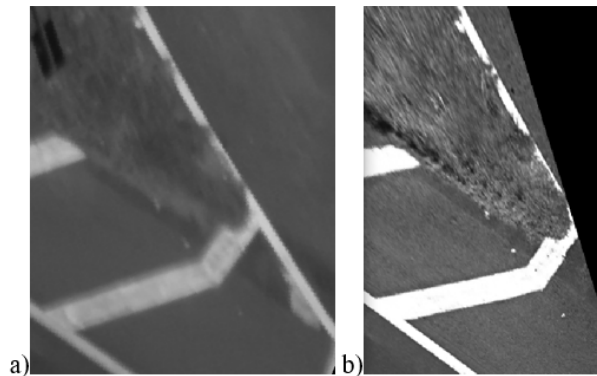


5.2.5 Outros resultados

5.2.5.1 Correspondência baseada em área

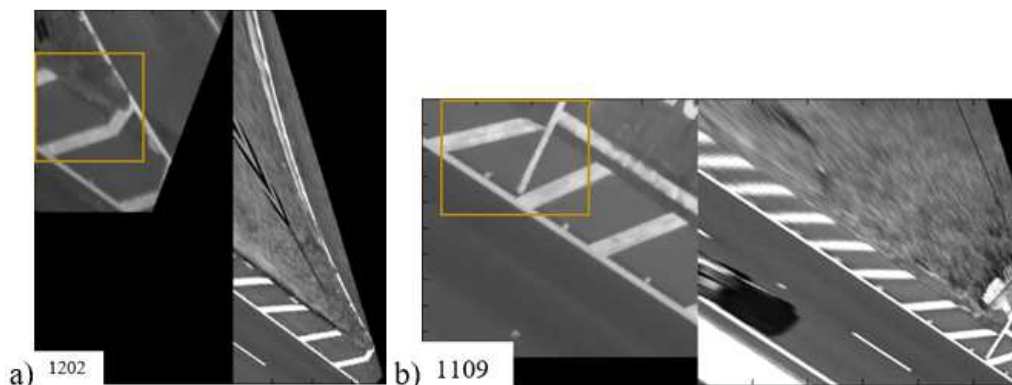
Para a retificação dos recortes terrestres e aéreos considera-se o mesmo plano de projeção – superfície horizontal em um plano médio do terreno – e também o mesmo GSD, acrônimo de *Ground Sample Distance*, de modo que se espera que eles sejam gerados na mesma escala ou em escalas próximas. A Figura 50 apresenta regiões correspondentes de um par de recortes retificados que reforça esta afirmação.

Figura 50. Regiões correspondentes em um par de recortes retificados, selecionadas em: a) recorte aéreo e b) recorte terrestre.



Com os recortes na mesma escala, espera-se que uma busca por correspondência baseada em área seja bem-sucedida. No entanto, a metodologia proposta neste trabalho garante que o ponto, localizado na imagem aérea, estará dentro da janela terrestre ao final da projeção, e não todo o recorte considerado. Deste modo, o recorte aéreo não pode servir como referência para o método de correspondência baseada em área, uma vez que não é possível garantir que o mesmo estará, em sua totalidade, presente no recorte terrestre. A fim de verificar a aplicação do método de correspondência baseada em área nos recortes retificados, dois pares foram selecionados e apenas uma pequena área foi considerada no caso aéreo, de modo que esta área estivesse completamente contida no respectivo recorte terrestre. A Figura 51 apresenta os pares de recortes selecionados para o experimento, com destaque para as áreas definidas como janela de referência em cada caso. A ordem apresentada traz o recorte aéreo seguido do terrestre, e o número que acompanha a figura indica o ponto de partida que originou aquele par de recortes.

Figura 51. Pares de recortes selecionados para o experimento, com destaque para a região definida como janela de referência no recorte aéreo.



Para a realização da busca por correspondência utilizou-se a função *normxcorr2* do Matlab, que calcula a correlação cruzada normalizada para as janelas fornecidas como referência e pesquisa. As Figuras 52 e 53 apresentam os resultados fornecidos pelo programa para os dois casos analisados, isto é, janelas de referência e pesquisa, sendo destacada na segunda a região definida como melhor correspondência pelo programa, e uma superfície gerada a partir da correlação calculada para cada posição da imagem de pesquisa, com destaque para o ponto que representa a correspondência verdadeira. Nas Figuras 52 e 53 os gráficos de superfície trazem os eixos X e Y representando, respectivamente, as colunas e linhas da imagem, e o eixo Z representando o valor da correlação calculada.

Figura 52. Resultado da correspondência baseada em área para o recorte 1202: a) janela de referência, b) janela de pesquisa, com destaque para a região que apresentou maior correlação, e c) superfície gerada a partir dos valores de correlação calculados para a imagem de pesquisa, com destaque para o ponto que representa a correspondência verdadeira.

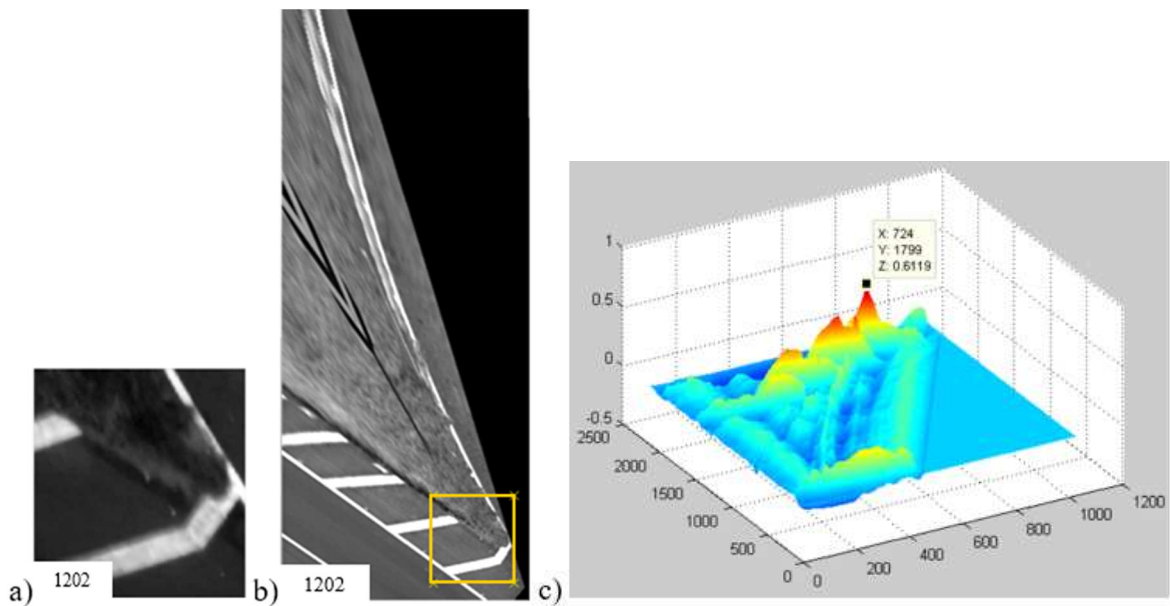
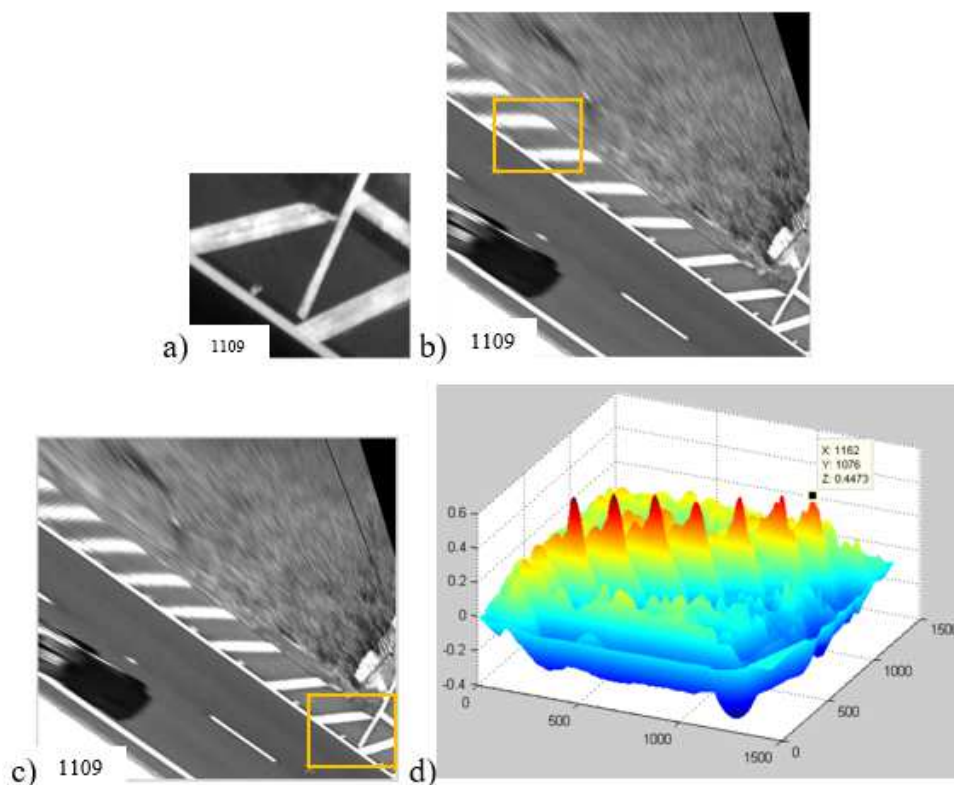


Figura 53. Resultado da correspondência baseada em área para o recorte 1109: a) janela de referência, b) janela de pesquisa, com destaque para a região que apresentou maior correlação, c) resultado esperado pelo programa, e d) superfície gerada a partir dos valores de correlação calculados para a imagem de pesquisa, com destaque para o ponto que representa a correspondência verdadeira.



A Figura 53.c) foi gerada plotando, na imagem de pesquisa, a posição que representa a correspondência verdadeira, identificada na Figura 53.d). Esta posição foi identificada nas superfícies manualmente.

Pode-se notar na Figura 53.d) a presença de muitos “picos” na superfície representada, que são os pontos que apresentaram alto valor de correlação. O mesmo não acontece no primeiro caso, no qual se identifica apenas um pico realmente sobressalente, como destacado na Figura 52.c), possibilitando a determinação precisa da região correspondente. Esta alta frequência de picos de correlação no segundo caso está relacionada com a dificuldade em se estabelecer correspondência devido à presença de padrões repetitivos na cena. A janela selecionada como referência possui um padrão de marcação rodoviário que aparece diversas vezes na imagem de pesquisa. Cada uma das vezes em que o padrão é encontrado a análise apresenta alta correlação. O padrão presente na janela de referência possui um traço a mais, proveniente de outra marcação rodoviária, o que permite diferenciá-lo visualmente dos demais. No entanto, esta pequena diferença não foi significativa para a busca automática, fazendo com que a função retornasse um resultado errado (Figura 53.b)). Outra

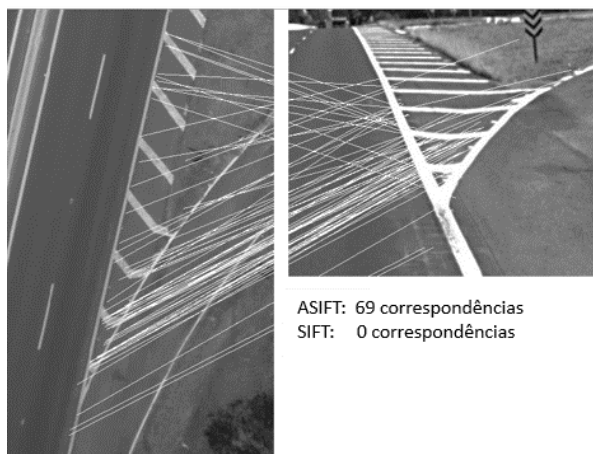
possível explicação seria uma pequena diferença de escala entre a janela de referência considerada no caso aéreo e a região de fato representada no caso terrestre.

Quanto à diferença de escala entre as imagens, o método de correspondência baseada em área não se aplica a esses casos. Já o problema causado por padrões repetitivos é bastante comum no processo de correspondência baseada em área, e pode ser solucionado, por exemplo, aumentando a janela de referência e incluindo mais detalhes à cena, facilitando assim a identificação da região. No entanto, o experimento foi suficiente para verificar a possibilidade de aplicação do método nos recortes retificados, uma vez que apresentem a mesma escala, e soluções relacionadas ao problema com padrões repetitivos não serão testadas neste trabalho.

5.2.5.2 *Métodos já existentes*

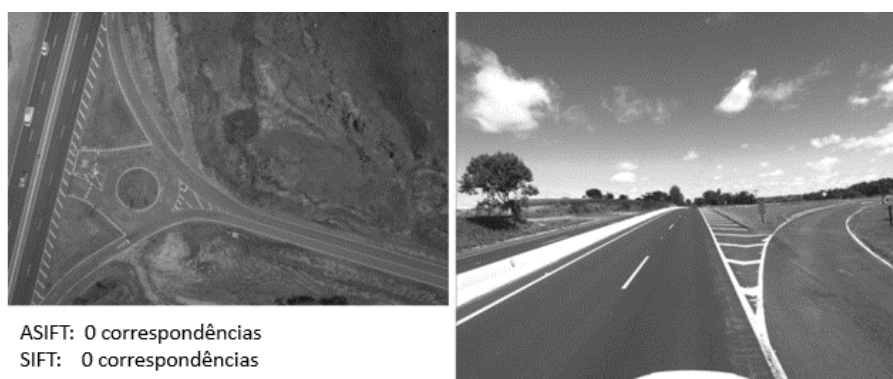
Em relação aos métodos já existentes, a opção considerada mais adequada para a correspondência entre imagens aéreas e terrestres são os operadores invariantes aos parâmetros de uma transformação afim, que, possivelmente, modelam mudanças de perspectiva. Uma versão modificada do SIFT que realiza esta modelagem foi proposta por Yu e Morel (2009) e recebe o nome de ASIFT, acrônimo de *Affine-SIFT*. Alguns testes foram realizados em uma versão demonstrativa deste operador, disponibilizada pelo *IPOL Journal: Image Processing On Line* (YU e MOREL, 2011) em conjunto com o SIFT. Os testes mostraram que o operador apenas consegue solucionar casos ideais, ou seja, quando já se insere recortes das imagens que representem regiões correspondentes, e esses recortes apresentem pouca variação de orientação e iluminação da cena, como o exemplo apresentado na Figura 54.

Figura 54. Resultado da busca por correspondência em um caso ideal, realizada com os operadores SIFT e ASIFT.



Para o caso apresentado na Figura 54, o ASIFT identifica um total de 69 correspondências, sendo 26 delas verdadeiras (aproximadamente 38% do total). Já quando se considera as imagens inteiras, ainda que haja uma alteração prévia na orientação dessas imagens para que o sentido considerado nelas seja o mesmo, o operador não consegue apresentar nenhuma solução (Figura 55).

Figura 55. Resultado da busca por correspondência nas imagens inteiras, realizada com os operadores SIFT e ASIFT.



Os recortes projetados não foram verificados com esta ferramenta, uma vez que não apresentam a variação de geometria esperada nas imagens para a utilização do operador ASIFT. Mais testes que verificam a aplicabilidade de operadores de interesse no contexto de correspondência de imagens aéreas e terrestres podem ser vistos em Gerke *et al.* (2016).

5.3 ANÁLISES

O experimento realizado de acordo com a metodologia proposta neste trabalho identificou, no conjunto de imagens, 182 correspondências, sendo 85 delas verdadeiras – 46,7% do total. As falsas correspondências foram identificadas de três maneiras: análise da inclinação das retas que ligam pontos homólogos; análise visual das correspondências plotadas nos recortes; e análise visual final – correspondências plotadas nas imagens originais.

Em relação à verificação baseada na inclinação das retas, pode-se concluir que o método teve pouca influência no resultado, excluindo apenas 7 correspondências incorretas. Este resultado era esperado, uma vez que a aplicação do método é bastante restrita, não se adequando a quase nenhum dos casos analisados. A primeira análise visual realizada excluiu um total de 12 correspondências incorretas. Embora este seja um valor pouco significativo, a identificação das falsas correspondências neste caso é muito simples de ser realizada, tornando viável a aplicação do método, uma vez que economiza, para os casos excluídos, o tempo de processamento das etapas futuras. Por último, a verificação final, realizada plotando os pares de pontos remanescentes nas imagens originais, excluiu um total de 78 correspondências incorretas e verificou que todas as demais representavam, realmente, correspondências verdadeiras. O Quadro 3 apresenta uma síntese dos resultados obtidos com a realização deste experimento.

Quadro 3. Síntese dos resultados obtidos com a metodologia proposta.

Par	Recortes gerados (pares)	Quant. Corresp. Identificadas	Quant. Corresp. Corretas
1	0	-	-
2	1	1	1
3	5	9	3
4	8	27	12
5	11	27	15
6	13	89	42
7	9	22	10
8	2	7	2
9	0	-	-
10	0	-	-
TOTAL	49	182	85

Pela análise do Quadro 3 e também da Figura 30, que apresenta a localização das imagens terrestres em relação à aérea, é possível notar uma influência desta localização na quantidade de correspondências encontradas, sendo maior a ocorrência na parte central da

imagem aérea. No Quadro 3 esta diferença é pouco evidente. Isto porque, como já mencionado, para os pares 5, 6 e 7, foi utilizado um valor mais restrito para o parâmetro *distRatio* considerado na busca com o SIFT, o que gerou uma diferença de 152 correspondências no total identificado quando se utilizou o mesmo valor do parâmetro para todas as imagens. A relação direta deste fator está na quantidade de recortes gerados, que também é bem maior para estes pares, mas não foi encontrada uma justificativa que explique esta maior ocorrência na parte central da imagem aérea. Pode-se observar ainda, que a quantidade de correspondências identificadas no par 6 foi bastante superior em relação aos demais casos, enquanto a quantidade de recortes gerados se manteve dentro do padrão. Estes comportamentos podem estar relacionados ao problema já mencionado por Heipke (1997), que qualifica a correspondência como um problema mal posto (*ill-posed*), ou mal condicionado (Galo, 2003).

De acordo com Heipke (1997), um problema é mal condicionado se não é possível garantir que existe uma solução, e que ela é única e/ou estável em relação a pequenas variações nos dados de entrada. Desta forma, a correspondência de imagens pode ser considerada um problema mal condicionado por vários motivos (HEIPKE, 1997): para um determinado ponto em uma imagem, por exemplo, um ponto correspondente pode não existir devido a oclusão; pode haver mais de uma possível correspondência devido à presença de padrões repetitivos na cena; e a solução pode ser instável em relação a ruídos. Ao observar o Quadro 3 e a Figura 29, é possível perceber que imagens muito semelhantes, como as terrestres 3 e 6, por exemplo, apresentam resultados muito diferentes para a correspondência, o que contribui com a afirmação apresentada por Heipke (1997). Outro fator que contribui com esta afirmação foi a utilização de valores diferentes para o parâmetro *distRatio*. Como mencionado, uma diferença de 0,1 neste parâmetro, considerado em apenas 3 pares de imagens, resultou em uma diferença de 152 correspondências no total identificado.

A respeito do método de correspondência baseado em área, o experimento realizado permitiu garantir a viabilidade da aplicação do método, apresentando apenas problemas já existentes neste tipo de correspondência, como é o caso dos padrões repetitivos, ou ainda, mostrando a inviabilidade da aplicação em casos que apresentam diferenças de escala. As Figuras 56 e 57 apresentam os resultados obtidos com a correlação e com o SIFT, respectivamente, para os dois recortes retificados que foram submetidos a estas análises.

Figura 56. Resultado da busca por correspondência no recorte 1202: a) correspondência baseada em área, com destaque para a região que apresentou maior correlação; e b) busca realizada com o SIFT, com a cor verde representando as correspondências corretas.

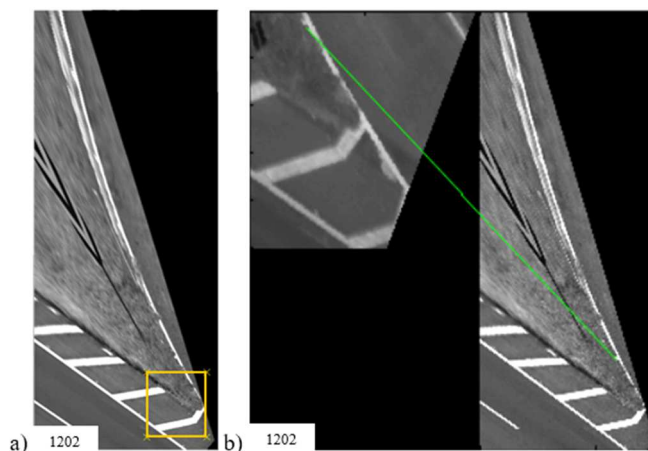
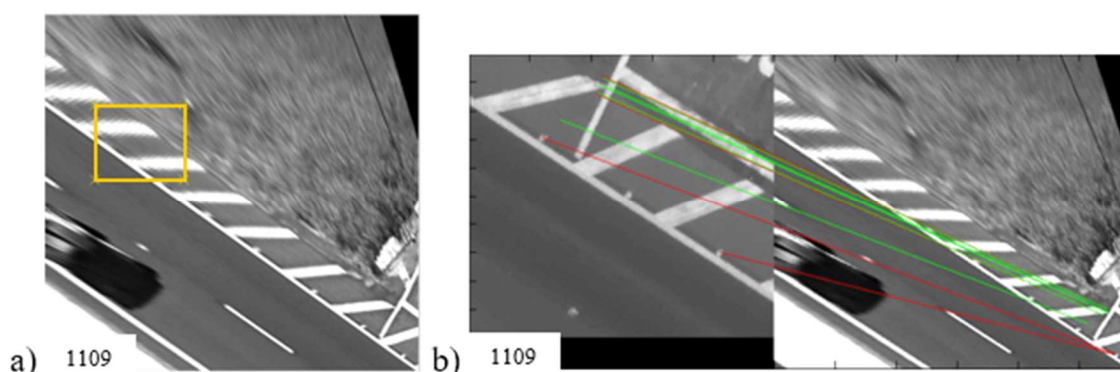


Figura 57. Resultado da busca por correspondência no recorte 1109: a) correspondência baseada em área, com destaque para a região que apresentou maior correlação; e b) busca realizada com o SIFT, com a cor verde representando as correspondências corretas e vermelho as incorretas.



Pelos resultados apresentados pode-se perceber que o SIFT, frequentemente, concentra a localização de correspondências em regiões comuns ou próximas, o que não é um resultado interessante para a finalidade de correspondência. Por outro lado, pode-se perceber que o método não é afetado pelo problema dos padrões repetitivos. Isto ocorre, pois, o operador considera características individuais das feições, como brilho, rotação e escala, e não o entorno, como acontece na correspondência baseada em área. No cenário rodoviário em que foi desenvolvido este trabalho, é muito comum a presença destes padrões, tornando interessante a utilização de um método baseado em características. Em contrapartida, o método de correspondência baseada em área pode buscar por pontos bem definidos, como os cantos selecionados na Etapa I da metodologia, por exemplo. Este pode ser um resultado interessante considerando um objetivo futuro deste trabalho, isto é, a fototriangulação

simultânea de imagens terrestres e aéreas. Desta forma, pode ser interessante o estudo de um método que combine as duas técnicas de correspondência mencionadas, agregando à metodologia as vantagens apresentadas por cada uma delas.

A respeito do teste com métodos já existentes, o experimento mostrou que os métodos disponíveis na literatura ainda não solucionam de forma satisfatória o problema de correspondência entre imagens aéreas e terrestres, apresentando solução apenas para casos ideais, como ao considerar recortes que já representem regiões correspondentes e com pouca variação na orientação e iluminação da cena. Isto permite concluir que a metodologia proposta neste trabalho, ainda que possua limitações, representa uma melhora significativa para o problema.

Por fim, uma pesquisa realizada na Universidade de Twente, Holanda, desenvolvida em paralelo com este trabalho, apresenta uma metodologia para a utilização da orientação de imagens aéreas em produtos terrestres que tem como uma das principais etapas a correspondência destas imagens. Para isto, o autor também realiza uma projeção das imagens terrestres como etapa inicial do processo. O resultado obtido por Jende *et al.* (2016) apresenta uma melhora na qualidade das correspondências identificadas por correlação em comparação com descritores clássicos de características. No entanto, para a verificação destas correspondências o autor realiza uma análise visual “bastante rigorosa”, de acordo com suas próprias palavras. Não foram mencionados os critérios utilizados nesta análise, mas se sabe que detectores de características como o SIFT e outros presentes na literatura não apresentam alta acurácia na identificação dos pontos, de modo que pontos homólogos identificados por estes métodos possam apresentar pequenos desvios em suas posições. Como no presente trabalho não foi verificada a aplicação de correspondência baseada em área para todo o conjunto de dados, não se pode apresentar conclusão a respeito de qual é o melhor método para o caso.

CAPÍTULO VI

CONCLUSÕES E RECOMENDAÇÕES

A integração de plataformas aéreas e terrestres representa uma solução promissora para diversas aplicações de mapeamento. Estes sistemas apresentam características complementares e, quando integrados, permitem a reconstrução completa de um objeto da superfície. Além desta característica de completude, a integração destes sistemas permite minimizar limitações presentes quando utilizados individualmente, representando assim uma potencial ferramenta para aplicações de mapeamento de maneira geral. No entanto, as abordagens existentes para correspondência entre imagens não apresentam solução satisfatória para este tipo de situação. A aquisição realizada a partir de diferentes pontos de vista gera imagens com perspectivas muito diferentes, o que se torna uma grande dificuldade no processo de correspondência.

Com esta motivação, o presente trabalho propõe uma metodologia para o estabelecimento de correspondência entre imagens aéreas e terrestres, baseada na alteração prévia da geometria destas imagens. Além disso, o trabalho traz um breve estudo sobre o estado da arte no que envolve este tipo de correspondência, e apresenta alguns experimentos com a finalidade principal de validar a metodologia proposta e verificar diferentes soluções para o problema, bem como analisar as que já estavam disponíveis na literatura. O primeiro experimento realiza a correspondência a partir de um método baseado em características, o SIFT, seguindo a metodologia proposta neste trabalho. O segundo, verifica a possibilidade de aplicação de um método de correspondência baseada em área nos resultados obtidos após a projeção das imagens. Por último, verifica-se a solução apresentada para o problema por métodos já existentes na literatura, sem que haja nenhum preparo prévio das imagens.

A respeito da metodologia proposta, pode-se dizer que a etapa de projeção é uma das mais importantes do processo, uma vez que permite reduzir as diferenças geométricas entre as imagens, principal problema no processo de correspondência tratado neste trabalho. Além disto, é importante que a extração dos pontos de partida seja realizada na imagem aérea, pois garante que as feições estarão presentes na imagem terrestre. O contrário não garantiria isto, uma vez que o nível de detalhe é inferior no caso aéreo, podendo não representar todo o conteúdo terrestre. As restrições impostas no programa também contribuíram de maneira significativa no processo de correspondência, reduzindo os erros envolvidos e a quantidade de correspondências identificadas. Este resultado é positivo considerando um objetivo futuro

deste trabalho, isto é, a fototriangulação simultânea de imagens aéreas e terrestres, que não requer uma grande quantidade de pontos.

Em relação aos experimentos, o primeiro deles valida a metodologia proposta com o descritor de características SIFT, identificando um total de 182 correspondências no conjunto de dados analisado, sendo 85 delas verdadeiras (aproximadamente 47% do total). O segundo experimento verifica a possibilidade de aplicação de um método de correspondência baseada em área, já que as imagens, após a projeção, apresentam geometrias bastante semelhantes. Neste segundo experimento foram testados apenas dois pares de imagens, um deles foi solucionado sem problemas e o segundo apresentou uma solução incoerente devido à presença de padrões repetitivos, problema este bastante comum neste tipo de correspondência, ou ainda, a uma possível diferença de escala existente nas imagens. O terceiro, e último, verifica a solução apresentada pela versão afim invariante do operador SIFT, o ASIFT, sendo esta abordagem considerada uma das mais adequadas para o problema entre os métodos já existentes. Os resultados obtidos mostraram que o operador consegue solucionar apenas casos ideais, isto é, quando se considera recortes das imagens que já representem regiões correspondentes, e esses recortes apresentam pouca variação de orientação e iluminação. Para um destes casos, o operador identificou 69 correspondências, sendo aproximadamente 38% delas verdadeiras. Quando se considera as imagens originais, inteiras, o operador não apresenta nenhuma solução.

Deste modo, pode-se concluir que o trabalho atendeu ao objetivo proposto, mostrando que é possível o estabelecimento de correspondência entre imagens aéreas e terrestre, e que este pode ser realizado tanto por métodos baseados em características como por métodos baseados em área, desde que haja uma alteração prévia na geometria destas imagens e que se tenham disponíveis POEs e MDT de boa qualidade. Os POEs podem ser fornecidos por georreferenciamento direto, o que é considerado um padrão na aquisição de imagens como as utilizadas neste trabalho, e o MDT pode ser pré-existente ou gerado a partir de imagens aéreas. Em relação à metodologia proposta, é importante mencionar que ela ainda apresenta muitas limitações, mas que, ainda assim, representa uma melhora considerável para o problema de correspondência tratado neste trabalho, uma vez que os métodos existentes só conseguem solucionar situações ideais e, na maioria das vezes, partindo de recortes ou regiões menores pré-identificadas nas imagens.

O método de correspondência baseada em área possui a vantagem de considerar pontos bem definidos, como cantos, por exemplo, o que representa um resultado positivo considerando um objetivo futuro desta técnica – a fototriangulação simultânea de imagens

terrestres e aéreas. Já o método baseado em características, em contrapartida, não é afetado pelos padrões repetitivos, problema muito comum no contexto rodoviário em que se insere este trabalho. Portanto, cada método agrega diferentes vantagens à abordagem proposta, não sendo possível afirmar qual deles é mais adequado para o caso, uma vez que o segundo experimento não foi realizado para todo o conjunto de dados. A integração sequencial da correspondência baseada em características, seguida de um refinamento por área, utilizando, por exemplo, o MMQ, é uma alternativa que sugere-se estudar no futuro.

Como recomendações para trabalhos futuros pode-se, ainda, mencionar a necessidade de uma estratégia que restrinja a identificação de pontos em padrões repetitivos, melhorando, assim, o cenário para a busca de correspondência baseada em área no contexto rodoviário. Jende *et al.* (2016) utiliza uma estrutura *kd-tree* para esta finalidade. Outras soluções podem ser estudadas. Além disto, a metodologia proposta considera uma estratégia de projeção que não é preparada para lidar com terrenos muito inclinados, podendo, com isto, comprometer o resultado da projeção naquela área. Estudos neste sentido podem ser desenvolvidos, uma vez que considerar apenas terrenos planos restringe muito a metodologia. Outra limitação que requer estudo é a verificação das correspondências, que neste caso é realizada de forma visual, impedindo a realização do método de maneira automática. Recomenda-se também o estudo de uma estratégia para lidar com a alta concentração de correspondências em locais próximos, identificada pelo SIFT.

Com base no resultado obtido no terceiro experimento, que verifica que um dos métodos existentes, no caso o ASIFT, consegue solucionar casos ideais em que se insere recortes de regiões correspondentes nas imagens, recomenda-se a utilização da estratégia de projeção dos pontos empregada neste trabalho como forma de restrição do espaço de busca e, em seguida, a aplicação de um operador afim invariante na busca por correspondência. Como os recortes, neste momento, já representarão regiões correspondentes, estes operadores podem apresentar solução, dispensando a necessidade de projeção desses recortes. Sugere-se ainda o estudo da viabilidade de aplicação da restrição epipolar nos recortes projetados, como forma de identificar correspondências incorretas. Por fim, recomenda-se a realização de uma fototriangulação simultânea entre imagens aéreas e terrestres, uma vez que a integração dos dados 3D gerados deve conferir alta acurácia e confiabilidade ao resultado. Este procedimento pode ser realizado considerando a abordagem proposta neste trabalho, mas ainda haverá a presença de muitas limitações que precisam ser solucionadas.

REFERÊNCIAS BIBLIOGRÁFICAS

- ABDULLAH, Q. A. et al. Photogrammetric platforms. In: MCGLONE, J. C. **Manual of Photogrammetry**. 6. ed. Bethesda: ASPRS, 2013. cap. 9. p. 719–784.
- ALCANTARILLA, P. F.; BARTOLI, A.; DAVISON, A. J. KAZE features. In: **Computer Vision–ECCV 2012**. Springer Berlin Heidelberg, p. 214-227, 2012.
- ALCANTARILLA, P. F.; SOLUTIONS, TrueVision. Fast explicit diffusion for accelerated features in nonlinear scale spaces. **IEEE Trans. Patt. Anal. Mach. Intell**, v. 34, n. 7, p. 1281-1298, 2011.
- ANDRADE, J. B. **Fotogrametria**. 2 ed. Curitiba: 2003.
- BAY, H.; ESS, A.; TUYTELAARS, T.; VAN GOOL, L. Speeded-up robust features (SURF). **Computer vision and image understanding**, v. 110, n. 3, p. 346-359, 2008.
- BERVEGLIERI, A. **Localização automática de pontos de controle em imagens aéreas baseada em cenas terrestres verticais**. Tese (Doutorado em Ciências Cartográficas) – Faculdade de Ciências e Tecnologia, Unesp, Presidente Prudente, 2014.
- BERVEGLIERI, A.; TOMMASELLI, A. M. G. Locating Control Points In Aerial Images With A Multi-Scale Approach Based On Terrestrial Image Patches. **The Photogrammetric Record**, v. 30, n. 149, p. 63-81, 2015.
- BISCHOF, H. e LEBERL, F. Digital image processing. In: MCGLONE, J. C. **Manual of Photogrammetry**. 6. ed. Bethesda: ASPRS, 2013. cap. 5. p. 451–5.
- CANON, S. 3D Least Squares Plane, **StackOverflow**, (2009). Disponível em: <<https://stackoverflow.com/questions/1400213/3d-least-squares-plane>>. Acesso em: nov/2016.
- CÓRDOBA, P. A. F. **Levantamento fotogramétrico de monumentos arquitetônicos para projetos de restauração – Estudo de caso: Palácio Itamaraty – Brasília, DF**. Dissertação (Mestrado em Estruturas e Construção Civil). Faculdade de Tecnologia. Universidade de Brasília. Brasília, DF, 2012.
- CRAMER, M.; STALLMANN, D.; HAALA, N. Direct georeferencing using GPS/inertial exterior orientations for photogrammetric applications. **International Archives of Photogrammetry and Remote Sensing**, v. 33, n. B3/1; PART 3, p. 198-205, 2000.
- FIETE, R. D. Elements of photogrammetry optics. In: MCGLONE, J. C. **Manual of Photogrammetry**. 6. ed. Bethesda: ASPRS, 2013. cap. 4. p. 359–450.
- FISCHLER, M. A.; BOLLES, R. C. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. **Communications of the ACM**, v. 24, n. 6, p. 381-395, 1981.

FONSECA, L. M. G.; MANJUNATH, B. S. Registration techniques for multisensor remotely sensed imagery. **PE & RS- Photogrammetric Engineering & Remote Sensing**, v. 62, n. 9, p. 1049-1056, sept. 1996.

FÖRSTNER, W. A feature based correspondence algorithm for image matching. **International Archives of Photogrammetry and Remote Sensing**, v. 26, n. 3, p. 150-166, 1986.

FÖRSTNER, W.; GULCH, E. A fast operator for detection and precise location of distinct points, corners and centres of circular features. In: **Proc. ISPRS intercommission conference on fast processing of photogrammetric data**. p. 281-305, 1987.

GALO, M. **Automação dos processos de correspondência e orientação relativa em visão estéreo**. Tese (Doutorado em Engenharia Elétrica). Faculdade de Engenharia Elétrica e de Computação. Universidade Estadual de Campinas. Campinas, SP, 2003.

GALO, M. **Introdução ao processamento e síntese de imagens: notas de aula**. UNESP, Departamento de Cartografia, 2015.

GALO, M. **Notas de aula de Fotogrametria**. UNESP, Departamento de Cartografia, 2014.

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

GEMAEL, C.; MACHADO A. M. L.; WANDRESEN R. **Introdução ao ajustamento de observações: aplicações geodésicas**. 2. ed. Paraná: Editora UFPR, 2015.

GERKE, M.; NEX, F.; JENDE, P. Co-Registration of Terrestrial and UAV-Based Images – Experimental Results. In: **Proc. EuroCOW the European Calibration and Orientation Workshop**. Lausanne, Switzerland, feb. 2016.

GRUEN, A. Development and status of image matching in photogrammetry. **The Photogrammetric Record**, v. 27, n. 137, p. 36-57, mar. 2012.

HABIB, A.; MORGAN, M.; LEE, Y. Bundle adjustment with self-calibration using straight lines. **The Photogrammetric Record**, v. 17, n. 100, p. 635-650, oct. 2002.

HARRIS, C.; STEPHENS, M. A combined corner and edge detector. In: **Alvey vision conference**. p. 50, 1988.

HEIPKE, C. Automation of interior, relative, and absolute orientation. **ISPRS journal of photogrammetry and remote sensing**, v. 52, n. 1, p. 1-19, feb. 1997.

JAUREGUI, M.; VÍLCHEZ, J.; CHACÓN, L. A procedure for map updating using digital mono-plotting. **Computers & geosciences**, v. 28, n. 4, p. 513-523, 2002.

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

JENDE, P.; PETER, M.; GERKE, M.; VOSSELMAN, G. Advanced tie feature matching for the registration of mobile mapping imaging data and aerial imagery. **International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences**, v. 41, 2016.

KÖTHE, U. Edge and junction detection with an improved structure tensor. In: **Joint Pattern Recognition Symposium**. Springer Berlin Heidelberg, 2003. p. 25-32.

KOVESI, P. **Autor do código implementado para Matlab do operador Harris**. Department of Computer Science & Software Engineering. The University of Western Australia (2002). Disponível em: <<http://web.engr.illinois.edu/~slazebni/spring16/harris.m>>. Acesso em: fev/2017.

KRAUS, K. **Photogrammetry: geometry from images and laser scans**. Second edition. Germany: Walter de Gruyter, 2007.

LEVI, G.; HASSNER, T. LATCH: learned arrangements of three patch codes. In: **2016 IEEE Winter Conference on Applications of Computer Vision (WACV)**. IEEE, p. 1-9. 2016.

LOPES R. F. **Formação de imagens multiespectrais por meio de fusão de imagens adquiridas por múltiplas câmaras**. Dissertação (Mestrado em Ciências Cartográficas) – FCT/UNESP, Presidente Prudente, 2010.

LOWE, D. G. Distinctive image features from scale-invariant keypoints. **International journal of computer vision**, v. 60, n. 2, p. 91-110, 2004.

LOWE, D. G. Object recognition from local scale-invariant features. In: **Computer vision, 1999. The proceedings of the seventh IEEE international conference on**. Ieee, p. 1150-1157, 1999.

LUXEN, M. **Autor do código implementado para Matlab do operador Förstner**. Department of Photogrammetry, Institute of Geodesy and Geoinformation, University of Bonn, Germany, 2005.

MAIR, E.; HAGER, G. D.; BURSCHKA, D.; SUPPA, M.; HIRZINGER, G. Adaptive and generic corner detection based on the accelerated segment test. In: **European conference on Computer vision**. Springer Berlin Heidelberg, 2010. p. 183-196.

MAYER, H. et al. Basic computer vision techniques. In: MCGLONE, J. C. **Manual of Photogrammetry**. 6. ed. Bethesda: ASPRS, 2013. cap. 6. p. 517–583.

MIKHAIL, E. M.; BETHEL, J. S.; MCGLONE, J. C. **Introduction to Modern Photogrammetry**. New York: John Wiley & Sons, Inc., 2001.

MIKOLAJCZYK, K.; SCHMID, C. Scale & affine invariant interest point detectors. **International journal of computer vision**, v. 60, n. 1, p. 63-86, oct. 2004.

MITISHITA E. A., **Monorestituição digital de aerofotos, associada com sistema de computação gráfica C.A.D., para fins de mapeamento na área florestal**. Tese (Doutorado em Ciências Florestais). Setor de Ciências Agrárias. Universidade Federal do Paraná. Curitiba, PR, 1997.

MOLINA, P.; BLÁZQUEZ, M.; SASTRE, J.; COLOMINA, I. A method for simultaneous aerial and terrestrial geodata acquisition for corridor mapping. **The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences**, v. 40, n. 1, p. 227, 2015.

MOLINA, P.; BLÁZQUEZ, M.; SASTRE, J.; COLOMINA, I. Precision analysis of point-and-scale photogrammetric measurements for corridor mapping: preliminary results. **ISPRS Annals of Photogrammetry, Remote Sensing & Spatial Information Sciences**, 2016.

MORAVEC, H. P. Towards automatic visual obstacle avoidance. In: **International Conference on Artificial Intelligence (5th: 1977: Massachusetts Institute of Technology)**. 1977.

MOREL, J.; YU, G. ASIFT: A new framework for fully affine invariant image comparison. **SIAM Journal on Imaging Sciences**, v. 2, n. 2, p. 438-469, 2009.

MUGNIER, C. J. et al. The mathematics of photogrammetry. In: MCGLONE, J. C. **Manual of Photogrammetry**. 6. ed. Bethesda: ASPRS, 2013. cap. 3. p. 235–357.

NEX, F.; GERKE, M.; REMONDINO, F.; PRZYBILLA, H. J.; BÄUMKER, M.; ZURHORST, A. ISPRS Benchmark for multi-platform photogrammetry. **ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences**, v. 2, n. 3, p. 135, 2015.

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

OLIVEIRA, R. A. **Geração automática de modelos digitais de superfície utilizando múltiplas imagens**. Dissertação (Mestrado em Ciências Cartográficas). Faculdade de Ciências e Tecnologia. Universidade Estadual Paulista Júlio de Mesquita Filho. Presidente Prudente, SP, 2013.

PARKER, J. A.; KENYON, R. V.; TROXEL, D. E. Comparison of interpolating methods for image resampling. **IEEE Transactions on medical imaging**, v. 2, n. 1, p. 31-39, 1983.

REMONDINO, F. Detectors and descriptors for photogrammetric applications. **International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences**, v. 36, n. 3, p. 49-54, 2006.

REMONDINO, F.; BARAZZETTI, L.; NEX, F.; SCAIONI, M.; SARAZZI, D. UAV photogrammetry for mapping and 3d modeling—current status and future perspectives. **International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences**, v. 38, n. 1, p. C22, 2011.

RODEHORST, V.; KOSCHAN, A. Comparison and evaluation of feature point detectors. In: **Proc. 5th International Symposium Turkish-German Joint Geodetic Days "Geodesy and Geoinformation in the Service of our Daily Life"**, Berlim, Alemanha. 2006.

ROSTEN, E.; DRUMMOND, T. Machine learning for high-speed corner detection. In: **Computer Vision—ECCV 2006**. Springer Berlin Heidelberg, p. 430-443, 2006.

RUBLEE, E.; RABAUD, V.; KONOLIGE, K.; BRADSKI, G. ORB: an efficient alternative to SIFT or SURF. In: **Computer Vision (ICCV), 2011 IEEE International Conference on**. IEEE, p. 2564-2571, 2011.

SMITH, S. M.; BRADY, J. M. SUSAN—a new approach to low level image processing. **International journal of computer vision**, v. 23, n. 1, p. 45-78, 1997.

TEMIZ M. S.; KÜLÜR S. Rectification of digital close range images: sensor models, geometric image transformations and resampling. **Int. Arch. Photogram. Rem. Sens. Spatial Inform. Sci.** Beijing, v. 37, 2008.

WOLF, P. R. **Elements of Photogrammetry: With Air Photo Interpretation and Remote Sensing**. Second Edition. Boston: McGraw-Hill, 1983

YAN, K. Image mosaic using SIFT, **MathWorks**, (2011). Disponível em: <<http://www.mathworks.com/matlabcentral/fileexchange/30849-image-mosaic-using-sift?focused=5180302&tab=function>>. Acesso em: fev/2017.

YU, G.; MOREL, J. A fully affine invariant image comparison method. In: **Acoustics, Speech and Signal Processing, 2009. ICASSP 2009. IEEE International Conference on**. IEEE, p. 1597-1600, 2009

YU, G.; MOREL, J. ASIFT: An Algorithm for Fully Affine Invariant Comparison, **Image Processing On Line**, 1 (2011). Disponível em: <<http://dx.doi.org/10.5201/ipol.2011.my-asift>>. Acesso em: jul/2016.