

**UNIVERSIDADE FEDERAL DO RIO DE JANEIRO**

**ESCOLA POLITÉCNICA**

**DEPARTAMENTO DE ENGENHARIA ELETRÔNICA E DE  
COMPUTAÇÃO**

**AUXÍLIO A ATERRISSAGEM DE UM HELICÓPTERO NÃO TRIPULADO  
PELA VISÃO**

**Rubens Felipe Quintanilha de Carvalho**

orientador: Prof. Eduardo Antônio Barros da Silva, Ph.D.

---

Rubens Felipe Quintanilha de Carvalho

Aprovada por:

---

Prof. Eduardo Antônio Barros da Silva, Ph.D.

---

Prof. José Gabriel Rodríguez Carneiro Gomes, Ph.D.

---

Prof. Gelson Vieira Mendonça, Ph.D.

RIO DE JANEIRO, RJ - BRASIL  
DEZEMBRO DE 2007

CARVALHO, RUBENS FELIPE QUINTA-  
NILHA DE

AUXÍLIO A ATERRISSAGEM DE UM  
HELICÓPTERO NÃO TRIPULADO PELA  
VISÃO[Rio de Janeiro] 2007

XII, 61 p. 29,7 cm (DEL/UFRJ, Engenharia Eletrônica e da Computação, 2007)

Universidade Federal do Rio de Janeiro,  
DEL

1. Visão computacional
2. Aeronave não tripulada
- 3.UAV

I. DEL/UFRJ II. Título ( série )

## DEDICATÓRIA

Dedico este trabalho aos meus pais e à minha irmã por caminharem sempre ao meu lado.

Aos meus avós pela lição de vida passada.

À Juliana por ser minha companheira de vida e estudos.

## AGRADECIMENTOS

Agradeço aos integrantes do projeto e meus orientadores no estágio, Patrick Fabiani e Roger Mampey. Gostaria de agradecer também a Guy Le Besnerais pelos conhecimentos transmitidos. A Mohsen Ardabilian, meu tutor na Ecole Centrale de Lyon, pelo apoio durante o estágio. Finalmente agradeço à toda equipe RESSAC, Alain Piquereau e Vincent Fuertes.

Agradeço a todos os professores do DEL/UFRJ por consolidarem meus conhecimentos em engenharia. Gostaria de agradecer em especial ao professor Luiz Wagner por despertar minha vontade para a pesquisa e ao meu orientador Eduardo Antônio Barros pela paciência, confiança e ensinamentos.

Resumo do projeto final apresentado ao DEL/UFRJ como parte dos requisitos necessários para a obtenção do grau de Bacharel em Engenharia Eletrônica e da Computação

## AUXÍLIO A ATERRISSAGEM DE UM HELICÓPTERO NÃO TRIPULADO PELA VISÃO

Rubens Felipe Quintanilha de Carvalho  
Dezembro/2007

Orientador : Eduardo Antônio Barros da Silva  
Programa: Engenharia Eletrônica e da Computação

Este trabalho se enquadra na parte de visão para a autonomia de um helicóptero que é equipado de sensores óticos. A pesquisa tem por missão o auxílio à aterrissagem do helicóptero não tripulado de forma automática, ou seja, realizar a escolha de uma região de aterrissagem, principalmente uma região plana, a partir de uma câmara embarcada. Para realizar esta tarefa com sucesso o dispositivo deve perceber seu ambiente imediato (particularmente os eventuais obstáculos presentes na zona de aterrissagem) e conhecer precisamente suas evoluções em relação ao solo, ou seja, realizar uma espécie de medida do seu próprio movimento.

Para solucionar o problema proposto, o helicóptero efetua uma trajetória de exploração ao mesmo tempo em que capta imagens do terreno. Em seguida essas imagens são tratadas seguindo técnicas de visão estéreo. Estas técnicas analisam o problema da reconstrução da informação tridimensional de objetos a partir de um par de imagens.

Por fim utilizamos as informações extraídas do processamento de cada par de imagens para projetar sobre um mapa do terreno os obstáculos detectados. Quando a câmara embarcada captura uma imagem a câmara associa à imagem, um estado relativo ao posicionamento do helicóptero no momento exato da captura. Essa informação da posição possibilita a projeção dos resultados sobre o mapa do terreno na posição correta em que cada obstáculo se encontra, corrigindo também as deformações oriundas da inclinação do helicóptero.

---

## SUMÁRIO

<b>1</b>	<b>Introdução</b>	<b>1</b>
1.1	Resumo do projeto RESSAC . . . . .	1
1.2	O projeto RESSAC . . . . .	2
1.2.1	Contexto . . . . .	2
Domínios científicos envolvidos	. . . . .	2
Domínios de aplicação	. . . . .	3
1.2.2	Estado da arte: equipamentos na França e no mundo . . . . .	4
1.3	Objetivos e interesses do projeto RESSAC . . . . .	6
1.3.1	Objetivos detalhados . . . . .	6
1.3.2	Interesses científicos e técnicos . . . . .	7
1.4	Organização do relatório . . . . .	9
<b>2</b>	<b>Visão</b>	<b>10</b>
2.1	Visão computacional . . . . .	10
2.2	Escolha de pontos da imagem . . . . .	12
2.2.1	Extração de pontos de interesse . . . . .	12

2.2.2	Associação de uma imagem à outra . . . . .	15
	Rastreamento KLT (Kanade, Lucas e Tomasi) . . . . .	16
	Correlação Ida e Volta . . . . .	18
2.3	Modelo global robusto . . . . .	21
2.3.1	Estimação robusta . . . . .	21
	RANSAC . . . . .	22
	LMedS (Least Median of Squares) . . . . .	23
	M-estimadores . . . . .	24
	IRLS . . . . .	24
2.3.2	Modelos de transformação vislumbrados . . . . .	25
	Afim . . . . .	25
	Quadrático 2D . . . . .	26
2.4	Tipos de movimento na imagem . . . . .	27
2.4.1	Campo de movimento real e aparente . . . . .	27
2.4.2	Campo de movimento aparente e estimado . . . . .	29
2.5	Processamento estéreo . . . . .	30
2.6	Detecção de zonas de relevo . . . . .	31
2.6.1	Modo esparsos . . . . .	31
2.6.2	Modo denso . . . . .	32
2.7	Fluxo ótico . . . . .	33
2.8	Detecção de Odobez-Bouthemy . . . . .	34
	Estimação do movimento . . . . .	35
	Algoritmo de detecção de movimento . . . . .	35
2.9	Trabalhos realizados . . . . .	39

<b>3 A cartografia do terreno</b>	<b>41</b>
3.1 O helicóptero . . . . .	41
3.2 Movimentos 3D do helicóptero . . . . .	42
3.2.1 Situações . . . . .	44
3.2.2 Os ângulos de Euler . . . . .	46
3.2.3 Projeções no solo . . . . .	47
3.3 Trabalhos realizados . . . . .	50
<b>4 Resultados</b>	<b>52</b>
4.1 Síntese dos resultados . . . . .	52
4.2 Análise dos filmes . . . . .	54
4.3 Conclusão . . . . .	60

---

## LISTA DE FIGURAS

1.1	Possíveis aplicações	3
1.2	O helicóptero Vigilant	7
1.3	Helicóptero Yamaha RMAX usado no projeto	8
2.1	Câmara posicionada sob a aeronave	10
2.2	Trajetória de exploração de uma zona (filme: Esperce80)	11
2.3	Idéia básica do detector de Harris	13
2.4	Extração de pontos de Harris	15
2.5	Rastreamento de primitivas	16
2.6	Ilustração de uma transformação afim.	25
2.7	Ilustração da transformação quadrática.	26
2.8	Movimento real e aparente num sistema ótico [Gra03]	28
2.9	À esquerda: informação inválida. À direita: caráter estéreo suficiente	30
2.10	Razão base (B) sobre altura (H) dos pares de imagens sucessivas da seqüência Esperce80	30
2.11	Imagens da seqüência Esperce80	32
2.12	Ilustração do problema de abertura.	34

2.13	Classificação de medidas . . . . .	39
3.1	Movimentos 3D de uma aeronave . . . . .	42
3.2	Helicóptero em posição estacionária . . . . .	43
3.3	Helicóptero em deslocamento . . . . .	43
3.4	Ângulos de Euler . . . . .	47
3.5	Projeção da câmera no solo . . . . .	48
3.6	Movimento em torno do eixo Z . . . . .	48
3.7	Movimento em torno do eixo X e projeção no solo . . . . .	49
3.8	Movimento em torno do eixo Y e projeção no solo . . . . .	50
4.1	Estimação do fluxo ótico residual entre as imagens 9 e 10 . . . . .	53
4.2	À esquerda a imagem de DFD, à direita a imagem com o resultado da detecção . . . . .	53
4.3	Detecção sobre o par de imagens 183-84, do filme Caylus000 . . . . .	55
4.4	À esquerda: par de imagens 177-178. À direita: par de imagens 192-193, filme Caylus002 . . . . .	56
4.5	Campo de teste Caylus . . . . .	57
4.6	Cartografia de Caylus001 . . . . .	58
4.7	Cartografia de Caylus001 com zoom . . . . .	59



# CAPÍTULO 1

## INTRODUÇÃO

Durante meu projeto final realizei um estágio no centro francês de pesquisas aeroespaciais, ONERA (Office National d'Etudes et Recherches Aérospatiales). As pesquisas foram feitas no Centro de Toulouse no sul da França, mais precisamente no departamento de comandas de sistemas e dinâmica do vôo, DCSD (Département Commande de Systèmes et Dynamique du vol).

O projeto final se enquadra na parte de decisão baseada em visão do projeto RESSAC (sigla para: busca e resgate por um sistema autônomo cooperante)<sup>1</sup>. Meu tema de pesquisa tem por missão o auxílio à aterrissagem de um helicóptero não tripulado, ou seja, realizar a escolha de zonas propícias à aterrissagem, zonas planas, através de uma câmara embarcada.

### 1.1 Resumo do projeto RESSAC

Será transcrito abaixo um resumo do projeto RESSAC segundo está escrito na página web do projeto<sup>2</sup>.

“As aeronaves não tripuladas atuais têm pouca capacidade de decisão embarcada. Sua implementação impõe grandes dificuldades para os operadores. Diversas mis-

<sup>1</sup>Recherche et Sauvetage par Système Autonome Coopérant

<sup>2</sup><http://www.cert.fr/dcsd/RESSAC>

sões em meio hostil podem ser consideradas como provas de fogo para as aeronaves inteligentes. As missões de busca e resgate de pessoas em meio hostil fazem parte de um desafio maior. Os possíveis usuários não fazem idéia das potencialidades das aeronaves autônomas pela falta de demonstrações.

Este projeto se propõe a avançar tecnologicamente abrindo inúmeras perspectivas em matéria de veículos não tripulados, principalmente permitindo ao par operador/veículo interromper um elo de comunicação. O objetivo é produzir uma demonstração das capacidades de autonomia de controle de vôo, desempenho da missão, aquisição e processamento de informação e decisão. Isto requer o desenvolvimento de ferramentas avançadas, harmonizadas e integradas às funcionalidades do protótipo experimental, para finalmente validá-las numa demonstração de vôo e em simulação sobre um cenário de busca e resgate de pessoas em meio hostil.

O principal interesse científico do projeto RESSAC é gerar pesquisas sobre os diferentes componentes de um sistema de controle autônomo da aeronave: controle do vôo, percepção e decisão autônoma. Buscamos dessa forma, na junção das linhas científicas algumas combinações inovadoras de métodos provenientes de domínios complementares permitindo rupturas tecnológicas quando se trata de percepção e decisão autônoma. De um ponto de vista técnico, almeja-se produzir uma tecnologia de vôo e ferramentas eficazes, tanto em metodologia quanto em programas, para futuros projetos ou contratos de pesquisa.”

## **1.2 O projeto RESSAC**

### **1.2.1 Contexto**

#### **Domínios científicos envolvidos**

Entre os domínios científicos envolvidos o principal é o controle de sistemas, integrando as dificuldades do comando automático, do processamento de informações e decisão - decisão embarcada - e auxílio à decisão do operador, unindo conhecimentos dos sensores óticos e radares.

O progresso tecnológico envolvendo diferentes elementos, da cadeia de aquisição

ao processamento da informação para a decisão embarcada e auxílio a operação, tais como avionica de bordo, estação solo e carga útil são promissores.

## Domínios de aplicação

O uso de aeronaves não tripuladas equipadas com sensores tem aplicações civis e militares inovadoras. É sem dúvida uma alternativa tecnológica para os satélites de observações. Ainda mais promissora seria a possibilidade de empregar esses dispositivos em contextos mais complexos ou inserí-los em redes ativas de informação e decisão.

Entre as possíveis aplicações, Figura 1.1, estão as missões atuais ou futuras de sistemas aeronáuticos não tripulados, dotados de capacidades de autonomia de decisão.



Figura 1.1: Possíveis aplicações

Um grande número de missões em meio hostil para os quais atualmente é preciso arriscar a vida de pessoas poderiam ser reconsideradas caso seja demonstrada a viabilidade das aeronaves dotadas de capacidade de decisão suficiente. As missões de vigilância são um bom exemplo, visto que são leves no que tange a pilotagem e orientação, pois o material humano pode fraquejar principalmente em longas missões. As missões de busca e salvamento de pessoas num meio hostil são um exemplo

ainda mais interessante: as aeronaves não tripuladas evitariam a exposição de um grande número de pessoas com o objetivo de salvar algumas outras.

O problema da inserção destes dispositivos no tráfego aéreo civil está longe de ser resolvido, pois ele implica em alguns desafios tecnológicos e regulamentares, tanto para a autonomia de decisão embarcada quanto para os sensores anticolisão, dificultando até mesmo a gestão do espaço aéreo seja ele civil ou militar.

A União Européia estando ciente deste problema financia desde 2001 o grupo temático UAV-NET, incumbindo a companhia israelense IAI de realizar os primeiros estudos visando solucionar os problemas de segurança e regulamentação envolvendo o uso civil destes aparelhos. Israel, um grande usuário e fabricante de aeronaves não tripuladas anunciou recentemente estar pronto para colocar no mercado uma versão civil da sua aeronave não tripulada Eagle para a vigilância anti-incêndio.

### **1.2.2 Estado da arte: equipamentos na França e no mundo**

Grandes dificuldades na construção de aeronaves não tripuladas eficazes são comuns na maioria das possíveis missões para sistemas aeronáuticos autônomos. Diversos estudos franceses e internacionais tratam sobre uma larga gama de possíveis missões de vigilância, de intervenção (militar ou para segurança civil), de reconhecimento, de observação (alternativa aos satélites), de inspeção, etc. Entre as dificuldades de tornar esses estudos operacionais, podemos citar:

1. autonomia de decisão embarcada;
2. integração da carga útil e do sistema de controle;
3. segurança do funcionamento e certificação;
4. sensores, capacidades em tempo integral, “ver e evitar”;
5. conexão de dados.

Um número crescente de projetos em robótica aérea autônoma acontecem nos Estados Unidos, na Austrália, no Japão e na Europa, envolvendo diversas universidades e empresas industriais.

Nas universidades o interesse principal é o desenvolvimento de pesquisas voltadas para a robótica, **visão computacional** ou inteligência artificial. Nesse contexto, a credibilidade dessas pesquisas é comprovada por uma demonstração experimental em vôo, o que vale mais que a simulação pura.

Os helicópteros são adequados, do ponto de vista experimental, pois podem evoluir em uma zona limitada por algumas dezenas de metros, o que simplifica a comunicação com a estação solo e também facilita a obtenção de autorizações para o vôo. A atual fase de compras de helicópteros não tripulados miniaturas (3m de diâmetro) nos laboratórios americanos, franceses e brasileiros simboliza bem essas vantagens.

Na França, a Universidade de Compiègne comprou helicópteros não tripulados para estudar o controle automático do vôo, com aterrissagem e decolagem automáticas numa zona preparada (GPS diferencial). O laboratório LAAS/CNRS em Toulouse trabalha igualmente nos comandos de vôo de um helicóptero sem piloto: eles compraram um pequeno dirigível para ser robotizado e realizar pequenas patrulhas. A ONERA foi contactada para cooperar com essas pesquisas, limitando-se à troca de informações e à realização de teses e estágios em parceria com a Dassault Aviação.

No Brasil, grandes universidades também adquiriram seus exemplares para desenvolver pesquisas. O Instituto Militar de Engenharia possui um helicóptero do mesmo modelo usado neste projeto, e o laboratório de controle da Universidade Federal do Rio de Janeiro também adquiriu um aeromodelo em escala de um helicóptero, modelo “Raptor” da Thunder Tiger, e um conjunto transmissor-receptor para o seu controle manual com o objetivo de instrumentá-lo e realizar o seu controle automático através de um computador.

Entretanto, os projetos acadêmicos universitários não buscam embarcar todos os componentes de decisão ou percepção presentes no sistema. Uma parte do processamento da informação e gestão da missão é realizado em terra, conectado à aeronave por um simples canal de radio Ethernet. Esses projetos querem principalmente demonstrar o desempenho dos algoritmos para o controle automático do vôo, reconhecimento de formas e aterrissagem autônoma.

Alguns projetos já foram implementados: a Universidade de Linköping na Suécia coordena o projeto WITAS de um helicóptero de vigilância de tráfego rodoviário, essencialmente em meio urbano. O projeto do dirigível autônomo AURORA, da Universidade de Campinas, vislumbra aplicações na segurança pública e no combate ao incêndio.

Outros projetos estão em andamento nos EUA, financiados pela marinha americana, o que reflete o interesse no uso desses dispositivos em missões de combate, tendo como sigla UCAV para os táticos e de combate e CSAR para os de busca e salvamento.

## **1.3 Objetivos e interesses do projeto RESSAC**

### **1.3.1 Objetivos detalhados**

A autonomia do controle de vôo, a realização de uma missão e a aquisição e processamento da informação são os principais objetivos do projeto.

Não se trata de uma autonomia completa de decisão, mas em casos em que a comunicação se interrompa, a aeronave deve ser hábil o suficiente para restabelecer contato e não colocar em risco a segurança das pessoas em terra. Isso permitiria ao operador romper a comunicação permanente com o sistema, aumentando o raio de operação, a autonomia de vôo, sem se limitar pela capacidade de análise e a habilidade do operador, diminuindo assim a vulnerabilidade do sistema.

O objetivo é demonstrar o salto tecnológico na integração e harmonização destas funcionalidades numa aeronave de busca e salvamento em meio hostil. Igualmente essencial é o desenvolvimento das capacidades de decisão autônoma e cooperativa.

Este projeto final aborda a parte de decisão ligada à escolha de zonas de aterrissagem, o que envolve uma série de dificuldades que serão analisadas no decorrer do texto.

### 1.3.2 Interesses científicos e técnicos

Os níveis de autonomia para uma aeronave podem ser definidos como:

- tele-pilotagem,
- pilotagem automática,
- navegação autônoma  
(com pontos de passagem e funções para evitar colisões),
- decisão autônoma (com a gestão da missão e da carga útil).

A maioria das aeronaves não tripuladas atuais são telepilotadas com uma navegação automática.

O sistemas atuais são dotados de navegação autônoma (pontos de passagem, mas geralmente sem funções para evitar obstáculos), porém ainda têm uma implementação complicada e onerosa.

Quando se trata de helicópteros não tripulados, a ONERA produziu em 1998 uma demonstração pioneira na França no controle avançado de missão num projeto denominado Vigilant, Figura 1.2. Foi nesse caso que se realizou uma das primeiras aplicações em tempo real de controle de vôo e navegação de um helicóptero não tripulado.



Figura 1.2: O helicóptero Vigilant

Dando prosseguimento, o projeto RESSAC visa a etapa tecnológica seguinte quando se trata de autonomia de decisão. Fala-se de autonomia:

- de controle de vôo e desempenho de missão,
- de aquisição e **processamento da informação**,
- de decisão embarcada.

Também faz parte do escopo o desenvolvimento de ferramentas de restituição da informação para auxiliar o operador.

Para garantir o uso genérico e a perenidade do trabalho, busca-se uma independência completa de fabricantes adquirindo um equipamento fechado, pronto para o uso desde o início do projeto.



Figura 1.3: Helicóptero Yamaha RMAX usado no projeto

Trata-se da escolha de um equipamento comercial, confiável e de custo relativo baixo, por isso a escolha do helicóptero sem piloto Yamaha RMAX, visto na Figura 1.3. Cerca de 1200 helicópteros sem piloto Yamaha R-50 e RMAX foram vendidos para uso agrícola no Japão. Este helicóptero é uma plataforma de demonstração apreciada com a qual diversos renomados centros de pesquisa já se equiparam, tais como o centro da NASA Ames na Califórnia, University of California at Berkeley, Carnegie Mellon University, Georgia Institute of Technology, a Universidade de Kyoto (Japão) e o projeto WITAS da Universidade Linköping (Suécia).

## 1.4 Organização do relatório

Como já foi mencionado, este relatório aborda minha contribuição na parte de decisão automática da aeronave, mais precisamente a decisão feita a partir de uma câmara embarcada. O objetivo é permitir a escolha de zonas planas e sem obstáculos do terreno, propícias a aterrissagem, evitando desse modo casas ou árvores.

O relatório está organizado da seguinte forma: o próximo capítulo descreve as técnicas de visão computacional e os algoritmos estudados e empregados no projeto, ressaltando os benefícios de cada alternativa e o porquê das escolhas. O Capítulo 3 descreve os métodos de transformação de coordenadas usados para compensar os movimentos 3D do helicóptero, para em seguida compor a cartografia do terreno. Finalmente, apresento no Capítulo 4 uma síntese comparativa entre os dois modos densos propostos, o fluxo ótico e a detecção de Odobez-Bouthemy. Apresento também a análise dos filmes reais do helicóptero aplicando a detecção de Odobez-Bouthemy, e para finalizar faço a conclusão sobre o projeto.

## CAPÍTULO 2

### VISÃO

#### 2.1 Visão computacional

Atualmente a visão artificial é utilizada em diversos setores desde aplicações médicas, militares, controle de tráfego rodoviário, robótica, ou na síntese de imagens. No caso do projeto RESSAC, onde um robô evolui num ambiente aéreo, a visão é utilizada para obter informações em relação ao terreno sobrevoado.

Neste capítulo, explicaremos rapidamente a parte visão do projeto, mostrando em detalhes as técnicas de processamento de imagens empregadas no processo de aterrissagem automática de uma aeronave num terreno desconhecido.



Figura 2.1: Câmara posicionada sob a aeronave

Sabendo que as informações sobre o terreno são obtidas essencialmente pela aqui-

sição de imagens, a visão é tida como aspecto chave. Primeiramente, foi necessário fazer a escolha em relação ao posicionamento da câmera, optou-se pela câmera colocada sob o helicóptero orientada verticalmente para o solo, conforme a Figura 2.1. Esta escolha foi feita em detrimento, por exemplo, de um posicionamento oblíquo que apresentava dificuldades na implementação devido à oclusão de certas regiões da imagem.

O equipamento usado é uma câmera digital modelo AVT 145-B2, sensor CCD do tipo 1/3 progressivo, preto e branco,  $1280 \times 960$ , dispositivo ótico de 6 mm e ângulo de abertura de  $45^\circ$ .

Para solucionar as dificuldades presentes na aterrissagem automática, o helicóptero efetua uma exploração do terreno, por exemplo, segundo a trajetória apresentada na Figura 2.2. Ele referencia as possíveis zonas de aterrissagem, para em seguida efetuar a trajetória de aterrissagem evitando os obstáculos elevados sobre uma zona selecionada.

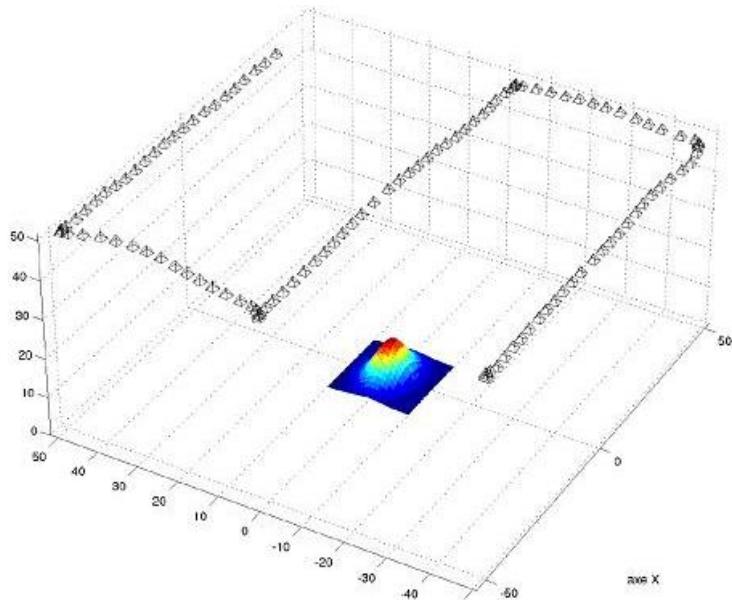


Figura 2.2: Trajetória de exploração de uma zona (filme: Esperce80)

O processamento desenvolvido na ONERA é baseado nas técnicas de estimação do relevo pela visão estéreo monocular [GLB07].

De forma mais concisa, procedemos em duas etapas :

- em um primeiro momento, as duas imagens são reajustadas globalmente por

um modelo paramétrico do movimento, segundo a hipótese de uma cena plana. Caso esse reajuste conduza a uma superposição perfeita entre as duas imagens, conclui-se que a cena sobrevoada é efetivamente plana, levando em conta o caráter estereoscópico do par de imagens considerado; ver Seção 2.5

- em seguida, caso a superposição não seja perfeita, buscamos a presença de zonas elevadas no relevo a partir dos deslocamentos residuais entre as duas imagens reajustadas.

## 2.2 Escolha de pontos da imagem

### 2.2.1 Extração de pontos de interesse

Uma etapa de extração de informações da imagem é feita *a priori* para que, em seguida, sejam realizados os outros processamentos necessários. Nesta seção, estuda-se a extração de pontos de interesse.

Entende-se por ponto de interesse um ponto qualquer da imagem para o qual o sinal muda nas duas dimensões. “Quinas”, definidas como as extremidades dos polígonos, as “junções” em T e Y, um ponto preto sobre um fundo branco, o fim de uma linha e qualquer outra manifestação com uma textura 2D significativa [CS00]. Os seguintes objetos satisfazem esse princípio, sendo considerados pontos de interesse.

Existe uma grande variedade de detectores de pontos e arestas na literatura. Eles são divididos em três categorias: métodos baseados no contorno, métodos baseados na intensidade e métodos baseados num modelo paramétrico. Primeiramente temos os métodos de contorno, que servem para extrair o contorno e em seguida procurar os pontos de curvatura ou inflexão máxima ao longo da linha de contorno; outra possibilidade seria fazer uma aproximação poligonal para só então procurar os pontos de interseção. O segundo caso trata dos métodos baseados na intensidade que calculam diretamente o nível de cinza de um indicador de presença ou um ponto de interesse. Por fim temos os métodos por modelo paramétrico que ajustam um modelo de intensidade paramétrico ao sinal. Normalmente eles permitem uma acuracidade de frações de pixel, mas somente para tipos específicos de pontos, como as arestas em L.

Dentre as propostas descritas acima opta-se pelo método baseado na intensidade, mais conhecido como detector de pontos de Harris [eMS88]. Entre suas qualidades podemos citar que este detector apresenta uma forte invariância à rotação, à troca de escala, à variação de iluminação e ao ruído da imagem.

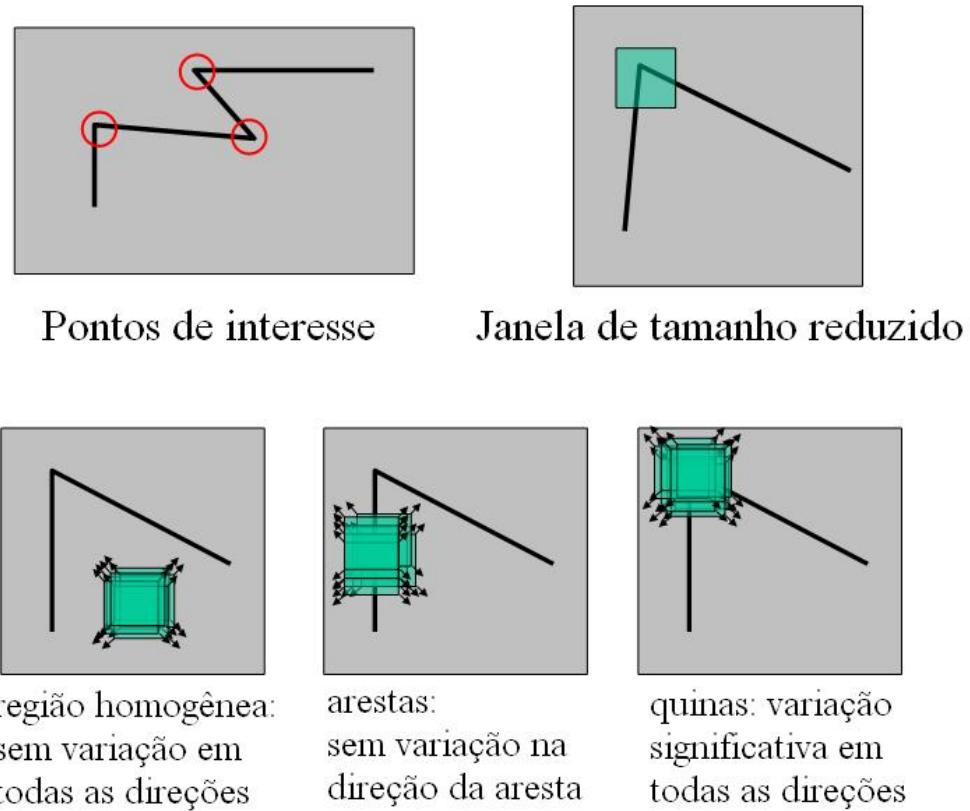


Figura 2.3: Idéia básica do detector de Harris

A extração de pontos de Harris, ver Figura 2.3, tem como idéia básica o fato de utilizar uma janela de tamanho reduzido para reconhecer facilmente um ponto específico, pois ao deslocar essa janela em qualquer direção observa-se uma grande variação de intensidade.

Variações de intensidade para um deslocamento de  $[u, v]$ :

$$E(u, v) = \sum_{x,y} w(x, y)[I(x + u, y + v) - I(x, y)]^2$$

onde  $I(x, y)$  é a intensidade do ponto  $(x, y)$  e  $I(x + u, y + v)$  é a intensidade deslocada,  $w(x, y)$  é a janela que pode ser uma gaussiana por exemplo.

Sendo  $E(u, v) = k$  uma elipse de valor constante, para pequenos deslocamentos  $[u, v]$  tem-se a aproximação bilinear:

$$E(u, v) \cong [u, v] \ M \begin{bmatrix} u \\ v \end{bmatrix}$$

onde  $M$  é uma matriz  $2 \times 2$  calculada segundo as derivadas da imagem:

$$M = \sum_{x,y} w(x, y) \begin{bmatrix} I_x^2 & I_x I_y \\ I_x I_y & I_y^2 \end{bmatrix} \quad (2.1)$$

Definimos  $\lambda_1$  e  $\lambda_2$  como sendo os autovalores de  $M$ .

As variações de intensidade serão detectadas através da análise de  $\lambda_1, \lambda_2$ . Então a classificação dos pontos é baseada nos autovalores de  $M$ .

Para facilitar a análise da resposta aplica-se:

$$R = \det M - k(\text{tr}(M))^2$$

onde o valor empírico de  $k$  é constante e igual a 0.06,  $\det M = \lambda_1 \lambda_2$  e  $\text{tr}(M) = \lambda_1 + \lambda_2$ .

A variável  $R$  depende exclusivamente dos autovalores de  $M$  - quando o valor de  $R$  é grande<sup>1</sup> o ponto em questão é uma quina, para um valor de  $R$  negativo com uma grande magnitude é uma aresta. Finalmente para um pequeno  $|R|$ , é uma região plana.

O algoritmo do detector consiste em encontrar as quinas,  $R > \text{limiar}$ . Depois disso são mantidos apenas os pontos de máximo local de  $R$ , um exemplo aparece na Figura 2.4.

---

<sup>1</sup>ultrapassa um limiar previamente estabelecido

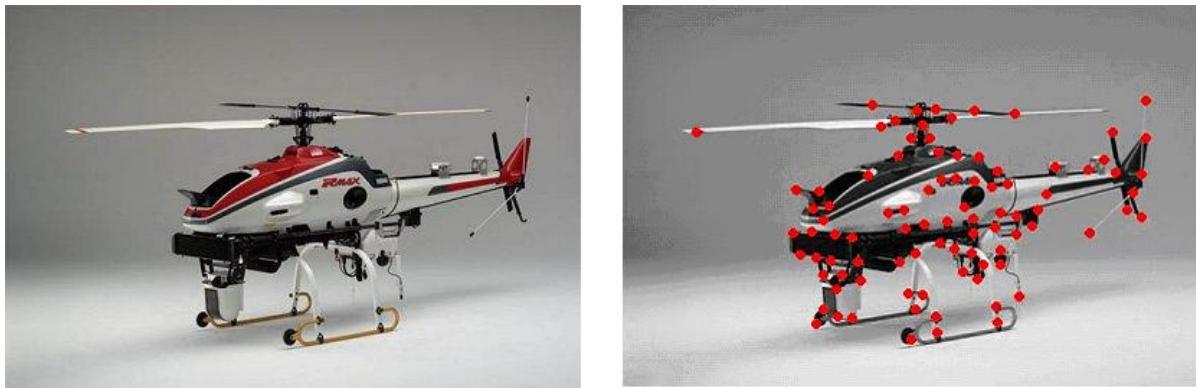


Figura 2.4: Extração de pontos de Harris

### 2.2.2 Associação de uma imagem à outra

A estimação de movimentos inter-imagem incita a utilizar o casamento ou rastreamento, que não é nada mais do que o casar de primitivas da primeira imagem com aquelas da segunda imagem.

Segundo Tomasi, [eTK91], um grande problema está em encontrar o deslocamento  $\mathbf{d}$  de um ponto de uma imagem à seguinte pois um único pixel não pode ser rastreado, salvo raras exceções em que o ponto tem uma luminância muito diferenciada dos demais, o que não corresponde ao caso real.

Quando a aeronave tira imagens sucessivas do terreno, partimos da hipótese que elas são projeções 2D de um mesmo detalhe 3D (falamos de “primitivas homólogas”): elas fazem referência à mesma cena com pequenas diferenças do ponto de vista. As imagens tiradas em curto intervalo de tempo têm, na maioria dos casos, uma forte correlação entre si.

Em seguida é preciso fazer a correspondência dos detalhes visíveis nas duas imagens como nos mostram as duas etapas descritas no fim da Seção 2.1. Esta operação é extremamente ambígua se as imagens são pouco texturadas ou muito ruidosas. De modo geral, o desempenho do casamento é relativo ao conteúdo textural da região considerada na imagem.

Realizar o casamento das primitivas de duas imagens de uma mesma cena ainda é um problema em aberto para os pesquisadores da comunidade de visão computacional.

O rastreamento pode ser realizado de diversas maneiras, por exemplo pelos os métodos que utilizam a similaridade entre as intensidades em nível de cinza dos pixels. As propostas abordadas podem ser classificadas em duas categorias:

1. Casamento por bloco: usando a hipótese de que os blocos “casados” são de tamanho reduzido ( $11 \times 11$ ). O procedimento realiza uma correlação entre os níveis de cinza da imagem, partindo da hipótese que existe uma certa similaridade entre as duas imagens e entre os níveis de cinza correspondentes a cada pixel. Esta hipótese é verificada em dois casos: zonas relativamente texturadas e pares de imagens com pequenos deslocamentos inter-imagem.
2. Aparecimento por primitivas: o algoritmo realiza a extração de primitivas, em seguida o casamento entre os dois pontos de vista é efetuado. O casamento pode ser igualmente feito sob um modelo paramétrico de movimento.

Quando se trata do projeto, inicialmente os pontos de interesse são extraídos, como já foi tratado na Seção 2.2.1, para em seguida realizar o casamento. Vejamos em mais detalhes o aparecimento de primitivas.

### Rastreamento KLT (Kanade, Lucas e Tomasi)

O método de rastreamento de pontos estima o deslocamento entre duas imagens sucessivas a partir de um modelo de movimento e pressupondo que um ponto deve ser reencontrado em uma janela de tamanho fixo na imagem seguinte.

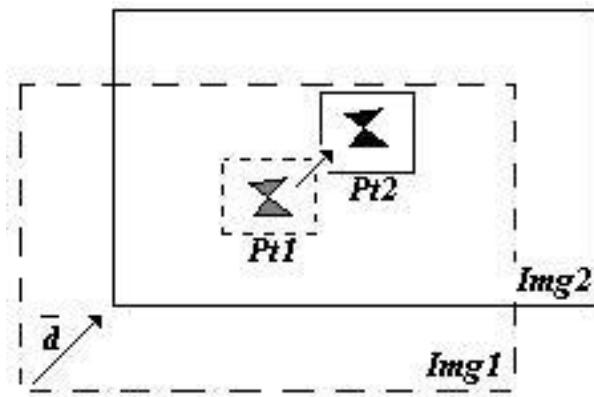


Figura 2.5: Rastreamento de primitivas

Na Figura 2.5,  $Pt1$  é um ponto da imagem  $I1$ ,  $Pt2$  o ponto correspondente na imagem  $I2$  e  $\bar{d}$  a translação ligando  $I1$  a  $I2$ . O método KLT proposto por Kanade,

Lucas e Tomasi [Luc81] é baseado numa extração de pontos na primeira imagem, e o rastreamento destes pontos nas imagens seguintes. Somente as vizinhanças dos pontos característicos são tratadas na imagem seguinte. Supomos que no caso de deslocamentos inter-imagem relativamente pequenos (câmera se movendo lentamente), que a vizinhança do ponto  $Pt1$  perseguido na imagem  $I1$  se encontra na imagem  $I2$  por uma translação:

$$I2(x, y) = I1(x - \alpha, y - \beta) + n(x, y)$$

onde o vetor de translação entre as imagens é  $\bar{d} = [\alpha, \beta]^T$  e  $n(x, y)$  é o ruído na posição  $(x, y)$ .

O vetor  $\bar{d}$  é então escolhido de forma a minimizar o erro  $\varepsilon$  calculado sobre uma janela de vizinhança  $W$  da seguinte maneira:

$$\varepsilon = \sum_{\bar{x} \in W} [I1(\bar{x} - \bar{d}) - I2(\bar{x})]^2 \omega(\bar{x}) \quad (2.2)$$

onde  $\omega(\bar{x})$  é uma função de ponderação e  $\bar{x} = [x, y]^T$ . Geralmente  $\omega(\bar{x}) = 1$  mas ela também pode tomar uma forma gaussiana caso seja do interesse dar mais importância ao centro da janela. Para minimizar  $\varepsilon$  nós diferenciamos a Equação (2.2). Escrevendo o desenvolvimento de Taylor em primeira ordem, obtemos:

$$I1(\bar{x} - \bar{d}) = I1(\bar{x}) - g g^T \bar{d}$$

$$\text{onde } g = \nabla I1 = \left[ \frac{\partial I1(\bar{x})}{\partial x}, \frac{\partial I1(\bar{x})}{\partial y} \right]^T$$

Desenvolvendo a Equação (2.2) obtemos uma equação do tipo:

$$Z \bar{d} = \bar{e} \quad (2.3)$$

$$\text{com : } \begin{cases} Z = \sum_W g g^T \omega(\bar{x}) \\ \bar{e} = \sum_W [I1(\bar{x}) - I2(\bar{x})] g(\bar{x}) \omega(\bar{x}) \end{cases}$$

O objetivo é estimar  $\bar{d}$  na Equação (2.3) conhecendo  $Z$  e  $\bar{e}$ . Para determinar a possibilidade de “rastrear” uma primitiva, é preciso que o sistema seja bem condicionado e robusto ao ruído. Para que isso aconteça, é necessário que a razão entre os dois autovalores de  $Z$  não seja muito grande (sistema bem condicionado) e que este valor seja superior a um limiar (robusto ao ruído). É importante notar que o critério de inversão do sistema seleciona aproximadamente os mesmos pontos de interesse que o detector de Harris. Mais precisamente a matriz  $M$ , da Equação (2.1), é uma versão de  $Z$  para um derivador do tipo gaussiano.

### Correlação Ida e Volta

No contexto do casamento de primitivas, colocamos em evidência o método escolhido para o projeto. Uma descrição sucinta deste método é: efetuamos o cálculo de similaridade entre as janelas englobando os pontos considerados nas duas imagens e escolhemos as associações que maximizam essa similaridade, utilizando uma estratégia de decisão ida/volta (associação de  $t - 1$  para  $t$ , depois de  $t$  para  $t - 1$ ).

O princípio é o seguinte:

1. detector de pontos de Harris é aplicado a cada imagem;
2. o casamento é realizado usando o modelo de translação;
3. *casamento ida*: para cada ponto de Harris  $x_1$  extraído na primeira imagem procuramos seu correspondente entre os pontos de Harris detectados na segunda imagem,  $x_2$ . Seja  $x_{2ap}$  o ponto no qual  $x_1$  apareceu.
4. *casamento volta*: o mesmo teste é efetuado porém no sentido inverso. Procuramos o correspondente do ponto  $x_{2ap}$  entre os pontos de Harris detectados na imagem  $I1$ . Seja  $x_{1ap}$  o ponto com o qual o ponto  $x_{2ap}$  casou. O casamento volta é uma etapa de verificação ou de validação do casamento ida. Na prática, a quantidade de casamentos volta é diferente da quantidade de casamentos ida.
5. 1<sup>a</sup> estimação do modelo afim (ver Seção 2.3.2) global baseado no rastreamento de pontos e na minimização de um erro quadrático.
6. rejeição de pontos aberrantes e 2<sup>a</sup> estimação robusta do modelo afim global utilizando o M-estimador IRLS com apenas uma ponderação.

O algoritmo extrai, utilizando um simples modelo de translação, duas listas de pontos de Harris  $n_1$  e  $n_2$  das imagens  $I_1$  e  $I_2$ , respectivamente. Para ter mais pontos rastreados na segunda imagem  $n_2 = \alpha n_1$ , com  $\alpha < 1$ . Além disso, podemos escolher o número de pontos que vamos reter, visando obter  $n_2 > n_1$ .

Definindo  $x_{1k}$  o ponto número  $k$  da lista extraída da imagem  $I_1$ , onde  $k = 1, \dots, n_1$ , o problema está em encontrar para cada ponto  $x_{1k}$  seu correspondente  $x_{2l}$  na imagem  $I_2$  entre todos os pontos  $x_{2l}$  extraídos anteriormente onde,  $l = 1, \dots, n_2$ .

Para buscar um candidato para o casamento de  $x_{1k}$  com um ponto de Harris  $x_{2l}$  da imagem  $I_2$  o método usa um critério de correlação calculado sobre uma zona de pesquisa na imagem  $I_2$ , centrado em  $x_{1k}$  de tamanho  $(2N + 1) \times (2N + 1)$ . A zona de pesquisa usa o conhecimento prévio sobre o movimento que separa  $I_1$  de  $I_2$ , por exemplo a translação inter-imagem global dada pelos sensores a bordo, neste caso um sensor GPS. A zona de pesquisa do ponto correspondente  $x_{2k}$  na imagem  $I_2$  se reduz a uma janela de pesquisa ao invés de varrer a imagem inteira. A zona de pesquisa inclui um número de pontos de Harris detectados sob um total de  $n_2$  pontos detectados sobre toda a imagem  $I_2$ . Seja  $x_{2k \in f}$  um ponto de Harris detectado na imagem  $I_2$ , compreendido nesta zona de pesquisa. Para encontrar um candidato para o aparecimento de  $x_{1k}$  com um ponto  $x_{2k \in f}$  incluído na zona de pesquisa estabelecida, o algoritmo minimiza a função “Score” a seguir:

$$Score(x_{1k}, x_{2k \in f}) = 1 - ZNCC(x_{1k}, x_{2k \in f})$$

Onde  $ZNCC$  é um critério de medida da semelhança entre duas primitivas. Ele deriva da  $NCC$  (*Normalized Cross Correlation*), o critério mais utilizado, que é a soma de dois produtos entre termos correspondentes, normalizados pelo produto das médias quadráticas calculadas para cada janela, mas no caso da  $ZNCC$  (*Zero Mean NCC*)  $I_1$  e  $I_2$  são valores médios nas regiões de interesse. Trata-se de um bom critério por não se alterar devido às variações uniformes de intensidade luminosa em uma das duas imagens. Caso consideremos  $r = (i, j)^T$  e  $V = \{(i, j) | -N \leq i \leq N, -N \leq j \leq N\}$ , dois pontos  $\mathbf{x}_1 = [x_1, y_1]^T$  e  $\mathbf{x}_2 = [x_2, y_2]^T$  pertencendo respectivamente às imagens  $I_1$  e  $I_2$  além de definir  $\bar{I}(\mathbf{x})$  como sendo o valor médio de  $I(\mathbf{x})$  na zona de pesquisa, obtemos a seguinte fórmula:

$$\text{ZNCC}(\mathbf{x}_1, \mathbf{x}_2) = \frac{\sum_{r \in V} [I_1(\mathbf{x}_1+r) - \bar{I}_1(\mathbf{x}_1)][I_2(\mathbf{x}_2+r) - \bar{I}_2(\mathbf{x}_2)]}{\left[ \sum_{r \in V} [I_1(\mathbf{x}_1+r) - \bar{I}_1(\mathbf{x}_1)]^2 \sum_{r \in V} [I_2(\mathbf{x}_2+r) - \bar{I}_2(\mathbf{x}_2)]^2 \right]^{1/2}}$$

Quanto mais o  $Score(x_{1k}, x_{2k \in f})$  estiver perto de 0 (ele é definido entre  $[0, 1]$ ), mais os pontos  $x_{1k}$  e  $x_{2k \in f}$  serão parecidos em relação a distribuição de intensidades em nível de cinza numa janela considerada. Podemos então notar que o ponto correspondente ao ponto  $x_{1,k}$  é o ponto dado por:

$$x_{2k \in f} = \operatorname{argmin} [Score(x_{1k}, x_{2k \in f})]$$

Seja  $A_1$  a quantidade de casamentos ida e volta. Temos que dentre os  $n_1$  pontos de Harris detectados na imagem  $I_1$ ,  $\bar{A}_1$  (com  $\bar{A}_1 = n_1 - A_1$ ) não foram casados. Isso acontece porque  $A_1$  pontos se encontram fora da zona de pesquisa pois os deslocamentos inter-imagem são superiores à zona de pesquisa. As soluções para combater esses inconvenientes são:

- lançar mão de um conhecimento prévio sobre o deslocamento que separa as duas imagens a partir de dados de bordo e compensá-los antes de realizar o casamento.
- aumentar a zona de pesquisa para ter mais pontos perseguidos. Isto acaba por aumentar consideravelmente o tempo de cálculo, o que não é do nosso interesse por se tratar de uma aplicação em tempo real.
- aumentar a cadênciça da tomada de imagens, ou seja as imagens serão separadas por deslocamentos menores, logo a zona de pesquisa será grande o suficiente para encontrar os pontos. Isto também não é interessante por aumentar o número de pares de imagem a serem tratados.

O casamento da volta consiste em encontrar para o ponto  $x_{2k \in f}$  seu correspondente na imagem  $I_1$ . Chamamos de  $x_{1k \in f}$  o ponto de Harris da imagem  $I_1$  correspondente de  $x_{2k \in f}$ . Para encontrar o ponto  $x_{1k \in f}$ , o método usa o mesmo critério de correlação calculado sobre uma zona centralizada em  $x_{2k \in f}$ , desta vez na imagem  $I_1$ .

Sendo  $A_2$  a quantidade de casamentos da volta, geralmente  $A_2$  é inferior a  $A_1$ . Assim decidimos quais pontos foram mal perseguidos. Se  $A_1 \neq A_2$  (o que acontece normalmente na prática) o método considera como casamentos válidos os primeiros  $A$  casamentos aparecidos dados por  $A = \min(A_1, A_2)$ .

## 2.3 Modelo global robusto

Sabendo que trabalhamos normalmente com um par de imagens, tratamos de homografias do plano, quer dizer em 2D.

Assumimos dois planos  $P$  e  $P'$  munidos respectivamente dos referenciais projetivos  $R$  e  $R'$ . Diz-se que uma homografia é uma aplicação linear bijetiva de  $P$  para  $P'$  definida por uma matriz  $H$  não nula de ordem  $3 \times 3$ , tal que a imagem de um ponto  $S(x, y, z)$  do plano  $P$  seja um ponto  $S'(x', y', z')$  do plano  $P'$ . Então podemos dizer que  $S' = HS$ .

Para aplicar homografias sobre as imagens, nós consideramos dois pixels  $p_1$  e  $p_2$  provenientes de um modelo em que dois pixels provenientes do mesmo ponto do espaço são projetados nas duas imagens. A geometria epipolar<sup>2</sup> se aplica a todos os pontos da imagem em correspondência. Façamos a hipótese de uma informação geométrica complementar, digamos que os pontos 3D são coplanares no espaço. Neste caso, podemos passar pontos de um plano ao seu correspondente através de uma transformação homográfica.

### 2.3.1 Estimação robusta

Um estimador é dito robusto se ele é insensível a pequenos desvios sob a hipótese para a qual ele foi otimizado. Existem dois sentidos para o termo “pequeno”: pequenas variações sobre todos os dados, ou desvios significativos sobre uma quantidade reduzida de dados. O segundo aspecto é o mais difícil de ser tratado pelos estimadores clássicos.

É difícil extrair informações visuais provenientes da imagem, ditas informações

---

<sup>2</sup>correspondente à estrutura geométrica entre duas vistas

de alto nível, com uma precisão considerável. Os parâmetros de alto nível como a profundidade, a posição da câmera em relação à cena, o deslocamento da câmera ou do robô entre outros parâmetros intrínsecos, não são facilmente extraídos se as câmeras fornecerem uma informação de baixo nível sobre o ambiente cheio de detalhes.

Os modelos geométricos são um exemplo de ferramentas bem adaptadas para a modelagem do movimento inter-imagem. Quando se trata de imagens reais (mais particularmente em aplicações robóticas em ambiente externo) a modelagem acaba sendo inexata, logo algoritmos robustos são necessários.

Apresentaremos os dois principais grupos de algoritmos que demonstram o estado da arte dos métodos de estimativa robusta usados em visão computacional [Mal05].

É comum que a robustez seja traduzida pela resistência da estimativa aos dados aberrantes. Nós a definimos matematicamente pela menor quantidade de dados extremos que modificam o valor da estimativa divididos pelo número total de amostras.

## RANSAC

O método de pesquisa aleatória, RANSAC [Fis81], é um método de voto probabilístico proposto com a pretensão de reduzir o tempo de cálculo dos métodos clássicos de voto. Partindo de um sub-grupo mínimo de  $s$  sinais medidos é possível calcular os parâmetros numa situação não degradada. Em seguida, calculamos a seguinte função de custo, onde  $r_k(x)$  representa os ruídos:

$$C(x) = \sum_k \rho(r_k(x))$$

$$\text{onde } \rho(r_k(x)) = \begin{cases} 0, & \text{se } r_k^2(x) \leq c \\ 1, & \text{se } r_k^2(x) > c \end{cases} \quad \text{e } c = 2.5\sigma$$

O valor de  $\sigma$  é uma estimativa robusta do desvio padrão do ruído sobre as boas medidas. Ele é normalmente tratado como uma variável de ajuste que é escolhida de forma manual em função de cada aplicação em particular. De modo geral, para apurar a precisão da detecção de dados aberrantes, usamos a MAD (Median Absolute

Deviation):

$$\sigma = 1.48 \text{mediana}(|r - \text{mediana}_{Res}|) \quad (2.4)$$

onde  $\text{mediana}_{Res}$  corresponde à mediana dos resíduos.

Seja  $p$  a probabilidade de encontrar a melhor solução,  $s$  o número mínimo de sinais necessários para a estimação dos parâmetros e  $r$  a porcentagem de “inliers”<sup>3</sup>. O número  $m$  de sorteios necessários para ter uma probabilidade  $p$  de encontrar os bons parâmetros do modelo são:

$$m = \frac{\log(1 - p)}{\log(1 - (1 - r)^s)}$$

Por exemplo, na presença de 20% de “outliers”<sup>4</sup>, apenas 5 sorteios são suficientes para obter uma probabilidade de 95% de se encontrar uma boa solução. Para 40% de “outliers”, são necessários 13 sorteios.

### LMedS (Least Median of Squares)

O método robusto LMedS minimiza a seguinte função de custo:

$$C(x) = \text{mediana}(r_1^2(x), r_2^2(x), \dots, r_n^2(x))$$

A mediana é vista como um estimador robusto porque ela não considera os 50% dos resíduos mais significativos. Há entretanto dois inconvenientes:

- a função de custo não é diferenciável na maioria das vezes. Os métodos de minimização baseados no gradiente da função de custo são difíceis de serem implementados;
- a velocidade de convergência dos algoritmos de minimização pode ser extremamente baixa, caso os resíduos sejam distribuídos de tal maneira que a mediana tenha um gradiente pequeno.

<sup>3</sup>medidas que confirmam o modelo

<sup>4</sup>medidas aberrantes

### M-estimadores

O princípio dos M-estimadores é modificar a função de custo do Least Squares penalizando os maiores resíduos. A função de custo é descrita por:

$$C(x) = \sum_{k=1}^n \rho(r_k(x))$$

onde a função  $\rho$  é ao menos  $C^0$ . Diferentes funções  $\rho$  foram propostas na literatura. Uma das mais usadas é a proposta por Huber. O ponto de breakdown dos M-estimadores é teoricamente o 0 como para o LMedS. Entretanto, trata-se do pior caso e veremos que é preciso ter medidas aberrantes antes de questionar o bom funcionamento dos M-estimadores. A função  $\rho$  proposta por Huber é a que segue abaixo:

$$\rho(r_k(x)) = \begin{cases} \frac{1}{2}r_k^2(x), & \text{if } r_k^2(x) \leq c \\ c(|r_k(x)| - \frac{c}{2}), & \text{if } r_k^2(x) > c \end{cases} \quad \text{e } c = 1.345\hat{\sigma}$$

onde  $\hat{\sigma}$  é um estimador robusto de desvio padrão. Geralmente usa a MAD, Equação (2.4).

### IRLS

Na prática os M-estimadores são implementados com o método dos mínimos quadrados ponderados iterados IRLS (Iteratively Reweighted Least Squares). O algoritmo visa resolver o seguinte sistema  $DAx = Db$  onde  $D = \text{diag}(w_1, \dots, w_k)$  é uma matriz diagonal. O cálculo do peso  $w_i$  associado a cada medida representa a confiança que temos em cada uma delas. Este algoritmo funciona da seguinte maneira: estimação do peso usando um dos múltiplos critérios robustos presentes na literatura (Huber, Tukey), estimação do valor de  $x$  resolvendo o sistema precedente e reiteração até a convergência. Os pesos  $w_i$ , elementos da matriz D, refletem a confiança em cada primitiva sendo definidos por Huber [Hub81]:

$$w_i = \frac{\varphi(ds)}{ds}$$

onde  $\varphi(ds) = \partial\rho(\delta_i)$ ,  $\delta_i$  é o resíduo normal dado por  $\delta_i = \Delta_i - \text{mediana}\Delta$  (mediana $\Delta$  corresponde ao valor da mediana dos resíduos). O parâmetro  $\sigma$  representa o valor do desvio padrão do ruído sobre as “boas” medidas, podendo ter grandes variações ao longo do procedimento de minimização. O valor  $\sigma$  é geralmente tratado como uma variável de ajuste escolhida manualmente de acordo com cada aplicação em particular. Além disso, também é possível, a fim de melhorar a precisão da detecção dos dados aberrantes, estimar o valor de  $\sigma$  paralelamente à minimização do erro usando uma estatística robusta. O MAD, Equação (2.4), é um destes casos estatísticos.

No caso do projeto RESSAC, utilizamos o método de penalização:

$$\{A, t\} = \operatorname{argmin} \sum_k w_k \|\hat{x}_k - (Ax_k^2 + t)\|$$

Em um primeiro momento colocamos todos os pesos  $w_k$  com o valor igual a 1 para contar com todas as amostras. Em seguida, numa segunda iteração atribuimos seu peso correspondente, os mais diferentes terão seu peso igual a 0, sendo assim eliminados.

### 2.3.2 Modelos de transformação vislumbrados

#### Afim

É uma aplicação entre dois espaços afins que envia as retas, planos, espaços, em outras retas, planos, espaços. Sempre preservando a noção de paralelismo, ver Figura 2.6.

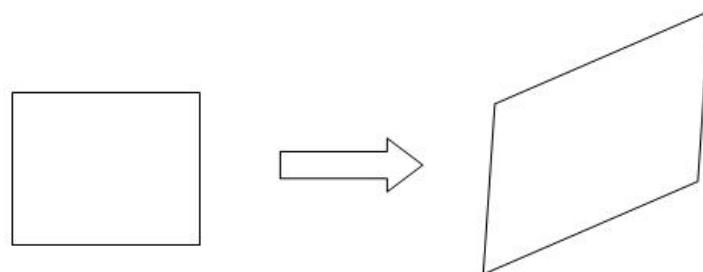


Figura 2.6: Ilustração de uma transformação afim.

$$x' = Ax + T$$

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} a_1 & a_2 \\ a_3 & a_4 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} a_5 \\ a_6 \end{bmatrix} \quad (2.5)$$

Hipótese : campo estreito, cena plana.

Procuramos a translação afim  $(A, T)$  para minimizar  $(I_1(k) - I_2(Ak + T))^2$ . Os inconvenientes deste tipo de translação são que muitas zonas da imagem são inúteis.

### Quadrático 2D

Derivada do modelo homográfico. É um modelo de transformação de parâmetros que representa uma boa aproximação em relação às deformações do plano imagem causadas pelos movimentos do helicóptero, ver Figura 2.7.

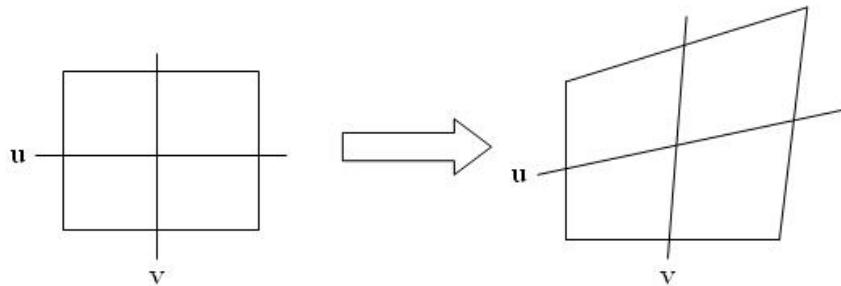


Figura 2.7: Ilustração da transformação quadrática.

$$x' = ax + by + cx^2 + dxy + e$$

Um dos pontos favoráveis é que ela se trata de uma transformação que depende de forma linear dos parâmetros. Sendo assim podemos calcular seus coeficientes mais facilmente.

## 2.4 Tipos de movimento na imagem

As imagens tratadas neste projeto representam projeções de cenas reais. Logo o movimento observado (ou movimento aparente) numa seqüência temporal de imagens vai representar a projeção do movimento 3D relativo à câmera no plano da imagem. O objetivo da estimativa do movimento é estimar o campo de movimento 2D ou 3D a partir de uma seqüência temporal de imagens 2D ou 3D nas quais o conteúdo varia em função do tempo. Para tal devemos distinguir:

- movimento real;
- movimento aparente;
- movimento estimado.

### 2.4.1 Campo de movimento real e aparente

O movimento real anima a cena real, ou seja, o espaço real 3D. Esse movimento real é observado por um ponto de vista que nos fornece o registro do movimento observado (aparente) sob a forma de uma seqüência temporal de imagens. Se nos basearmos apenas na intensidade dos pixels normalmente, não é possível medir o campo de movimento real diretamente: o que observamos de fato são as mudanças de distribuição espacial da intensidade luminosa. O movimento assim constatado é chamado de “movimento aparente” ou “fluxo ótico” que, em geral, é diferente do campo de movimento real.

O campo de movimento aparente representa, em geral, a projeção (ortográfica ou projetiva) do movimento real 3D no plano imagem. É por isso que o movimento aparente também é chamado de “movimento projetado”: ele representa uma aproximação do movimento real, veja a Figura 2.8.

Suponhamos que o ponto  $P_t$  de um objeto real 3D, no instante  $t$ , vem a ser o ponto  $P_{t+1}$  no instante  $t+1$ . As projeções perspectivas dos pontos  $P_t$  e  $P_{t+1}$  no plano imagem são representadas por  $p_t$  e  $p_{t+1}$ , respectivamente, com  $I_t(x, y)$  e  $I_{t+1}(x, y)$  sendo suas intensidades. Na Figura 2.8 notamos que o deslocamento aparente  $\vec{p_t p_{t+1}}$  corresponde à projeção perspectiva do movimento real  $\vec{P_t P_{t+1}}$ . Todos os vetores

de deslocamento em que suas extremidades pertencem a retas pontilhadas correspondendo a operação de projeção têm o mesmo vetor 2D de deslocamento aparente (ou projetado). A ilustração presente na Figura 2.8 é válida para uma seqüência de imagens adquirida com uma câmara de vídeo. Lembramos que ter uma câmara móvel num ambiente fixo é interpretado da mesma forma que ter uma câmara fixa num ambiente móvel quando se trata de movimentos gerados.

No caso contínuo, o campo de deslocamento aparente ou projetado é definido por todos os pontos  $p(x, y, t)$  do plano imagem por uma função vetorial de deslocamento,  $d(x, y, t)$ , com valores em  $\mathbb{R}^2$ .

No caso discreto, o campo de vetores de deslocamento entre os instantes  $t_k$  e  $t_{k+\Delta t} = t_0 + k \cdot \Delta t$ , onde  $k \in \mathbb{Z}$  e  $\Delta t$  é o passo de amostragem temporal, constitui uma representação amostrada da função definida no caso contínuo.

Em suma, o campo de deslocamento representa a totalidade dos vetores de deslocamento,  $d(i, j, t_k)$  onde  $(i, j, t_k) \in \Lambda^3$ ,  $\Lambda^3$  sendo um sub-espaco discreto do espaço (2D+tempo).

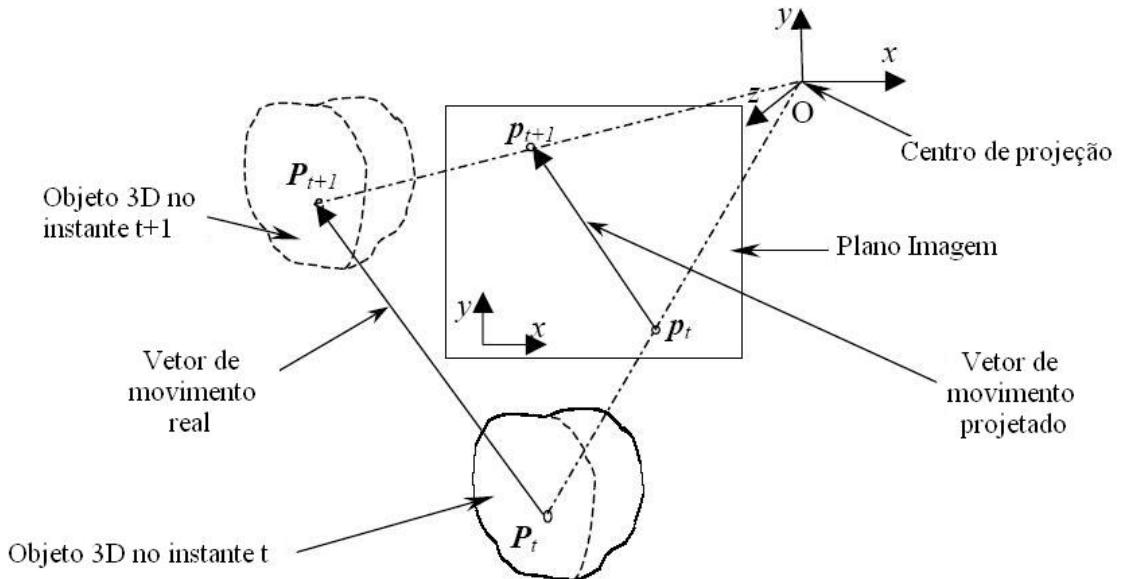


Figura 2.8: Movimento real e aparente num sistema ótico [Gra03]

### 2.4.2 Campo de movimento aparente e estimado

O vetor de deslocamento estimado  $d(p) = (d_x(p), d_y(p))$  corresponde ao movimento do ponto  $p = (x, y)$  no plano imagem. Ele é determinado a partir do campo de movimento aparente, quer dizer a partir de variações locais de intensidade luminosa  $I_t(x, y)$  entre os instantes  $t$  e  $t + \Delta t$ , onde  $\Delta t$  é o intervalo de tempo entre as imagens. O vetor de velocidade estimada  $\nu = (\nu_x(p), \nu_y(p))$  é definido como a variação temporal do deslocamento por unidade de tempo,  $(\nu_x, \nu_y) = \left(\frac{d_x}{d_t}, \frac{d_y}{d_t}\right)$ . Isto explica porque uma seqüência temporal de imagens, permite apenas estimar o campo de movimento (velocidade ou deslocamento) aparente observável em uma sequência e não o campo de velocidade real. Determina-se o campo de deslocamento (velocidade) estimado com um método qualquer.

Em ótica por exemplo, a diferença entre o campo de velocidade estimado e o aparente, pode existir se :

**O gradiente espacial é muito fraco :** Para que o movimento real seja bem observado, é preciso que haja variações em nível de cinza (ou em cores) suficientemente grandes nas regiões onde há movimento. O exemplo clássico que comprova esse fator, seria uma esfera com distribuição uniforme de intensidade (iluminação homogênea), realizando uma rotação em torno de seu próprio eixo, numa cena também com iluminação constante. Mesmo existindo movimento na realidade, ele não pode ser observado a partir de uma seqüência temporal de imagens em nível de cinza, porque não há nenhuma variação aparente em nível de cinza interferindo durante o movimento.

**A iluminação da cena varia :** Um movimento observável numa seqüência de imagens de intensidade nem sempre corresponde a um movimento real. A mesma esfera do exemplo anterior, desta vez imóvel submetida a iluminação variável ao longo de uma seqüência gera movimento aparente artificial (ou seja uma variação de intensidade).

## 2.5 Processamento estéreo

A estereoscopia monocular aplicada no projeto, consiste em explorar conjuntamente duas imagens sucessivas de uma cena supostamente rígida, para estimar as elevações do terreno a partir do cálculo da velocidade de deslocamento dos pontos (fluxo ótico).

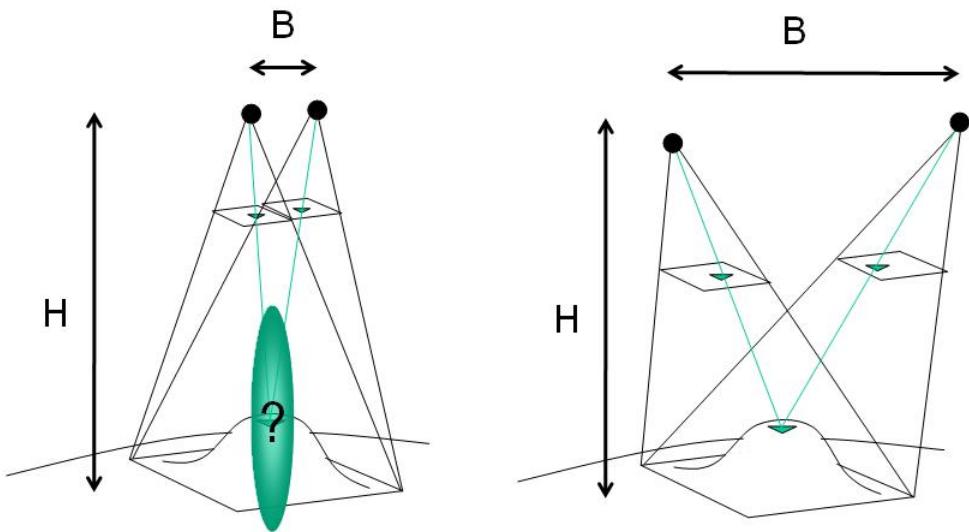


Figura 2.9: À esquerda: informação inválida. À direita: caráter estéreo suficiente

A precisão desta estimativa, ou caráter estéreoscópico do par de imagens considerado, em primeira ordem, é dado pela razão entre a distância horizontal entre as imagens no terreno (a base  $B$ ) e a altitude de vôo (altura  $H$ ), ver Figura 2.9.

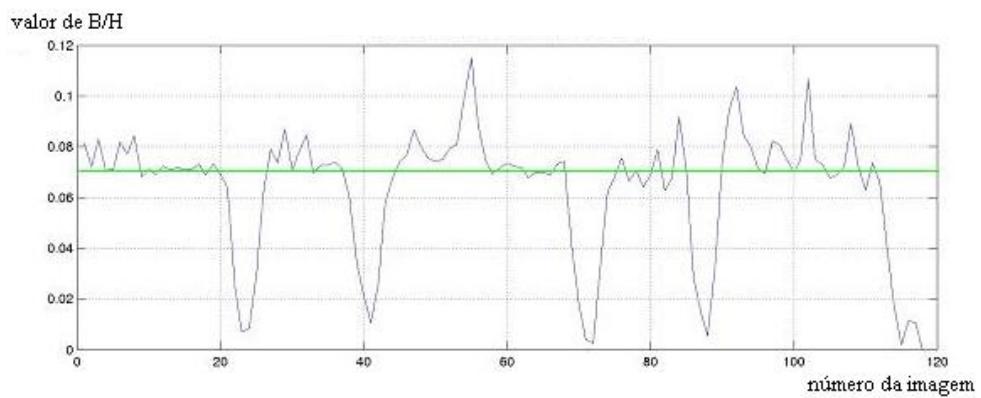


Figura 2.10: Razão base ( $B$ ) sobre altura ( $H$ ) dos pares de imagens sucessivas da seqüência Esperce80

Analisando a Figura 2.10, percebemos os quatro vales que estão abaixo do limiar constante 0.07 representam as curvas realizadas pelo helicóptero durante sua trajetória descrita na figura 2.2 (quinas do “S”): nesses casos o movimento 3D do helicóptero é essencialmente uma rotação, tendo o caráter estereoscópico é muito frágil. Dessa forma, esses pares de imagem serão descartados não fazendo parte do processamento.

## 2.6 Detecção de zonas de relevo

O princípio da detecção de zonas de relevo se baseia nos pixels “fora do chão”, que são mal reajustados pelos métodos descritos até então, essa característica nos permitirá classificá-los como zonas não-planas.

Neste caso, podemos pensar de duas formas distintas para estimar os movimentos residuais, sendo elas: o modo esparso, ou o modo denso.

No caso esparso apenas os pontos identificados como pontos de interesse serão rastreados. Já no caso denso todos os pontos da imagem são considerados, seja fazendo a estimativa do fluxo ótico ou a detecção de movimentos residuais pelo método de Odobe-Bouthemy [ePB97].

### 2.6.1 Modo esparso

O casamento esparso corresponde a eleger os pixels das duas imagens, contendo a textura mais heterogênea: os pontos de interesse. Usamos neste caso o detector de pontos de Harris, apresentado na Seção 2.2.1, parametrizado segundo as indicações presentes no artigo de C. Schmid [CS00]. Em seguida procuramos as associações entre esses pontos, como visto na Seção 2.2.2. Depois efetuamos um cálculo de similaridade entre janelas englobando os pontos considerados nas duas imagens e escolhemos as associações que maximizam esta similaridade, utilizando a estratégia de decisão ida e volta (associação de  $t-1$  para  $t$ , depois de  $t$  para  $t-1$ ). Um exemplo de pontos extraídos e rastreados é apresentado na Figura 2.11.

Para identificar se cada ponto faz parte ou não do solo, é estabelecido um valor máximo para o seu deslocamento inter-imagens, quando o ponto em questão

ultrapassa esse valor é classificado como um ponto que faz parte do relevo. Esse valor máximo permitido é atribuído após a análise de uma série de fimes de forma empírica.

Como processamos um número reduzido de pontos da imagem, o modo esparso se torna vantajoso pela velocidade. Seu resultado é válido uma vez que os pontos de interesse encontrados fazem parte do relevo, como no caso da Figura 2.11 à esquerda, o arbusto que foi bem classificado como relevo por ter diversos pontos de interesse (indicado pelos círculos verdes). Entretanto na imagem à direita, em que temos uma árvore sem folhas (inverno europeu), podemos constatar que há poucos pontos de interesse detectados sobre a árvore, logo ela não é identificada como relevo. Por esse fator esse método foi considerado ineficaz para o projeto, sendo descartado.

### 2.6.2 Modo denso

Após o processo de reajuste pelo modelo quadrático robusto passamos a uma fase de processamento sobre os pares de imagem através das imagens de diferença<sup>5</sup>, realizando o fluxo ótico por casamento de blocos, ou “block-matching”, a correlação de janelas para todo pixel  $k$ , ou pela detecção de Odobe, [ePB97].

<sup>5</sup>mais conhecida como DFD : *Displaced Frame Difference*

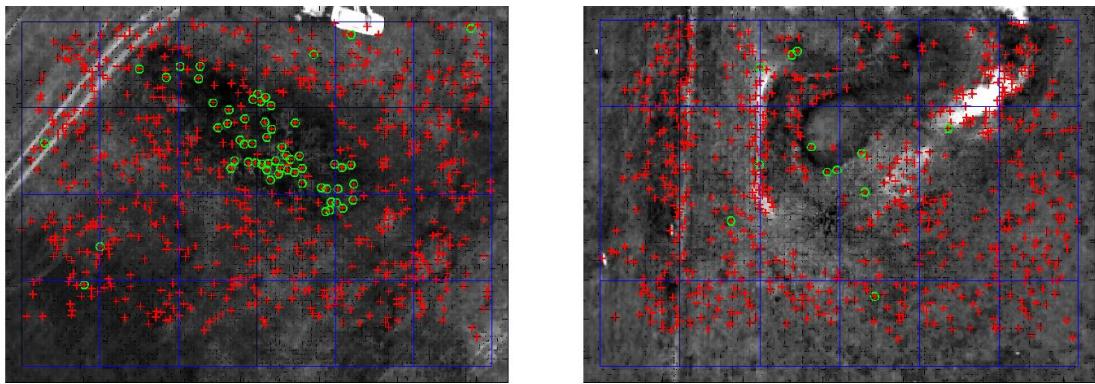


Figura 2.11: Imagens da seqüência Esperce80

## 2.7 Fluxo ótico

A detecção e estimação do movimento em seqüências temporais de imagens bidimensionais (2D) e tridimensionais (3D) é um dos problemas fundamentais no processamento e análise de imagens. Este domínio recente de pesquisa tem sua origem em diversos problemas em aberto para os quais não existe nenhuma solução geral e satisfatória.

O fluxo ótico é uma aproximação do movimento da imagem, definido como a projeção da velocidade dos pontos das superfícies 3D no plano imagem do captor visual [BB95].

A hipótese inicial para medir o movimento da imagem é ter iluminação constante, ou seja, a luminosidade aparente de um ponto  $\mathbf{p}(t) = (p_1(t), p_2(t))$  de uma seqüência temporal de imagens é constante ao menos por um curto intervalo de tempo. Então a observação de uma variação de intensidade na imagem é traduzida necessariamente pela existência de movimento na cena.

Formalmente, dizemos que a luminosidade aparente de um ponto  $\mathbf{p}(t)$  não varia com tempo  $t$ , o que nos leva a escrever:

$$I(t; \mathbf{p}(t)) = I_0$$

Chegando a equação do fluxo ótico descrita abaixo,

$$\frac{\partial I}{\partial t} + \nabla I \cdot \frac{d\mathbf{p}}{dt} = 0$$

ou ainda,

$$v \cdot \nabla I + \frac{\partial I}{\partial t} = 0 \quad (2.6)$$

A Equação (2.6) não pode ser resolvida pontualmente, pois em cada ponto dispomos apenas de uma única grandeza escalar para encontrar duas incógnitas ( $v_1, v_2$ ) ou mais. É o que chamamos de problema de abertura, ver Figura 2.12. Enquanto nenhuma informação suplementar sobre o fluxo a estimar estiver disponível, a única equação que dispomos é a equação do fluxo ótico. Logo o problema está mal posto.

Podemos considerar uma imagem regular como localmente invariante por translação. Somente a componente paralela ao gradiente  $\nabla I$  do deslocamento entre  $t = 0$  e  $t = 1$  pode ser estimada. A componente ortogonal não pode ser medida.

A única maneira de impedir esses problemas reduzindo o número de soluções, é fazer uma hipótese suplementar sobre o fluxo. Isto se aplica a todas as técnicas de medida do fluxo, sem exceção. Por exemplo, Horn e Schunck [eBGS81] substituiram o sistema linear não inversível constituído por todas as equações do fluxo para todos os pontos de medida  $\mathbf{p}$  possíveis em um sistema proposto fazendo a hipótese de que a “melhor” solução deste sistema é a solução mais regular.

O que quer dizer, minimizar a função:

$$M[v] = \iint \left( v \cdot \nabla I + \frac{\partial I}{\partial t} \right)^2 dp_1 dp_2$$

No caso do projeto RESSAC foi desenvolvido no centro ONERA de Chatillon, o algoritmo FOLKI [eFC05]. Se trata de um algoritmo rápido, iterativo e convergente. Por outro lado, para aplicações em tempo real, a aplicação do fluxo ótico demandou um alto tempo de cálculo. Para contornar problemas oriundos do algoritmo de fluxo ótico vejamos o método explicitado a seguir.

## 2.8 Detecção de Odobe-Bouthemy

Um método recente proposto por Odobe-Bauthemy, [ePB97], consiste em detectar as regiões onde o movimento aparente não está em conformidade ao movimento dominante do fundo decorrente da movimentação da câmera. Atualmente, é a solução

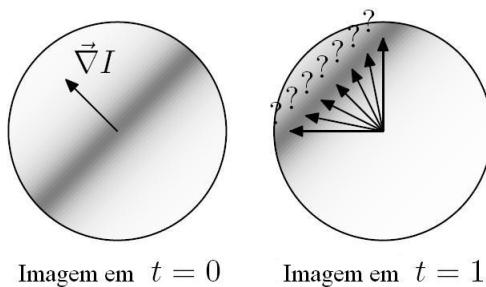


Figura 2.12: Ilustração do problema de abertura.

ção implementada nos helicópteros do projeto RESSAC.

Ele se baseia no fato de o movimento aparente induzido pela câmara móvel ser representado por um modelo de movimento paramétrico 2D, compensado pelo uso de valores de parâmetros do modelo de movimento estimado por alguma das técnicas robustas estatísticas de multiresolução. Em seguida, as regiões sob as quais o movimento não é descrito por este modelo global, estimado sobre toda a imagem, são extraídas.

### Estimação do movimento

Para estimar o movimento dominante entre duas imagens sucessivas  $I_t$  e  $I_{t+1}$ , utilizamos uma estimação robusta multiresolução baseada no gradiente. A dificuldade consiste em manter a luminosidade de uma projeção dos elementos da superfície constante segundo sua trajetória 2D [eBGS81]. O vetor do parâmetro estimado é dado por:

$$\hat{\theta} = \operatorname{argmin}_{\theta} E(\theta) = \operatorname{argmin}_{\theta} \sum_{p \in R(t)} \rho(DFD_{\theta}(p))$$

$$\text{sendo} \quad DFD_{\theta}(p) = I_{t+1}(p + \vec{w}_{\theta}(p)) - I_t(p)$$

onde  $\vec{w}_{\theta}$  é o modelo paramétrico de movimento,  $R(t)$  é toda a imagem e  $\rho(x)$  é uma função compreendida pelos autovalores de  $x$  (usamos uma função de Tukey duplamente ponderada). Se queremos detectar os objetos em movimento, a estimação baseada em  $R(t)$  leva em conta toda a imagem.

### Algoritmo de detecção de movimento

Depois de estimar o modelo dominante de movimento  $\hat{\theta}_t$ , entre as imagens  $t$  e  $t + 1$ , é preciso encontrar todos os pontos  $p$  da imagem onde o verdadeiro vetor de movimento  $\vec{w}_{true}(p)$  é diferente do vetor de fluxo estimado  $\vec{w}_{\hat{\theta}_t}(p)$ . Em seguida, antes de chamá-los de “ponto móveis”, os chamamos por um instante de “pontos não conformes”.

Introduzimos um limiar para poder definir estes tipos de pontos:

$$\left\{ \begin{array}{l} \text{se } \|\vec{w}_{err}(p)\| > \delta \text{ então } p \text{ é considerado um ponto não conforme} \\ \text{com } \vec{w}_{err}(p) = \vec{w}_{true}(p) - \vec{w}_{\hat{\theta}_t}(p) \end{array} \right.$$

Uma simples maneira de avaliar a estimativa do modelo de movimento dominante consiste em deformar uma imagem sobre a outra. Mais precisamente nós definimos  $\tilde{I}_t(p)$  e  $\tilde{I}_{t+1}(p)$  segundo:

$$\tilde{I}_t(p) = I_t(p) \text{ e } \tilde{I}_{t+1}(p) = I_{t+1}(p + \vec{w}_{\hat{\theta}_t}(p))$$

Os valores de  $I_{t+1}(p + \vec{w}_{\hat{\theta}_t}(p))$  são obtidos por uma interpolação linear pois  $p + \vec{w}_{\hat{\theta}_t}(p)$  normalmente não faz parte da malhagem da imagem. Então o campo de movimento entre  $\tilde{I}_t$  e  $\tilde{I}_{t+1}$  é exatamente  $\vec{w}_{err}$ . Se assimilamos a derivada temporal  $\frac{\tilde{I}}{\partial t}$  com a diferença entre  $\tilde{I}_{t+1} - \tilde{I}_t$ , a equação de intensidade luminosa nos dá:

$$FD_t(p) = \tilde{I}_{t+1}(p) - \tilde{I}_t(p) \simeq -\vec{\nabla} \tilde{I}_t(p) \cdot \vec{w}_{err}(p) \quad (2.7)$$

A diferença de quadros<sup>6</sup>  $FD_t$  não é uma medida local apropriada para a detecção de movimento. Uma boa medida deve ser mais significativa nas regiões não conformes e menos significativa nas regiões bem compensadas, de preferência com um comportamento contínuo ao meio. Considerando a Equação (2.7), usando a diferença de quadros:

1. a resposta é pequena em regiões de mesma intensidade em qualquer região,
2. a medida é geralmente grande em arestas com alto contraste, mesmo se existe um pequeno erro de compensação.

Tudo isso nos indica que a resposta é dada sobretudo graças ao gradiente espacial de intensidade e não graças aos movimentos residuais. Sendo assim faz-se necessário considerar uma outra medida local atrelada aos erros de compensação de movimento.

---

<sup>6</sup>FD: frame difference

Para tal consideremos o fluxo normal de erro residual do deslocamento par um ponto  $p$ :

$$v_n = \frac{FD_t(p)}{\|\vec{\nabla} \tilde{I}_t(p)\|}$$

Nesta etapa ainda temos problemas. A detecção ainda é incerta em regiões uniformes e caso o deslocamento residual  $\vec{w}_{err}(p)$  seja perpendicular à direção do gradiente espacial de intensidade, a medida será igual a zero (problema de abertura), mesmo se existe movimento.

Por outro lado essas dificuldades podem ser amenizadas a partir da análise de diversos tipos de estimativa do fluxo ótico [BFB]. Foi demonstrado que  $\|\vec{\nabla} \tilde{I}(p)\|$  é uma medida mais confiável de estimativa do fluxo normal  $v_n$ . Neste caso ao invés de calcular uma simples média do fluxo normal, usamos a média ponderada proposta por Irani e Russo [Ira94]:

$$\begin{aligned} \text{Mes}_{\hat{\theta}_t}(p) &= \frac{\sum_{q \in \mathcal{F}(p)} (\|\vec{\nabla} \tilde{I}(q)\|^2 \cdot v_n(q))}{\text{Max}(\sum_{q \in \mathcal{F}(p)} \|\vec{\nabla} \tilde{I}(q)\|^2, n \cdot G_m^2)} \\ &= \frac{\sum_{q \in \mathcal{F}(p)} (\|\vec{\nabla} \tilde{I}(q)\| \cdot |FD_t(q)|)}{\text{Max}(\sum_{q \in \mathcal{F}(p)} \|\vec{\nabla} \tilde{I}(q)\|^2, n \cdot G_m^2)} \end{aligned} \quad (2.8)$$

onde  $\mathcal{F}(p)$  é uma vizinhança (normalmente  $3 \times 3$ ) do ponto  $p$  que contém  $n$  pontos, e  $G_m$  é uma constante que representa o ruído das zonas uniformes. Supondo que o pixel  $p$  e seus vizinhos realizam o mesmo deslocamento de módulo  $\delta$  e direção  $\vec{u}$ , a partir do estudo das variações desta medida em relação a direção de  $\vec{u}$ , podemos delimitar duas fronteiras  $l_p$  e  $L_p$  pelas quais para qualquer direção  $\vec{u}$  a seguinte inequação é respeitada.

$$o \leq l(p) \leq \text{Mes}_{\hat{\theta}_t}(p) \leq L(p)$$

Em [Odo94], o autor determina as duas fronteiras. Considerando a Equação de intensidade (2.7), temos:

$$\begin{cases} L(p) = \delta \\ l(p) = \eta \times \delta \times \lambda'_{min} \end{cases}$$

sendo  $\eta = \frac{\sum_{q \in \mathcal{F}(p)} \|\vec{\nabla} \tilde{I}(q)\|^2}{\text{Max}(n \times G_m^2, \sum_{q \in \mathcal{F}(p)} \|\vec{\nabla} \tilde{I}(q)\|^2)}$

$$\text{e } \lambda'_{min} = \frac{\lambda_{min}}{\lambda_{max} + \lambda_{min}}$$

onde  $\lambda_{min}$  e  $\lambda_{max}$  são respectivamente o menor e o maior autovalor da seguinte matriz ( $\nabla \tilde{I}(q) = (\tilde{I}_x(q), \tilde{I}_y(q))$ ):

$$M = \begin{pmatrix} \sum_{q \in \mathcal{F}(p)} \tilde{I}_x(q)^2 & \sum_{q \in \mathcal{F}(p)} \tilde{I}_x(q) \tilde{I}_y(q) \\ \sum_{q \in \mathcal{F}(p)} \tilde{I}_x(q) \tilde{I}_y(q) & \sum_{q \in \mathcal{F}(p)} \tilde{I}_y(q)^2 \end{pmatrix}$$

Para obter fronteiras mais finas, os autores modelaram uma iso-intensidade de contorno, que passa através do pixel  $p$ , como uma união de dois segmentos que se reencontram em  $p^3$ . A partir deste modelo, eles definiram novas fronteiras.

$$\begin{cases} l(p) = \eta \delta \sqrt{\lambda'_{min}(1 - \lambda'_{min})} \\ L(p) = \delta \sqrt{1 - \lambda'_{min}} \end{cases}$$

Considerando  $l$  e  $L$  dependentes da distribuição local de direções do gradiente espacial de intensidade, isto indica que a estrutura local de intensidade da imagem nos permite tratar o problema de abertura. Se a iso-intensidade linear do contorno desliza sobre si mesma, a medida dada pela Equação (2.8) será próxima de zero, mesmo existindo movimento real. Entretanto, neste caso a fronteira  $l$  também estará muito próxima de zero. Podemos concluir que uma pequena medida vai indicar uma zona segura em conformidade no caso de ser menor que a fronteira  $l$ .

Querendo detectar o movimento residual  $\vec{w}_{err}$  de uma magnitude maior que o valor de  $\delta$ , as medidas locais dadas pela Equação (2.8) serão classificadas em três classes, ver Figura 2.13. Primeiramente uma observação menor que  $l(p)$  nos assegura que a magnitude do deslocamento residual  $\|\vec{w}_{err}\|$  é menor que  $\delta$ , classificando o ponto como solo. Em segundo lugar, quando a observação local é maior que  $L$

podemos dizer que a magnitude do deslocamento residual é maior que  $\delta$ , classificando o ponto em questão como relevo. Finalmente, quando os valores observados cairem entre as duas fronteiras, não podemos tirar nenhuma conclusão. Neste caso, poderíamos usar outro contexto espacial e temporal para classificar corretamente o pixel correspondente.

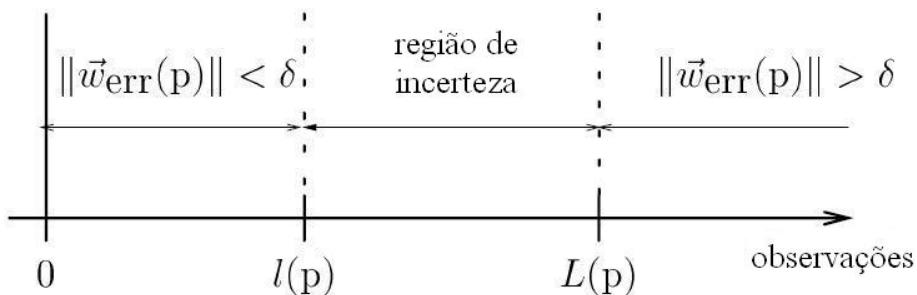


Figura 2.13: Classificação de medidas.

## 2.9 Trabalhos realizados

Minha contribuição com a parte de visão do projeto se deu inicialmente com o estudo completo do problema. Em seguida implementei um detector de pontos de Harris usando a biblioteca de visão computacional OpenCV. Essa implementação serviu para aprimorar o detector existente e determinar um bom compromisso entre o número de pontos detectados e o reajuste entre as imagens.

Como o modo esparsa foi tema de uma pesquisa anterior e não apresentou resultados satisfatórios, concentrei minha pesquisa na modelagem densa do ambiente, havendo duas possibilidades de escolha, o fluxo ótico ou a detecção Odobez-Bouthemy.

Para o fluxo ótico havia uma solução já implementada denominada FOLKI. Para a solução que compreende a detecção de Odobez-Bouthemy foi preciso escrever rotinas complementares de forma a adequá-la à realidade do projeto. Ficou constatado que o algoritmo de Odobez-Bouthemy é mais eficiente para o nosso objetivo, visto que usa parte dos cálculos de reajuste da imagem realizados anteriormente.

O processamento estereoscópico envolve diversas etapas, cada uma com parâmetros variáveis, a exemplo do tamanho da janela do detector de Harris e do número de pontos usados na correlação ida e volta. Para o bom funcionamento da solução, realizei diversos testes com imagens reais para escolher os melhores parâmetros de forma a obter o bom resultado para a nossa aplicação. Além disso, esses testes serviram para determinar o valor ideal para a razão B sobre H, medindo o caráter estéreo dos pares de imagem.

## CAPÍTULO 3

### A CARTOGRAFIA DO TERRENO

A necessidade de informações cartográficas em meio urbano ou rural, em matéria de urbanismo ou planejamento territorial, resgate, telecomunicações ou defesa são extremamente importantes [Cor04].

As técnicas de decisão embarcada estão no coração do projeto, dessa forma buscamos a operação completamente independente do helicóptero em relação a seus objetivos.

Neste capítulo, detalharemos os procedimentos para a constituição da cartografia da zona explorada pela aeronave. Após realizar o processamento de imagens explicitado no capítulo anterior, é preciso acumular as decisões feitas sobre os pares de imagem em um modelo da zona explorada.

#### 3.1 O helicóptero

Segundo a definição oficial, um helicóptero é “*uma aeronave de asas rotativas propulsada por um ou mais rotores horizontais (propulsores) que, quando girados pelo motor, criam sustentação e propulsão necessárias para o vôo.*”

Cada rotor de sustentação, com o eixo na direção vertical, é uma espécie de hélice com variação de passo, e comporta duas (no caso do nosso projeto) superfícies

aerodinâmicas mais conhecidas como pás, que servem para deslocar o aparelho tanto verticalmente como horizontalmente.

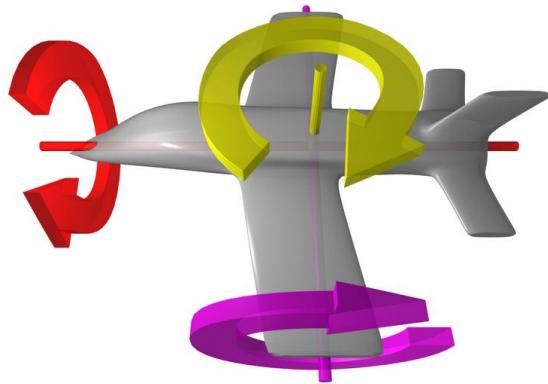


Figura 3.1: Movimentos 3D de uma aeronave

Comparando com os aviões, os helicópteros são muito mais complexos, mais caros na aquisição e manutenção e operam com velocidade reduzida, com pouca autonomia e com pouca capacidade de carga. A vantagem obtém-se na capacidade de manobra: helicópteros podem pairar, inverter a trajetória e, acima de tudo, podem decolar e pousar com vôo vertical. Dependendo do reabastecimento e da quantidade de carga, um helicóptero pode viajar para qualquer lugar e aterrissar em qualquer lugar desde que haja espaço no local.

Como mencionado acima, ele possui uma vantagem considerável sobre o avião: sua possibilidade de realizar um vôo estacionário<sup>1</sup> e realizar um vôo com velocidade reduzida, o que é importante para a nossa aplicação já que visamos explorar uma determinada área do terreno.

## 3.2 Movimentos 3D do helicóptero

Antes de considerar os resultados dos algoritmos de processamento de imagem, iremos esclarecer as possíveis situações e referenciais. Uma vez que a câmera é fixada perpendicularmente à parte inferior do helicóptero, a tomada de imagens é feita sob posições e ângulos diferentes, ver Figura 3.1. As condições do vôo nunca

<sup>1</sup>manter uma posição fixa em vôo

serão idênticas, mesmo se a missão carregada for igual tendo os mesmos pontos de passagem, pois teremos influências externas como o vento por exemplo.

No vôo estacionário, Figura 3.2, o plano da câmara está horizontal ao solo, sendo o caso ideal para o processamento de imagens a ser realizado *a posteriori*.

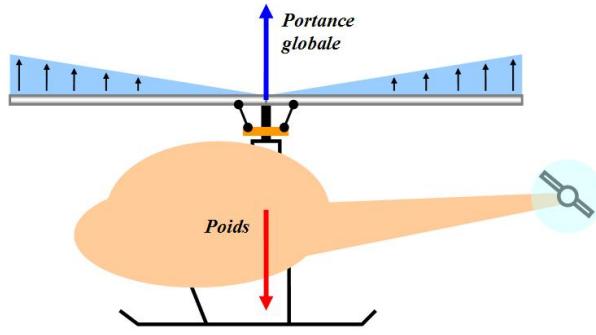


Figura 3.2: Helicóptero em posição estacionária

No vôo horizontal, o controle de passo de hélice é ajustado para que o ângulo de cada lâmina aumente à medida que ela se move para a parte de trás de sua varredura. Isto faz com que a aeronave se incline para a frente, dando-lhe uma componente de empuxo horizontal, além de sustentação. Neste tipo de vôo para que o helicóptero avance, é preciso que haja uma componente de força horizontal, ver Figura 3.3.

O princípio é o mesmo para qualquer movimento desejado.

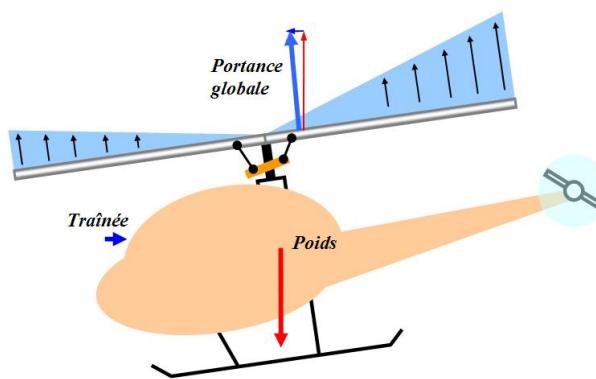


Figura 3.3: Helicóptero em deslocamento

Além disso nestes movimentos o helicóptero ainda é suscetível a inclinações laterais e para trás, além do fenômeno de derrapagem que é intrínseco ao vôo destes equipamentos.

Para tentar compensar essas características no momento que cada imagem é tirada, armazenamos também informações referentes a inclinação e posição da aeronave. Estes dados auxiliares serão em seguida usados para fazer as transformações de coordenadas necessárias, de modo a reajustar no mesmo referencial o par de imagem a ser processado.

O movimento do helicóptero em relação a um referencial, engloba seis coordenadas descrevendo a posição espacial do seu centro de massa, além de três ângulos, chamados ângulos de Euler, que descrevem sua orientação.

### 3.2.1 Situações

No que tange a representação das situações, por definição, a situação  $S_{i,j}$  do referencial  $R_j$  é a expressão no referencial  $R_i$  do deslocamento  $D_{i,j}$  que leva ao referencial  $R_i$  sobre o referencial  $R_j$ . O conjunto de deslocamentos de uma espaço euclidiano afim é um grupo contínuo; de dimensão 6 se o espaço se trata de  $\mathbb{R}^2$ , então sua expressão no referencial depende de 6 parâmetros.

De modo geral, consideramos todo deslocamento  $D$  composto por uma rotação  $R$  e uma translação  $T$ . Dizer que o ponto  $M_2$  é imagem do deslocamento  $D$  do ponto  $M_1$ , é o mesmo que:

$$M_2 = D(M_1)$$

sendo equivalente a dizer que o vetor  $V_2 = OM_2$  é obtido do vetor  $V_1 = OM_1$  pela expressão:

$$V_2 = R(V_1) + T \tag{3.1}$$

qualquer que seja o ponto  $O$  escolhido em  $\mathbb{R}^3$ .

O deslocamento  $D_{i,j}$  é composto de uma translação  $T_{i,j}$ , que desloca a origem de  $R_i$  sobre a de  $R_j$  a ainda de uma rotação  $R_{i,j}$  que alinha os dois referenciais. Em consequência, a situação  $S_{i,j}$  é composta por uma posição  $P_{i,j}$  que exprime  $T_{i,j}$  em  $R_i$  e uma orientação  $O_{i,j}$  que exprime  $R_{i,j}$  em  $R_i$ .

Classicamente  $P_{i,j}$  é representado por uma matriz coluna de 3 parâmetros, e  $O_{i,j}$  por uma matriz quadrada  $3 \times 3$  da qual as colunas são vetores de base  $R_j$  expressos em  $R_i$ .

A translação  $T_{i,j}$  e a rotação  $R_{i,j}$  são objetos geométricos intrínsecos que talvez possam ser necessários fora de  $R_i$ . Se fizermos a notação - provisória -  $A_{i,j}$  como a matriz de mudança de base entre  $R_i$  e  $R_j$  (levadas a mesma origem), e  $T_{/i}$  ( $R_{/i}$ ) expressão em  $R_i$  do vetor de translação  $T$  (do operador de rotação  $R$ ), o elo entre a situação e o deslocamento se entende por:

$$P_{i,j} = T_{i,j/i}$$

$$O_{i,j} = R_{i,j/i}$$

e as fórmulas de mudança de base:

$$T_{i,j/k} = A_{k,l}T_{i,j/l}$$

$$R_{i,j/k} = A_{k,l}R_{i,j/l}A_{l,k}$$

As definições da matriz de orientação  $O_{i,j}$  e da matriz de mudança de base  $A_{i,j}$ , nos conduzem a identificá-las e obtemos:

$$T_{i,j/k} = O_{k,l}T_{i,j/l}$$

$$R_{i,j/k} = O_{k,l}R_{i,j/l}O_{k,l}^{-1}$$

Notaremos  $M_i$  a posição do ponto  $M$  em  $R_i$ , a Equação (3.1) - escolhendo como origem  $O$  de  $R_i$  - expresso em  $R_i$  leva a:

$$M_i = O_{i,j}M_j + P_{i,j}$$

Mostrando que a matriz de representação é uma versão melhor adaptada para avaliar a posição de um ponto qualquer pertencente ao objeto  $obj_1$ , do ponto de vista do referencial agregado ao objeto  $obj_2$ , se conhecemos a situação do  $obj_1$  do ponto de vista do  $obj_2$ .

As matrizes de orientação são definidas por 9 parâmetros, ou elas são unitárias dependendo apenas de 3 parâmetros que fornecem uma representação alternativa das orientações.

Propomos a utilização dos ângulos de Euler associados as três rotações sucessivas em torno dos eixos  $ZXY$ .

### 3.2.2 Os ângulos de Euler

Os ângulos de Euler, ver Figura 3.4, servem para representar a orientação do helicóptero em relação a um referencial fixo. No caso do projeto RESSAC usamos os ângulos associados às três rotações sucessivas em torno dos eixos  $ZXY$ . Outras orientações podem ser efetuadas ( $ZXY$  para os giroscópios, ou  $ZYX$  em aeronáutica), mas este trio de ângulos  $(l, t, r)$  -  $l$  para *lacê*,  $t$  para *tangagem* e  $r$  para *rulis* - associado ao trio de rotações  $ZXY$  é clássico em robótica.

Fazendo a notação  $c_x$  para  $\cos(x)$  e  $s_x$  para  $\sin(x)$ , a matriz de orientação  $O$  associada ao trio  $(l, t, r)$  é:

$$\mathbf{X} = \begin{bmatrix} c_l c_r - s_l s_t s_r & -s_l c_t & c_l s_r + s_l s_t c_r \\ s_l c_r + c_l s_t s_r & c_l c_t & s_l s_r - c_l s_t c_r \\ -c_t s_r & s_t & c_t c_r \end{bmatrix}$$

Reciprocamente fazemos a notação  $o_{i,j}$  os elementos da matriz  $O$  obtemos:

$$l = \text{arctan2}(-o_{12}, o_{22})$$

$$t = \arcsin(o_{32})$$

$$r = \text{arctan2}(-o_{31}, o_{33})$$

na qual a função  $\text{arctan2}(y, x) = \arctan(y/x)$  fornece um ângulo definido por cerca de  $2\pi$ .

Fora a singularidade  $t = \pm\pi/2$ , a correspondência entre a matriz  $O$  e o trio  $(l, t, r)$  é bijetora (porque a função  $\arcsin$  fornece  $t \in [-\pi/2, \pi/2]$ ) e o termo de orientação recobre uma das duas representações segundo o contexto. Porém, recomenda-se

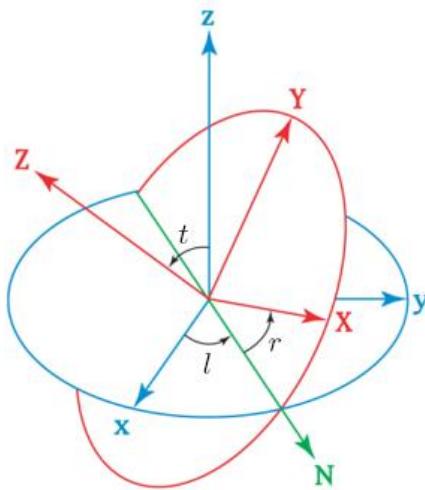


Figura 3.4: Ângulos de Euler

que no caso de uma orientação suscetível a variações deve-se evitar a vizinhança de singularidade e realizar a escolha em virtude do eixo  $Y$  do segundo referencial o mais longe possível do eixo  $Z$  do primeiro.

### 3.2.3 Projeções no solo

Para compor corretamente o mapa topográfico deve-se levar em conta o estado do helicóptero no momento da tomada das imagens. Quando o dispositivo embarcado captura uma imagem, são associados os dados referentes aos seis parâmetros relativos da aeronave ao mesmo tempo. Os dados serão registrados referentes às posições  $x, y, z$ , e também aos ângulos  $l, t, r$ , além disso é essencial gravar o momento exato em que a imagem foi tirada para a parte de processamento. O intervalo de tempo servirá para calcular o caráter estéreo através da razão base (B) sobre altura (H).

A Figura 3.5 mostra o ângulo de visão da câmera projetada no solo. A Figura 3.6 mostra a rotação em torno do eixo  $Z$ , ângulo *lacê*. É possível perceber o movimento de rotação da projeção da câmera.

Supondo uma variação no ângulo de *tangagem*, em torno do eixo  $X$ , quando o helicóptero se inclina para frente ou para trás tem-se uma deformação da projeção feita pela câmera no solo, como na Figura 3.7

Para finalizar, quando o helicóptero se inclina de lado tem-se uma variação no ângulo *rulis* como observado na Figura 3.8.

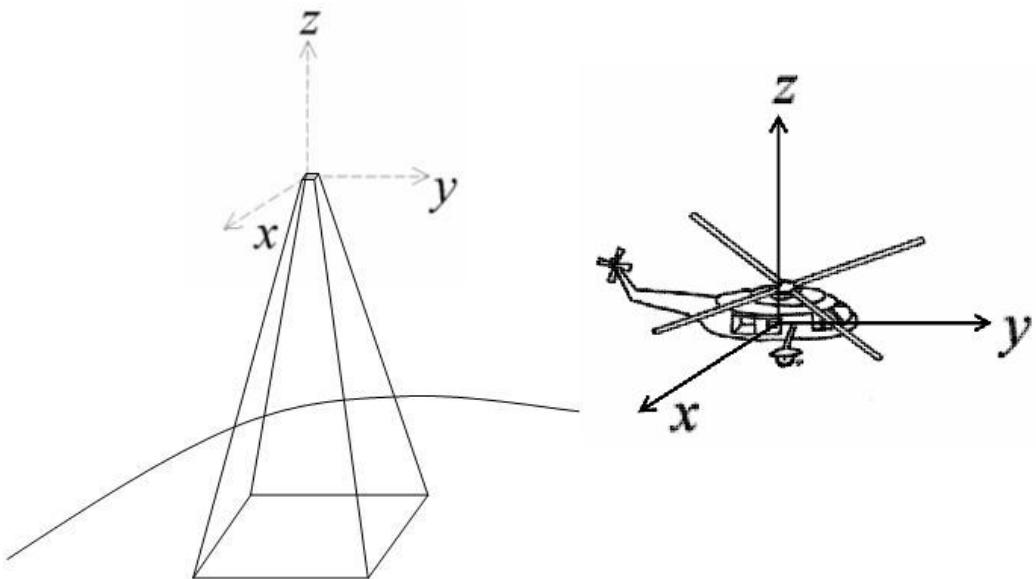


Figura 3.5: Projeção da câmera no solo

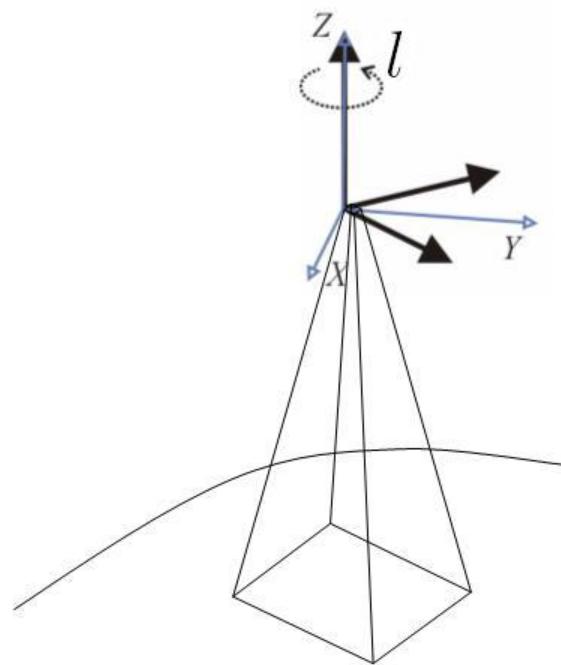


Figura 3.6: Movimento em torno do eixo Z

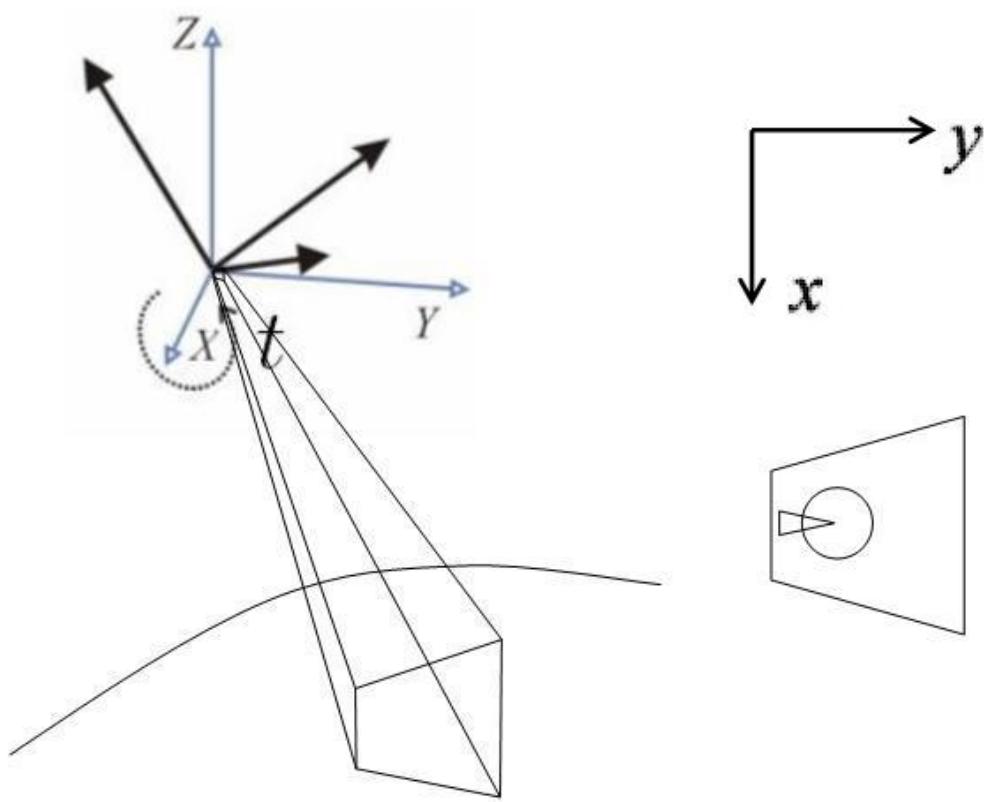


Figura 3.7: Movimento em torno do eixo X e projeção no solo

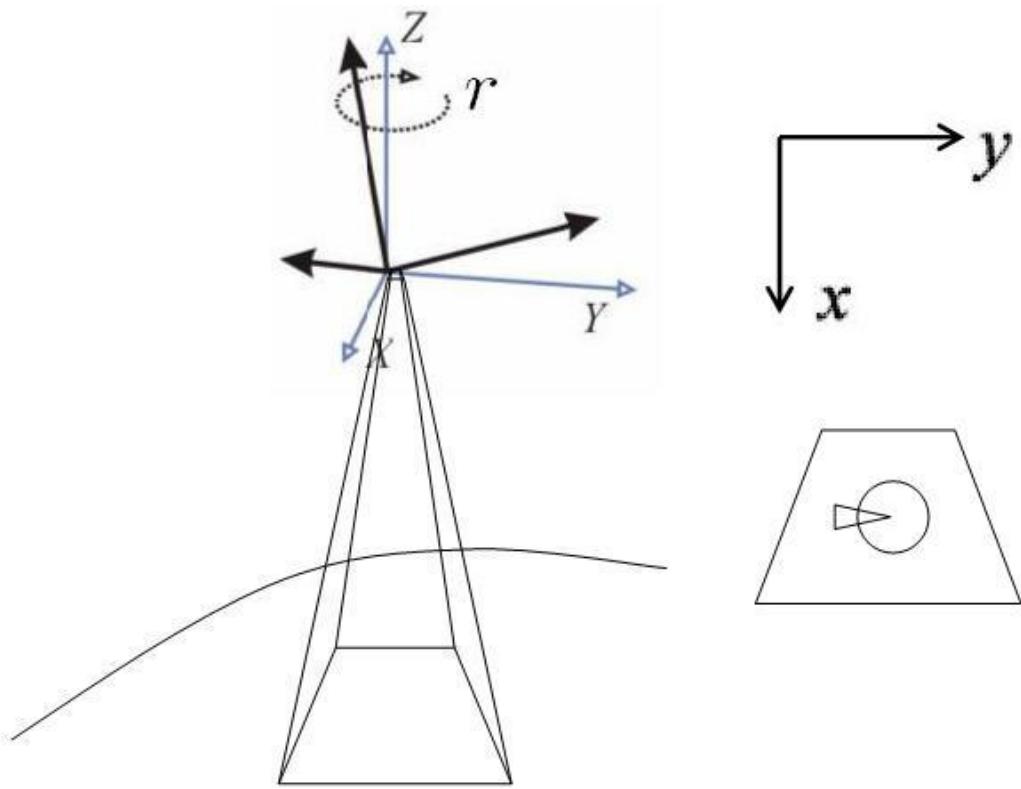


Figura 3.8: Movimento em torno do eixo Y e projeção no solo

O algoritmo Odobe-Bouthemy, associa a cada pixel uma variável que pode ter 3 valores. Tais valores serão projetados no solo suposto  $Z = 0$  sobre uma malhagem fixa (células de  $0,5 \times 0,5 m^2$ ), compreendendo a situação do helicóptero, ou seja, a situação da câmera no momento da tomada de imagens.

As células sobre esta malhagem são estocadas em urnas (aceitável se  $< l$ , inaceitável se  $> L$ ) que fornecem uma cartografia bruta da zona sobrevoada.

A partir da cartografia bruta são calculadas as zonas de interesse e de obstáculos.

### 3.3 Trabalhos realizados

Minha contribuição com a parte da composição dos mapas, se deu na implementação da interface gráfica para o campo de testes de Caylus usando o toolkit multi-plataforma GTK+<sup>2</sup>. Essa implementação consistiu em usar uma foto aérea do terreno para projetar os resultados obtidos na análise de cada par de imagens.

<sup>2</sup>[www.gtk.org](http://www.gtk.org)

Em seguida realizei testes para os diversos filmes registrados pelo helicóptero. De acordo com a classificação de cada célula do terreno obtem-se a cartografia bruta. Esse processo, assim como na parte visão, também serve para validar o algoritmo.

Partindo dessa cartografia bruta, separo duas zonas: as classificadas como obstáculos e as classificadas com planas. Na seqüência, a partir de um esqueleto das zonas planas é adicionado uma margem de segurança para impedir que o helicóptero se aproxime dos obstáculos. Tal procedimento será ilustrado no próximo capítulo.

## CAPÍTULO 4

---

### RESULTADOS

#### 4.1 Síntese dos resultados

Como já mencionado ao longo do texto, o inconveniente do modo esparso é não detectar zonas de relevo quando não há pontos de interesse sobre a superfície em relevo a serem reastreados. Diversas árvores presentes nas cenas não foram detectadas em virtude dos pontos a serem rastreados apresentarem baixo contraste, além disso, os fenômenos de ocultação entre a árvore e o solo atrapalhavam o casamento. Para definir a qualidade de cada pixel (no solo, em relevo ou indefinido) pode-se pensar em um teste sobre todos os pixels da imagem. Trata-se do modo denso que fornece uma decisão por pixel.

Na Figura 4.1 aplicou-se o algoritmo de cálculo do fluxo ótico, FOLKI , entre as imagens 9 e 10 (reajustadas pelo movimento dominante) da seqüência Esperce80. O algoritmo realiza a estimativa do fluxo ótico residual. O deslocamento e o modo de estimativa local é um indicador do relevo, sobre certas hipóteses. À esquerda há a representação em campo vetorial (percebe-se que os vetores próximos apresentam a mesma direção, o que é resultado essencialmente de um efeito de paralaxe), à direita está a norma do deslocamento, onde as árvores estão distintas, e ao centro da imagem há uma pequena depressão em relação ao solo estimado, representada pela norma mais baixa.

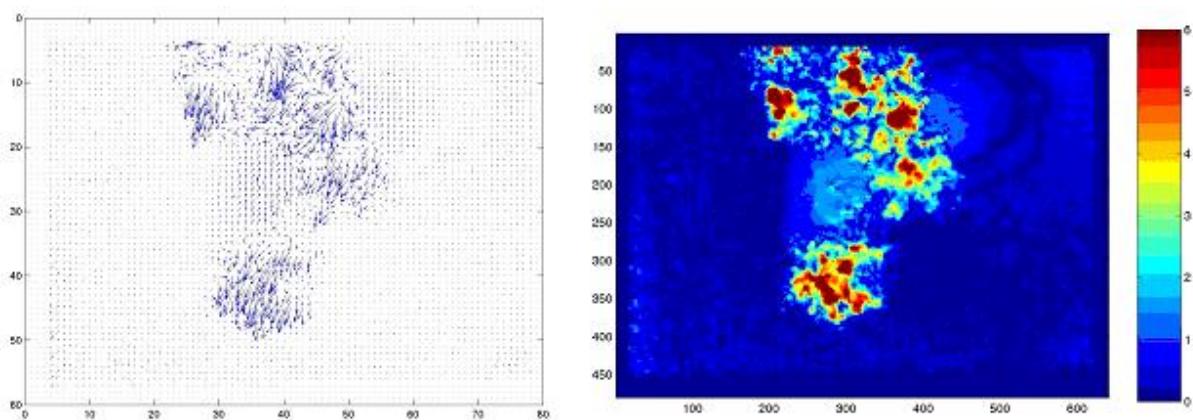


Figura 4.1: Estimação do fluxo ótico residual entre as imagens 9 e 10

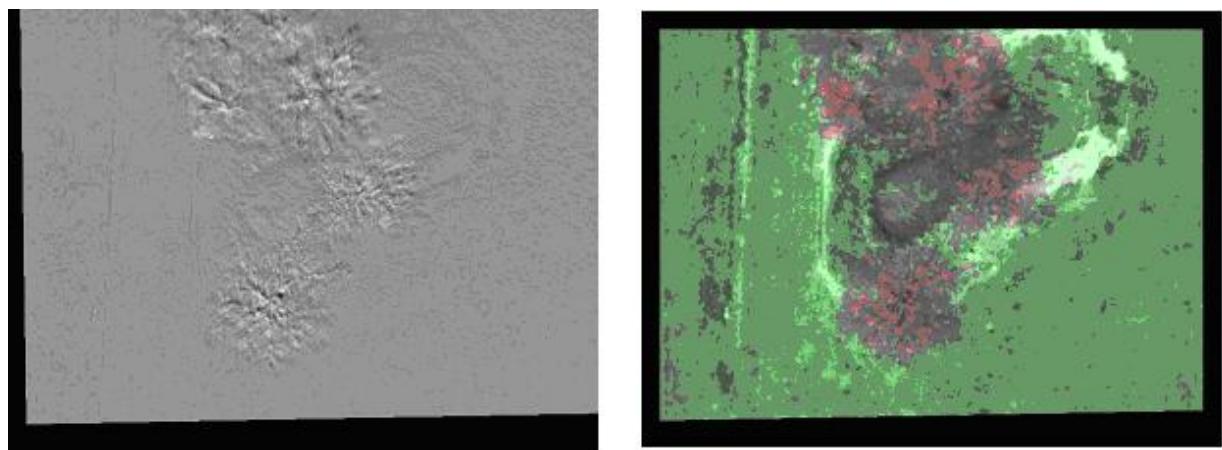


Figura 4.2: À esquerda a imagem de DFD, à direita a imagem com o resultado da detecção

Na Figura 4.2, temos à esquerda a imagem de DFD entre as mesmas imagens 9 e 10 (trata-se da imagem de diferenças reajustada por uma transformação quadrática dominante). À direita está a imagem com o resultado da detecção. Os pontos verdes indicam o solo, o relevo está indicado pelos pontos vermelhos e os indefinidos estão indicados em cinza. O limiar de deslocamento escolhido é de 1.5 pixel.

## 4.2 Análise dos filmes

Diversos ensaios de campo foram realizados até o momento. Eles são essenciais para o desenvolvimento dos dispositivos de navegação, e também para o ajuste dos parâmetros relativos a visão. O objetivo é realizar os testes do *hardware* embarcado que sejam necessários para validá-los numa situação real permitindo a repetibilidade das manobras.

Os vôos do helicóptero foram feitos principalmente no aeródromo denominado “Esperce”, localizado a alguns quilômetros da cidade de Toulouse. O helicóptero também realizou algumas experiências numa zona militar denominada “Caylus”. As seqüências dadas como exemplo têm esses nomes em referência ao local onde as imagens foram registradas. O campo de Caylus é muito interessante para a parte do projeto compreendida neste relatório, a visão, pois somente neste local foi possível sobrevoar uma região composta por casas, estradas e vegetação. A limitação imposta pelos orgãos reguladores do espaço aéreo dificultam a realização de testes em áreas habitadas.

Nas Figuras 4.3 e 4.4, vê-se a saída do algoritmo de detecção implementado. As zonas vermelhas são consideradas como obstáculos, as zonas verdes são consideradas planas e aquelas em que não há cor foram consideradas indefinidas.

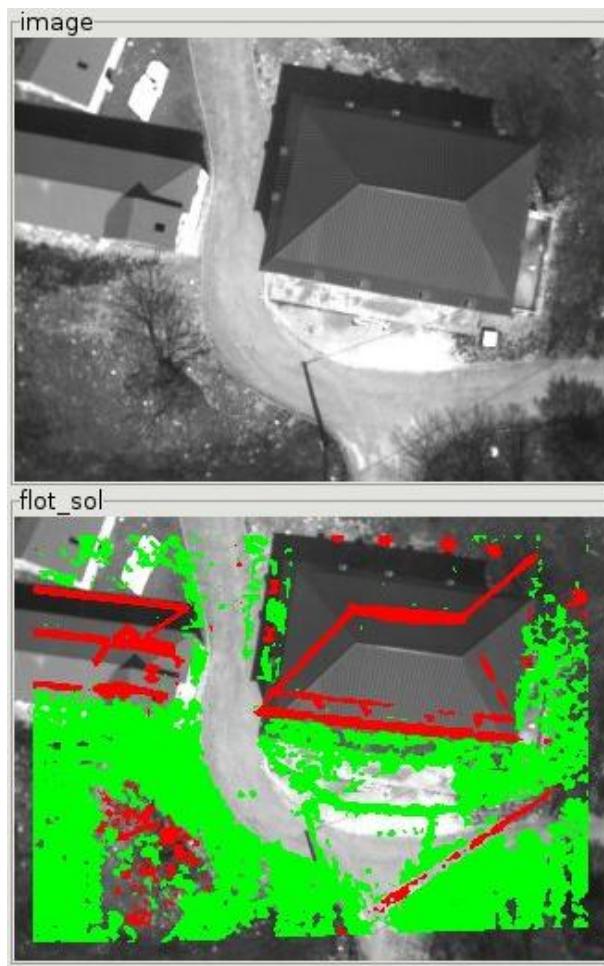


Figura 4.3: Detecção sobre o par de imagens 183-84, do filme Caylus000

É possível observar na Figura 4.3 que as quinas do telhado, assim como a fiação elétrica, foram detectados como obstáculos. Por outro lado, regiões como a estrada e o meio de telhado foram consideradas como indefinidas devido a homogeneidade dos pontos.

Na Figura 4.4 observa-se a detecção de árvores em dois pares distintos de imagem do mesmo filme Caylus002.

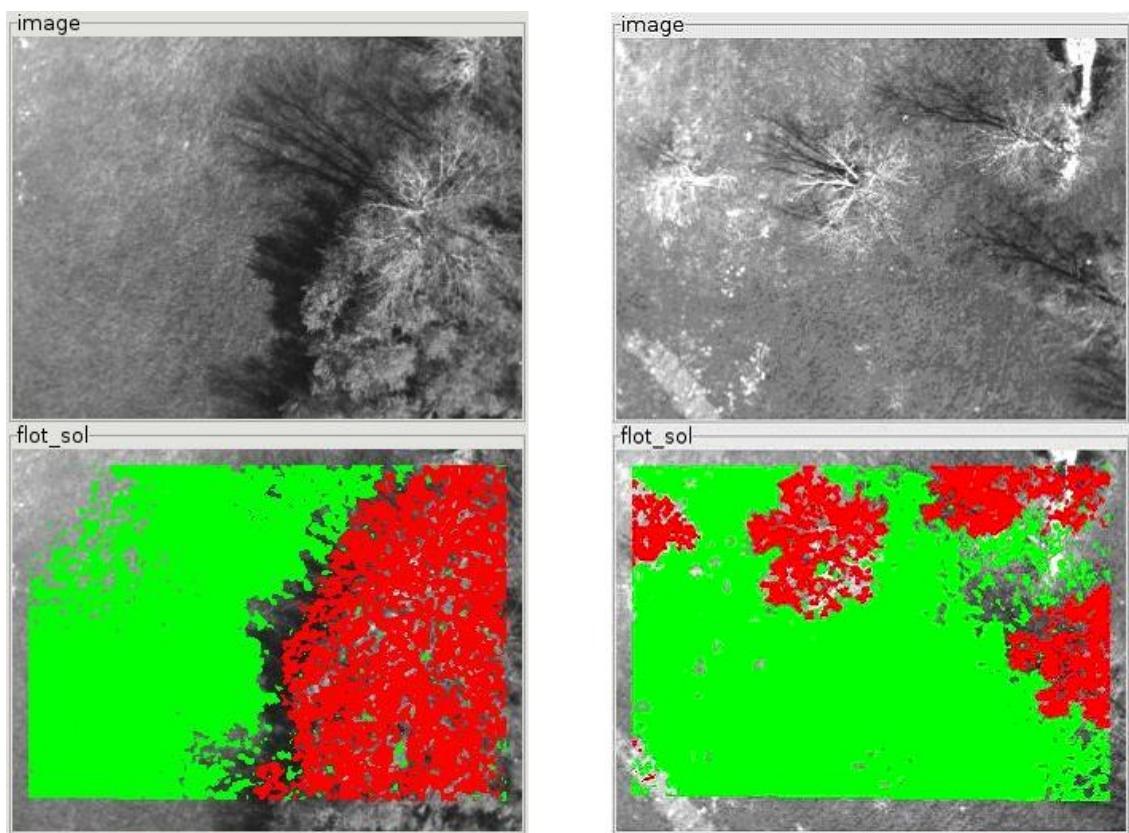


Figura 4.4: À esquerda: par de imagens 177-178. À direita: par de imagens 192-193, filme Caylus002

Na etapa seguinte de cartografia do terreno, a imagem aérea do campo de testes Caylus, ver Figura 4.5, serve de base para a projeção da saída do algoritmo de detecção aplicado a cada par de imagens, chegando finalmente à composição das zonas de aterrissagem, Figura 4.6.

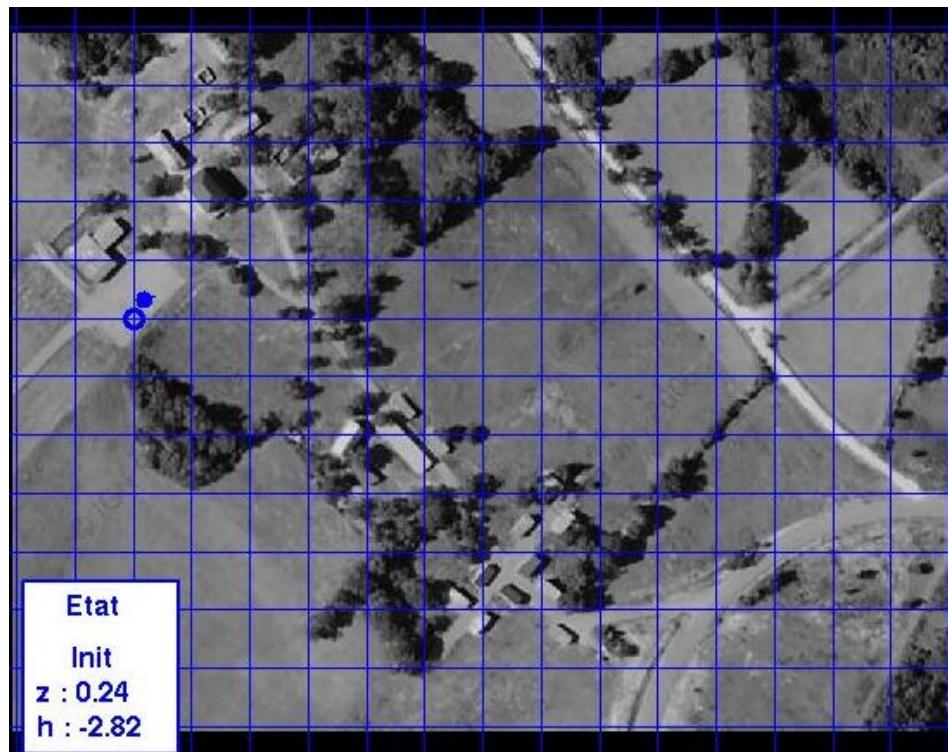


Figura 4.5: Campo de teste Caylus

A cartografia só pode ser feita nas zonas varridas pela aeronave. As regiões coloridas da Figura 4.6 representam o traço da câmara no solo. Para cada ponto que faz parte do traço da câmara projeta-se no terreno o resultado do algoritmo de detecção. Como existe superposição dos pares de imagem no solo, a classificação foi determinada em função do número de vezes que cada ponto foi classificado como um obstáculo ou como uma zona plana, para finalmente atribuir a cor vermelha para o obstáculo e a cor amarela para a zona plana.

Constata-se ainda para as imagens da mesma Figura 4.6, no alto à esquerda, a foto aérea da região sobrevoada. No alto à direita, os quadrados vermelhos representam os obstáculos identificados e a zona amarela é vista como sem obstáculos. Em baixo à esquerda, realiza-se um zoneamento de regiões de aterrissagem a partir de um esqueleto das zonas amarelas. Finalmente, em baixo à direita, mostra-se as possíveis zonas de aterrissagem em verde.

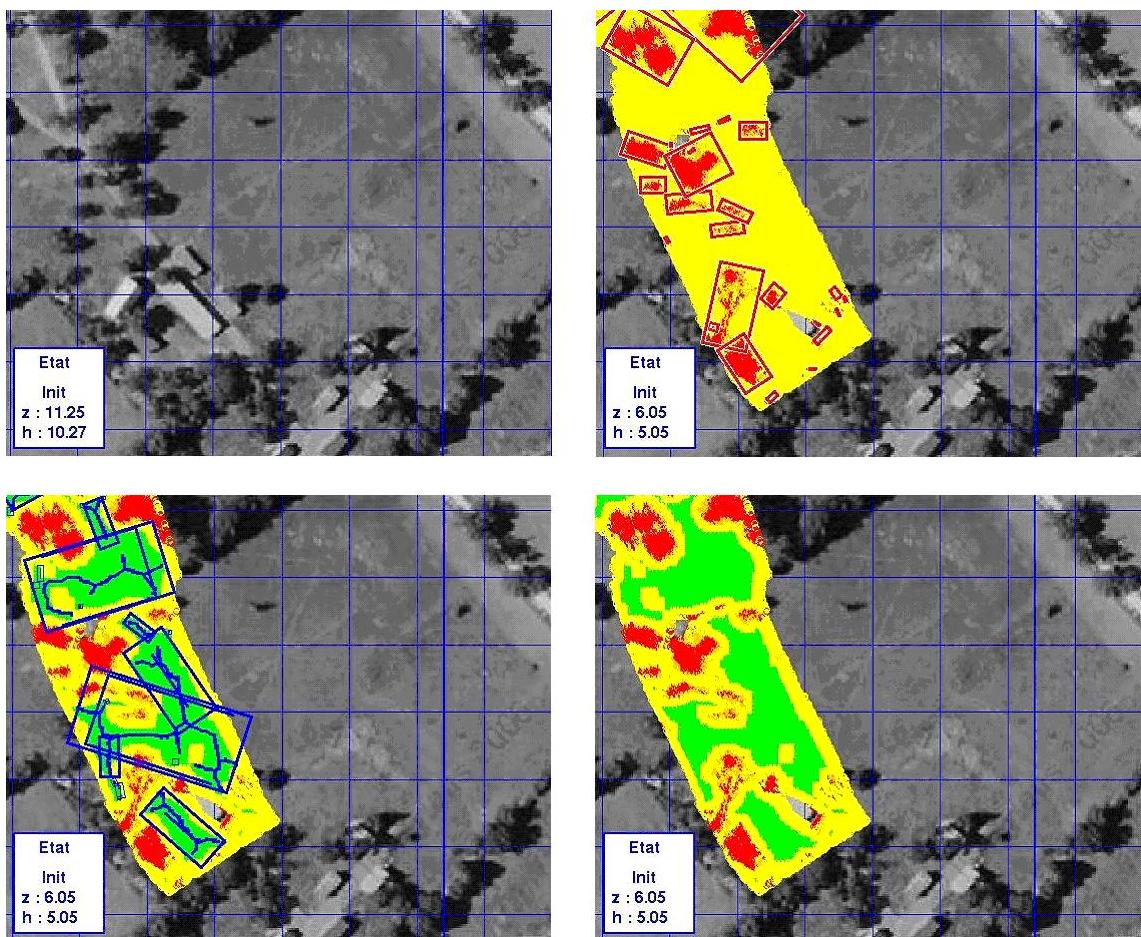


Figura 4.6: Cartografia de Caylus001

A borda amarela em torno das regiões verdes representa o raio de segurança, ou seja, atribui-se um limite de proximidade dos obstáculos para impedir que a aeronave vá além dessa marcação.

Agora sobre a Figura 4.7, é realizado uma aproximação sobre a região explorada para vê-la em mais detalhes. Continuamos tendo no alto à direita os quadrados que identificam os obstáculos, em baixo à esquerda o esqueleto da zona sem obstáculos, e, finalmente, em baixo à direita as zonas verdes de aterrissagem.

É interessante observar que o telhado da casa não é identificado como um obstáculo pela solução proposta. Como os pixels são homogêneos nosso algoritmo os classifica como uma região indefinida.

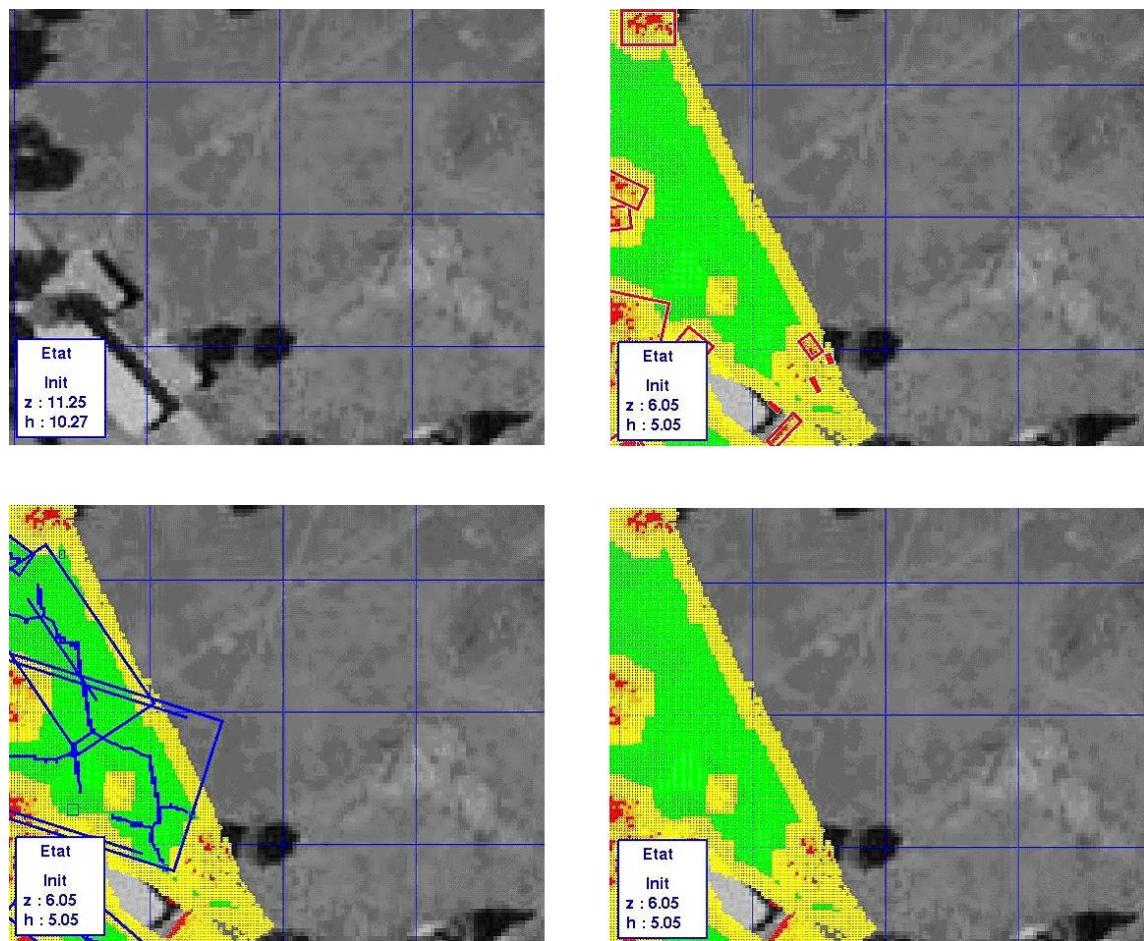


Figura 4.7: Cartografia de Caylus001 com zoom

## 4.3 Conclusão

Após explicar o princípio da parte visão do projeto RESSAC, principalmente a escolha de uma zona de aterrisagem com uma modelagem densa do ambiente, explicitamos as técnicas estudadas e implementadas.

Em suma, o modelo proposto é capaz de discernir corretamente os obstáculos e determinar as zonas de aterrisagem, chegando a uma solução operacional que permite resultados satisfatórios. Atualmente o projeto encontra-se em fase final, espera-se que o conhecimento adquirido possa solucionar problemas semelhantes no futuro.

Uma limitação a ser trabalhada é o elevado tempo de cálculo. A taxa atual não é alta o suficiente para um processamento em tempo real, o que era uma das metas especificadas.

Outra perspectiva para trabalhos futuros é a melhoria da precisão da detecção de forma a obter uma estimativa da profundidade em  $Z$ . Além disso, é possível integrar outros sensores, a exemplo do laser, para melhorar a detecção.

Este trabalho me permitiu adquirir sólidos conhecimentos em visão computacional e processamento de imagens. O fato de ter participado de tal projeto em um laboratório de grande importância como a ONERA é, na minha opinião, uma experiência de grande valor para o futuro.

---

## REFERÊNCIAS BIBLIOGRÁFICAS

- [BB95] S. S. Beauchemin and J. L. Barron. The computation of optical flow. *ACM Computing Surveys*, 27(3):433–467, 1995.
- [BFB] J.L. Barron, D.J. Fleet, and T.A. Beauchemin, S.S. e Burkitt. Performance of optical flow techniques. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR'92)*, 92:236–242.
- [Cor04] M. Cord. *Traitemet d'images: application à la reconstruction de surfaces et à la recherche interactive d'images*. PhD thesis, Université Cergy-Pontoise, 2004.
- [CS00] Roger Mohr e Christian Bauckhage Cordelia Schmid. Evaluation of interest point detectors. *International Journal of Computer Vision*, 37(2):151–172, 2000.
- [eBGS81] B.K. P. Horn e B. G. Schunck. Determining optical flow. *Artificial Intelligence*, 17:185–203, 1981.
- [eFC05] Guy Le Besnerais e Frédéric Champagnat. Dense optical flow by iterative local window registration. In *International Conference on Image Processing, ICIP (1)*, pages 137–140, 2005.
- [eMS88] C. Harris e M. Stephens. A combined corner and edge detector. In *In ALVEY Vision Conference*, pages 147-151, 1988.

- [ePB97] J. Odobez e P. Bouthemy. Separation of moving regions from background in an image sequence acquired with a mobile camera, 1997.
- [eTK91] Carlo Tomasi e Takeo Kanade. Detection and tracking of point features. Technical Report CMU-CS-91-132, Carnegie Mellon University, April 1991.
- [Fis81] Robert C. Fischler, Martin A. e Bolles. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Commun. ACM*, 24(6):381–395, June 1981.
- [GLB07] M. Sanfourche e F. Champagnat G. Le Besnerais. Dense height map estimation from oblique aerial image sequence. *submetido a Computer Vision and Image Understanding, 2nd rev*, April 2007.
- [Gra03] Cristian Grava. *Compensation de Mouvement par Réseaux Neuronaux Cellulaires*. PhD thesis, INSA-Lyon, 2003.
- [Hub81] P. J. Huber. *Robust Statistics*. Wiley, New York, 1981.
- [Ira94] Benny e Peleg Shmuel Irani, Michal e Roussou. Computing Occluding and Transparent Motions. *International Journal of Computer Vision*, 12(1):5–16, January 94.
- [Luc81] T. Lucas, B.D. e Kanade. An iterative image registration technique with an application to stereo vision. In *International Joint Conference on Artificial Intelligence, IJCAI81*, pages 674–679, 1981.
- [Mal05] E. Malis, E. e Marchand. Méthodes robustes d'estimation pour la vision robotique. In *Journées nationales de la recherche en robotique, JNRR'05*, Guidel, France, October 2005.
- [Odo94] J.-M. Odobez. *Estimation, détection et segmentation du mouvement: une approche robuste et markovienne*. PhD thesis, Université de Rennes 1, Mention Traitement du Signal et Télécommunications, December 1994.