



UNIVERSIDADE TÉCNICA DE LISBOA
INSTITUTO SUPERIOR TÉCNICO



ESTIMAÇÃO DO MOVIMENTO PRÓPRIO
DE UM OBSERVADOR MONOCULAR

CÉSAR AUGUSTO DOS SANTOS SILVA

(Licenciado)

Dissertação para obtenção do Grau de Mestre em
Engenharia Electrotécnica e de Computadores

Orientador Científico:

Doutor José Alberto Rosado dos Santos Victor

Constituição do Júri:

Doutor João José dos Santos Sentieiro

Doutor Jorge Manuel Miranda Dias

Doutor José Alberto Rosado dos Santos Victor

Lisboa, Dezembro de 1996

Agradecimentos

A tarefa de construção de um projecto é por vezes fruto de um conjunto de acasos, nunca de certezas definitivas ou de metas rigorosamente pré-estabelecidas. Os acasos são um encontro inesperado com um livro ou um amigo, são um “passeio” caótico pelo BA, são conversas intermináveis com amigos, são idas consecutivas ao -1... e todos estes acasos, sem excepção, são ideias que surgem repentinamente, alegremente, sem que a razão as aponte ou justifique.

Se hoje termino mais uma etapa de construção do meu projecto, foi porque, por um simples e maravilhoso acaso, conheci homens como João Sentieiro ou José Alberto.

Se hoje estou feliz por fazer parte da construção de outros projectos, foi porque, por um simples e maravilhoso acaso, conheci homens como Alexandre José, António, NunoP, Nuno Sena, Nuno Gracias, Etienne Grossmann, Pedro Encarnação, Artur Arsénio, Stefan Rolfes, Pedro Aguiar, António Pascoal, Carlos Carreira, João Xavier, Victor Barroso, João Paulo, Gaspar, Luís Sebastião ou João Pedro.

A todos eles, cabe aqui o meu mais profundo agradecimento.

Sobre a Tese: Eis aqui, este molho de papel, com cento e tal páginas, com um título, com ideias, capítulos, figuras e tabelas. Um pormenor no interior de um enorme quadro, sobre o qual não me atrevo falar, por ser tão grande e complexo e pessoal. Porventura, este singelo pormenor, que agora vos apresento, é a única parte desse quadro que se pretende organizada, com estrutura rígida e visível: a esse pormenor do quadro chamaram Tese... Pois tudo o resto, a Antítese (ou Anti-tese), permanece e permanecerá na neblina da minha memória, intraduzível, ilegível, definitivamente desorganizada, preciosa.

Aos homens e mulheres que viveram na Anti-tese, agradeço secretamente. Sei que alguns nunca lerão a Tese... nem na Tese jamais caberá o meu agradecimento.

Sobre a Tese, a última linha:

a José Alberto não posso eu agradecer, porque foi ele, também, seu autor.

**Esta tese foi realizada no âmbito do respectivo Curso de Mestrado, apoiado financeiramente pela bolsa PRAXIS XXI/BM/1672/94 conferida pela Junta Nacional de Investigação Científica e Tecnológica.*

Resumo

A Visão é uma capacidade sensorial privilegiada na *percepção* do meio envolvente. A sua enorme importância revela-se em aplicações robóticas tais como a locomoção, a manipulação ou o reconhecimento.

Nesta Tese, propõe-se aplicar a visão na estimação da velocidade de um observador móvel. Assume-se que o observador é monocular e que pode mover-se livremente em qualquer direcção com qualquer rotação, num ambiente estático arbitrário composto por objectos texturados. O estudo deste problema tem merecido uma grande atenção nos últimos anos da parte de diversos investigadores, e o impacte desse estudo faz-se sentir essencialmente em tarefas de navegação de sistemas autónomos.

Um dos pontos essenciais do método de estimação que propomos desenvolver, consiste no uso exclusivo das derivadas espaço-temporais da imagem, ou, equivalentemente, do fluxo normal. Evita-se, deste modo, o cálculo do fluxo óptico completo, que, devido ao *problema da abertura*, se traduz num problema mal posto.

Adicionalmente, introduz-se um novo paradigma de procura baseado nas propriedades geométricas do fluxo normal e que consiste na subdivisão do espaço das estimativas e das observações. É apresentado um conjunto de algoritmos de procura que completam o processo de estimação do movimento próprio do observador. No sentido de diminuir a sensibilidade do método ao ruído nas observações, utilizam-se ferramentas de tratamento estatístico, inspiradas em métodos de regressão robusta.

Observa-se que algumas restrições geométricas do fluxo normal, consideradas no espaço cartesiano, têm uma representação natural em outros tipos de representação da imagem, nomeadamente no espaço log-polar. Em consequência, propõe-se proceder à extensão do método para imagens log-polares.

Nesta tese, apresentam-se várias experiências com imagens reais e sintéticas para vários tipos de movimento que permitem ilustrar o desempenho do método proposto.

Palavras Chave: Visão Robótica, Análise do Movimento, Recuperação Visual do Movimento, Fluxo Óptico e Normal, Estimação Robusta, Imagem Log-Polar.

Abstract

Vision is a powerful sense to *perceive* the surrounding environment. Vision is determinant in various robotic applications like locomotion, manipulation, self-location, recognition, etc.

This Thesis addresses the problem of **egomotion estimation** for a monocular moving observer, using visual information. We assume that the observer moves under arbitrary translation and rotation, in an unknown environment, with textured objects. This problem has attracted a good deal of attention in the past and has a large impact in navigation tasks of autonomous systems.

One of the most distinctive aspects of our approach is the exclusive use of spatio-temporal image derivatives or normal flow. Thus, we avoid computing the complete optical flow field, which is an ill-posed problem due to the aperture problem.

The proposed method introduces a new search paradigm which is based on geometric properties of the normal flow field, and consists in considering a family of search subspaces to estimate the egomotion parameters. Various search algorithms are proposed within this framework. In order to decrease the noise sensitivity of the estimation methods, we carefully formulate the estimation problem, and use statistical tools, based on robust regression theory.

Additionally, we propose to extend the method for other image representations, namely the log-polar representation. Consequently, we show that some search algorithms are particularly well-suited for the log-polar image geometry, as we use a selection of special normal flow vectors with natural representation in log-polar images.

Throughout the thesis, we present and discuss a wide variety of illustrative experiments with synthetic and real images, for various kinds of camera motions.

Keywords: Computer Vision, Motion Analysis, Egomotion Estimation, Optical and Normal Flow, Robust Estimation, Log-Polar Imaging.

Índice

1	Introdução	5
1.1	Estrutura do problema	6
1.2	Trabalho relacionado	10
1.2.1	Métodos para deslocamento discreto vs. contínuo	10
1.2.2	Algoritmos iterativos vs. não iterativos	11
1.3	Estrutura da tese	12
2	Definição do Problema	15
2.1	Fluxo óptico vs. Movimento da câmara	16
2.2	Caracterização dos parâmetros a estimar	19
2.3	Estimação de parâmetros a partir do fluxo óptico	21
2.3.1	Primeiro método de estimação	21
2.3.2	Segundo método de estimação	23
2.4	A restrição do fluxo normal	25
2.5	Discussão	28
3	Subespaços de Procura	31
3.1	Definição de subespaços de procura	31
3.2	Propriedades geométricas do fluxo normal	35
3.2.1	Fluxo normal radial	35
3.2.2	Fluxo normal circular	36
3.3	Parametrização dos subespaços de procura	37
3.4	Síntese	40

4	Estimação dos Parâmetros de Velocidade da Câmara	41
4.1	Função de medida dos resíduos	41
4.2	Algoritmos de Procura	42
4.2.1	O algoritmo de procura da recta- Ψ	42
4.2.2	O algoritmo de procura da recta- Φ	45
4.2.3	O algoritmo de procura da circunferência- Γ	46
4.2.4	Algoritmo baseado em histogramas	48
4.3	Algoritmos de Procura Sequencial e Paralela	50
5	Resultados Experimentais	55
5.1	As sequências de imagens	55
5.2	Algoritmo paralelo	58
5.3	Primeiro passo numa abordagem sequencial	60
5.4	Segundo passo do algoritmo sequencial	63
5.5	Cálculo do tempo de colisão	67
5.6	Síntese	68
6	Extensão para Imagens Log-Polares	71
6.1	Método de Estimação	74
6.2	Resultados Experimentais	77
7	Conclusões	81
A	Estimação Robusta	85
A.1	Estimação LTS bidimensional	89
A.2	Aplicação do estimador LTS aos algoritmos de estimação	90
B	Hipóteses simplificativas impostas no sistema	93
B.1	Hipóteses ao nível das observações	94
B.2	Hipóteses ao nível do movimento e da cena visualizada	96
B.3	Hipóteses ao nível dos algoritmos de estimação	100

— Agora — disse eu —, imagina a nossa natureza [...] de acordo com o quadro seguinte. Imagina uns homens numa morada subterrânea em forma de caverna, cuja entrada, aberta à luz, se estende ao longo de toda a fachada; eles estão ali desde a infância, as pernas e o pescoço presos a correntes, de forma que eles não podem mudar de lugar, nem olhar para outro lado senão em frente; [...] a luz de uma fogueira acesa ao longe [...] brilha atrás deles; entre o fogo e os prisioneiros há um caminho; ao longo deste caminho imagina um pequeno muro [...].

— Vejo isso — disse ele.

— Imagina agora, ao longo deste pequeno muro, uns homens com toda a espécie de utensílios, que ultrapassam a altura do muro, e figuras de homens e de animais, em pedra, em madeira, de toda a espécie de formas [...].

— Estranho quadro e estranhos prisioneiros são esses — disse ele.

— Eles parecem-se connosco — respondi eu. — [...] pensas que nesta situação eles pudessem ver de si mesmos e dos seus vizinhos qualquer coisa senão as sombras projectadas pelo fogo sobre a parte da caverna que está em frente deles?

— Poderia ser de outro modo — disse — se eles são obrigados a ficar com a cabeça imóvel durante toda a vida?

[...]

— [...] se pudessem conversar entre eles, não pensas que acreditariam nomear os próprios objectos reais, ao nomearem as sombras que veriam?

— Necessariamente.

— E se houvesse também um eco que reenviasse os sons do fundo da prisão, de todas as vezes que um dos passantes viesse falar, não acreditas que eles tomariam a sua voz pela da sombra que desfilasse?

— Sim [...] — disse.

— É indubitável — afirmei eu — que aos olhos daquelas pessoas a realidade não poderia ser outra coisa senão as sombras dos objectos.

Platão in *A Alegoria da Caverna*, *Rep.* VII.

1. Introdução

A percepção é um processo fundamental para a sobrevivência dos seres vivos. Porque da percepção nasce uma representação do meio envolvente, e só a partir duma representação do meio, pode um ser vivo exercer acção sobre esse meio.

Percepção é a primeira forma de interacção com o mundo. Percepção significa observação dos fenómenos que nos envolvem. Dentro de todas as formas de percepção conhecidas, uma das mais poderosas e complexas é a Visão. A Visão fornece uma parte importante da informação necessária para que possamos cumprir, com sucesso, uma determinada acção. **A Visão oferece-nos uma representação espacial e temporal do mundo:** Através da visão e de uma representação visual do mundo, não só reconhecemos objectos, como também dispomos de uma “medida” qualitativa de distâncias entre corpos, estimamos a nossa própria velocidade ou a velocidade dos corpos circundantes.

A Visão tem sido alvo, ao longo dos tempos, de diversas abordagens, dando origem a teorias tão importantes como as desenvolvidas por:

- **Gibson**, que aborda a visão como um sistema perceptual que interage directamente com o meio ambiente e que é aproveitado passivamente pelo observador [Gibson, 1950];
- **Marr**, que marcou definitivamente a visão por computador e aborda a visão como um construtor de modelos internos do mundo [Marr, 1982];
- **Bajcsy e Aloimonos**, que lançam as bases da Visão Activa, assumindo que o observador se envolve activamente no próprio processo de percepção [Bajcsy, 1985, Aloimonos et al., 1988].

Nesta tese, propomos usar a informação visual de forma passiva para resolver um problema específico. A visão é um sensor que “observa” um determinado **estado** do sistema, e o estado que se propõe observar é um fenómeno interno do sistema: o seu próprio movimento. Queremos responder à seguinte questão: **é possível determinar o movimento próprio de um sistema autónomo, usando exclusivamente o sentido da visão?**

Para responder a este problema, devemos procurar na informação visual elementos que nos levem a identificar os parâmetros de movimento do observador. O problema da recuperação da velocidade própria a partir da informação visual tem atraído a atenção de um elevado número de investigadores, nomeadamente na área da visão robótica. O seu impacto é sentido em várias aplicações, especialmente nos sistemas autónomos, capazes de estimar e/ou controlar os seus parâmetros de velocidade (de forma a efectuar tarefas de alto nível — por exemplo, tarefas de atracagem ou de deslocamento entre pontos).

Para levar a cabo o estudo deste problema de estimação, define-se o sensor visual que desejamos usar. Assim, o observador móvel corresponde a um observador **monocular**, e a informação visual é adquirida por uma câmara. Neste caso, a representação do mundo visual é uma sequência de imagens; e uma imagem consiste na **projectão perspectiva** dos objectos visualizados sobre um plano.

1.1 Estrutura do problema

O problema de estimação de velocidade própria de um observador móvel monocular consiste em recuperar a sua velocidade, por observação de uma sequência de imagens adquirida ao longo do tempo.

O **primeiro passo** para a estimação é o cálculo do deslocamento dos pontos entre duas imagens consecutivas. As abordagens usuais neste campo dividem-se em três técnicas principais:

- cálculo da correspondência directa entre pontos de uma sequência de imagens [Faugeras et al., 1987, Weng et al., 1989, Weng et al., 1993];

- estimação do fluxo óptico da imagem [Horn and Shunck, 1981, Horn, 1986, Ballard and Kimball, 1983, Heeger and Jepson, 1992];
- métodos directos [Horn and Weldon, 1988, Aloimonos and Duric, 1994, Fermuller, 1995a, Fermuller and Aloimonos, 1993, Sinclair et al., 1994].

O cálculo da correspondência e a estimação do fluxo óptico implicam um conjunto de hipóteses extremamente restritivas quanto à cena observada [Terzopoulos, 1986, Tikhonov and Arsenin, 1977], e correspondem, geralmente, a problemas mal postos [Aloimonos et al., 1988, Bertero et al., 1988]. Para além disso, as soluções obtidas são, por vezes, instáveis e requerem complexos processos computacionais. No sentido de ultrapassar estas dificuldades e atingir soluções robustas, muitos investigadores optam por usar uma grande quantidade de informação redundante, adquirida no espaço e no tempo.

Por outro lado, os métodos directos são menos exigentes em termos computacionais e usam exclusivamente a informação local do brilho da imagem para recuperar os parâmetros de velocidade da câmara. A única componente do fluxo óptico calculável a partir de medidas locais do brilho é a sua componente vectorial ao longo da direcção do gradiente, conhecida por *fluxo normal* [Horn and Shunck, 1981]. Assim, os métodos directos baseiam-se usualmente no fluxo normal ou, equivalentemente, nas derivadas espaço-temporais da imagem [Santos-Victor et al., 1995, Negahdaripour and Lee, 1992]. Da mesma forma, o método que iremos propôr nesta tese passa pelo uso directo e exclusivo do fluxo normal. As razões fundamentais para esta escolha residem, por um lado, na simplicidade computacional do método, por outro lado, na possibilidade de criar um estimador de velocidade teoricamente exacto, exclusivamente a partir da informação contida no fluxo normal.

Dispondo da informação do deslocamento dos pontos entre imagens consecutivas, o **segundo passo** do processo de estimação consiste no algoritmo de recuperação do movimento tridimensional. Isto pode ser conseguido através de dois tipos de abordagens:

- a primeira abordagem está assente na resolução de equações não-lineares, usualmente através de métodos numéricos iterativos;

- a segunda abordagem resolve os mesmos problemas através de equações lineares com solução fechada ou através de algoritmos de procura finita.

Nesta tese, é adoptada a segunda abordagem. O estimador que propomos desenvolver consiste num conjunto de algoritmos de procura de complexidade reduzida.

A ideia chave do método consiste na selecção de vectores de fluxo normal com propriedades geométricas particulares, intimamente relacionadas com os parâmetros de velocidade própria do observador [Silva and Santos-Victor, 1996b, Silva and Santos-Victor, 1996a, Silva and Santos-Victor, 1996c]. Nesse sentido, em vez de se usar toda a informação visual disponível, usamos apenas um subconjunto de informação relevante no que se refere ao movimento do observador.

Como foi proposto por Heeger e Jepson [Heeger and Jepson, 1992], pode-se calcular separadamente a translação e a rotação do observador, fazendo uso das propriedades algébricas das equações do fluxo óptico. Nesta tese, partimos da mesma motivação: procuramos subdividir o problema em vários subproblemas que nos permitem estimar gradualmente os parâmetros de velocidade. No entanto, o processo de subdivisão é feito não só ao nível das estimativas, mas também ao nível das observações — em que uma observação consiste num vector de fluxo normal. Mais especificamente, o domínio das observações pode ser reagrupado em subconjuntos mais reduzidos, com características geométricas especiais. Para cada subconjunto, calcula-se uma estimativa dos parâmetros de velocidade.

No trabalho de estimação de movimento próprio realizado por Fermuller e Aloimonos [Fermuller, 1993, Fermuller and Aloimonos, 1993, Fermuller, 1995a, Fermuller, 1995b, Fermuller and Aloimonos, 1995, Fermuller and Aloimonos, 1994], exploram-se as restrições devido ao movimento, definidas nas observações do fluxo normal. Estes autores optam por uma classificação baseada nas orientações dos vectores de fluxo normal, permitindo uma selecção de vectores, de forma a construir padrões globais no plano de imagem. Como se demonstra, as posições destes padrões podem relacionar-se com o movimento 3-D do observador. Nesta tese, propomos explorar igualmente algumas associações entre a geometria do fluxo normal no plano de imagem e os parâmetros de velocidade. No

entanto, não é nosso objectivo abordar este problema segundo a perspectiva de um problema de reconhecimento de padrões. Na abordagem que aqui se propõe, pretende-se, por um lado, aproveitar mais informação existente nas observações (no sentido de usar tanto a orientação como o módulo do fluxo normal), e, por outro lado, investir sobre a redução da complexidade dos algoritmos associados.

Em oposição às abordagens dominantes que consideram os aspectos algorítmicos de estimação dissociados da geometria do sensor visual, iremos neste trabalho estudar a relação especial que se pode estabelecer entre o desenho do sensor e a concepção do estimador. Com base em exemplos da área da biologia, alguns investigadores têm adoptado mecanismos de amostragem não-uniforme da imagem, e, entre estes, o método mais popular corresponde à transformação log-polar [Weiman, 1990, Griswald et al., 1992, Tistarelli and Sandini, 1993, Wallace et al., 1995, Bernardino and Santos-Victor, 1996]. As vantagens deste tipo de representação estão essencialmente relacionadas com propriedades perceptuais e algorítmicas da geometria log-polar:

- mostra-se que os sistemas de visão, que usam imagens log-polares, estão especialmente vocacionados para certas tarefas perceptuais tais como a vergência binocular [Bernardino and Santos-Victor, 1996] ou a estimação de velocidade própria [Silva and Santos-Victor, 1997];
- a informação disponível numa imagem log-polar é menor (comparativamente com o caso cartesiano), o que significa uma redução significativa relativamente ao tempo de computação dos algoritmos associados.

Iremos mostrar, neste trabalho, que algumas restrições geométricas do fluxo normal, consideradas no espaço cartesiano, têm uma representação natural no espaço log-polar. Em consequência disso, iremos propôr um método para estimação do movimento da câmara especialmente bem adaptado à geometria log-polar.

Concluindo, os aspectos mais relevantes do método proposto nesta tese são:

- uso exclusivo do fluxo normal;

- introdução de um novo paradigma de procura, subdividindo o espaço das estimativas e das observações;
- caracterização de algumas propriedades geométricas relevantes quanto ao fluxo normal;
- introdução de ferramentas de tratamento estatístico do ruído, inspiradas em métodos de regressão robusta;
- extensão do método de estimação para geometria log-polar.

Assume-se que o observador pode mover-se livremente em qualquer direcção com qualquer rotação, num ambiente estático arbitrário composto por objectos texturados, e a imagem é formada segundo um modelo de projecção perspectiva.

1.2 Trabalho relacionado

Um largo espectro de investigação tem sido proposto no campo da estimação de parâmetros de velocidade para sistemas de visão robótica. Uma revisão exhaustiva deste tópico pode ser encontrada em [Heeger and Jepsen, 1992]. Nesta secção, propomos proceder a uma classificação simples das abordagens existentes, revendo, deste modo, os principais métodos de estimação.

Numa primeira fase, os métodos podem ser divididos em *métodos para deslocamento discreto* e *métodos para deslocamento contínuo*, consoante as observações visuais são referidas a posições discretas ou a um movimento contínuo do observador. Numa segunda fase, dependendo da abordagem algorítmica, os métodos podem ter um carácter *iterativo* ou *não iterativo*.

1.2.1 Métodos para deslocamento discreto vs. contínuo

Os *métodos para deslocamento discreto* usam uma colecção de pontos característicos extraídos duma sequência discreta de imagens (correspondentes a diferentes vistas), de forma

a construir um sistema de equações sobredeterminadas, a partir do qual é possível calcular a matriz de transformação que relaciona as diversas vistas. Os métodos para deslocamento discreto propostos em [Faugeras et al., 1987, Weng et al., 1989, Weng et al., 1993] são baseados na **matriz essencial** [Faugeras, 1993], calculada a partir de pontos correspondentes entre imagens. Os algoritmos que lhe estão associados são de baixa complexidade — sendo, no entanto, muito sensíveis ao ruído de entrada. Por outro lado, os processos de correspondência entre pontos, necessários para calcular a matriz essencial, são normalmente problemas mal postos e instáveis.

Os *métodos para deslocamento contínuo* pretendem estimar a velocidade instantânea tridimensional do observador a partir da informação do fluxo óptico [Horn, 1986, Ballard and Kimball, 1983, Heeger and Jepson, 1992]. O fluxo óptico pode ser calculado através de várias técnicas — entre as quais, as mais importantes baseiam-se na correlação e na filtragem espaço-temporal. Estas técnicas podem ser muito elaboradas e exibem diferentes graus de sucesso. No entanto, como se nota em [Aloimonos et al., 1988], é irrealista considerar o problema de estimação do fluxo óptico um problema resolvido.

De forma a ultrapassar as dificuldades associadas à estimação do fluxo óptico, muitos investigadores [Horn and Weldon, 1988, Aloimonos and Duric, 1994, Fermuller, 1995a, Fermuller and Aloimonos, 1993, Sinclair et al., 1994] têm usado directamente as derivadas espaço-temporais do brilho (ou alternativamente o fluxo normal) de uma sequência de imagens. O fluxo normal representa a única componente do fluxo óptico que pode ser estimada exclusivamente a partir de medidas locais do brilho, devido ao conhecido *problema da abertura* [Horn and Shunck, 1981, Horn, 1986].

Os métodos que utilizam o fluxo normal, conhecidos por métodos directos, não requerem algoritmos aproximativos ou iterativos para o cálculo das suas observações. Em geral, os métodos directos são abordagens que ignoram a natureza estatística das observações, limitando-se a fazer uso da informação global do fluxo e/ou a aplicar estimadores lineares de mínimos quadrados. Neste trabalho, iremos mostrar como certos desvios inesperados nas observações, podem contaminar as soluções de um estimador de mínimos quadrados — pelo que se exige alguma atenção no que se refere aos processos de

selecção e rejeição de eventuais desvios¹ no espaço das observações.

1.2.2 Algoritmos iterativos vs. não iterativos

Nas abordagens para deslocamento discreto, a maioria dos algoritmos usam a conhecida restrição da linha epipolar [Faugeras, 1993]. Normalmente, os parâmetros de deslocamento do observador são calculados directamente a partir da matriz essencial, que, por sua vez, é determinada de forma linear a partir de alguns pontos das imagens. Como se disse, estes algoritmos são particularmente sensíveis ao ruído de entrada e, para aumentar a robustez do método, alguns autores [Weng et al., 1993] propõem minimizar iterativamente um dado funcional de custo, após a introdução de algum conhecimento quanto a modelos de distribuição probabilística do ruído.

Horn em [Horn, 1986] parte da informação do fluxo óptico e propõe três abordagens usando estimadores do tipo mínimos quadrados. Nesse trabalho são derivadas formas fechadas para a solução da estimação de velocidade, para o caso do movimento ser puramente translacional ou puramente rotacional. Para o caso geral, Horn deriva um conjunto de equações não-lineares que são resolvidas iterativamente.

Heeger e Jepson [Heeger and Jepson, 1992] transformam esta aproximação num algoritmo de procura finita, subdividindo o problema em dois passos: no primeiro passo, o método estima a translação da câmara, e, no segundo passo, determina a rotação. Partindo do conhecimento completo do fluxo óptico, este método é exacto (sem hipóteses simplificativas) e não faz uso de técnicas numéricas de natureza iterativa.

Vários métodos directos (baseados no fluxo normal) têm sido propostos: a maioria deriva directamente dos métodos que usam o fluxo óptico [Horn, 1986, Heeger and Jepson, 1992]. Excluem-se deste grupo outras abordagens, nomeadamente a desenvolvida por Fermuller e Aloimonos [Fermuller and Aloimonos, 1993, Fermuller, 1995a, Fermuller, 1995b], que transformam o problema de estimação a partir do fluxo normal num problema de reconhecimento de padrões. O problema da estimação

¹ *outliers*, na nomenclatura inglesa

de velocidade pode ser visto como um algoritmo não iterativo. Também neste caso, o uso dos dados globais deve garantir a robustez do algoritmo à influência do ruído nas observações.

Concluindo, os algoritmos iterativos são computacionalmente pesados e requerem a escolha prévia de um estado de inicialização. Estão igualmente dependentes de critérios de paragem, que, usalmente, são definidos pela relação entre um nível de erro e um limite de natureza heurística.

Por outro lado, os algoritmos não iterativos são extremamente sensíveis ao ruído. Para ultrapassar este problema, usam-se essencialmente dois tipos de abordagens: combinar as observações globais ao longo do espaço e/ou do tempo; ou usar ferramentas provenientes da estatística, de forma a melhorar os procedimentos de estimação.

1.3 Estrutura da tese

No Capítulo 2, introduzem-se os principais aspectos teóricos do sistema de visão adoptado. Exploramos a relação algébrica e geométrica entre o fluxo óptico e a velocidade do observador, e apresentamos uma primeira abordagem na estimação dos parâmetros de velocidade. Posteriormente, o método é extendido para o caso em que se introduz a restrição do fluxo normal. Analisam-se as principais dificuldades dos métodos existentes, e definem-se os caminhos para uma nova abordagem.

No Capítulo 3, estudam-se algumas propriedades geométricas do fluxo normal, procedendo à selecção das localizações do fluxo que estão intimamente associadas à velocidade do observador. Introduce-se um conjunto de subespaços de procura definidos no plano da imagem que permitem reduzir significativamente a complexidade algorítmica dos métodos de estimação.

O Capítulo 4 é dedicado aos algoritmos de estimação propostos. Define-se um conjunto de algoritmos de procura ao longo de diferentes subespaços. Cada algoritmo de procura pretende estimar uma restrição (ou conjunto de restrições) no espaço das estimativas. Estuda-se a organização geral dos algoritmos propostos, de forma a recuperar completa-

mente os parâmetros de velocidade da câmara. No Capítulo 5, verificamos o desempenho do método, através da apresentação de alguns resultados experimentais, efectuados a partir de imagens sintéticas e reais, e para diversos tipos de movimento translacional e rotacional.

No Capítulo 6, apresenta-se uma extensão do método para uma nova geometria do plano de imagem: a geometria log-polar. Mostra-se que alguns algoritmos de procura propostos para coordenadas cartesianas, têm uma representação natural em coordenadas log-polares.

Finalmente, o Capítulo 7 é dedicado às conclusões retiradas deste trabalho. Estabelecem-se, adicionalmente, as direcções principais para um trabalho futuro.

2. Definição do Problema

Computational vision researchers have been trying to solve the problem for a number of years with only limited success. It is a difficult problem to solve because the optical flow field is nonlinearly related to the 3D motion and depth parameters.

Heeger & Jepson in [Heeger and Jepson, 1992]

Neste capítulo introduzem-se as principais propriedades do fluxo óptico, nomeadamente a sua dependência com o movimento da câmara e com a estrutura da cena. Procuraremos adicionalmente introduzir alguns caminhos para a estimação da velocidade da câmara, tendo por base o conhecimento total ou parcial do fluxo óptico.

Inicialmente apresentaremos dois métodos de estimação, sem impôr restrições quanto ao conhecimento do fluxo óptico.

O passo seguinte consiste em conceber um método no qual se introduz a restrição do fluxo normal. O conhecimento parcial do fluxo óptico constitui uma limitação importante. Como iremos ver neste capítulo, o método daqui resultante está associado a um algoritmo de elevada complexidade, pelo que se torna necessário um esforço de simplificação do método. Essa simplificação só é possível se soubermos algo mais acerca da estrutura do problema, nomeadamente no que se refere às relações geométricas entre o fluxo normal e as características do movimento.

2.1 Fluxo óptico vs. Movimento da câmara

Consideremos um sistema de coordenadas com origem no centro óptico da câmara, como se ilustra na Figura 2.1. Os eixos x e y são paralelos aos eixos da câmara correspondentes.

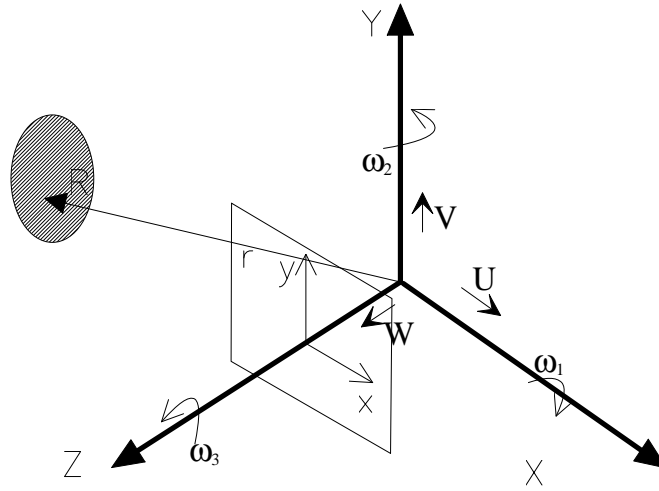


Fig. 2.1: Representação do movimento e da geometria de projecção no plano de imagem.

O eixo z coincide com o eixo óptico. Supõe-se que um ponto genérico \mathbf{R} definido no espaço \mathcal{R}^3 pelas coordenadas $[X \ Y \ Z]^T$, é projectado num ponto da imagem \mathbf{r} , segundo um modelo de projecção perspectiva. Os pontos \mathbf{R} e \mathbf{r} relacionam-se do seguinte modo :

$$\mathbf{r} = [x \ y \ f]^T = f \frac{\mathbf{R}}{\mathbf{R} \cdot \hat{\mathbf{z}}}, \quad (2.1)$$

onde f corresponde à distância focal da câmara, $\hat{\mathbf{z}}$ representa o versor ao longo do eixo z e o operador “ \cdot ” traduz o produto interno entre vectores.

Supondo que a câmara adquire uma velocidade linear $\mathbf{t} = [U \ V \ W]^T$ e uma velocidade angular $\boldsymbol{\omega} = [\omega_1 \ \omega_2 \ \omega_3]^T$, pode-se calcular o movimento do ponto \mathbf{r} induzido no plano de imagem [Longuet-Higgins and Prazdny, 1980]:

$$\frac{d\mathbf{r}}{dt} = \begin{bmatrix} u \\ v \\ 0 \end{bmatrix} = \begin{bmatrix} \frac{-Uf+xW}{Z} + \omega_1 \frac{xy}{f} - \omega_2 \left(\frac{x^2}{f} + f \right) + \omega_3 y \\ \frac{-Vf+yW}{Z} + \omega_1 \left(\frac{y^2}{f} + f \right) - \omega_2 \frac{xy}{f} - \omega_3 x \\ 0 \end{bmatrix}, \quad (2.2)$$

onde u e v representam as derivadas de \mathbf{r} na direcção do eixo x e y respectivamente. As componentes u e v formam um vector no plano de imagem habitualmente denominado fluxo óptico.

Observando as equações, pode-se concluir de imediato que u e v , para além de dependerem das componentes da velocidade linear e angular, dependem explicitamente da distância Z ao ponto \mathbf{R} . Por consequência, o fluxo óptico depende, em cada ponto da imagem, de 7 parâmetros desconhecidos (3 componentes da velocidade linear, 3 componentes da velocidade angular e 1 valor de profundidade Z).

A questão fundamental que se põe é a seguinte: será possível recuperar todos os parâmetros da velocidade linear e angular a partir das equações (2.2)? Esta é uma pergunta central que requer alguma atenção. Primeiro, a resposta está directamente ligada ao número total de **restrições** inerentes ao próprio processo de estimação. Estas restrições podem ser agrupadas em três classes: a restrição física, a restrição do fluxo normal e a restrição visual.

- **Restrição física:** esta restrição provém da própria estrutura física do sistema óptico e deriva das propriedades associadas à projecção perspectiva; mais adiante iremos ver que, devido a esta restrição, é impossível recuperar todas as componentes da velocidade linear a partir do fluxo óptico. Esta limitação advém do simples facto de todas as componentes de translação, na equação de fluxo óptico (2.2), serem escaladas pelo mesmo parâmetro (desconhecido) Z .
- **Restrição do fluxo normal:** este tipo de restrição deriva directamente do conhecido problema da abertura, que se traduz na impossibilidade de conhecer completamente o fluxo óptico em cada ponto da imagem — prova-se que, a partir de medições locais, só é possível calcular a projecção do vector (u, v) na direcção do gradiente da imagem; esta projecção é conhecida por fluxo normal. Neste capítulo, mostra-se que esta importante restrição vai afectar profundamente as técnicas de estimação dos parâmetros de movimento.
- **Restrição visual:** esta restrição está associada ao estudo da capacidade do es-

timador de recuperar os parâmetros de velocidade a partir da informação visual disponível. Neste caso, discute-se o que, na linguagem do controlo, é usual chamar-se a observabilidade do sistema. É fácil verificar que a observabilidade está dependente da textura e da forma das superfícies visualizados. Daremos aqui dois exemplos muito simples que permitem ilustrar este tipo de restrição:

(1) quanto à textura: sabendo que só há observações em pontos (ou zonas) da imagem com textura, então a textura condiciona necessariamente a observabilidade do sistema. Na realidade, a observabilidade devido à textura da imagem é directamente afectada pela restrição do fluxo normal existente em cada observação¹: Por exemplo, supondo que a cena visualizada é composta por uma faixa vertical negra sobre um fundo branco e que a câmara sofre um movimento de translação, então nada se pode concluir quanto ao movimento vertical da câmara — existe claramente um estado do sistema não observável.

(2) quanto à forma: existem ambiguidades na detecção de soluções, que podem estar relacionadas com a forma da superfície observada. Por exemplo, quando essa superfície é planar, prova-se que, para o mesmo fluxo óptico e em determinadas condições, o problema de estimação do movimento tem duas soluções distintas. Uma das soluções corresponde à velocidade exacta da câmara; a outra está relacionada com os parâmetros do próprio plano visualizado (ver Apêndice B).

Em alguns casos, as **restrições visuais** produzem, soluções instáveis e/ou indeterminadas (que sob certas condições, podem ser recuperadas com exactidão, usando técnicas específicas de estimação). Contudo, numa primeira análise do problema, vamos assumir que a informação visual é suficientemente “rica” para produzir soluções exactas. Isto significa que vamos centrar objectivamente a nossa atenção nas duas primeiras restrições.

¹De facto, as restrições visuais quanto à textura tem a sua origem na restrição do fluxo normal em cada observação. É preciso sublinhar, no entanto, que as duas restrições tratam problemas diferentes: a restrição visual aborda o problema da observabilidade do **conjunto** das observações, enquanto a restrição normal foca o problema que relaciona **cada** observação à textura da imagem.

2.2 Caracterização dos parâmetros a estimar

Como vimos, a restrição física equivale à impossibilidade algébrica de recuperar completamente os parâmetros de movimento unicamente a partir do fluxo óptico. Se assim é, podemos saber com exactidão que parâmetros podemos estimar e qual a sua relação com o movimento da câmara.

Considere-se a seguinte mudança de variáveis:

$$\begin{cases} \sigma = fU/W \\ \eta = fV/W \\ \rho = W/Z \end{cases} \quad (2.3)$$

Estas novas variáveis têm um significado físico preciso, que pode ser descrito do seguinte modo:

- a) o vector (σ, η, f) é colinear com o vector da velocidade linear; logo, o ponto (σ, η) corresponde à projecção da velocidade linear sobre o plano da imagem — este ponto especial denomina-se Foco de Expansão²;
- b) o parâmetro ρ corresponde ao inverso do tempo de colisão da câmara com um dado ponto observado no espaço; isto implica que existe um valor ρ para cada ponto da imagem (pelo que ρ é função das coordenadas da imagem (x, y)).

A mudança de variáveis definida em (2.3) é válida para $W \neq 0$ e $Z \neq 0$: A primeira condição significa que a velocidade linear no sentido do eixo óptico é não nula³. A segunda restrição, dada por $Z \neq 0$, é garantida à partida por razões construtivas do próprio sistema óptico.

²A explicação para a denominação “Foco de Expansão” advém de um facto recorrente da sua própria definição: dada uma câmara com um movimento composto por translação pura, observa-se que todos os vectores de fluxo óptico registados na imagem “partem de” ou “dirigem-se para” um único ponto da imagem. Este ponto é definido pelo Foco de Expansão. Na nomenclatura inglesa denomina-se *Focus of Expansion* e é usualmente designado pela sigla FOE.

³O caso $W = 0$ será abordado no Apêndice B, como extensão do caso $W \neq 0$.

Aplicamos então a mudança de variáveis enunciada em (2.3) nas equações (2.2). Obtemos novas equações para o fluxo óptico:

$$\begin{cases} u = \rho(x, y)(x - \sigma) + \omega_1 \frac{xy}{f} - \omega_2 \left(\frac{x^2}{f} + f \right) + \omega_3 y \\ v = \rho(x, y)(y - \eta) + \omega_1 \left(\frac{y^2}{f} + f \right) - \omega_2 \frac{xy}{f} - \omega_3 x \end{cases} \quad (2.4)$$

Verifica-se que, para cada ponto da imagem, o fluxo depende explicitamente, não de 7 mas de 6 parâmetros. Conclui-se ainda que não é possível obter simultaneamente as três componentes da velocidade linear \mathbf{t} , a partir da informação oferecida pelo fluxo óptico (restrição física).

As equações desenvolvidas em (2.4) podem ser escritas de uma forma vectorial, onde se torna mais clara a relação entre os parâmetros envolvidos:

$$\mathbf{v}(\mathbf{x}) = \rho(\mathbf{x})(\mathbf{x} - \boldsymbol{\Sigma}) + \mathbf{B}(\mathbf{x})\boldsymbol{\omega} \quad (2.5)$$

onde $\mathbf{x} = [x \ y]^T$ indica o ponto da imagem, $\mathbf{v}(\mathbf{x}) = [u(\mathbf{x}) \ v(\mathbf{x})]^T$ representa o vector do fluxo óptico, $\boldsymbol{\Sigma} = [\sigma \ \eta]^T$ o Foco de Expansão, e o vector de rotação é multiplicado pela matriz

$$\mathbf{B}(\mathbf{x}) = \begin{bmatrix} \frac{xy}{f} & -\left(\frac{x^2}{f} + f\right) & y \\ \left(\frac{y^2}{f} + f\right) & -\frac{xy}{f} & -x \end{bmatrix} \quad (2.6)$$

Definiu-se um conjunto restrito de parâmetros que caracterizam o movimento da câmara: $\boldsymbol{\omega}$, ρ e $\boldsymbol{\Sigma}$. A questão que se põe agora é como estimar esses parâmetros usando o fluxo óptico de N pontos da imagem.

A resposta seria trivial se o vector do fluxo óptico fosse uma função linear dos parâmetros que queremos estimar. Conclui-se claramente que não é, observando a equação vectorial (2.5): na realidade, o fluxo óptico \mathbf{v} é linear em $\boldsymbol{\omega}$ e bilinear em ρ e $\boldsymbol{\Sigma}$.

Na próxima secção, iremos procurar responder a esta questão através de dois conhecidos métodos de estimação [Faugeras, 1993, Heeger and Jepson, 1992]. O primeiro é aplicado ao conjunto das observações do fluxo óptico e deriva directamente da equação vectorial (2.5); o segundo método, mais elaborado, é independente de $\rho(\mathbf{x})$ e faz uso de algumas propriedades do fluxo.

2.3 Estimação de parâmetros a partir do fluxo óptico

O primeiro método de estimação baseia-se num pressuposto muito simples aplicado a equações bilineares: se fixarmos a variável Σ , a equação (2.5) torna-se linear em relação aos restantes parâmetros ω e ρ . Logo, se juntarmos $2N$ equações lineares (2.5) correspondentes a N pontos de observação da imagem, podemos obter um sistema linear sobredeterminado.

Neste método, propõe-se considerar o parâmetro Σ como uma variável que varia numa dada gama de vectores definidos no plano de imagem. O método baseia-se num simples algoritmo de procura: o valor estimado do Foco de Expansão corresponde ao vector Σ que minimiza um determinado funcional de custo. Mais especificamente, propõe-se estudar a variância da solução do sistema linear associado a cada hipótese para localização do Foco de Expansão, Σ . Espera-se que a solução exacta (o Foco de Expansão) coincida com a solução com menor variância para uma família (infinita!) de valores Σ .

Iremos, para já, formular o primeiro método de estimação.

2.3.1 Primeiro método de estimação

Para N pontos de observação, temos o seguinte sistema de equações:

$$\underbrace{\begin{bmatrix} v_1 \\ \vdots \\ v_N \end{bmatrix}}_{\bar{v}} = \underbrace{\begin{bmatrix} x_1 - \Sigma & \dots & 0 & B_1 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \dots & x_N - \Sigma & B_N \end{bmatrix}}_{\mathbf{A}(\Sigma)} \underbrace{\begin{bmatrix} \rho_1 \\ \vdots \\ \rho_N \\ \omega \end{bmatrix}}_{\mathbf{p}} \quad (2.7)$$

onde \bar{v} é um vector de dimensão $2N$, contendo o fluxo óptico dos N pontos observados; \mathbf{p} é o vector de parâmetros a estimar, de dimensão $N + 3$, e inclui os 3 parâmetros de rotação e os N valores de ρ . Finalmente, $\mathbf{A}(\Sigma)$ é uma matriz de dimensão $2N$ por $(N + 3)$, calculável se o Foco de Expansão, Σ , for conhecido.

Para um dado vector Σ , o sistema de equações (2.7) é linear e tem uma solução sobredeterminada (associada a uma variância) se⁴ $\text{car}(\mathbf{A}(\Sigma)) = N + 3$ e

$$2N > N + 3 \Leftrightarrow N > 3.$$

Isto significa que, para obter uma estimativa dos parâmetros a partir do fluxo óptico (supondo que é conhecido), é necessário dispôr de pelo menos 4 pontos da imagem com informação de fluxo e, adicionalmente, a matriz $\mathbf{A}(\Sigma)$ do sistema deve ter característica igual a $N + 3$. Esta é uma restrição visual importante e depende dela o desempenho do algoritmo de estimação. No entanto, vamos mais uma vez partir do princípio que os pressupostos visuais quanto à observabilidade do sistema estão garantidos.

Assim, para um dado vector Σ , a solução do sistema (2.7) é determinada de forma explícita, se, para tal, recorrermos ao método dos mínimos quadrados

$$\min_{\hat{\mathbf{p}}} \{ \|\bar{\mathbf{v}} - \mathbf{A}(\Sigma)\hat{\mathbf{p}}\|^2 \}$$

Assim, no caso do sistema ser sobredeterminado, a solução é dada por:

$$\hat{\mathbf{p}} = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \bar{\mathbf{v}}$$

Por outro lado, o valor exacto de Σ é aquele que minimiza a variância da estimativa $\hat{\mathbf{p}}$ dada por

$$\|\bar{\mathbf{v}} - \mathbf{A}(\Sigma)\hat{\mathbf{p}}\|^2$$

Resumindo, o método de estimação resulta do seguinte critério de minimização:

$$\min_{\hat{\Sigma}} \left\{ \min_{\hat{\mathbf{p}}} \left\{ \|\bar{\mathbf{v}} - \mathbf{A}(\hat{\Sigma})\hat{\mathbf{p}}\|^2 \right\} \right\} \quad (2.8)$$

Uma das características estruturais que torna este método de estimação instável (no sentido de produzir soluções que dependem fortemente do ruído nas observações), é o elevado número de parâmetros a estimar ($N + 3$). Na realidade, acrescentar um novo ponto de observação ao sistema não beneficia tanto quanto seria desejável a robustez do

⁴ $\text{car}(\mathbf{A})$ corresponde à característica da matriz \mathbf{A} .

processo de estimação face ao ruído nas observações, já que isso implica acrescentar um novo parâmetro desconhecido $\rho(\mathbf{x})$ ao vector \mathbf{p} .

Seria por isso altamente vantajoso se pudéssemos eliminar das equações de fluxo a sua dependência em $\rho(\mathbf{x})$ e, a partir daí, construir um novo sistema linear cujo vector de parâmetros a estimar fosse de dimensão reduzida e fixa para qualquer número N de observações. Verifica-se que isso é possível operando uma simples projecção nos vectores de fluxo óptico, dando origem ao segundo método de estimação.

2.3.2 Segundo método de estimação

Seja $\mathbf{d}(\mathbf{x}, \Sigma)$ um vector unitário perpendicular à direcção definida pelo vector $(\mathbf{x} - \Sigma)$:

$$\mathbf{d}(\mathbf{x}, \Sigma) \cdot (\mathbf{x} - \Sigma) = 0 \quad (2.9)$$

$$\|\mathbf{d}(\mathbf{x}, \Sigma)\| = 1$$

Note-se que, para um dado vector Σ , $\mathbf{d}(\mathbf{x}, \Sigma)$ está bem definido para todos os pontos da imagem. Se fizermos o produto interno de $\mathbf{d}(\mathbf{x}, \Sigma)$ com os termos da equação (2.5), obtemos uma equação independente de $\rho(\mathbf{x})$:

$$\mathbf{d}(\mathbf{x}, \Sigma) \cdot \mathbf{v}(\mathbf{x}) = \mathbf{d}(\mathbf{x}, \Sigma) \cdot [\mathbf{B}(\mathbf{x})\boldsymbol{\omega}] \quad (2.10)$$

Este resultado está directamente associado a uma propriedade do fluxo muito importante que convém reter: **a componente do fluxo óptico perpendicular à linha que une um ponto \mathbf{x} ao Foco de Expansão não depende do movimento linear da câmara.** Essa componente corresponde precisamente à projecção do fluxo sobre o vector $\mathbf{d}(\mathbf{x}, \Sigma)$.

O segundo método de estimação é semelhante ao primeiro e pode ser descrito da seguinte forma:

- Para cada valor Σ , define-se o sistema linear

$$\underbrace{\begin{bmatrix} \mathbf{d}(\mathbf{x}_1, \Sigma) \cdot \mathbf{v}_1 \\ \vdots \\ \mathbf{d}(\mathbf{x}_N, \Sigma) \cdot \mathbf{v}_N \end{bmatrix}}_{\bar{\mathbf{v}}_d(\Sigma)} = \underbrace{\begin{bmatrix} \mathbf{d}(\mathbf{x}_1, \Sigma) \cdot \mathbf{B}_1 \\ \vdots \\ \mathbf{d}(\mathbf{x}_N, \Sigma) \cdot \mathbf{B}_N \end{bmatrix}}_{\mathbf{A}_d(\Sigma)} \boldsymbol{\omega} \quad (2.11)$$

onde $\bar{\mathbf{v}}_{\mathbf{d}}(\boldsymbol{\Sigma})$ é um vector de dimensão N , $\mathbf{A}_{\mathbf{d}}(\boldsymbol{\Sigma})$ é uma matriz de dimensão $N \times 3$ e os parâmetros a estimar são as 3 coordenadas do vector de rotação. O sistema tem solução sobredeterminada, se $\text{car}(\mathbf{A}_{\mathbf{d}}(\boldsymbol{\Sigma}))=3$ e $N > 3$.

- Obtemos uma estimativa dos parâmetros $\boldsymbol{\Sigma}$ e $\boldsymbol{\omega}$, a partir do seguinte critério de minimização:

$$\min_{\hat{\boldsymbol{\Sigma}}} \left\{ \min_{\hat{\boldsymbol{\omega}}} \left\{ \|\bar{\mathbf{v}}_{\mathbf{d}}(\hat{\boldsymbol{\Sigma}}) - \mathbf{A}_{\mathbf{d}}(\hat{\boldsymbol{\Sigma}})\hat{\boldsymbol{\omega}}\|^2 \right\} \right\} \quad (2.12)$$

Recapitulando, apresentámos dois métodos de estimação de parâmetros de velocidade a partir do fluxo óptico. Baseiam-se na procura exaustiva do Foco de Expansão num dado domínio, de forma a minimizar a variância da solução de um sistema linear. O segundo método não estima directamente os parâmetros ρ , ou seja, restringe-se à estimação dos parâmetros de rotação. Pode-se deduzir que o segundo método é numericamente menos complexo e mais imune à existência de ruído nas observações, dado que limita o espaço das estimativas. No entanto, os dois métodos de estimação apresentados partem de dois pressupostos muito fortes que gostaríamos de ultrapassar:

- requer-se uma busca exaustiva de $\boldsymbol{\Sigma}$ ao longo de um domínio necessariamente limitado (por razões algorítmicas); este pressuposto implica pesados custos computacionais;
- exige-se o conhecimento completo do fluxo óptico.

Se, por um lado, o primeiro pressuposto pode ser ultrapassado com alguma manipulação algorítmica, o segundo pressuposto oferece maior dificuldade: como já tínhamos enunciado anteriormente, o fluxo óptico não é completamente conhecido devido à restrição do fluxo normal. Isto leva-nos a um novo problema: obter um método de estimação dos parâmetros de velocidade que não exija o conhecimento completo do fluxo óptico. Para responder a esta questão, precisamos definir com precisão em que consiste a restrição do fluxo normal.

2.4 A restrição do fluxo normal

Assume-se que a imagem é composta por superfícies texturadas e que as condições de iluminação variam lentamente ao longo do tempo.

Supondo que o brilho de um ponto da superfície não varia ao longo do tempo — hipótese do brilho constante [Horn and Shunck, 1981, Horn, 1986] — tem-se:

$$\begin{aligned} \frac{dE(x, y)}{dt} &= 0 \\ \Leftrightarrow E_t + uE_x + vE_y &= 0, \end{aligned} \quad (2.13)$$

onde $E(x, y)$, E_x , E_y e E_t correspondem respectivamente ao nível de brilho e às suas derivadas de primeira ordem relativamente a x , y e t . Vamos partir do princípio que a hipótese é verdadeira em quase todos os pontos da imagem⁵. Assim sendo, a equação (2.13) expressa aquilo que na realidade é possível conhecer do fluxo óptico: uma combinação linear das suas duas componentes vectoriais. Esta restrição no conhecimento do fluxo óptico é conhecida por problema da abertura [Horn, 1986]. De facto, partindo da hipótese do brilho constante, só é possível calcular a componente do fluxo óptico, \mathbf{v} , ao longo da direcção do gradiente da imagem. Essa componente é dada por:

$$v_n = \mathbf{n} \cdot \mathbf{v} = -\frac{E_t}{\|\nabla E\|} \quad (2.14)$$

onde ∇E é o gradiente espacial da imagem, \mathbf{n} é um vector unitário colinear com a direcção do gradiente. A projecção $\mathbf{n} \cdot \mathbf{v}$ é conhecida por *fluxo normal*, v_n .

Usando a relação (2.5) entre o fluxo óptico e os parâmetros de movimento da câmara, obtemos a seguinte equação definida em cada ponto da imagem:

$$v_n(\mathbf{x}) = \rho(\mathbf{x})\mathbf{n}^t(\mathbf{x})(\mathbf{x} - \boldsymbol{\Sigma}) + \mathbf{n}^t(\mathbf{x})\mathbf{B}(\mathbf{x})\boldsymbol{\omega} \quad (2.15)$$

Sabemos que em cada pixel existem teoricamente duas equações de fluxo óptico, mas na realidade dispomos de uma única equação, devido à restrição associada ao fluxo normal.

⁵Remetemos a discussão sobre as condições de validade da hipótese do brilho constante para o Apêndice B.

Este facto inviabiliza de imediato a utilização do primeiro método de estimação colocado em (2.7), dado que em N pontos dispomos de N equações (em vez de $2N$), logo o sistema linear é subdeterminado em quaisquer condições.

Um caminho possível pode ser a segunda abordagem baseada na minimização (2.12), que coloca menores restrições quanto ao número de equações associadas ao sistema linear (2.11). Começemos por analisar a relação entre o segundo método de estimação e o problema de estimação levantado pela restrição do fluxo normal:

- só é necessária uma equação por pixel;
- cada observação corresponde a uma projecção do fluxo óptico numa direcção normalizada.

Recorde-se que no método de estimação (2.12) a direcção normalizada é determinada pelo vector $\mathbf{d}(\mathbf{x}, \Sigma)$, enquanto, segundo a restrição do fluxo normal, essa direcção é determinada pelo vector \mathbf{n} . Esta analogia permite-nos derivar um novo método de estimação por adaptação do segundo método de estimação.

A ideia consiste em *seleccionar unicamente o fluxo normal que é colinear com a direcção $\mathbf{d}(\mathbf{x}, \Sigma)$, para cada vector Σ* — em vez de usar todas as observações de fluxo. Define-se assim o conjunto $\mathcal{U}(\Sigma)$ como o conjunto de pontos da imagem onde o fluxo normal é colinear com o vector $\mathbf{d}(\mathbf{x}, \Sigma)$, ou, equivalentemente, onde o fluxo normal é perpendicular à linha que passa pelo ponto da imagem e pelo Foco de Expansão. A Figura 2.2 ilustra um conjunto $\mathcal{U}(\Sigma)$ para um dado Σ .

O método vem descrito nos seguintes passos:

- Para cada hipótese de localização do Foco de Expansão Σ , constrói-se o seguinte funcional de custo:

$$\mathbf{E}(\Sigma, \boldsymbol{\omega}) = \sum_{\mathbf{x} \in \mathcal{U}(\Sigma)} \left[v_n(\mathbf{x}) - \mathbf{n}^t(\mathbf{x})\mathbf{B}(\mathbf{x})\boldsymbol{\omega} \right]^2, \quad (2.16)$$

onde $\mathcal{U}(\Sigma) = \{\forall \mathbf{x} : \mathbf{n}(\mathbf{x}) \cdot (\mathbf{x} - \Sigma) = 0\}$.

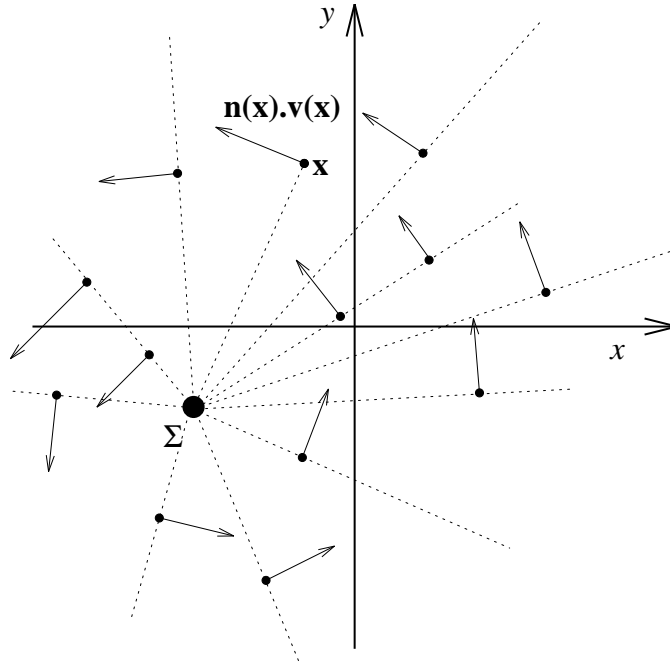


Fig. 2.2: $\mathcal{U}(\Sigma)$ para uma dada localização do Foco de Expansão.

- Para cada Σ , a minimização de $\mathbf{E}(\Sigma, \omega)$ em função de ω é um problema de mínimos quadrados, com solução explícita.
- O método de estimação da posição do Foco de Expansão consiste em procurar o mínimo da função $\mathbf{E}(\Sigma, \omega)$, para uma gama de hipóteses de Σ :

$$\min_{\hat{\Sigma}} \left\{ \min_{\hat{\omega}} \{ \mathbf{E}(\hat{\Sigma}, \hat{\omega}) \} \right\} \quad (2.17)$$

Note-se que o método com restrição do fluxo normal requer uma selecção prévia de observações para cada vector Σ , enquanto o segundo método de estimação conjuga em qualquer caso todas as observações disponíveis. Sendo assim, o método com restrição do fluxo normal é necessariamente mais limitado, dado que, para cada vector Σ , a solução que minimiza $\mathbf{E}(\Sigma, \omega)$ conta com menos observações.

2.5 Discussão

Neste capítulo estudámos a estimação dos parâmetros de movimento a partir da informação de fluxo na imagem. Constatámos a existência de 3 restrições associadas à estimação: a restrição física, a restrição do fluxo normal e a restrição visual. A partir dos métodos de estimação convencionais para o fluxo óptico, formulámos um método de estimação fazendo uso exclusivo do fluxo normal. Emergem deste método duas ideias essenciais que requerem discussão: robustez e complexidade.

- Será o método robusto face às incertezas nas observações? Supondo que existe informação suficiente nas observações de cada conjunto $\mathbf{U}(\Sigma)$, será que essa informação (necessariamente mais limitada, como vimos) é suficiente para produzir uma “boa” estimativa de ω ? O problema é que, ao diminuir-se o número de observações disponíveis em cada $\mathbf{U}(\Sigma)$, deve-se aumentar a qualidade das observações, não no sentido de tornar cada observação mais fiável (menos ruidosa), mas no sentido de aproveitar “bem” as observações mais fiáveis. Por isso, uma peça fundamental no método, e que propomos estudar no Apêndice A, é a técnica de estimação: conclui-se aí que a técnica dos “mínimos quadrados” aplicada à estimação de ω em (2.16) não é a técnica de estimação mais adequada; este facto, leva-nos ao estudo de técnicas de estimação robusta [Rousseeuw and Leroy, 1987], que servem muito melhor os nossos objectivos. A discussão deste tema é remetido para o Apêndice A, dedicado às técnicas de estimação.
- É a complexidade do algoritmo um entrave à estimação dos parâmetros desejados? Na prática, procurar uma solução em toda a gama de vectores Σ (que teoricamente é infinita) é um processo extremamente exigente (ou, no limite, irrealista) em termos computacionais. No entanto, muitas das operações efectuadas neste processo são redundantes, ou então, não contribuem decididamente para a estabilidade da solução. Por conseguinte, talvez seja uma boa opção restringir criteriosamente o espaço das observações e (se necessário) das estimativas, o que poderá implicar uma diminuição substancial na complexidade algorítmica do método.

Assim, iremos dedicar o próximo capítulo a interpretar o espaço das observações e das estimativas, com o intuito de construir um método de baixa complexidade que produza soluções teoricamente exactas⁶, a partir de um conjunto restrito de observações.

⁶ou seja, o método que se pretende desenhar não vai partir de mais pressupostos aproximativos para além dos que foram anunciados.

3. Subespaços de Procura

Assuming that we live in a finite world and that we have a finite number of possibilities for performing computations, any vision problem might be formulated as a simple search problem in a very high dimensional space.

Fermuller & Aloimonos in [Fermuller and Aloimonos, 1995]

No capítulo anterior, debruçámo-nos sobre o problema da estimação dos parâmetros de movimento a partir da informação de fluxo na imagem. Dada a complexidade algorítmica dos métodos associados, sentiu-se a necessidade de subdividir o espaço das observações e das estimativas em subespaços mais simples, e de criar um método de estimação *gradual*, formado por um conjunto de algoritmos que exploram diferentes subespaços de procura.

Neste capítulo, iremos derivar alguns subespaços de procura com descrições geométricas muito simples e facilmente parametrizáveis em coordenadas cartesianas. De forma a facilitar a aplicação dos algoritmos de estimação, observaremos ainda que a concepção dos respectivos subespaços pode ser feita em função das propriedades geométricas do fluxo normal. Finalmente, estaremos na posse das ferramentas necessárias para definir um conjunto de algoritmos de procura de baixa complexidade e com eles construir um método geral de estimação.

3.1 Definição de subespaços de procura

Recorde-se a definição do conjunto $\mathbf{U}(\Sigma) = \{\forall \mathbf{x} : \mathbf{n}(\mathbf{x}) \cdot (\mathbf{x} - \Sigma) = 0\}$. Começemos por estabelecer duas propriedades dos conjuntos $\mathbf{U}(\Sigma)$ que serão essenciais para a definição

de um algoritmo de estimação mais simples:

- para dois valores de Σ , tais que $\Sigma_1 \neq \Sigma_2$, os conjuntos $\mathcal{U}(\Sigma_1)$ e $\mathcal{U}(\Sigma_2)$ têm pontos comuns na recta que passa por Σ_1 e Σ_2 (basta observar a definição de conjunto $\mathcal{U}(\Sigma)$);
- todos os subconjuntos de $\mathcal{U}(\Sigma)$ produzem idealmente, para o mesmo estimador de ω , a mesma estimativa $\hat{\omega}^1$.

Dadas as propriedades acima indicadas, estamos prontos a seleccionar observações através duma geração criteriosa de subconjuntos de $\mathcal{U}(\Sigma)$.

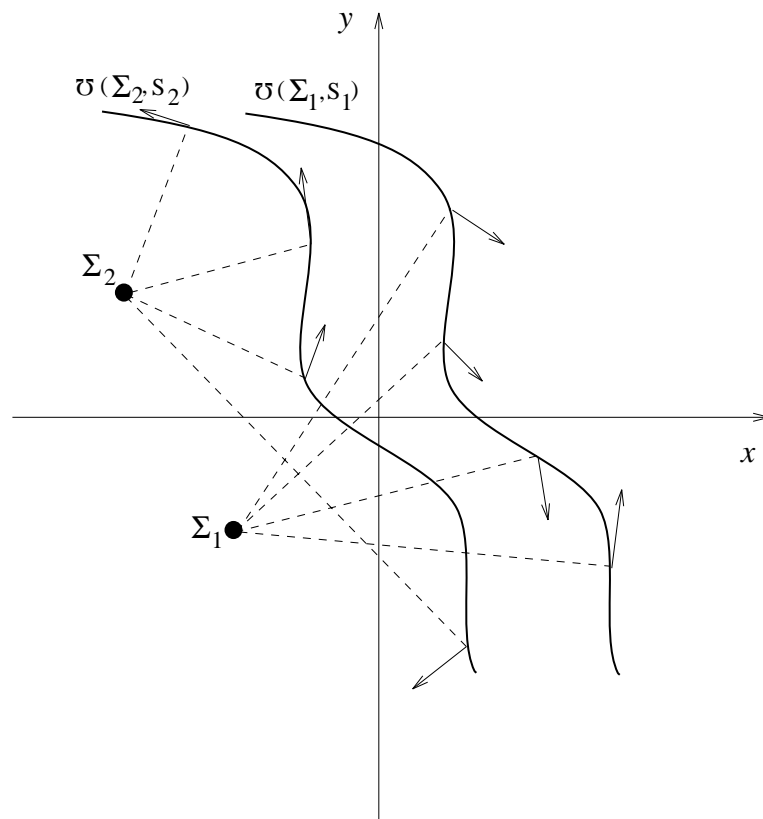


Fig. 3.1: Exemplo de subespaços de procura.

¹Válido para os subconjuntos de $\mathcal{U}(\Sigma)$ com solução sobredeterminada, e áparte um termo devido às incertezas e ao ruído nas observações.

O maior subconjunto de $\mathbf{U}(\Sigma)$ corresponde aos pontos de fluxo localizados em toda a imagem. O subespaço de suporte é o próprio plano da imagem.

Consideremos agora um subespaço mais reduzido, por exemplo uma curva arbitrária definida no plano da imagem. Note-se que o subconjunto de pontos de $\mathbf{U}(\Sigma)$ localizados nesse subespaço correspondem às localizações nesse subespaço dos vectores do fluxo normal perpendiculares às rectas que radiam do Foco de Expansão.

A Figura 3.1 permite exemplificar esta situação, apresentando dois subespaços distintos para dois candidatos a Foco de Expansão. Neste exemplo, definimos $\mathbf{U}(\Sigma_1, \mathbf{S}_1)$ como o subconjunto correspondente a Σ_1 no subespaço \mathbf{S}_1 e $\mathbf{U}(\Sigma_2, \mathbf{S}_2)$ o subconjunto correspondente a Σ_2 no subespaço \mathbf{S}_2 . Usando os vectores do fluxo normal de $\mathbf{U}(\Sigma_1, \mathbf{S}_1)$ podemos estimar os parâmetros de rotação e calcular a variância associada a essa estimativa. Em seguida, repetimos o processo para o subconjunto $\mathbf{U}(\Sigma_2, \mathbf{S}_2)$.

O processo de procura pode continuar para outras localizações de Σ_i e outras curvas \mathbf{S}_i , para os quais se definem os correspondentes subconjuntos $\mathbf{U}(\Sigma_i, \mathbf{S}_i)$. Quando Σ_i corresponder à verdadeira localização do Foco de Expansão, então a variância da estimativa calculada para cada subconjunto será mínima. Assim, para cada candidato, podemos escolher um subespaço arbitrário do plano de imagem, ao longo do qual recolhemos as nossas observações. Isto reduzirá significativamente a complexidade do nosso algoritmo de procura.

Existe uma ressalva a fazer quanto à arbitrariedade na escolha dos subespaços. De facto, existem alguns subespaços que não nos permitem tirar conclusões precisas quanto à localização do Foco de Expansão. Voltemos ao exemplo anterior: suponha que escolhemos outro subespaço de selecção; em vez de optarmos pelas duas curvas desenhadas na Figura 3.1, escolhemos uma recta que intersecte Σ_1 e Σ_2 ; se atendermos à primeira propriedade dos conjuntos $\mathbf{U}(\Sigma)$ apresentada no princípio desta secção, concluiremos que tanto Σ_1 como Σ_2 (tal como todos os candidatos Σ contidos nessa recta) produzem a mesma variância na estimativa de ω^2 . Assim, se o subespaço consistir numa linha recta \mathbf{S} contendo Σ , então, a partir da informação relativa à variância das estimativas para o

²dado que a estimação é feita com as mesmas observações.

conjunto das observações nos pontos $\mathcal{U}(\Sigma, \mathbf{S})$, só podemos tirar conclusões sobre a existência do Foco de Expansão nessa recta (não sendo possível precisar a sua localização exacta sobre a recta).

Contudo, encontrar uma recta que contenha o Foco de Expansão corresponde a uma informação útil (apesar de parcial) da estimativa $\hat{\Sigma}$, na medida que, com essa informação, podemos reduzir o espaço de procura de algoritmos subsequentes. Esta propriedade dos subespaços definidos por rectas levar-nos-á a uma redução importante na complexidade algorítmica do método de estimação, nomeadamente no que se refere à separação do espaço das estimativas.

Nesta secção apresentámos um caminho na divisão do espaço das observações e das estimativas, de forma a simplificar o método de estimação dos parâmetros de velocidade. Ficam aqui algumas notas relevantes quanto ao desenho e concepção de subespaços de procura:

- A família de curvas associadas aos subespaços \mathbf{S} pode ser parametrizada, tornando o processo algorítmico mais eficiente em termos computacionais.
- Os subespaços podem consistir em curvas abertas, rectas, curvas fechadas, ou regiões do plano de imagem.
- O fluxo normal de interesse é seleccionado durante o próprio processo de procura, isto é, para cada par {subespaço \mathbf{S} , Σ }, construímos o subconjunto $\mathcal{U}(\Sigma, \mathbf{S})$ correspondente. Contudo, a possibilidade de detectar as observações de interesse *a priori* conduz à redução da complexidade do problema. Nesse sentido, seria vantajoso conjugar as propriedades do fluxo normal com a forma dos subespaços gerados.

Na próxima secção, apresentam-se alguns critérios para a concepção de subespaços, a partir de características geométricas do fluxo normal.

3.2 Propriedades geométricas do fluxo normal

Nesta secção, analisaremos propriedades geométricas de dois subconjuntos do fluxo normal:

- *o fluxo normal radial*, definido como o fluxo normal colinear com as rectas que radiam da origem do plano de imagem;
- *o fluxo normal circular*, definido como o fluxo normal perpendicular às rectas que radiam da origem do plano de imagem.

É claro que existem outros tipos de fluxo, tantos quantas as definições que queiramos formular. Em [Fermuller, 1995a], pode-se encontrar um estudo alargado sobre as propriedades do fluxo normal definido para as rectas que radiam de um ponto genérico da imagem. Contudo, existem duas boas razões para considerar os dois tipos de fluxo normal anunciados acima:

- os subespaços subjacentes são muito mais simples de parametrizar (como iremos ver nesta secção);
- as propriedades geométricas do fluxo normal radial e circular permitem desacoplar facilmente a estimação dos parâmetros de velocidade da câmara (o que implica que o cálculo desses parâmetros pode ser feito separadamente).

3.2.1 Fluxo normal radial

O *fluxo normal radial* é o conjunto dos vectores de fluxo normal com direcção radial, \mathbf{n}_r . Consideremos o subconjunto dos vectores do fluxo normal radial tal que:

$$\mathbf{n}_r \cdot (\mathbf{x} - \Sigma) = 0 \quad (3.1)$$

Observando a equação (2.15), conclui-se que, para este subconjunto, o fluxo normal não depende da componente translacional do movimento da câmara. É possível determinar

os pontos onde se localizam os vectores do fluxo normal radial que obedecem à restrição (3.1), reescrevendo a equação (3.1) em coordenadas polares ($x = r \cos \psi$, $y = r \sin \psi$) :

$$\begin{aligned} (\cos \psi, \sin \psi) \cdot (r \cos \psi - \sigma, r \sin \psi - \eta) &= 0 \\ \Leftrightarrow r &= (\sigma, \eta) \cdot (\cos \psi, \sin \psi) \end{aligned} \quad (3.2)$$

Recorde-se que $\Sigma = (\sigma, \eta)$. Assim, os pontos desejados descrevem uma circunferência, denominada **circunferência- Γ** , definida exactamente pela origem e pelo Foco de Expansão, Σ , localizados em pontos diametralmente opostos da circunferência, como se ilustra na Figura 3.2a. **A circunferência- Γ corresponde à localização geométrica onde o fluxo normal radial não tem componente translacional (só depende da rotação)**. A componente rotacional do fluxo normal radial é dada por v_r , em que

$$v_r = \mathbf{n}_r \cdot \mathbf{B}(\mathbf{x})\boldsymbol{\omega} \quad (3.3)$$

pode ser representada com as coordenadas polares do ponto \mathbf{x} :

$$v_r = \left(\frac{r^2}{f} + f \right) (\omega_1, \omega_2) \cdot (\sin \psi, -\cos \psi) \quad (3.4)$$

A componente rotacional do fluxo normal radial tem um comportamento sinusoidal em ψ para um raio fixo, r , e anula-se numa recta (Figura 3.2a) que passa pela origem e pelo ponto definido por $\left(f \frac{\omega_1}{\omega_3}, f \frac{\omega_2}{\omega_3} \right)$ (denominado Foco de Rotação por analogia ao Foco de Expansão, *Focus of Rotation* na nomenclatura inglesa — habitualmente designado pela sigla FOR).

3.2.2 Fluxo normal circular

Podemos analisar da mesma forma o *fluxo normal circular*, definido como o conjunto de vectores de fluxo normal com direcção \mathbf{n}_c , perpendicular às rectas radiais. Consideremos o subconjunto de *fluxo normal circular* que verifica

$$\mathbf{n}_c \cdot (\mathbf{x} - \Sigma) = 0 \quad (3.5)$$

O fluxo normal circular pertencente a este subconjunto é perpendicular às rectas que passam pelo Foco de Expansão, ou seja, não depende da translação da câmara (ver Figura 3.2b). Usando novamente coordenadas polares, temos:

$$(\sigma, \eta) \cdot (\sin \psi, -\cos \psi) = 0 \quad (3.6)$$

Isto significa que o fluxo normal circular que não depende do movimento de translação está sobre uma determinada recta, a **recta- Ψ** , definida pelo Foco de Expansão e pela origem da imagem. A componente rotacional do fluxo normal circular é dada por:

$$v_r = f(\omega_1, \omega_2) \cdot (\cos \psi, \sin \psi) - \omega_3 r \quad (3.7)$$

A componente, v_r , é afim em r ao longo de todas as rectas contendo a origem (Figura 3.2b). Estas rectas definem subespaços cuja componente de rotação do fluxo normal circular é afim. Em particular, **o subespaço correspondente à recta- Ψ define um subespaço onde se anula a componente de translação**. Considerando $\omega_3 \neq 0$, v_r anula-se quando

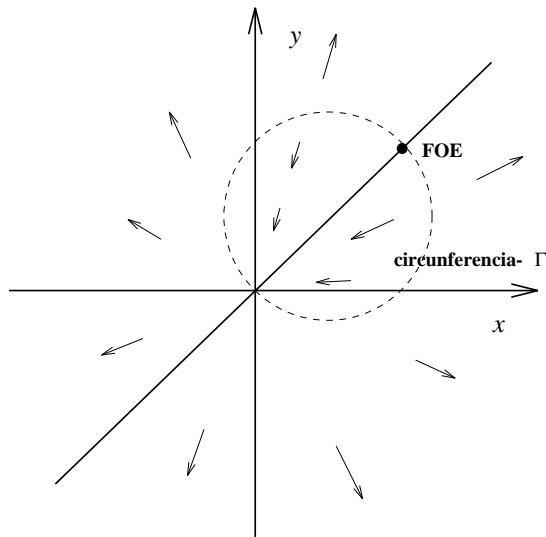
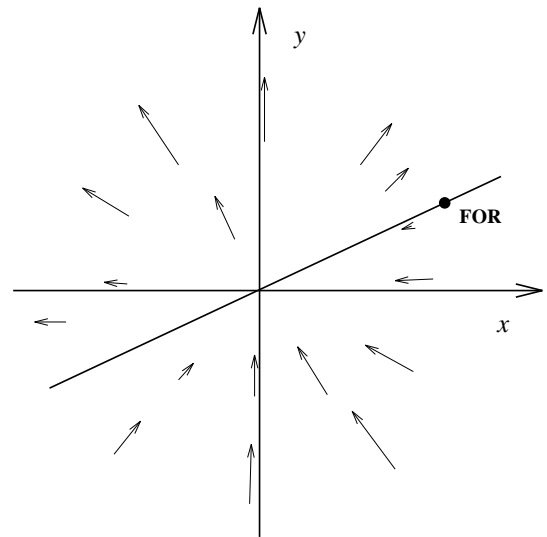
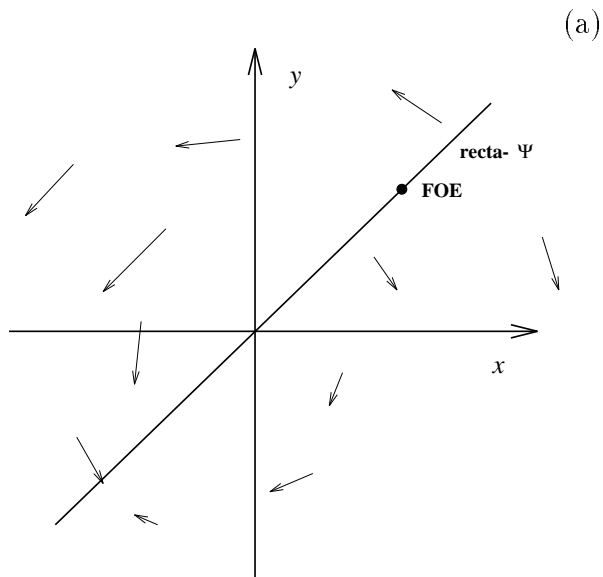
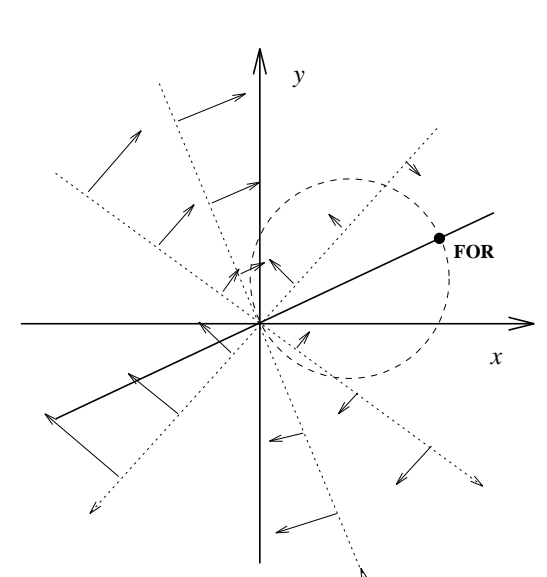
$$r = (\xi_1, \xi_2) \cdot (\cos \psi, \sin \psi), \quad (3.8)$$

onde (ξ_1, ξ_2) são as coordenadas do Foco de Rotação. Logo, a componente rotacional do fluxo normal circular anula-se numa circunferência cujos Foco de Rotação e a origem da imagem estão em pontos diametralmente opostos.

Comparando a equação (3.2) com (3.8) e a equação (3.4) com (3.6), concluímos que as propriedades geométricas encontradas para o fluxo normal circular e radial são duais, como ilustra a Figura 3.2. Resumindo, a componente de translação do fluxo normal radial / circular anulam-se nos subespaços correspondentes à circunferência- Γ e à recta- Ψ , respectivamente. Por outro lado, a componente de rotação é sinusoidal em ψ (para um r fixo) ou afim em r (para um ψ fixo), respectivamente no fluxo normal radial e circular.

3.3 Parametrização dos subespaços de procura

Na secção anterior, definimos alguns subespaços de interesse, que tinham em conta determinadas propriedades do fluxo normal. Nesta secção, vamos estabelecer uma metodologia

Componente de translação do *fluxo normal radial*Componente de rotação do *fluxo normal radial*Componente de translação do *fluxo normal circular*Componente de rotação do *fluxo normal circular*

(b)

Fig. 3.2: Propriedades geométricas do fluxo normal (a) radial e (b) circular.

para os parametrizar no espaço cartesiano.

Seja

$$F(\mathbf{x}; k) = 0 \quad (3.9)$$

a equação duma variedade no espaço cartesiano parametrizada pelo parâmetro k . Apresentaremos de seguida o exemplo de 3 famílias de subespaços e respectiva parametrização:

- **Rectas paralelas** (Figura 3.3a): $F_0(\mathbf{x}; k) = A_0x + B_0y + k$, onde $\|(A_0, B_0)\| = 1$.
- **Rectas radiais** (Figura 3.3b): $F_r(\mathbf{x}; k) = x \sin k + y \cos k$
- **Circunferências tangenciais** (Figura 3.3c): $F_c(\mathbf{x}; k) = x^2 + y^2 - k(x \cos \Psi + y \sin \Psi)$.

Nota: a função F_c depende duma restrição em Σ , definida por: $(\sigma, \eta) \cdot (-\sin \Psi, \cos \Psi) = 0$. Como se observa na figura, $F_c = 0$ define um conjunto de circunferências tangenciais no centro da imagem, cujo diâmetro de comprimento k é medido sobre a recta- Ψ definida na Secção 3.2.2. A circunferência que contém o Foco de Expansão corresponde à circunferência- Γ .

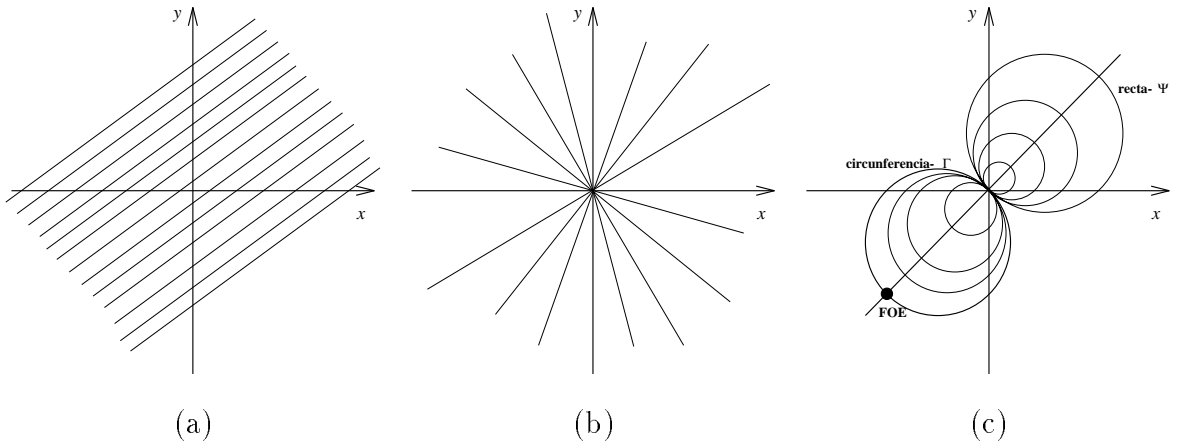


Fig. 3.3: a) Rectas paralelas descritas por $F_0(x; k)$. b) Rectas radiais descritas por $F_r(x; k)$. c) Circunferências tangenciais descritas por $F_c(x; k)$.

Estas três famílias de subespaços estão bem definidas e são facilmente parametrizáveis. Por um lado, a família das rectas paralelas corresponde à parametrização mais simples

de uma família de subespaços em coordenadas cartesianas. Por outro lado, as famílias das rectas radiais e das circunferências tangenciais têm uma construção mais complexa, claramente motivada pelas propriedades geométricas extraídas do fluxo normal radial e circular³.

3.4 Síntese

Neste capítulo, derivámos dois subespaços a partir de dois tipos de fluxo normal: o fluxo normal radial e circular. Por um lado, estes subespaços têm uma descrição geométrica muito simples, que podem ser facilmente parametrizáveis. Por outro lado, o fluxo normal radial e circular são detectáveis *a priori* de forma trivial no conjunto do fluxo normal da imagem.

Neste momento, parece sugerida uma via no sentido da simplificação do método de estimação com restrição do fluxo normal:

- subdivide-se o espaço das observações em subespaços mais simples;
- conseqüentemente, para cada par $\{\text{subespaço}, \Sigma\}$, calcula-se uma estimativa da rotação ω .

No entanto, ainda não é claro como se deve proceder a procura do Foco de Expansão, de forma a minimizar a variância da estimativa de ω . Procuraremos uma resposta para esta questão no próximo capítulo, desenvolvendo um conjunto paradigmático de algoritmos de procura, baseados nos subespaços de procura das rectas paralelas, das rectas radiais e das circunferências tangenciais.

³Sugere-se a comparação entre as Figuras 3.3b, 3.3c e a Figura 3.2.

4. Estimação dos Parâmetros de Velocidade da Câmara

Em geral, os métodos de estimação apresentados no Capítulo 2 envolvem algoritmos de procura computacionalmente pesados. No sentido de reduzir a complexidade desses algoritmos, subdividimos o problema usando subespaços de procura. Neste capítulo, apresentamos exemplos de estimadores baseados nas propriedades geométricas do fluxo normal. Inicialmente, abordam-se alguns aspectos quanto às funções de medida dos resíduos. Em seguida, os subespaços de procura, apresentados no capítulo anterior, levar-nos-ão à concepção de um conjunto de algoritmos de procura de baixa complexidade. Finalmente, teremos oportunidade de apresentar algumas estratégias de estimação quer sequenciais quer paralelas.

4.1 Função de medida dos resíduos

À operação de estimação está inevitavelmente associada a minimização da medida de um dado conjunto de resíduos. Define-se resíduo $\mathbf{R}(\mathbf{x}_i)$ no ponto \mathbf{x}_i como a diferença entre o resultado duma observação e a respectiva estimativa (neste caso específico: entre o fluxo normal observado e o fluxo calculado para as estimativas dos parâmetros de movimento). A medida do conjunto de resíduos pode ser traduzida pela expressão $\mathcal{M}_{\mathbf{L}}(\mathbf{R}(\mathbf{x}_i)|\mathbf{U})$, que significa a medida dos resíduos, calculados nos pontos \mathbf{x}_i ao longo do domínio \mathbf{U} . A medida, \mathcal{M} , está associada a uma métrica genérica, \mathbf{L} , que pode consistir na média dos quadrados, na média pesada dos quadrados, na mediana dos quadrados, ou outra.

Apresentamos aqui um exemplo de aplicação desta notação no método de estimação com restrição do fluxo normal.

A função definida em (2.16) é uma medida dos resíduos $\mathbf{R}(\mathbf{x}_i) = v_n(\mathbf{x}_i) - \mathbf{n}^t(\mathbf{x}_i)\mathbf{B}(\mathbf{x}_i)\boldsymbol{\omega}$ para o conjunto dos pontos $\mathbf{x}_i \in \mathbf{U}(\boldsymbol{\Sigma})$. Recorde-se que $v_n(\mathbf{x}_i)$ corresponde à observação da magnitude do vector de fluxo normal no ponto \mathbf{x}_i . Assim, a estimativa $\hat{\boldsymbol{\omega}}$ é calculada a partir do seguinte estimador:

$$\hat{\boldsymbol{\omega}} = \arg \left\{ \min_{\boldsymbol{\omega}} \{ \mathcal{M}_{\mathbf{S}\mathbf{Q}}(\mathbf{R}(\mathbf{x}_i) | \mathbf{U}(\boldsymbol{\Sigma})) \} \right\} \quad (4.1)$$

onde a métrica $\mathbf{S}\mathbf{Q}$ corresponde ao somatório do quadrado dos resíduos $\mathbf{R}(\mathbf{x}_i)$. Este estimador é naturalmente equivalente ao estimador dos mínimos quadrados aplicado aos resíduos calculados para o conjunto $\mathbf{U}(\boldsymbol{\Sigma})$.

A definição de uma métrica para a medida dos resíduos é fundamental na estimação de parâmetros, e depende da interpretação da distribuição estatística das observações. Neste trabalho, usamos uma métrica robusta para a função de medida \mathcal{M} , em grande parte inspirada no trabalho de Rousseeuw [Rousseeuw and Leroy, 1987]. O estimador robusto que propomos nesta tese baseia-se num estimador não iterativo para regressão simples e inclui mecanismos de detecção e rejeição de *outliers*. Apresenta-se no Apêndice A um estudo alargado sobre as ferramentas de estimação utilizadas.

4.2 Algoritmos de Procura

Nesta secção, propomos estudar um conjunto restrito de algoritmos de procura de complexidade reduzida, fazendo uso dos subespaços estudados no capítulo anterior. Cada algoritmo explora um dado subespaço, para um subconjunto específico de observações, e estima uma determinada restrição (ou restrições) definida no espaço das estimativas.

4.2.1 O algoritmo de procura da recta- Ψ

Muitas das expressões usadas neste capítulo são similares às expressões definidas para o fluxo normal radial e circular. A sua reformulação serve essencialmente os propósitos

algorítmicos, pelo que não repetiremos as propriedades que lhes estão associadas.

Recordemos então a noção de *fluxo normal circular* descrita na Secção 3.2.2, definido para as rectas radiais com orientação ψ :

$$\begin{aligned} \forall(x, y) & : (x, y) = r(\cos \psi, \sin \psi), \\ U_\psi(r) & = \mathbf{n}_c \cdot \mathbf{v} = (u, v) \cdot (-\sin \psi, \cos \psi), \end{aligned} \quad (4.2)$$

onde r e ψ representam as coordenadas polares de (x, y) . Usando as expressões que relacionam o fluxo óptico com o movimento, temos:

$$\begin{aligned} \forall(x, y) & : (x, y) = r(\cos \psi, \sin \psi) \\ U_\psi(r) & = -\rho(x, y)(\sigma, \eta) \cdot (-\sin \psi, \cos \psi) + f(\omega_1 \cos \psi + \omega_2 \sin \psi) - \omega_3 r \end{aligned} \quad (4.3)$$

Na Secção 3.2.2 introduzimos a *recta- Ψ* , definida pela origem da imagem e pelo Foco de Expansão, fazendo um ângulo Ψ com o eixo horizontal. O fluxo normal circular, U_ψ , é composto por um termo rotacional afim em r , e um termo translacional que, geralmente, é não-linear em r^1 . Contudo, esta componente não-linear anula-se quando ψ é o ângulo associado à *recta- Ψ* (i.e. $\psi = \Psi$), ou quando o Foco de Expansão está localizado no centro da imagem.

Seja U_Ψ , a função U_ψ definida em $\psi = \Psi$, afim em r , e dada por:

$$\begin{aligned} \forall(x, y) & : (x, y) = r(\cos \Psi, \sin \Psi), \\ U_\Psi(r) & = f(\omega_1 \cos \Psi + \omega_2 \sin \Psi) - \omega_3 r. \end{aligned} \quad (4.4)$$

O algoritmo proposto consiste basicamente em procurar uma recta radial ψ^k , tal que

¹Pode-se provar que existem algumas superfícies para as quais este termo é afim em r . Ver Apêndice B.

$U_{\psi^k}(r)$ seja afim em r , ou seja, detectar a recta radial que minimiza² \mathcal{M} :

$$\min_{\psi^k} \left\{ \min_{p, \omega_3} \left\{ \mathcal{M} \left(v_n - (p - \omega_3 r) | \psi^k \right) \right\} \right\} \quad (4.5)$$

onde $p = f(\omega_1 \cos \psi + \omega_2 \sin \psi)$.

A tarefa de procura é feita ao longo da família de rectas radiais definidas por $F_r(\mathbf{x}; k) = x \sin k + y \cos k = 0$. Para cada recta ψ^k , são necessárias pelo menos 3 observações para obter uma estimativa de p e ω_3 .

A Figura 4.1 ilustra o algoritmo apresentado, para uma gama de rectas radiais $\psi \in [-\pi/2, \pi/2[$. A variância mínima do resíduo dá a indicação da direcção exacta do Foco de Expansão. A Figura 4.1 mostra ainda a *recta- Ψ* , as observações usadas pelo algoritmo (os vectores de fluxo normal perpendiculares à recta), e o comportamento afim da função $U_{\Psi}(r)$, medida sobre a *recta- Ψ* . Assim que a *recta- Ψ* é detectada, ω_3 é dado pelo declive

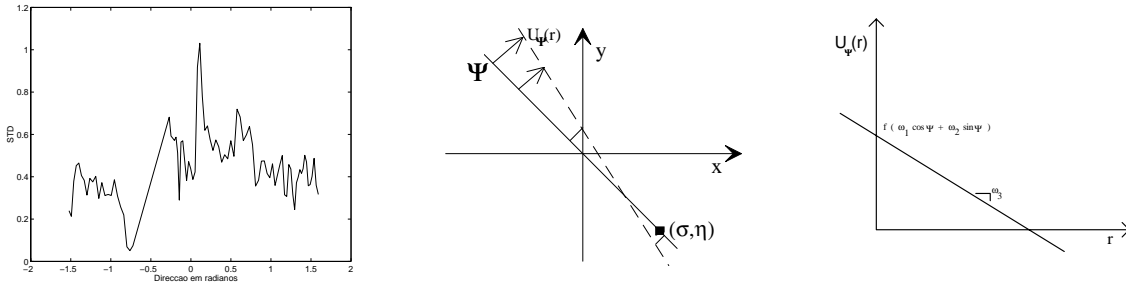


Fig. 4.1: À esquerda: variância dos resíduos para cada direcção ψ ; No centro: a *recta- Ψ* que contém o Foco de Expansão; À direita: comportamento linear de $U_{\Psi}(r)$ associado à *recta- Ψ* .

da função $U_{\Psi}(r)$. Por outro lado, o valor de $U_{\Psi}(0)$ é uma combinação linear de ω_1 e ω_2 , e corresponde à projecção de (ω_1, ω_2) sobre a *recta- Ψ* (ver equação (4.4)).

²Por razões de simplificação de notação faz-se

$$\mathcal{M} \left(v_n - (p - \omega_3 r) | \mathcal{U}(\Sigma, \psi^k) \right) \equiv \mathcal{M} \left(v_n - (p - \omega_3 r) | \psi^k \right).$$

Note-se que Σ está definido sobre a recta radial, logo a localização de Σ nessa recta não influencia o processo de minimização. Em consequência disso, a indicação da recta ψ^k é suficiente para a completa caracterização da função de medida que se pretende minimizar.

É importante sublinhar que este algoritmo usa um conjunto limitado das observações (o fluxo normal circular) e estima apenas um subespaço das estimativas:

- uma restrição em Σ (a direcção do vector (σ, η));
- duas restrições em ω (ω_3 e a projecção de (ω_1, ω_2) sobre a recta Ψ).

4.2.2 O algoritmo de procura da recta- Φ

O algoritmo de procura da recta- Φ usa o subespaço definido pelas rectas paralelas e pode ser sintetizado em quatro passos :

- (i) Consideremos a família das rectas paralelas ϕ^k , descritas pela equação $F_0(\mathbf{x}; k) = 0$. Por definição, a recta- Φ é a recta paralela que contém o Foco de Expansão.
- (ii) Consideremos o subconjunto do fluxo normal perpendicular às rectas ϕ^k . Note-se que, para a recta- Φ , os vectores deste subconjunto dependem exclusivamente do movimento de rotação da câmara.
- (iii) Para os vectores seleccionados em (ii), aplica-se o seguinte estimador de ω :

$$\hat{\omega} = \arg \left\{ \min_{\omega} \left\{ \mathcal{M} \left(v_n - \mathbf{n}^t(\mathbf{B}\omega) | \phi^k \right) \right\} \right\} \quad (4.6)$$

Se a direcção definida pelas rectas ϕ não for colinear com a recta- Ψ , podemos introduzir no estimador a informação de velocidade proveniente do algoritmo de procura da recta- Ψ .

- (iv) A recta- Φ é aquela que minimiza a medida \mathcal{M} dos resíduos.

Note-se que, no caso de se aplicar *a priori* o algoritmo de procura da recta- Ψ , a posição do Foco de Expansão corresponde à intersecção das rectas Ψ e Φ (naturalmente, isto só é válido se estas duas rectas não forem colineares).

A Figura 4.2 mostra o subconjunto do fluxo normal cuja direcção é perpendicular a ϕ ; mostra igualmente a sequência de rectas paralelas com orientação ϕ .

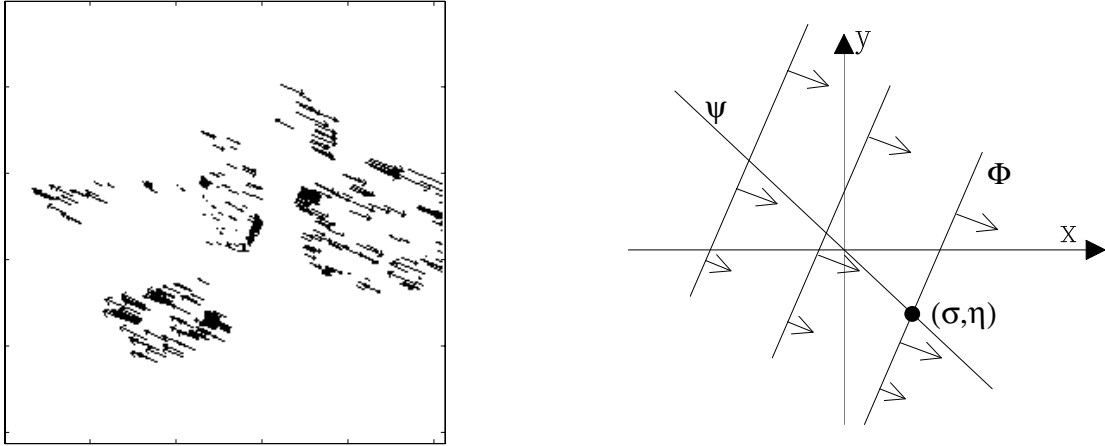


Fig. 4.2: À esquerda: vectores de fluxo normal perpendiculares a uma dada recta seleccionada, ϕ . À direita: sequência de rectas usadas para estimar o movimento da câmara.

4.2.3 O algoritmo de procura da circunferência- Γ

O algoritmo que iremos apresentar, ao contrário dos anteriores, está dependente de algum conhecimento prévio quanto aos parâmetros de velocidade. Este algoritmo serve essencialmente para complementar informação conseguida após a aplicação do algoritmo de procura da recta- Ψ . Sendo assim, supõe-se que é conhecida a recta radial Ψ onde se localiza o Foco de Expansão.

Consideremos o *fluxo normal radial*, $V_\psi(r)$, calculado da seguinte forma:

$$\begin{aligned} \forall(x, y) \quad &: (x, y) = r(\cos \psi, \sin \psi) \\ V_\psi(r) &= \mathbf{n}_r \cdot \mathbf{v} \\ &= \underbrace{\rho(x, y)(r - (\sigma, \eta) \cdot (\cos \psi, \sin \psi))}_{\text{componente translacional}} + \underbrace{\left(\frac{r^2}{f} + f\right) (\omega_1 \sin \psi - \omega_2 \cos \psi)}_{\text{componente rotacional}} \quad (4.7) \end{aligned}$$

Na Secção 3.2.1, observámos que a componente translacional de V_ψ anula-se no subespaço correspondente à circunferência- Γ . Dado o conhecimento da recta Ψ , define-se a família de circunferências tangenciais Γ^k que contêm a origem da imagem e têm diâmetro k , descritas pela equação $F_c(\mathbf{x}; k) = 0$ — ver Figura 4.3. Suponha que (4.7) é calculado para cada elemento da família das circunferências Γ^k . Se dividirmos ambos os termos da

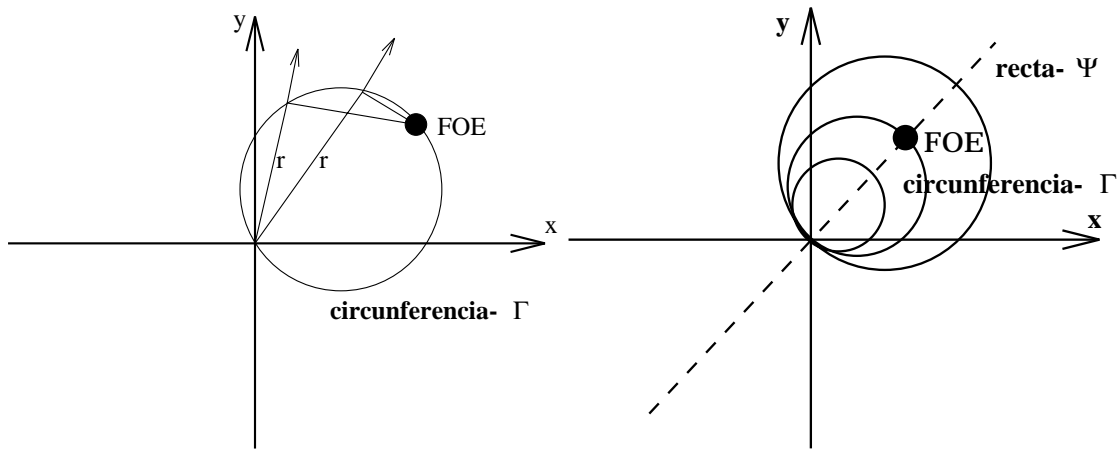


Fig. 4.3: À esquerda: fluxo normal radial na circunferência- Γ . À direita: seqüência de circunferências Γ^k .

equação por $\frac{r^2}{f} + f$, obtemos:

$$\begin{aligned} \forall(x, y) & : (x, y) \in \Gamma^k, \\ T^k(\psi) & = \frac{V_\psi(r)}{\frac{r^2}{f} + f} \\ & = \rho(x, y) \frac{r - (\sigma, \eta) \cdot (\cos \psi, \sin \psi)}{\frac{r^2}{f} + f} + \omega_1 \sin \psi - \omega_2 \cos \psi, \end{aligned} \quad (4.8)$$

onde r é função de k e ψ , de forma a garantir que $T^k(\psi)$ é calculado ao longo da circunferência Γ^k .

A expressão (4.8) mostra que se $\hat{\Gamma}^k$ corresponde à circunferência- Γ , então $T^k(\psi)$ é sinusoidal em ψ . O algoritmo que propomos aqui baseia-se exactamente nesta propriedade e consiste em procurar a circunferência Γ^k que torna $T^k(\psi)$ sinusoidal em ψ . Para estimar os parâmetros da sinusóide precisamos de pelo menos três pontos por cada subespaço Γ^k . Assim sendo, a circunferência- Γ minimiza o seguinte critério:

$$\min_{\omega_1, \omega_2} \left\{ \mathcal{M} \left(\frac{v_n}{\frac{r^2}{f} + f} - (\omega_1 \sin \psi - \omega_2 \cos \psi) | \Gamma^k \right) \right\} \quad (4.9)$$

Este algoritmo usa exclusivamente o fluxo normal radial, necessita do conhecimento prévio da recta Ψ e estima os valores individuais de ω_1 e ω_2 , para além da localização exacta do Foco de Expansão.

4.2.4 Algoritmo baseado em histogramas

Recordemos uma vez mais o método exposto no Capítulo 2 para estimação de parâmetros de velocidade dada a restrição do fluxo normal e que se baseia em duas ideias fundamentais:

- ao conjunto \mathbf{U} do fluxo normal definido para cada Σ_i associa-se uma variância para a estimativa $\hat{\omega}$;
- o valor exacto de Σ corresponde à menor variância da estimativa $\hat{\omega}$.

O método de procura é complexo e computacionalmente pesado.

Nesta secção, iremos analisar o problema da procura inerente ao método, segundo uma abordagem alternativa baseada nas ferramentas da estimação estatística. Como vimos, existe um conjunto de pontos na imagem para os quais o fluxo normal tem componente translacional nula. A esse conjunto chamámos $\mathbf{U}(\Sigma)$. Em N pontos da imagem, o valor exacto de ω é solução para um conjunto restrito P de equações (com $P \leq N$) do sistema seguinte:

$$\begin{bmatrix} \mathbf{n}_1^t \mathbf{v}_1 \\ \vdots \\ \mathbf{n}_N^t \mathbf{v}_N \end{bmatrix} = \begin{bmatrix} \mathbf{n}_1^t \mathbf{B}(\mathbf{x}_1) \\ \vdots \\ \mathbf{n}_N^t \mathbf{B}(\mathbf{x}_N) \end{bmatrix} \omega \quad (4.10)$$

Cada equação do sistema só é válida se, nesse ponto, o fluxo tem parte translacional nula, ou, equivalentemente, o fluxo só depende da rotação da câmara. Os pontos com estas características são os pontos do conjunto $\mathbf{U}(\Sigma)$. Coloquemos a hipótese de que todos os N pontos da imagem pertencem a $\mathbf{U}(\Sigma)$, ou, equivalentemente, que *em todos os pontos a componente translacional do fluxo é nula*. Claro que esta hipótese não é válida para todos os pontos da imagem. Logo, para alguns pontos, a aplicação directa das equações (4.10) na estimação de ω resulta em estimativas erradas. No entanto, podemos esperar o seguinte: **o maior conjunto de pontos da imagem que verifica a equação (4.10) do sistema para um único $\hat{\omega}$, corresponde ao conjunto dos pontos cuja componente translacional do fluxo é nula**. Este resultado não é demonstrado, pelo que é apresentado como resultado empírico.

A detecção dos pontos que observam a hipótese pode ser feito através da aplicação da Transformada de Hough [Ballard and Brown, 1982]. Para cada equação do sistema (4.10) construímos o subespaço das estimativas $\hat{\omega}$. Estas estimativas são introduzidas num acumulador de ω . A Transformada de Hough conduz-nos à selecção das soluções mais “votadas” nesse acumulador. No entanto, este algoritmo implica a criação de um acumulador, que corresponde ao espaço das estimativas ω . Isto significa duas coisas:

- o espaço das estimativas tem que ser discretizado e de dimensão finita, para que a Transformada de Hough seja computacionalmente viável;
- o acumulador de ω é tridimensional.

As características do acumulador dificultam a concretização algorítmica desta abordagem. Contudo, o algoritmo pode ser substancialmente simplificado se inserirmos algumas condições *a priori*. É com este pressuposto que iremos formular o próximo algoritmo de estimação.

Supõe-se então que foi aplicado o algoritmo de procura da recta Ψ , que determina duas restrições em ω e uma restrição em Σ . Os parâmetros de velocidade podem ser completamente estimados para cada ponto que verifique:

$$\begin{cases} \mathbf{n}(\mathbf{x}_i) \cdot (\mathbf{x}_i - \Sigma) = 0 \\ \mathbf{x}_i \notin \text{recta-}\Psi \end{cases} \quad (4.11)$$

Como vimos, o fluxo normal nestes pontos só depende da rotação da câmara e verifica a seguinte equação:

$$v_n(\mathbf{x}_i) = \mathbf{n}(\mathbf{x}_i) \cdot \mathbf{B}(\mathbf{x}_i)\omega \quad (4.12)$$

Assim, para cada ponto da imagem \mathbf{x}_i , podemos estimar os valores de ω_{1i} and ω_{2i} usando a equação (4.12) e as duas restrições conhecidas de ω . Dado que, só os pontos que obedecem à restrição (4.11) poderão verificar (4.12), espera-se que as estimativas de ω_{1i} e ω_{2i} incidirão numa larga gama de valores, concentrando-se especialmente em torno dos valores correctos de ω_1 e ω_2 .

Deste modo, podemos calcular os valores de $(\omega_{1i}, \omega_{2i})$ para todos os N pontos, determinar os histogramas respectivos e seleccionar os valores mais “votados” (ou seja, a moda³ de $(\omega_{1i}, \omega_{2i})$). Para os pontos da imagem associados aos $(\omega_{1i}, \omega_{2i})$ mais “votados”, as rectas perpendiculares a $\mathbf{n}(\mathbf{x}_i)$ deverão intersectar a recta- Ψ no Foco de Expansão.

Observa-se que, determinar a moda de ω_{1i} e ω_{2i} é equivalente a estimar os valores de (ω_1, ω_2) que minimizam a moda dos respectivos resíduos quadráticos (calculados a partir das restrições conhecidas do algoritmo de procura da recta- Ψ). O problema pode ser formulado como se segue:

$$\min_{\omega_1} \{ \mathcal{M}_{\text{moda}} (v_n - \mathbf{n}^t \mathbf{B} \boldsymbol{\omega} | I) \} \quad (4.13)$$

$$\min_{\omega_2} \{ \mathcal{M}_{\text{moda}} (v_n - \mathbf{n}^t \mathbf{B} \boldsymbol{\omega} | I) \} \quad (4.14)$$

onde I é o conjunto dos pontos da imagem.

Conclui-se assim que o algoritmo proposto calcula completamente os parâmetros de velocidade da câmara. Observa-se que essa estimação não se faz através da escolha explícita de subespaços da imagem onde o fluxo normal possui determinadas propriedades, mas sim através da concepção duma métrica específica da função de medida \mathcal{M} .

Relativamente a este método, é de sublinhar que:

- a métrica usada (a moda) implica a escolha implícita do subconjunto $\mathcal{U}(\boldsymbol{\Sigma})$ e é aplicada a todos os pontos da imagem;
- o algoritmo desenvolvido tem como pressuposto a aplicação prévia do algoritmo de procura da recta- Ψ — no entanto, esta restrição existe apenas por razões computacionais.

4.3 Algoritmos de Procura Sequencial e Paralela

Neste momento, dispomos de um conjunto de algoritmos de reduzida complexidade que permitem estimar parâmetros de velocidade da câmara, e que podem ser organizados

³nesta tese, aplicamos o operador estatístico *moda* a um espaço discreto e finito de valores reais, resultado de uma quantificação do espaço real contínuo.

de forma a recuperar completamente o movimento. Como vimos anteriormente, existem algoritmos que necessitam de informação de velocidade *a priori*, outros não necessitam dessa informação. Este facto sugere uma organização dos algoritmos globais de estimação que pode ser paralela ou sequencial.

A primeira abordagem consiste numa disposição paralela dos algoritmos de procura da recta- Φ . O domínio de procura deste algoritmo é descrito pela equação $F_0(\mathbf{x}; k) = 0$ (rectas paralelas) — ilustrado na Figura 3.3a. Cada algoritmo de procura minimiza a medida dos resíduos para uma determinada família de rectas paralelas associada a uma dada direcção:

$$\begin{array}{l} \boxed{F_0(\mathbf{x}; k) \rightarrow \phi_0^k} \Rightarrow \min_{\phi_0^k} \left\{ \min_{\omega} \mathcal{M}(\mathbf{R} | \phi_0^k) \right\} \Rightarrow \hat{\phi}_0 \\ \boxed{F_1(\mathbf{x}; k) \rightarrow \phi_1^k} \Rightarrow \min_{\phi_1^k} \left\{ \min_{\omega} \mathcal{M}(\mathbf{R} | \phi_1^k) \right\} \Rightarrow \hat{\phi}_1 \\ \vdots \\ \boxed{F_n(\mathbf{x}; k) \rightarrow \phi_n^k} \Rightarrow \min_{\phi_n^k} \left\{ \min_{\omega} \mathcal{M}(\mathbf{R} | \phi_n^k) \right\} \Rightarrow \hat{\phi}_n \end{array}$$

A intersecção das rectas $\{\hat{\phi}_0, \hat{\phi}_1, \dots, \hat{\phi}_n\}$ corresponde à localização do Foco de Expansão⁴.

Cada algoritmo paralelo calcula uma estimativa para a componente de rotação da câmara. A estimativa global de ω é calculada segundo uma média pesada dessas estimativas (os pesos podem ser atribuídos com base nas variâncias das estimativas resultantes de cada algoritmo).

Poderemos transformar a estrutura paralela numa estrutura sequencial, se a informação de um algoritmo for utilizada por outro consequente. Assim, a segunda abordagem consiste num algoritmo sequencial descrito em dois passos:

Primeiro Passo: O algoritmo de procura da recta- Ψ determina a direcção do Foco

⁴Para aumentar a robustez do algoritmo, pode-se construir um acumulador com a dimensão da imagem, similar ao usado na Transformada de Hough. A recta- Φ encontrada por cada algoritmo é adicionada no acumulador e o ponto que maximiza esse acumulador corresponde ao Foco de Expansão.

de Expansão, calcula ω_3 e uma restrição em ω_1 e ω_2 .

$$\boxed{F_r(\mathbf{x}; k) \rightarrow \psi^k} \Rightarrow \min_{\psi^k} \left\{ \min_{\omega} \mathcal{M}(\mathbf{R}|\psi^k) \right\} \Rightarrow \hat{\psi}$$

Segundo Passo: O segundo passo do algoritmo sequencial usa os resultados do algoritmo anterior e estima os restantes parâmetros de velocidade. Apresentámos três algoritmos que completam o processo de estimação:

1. O **algoritmo de procura da recta- Φ** selecciona a recta (a recta- Φ) que contém o Foco de Expansão, numa família de rectas paralelas com direcção não colinear com a recta- Ψ .

$$\hat{\psi} \Rightarrow \boxed{F_0(\mathbf{x}; k) \rightarrow \phi_1^k} \Rightarrow \min_{\phi_1^k} \left\{ \min_{\omega} \mathcal{M}(\mathbf{R}|\phi_1^k) \right\} \Rightarrow \hat{\phi}_1$$

2. O **algoritmo de procura da circunferência- Γ** usa a direcção conhecida da recta- Ψ .

$$\hat{\psi} \Rightarrow \boxed{F_c(\mathbf{x}; k) \rightarrow \Gamma^k} \Rightarrow \min_{\Gamma^k} \left\{ \min_{\omega} \mathcal{M}(\mathbf{R}|\Gamma^k) \right\} \Rightarrow \hat{\Gamma}$$

3. O **algoritmo baseado em histogramas** é aplicado em todos os pontos da imagem I .

$$\psi \Rightarrow \min_{\omega} \mathcal{M}_{\text{moda}}(\mathbf{R}|I) \Rightarrow \hat{\Sigma}, \hat{\omega}$$

Neste capítulo apresentámos um conjunto de algoritmos de estimação que assentam em dois conceitos fundamentais:

- a subdivisão do espaço das observações e das estimativas reduz a complexidade algorítmica do método de estimação apresentado em (2.17);
- a robustez do algoritmo só é possível se impusermos critérios rigorosos na métrica das funções de medida residual \mathcal{M} . Este ponto é desenvolvido no Apêndice A.

Propomos, de seguida, apresentar um conjunto de resultados experimentais, obtidos com imagens sintéticas e reais, para vários tipos de movimento de translação e rotação.

5. Resultados Experimentais

Neste capítulo, testamos o método de estimação com restrição do fluxo normal com um conjunto de sequências de imagens de natureza real ou sintética. As sequências de imagens foram escolhidas no sentido de ilustrar as características de cada algoritmo de procura proposto nesta tese, e correspondem a diferentes movimentos de câmara. Começa-se por apresentar cada uma dessas sequências.

5.1 As sequências de imagens

As sequências de imagens foram numeradas de 1 a 4 e têm as características apresentadas na Tabela 5.1. Todas as imagens são pré-filtradas com um filtro gaussiano passa-baixo

Sequência	Dimensão (pixel)	Dist. focal(pixel)	Movimento	Obs.
1	256×256	302	Linear e Angular	Sintética
2	256×256	302	Linear e Angular	Sintética
3	322×244	229	Linear	Real
4	256×340	427	Linear e Angular	Real

Tabela 5.1: Sequências de imagens.

(máscara 5×5 , desvio padrão: 1.0). As derivadas temporal e espacial, necessárias para o cálculo do fluxo, são obtidas aplicando uma máscara de derivação de 5 pontos:

$$[1 \quad -8 \quad 0 \quad 8 \quad -1]/12.$$

Nas imagens sintéticas, a localização do Foco de Expansão é determinada a partir do conhecimento dos parâmetros de velocidade da câmara. Nas imagens reais, o Foco de Expansão é medido usando pontos característicos da cena, e é indicado nas figuras pela sigla FOE.

Sequência 1: Movimento linear ao longo do eixo Z e rotação em torno de Z

Nesta sequência de imagens sintéticas, a câmara move-se ao longo do eixo óptico e sofre uma rotação constante em torno desse eixo de 1° /imagem. A Figura 5.1 ilustra a sequência de imagens obtida.



Fig. 5.1: Sequência 1, com rotação em torno do eixo óptico.

Sequência 2: Movimento linear e angular

Na segunda experiência, o movimento é mais complexo, dado que a câmara sofre uma rotação em torno do eixo Y de $\omega_2 = -0.00785\text{rad/imagem}$ (-0.45° /imagem), enquanto translada numa direcção constante não colinear com qualquer dos seus eixos. A cena corresponde a uma superfície planar texturada, como se mostra na Figura 5.2.

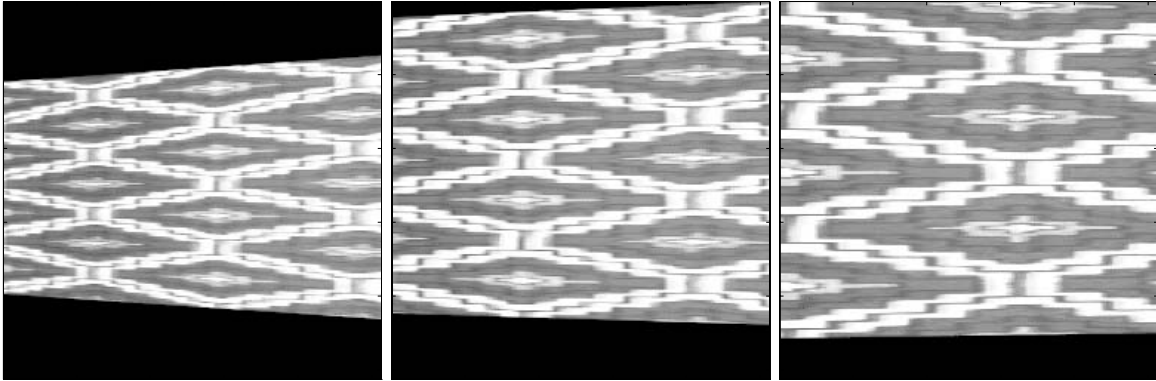


Fig. 5.2: Sequência 2, com movimento de translação e rotação.

Sequência 3: Experiência real com movimento linear

Nesta experiência com imagens reais, a câmara tem um movimento puramente translacional, como ilustra a Figura 5.3. O movimento rotacional associado é aproximadamente

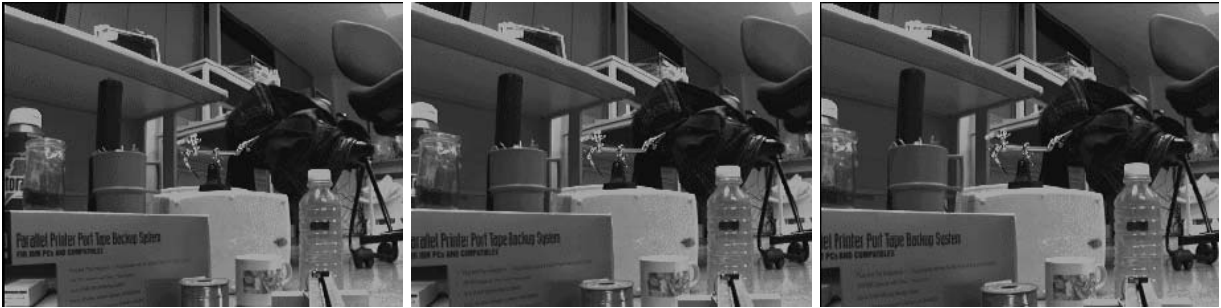


Fig. 5.3: Sequência 3 com imagens reais.

nulo.

Sequência 4: Experiência real com movimento linear e angular

Na experiência final, igualmente com imagens reais, a câmara sofre um movimento translacional e rotacional — Figura 5.4. A rotação da câmara em torno do eixo Y é de $\omega_2 = -0.005$ rad/imagem (cerca de 0.3° /imagem). O Foco de Expansão e o Foco de Rotação (como são definidos nesta tese) estão sobre o eixo x do plano de imagem. O vector de velocidade linear tem componente não nula na direcção do eixo óptico.



Fig. 5.4: Sequência 4, com imagens reais.

5.2 Algoritmo paralelo

Disponemos de um conjunto de algoritmos de procura da recta- Φ , organizados segundo a forma de um algoritmo paralelo. O algoritmo é aplicado à Sequência 2 e fica definido com as seguintes leis de minimização:

$$\begin{aligned}
 & \min_{\Phi_{0.2}} \left\{ \min_{\omega} \left\{ \mathcal{M}_{\text{LS}}(R_i | \phi_{0.2}^k) \right\} \right\} \\
 & \min_{\Phi_{0.4}} \left\{ \min_{\omega} \left\{ \mathcal{M}_{\text{LS}}(R_i | \phi_{0.4}^k) \right\} \right\} \\
 & \min_{\Phi_{0.8}} \left\{ \min_{\omega} \left\{ \mathcal{M}_{\text{LS}}(R_i | \phi_{0.8}^k) \right\} \right\} \\
 & \min_{\Phi_{1.6}} \left\{ \min_{\omega} \left\{ \mathcal{M}_{\text{LS}}(R_i | \phi_{1.6}^k) \right\} \right\} \\
 & \min_{\Phi_{2.4}} \left\{ \min_{\omega} \left\{ \mathcal{M}_{\text{LS}}(R_i | \phi_{2.4}^k) \right\} \right\}
 \end{aligned} \tag{5.1}$$

onde cada índice de Φ representa a direcção em radianos da família de rectas paralelas — para um índice de 0.2 radianos, a família de rectas é descrita pela equação

$$F_{0.2}(\mathbf{x}; k) = -x \sin 0.2 + y \cos 0.2 + k.$$

A métrica de \mathcal{M} corresponde à soma do quadrado dos resíduos (LS).

A estimativa de cada recta Φ é introduzida na Figura 5.5, sobreposta à imagem original. As estimativas de ω são apresentadas na Tabela 5.2. Conclui-se que a estimativa do Foco de Expansão (pontos da imagem onde as rectas Φ se intersectam mais vezes) encontra-se próxima da sua localização exacta representada na Figura por um quadrado preto (FOE). Escolhendo a estimativa de rotação da Tabela 5.2 com menor variância, obtém-se:

$$\hat{\omega} = [-0.000324 \quad -0.00915 \quad -0.000632]^T \text{ rad/imagem.}$$

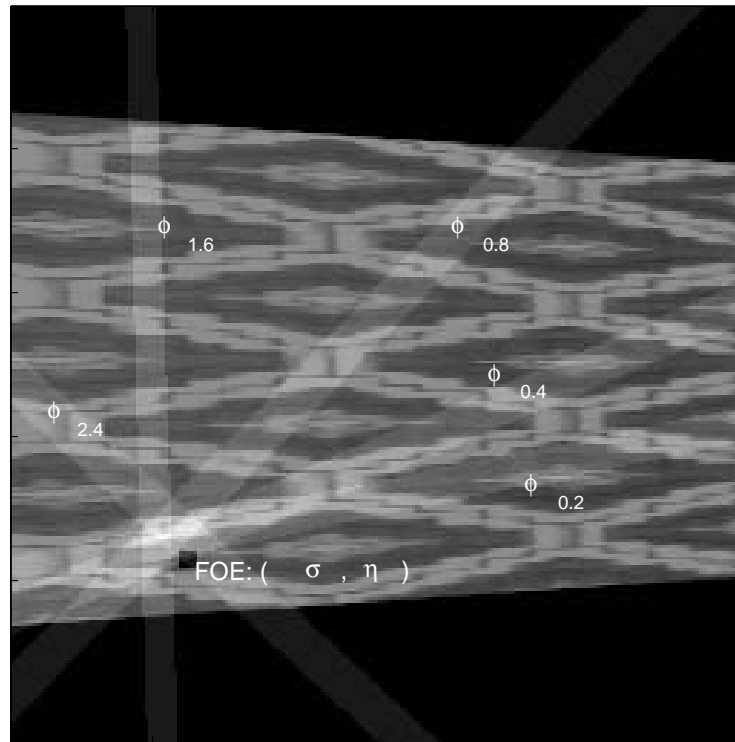


Fig. 5.5: Algoritmo paralelo aplicado à Sequência 2.

ângulo- Φ (rad)	ω_1 (rad/fr)	ω_2 (rad/fr)	ω_3 (rad/fr)	variância
0.2	-0.000324	-0.00915	-0.000632	0.3159
0.4	0.005042	-0.01683	-0.00479	0.4238
0.8	-0.000168	-0.00798	-0.004022	0.6099
1.6	0.003616	-0.007918	-0.000715	0.3427
2.4	0.02886	0.02338	0.006928	0.5012

Tabela 5.2: Estimação dos parâmetros de rotação para um conjunto de cinco algoritmos paralelos de procura da recta- Φ , usando famílias de rectas- ϕ com diferentes direcções.

Observa-se que, por um lado, os parâmetros ω_1 and ω_3 são extremamente baixos, como era esperado, por outro lado, a estimativa de ω_2 corresponde à recuperação do seu valor verdadeiro $\omega_2 = -0.00785\text{rad/imagem}$ com um erro relativo de cerca de 15%. Verifica-se igualmente uma apreciável contaminação de ruído no conjunto das estimativas, essencialmente porque se aplica o estimador LS a um espaço limitado de observações. Este facto pode ser “minimizado” usando um conjunto mais extenso de algoritmos de procura (para outras direcções Φ).

5.3 Primeiro passo numa abordagem sequencial

De forma a realizar o algoritmo sequencial proposto nesta tese, aplica-se a todas sequências o primeiro passo do algoritmo: **o algoritmo de procura da recta- Ψ** .

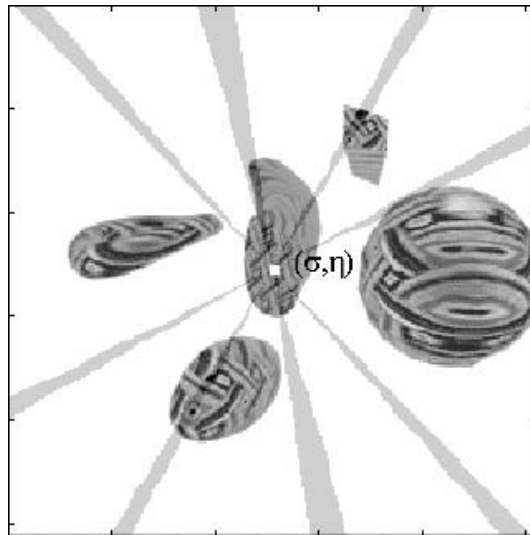


Fig. 5.6: Algoritmo de procura da recta- Ψ aplicado à Sequência 1.

Na primeira experiência, a câmara move-se ao longo do eixo óptico e sofre uma rotação em torno desse eixo de $1^\circ/\text{imagem}$. No caso particular do Foco de Expansão se localizar na origem do plano de imagem, o algoritmo de procura da recta- Ψ é suficiente para recuperar completamente os parâmetros de velocidade da câmara. De facto, neste caso, todas as rectas que passam pela origem do plano de imagem verificam as propriedades da recta-

Ψ , ou seja, quando o Foco de Expansão corresponde ao centro da imagem, existe uma infinidade de rectas- Ψ .

Por conseguinte, ao aplicar-se o algoritmo proposto, foram detectadas múltiplas rectas- Ψ (isto é, foram detectadas múltiplas rectas radiais com as propriedades da recta- Ψ), o que permite concluir experimentalmente que o Foco de Expansão se encontra na origem do plano de imagem. A Figura 5.6 ilustra um conjunto de rectas- Ψ detectadas. Pode-se mostrar que, se seleccionarmos duas rectas Ψ com direcções genéricas Ψ_1 e Ψ_2 , então todos os parâmetros de rotação podem ser explicitamente estimados, usando as seguintes expressões:

$$U_{\Psi_1}(r) = f(\omega_1 \cos \Psi_1 + \omega_2 \sin \Psi_1) - \omega_3 r$$

$$U_{\Psi_2}(r) = f(\omega_1 \cos \Psi_2 + \omega_2 \sin \Psi_2) - \omega_3 r$$

Nesta experiência, a função $U_{\Psi}(r)$ calculada para cada recta Ψ é uma função afim com a seguinte forma:

$$U_{\Psi}(r) = -0.0178r \quad (5.2)$$

O termo constante (correspondente ao valor $\omega_1 \cos \Psi_1 + \omega_2 \sin \Psi_1$) é, para todas as rectas seleccionadas, um valor da ordem de 10^{-4} rad/imagem. Conclui-se então que $\hat{\omega}_1$ e $\hat{\omega}_2$ são aproximadamente nulos e $\hat{\omega}_3 = 1.02^\circ/\text{imagem}$, o que representa um erro de 2%.

Para as restantes sequências, o algoritmo calcula um conjunto de restrições no espaço das estimativas, como se segue:

- O valor exacto do Foco de Expansão de cada sequência é representado na Figura 5.7 por uma marca quadrada.
- É aplicado o estimador LTS (descrito no Apêndice A.1).
- A Figura 5.7 apresenta, para cada sequência de imagens, uma estimativa da recta- Ψ .
- A Figura 5.7 mostra igualmente as respectivas variâncias das estimativas (produzidas pelo estimador LTS) ao longo de cada direcção ψ . Note-se que a variância é calculada após a rejeição dos *outliers*.

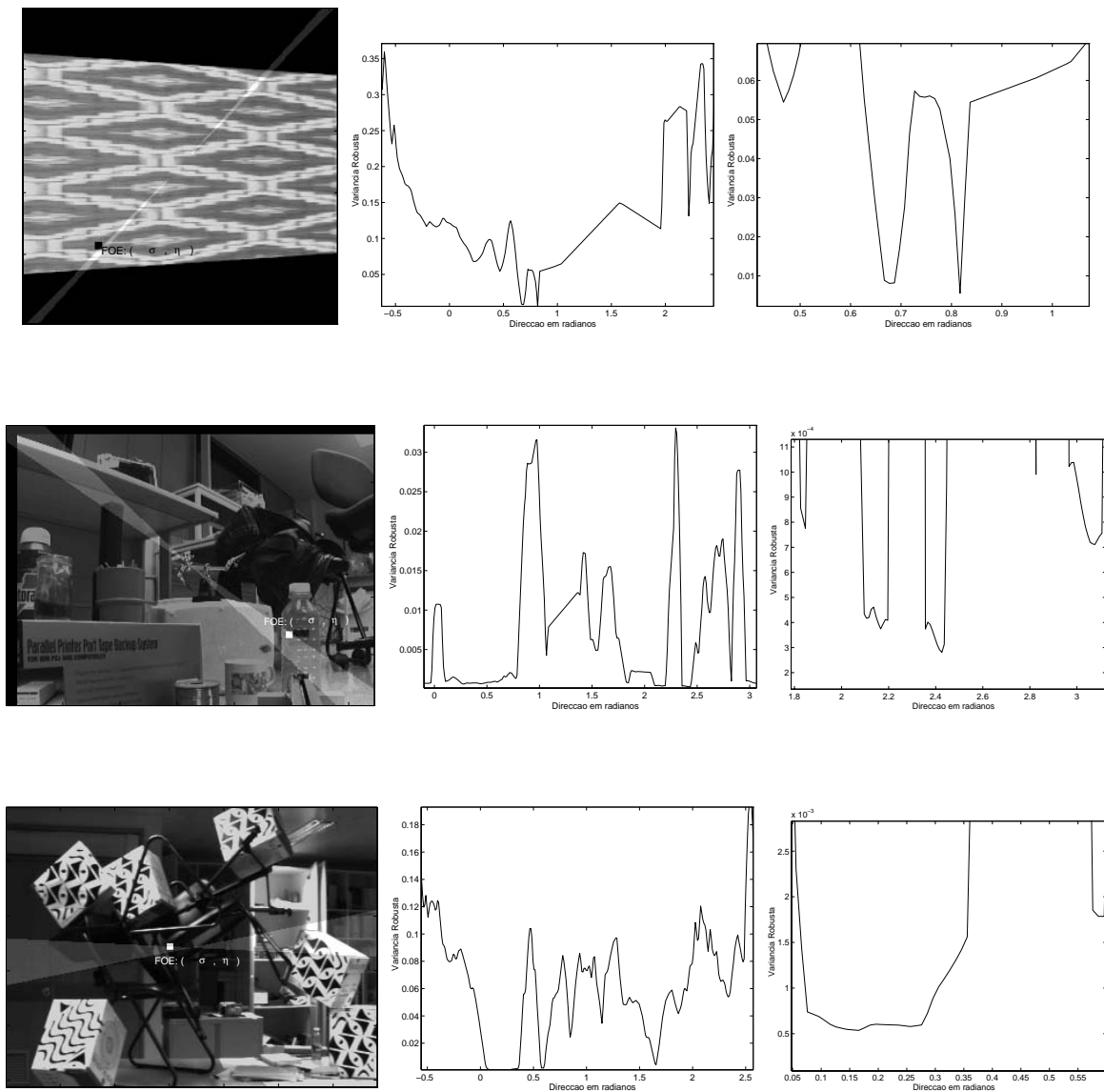


Fig. 5.7: Resultados de estimação do algoritmo de procura da recta- Ψ aplicados às sequências 2, 3 and 4. À esquerda: a recta- Ψ encontrada pelo primeiro passo do algoritmo. No centro: variâncias das estimativas ao longo de cada recta com orientação ψ . À direita: área de baixa variância (aumentada) em torno da direcção do vector do Foco de Expansão.

- Como se observa, as variâncias são significativamente baixas em torno da direcção Ψ^1 e formam uma região angular na qual se espera que se localize o Foco de Expansão.

A Tabela 5.3 mostra os respectivos parâmetros de velocidade estimados através do algoritmo de procura da recta- Ψ , para cada sequência proposta. Conclui-se que as estimativas rotacionais são próximas dos valores esperados.

Sequência	ângulo da recta- Ψ	$\omega_1 \cos \Psi + \omega_2 \sin \Psi$	ω_3
2	0.817 rad	-0.005717	-0.001496
3	-0.707 rad	-0.000135	0.000125
4	0.166 rad	0.00000	-0.000011

Tabela 5.3: Estimativas rotacionais usando o algoritmo de procura da recta- Ψ . Todos os valores de rotação são em *rad/imagem*.

5.4 Segundo passo do algoritmo sequencial

De forma a completar o processo de estimação de parâmetros, aplicam-se três algoritmos que realizam a segunda parte do algoritmo sequencial de estimação: o algoritmo de procura da recta- Φ , o algoritmo de procura da circunferência- Γ e o algoritmo baseado em histogramas. A Tabela 5.4 representa as respectivas estimativas de (ω_1, ω_2) . Note-se que a

Seq.	(ω_1, ω_2)	Alg. recta- Φ	Alg. da circ.- Γ	Alg. Histogr.
2	(0, -0.0079)	(0.00059, -0.0066)	(-0.00079, -0.00886)	(0.00029, -0.0072)
3	(0, 0)	(-0.00013, -0.00020)	(0.00000, -0.00000)	(0.00026, 0.00047)
4	(0, -0.0050)	(0.00000, -0.0055)	(0.00002, -0.0054)	(0.00039, -0.0048)

Tabela 5.4: Estimativas rotacionais resultantes da segunda parte do algoritmo. Todos os valores de rotação são em *rad/imagem*.

¹ou seja, da direcção da recta radial que atravessa o Foco de Expansão

direcção do Foco de Expansão (a recta- Ψ) e a terceira componente da velocidade angular ω_3 são conhecidos do primeiro passo do algoritmo (o algoritmo de procura da recta- Ψ).

Ao aplicar o algoritmo de procura da recta- Φ , o Foco de Expansão é dado pela intersecção da recta- Ψ com a recta- Φ . A estimativa da recta- Φ para cada sequência de imagens é apresentada nas Figuras 5.8a, 5.9a, 5.10a.

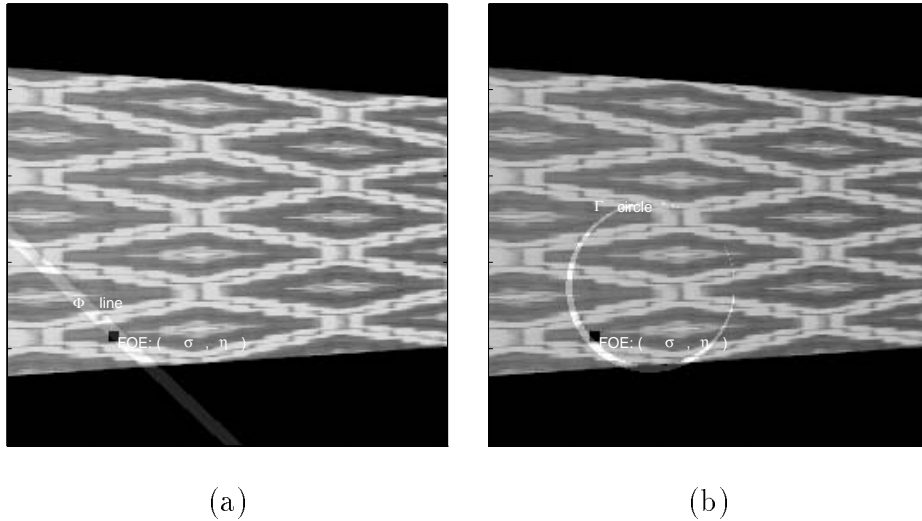


Fig. 5.8: Resultados da estimação através do (a) algoritmo de procura da recta- Φ e do (b) o algoritmo de procura da circunferência- Γ , aplicados à Sequências 2.

O algoritmo de procura da circunferência- Γ estima uma circunferência que intersecta a recta- Ψ numa localização próxima do ponto exacto do Foco de Expansão, como mostram as Figuras 5.8b, 5.9b, 5.10b.

No que se refere à aplicação do algoritmo baseado em histogramas, obtém-se um conjunto de histogramas para os valores de ω_1 e ω_2 , calculados nos pontos da imagem. Observa-se na Figura 5.11 que o máximo de cada histograma está bem definido e determina uma estimativa para os valores rotacionais, apresentada na Tabela 5.4. Os vectores de fluxo normal que correspondem aos valores (ω_1, ω_2) seleccionados (os máximos dos histogramas) têm uma direcção que intersecta a recta- Ψ em diferentes pontos. O número de intersecções ao longo da recta- Ψ é apresentado na Figura 5.11. O número máximo de intersecções fica próximo da localização real do Foco de Expansão, que, na figura, é representada pela sua projecção radial na recta- Ψ estimada (linha tracejada).



Fig. 5.9: Resultados da estimação através do (a) algoritmo de procura da recta- Φ e do (b) o algoritmo de procura da circunferência- Γ , aplicados à Sequências 3.

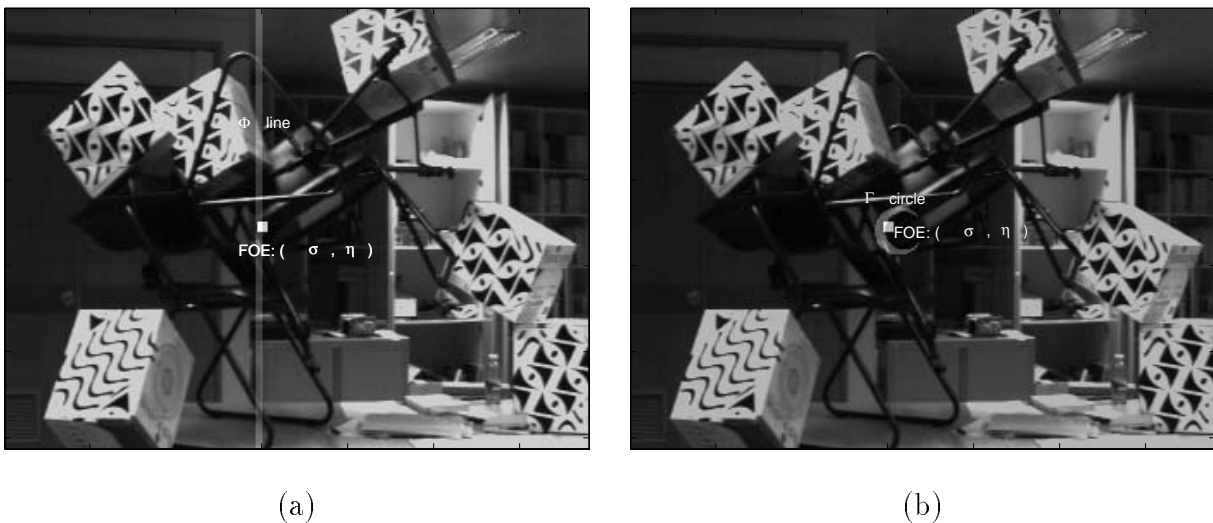


Fig. 5.10: Resultados da estimação através do (a) algoritmo de procura da recta- Φ e do (b) o algoritmo de procura da circunferência- Γ , aplicados à Sequências 4.

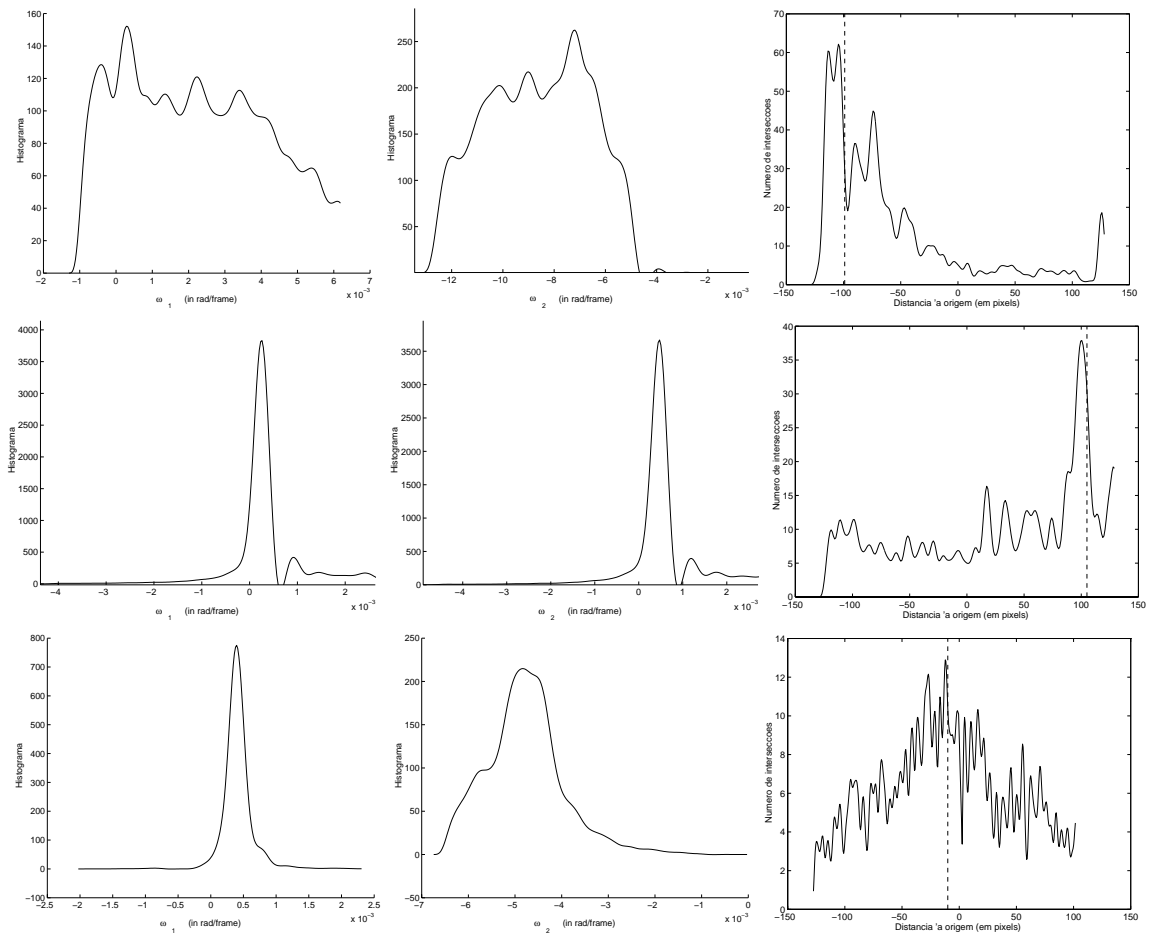


Fig. 5.11: Resultados da estimação do algoritmo baseado em histogramas, aplicado às seqüências 2, 3 e 4.

5.5 Cálculo do tempo de colisão

Dispondo do fluxo normal definido para cada ponto da imagem e dos parâmetros de velocidade da câmara, é então possível recuperar o tempo de colisão da câmara para cada ponto da imagem, dado pelos parâmetros $\rho(\mathbf{x}) = W/Z(\mathbf{x})$, por aplicação da equação (2.4). Note-se que, enquanto os parâmetros de velocidade dispõem de um conjunto vasto de observações que asseguram a robustez da própria estimação, o cálculo de cada parâmetro $\rho(\mathbf{x})$ (definido em cada ponto da imagem) dispõe de uma única observação, dado que não se introduzem pressupostos quanto à forma tridimensional observada.

Assim, a determinação do parâmetro $\rho(\mathbf{x})$ está sujeita ao ruído existente em cada observação. Por isso, é conveniente limitar o cálculo de $\rho(\mathbf{x})$ para observações bem comportadas (ou seja, para observações que estejam dentro duma determinada gama de amplitudes vectoriais) em zonas de elevada textura (com módulo do gradiente elevado).



Fig. 5.12: a) Cena correspondente à Sequência 3, composta por dois *planos de profundidade* distintos. b) Zonas de profundidade dominantes: a zona mais escura corresponde a uma zona mais próxima do observador; a zona mais clara corresponde a uma zona mais afastada.

Para ilustrar este procedimento, optou-se por calcular $\rho(\mathbf{x})$ para uma sequência muito simples, com imagens reais: a Sequência 3. Esta sequência tem uma característica inte-

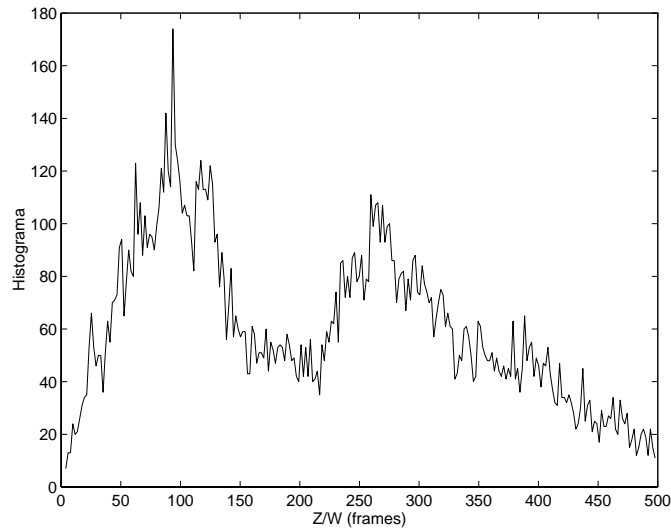


Fig. 5.13: Histograma de Z/W para a Sequência 3.

ressante que, em relação às restantes sequências, permite ilustrar melhor a reconstrução qualitativa do meio: a cena visualizada é composta por dois *planos de profundidade* marcadamente distintos (a Figura 5.12a ilustra estas duas zonas, divididas por um traço branco). Espera-se, assim, que os valores de $\rho(\mathbf{x})$ estejam “agrupados” em duas zonas distintas, correspondentes a duas profundidades dominantes.

Na Figura 5.13, podemos observar o histograma dos parâmetros $1/\rho(\mathbf{x})$ calculados para toda a imagem, onde se confirma a existência dos dois *planos de profundidade* dominantes. A Figura 5.12b mostra a localização dessas zonas, marcadas com dois níveis de cinzento distintos. Obtém-se deste modo uma estimativa bastante satisfatória da localização espacial dos objectos observados — com clara aplicação em tarefas de detecção de obstáculos e/ou de zonas desobstruídas.

5.6 Síntese

Conclui-se que os três algoritmos recuperam com um baixo nível de erro os parâmetros de rotação da câmara e de localização do Foco de Expansão. Observa-se ainda que o erro de estimação varia de sequência para sequência e depende do método de estimação.

Como seria de esperar, o desempenho de cada algoritmo depende dos dados oferecidos pela imagem e dos parâmetros intrínsecos associados a cada algoritmo. No Apêndice B discute-se algumas alterações de ordem prática efectuadas nos algoritmos de procura aplicados neste capítulo, no sentido de garantir a existência de um número suficientemente elevado de observações em cada subespaço de procura (tornando fiável cada estimativa realizada em cada subespaço). Neste trabalho, uma estimativa só é considerada válida, se o número das respectivas observações for superior a um determinado limiar de natureza empírica — que nos estimadores utilizados foi de 10 observações. Nas experiências efectuadas, o número de observações por subespaço de procura varia tipicamente num intervalo compreendido entre as 20 e as 80 observações.

No final deste capítulo, recuperámos ainda alguns parâmetros associados à estrutura do meio, através do cálculo de $\rho(\mathbf{x})$, que, aparte um termo constante multiplicativo (dado por W), oferece uma estimativa da coordenada Z dos objectos visualizados.

6. Extensão para Imagens

Log-Polares

Em oposição às abordagens dominantes que consideram os aspectos algorítmicos de estimação dissociados da geometria do sensor visual, propomos explorar a relação que se pode estabelecer entre o desenho do sensor e a concepção do estimador. Iremos mostrar que pode ser vantajoso abordar o problema de estimação no plano dos algoritmos e da especificação da geometria ligada ao sensor visual.

Neste sentido, tem-se adoptado, nos últimos anos, mecanismos de amostragem da imagem não-uniforme em sistemas de controlo por visão, com inspiração nos sistemas biológicos. Como se sabe, a sua aplicação tem implicações fundamentalmente a nível (1) da redução da complexidade dos algoritmos inerentes ao processamento da imagem ou (2) do desempenho de certas tarefas visuais. O método mais usual de amostragem não-uniforme (aplicado à imagem) corresponde à transformação log-polar [Wallace et al., 1995, Weiman, 1990, Griswald et al., 1992, Tistarelli and Sandini, 1993, Bernardino and Santos-Victor, 1996].

Uma imagem log-polar tem três características essenciais:

1. possui maior resolução na zona central da imagem (e menor resolução nas zonas periféricas);
2. contém menor quantidade total de informação;
3. a transformação que lhe está associada é rotacionalmente simétrica.

Estas três características têm uma relação directa com as principais razões pelas quais a transformação log-polar é introduzida nesta tese como extensão ao caso cartesiano:

- A representação log-polar de uma imagem implica uma redução na dimensão dos dados disponíveis (ou, equivalentemente, implica menos tempo de cálculo).
- A geometria log-polar parece mais adaptada a outras tarefas visuais: apresentamos como exemplo a tarefa de vergência para sistemas binoculares em [Bernardino and Santos-Victor, 1996].
- A sua geometria rotacionalmente simétrica oferece condições ideais para seleccionar o fluxo normal circular e radial, ou para efectuar certos algoritmos de procura, tal como o que se aplica em subespaços de procura da recta Ψ .

Neste capítulo, faz-se a extensão dos métodos de estimação apresentados em coordenadas cartesianas para coordenadas log-polares. Define-se transformação log-polar como a transformação conforme do plano cartesiano definido pelos eixos (x, y) para o plano log-polar definido pelos eixos (λ, γ) , descrita do seguinte modo:

$$\begin{cases} \lambda = \log_l r \\ \gamma = \psi \end{cases} \quad (6.1)$$

onde l é o factor logarítmico associado à transformação e $(x, y) = r(\cos \psi, \sin \psi)$, como é usual.

A transformação log-polar exhibe uma singularidade na origem, que normalmente é evitada limitando o domínio da transformação a uma coroa circular. Em consequência, define-se um limite inferior e um limite superior para a coordenada log-polar λ :

$$\lambda \in [\log_l r_{min}, \log_l r_{max}] \quad (6.2)$$

onde r_{max} é determinado pela dimensão da imagem e r_{min} descreve uma vizinhança em torno da origem para a qual a transformação não está definida. A Figura 6.1 mostra o domínio de transformação cartesiano e o plano log-polar correspondente. A Figura 6.2

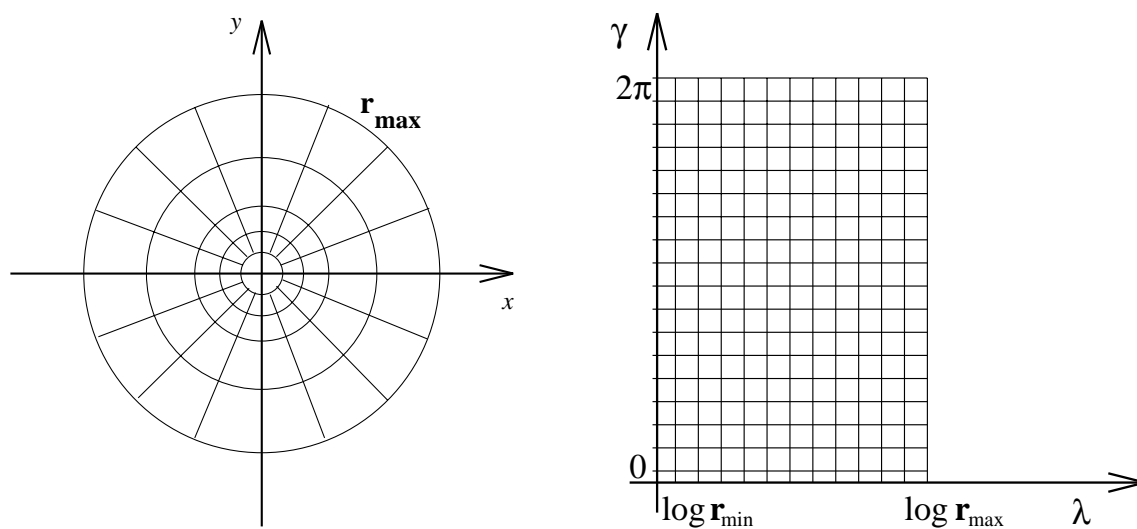


Fig. 6.1: Transformação log-polar.

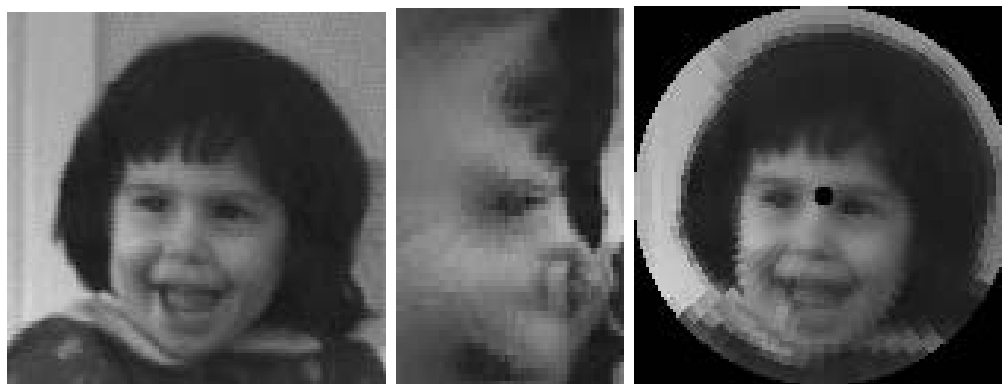


Fig. 6.2: A imagem original cartesiana (à esquerda) é “mapeada” segundo uma representação log-polar (no centro). À direita, a imagem é “remapeada” no plano cartesiano.

permite ilustrar uma transformação log-polar de uma imagem real adquirida em coordenadas cartesianas. Observamos nesta figura as principais características da representação log-polar: redução da quantidade de informação na imagem log-polar; preservação da estrutura da imagem; menor resolução nas áreas periféricas; existência de uma singularidade na origem; e a natureza rotacionalmente simétrica da transformação.

Em coordenadas log-polares as definições do fluxo óptico e normal mantêm-se naturalmente. Assim, $E(\lambda, \gamma)$ corresponde ao nível de brilho na imagem; E_λ e E_γ são as derivadas espaciais do brilho; e a restrição do fluxo normal é definida pela projecção do fluxo óptico $\mathbf{v}(\lambda, \gamma)$ sobre o gradiente normalizado $\mathbf{n} = \frac{\nabla \mathbf{E}}{\|\nabla \mathbf{E}\|}$, onde E_x e E_y se relacionam com E_λ e E_γ da seguinte forma:

$$\begin{cases} E_x = \frac{1}{r} (E_\lambda / \log l, E_\gamma) \cdot (\cos \psi, -\sin \psi) \\ E_y = \frac{1}{r} (E_\lambda / \log l, E_\gamma) \cdot (\sin \psi, \cos \psi) \end{cases} \quad (6.3)$$

Começemos por verificar que o fluxo normal circular e radial tem uma representação bastante simples em coordenadas log-polares. De facto, observa-se que o fluxo normal radial tem derivada angular E_γ nula; da mesma forma, a derivada radial E_λ do fluxo normal circular é zero. Isto significa que os vectores de fluxo radial e circular têm respectivamente uma orientação **horizontal** (ao longo do eixo λ) e **vertical** (ao longo do eixo γ).

Assim, a representação log-polar parece estar bem adaptada à aplicação do algoritmo de procura da recta- Ψ , baseado no fluxo normal circular, ou à aplicação do algoritmo de procura da circunferência- Γ , baseado no fluxo normal radial. Em consequência disso, iremos em seguida desenvolver um método de estimação aplicado a imagens log-polar segundo um algoritmo sequencial constituído pelos algoritmos de procura da recta- Ψ e da circunferência- Γ . Concluiremos então que os algoritmos em questão têm uma construção natural em geometria log-polar.

6.1 Método de Estimação

A abordagem que propomos é muito semelhante à descrita para coordenadas cartesianas, e tem a seguinte organização sequencial:

$$\boxed{F_r^*(\lambda, \gamma; k) \rightarrow \psi^k} \Rightarrow \min_{\hat{\psi}^k} \left\{ \min_{\hat{\omega}} \mathcal{M}(\mathbf{R}|\psi^k) \right\} \Rightarrow \hat{\psi} \Rightarrow$$

$$\hat{\psi} \Rightarrow \boxed{F_c^*(\lambda, \gamma; k) \rightarrow \Gamma^k} \Rightarrow \min_{\hat{\Gamma}^k} \left\{ \min_{\hat{\omega}} \mathcal{M}(\mathbf{R}|\Gamma^k) \right\} \Rightarrow \hat{\Gamma}$$

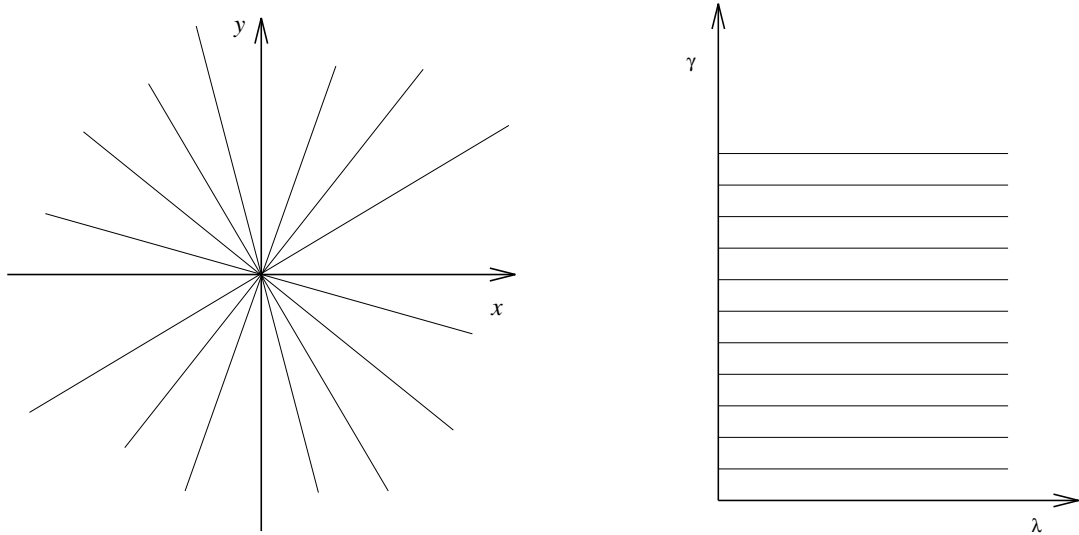
Como se pode observar (por comparação directa), a principal diferença deste método em relação ao caso cartesiano consiste nas variedades definidas no espaço log-polar (em vez de definidas no espaço cartesiano): estas variedades são dadas por $F_r^*(\lambda, \gamma; k) = 0$ e $F_c^*(\lambda, \gamma; k) = 0$, e correspondem, naturalmente, a uma transformação log-polar das variedades cartesianas correspondentes (ver Figura 6.3):

$$\begin{cases} F_r(\mathbf{x}; k) = x \sin k + y \sin k & \Rightarrow F_r^*(\lambda, \gamma; k) = \gamma + k \\ F_c(\mathbf{x}; k) = x^2 + y^2 - k(x \cos \Psi + y \sin \Psi) & \Rightarrow F_c^*(\lambda, \gamma; k) = \lambda - \log_l |\cos(\Psi - \gamma)| + k \end{cases} \quad (6.4)$$

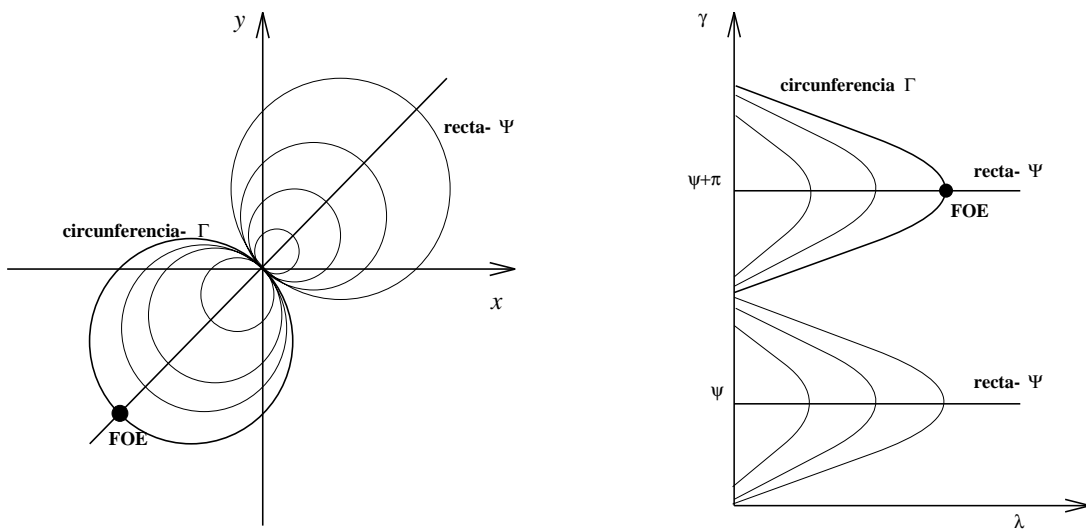
Não iremos, nesta secção, deduzir novamente a estrutura inerente aos algoritmos de procura da recta Ψ e da circunferência Γ , dado que essa estrutura é rigorosamente a mesma, relativamente ao caso cartesiano. O que se altera, de facto, são os subespaços de procura sobre os quais se aplicam esses algoritmos. Assim, a restrição em coordenadas log-polares descrita por $F_r^*(\lambda, \gamma; k) = 0$ implica uma procura em rectas horizontais da imagem log-polar (ver Figura 6.3b), enquanto a restrição definida por $F_c^*(\lambda, \gamma; k) = 0$ implica uma procura nas curvas igualmente apresentadas na Figura 6.3b.

Neste contexto, o algoritmo de procura da recta- Ψ realiza uma procura em rectas horizontais de forma a recuperar a recta- Ψ e estimar o valor de ω_3 e uma combinação linear entre ω_1 e ω_2 ; o segundo algoritmo executa a procura da circunferência- Γ^1 , no sentido de estimar os restantes parâmetros de velocidade.

¹Em coordenadas log-polares, a circunferência- Γ não corresponde naturalmente a uma circunferência



(a)



(b)

Fig. 6.3: Restrições nos planos cartesiano e log-polar. a) Transformação log-polar das rectas radiais. b) Transformação log-polar das circunferências Γ .

Na próxima secção mostraremos os resultados experimentais conseguidos a partir de algumas seqüências de imagens conhecidas, sobre as quais efectuámos uma transformação log-polar.

6.2 Resultados Experimentais

O algoritmo sequencial proposto foi testado para as seqüências 2 e 3, apresentadas anteriormente. Recordemos que, a Sequência 2 é sintética, com dimensão 256×256 pixels, distância focal $f = 302$ pixels. As imagens da seqüência real 3 têm a dimensão 322×244 pixels, sendo 229 pixels a distância focal. Aplica-se uma transformação log-polar sobre as imagens originais com os seguintes parâmetros: $r_{min} = 28.65$ pixels, o factor logarítmico $l = 1.0151$ e a dimensão da imagem log-polar é 180×100 . Note-se que a dimensão de uma imagem log-polar é aproximadamente 4 vezes menor que a dimensão da correspondente imagem cartesiana. As Figuras 6.4 e 6.5 permitem visualizar a transformação descrita.

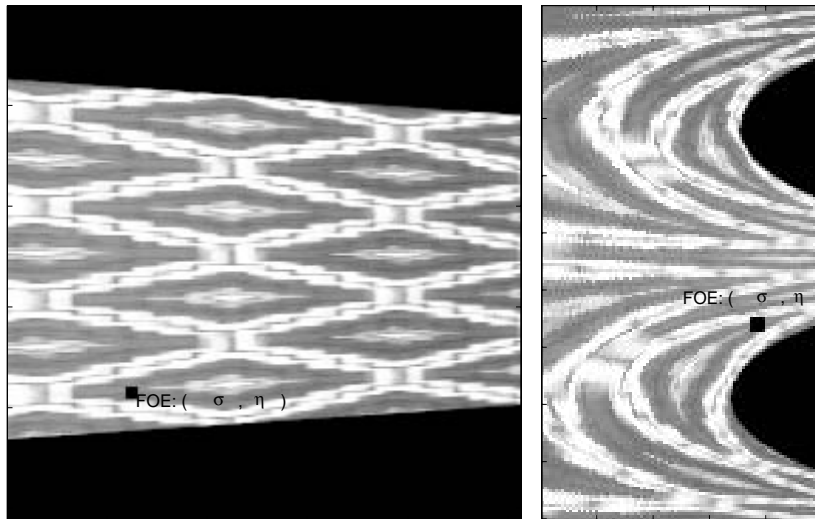


Fig. 6.4: Transformação log-polar duma imagem da Sequência 2. À esquerda, a imagem em coordenadas cartesianas e, à direita, a mesma imagem em coordenadas log-polares.

(como se mostra na Figura 6.3), mas por razões de simplicidade de notação mantemos o nome da sua transformada cartesiana.



Fig. 6.5: Transformação log-polar numa imagem da Sequência 3. À esquerda, a imagem em coordenadas cartesianas e, à direita, a mesma imagem em coordenadas log-polares.

Na sequência de imagens sintéticas, a câmara sofre uma rotação em torno do eixo Y de $-0.00785\text{rad/imagem}$, enquanto se move com um movimento de translação conhecido (ver capítulo anterior). Os algoritmos de procura da recta- Ψ e da circunferência- Γ foram usados para estimar os parâmetros de rotação apresentados na Tabela 6.1. A

	ω_1	ω_2	ω_3
Valores exactos	0	-0.00785	0
Estimativas	-0.001625	-0.008618	-0.000214

Tabela 6.1: Estimativas rotacionais da Sequência 2, usando uma representação log-polar.

Figura 6.6 mostra a localização exacta do Foco de Expansão na imagem log-polar, representando simultaneamente a recta- Ψ e a circunferência- Γ encontradas pelos algoritmos de procura respectivos. Confrontando estes resultados com os obtidos com imagens cartesianas, observa-se que os erros associados à estimativas são maiores devido ao efeito da amostragem não-uniforme imposta à transformação log-polar. No entanto, as estimativas são uma boa aproximação dos valores esperados.

Na segunda experiência, com imagens reais, a câmara move-se com translação pura

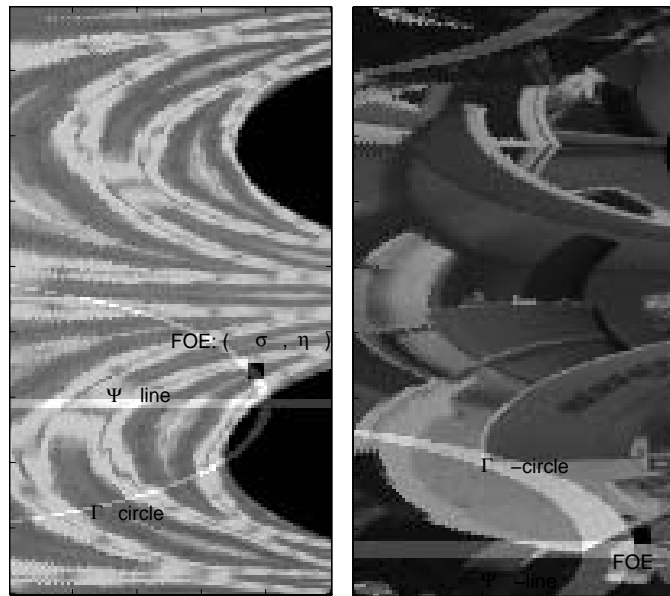


Fig. 6.6: Resultados da estimação para imagens log-polares, usando a Sequência 2 (à esquerda) e a Sequência 3 (à direita).

(ou, mais correctamente, com rotação aproximadamente nula). A aplicação dos algoritmos de procura da recta- Ψ e da circunferência- Γ dá-nos uma estimativa dos parâmetros de rotação: $(\omega_1, \omega_2, \omega_3) = (0.00075, 0.00023, 0.0019)$ rad/imagem, que correspondem a valores aproximadamente nulos, como era esperado. Da mesma forma, a recta Ψ e a circunferência Γ , encontrados pelo método de estimação proposto, intersectam-se numa localização próxima da localização exacta do Foco de Expansão, como mostra a Figura 6.6.

A partir das experiências efectuadas, constata-se que o tempo de computação foi drasticamente reduzido. Este facto deveu-se a duas razões fundamentais: o problema de estimação e os subespaços de procura foram subdivididos, diminuindo a complexidade algorítmica do próprio método; os algoritmos de procura da recta- Ψ e da circunferência- Γ foram aplicados a imagens log-polares. Apesar de existir, para determinadas aplicações, uma natural deterioração das estimativas, estas estimativas revelam-se, ainda, bastante satisfatórias e a sua deterioração pode não representar um obstáculo à aplicação do método — especialmente se o objectivo primordial de um dado sistema activo controlado por

visão consiste, não na estimação exacta dos parâmetros de velocidade, mas simplesmente, na obtenção de estimativas aproximadas (e rápidas) dos estados desses parâmetros (por exemplo, se as estimativas forem usadas na retroação de estado do sistema de controlo que lhe está associado).

7. Conclusões

Nesta tese, debruçámo-nos sobre um problema específico da visão robótica: a estimação do movimento próprio de um observador a partir da sua informação visual. A nossa abordagem centra-se num conjunto de pressupostos básicos que podem ser enumerados da seguinte forma:

1. O observador é monocular e as imagens são formadas segundo um modelo de projecção perspectiva.
2. As cenas observadas são compostas por objectos estáticos texturados.
3. O movimento do observador é arbitrário, com componentes de translação e rotação.
4. As observações são baseadas exclusivamente nas derivadas espaço-temporais da imagem.

Partindo destes pressupostos, desenvolvemos um método de estimação de movimento, com três características fundamentais:

- Apresentamos um paradigma de estimação de velocidade baseado na subdivisão do espaço das estimativas e das observações. O método é composto por um conjunto de algoritmos de baixa complexidade, que procuram uma solução ao longo de diversas famílias de subespaços. Cada família de subespaços de procura depende de restrições geométricas específicas do fluxo normal. O resultado da procura de cada algoritmo pode ser usado por outros algoritmos, definindo uma estrutura global que pode ser

paralela ou sequencial. A utilização de uma grande variedade de subespaços de procura aumenta a robustez de estimação do método¹.

- São aplicados estimadores que detectam e rejeitam *outliers* do espaço das observações, inspirados na teoria da regressão robusta.
- O método pode ser estendido para outros tipos de representação geométrica da imagem, particularmente a representação log-polar, para a qual alguns algoritmos de procura estão bem adaptados. Procede-se, assim, ao projecto conjunto de um processo visual (o estimador de movimento próprio) e da geometria do sensor correspondente (geometria log-polar).

Nesta Tese, apresentamos uma série de experiência com imagens sintéticas e reais, para vários tipos de movimento, que permitem ilustrar a robustez e eficácia do método proposto. Propomos avaliar esta abordagem à luz dos seguintes critérios:

Estabilidade: Um método de estimação deve ser pouco sensível ao erro nas observações.

No sentido de resolver este problema, aplicamos um método directo no cálculo das observações (adquirindo unicamente o fluxo normal) e introduzimos estimadores robustos para a selecção posterior dessas observações. Para além disso, o método torna-se mais robusto na recuperação de uma solução, se diversificarmos os subespaços de procura e reforçarmos o conjunto das observações, alargando o tempo de aquisição (ver Apêndice B).

Generalidade: O método proposto lida com casos de movimento arbitrário para cenas arbitrárias. Apresentamos no Apêndice B um estudo sobre as hipóteses simplificativas impostas no sistema, onde se propõem algumas soluções.

¹A subdivisão do espaço das observações depende das exigências impostas ao sistema: se desejamos robustez na estimação completa do movimento, devemos diversificar os subespaços de procura — então, o estimador torna-se lento; se desejamos ter uma ideia aproximada dos parâmetros de velocidade ou estimar unicamente uma dada restrição desses parâmetros, então o método pode-se restringir a um número limitado de subespaços de procura — os resultados desejados são obtidos mais rapidamente.

Complexidade: Neste trabalho, adoptamos um método não-iterativo de baixa complexidade. Sabendo que cada algoritmo de procura “visita” cada *pixel* da imagem uma só vez e a complexidade do estimador robusto que lhe está associado (ver Apêndice A) é de $O(N^2)$ (onde N é o número de observações), logo o algoritmo tem complexidade $O(\max\{n^2, N^2\})$, para uma imagem de dimensão $n \times n$.

Existem alguns tópicos que merecem a nossa atenção para trabalho futuro. Começamos pelo desenvolvimento de métodos de estimação mais sofisticados de forma a melhorar a robustez e o desempenho do sistema. Esse desenvolvimento passa por: (1) investigar estimadores robustos mais rápidos e que tomem em conta medidas estatísticas das observações; (2) estudar novas famílias de subespaços de procura, tendo em conta um número mais abrangente de propriedades geométricas do fluxo normal; (3) analisar uma generalização da abordagem por via de histogramas, o que pode representar, em determinados casos, um caminho alternativo no sentido da estimação de movimento.

Apontam-se também direcções de trabalho futuro, no estudo mais aprofundado dos casos degenerados ou patológicos, com múltipla solução, apresentados por Horn em [Horn, 1987]. Neste estudo, pretende-se desenvolver uma medida de “proximidade” dos casos com múltipla solução — medida essa que permitirá prever, com mais rigor, o nível de confiança das soluções detectadas. O mesmo estudo pode ser extendido para casos com solução instável, como é o exemplo (apresentado no Apêndice B) do caso em que o Foco de Expansão se localiza fora do plano de imagem.

No sentido das actuais linhas de investigação em Visão Robótica, propomos considerar no futuro um observador activo, introduzindo o processo de estimação num sistema autónomo de navegação. Por um lado, isto permite-nos incluir restrições adicionais no próprio processo de estimação, aumentando a estabilidade das soluções do estimador. Por outro lado, o sistema de controlo deve ser desenvolvido no sentido de assegurar integralmente um bom desempenho do sistema e razoáveis níveis de robustez na rejeição de incertezas e de ruído. A reduzida complexidade do método de estimação e a sua grande flexibilidade (dado que a concepção do estimador pode ser feita à medida das necessidades do projectista) irão aumentar os graus de liberdade do desenho do controlador.

Assim, a íntima relação entre a percepção do movimento próprio e a acção de um sistema autónomo representa na área da visão um dos temas futuros mais promissores no estudo da estimação de velocidade.

A. Estimação Robusta

The method of Least Squares is seen to be our best course when we have thrown overboard a certain portion of our data — a sort of sacrifice which has often to be made by those who sail upon the stormy seas of Probability.

Edgeworth in 1887

A estimação estatística é uma ferramenta essencial em visão robótica. O método dos mínimos quadrados, *Least Squares* (LS) na nomenclatura inglesa, é um dos métodos tradicionalmente usados na estimação estatística e domina quase por completo as técnicas de estimação em visão robótica. O sucesso do estimador LS tem dois motivos: (1) o método traz claras vantagens computacionais, dado que as estimativas são calculadas explicitamente a partir das observações; (2) quando a distribuição do erro nas observações é gaussiano, o método LS é ótimo.

O estimador LS tem por base uma sólida teoria matemática com quase dois séculos de existência. No entanto, mais recentemente, alguns investigadores têm constatado que as observações e os dados reais, tais como os que se nos oferece em visão robótica, raramente obedecem aos pressupostos de optimalidade do método LS, ou seja, raramente o ruído nas observações e nos dados é puramente gaussiano. Na realidade, é facilmente demonstrável que uma única observação marginal (*outlier* em inglês) pode “destruir” completamente a qualidade das estimativas.

Os *outliers* são fenómenos frequentes em visão robótica. Um *outlier* pode provir de pontos de imagem espúrios, gradientes de baixa amplitude, artefactos de iluminação,

etc. Assim sendo, nos últimos anos, tem-se procurado métodos que anulem eficazmente a influência de um *outlier* na estimação de parâmetros. Rousseeuw oferece-nos um estudo minucioso sobre esta matéria, no seu livro “Robust Regression & Outlier Detection” [Rousseeuw and Leroy, 1987], para o qual remetemos os principais aspectos teóricos de estimação abordados neste Apêndice.

Começemos com um exemplo específico que ilustra e compara os métodos de estimação usados neste trabalho. Para tal, usaremos o problema de regressão associado ao *algoritmo de procura da recta- Ψ* (ver equação (4.4)), que pode ser expresso da seguinte forma:

$$b_i = a_i\theta_1 + \theta_2 + \epsilon_i \quad (\text{A.1})$$

onde $a_i = -r_i$, $\theta_1 = \omega_3$, $\theta_2 = f(\omega_1 \cos \Psi + \omega_2 \sin \Psi)$, b_i é a observação da amplitude do vector de fluxo e ϵ_i é o erro associado a essa observação. O problema consiste em estimar os parâmetros θ_1 e θ_2 ; θ_1 corresponde ao declive duma recta no espaço $\{a, b\}$; θ_2 corresponde à ordenada da intersecção da recta com o eixo vertical. A Figura A.1 mostra 7 medidas do fluxo normal tirados ao longo da recta- Ψ^1 , onde os pontos (a_2, b_2) e (a_5, b_5) correspondem a *outliers* nas direcções a e b respectivamente.

A Figura A.2 ilustra bem a diferença entre (1) o resultado da aplicação do estimador LS usando todos os pontos disponíveis e (2) o resultado do estimador LS usando o conjunto de pontos após a rejeição dos *outliers* (a_2, b_2) e (a_5, b_5) (para diferenciarmos as duas situações, o último estimador é designado por LS*). Observa-se que a diferença entre as estimativas é significativa. Demonstra-se que mesmo um número reduzido de *outliers* nos dados ou nas observações pode alterar significativamente o valor das estimativas calculadas pelo método dos mínimos quadrados.

Existem vários estimadores robustos que permitem detectar e remover *outliers*. Em [Rousseeuw and Leroy, 1987] podemos encontrar uma extensa introdução a métodos de estimação robusta. Neste trabalho iremos adoptar dois métodos de estimação robusta.

¹Neste exemplo, a recta- Ψ tem uma direcção horizontal por motivos de simplificação na representação das observações. Para qualquer outra direcção, o exemplo continuaria válido.

O **primeiro método** consiste no estimador da mínima mediana dos quadrados, *Least Median of Squares* (LMS) em inglês, e baseia-se no seguinte problema de minimização:

$$\min_{\hat{\boldsymbol{\theta}}} \text{med}_i R_i^2 \quad (\text{A.2})$$

onde o resíduo $R_i = b_i - a_i\theta_1 - \theta_2$ e $\boldsymbol{\theta} = [\theta_1 \ \theta_2]^T$. O estimador LMS pode tolerar até 50% de *outliers* no conjunto dos dados. Esta vantagem do estimador na detecção e rejeição de *outliers* é contraposta à desvantagem de ser um método iterativo de convergência lenta — exceptua-se o caso de estimação unidimensional ou bidimensional, no qual é possível aplicar um algoritmo alternativo de estimação LMS não iterativo.

O **segundo método** proposto consiste no estimador dos mínimos quadrados com rejeição de *outliers*, (*Least Trimmed Squares*, LTS, na nomenclatura inglesa):

$$\min_{\hat{\boldsymbol{\theta}}} \sum_{i=1}^h [R^2]_{i:N} \quad (\text{A.3})$$

onde os quadrados dos resíduos são ordenados de forma a que $[R^2]_{1:N} \leq \dots \leq [R^2]_{N:N}$; e h é um inteiro tal que $h \leq N$. O estimador LTS equivale ao estimador LS após a rejeição dos termos com maior valor residual.

O estimador LTS tem três características fundamentais:

- possui um mecanismo de rejeição de *outliers*;
- pode ser implementado através de um algoritmo de convergência rápida;
- tal como o estimador LS, é capaz de rejeitar ruído nas observações com distribuição gaussiana.

Na Figura A.2 compara-se a *performance* destes estimadores. Observa-se que os estimadores LMS e LTS rejeitam de facto a influência dos dois *outliers*, aproximando as estimativas ao seu valor real. Na próxima secção, propõe-se um algoritmo não iterativo para um problema de estimação bidimensional (regressão simples) que combina os estimadores LMS e LTS.

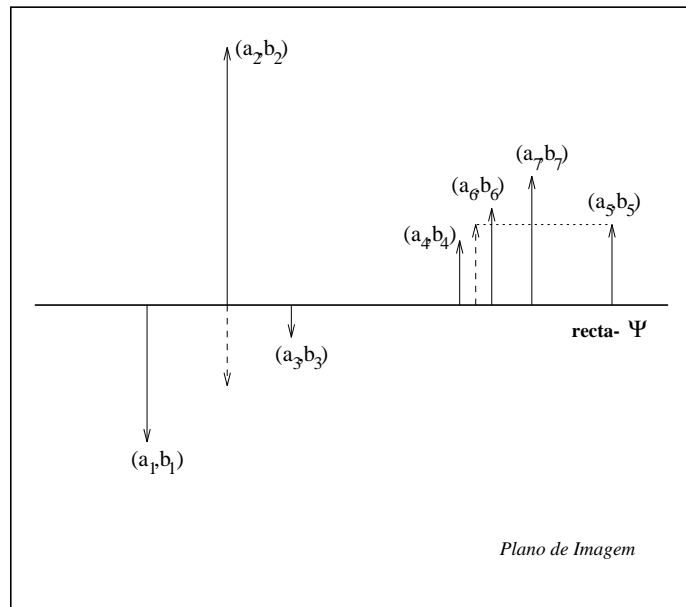


Fig. A.1: Observações medidas ao longo da recta- Ψ , com dois *outliers*.

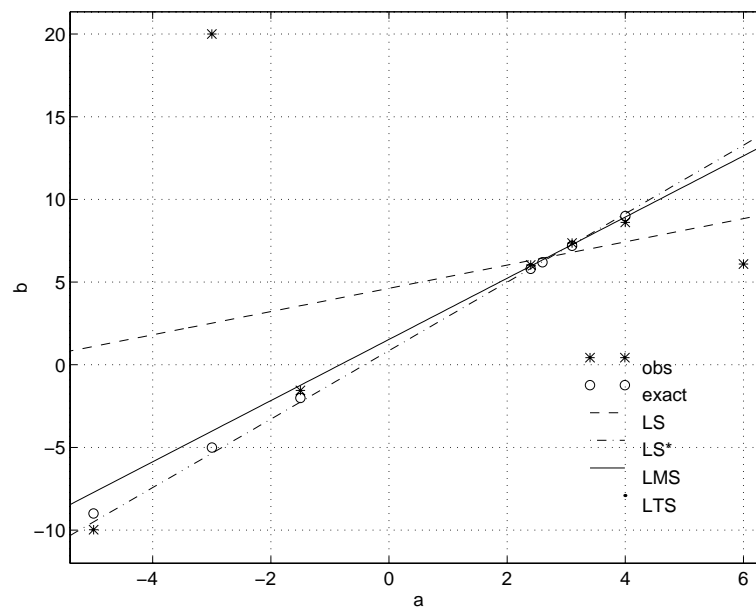


Fig. A.2: Aplicação dos estimadores LS (*least squares*), LS* (*least squares* após rejeição de *outliers*), LMS (*least median squares*) e LTS (*least trimmed squares*).

A.1 Estimação LTS bidimensional

Neste trabalho, propomos um método LTS para resolver um problema de regressão simples. O estimador é aplicado em duas fases :

- Detecção e rejeição dos *outliers*.
- Aplicação do estimador LS aos restantes dados e observações.

A detecção dos *outliers* é conseguido através do estimador LMS para regressão simples. Em regressão simples, pretende-se estimar os dois parâmetros θ_1 e θ_2 numa recta, usando N pontos do espaço. Neste caso particular, é possível aplicar um estimador LMS não iterativo, de complexidade $O(N^2)$, proposto por Rousseeuw em [Rousseeuw and Leroy, 1987]. Assim, após uma primeira estimativa de θ , é possível calcular o valor dos resíduos R_i .

Neste momento temos um conjunto de resíduos R_i para um dado $\hat{\theta}$ calculado através de um estimador LMS. Os *outliers*, no caso de existirem, corresponderão aos pontos com resíduo mais elevado. Se quisermos rejeitar os *outliers*, teremos que estabelecer um critério de rejeição, isto é, introduzir um limiar robusto acima do qual o resíduo é considerado *outlier*. Esse limiar deverá depender da mediana dos valores residuais. Seja

$$\hat{\sigma} = C \sqrt{\text{med}_i R_i^2}$$

onde C é um factor de correcção de natureza empírica que garante a consistência dos dados perante uma distribuição gaussiana do erro². O ponto (a_i, b_i) é considerado *outlier* quando o resíduo R_i ultrapassa o limiar³ $2.5\hat{\sigma}$.

Após a rejeição dos *outliers*, aplica-se o segundo passo do algoritmo que resulta numa nova estimativa $\hat{\theta}$. Neste caso, a aplicação do estimador LS aos pontos não rejeitados implica a anulação do ruído gaussiano nesses pontos.

²valor de C calculado segundo uma fórmula apresentada em [Rousseeuw and Leroy, 1987]:

$$C = 1.4826 \times \left(1 + \frac{5}{N-2}\right).$$

³Este limiar é atribuído tendo em conta que, nas condições de ruído gaussiano, existem “poucos” resíduos acima de $2.5\hat{\sigma}$.

Em resumo, o estimador proposto nesta secção é um estimador LTS, com as seguintes características: é composto por uma única iteração; a inicialização é feita com o estimador LMS; o valor h depende dos valores residuais produzidos pelo estimador LMS.

Na Figura A.2 podemos comparar a estimativa $\hat{\theta}$ usando o estimador LTS proposto nesta secção e o estimador LMS. Podemos observar igualmente na Tabela A.1 que o estimador LTS tem o melhor desempenho entre os três estimadores referidos anteriormente (LS, LMS, LTS).

Parâmetro	Valor exacto	LS	LS*	LMS	LTS
θ_1	2	0.7024	2.071	1.850	1.882
θ_2	1	4.6223	0.853	1.531	1.350

Tabela A.1: Estimação de θ_1 e θ_2 a partir do conjunto de sete pontos observados no exemplo, aplicando três estimadores. Note-se que o estimador LS* foi aplicado exclusivamente aos “melhores” cinco pontos, ilustrando o efeito de contaminação dos *outliers* no estimador LS.

A.2 Aplicação do estimador LTS aos algoritmos de estimação

Os algoritmos de procura de baixa complexidade para estimação dos parâmetros de velocidade da câmara limitam significativamente o espaço das observações. Os algoritmos apresentados são essencialmente três:

- o algoritmo de procura da recta- Ψ ;
- o algoritmo de procura da recta- Φ ;
- o algoritmo de procura da circunferência- Γ .

Qualquer um dos três algoritmos de procura requerem uma métrica robusta L para a medida dos resíduos nas observações dum subespaço S : $\mathcal{M}_L(\mathbf{R}|S)$. No entanto, um esti-

mador robusto só é computacionalmente viável se não exigir uma implementação iterativa, isto é, no caso do problema de estimação ser unidimensional ou bidimensional⁴.

Em seguida, apresentamos os estimadores aplicados a cada um dos algoritmos anunciados:

- Algoritmo de procura da recta- Ψ :
 - estimação de 2 parâmetros: ω_3 , projecção de (ω_1, ω_2) sobre a recta Ψ — **estimador LTS**
- Algoritmo de procura da recta- Φ :
 - *inserido num Algoritmo paralelo*: estimação dos três parâmetros de rotação (problema de dimensão tridimensional) — **estimador LS**
 - *após a aplicação do Algoritmo de procura da recta- Ψ* : estimação de uma restrição sobre os parâmetros de rotação — **estimador LTS**
- Algoritmo de procura da circunferência- Γ :
 - *após a aplicação do Algoritmo de procura da recta- Ψ* : estimação de (ω_1, ω_2) — **estimador LTS**

Observa-se aqui a principal vantagem de aplicar o algoritmo de procura da recta- Ψ como primeiro passo para uma estimação sequencial dos parâmetros de velocidade da câmara: dado ser um problema de estimação bidimensional, possibilita a aplicação de um estimador robusto LTS. O segundo passo do algoritmo de estimação de velocidade continua a ser caracterizado por problemas de estimação unidimensional ou bidimensional.

Observa-se igualmente que o Algoritmo de procura da recta- Φ , quando aplicado sem informação prévia quanto à velocidade, está associado a um estimador LS, que, para alguns casos, é um mau estimador. Esta característica deficitária é em parte compensada pela natureza redundante do algoritmo paralelo inerente ao método, nomeadamente quando

⁴Excluimos deste estudo o algoritmo baseado em histogramas, dado que, para além de ter em conta o total das observações da imagem, faz uso de um estimador robusto (o mínimo da moda dos resíduos).

se garante a procura de um conjunto suficientemente elevado de rectas- Φ com direcções distintas.

B. Hipóteses simplificativas impostas no sistema

Whenever we perform computations, we design models on the basis of assumptions, which in the case of visual processing are constraints on the space-time in which the system is acting, on the system itself, and on their relationship.

Yiannis Aloimonos in [Fermuller and Aloimonos, 1995]

Neste Apêndice, propomos discutir algumas das hipóteses simplificativas impostas no sistema de estimação proposto nesta tese. Essas hipóteses podem existir a vários níveis: ao nível das observações, ao nível do movimento e da cena visualizada ou ao nível dos algoritmos de estimação propriamente ditos. Temos, portanto, um sistema de estimação, na qual vamos encontrar os seguintes tipos de problemas (que, por motivos de simplificação, viémos a desprezar no processo de construção do método):

- **Ao nível das observações:**
 - Ruído no plano da imagem.
 - Hipótese do brilho constante.
- **Ao nível do movimento e da cena visualizada:**
 - O caso em que a velocidade ao longo do eixo óptico é nulo ($W = 0$).
 - O caso em que a translação é nula.

- O caso em que a cena é planar.

- **Ao nível dos algoritmos de estimação:**

- Distribuição das observações nos subespaços de procura.

- Unicidade das soluções dos algoritmos de procura.

Nas próximas secções, iremos explorar o impacto destes problemas na realização de um sistema de estimação e propor algumas soluções.

B.1 Hipóteses ao nível das observações

Ruído no plano da imagem

O ruído na intensidade luminosa num determinado pixel do plano de imagem advém (fundamentalmente) do ruído na projecção de um feixe luminoso proveniente de várias fontes de luz, influenciado pelo conjunto de vários factores de natureza aleatória. Pelo teorema do limite central, pode-se considerar a distribuição do ruído aproximadamente gaussiana.

A assumpção duma distribuição deste tipo para o erro nas observações tem gerado alguma discussão, por duas razões fundamentais: por um lado, a hipótese é aproximada, e por outro, é difícil de caracterizar. À partida, a média e a variância da distribuição gaussiana em questão não é calculável, dado o desconhecimento prévio da própria cena projectada no plano de imagem. De facto, as propostas mais frequentes neste domínio consistem em assumir que a média da distribuição é nula em cada ponto e que o ruído entre os pontos da imagem é independente.

É, contudo, importante sublinhar que a hipótese de que *o ruído em cada ponto é gaussiano, independente e de média nula* pode, em alguns casos, estar longe da realidade. Mas perante a falta de modelos mais consistentes (e, principalmente, tão fáceis de manipular estatisticamente), pareceu-nos razoável adoptar o modelo proposto. Em consequência, os

métodos de estimação apresentados nesta tese lidam com este modelo de ruído, dado que têm por base um estimador LS, que, para ruído gaussiano, conduz a estimativas óptimas.

Hipótese do brilho constante

Nesta secção iremos mostrar que a hipótese do brilho constante para a definição de fluxo normal é válida sob certas condições rígidas do sistema [Horn, 1986, Faugeras, 1993].

Recordemos a expressão que relaciona a variação do brilho $\frac{dE}{dt}$ com as coordenadas do fluxo:

$$\frac{\partial E}{\partial x}u + \frac{\partial E}{\partial y}v + \frac{\partial E}{\partial t} = \frac{dE}{dt} \quad (\text{B.1})$$

Esta relação não envolve qualquer tipo de aproximação. No entanto, o termo $\frac{dE}{dt}$ não pode ser calculado dispondo apenas duma sequência de imagens.

Para o fazer, precisaríamos de introduzir adicionalmente conhecimento quanto à iluminação e aos modelos de reflectância da cena projectada na imagem. À partida, esses modelos são desconhecidos. Espera-se então que o brilho em cada ponto não se altere significativamente com o movimento do observador, recorrendo-se, assim, à hipótese do brilho constante, $\frac{dE}{dt} = 0$. Encontramos uma discussão mais pormenorizada deste assunto em [Horn and Weldon, 1988].

Define-se vector do fluxo normal \mathbf{v}_n num ponto da imagem como a projecção do fluxo óptico segundo o gradiente ∇E calculado nesse ponto. Partindo da equação (B.1), \mathbf{v}_n calcula-se do seguinte modo:

$$\mathbf{v}_n = \frac{\frac{dE}{dt} - \frac{\partial E}{\partial t}}{\|\nabla E\|} \frac{\nabla E}{\|\nabla E\|},$$

onde $\frac{\nabla E}{\|\nabla E\|}$ define a direcção do vector \mathbf{v}_n . Note-se, no entanto, que a definição de fluxo normal proposta nesta tese parte da hipótese do brilho constante e é calculada da seguinte forma:

$$\mathbf{v}_n = \frac{-\frac{\partial E}{\partial t}}{\|\nabla E\|} \frac{\nabla E}{\|\nabla E\|}$$

O erro de observação do fluxo normal existe ao nível da amplitude do vector que lhe está associado e pode-se considerar nulo para certos modelos de reflectância ideais em

ambientes estáticos. É o caso duma superfície lambertiana. Como se sabe, a intensidade luminosa reflectida num ponto duma superfície lambertiana, projectada no plano da imagem, mantém-se, independentemente da posição da câmara em relação ao ponto observado. Logo, se o observador estiver em movimento e o ponto observado for estático, $\frac{dE}{dt}$ é naturalmente nulo. No caso do objecto ter movimento (nomeadamente movimento de rotação) prova-se em [Faugeras, 1993] que a aproximação deixa de ser válida.

Para modelos de reflectância onde a direcção de observação influencia a intensidade luminosa projectada (que é o que acontece, por exemplo, em relação a uma superfície especular), então a aproximação $\frac{dE}{dt} = 0$ deixa de ser válida.

É importante referir que considerar o termo $\frac{dE}{dt}$ na expressão do fluxo normal é uma tarefa difícil, dado o desconhecimento prévio quer da própria estrutura visualizada quer da posição relativa do observador em cada instante. No entanto, este termo pode ser considerado ruído na observação do fluxo normal, que, por motivos de simplificação, é aproximadamente gaussiano com média nula — basta considerar que existe na cena visualizada uma gama variada de superfícies com diversos modelos de reflectância, iluminadas por diferentes fontes de luz. Neste caso, e como se viu anteriormente, um estimador LS é óptimo para uma distribuição de erro gaussiano.

Quando um ponto de superfície diverge consideravelmente do modelo gaussiano de média nula (ou seja, o valor esperado de $\frac{dE}{dt}$ é consideravelmente diferente de zero numa superfície com um dado modelo de reflectância), então um estimador robusto pode eliminar o ruído que lhe está associado, considerando-o *outlier* no conjunto das observações.

B.2 Hipóteses ao nível do movimento e da cena visualizada

Caso $W = 0$

No caso de não haver translação ao longo do eixo óptico (com translação global da câmara não nula), a transformação efectuada em (2.3) não pode ser efectuada e o Foco de Ex-

pansão encontra-se no infinito. Assim, o máximo que se pode retirar da componente translacional da câmara é a direcção do vector (U, V) , ou seja, a direcção da recta de suporte do Foco de Expansão. Neste caso, todas as rectas paralelas à recta- Ψ intersectam o Foco de Expansão no infinito.

Deste modo, as condições de validade dos algoritmos de procura alteram-se para o caso em que $W = 0$. De facto, ficamos restringidos a uma situação típica de procura da recta- Ψ , pelo que um algoritmo de procura em rectas radiais parece o mais indicado, neste caso específico. Este facto parece evidenciar que um algoritmo de procura ao longo de rectas paralelas (o algoritmo de procura da recta- Φ) só funciona razoavelmente quando o Foco de Expansão se encontra dentro duma vizinhança junto à origem do plano de imagem (vizinhança essa que pode ou não estar confinada aos limites da própria imagem).

Mais genericamente, quando o Foco de Expansão tende para infinito numa dada direcção, passa a ter sentido falar-se de estimar essa direcção e não a sua localização exacta. Como se sabe, a componente translacional do fluxo é função da distância entre o pixel considerado, dado por \mathbf{x} , e o Foco de Expansão. Assim, um efeito imediato de o Foco de Expansão se afastar da zona limitada da imagem consiste em diminuir a variação da componente translacional do fluxo entre pixels diferentes. No caso limite em que $W = 0$, então esta variação é nula.

Conclui-se assim que o algoritmo de procura da recta- Ψ é um bom algoritmo de detecção da direcção do vector Σ , mesmo para o caso em que $W = 0$. Os algoritmos consequentes (correspondentes ao algoritmo sequencial apresentado em 4.3) dão uma indicação sobre a localização do Foco de Expansão, mesmo quando se encontra fora da região da imagem — note-se que, neste caso, é difícil determinar com precisão a sua localização exacta, já que, como se mostrou anteriormente, a estimação do Foco de Expansão, com as observações disponíveis, pode ser inconclusiva.

Caso translação nula

Considere-se o caso em que a componente de translação no fluxo é pequena ou nula, o que pode acontecer quando a translação da câmara é baixa ou quando a cena observada está a uma grande distância da imagem (Z médio elevado). Em qualquer dos casos, a componente rotacional é dominante sobre a componente translacional, e a aplicação de um algoritmo que procure o ponto $\hat{\Sigma}$ que minimize o resíduo devido à componente de translação,

$$\min_{\hat{\Sigma}} \left\{ \min_{\hat{\omega}} \{ \mathcal{M}(\mathbf{R}|\mathbf{S}) \} \right\} \quad (\text{B.2})$$

irá forçosamente encontrar diversos mínimos locais sem significado, provocados, essencialmente, pelo ruído existente nas observações. No caso limite em que a translação é nula, o resíduo é idealmente zero qualquer que seja o par de procura $(\hat{\Sigma}, \mathbf{S})$ escolhido, ou seja, não existe um mínimo residual bem definido no conjunto das estimativas.

Este caso pode ser ultrapassado da seguinte forma: primeiro, é preciso notar que as equações do fluxo são lineares (aparte um termo de ruído) em relação à velocidade angular, pelo que o algoritmo de estimação aplicado à imagem estima robustamente os parâmetros de rotação (para qualquer estimativa de Σ). Assim, uma vez recuperada a rotação da câmara, testa-se então se existe ou não translação da câmara: recalculam-se os vectores de fluxo normal a partir do valor de rotação (supondo que o movimento é de rotação pura); determina-se o termo residual resultante da diferença entre o fluxo estimado e observado; e, se esse termo for significativamente reduzido, então considera-se que a translação é nula.

Caso cena planar

Uma área importante de investigação em análise de movimento consiste em verificar se um dado campo de fluxo óptico pode ser gerado a partir de dois movimentos de câmara diferentes. Este facto depende, naturalmente, da superfície observada. Como veremos nesta secção, um dos casos mais simples (a superfície planar) corresponde exactamente a um problema com duas soluções possíveis. Horn [Horn, 1987] demonstra que só um hiperbolóide

de uma folha (ou qualquer forma degenerada derivada de um hiperbolóide), observado segundo um ponto de vista específico, pode produzir um campo de fluxo ambíguo. No entanto, Horn faz notar que este tipo de superfície é um subconjunto de medida nula no espaço de todas as superfícies suaves e de todos os pontos de vista, pelo que **a probabilidade de uma determinada cena corresponder a uma superfície com solução ambígua é nula.**

A título de exemplo, iremos referir o caso de uma superfície planar. Suponha que o declive do plano é caracterizado pelo vector (p, q) , através da relação $Z = Z_o(1 - px - qy)^{-1}$, e o movimento da câmara é dado pelos vectores \mathbf{t} e $\boldsymbol{\omega}$. Prova-se [Waxman et al., 1987] que existe uma segunda solução $(p', q', \mathbf{t}', \boldsymbol{\omega}')$ que produz o mesmo fluxo óptico, e que pode ser escrita explicitamente a partir da primeira solução, da seguinte forma:

$$\begin{aligned}
 p' &= -U/W, \\
 q' &= -V/W, \\
 U' &= -pW, \\
 V' &= -qW, \\
 W' &= W; \\
 \omega'_1 &= \omega_1 - (V + qW), \\
 \omega'_2 &= \omega_2 + (U + pW), \\
 \omega'_3 &= \omega_3 - (pV - qU)
 \end{aligned} \tag{B.3}$$

onde a distância focal f e o parâmetro Z_o são igualados a 1 por razões de simplificação algébrica.

O mecanismo mais simples de detectar a solução exacta do movimento da câmara é garantir a continuidade dessa solução ao longo do tempo. Por outras palavras, a desambiguidade é resolvida porque só uma solução representa uma interpretação consistente no tempo. Nesse sentido, introduzimos nos algoritmos de estimação o conjunto das observações recolhidas ao longo de um intervalo de tempo pré-definido — supondo que as

condições de movimento da câmara não se alteram nesse intervalo. Deste modo, reforça-se a robustez do método (contando com mais observações) e evitam-se os casos com solução ambígua como o da cena planar.

B.3 Hipóteses ao nível dos algoritmos de estimação

Distribuição das observações nos subespaços de procura

Recorde-se o método de estimação do movimento usando a restrição do fluxo normal, apresentada no Capítulo 2. Como se viu, o método baseia-se na selecção de um subconjunto de vectores de fluxo normal com determinadas propriedades: selecciona-se o fluxo normal que é perpendicular às rectas que radiam de um ponto candidato a Σ .

Comecemos por verificar que, para um dado Σ , é muito difícil encontrar, no conjunto das observações disponíveis em toda a imagem, um vector de fluxo normal $\mathbf{n}(\mathbf{x})$ que seja exactamente perpendicular à recta que une \mathbf{x} a Σ , ou seja, encontrar um vector unitário $\mathbf{n}(\mathbf{x})$ tal que $\mathbf{n}(\mathbf{x}) \cdot (\mathbf{x} - \Sigma) = 0$.

Mais rigorosamente, o problema pode ser exposto nos seguintes passos:

- O conjunto das observações \mathbf{B} é finito e contável, dado que as observações só podem existir num determinado domínio discreto e finito, representado pela grelha do sensor visual da câmara, cujas unidades elementares são conhecidas por *pixels*.
- Assim, o conjunto \mathbf{B} pode ser posto como um vector de N variáveis aleatórias $[\theta_{\mathbf{x}_1} \cdots \theta_{\mathbf{x}_N}]$, onde $\theta_{\mathbf{x}_i}$ representa o ângulo do vector unitário $\mathbf{n}(\mathbf{x}_i)$. Considere-se que estas variáveis aleatórias são independentes e com distribuição uniforme entre -180 e 180 graus¹.

¹Note-se, no entanto, que o domínio das orientações de um vector de fluxo está, na realidade, limitado a um domínio contável, já que resulta duma sequência finita de operações efectuadas sobre um conjunto limitado de níveis de luminosidade quantificados. Contudo este domínio é extremamente denso e difícil de determinar, pelo que se considera que os eventos se distribuem uniformemente ao longo de um espaço contínuo.

- A cada ponto \mathbf{x}_i associa-se dois vectores unitários $\mathbf{n}(\mathbf{x}_i)$ e $\mathbf{n}^*(\mathbf{x}_i)$ (com uma diferença angular de 180 graus) que são solução da equação $\mathbf{n} \cdot (\mathbf{x}_i - \Sigma) = 0$. Os dois vectores encontrados correspondem a dois ângulos desejados $\theta_d(\mathbf{x}_i)$ e $\theta_d^*(\mathbf{x}_i)$.
- No espaço da variável aleatória $\theta_{\mathbf{x}_i}$, os pontos $\theta_d(\mathbf{x}_i)$ e $\theta_d^*(\mathbf{x}_i)$ têm medida de Lebesgue nula.
- Assim, a menos de um conjunto de medida Lebesgue nula, nunca se observa um vector de fluxo $\mathbf{n}(\mathbf{x})$ tal que $\mathbf{n}(\mathbf{x}) \cdot (\mathbf{x} - \Sigma) = 0$. Recordando a definição de conjunto $\mathcal{U}(\Sigma)$, conclui-se que a probabilidade do conjunto $\mathcal{U}(\Sigma)$ ser um conjunto vazio é exactamente um.

Este resultado obriga-nos a introduzir uma tolerância relativamente aos ângulos desejados, no espaço das variáveis aleatórias $\theta_{\mathbf{x}_i}$. Isto significa que uma observação passa a ser válida quando é perpendicular à recta $\overline{\mathbf{x} \Sigma}$ a menos de um ângulo de tolerância ϵ . Neste caso, para um parâmetro de tolerância $\epsilon > 0$, a probabilidade de existir uma observação válida é diferente de zero. No caso da distribuição de cada observação ser uniforme, essa probabilidade é igual a $4\epsilon/180$.

Temos um problema semelhante quando o subespaço de procura se limita a uma curva contida no plano (subespaço de dimensão 1). Até aqui, considerámos que as observações podiam ser recolhidas em qualquer ponto do plano de imagem. No caso do subespaço ser uma curva, deve-se introduzir uma tolerância, não só ao nível das orientações dos vectores de fluxo, mas também ao nível do espaço \mathbf{x} de procura. Assim, o subespaço de procura deixa de ser uma simples curva, para passar a ser uma região da imagem que acompanha o desenho da curva original — como mostram as figuras correspondentes aos resultados experimentais do Capítulo 5.

Unicidade das soluções dos algoritmos de procura

Como vimos anteriormente, algumas superfícies produzem um campo de fluxo com solução de velocidade ambígua. Nesses casos, não é possível garantir a unicidade da solução, a

menos que se integrem as observações obtidas ao longo do tempo. É uma consequência natural deste resultado que, para os algoritmos propostos nesta tese, não se garante igualmente a unicidade das soluções perante determinadas configurações da cena observada. De facto, essas configurações derivam directamente das superfícies referidas anteriormente, que, como se viu, estão restringidas à família das hiperbolóides de uma folha e suas formas degeneradas.

Recordemos que a lógica subjacente aos algoritmos de estimação propostos consiste em procurar um subespaço que anule uma componente não linear das observações. Isto significa que devemos garantir que isso só acontece para o subespaço desejado S_d (ou seja, para o subespaço que contém o Foco de Expansão). Se existe um subespaço de procura S_o , diferente do desejado, para o qual a componente não linear é nula, então teremos duas soluções para o problema de estimação. O subespaço S_o é a projecção de uma linha do espaço com determinadas características.

A título de exemplo, propomos fazer uma análise simples quanto à unicidade de solução relativamente ao algoritmo de procura da recta- Ψ . Assim, reescrevemos a equação (4.3)

$$\begin{aligned} \forall(x, y) \quad &: \quad (x, y) = r(\cos \psi, \sin \psi) \\ U_\psi(r) \quad &= \quad -\rho(x, y)(\sigma, \eta) \cdot (-\sin \psi, \cos \psi) + f(\omega_1 \cos \psi + \omega_2 \sin \psi) - \omega_3 r \end{aligned} \quad (\text{B.4})$$

Suponha que no subespaço S_o definido pela recta radial com ângulo $\psi_o \neq \Psi$ (onde Ψ é naturalmente a orientação exacta da recta- Ψ), a componente

$$\delta(r) = -\rho(x, y)(\sigma, \eta) \cdot (-\sin \psi_o, \cos \psi_o)$$

é afim em r , ou seja, o subespaço S_o é solução do algoritmo de procura. O termo $(\sigma, \eta) \cdot (-\sin \psi_o, \cos \psi_o)$ é constante para qualquer ponto do subespaço S_o , pelo que o problema reside precisamente na função $\rho(x, y)$. Queremos saber, então, qual deve ser a forma da função $\rho(x, y) = W/Z(x, y)$ para que $\delta(r)$ seja afim em r .

Suponha que a linha do espaço tridimensional obedece à lei

$$W/Z = ar + b,$$

onde a e b são parâmetros duma função afim (com $ar + b > 0$ para todo o plano da imagem). Fazendo uso das relações $x = fX/Z$ e $y = fY/Z$ (expressões usuais que traduzem a projecção de cada ponto tridimensional num ponto da imagem), obtemos a expressão de Z da linha tridimensional medida no subespaço S_o :

$$Z = \frac{W}{b} - \frac{a}{b}fR,$$

onde $R = \sqrt{X^2 + Y^2}$.

Conclui-se assim que a solução do algoritmo de procura da recta- Ψ não é única sempre que, para uma recta radial com orientação diferente de Ψ , a equação de Z medido nesta recta é afim em R . Logo, o algoritmo de procura pode conduzir não só a uma, mas a múltiplas soluções de estimação.

No entanto, há que notar o seguinte:

- a linha do espaço que dá origem a um subespaço do tipo S_o está contida numa família de superfícies especiais, que muito raramente se verificam;
- se os parâmetros de velocidade se mantiverem num intervalo de tempo, o comportamento de Z nesse subespaço deve alterar-se nesse período (devido à mudança de ponto de vista do observador); logo, a ambiguidade pode ser levantada se o algoritmo tiver em conta as observações recolhidas ao longo desse intervalo de tempo.

O problema da não unicidade de solução raramente é um entrave à correcta estimação de velocidade, através do método proposto nesta tese. Deste modo, assume-se que a solução dos algoritmos de procura utilizados é única, a menos de um conjunto de casos que raramente se observam, e que, por isso mesmo, não são directamente considerados nos respectivos algoritmos.

Bibliografia

- [Aloimonos and Duric, 1994] Aloimonos, Y. and Duric, Z. (1994). Estimating the heading direction using normal flow. *Int. Journal of Computer Vision*, 13(1):33–56.
- [Aloimonos et al., 1988] Aloimonos, Y., Weiss, I., and Banddophaday, A. (1988). Active vision. *Int. Journal of Computer Vision*, 1(4):333–356.
- [Bajcsy, 1985] Bajcsy, R. (1985). Active perception vs. passive perception. In *Proc. IEEE Workshop on Computer Vision: Representation and Control*, pages 55–62, Bellaire MI.
- [Ballard and Brown, 1982] Ballard, D. H. and Brown, C. M. (1982). *Computer vision*.
- [Ballard and Kimball, 1983] Ballard, D. H. and Kimball, O. A. (1983). Rigid body motion from depth and optical flow. *Comput. Vision Graphics Image Processing*, 22(1).
- [Bernardino and Santos-Victor, 1996] Bernardino, A. and Santos-Victor, J. (1996). Vergence control for robotic heads using log-polar images. In *Proc. of the 1996 IEEE/RSJ Int. Conference on Intelligent Robots and Systems*, Osaka, Japan.
- [Bertero et al., 1988] Bertero, M., Poggio, T., and Torre, V. (1988). Ill-posed problems in early vision. *Proceedings of the IEEE*, 76(8):869–889.
- [Faugeras, 1993] Faugeras, O. (1993). *Three-Dimensional Computer Vision*. The MIT Press.
- [Faugeras et al., 1987] Faugeras, O., Lustman, F., and Toscani, G. (1987). Motion and structure from motion from point and line matches. In *Proc. 1st Intern. Conf. Comput. Vision*, London.

- [Fermuller, 1993] Fermuller, C. (1993). Basic visual capabilities. Computer Vision Laboratory, University of Maryland, USA.
- [Fermuller, 1995a] Fermuller, C. (1995a). Passive navigation as a pattern recognition problem. *Int. Journal of Computer Vision*, 14(2):147–158.
- [Fermuller, 1995b] Fermuller, C. (1995b). Qualitative egomotion. *Int. Journal of Computer Vision*, 15(1/2):7–29.
- [Fermuller and Aloimonos, 1993] Fermuller, C. and Aloimonos, Y. (1993). The role of fixation in visual motion analysis. *Int. Journal of Computer Vision*, 11(2):165–186.
- [Fermuller and Aloimonos, 1994] Fermuller, C. and Aloimonos, Y. (1994). On the geometry of visual correspondence. CAR-TR-732, CS-TR-3341.
- [Fermuller and Aloimonos, 1995] Fermuller, C. and Aloimonos, Y. (1995). Vision and action. *Image and Vision Computing*, 13(10):725–744.
- [Gibson, 1950] Gibson, J. (1950). *The perception of the visual world*. Houghton-Mifflin.
- [Griswald et al., 1992] Griswald, N., Lee, J., and Weiman, C. (1992). Binocular fusion revisited utilizing a log-polar tessellation. *Computer Vision and Image Processing*, pages 421–457.
- [Heeger and Jepson, 1992] Heeger, D. and Jepson, A. (1992). Subspace methods for recovering rigid motion i: Algorithm and implementation. *Int. Journal of Computer Vision*, 7(2):95–117.
- [Horn, 1986] Horn, B. (1986). *Robot Vision*. The MIT Press.
- [Horn, 1987] Horn, B. (1987). Motion fields are hardly ever ambiguous. *International Journal of computer Vision*, 1(3):259–274.
- [Horn and Shunck, 1981] Horn, B. P. and Shunck, B. (1981). Determining optical flow. *Artificial Intelligence*, 17:185–203.

- [Horn and Weldon, 1988] Horn, B. P. and Weldon, E. (1988). Direct methods for recovering motion. *International Journal of computer Vision*, 2(1):51–76.
- [Longuet-Higgins and Prazdny, 1980] Longuet-Higgins, H. C. and Prazdny, K. (1980). The interpretation of a moving retinal image. In *Proc. Roy. Soc. London*, London B 208.
- [Marr, 1982] Marr, D. (1982). *Vision*. W. H. Freeman.
- [Negahdaripour and Lee, 1992] Negahdaripour, S. and Lee, S. (1992). Motion recovery from images sequences using only first order optical flow information. *International Journal of computer Vision*, 9(3):163–184.
- [Rousseeuw and Leroy, 1987] Rousseeuw, P. and Leroy, A. (1987). *Robust Regression & Outlier Detection*. John Wiley & Sons, Inc.
- [Santos-Victor et al., 1995] Santos-Victor, J., Sandini, G., Curotto, F., and Garibaldi, S. (1995). Divergent stereo in autonomous navigation : From bees to robots. *Int. Journal of Computer Vision*, 14(2):159–177.
- [Silva and Santos-Victor, 1996a] Silva, C. and Santos-Victor, J. (1996a). Direct egomotion estimation. In *Proc. of the 13th Int. Conference on Pattern Recognition*, Vienna, Austria.
- [Silva and Santos-Victor, 1996b] Silva, C. and Santos-Victor, J. (1996b). Geometric approach for egomotion estimation using normal flow. In *Proc. of the 4th Int. Symposium on Intelligent Robotic Systems*, Lisbon, Portugal.
- [Silva and Santos-Victor, 1996c] Silva, C. and Santos-Victor, J. (1996c). Robust egomotion estimation from the normal flow using search subspaces. Submitted to IEEE Transactions on PAMI in 1996.
- [Silva and Santos-Victor, 1997] Silva, C. and Santos-Victor, J. (1997). Egomotion estimation using log-polar images. Submitted to IEEE Int. CVPR-97.

- [Sinclair et al., 1994] Sinclair, D., Blake, A., and Murray, D. (1994). Robust estimation of egomotion from normal flow. *Int. Journal of Computer Vision*, 13(1):57–69.
- [Terzopoulos, 1986] Terzopoulos, D. (1986). Regularization of inverse visual problems involving discontinuities. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8(4):413–424.
- [Tikhonov and Arsenin, 1977] Tikhonov, A. and Arsenin, V. (1977). *Solution of ill-posed problems*. Washington DC: Winston.
- [Tistarelli and Sandini, 1993] Tistarelli, M. and Sandini, G. (1993). On the advantages of polar and log-polar mapping for direct estimation of the time-to-impact from optical flow. *IEEE Trans. on PAMI*, 15(8):401–411.
- [Wallace et al., 1995] Wallace, R., Ong, P., Bederson, B., and Schwartz, E. (1995). Space variant image processing. *Int. Journal of Computer Vision*, 13(1):71–90.
- [Waxman et al., 1987] Waxman, A. M., Kamgar-Parsi, B., and Subbarao, M. (1987). Closed-form solutions to image flow equations for 3d structure and motion. *Int. Journal of Computer Vision*, 1(3):239–258.
- [Weiman, 1990] Weiman, C. (1990). Log-polar vision for mobile robot navigation. In *Proc. of Electronic Imaging Conference*, pages 382–385, Boston, USA.
- [Weng et al., 1993] Weng, J., Ahuja, N., and Huang, T. (1993). Optimal motion and structure estimation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 15(9):864–884.
- [Weng et al., 1989] Weng, J., Huang, T., and Ahuja, N. (1989). Motion and structure from two perspective views: Algorithms, error analysis and error estimation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(5):451–476.