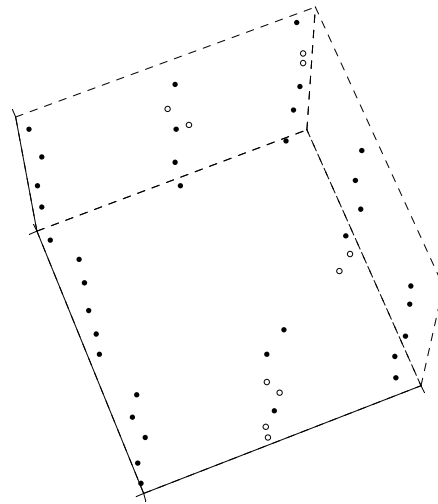




INSTITUTO
SUPERIOR
TÉCNICO

UNIVERSIDADE TÉCNICA DE LISBOA
INSTITUTO SUPERIOR TÉCNICO



MODELOS 3D A PARTIR DE VÍDEO

Bruno Braz Gonçalves, nº 42327, AE de Telecomunicações

LICENCIATURA EM ENGENHARIA ELECTROTÉCNICA E DE COMPUTADORES
Relatório de Trabalho Final de Curso
43/2002/L

Prof. Orientador: Pedro Manuel Quintas Aguiar

Novembro de 2003

MODELOS 3D A PARTIR DE VÍDEO

Universidade Técnica de Lisboa, Instituto Superior Técnico
Relatório de Trabalho Final de Curso

Aluno

Bruno Braz Gonçalves
Número 42327
Licenciatura de
Engenharia Electrotécnica e de Computadores
Área de Especialização de Telecomunicações
Instituto Superior Técnico

Orientador

Professor Doutor
Pedro Manuel Quintas Aguiar
Departamento de
Engenharia Electrotécnica e de Computadores
Instituto de Sistemas e Robótica
Instituto Superior Técnico

Agradecimentos

Agradeço à minha família e amigos pelo apoio e paciência.
Agradeço também ao Prof. Pedro Aguiar pela total disponibilidade e apoio.

Resumo

Este trabalho trata a recuperação da estrutura tridimensional (3D) de objectos a partir de vídeo. Para recuperar a estrutura 3D utiliza-se habitualmente o movimento no plano da imagem e por isso o problema é denominado na literatura anglo-saxónica por *structure from motion* (SFM). O SFM é um problema que se insere na área da análise do movimento, sendo esta área bastante tratada no domínio da visão por computador.

Para resolver o SFM utiliza-se habitualmente o método da factorização matricial introduzido, no início dos anos 90, por Tomasi e Kanade. Confirma-se neste trabalho que o método da factorização estima a forma do objecto e o movimento da câmara de uma forma robusta e eficiente. Neste método, o objecto é representado por um conjunto de pontos característicos no espaço 3D. Num primeiro passo deste processo SFM estima-se o movimento 2D da projecção destes pontos no plano da imagem. Neste âmbito apresentam-se dois métodos de seguimento de pontos característicos: o *Seguidor de Intensidades* e o *Seguidor de Pontos Característicos*. No segundo passo deste processo estima-se a estrutura 3D factorizando a matriz que coleciona as trajectórias dos pontos seguidos.

Naturalmente, a reconstrução 3D é tanto mais precisa quanto maior for o número de ângulos de vista do objecto no vídeo. Porém, como em cada ângulo de vista apenas é possível visualizar alguns pontos característicos do objecto, é normal que um ponto esteja visível num grupo de imagens e depois desapareça, ficando então ocluso. E é também normal que o ponto volte a aparecer, denominando-se este acontecimento por reencontro. Assim, num vídeo considerado bom para a reconstrução 3D existem sempre períodos de oclusão. Estes períodos de oclusão originam informação desconhecida, acerca das posições dos pontos característicos no vídeo.

Métodos que estendem a factorização a matrizes incompletas foram alvo de trabalhos recentes. Contudo, esses trabalhos não consideram os reencontros, i.e., interpretam o aparecimento de pontos característicos como se de pontos nunca previamente vistos se tratassem. Neste trabalho procura-se contribuir para colmatar essa lacuna, propondo-se métodos para identificação automática de reencontros. Mostra-se que a identificação dos reencontros é importante para o SFM na medida em que permite melhorar quer a estimação da informação desconhecida quer a estimação da estrutura 3D. Propõem-se dois métodos originais: *Inspecção da Forma* e *Análise de Custo*. O método *Inspecção da Forma* identifica reencontros com base na distância entre os pontos característicos no espaço 3D; a eficiência deste método depende do desempenho do processo SFM que é feito inicialmente. O método *Análise do Custo* é inspirado no princípio da Descrição de Comprimento Mínimo. Neste método, os reencontros são identificados minimizando uma função de custo global. A função de custo utilizada é uma função de Verosimilhança Penalizada, i.e., o custo combina dois termos: um mede a distância aos dados observados; outro penaliza a complexidade. Mostra-se que o método *Análise de Custo* resolve eficientemente o problema SFM com oclusão e reencontros. Embora seja sensível ao nível de ruído e à quantidade de informação conhecida, apresenta bom desempenho para uma gama pertinente de valores.

O trabalho contemplou uma parte experimental de validação e avaliação de desempenho dos diversos algoritmos implementados. Em particular, estudaram-se os vários passos do método de seguimento, o efeito do ruído no método da factorização e nos métodos de identificação de reencontros e o impacto de diferentes expressões para o cálculo do custo. Finalmente, são apresentadas experiências de modelação 3D a partir de vídeo com os métodos desenvolvidos.

Palavras Chave: Processamento de Imagem; Visão por Computador; Estrutura 3D a Partir do Movimento; Factorização Matricial com Dados Desconhecidos; Verosimilhança Penalizada; Descrição de Comprimento Mínimo.

Índice

Resumo	ii
Índice	iii
Lista de Figuras.....	v
Lista de Tabelas	vi
Lista de Siglas.....	vii
Lista de Programas	vii
1. Introdução	1
1.1. Motivação e Objectivos	1
1.2. Metodologia.....	2
1.3. Contribuição Original	3
1.4. Organização do Relatório	3
2. Selecção e Seguimento de Pontos Característicos.....	5
2.1. Introdução	5
2.2. Selecção	6
2.3. Seguimento	8
2.3.1. Seguidor de Intensidades	8
2.3.2. Seguidor de Pontos Característicos	9
2.3.3. Análise e Escolha do Método de Seguimento	9
2.4. Oclusões.....	10
3. Estimção de Estrutura 3D Usando Factorização Matricial.....	12
3.1. Introdução	12
3.2. Definições	12
3.3. Factorização Matricial	14
3.4. Factorização com Dados Desconhecidos.....	15
3.4.1. Introdução	15
3.4.2. Análise dos métodos de estimção de informação desconhecida.....	16
4. Identificação de Reencontros.....	17
4.1. Introdução	17
4.2. Inspeção da Forma	19
4.2.1. Descrição do Método	19
4.2.2. Cálculo da Distância de Reencontro.....	20
4.3. Análise do Custo.....	22
4.3.1. Descrição do Método	22
4.3.2. Implementação do método.....	23
4.3.3. Cálculo do Factor de Custo.....	24
4.4. Análise dos Métodos de Identificação de Reencontros	27
5. Resultados Experimentais.....	28
5.1. Introdução	28
5.2. Selecção de Pontos Característicos.....	28
5.3. Seguimento de Pontos Característicos.....	31
5.4. Factorização Matricial	33
5.5. Importância da Identificação de Reencontros.....	35
5.6. Métodos de Identificação de Reencontros	37
5.6.1. Experiência Com Dados Artificiais	37
5.6.2. Experiência Com Vídeo Real	38
5.6.3. Custo Computacional.....	39

5.7. Modelação 3D a Partir de Vídeo	41
6. Análise do Desempenho com Ruído	43
6.1. Introdução	43
6.2. Factorização Matricial	44
6.3. Identificação dos Reencontros	47
7. Conclusões	48
7.1. Sumário	48
7.2. Direcções Futuras	49
Anexo A. Descrição do Método da Factorização	i
A.1. Introdução	i
A.2. Matriz com Característica 3	i
A.3. Matriz com Característica 4	iv
Anexo B. Descrição da Estimação de Matrizes Completas.....	vi
B.1. Introdução	vi
B.2. Estimativa Inicial	vi
B.3. Algoritmo Iterativo Expectation-Maximization	viii
B.4. Algoritmo Iterativo Two-Steps	viii
Anexo C. Expressão do Custo de Reencontro	x
C.1. Introdução	x
C.2. Cálculo do Custo do Reencontro	xi
C.3. Conclusões	xiv
Anexo D. Efeito dos Reencontros ao Estimar Matrizes Completas	xv
D.1. Introdução	xv
D.2. Evolução do Erro da Estimação	xv
D.3. Relações Entre os Erros dos Reencontros Possíveis.....	xviii
D.4. Conclusões	xx
Anexo E. Referências.....	xxi

Lista de Figuras

Fig. 1.1: Processo <i>Structure From Motion</i>	2
Fig. 1.2: Factorização com dados desconhecidos	2
Fig. 2.1: Exemplos de pontos característicos e não característicos	5
Fig. 2.2: Problema da Abertura na decisão de um ponto ser característico.....	6
Fig. 2.3: As 3 fases do processo de selecção de pontos característicos.....	6
Fig. 2.4: Seguidor de Intensidades.....	8
Fig. 2.5: Seguidor de Pontos Característicos	9
Fig. 2.6: Ocorrência de reencontro de um ponto característico	11
Fig. 3.1: Modelo de projecção ortográfica.....	13
Fig. 3.2: Matriz de observação \mathbf{W} com informação desconhecida (\mathbf{W}_0).....	16
Fig. 4.1: Tratamento das Oclusões	17
Fig. 4.2: Matriz de observação com informação desconhecida e reencontros (identificados pela cor clara).....	18
Fig. 4.3: Reencontros possíveis na matriz \mathbf{W}_0	19
Fig. 4.4: Método da Inspeção da Forma	19
Fig. 4.5: Escolha da distância de decisão do reencontro.	21
Fig. 4.6: Processo de escolha do modelo	22
Fig. 4.7: Método da Análise do Custo	22
Fig. 4.8: Transformação pretendida com a utilização do custo de reencontro	25
Fig. 4.9: Evolução do número de reencontros identificados em do factor de custo α	26
Fig. 5.1: Resultados do passo Gradiente do processo de <i>Seleção</i>	29
Fig. 5.2: Resultados do processo de <i>Seleção</i> com ϵ	30
Fig. 5.3: Resultados do Seguidor de Pontos Característicos	32
Fig. 5.4: Desempenho do método da factorização com dados conhecidos e sem ruído... 33	
Fig. 5.5: Desempenho do método da factorização com informação desconhecida e sem ruído.	34
Fig. 5.6: Matrizes de observação utilizadas na experiência que ilustra a importância dos reencontros	35
Fig. 5.7: Importância da identificação dos reencontros no processo SFM.....	36
Fig. 5.8: Forma do objecto antes e depois da identificação de reencontros	37
Fig. 5.9: Trajectórias dos pontos característicos no plano da imagem na experiência com a pilha de discos	38
Fig. 5.10: Seguimento no vídeo, reencontros correctos e reencontros encontrados na experiência com a pilha de discos	39

Fig. 5.11: Tempo de processamento dos métodos <i>Inspecção da Forma e Análise do Custo</i>	40
Fig. 5.12: Seguimento no vídeo e reencontro na experiência com o cubo	41
Fig. 5.13: Matrizes de observação da experiência com o cubo	42
Fig. 5.14: Forma 3D reconstruída na experiência com o cubo	42
Fig. 6.1: Modelo de projecção em perspectiva	43
Fig. 6.2: Desempenho do método da Factorização com dados conhecidos e com ruído ($\sigma = 4pixels$)	44
Fig. 6.3: Efeito do ruído no processo <i>Structure From Motion</i>	45
Fig. 6.4: Importância do número de imagens no combate ao ruído	46
Fig. 6.5: Sensibilidade dos métodos <i>Inspecção da Forma e Análise do Custo</i> ao ruído.	47
Fig. B.1: Matriz de observação \mathbf{W} com informação desconhecida (\mathbf{W}_0).....	vi
Fig. B.2: Matriz \mathbf{W}_0 e a estimação do subespaço	vi
Fig. B.3: Matriz \mathbf{W}_0 e a combinação de subespaços	vii
Fig. C.1: Exemplos das matrizes de observação utilizadas	xi
Fig. C.2: Estimar o factor de custo $\alpha - 2^a$ expressão	xii
Fig. C.3: Estimar o factor de custo $\alpha - 3^a$ expressão	xiii
Fig. C.4: Estimar o factor de custo $\alpha - 4^a$ expressão	xiii
Fig. C.5: Estimar o factor de custo $\alpha - 5^a$ expressão	xiv
Fig. D.1: Exemplos das matrizes de observação utilizadas.....	xv
Fig. D.2: Erros da estimação de informação desconhecida com reencontros – 1 ^a experiência	xvii
Fig. D.3: Erros da estimação de informação desconhecida com reencontros – 2 ^a experiência	xviii
Fig. D.4: Erro na estimação de matrizes de observação completas com reencontros	xx

Lista de Tabelas

Tab. 5.1: Resultados do passo Gradiente do processo de selecção	30
Tab. 5.2: Variação do número de reencontros descobertos com a distância de reencontro no método <i>Inspecção da Forma</i>	38
Tab. 5.3: Tempo de processamento dos métodos <i>Inspecção da Forma e Análise do Custo</i>	39

Tab. D.1: Erros da estimação de informação desconhecida com reencontros – 1ª experiência.....	xvi
Tab. D.2: Erros da estimação de informação desconhecida com reencontros – 2ª experiência.....	xvii
Tab. D.3: Número de testes em que as relações entre os vários erros se verificam.....	xix

Lista de Siglas

EM-----	Expectation-Maximization
IEEE-----	Institute of Electrical and Electronics Engineers
LS-----	Least Squares
MDL-----	Minimum Description Length
ML-----	Maximum Likelihood
PL-----	Penalized Likelihood
SFM -----	Structure From Motion
SVD -----	Singular Value Decomposition
TS-----	Two-Step

Lista de Programas

MatLab v6.5 R13 – utilizado na implementação dos algoritmos
3D Studio Max 4 – utilizado para produzir vídeo artificial

1. Introdução

1.1. Motivação e Objectivos

Este trabalho tem três objectivos: (i) compreender a recuperação da estrutura tridimensional (3D) de um objecto inanimado a partir de vídeo, estudando as várias fases do seu processamento; (ii) contribuir para o desenvolvimento de soluções eficientes para este problema; e (iii) implementar um sistema de recuperação da estrutura 3D a partir de vídeo. Estes tipos de sistemas têm aplicações em robótica, vídeo digital e realidade virtual. Aplicações como navegação de veículos, manipulação de robots ou reconhecimento de objectos encontram-se limitadas pela percepção do espaço 3D. A utilização de representações 3D na codificação de vídeo permite taxas de compressão de dados elevadas, além de permitir criar vídeo interactivo. A percepção da realidade 3D é cada vez mais importante no desenvolvimento de ambientes multimédia. O facto de se utilizar diariamente o vídeo para representar a realidade 3D a duas dimensões motiva a resolução do problema inverso: recuperar a estrutura 3D a partir de vídeo.

Por estrutura 3D entende-se a forma dos objectos filmados e o movimento efectuado pela câmara. Na recuperação da estrutura 3D utiliza-se habitualmente o movimento no plano bidimensional (2D) – o plano da imagem. Este problema é por isso usualmente denominado por *structure from motion* (SFM) na literatura anglo-saxónica e insere-se na área da análise de movimento, bastante tratada em visão por computador. O SFM tem sido objecto de maior estudo nas últimas duas décadas. O seu estudo inicia-se com Ullman [1] em 1979 mas só em 1992 foi introduzido o método da factorização por Tomasi e Kanade [2], que resolve o problema de uma forma simples, fiável e robusta. Este método assenta num modelo de projecção ortográfica, sendo depois estendido a outros modelos mais gerais no artigo [3]. Em [4] aplica-se o SFM para a segmentação a partir do movimento, permitindo a recuperação da estrutura 3D de vários objectos num vídeo.

O movimento da câmara em relação ao objecto é importante pois quanto mais ângulos de vista do objecto existirem, melhor a recuperação da estrutura 3D. Por outro lado, mais movimento câmara-objecto implica que partes do objecto se tornem visíveis e outras desapareçam, existindo assim informação acerca do objecto que é desconhecida. O estudo das oclusões é por isso importante, pois estas sempre existirão na informação extraída do vídeo. Este problema denomina-se por SFM com informação desconhecida. Utilizar a factorização com informação desconhecida é objecto de estudo desde a apresentação do método da factorização em [2]. Neste mesmo artigo é apresentada uma solução muito custosa em termos computacionais. Desde então outros trabalhos surgiram para resolver este problema de forma mais eficiente [3, 5, 6, 7].

Os trabalhos mencionados [3, 5, 6, 7] e outros que tratam a factorização com informação desconhecida desprezam a ocorrência de reencontros, i.e., de partes do objecto que estiveram visíveis durante alguns instantes, ficando depois oclusas e que voltam a aparecer no vídeo. Ao ignorar os reencontros, consideram-se partes do objecto já anteriormente vistas como sendo novas regiões, que não se relacionam com a informação já reunida. O estudo das oclusões não deve ignorar o acontecimento de reencontros pois além de ocorrerem com frequência são bastante desejáveis, dada a necessidade de vários ângulos de vista para se conhecer bem o objecto. A identificação dos reencontros é importante pois melhora a estimação da informação desconhecida e permite que o cálculo da estrutura 3D seja mais rigorosa, assente em informação mais precisa. Enquanto que a estimação da informação desconhecida tem sido objecto de estudo, a identificação dos reencontros não. Após compreender e implementar os

métodos SFM escolhidos, este trabalho centra-se na resolução do problema das oclusões, em particular no desenvolvimento de métodos para identificação dos reencontros.

1.2. Metodologia

Utiliza-se o método de factorização de Tomasi e Kanade [2] por ser robusto e computacionalmente simples. O processo SFM é então dividido em dois blocos, como ilustra a Fig. 1.1.

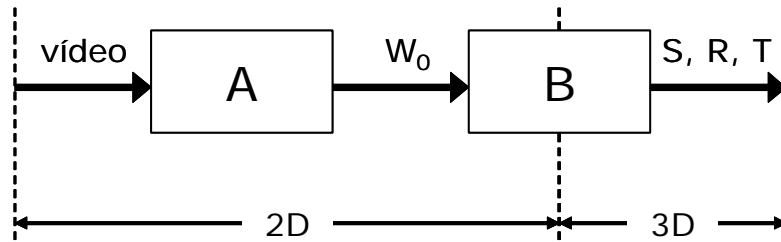


Fig. 1.1: Processo *Structure From Motion*

O bloco A da Fig. 1.1 extrai a informação relevante das imagens do vídeo: o movimento 2D do objecto. Tanto a saída como a entrada deste primeiro bloco são representações 2D no plano da imagem.

O bloco B da Fig. 1.1 é o responsável pela recuperação da estrutura 3D dos objectos. É neste que o método da factorização é aplicado. Tem como entrada a descrição do movimento 2D dos objectos e como saída as informações relativas à forma do objecto (**S**) e rotação (**R**) e translação (**T**) da câmara.

Para estimar o movimento 2D no bloco A da Fig. 1.1 são utilizados pontos que caracterizam o objecto na sequência de imagens – pontos característicos. No estudo da selecção destes pontos numa imagem são aplicados conhecimentos de [8]. Ao seguir as trajectórias destes pontos característicos na sequência de imagens (*tracking*, na literatura anglo-saxónica) obtém-se a informação do movimento 2D do objecto. Neste relatório apresentam-se dois métodos de *tracking* dos pontos característicos pelo vídeo: o *Seguidor de Intensidades* e o *Seguidor de Pontos Característicos*. Dos métodos de *tracking* resulta uma matriz W_0 , denominada matriz de observação, com as coordenadas dos pontos característicos em cada imagem. Devido à oclusão, verifica-se que cada ponto é observado continuamente, i.e., sem interrupções de seguimento, por períodos limitados de tempo. Assim, a matriz W_0 tem entradas desconhecidas.

A análise do seguimento dos pontos característicos no vídeo conduz ao estudo das oclusões e inclusões pois para a factorização é necessária uma matriz de observação **W** completamente conhecida. Na Fig. 1.2 mostra-se como o bloco B da Fig. 1.1 é então subdividido em dois blocos: um para o tratamento das oclusões e outro que aplica o método da factorização. Para processar o SFM com informação desconhecida utilizam-se os algoritmos apresentados em [7], que permitem estimar as posições desconhecidas.

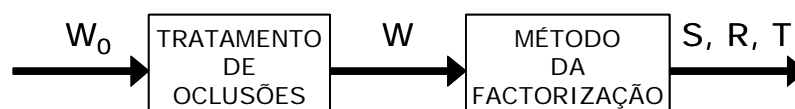


Fig. 1.2: Factorização com dados desconhecidos

Todos os trabalhos mencionados acima ignoram a ocorrência de reencontros, i.e., tratam o reaparecimento de um ponto previamente observado como se de um novo ponto se tratasse. Este trabalho pretende colmatar essa falta, introduzindo dois métodos para os identificar: *Inspecção da Forma* e *Análise do Custo*.

Inspecção da Forma

O método de *Inspecção da Forma* assenta no facto de a não identificação de um reencontro resultar em dois pontos característicos muito próximos na forma 3D. O reencontro é identificado avaliando a distância entre dois pontos característicos no espaço 3D. O desempenho deste método é dependente do desempenho do processo SFM que é feito inicialmente, bem como da definição de uma distância de decisão. Enquanto que o método da factorização como processo SFM é fiável, como se mostra neste relatório, a definição do elemento de decisão – distância entre pontos característicos – não é fácil, dependendo fortemente de algo que é desconhecido, a distância do plano da imagem ao objecto. O reencontro de um ponto característico por inspecção da forma 3D é um método intuitivo. Este serve para inicializar o estudo dos reencontros e estabelecer uma base de comparação para o segundo método.

Análise do Custo

O método *Análise do Custo* é um método mais elegante e robusto que decide os reencontros pelo seu impacto num custo global. Este método, inspirado no princípio de Descrição de Comprimento Mínimo – *Minimum Description Length* (MDL) na literatura anglo-saxónica –, integra o processamento das oclusões e antecede a reconstrução 3D. Ao minimizar uma função de custo global procura-se, por um lado, escolher uma solução que minimize a distância do modelo aos dados observados, por outro, penalizar soluções complexas, i.e., soluções com mais pontos característicos. Funções de custo que contemplam estes dois termos são habitualmente designadas por Verosimilhança Penalizada – *Penalized Likelihood* (PL) na literatura anglo-saxónica [9]. Naturalmente, na definição da função de custo, influi um parâmetro que quantifica a influência do número de pontos. Neste trabalho analisa-se a expressão do custo de reencontro e apresentam-se valores para o parâmetro referido, que se mostra serem válidos para um grande número de situações.

1.3. Contribuição Original

Todos os trabalhos no âmbito da factorização com informação desconhecida desprezam o facto de o mesmo ponto característico poder ter vários períodos de seguimento [Tomasi&Kanade, Jacobs, Maruyama&Kurumi, Guerreiro&Aguiar]. Este trabalho pretende iniciar o estudo dos reencontros e de propor soluções para a sua identificação. Para isso são introduzidos dois métodos: *Inspecção da Forma* e *Análise do Custo*. O método *Inspecção da Forma* assenta no facto de a ocorrência de um reencontro resultar em dois pontos característicos muito próximos. O método *Análise do Custo*, inspirado em MDL, decide reencontros com base no seu impacto numa função de custo global.

1.4. Organização do Relatório

O relatório está organizado em 8 capítulos. No capítulo 2 apresenta-se o método de selecção de pontos característicos e os dois métodos de seguimento. Também é descrito o aparecimento das oclusões. No capítulo 3 é descrita a solução de Tomasi e Kanade, bem como os algoritmos que resolvem o SFM com informação desconhecida. O capítulo 4

formaliza o problema dos reencontros e apresenta as duas soluções propostas neste trabalho: os métodos *Inspecção da Forma* e *Análise do Custo*. O capítulo 5 condensa os resultados experimentais dos três capítulos anteriores. Nas experiências realizadas são utilizados vídeos reais e outros criados artificialmente, bem como dados artificiais sobre o movimento 2D de modo a ter controlo sobre as coordenadas dos pontos característicos e do nível de ruído. No capítulo 6 analisa-se o desempenho dos métodos propostos em função do nível de ruído dos dados. O capítulo 7 conclui o relatório. Apresentam-se ainda quatro anexos. Os anexos A e B servem de apoio ao capítulo 3 e descrevem a implementação dos métodos de factorização matricial. Os anexos C e D apoiam o capítulo 4. No anexo C analisa-se a expressão do custo de reencontro e no anexo D o efeito dos reencontros ao estimar matrizes de observação completas.

2. Selecção e Seguimento de Pontos Característicos

2.1. Introdução

O bloco A da Fig. 1.1 tem o objectivo de extrair o movimento 2D dos objectos no vídeo. Isso é conseguido analisando o padrão de intensidades de cada imagem no vídeo, a partir do qual são escolhidos pontos que caracterizam o objecto na sequência de imagens – pontos característicos. Ao seguir as trajetórias destes pontos característicos na sequência de imagens obtém-se a informação do movimento 2D do objecto.

Cada imagem do vídeo é composta de pontos (*pixels* na literatura anglo-saxónica). Um ponto é considerado característico se for passível de ser seguido na sequência de imagens. As coordenadas dos pontos característicos em cada imagem do vídeo arrumam-se numa matriz W , denominada matriz de observação. Esta matriz é a saída do primeiro bloco da Fig. 1.1.

A tarefa do primeiro bloco é seleccionar e seguir os pontos característicos no vídeo identificando as suas coordenadas no plano da imagem. Esta tarefa divide-se então em:

- a selecção de pontos característicos numa imagem
- o seguimento de pontos característicos entre imagens (*tracking*, na literatura anglo-saxónica),

Selecção de Pontos Característicos

Os pontos característicos seleccionados numa imagem devem ser relevantes para a descrição do movimento 2D. Por isso têm de ser possíveis de identificar nas imagens posteriores (*tracking*).

Para conseguir seguir um ponto de imagem para imagem é necessário definir como o destacar dos demais. Um ponto característico tem em seu redor (região de pontos adjacentes) um padrão de intensidades que varia equilibradamente na vertical e na horizontal. São assim escolhidos pontos da imagem como cantos, sinais ou marcas. Esta selecção permite obter uma boa representação dos objectos, sendo excepção os que combinem as características de: linhas arredondadas, textura plana e padrões lisos.

Na Fig. 2.1 observa-se a diferença entre boas e más decisões do ponto do centro ser característico. Nas Fig. 2.1(a,b) o ponto será dificilmente reconhecível nas imagens posteriores dado que outros em seu redor têm as mesmas características. Nas Fig. 2.1(c,d) o ponto destaca-se dos outros, logo é característico.

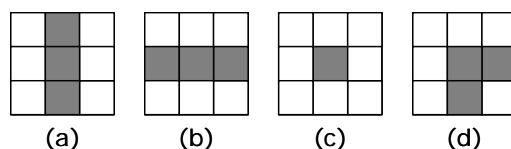


Fig. 2.1: Exemplos de pontos característicos e não característicos
Em (a) e (b) o ponto do centro não é característico pois tem características semelhantes aos em seu redor, não sendo possível segui-lo. Pela razão contrária, em (c) e (d) o ponto do centro é característico.

Na escolha de um ponto característico tem também relevância o Problema da Abertura. O tamanho e forma da região envolvente a analisar (a abertura) vai condicionar o modo como o ponto se destaca. O exemplo da Fig. 2.2 mostra o efeito da abertura na decisão do ponto negro ser característico. Deve considerar-se uma região em redor do ponto suficiente para analisar

se este se destaca ou não. Na Fig. 2.2(a) observa-se uma situação extrema: é impossível discernir se um ponto se destaca sem o comparar com os restantes. Na Fig. 2.2(b) apenas se compara o ponto com alguns pontos. Finalmente, na Fig. 2.2(c) é possível decidir que o ponto se destaca dos restantes.

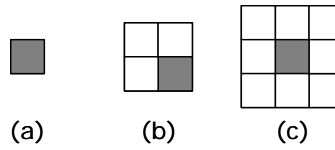


Fig. 2.2: Problema da Abertura na decisão de um ponto ser característico
Regiões pequenas como (a) e (b) não permitem comparar o ponto escuro de modo a decidir eficientemente se este é característico, ao contrário de (c).

A maioria dos pontos característicos seleccionados faz parte das linhas, das texturas e dos padrões visíveis na imagem. É possível assim reduzir o conjunto dos pontos a processar considerando apenas os pontos pertencentes aos contornos visíveis na imagem.

Na secção 2.2 descreve-se o método de selecção utilizado, que aplica estas definições e conclusões.

Seguimento de Pontos Característicos

Na secção 2.3 são propostos dois métodos de *tracking*: o *Seguidor de Intensidades* e o *Seguidor de Pontos Característicos*. Estes métodos permitem construir a matriz de observação W utilizando a selecção e executando o seguimento de maneiras diferentes.

2.2. Selecção

No processo de selecção pretende-se identificar os pontos característicos de uma imagem tal como referido acima. De modo a fazer uma selecção eficiente dos pontos de uma imagem, a escolha é então feita em três fases, como ilustrado na Fig. 2.3. Este método de selecção baseia-se no estudo apresentado em [8].

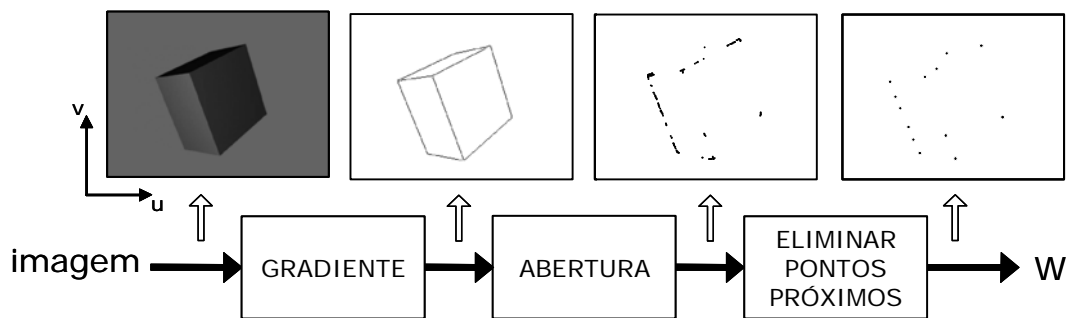


Fig. 2.3: As 3 fases do processo de selecção de pontos característicos

Gradiente

Aplica-se um método de detecção de contornos [10].

Calcula-se o gradiente da matriz I , que contém o padrão de intensidades de toda a imagem.

$$\nabla I = \frac{\partial I}{\partial u} \mathbf{i} + \frac{\partial I}{\partial v} \mathbf{j} \quad \text{Eq. 2.1}$$

Detectam-se os contornos visíveis na imagem encontrando os máximos relativos do gradiente. Uma condição mais geral é a definição de um *threshol* mínimo para a norma do gradiente,

$$\|\nabla I\| > \xi_1 \quad \text{Eq. 2.2}$$

Com a Eq. 2.2 variamos a sensibilidade à textura da imagem. Pretende-se assim encontrar um conjunto de pontos que se destaquem na imagem. Este conjunto é muito menor do que o conjunto de todos os pontos da imagem e por isso facilitará o processamento das restantes fases. Por isso esta fase tem benefícios em termos de tempo de processamento. A segunda fase pode ser utilizada sem encontrar os contornos da imagem, processando cada ponto da imagem, mas dessa maneira o tempo de processamento aumenta muito.

Para obter os melhores resultados utiliza-se habitualmente um método de aproximação da derivada como o método de Sobel, de Prewitt ou de Roberts [11].

Abertura

De modo a seleccionar o ponto central da região R_1 onde padrão de intensidades varia tanto verticalmente como horizontalmente, utiliza-se o método descrito em [8].

Constrói-se a matriz \mathbf{M} , de tamanho (2×2) , dada por

$$\mathbf{M} = \begin{bmatrix} \sum_{(u,v) \in R} \left(\frac{\partial \mathbf{I}}{\partial u} \right)^2 & \sum_{(u,v) \in R} \frac{\partial \mathbf{I}}{\partial u} \frac{\partial \mathbf{I}}{\partial v} \\ \sum_{(u,v) \in R} \frac{\partial \mathbf{I}}{\partial u} \frac{\partial \mathbf{I}}{\partial v} & \sum_{(u,v) \in R} \left(\frac{\partial \mathbf{I}}{\partial v} \right)^2 \end{bmatrix} \quad \text{Eq. 2.3}$$

Em [8] mostra-se que a estimação do movimento 2D depende do condicionante de \mathbf{M} . O condicionante $k(\mathbf{M})$ da matriz \mathbf{M} é dado pelo rácio

$$k(\mathbf{M}) = \frac{\lambda_1(\mathbf{M})}{\lambda_N(\mathbf{M})} \quad \text{Eq. 2.4}$$

em que $\lambda_1(\mathbf{M})$ é o seu maior valor singular e $\lambda_N(\mathbf{M})$ o menor. Se um dos componentes do gradiente do padrão de intensidades da imagem for muito maior que o outro tem-se

$$\lambda_1(\mathbf{M}) \gg \lambda_N(\mathbf{M})$$

logo $k(\mathbf{M})$ é grande. Se $k(\mathbf{M})$ for pequeno sabe-se que a variação numa direcção do padrão de intensidades não é muito maior do que a variação na direcção perpendicular [8].

Escolhem-se então pontos que são centro de regiões R_1 onde $k(\mathbf{M})$ está abaixo de um limite ξ_2 ,

$$k(\mathbf{M}) < \xi_2 \quad \text{Eq. 2.5}$$

Eliminar Pontos Próximos

Neste passo eliminam-se pontos característicos próximos entre si (dentro de uma região R_2), escolhendo o de $k(\mathbf{M})$ menor. A necessidade deste passo deve-se ao facto dos pontos característicos próximos:

- serem difíceis de seguir nas imagens posteriores
- não contribuírem relevantemente para definição dos objectos na imagem

Valores Utilizados na Implementação

Para detecção dos contornos utiliza-se o método de Sobel [11] para aproximar a derivada e escolhidos os pontos correspondentes aos máximos relativos do gradiente.

As regiões definidas em redor de um ponto (u_{fp}, v_{fp}) devem ser simétricas. Neste trabalho utilizam-se regiões quadradas como nas Fig. 2.1 e Fig. 2.2, sendo a escolha do comprimento do lado L importante em cada situação. Como o valor L pode não ser conseguido para todos os pontos, nomeadamente nos pontos próximos dos limites da imagem, a Eq. 2.6 considera os limites $L_{u\max}$, $L_{u\min}$, $L_{v\max}$ e $L_{v\min}$, todos valores inteiros positivos.

$$R = \left\{ (u_{fp} + k_u, v_{fp} + k_v) ; k_u \in [-L_{u\min}; L_{u\max}] , k_v \in [-L_{v\min}; L_{v\max}] \right\} \quad \text{Eq. 2.6}$$

Na análise de cada ponto utiliza-se uma região R_1 de lado $L(R_1) = 5 \text{ pixels}$ e $\xi_2 = 45$. O valor de $L(R_1)$ é escolhido pesando a eficiência e o custo computacional da Eq. 2.5. O valor de ξ_2 escolhido é resultado de experiências para esse fim e com base em [8]. Eliminam-se os pontos próximos numa região com $L(R_2) = 21 \text{ pixels}$, sendo este valor escolhido qualitativamente por observação do vídeo.

2.3. Seguimento

2.3.1. Seguidor de Intensidades

O primeiro método de *tracking* proposto é o *Seguidor de Intensidades*. Neste método os pontos característicos são seleccionados na primeira imagem, sendo seguidos nas restantes, como ilustra a Fig. 2.4.

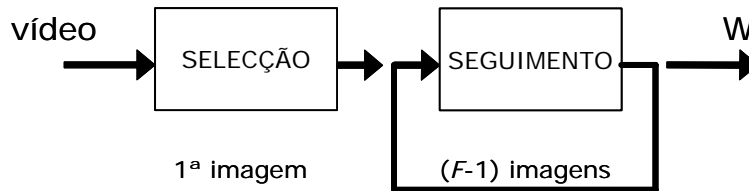


Fig. 2.4: Seguidor de Intensidades

Para seguir os pontos seleccionados procura-se nas imagens posteriores por pontos em seu redor com intensidades parecidas. É definida uma região R_3 na imagem em redor das coordenadas do ponto característico anterior, que agrupa os pontos possíveis de o corresponder.

Para escolher o ponto de intensidade mais parecida é imperativo comparar porções de imagem e não pontos. Isto porque a intensidade de um ponto pode variar rapidamente de imagem para imagem – por efeito da luz, reflexos – enquanto uma região da imagem não é normalmente confundida com outra na imagem seguinte.

São definidas duas regiões de igual forma e tamanho para comparar cada dois pontos: R_{4_1} com centro no ponto anterior e R_{4_2} com centro numa hipótese. O módulo da diferença de intensidades é calculado ponto-a-ponto,

$$\mathbf{I}_{f,(f-1)} = \left| \mathbf{I}_f(u_{fp} + k_u, v_{fp} + k_v) - \mathbf{I}_{(f-1)}(u_{(f-1)p} + k_v, v_{(f-1)p} + k_u) \right| \text{ com } \begin{cases} k_u \in [-L_{u\min}; L_{u\max}] \\ k_v \in [-L_{v\min}; L_{v\max}] \end{cases}$$

sendo depois feita a média,

$$\overline{\mathbf{I}}_{f,(f-1)} = \frac{1}{1 + L_{v \max} + L_{v \min}} \sum_{k_v=L_{v \min}}^{L_{v \max}} \left[\frac{1}{1 + L_{u \max} + L_{u \min}} \sum_{k_u=L_{u \min}}^{L_{u \max}} \mathbf{I}_{f-(f-1)}(k_v, k_u) \right]$$

É feito o seguimento como ponto que apresentar a diferença de intensidades mínima, tendo esta que estar abaixo de um valor ζ_3 , como na Eq. 2.7. Esta condição define o valor ζ_3 a partir do qual não se considera o seguimento, iniciando o período de oclusão.

$$\min \left\{ \overline{\mathbf{I}}_{f,(f-1)} \right\} < \zeta_3 \quad \text{Eq. 2.7}$$

O cálculo da média das diferenças das intensidades é importante pois as regiões das diferentes hipóteses podem não ter o mesmo tamanho, nomeadamente ao definir regiões em redor de pontos próximos dos limites da imagem.

2.3.2. Seguidor de Pontos Característicos

O segundo método de *tracking* proposto é o *Seguidor de Pontos Característicos*. Neste método os pontos característicos são seleccionados em todas as imagens, sendo de seguida feita a correspondência entre os pontos novos e os antigos.

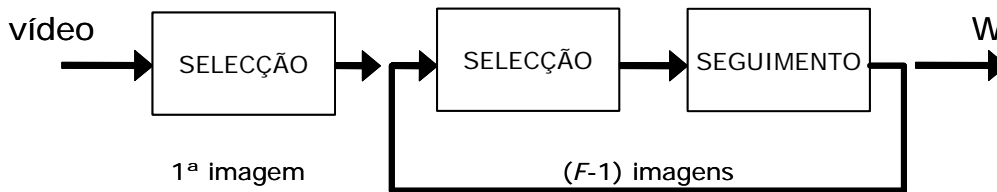


Fig. 2.5: Seguidor de Pontos Característicos

Ao fazer a correspondência de imagem para imagem é definida uma região R_3 em redor do ponto característico anterior de modo a encontrar correspondentes no conjunto da nova selecção. Podem ocorrer três tipos de situações:

- existir um ponto na região – fazer a correspondência
- existirem vários pontos na região – escolher o de intensidade mais parecida
- não existir nenhum ponto na região – o ponto não é mais seguido (occlusão)

Todos os pontos característicos da nova selecção que escaparam às regiões de correspondência são considerados, iniciando-se assim o seu seguimento.

Para comparar intensidades na segunda situação é necessário comparar porções de imagem e não pontos, como explicado acima. Assim, apenas no caso de existir mais do que uma hipótese para o seguimento de um ponto característico, são definidas duas regiões de igual forma e tamanho para comparar cada dois pontos: R_{4_1} com centro no ponto anterior e R_{4_2} com centro numa hipótese. Define-se ainda o valor ζ_3 como na Eq. 2.7, a partir do qual se deixa de seguir o ponto (occlusão).

2.3.3. Análise e Escolha do Método de Seguimento

Os dois métodos apresentados acima são dois extremos do *tracking* possível. Enquanto no primeiro a selecção de pontos característicos apenas é feita uma vez, no segundo método é feita em todas as imagens. Por esta razão o custo computacional do *Seguidor de Pontos Característicos* é superior ao *Seguidor de Intensidades*.

Os pontos característicos utilizados têm influência na reconstrução 3D resultante. O *Seguidor de Intensidades* executa o seguimento apenas comparando a intensidade, ignorando a definição de ponto característico descrita na secção 2.1. Assim, permite que pontos numa imagem que não respeitem esta definição sejam considerados característicos. Por esta razão o *Seguidor de Pontos Característicos* é mais eficiente ao realizar o *tracking*.

Um terceiro método possível seria intermédio aos dois descritos: um algoritmo que segue as intensidades e que executa a selecção de pontos característicos periodicamente. Este é mais eficiente do que o *Seguidor de Intensidades*. É também importante referir que, tal como o *Seguidor de Pontos Característicos*, permite iniciar o seguimento de novos pontos característicos a meio do vídeo, algo que o *Seguidor de Intensidades* não consegue. Outra razão para a implementação deste método intermédio seria o equilíbrio do custo computacional *versus* eficiência.

Neste trabalho utiliza-se o *Seguidor de Pontos Característicos* pois o aumento de eficiência justifica o maior custo computacional.

Valores Utilizados na Implementação

Executa-se o seguimento em regiões quadradas com lado $L(R_3) = 21\text{pixels}$. Este valor depende da velocidade do movimento 2D do objecto no vídeo e deve ser corrigido por observação dos resultados. Para comparar intensidades utilizam-se regiões com $L(R_{4_1}) = L(R_{4_2}) = 5\text{pixels}$, considerando-se este valor suficiente para uma comparação eficiente.

2.4. Oclusões

É o seguimento dos pontos característicos que permite a estimação do movimento 2D do objecto. Para conhecer bem o objecto é importante ter informação relativa ao maior número de ângulos de vista. Quando a câmara se move alguns pontos característicos aparecem e desaparecem da imagem, tendo normalmente períodos em que estão visíveis e outros em que estão oclusos. E é também normal que o ponto volte a aparecer, denominando-se este acontecimento por reencontro, como ilustrado na Fig. 2.6. Além disso, um ponto característico pode deixar de ser seguido devido aos métodos de seguimento não serem totalmente fiáveis.

No que diz respeito ao seguimento de um ponto característico num vídeo identificam-se então três situações distintas:

- o ponto é seguido por todo o vídeo
- o ponto é seguido num período de tempo
- o ponto é seguido em vários períodos disjuntos no tempo, acontecendo o seu reencontro como ilustrado na Fig. 2.6

Pelas duas razões acima enunciadas, primeira situação raramente se verifica. A segunda situação introduz o problema das oclusões no estudo do SFM, criando o problema SFM com informação desconhecida. A terceira situação motiva o problema de reconhecer o reencontro do ponto.

Neste trabalho estuda-se em pormenor a terceira situação: como identificar um reencontro, i.e., como decidir que um ponto característico seguido num período de tempo é o mesmo que o seguido noutra período de tempo do vídeo?

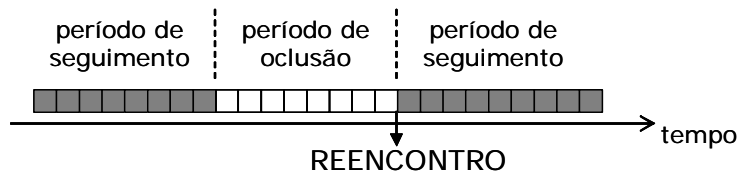


Fig. 2.6: Ocorrência de reencontro de um ponto característico
O ponto está visível e depois ocluso. O reencontro acontece quando o ponto volta a ser seguido.

O *Seguidor de Intensidades* ou o *Seguidor de Pontos Característicos* não identificam os reencontros. Quando um reencontro ocorre é iniciado o seguimento de um novo ponto característico a meio do vídeo. Assim, com os métodos de *tracking* acima propostos um ponto característico no vídeo resulta em tantos pontos característicos quantos os seus períodos de seguimento.

3. Estimação de Estrutura 3D Usando Factorização Matricial

3.1. Introdução

Neste capítulo descreve-se o bloco **B** da Fig. 1.1. Este bloco utiliza a factorização matricial para recuperar a estrutura 3D. Nas secções 3.2 e 3.3 expõem-se os conceitos e conclusões do método de factorização de [2]. Para resolver o problema SFM com informação desconhecida utilizam-se os algoritmos de [7], que são tratados na secção 3.4.

3.2. Definições

Observação

Existe uma matriz de observação do vídeo que tem a informação sobre as trajectórias dos P pontos característicos nas F imagens. Esta informação é o conjunto das coordenadas dos pontos característicos, em cada imagem do vídeo,

$$\{(u_{fp}, v_{fp}) \mid f = 1, \dots, F, p = 1, \dots, P\} \quad \text{Eq. 3.1}$$

A matriz de observação \mathbf{W} é formada pela matriz das coordenadas horizontais \mathbf{U} e pela matriz das coordenadas verticais \mathbf{V} ,

$$\begin{aligned} \mathbf{U} &= \begin{bmatrix} u_{fp} \end{bmatrix} \\ \mathbf{V} &= \begin{bmatrix} v_{fp} \end{bmatrix} \end{aligned} \rightarrow \mathbf{W} = \begin{bmatrix} \mathbf{U} \\ \mathbf{V} \end{bmatrix} \quad \text{Eq. 3.2}$$

Cada linha corresponde a uma imagem e cada coluna à coordenada de um ponto característico. A matriz \mathbf{W} tem por isso o tamanho $2F \times P$.

A normalização de \mathbf{U} e \mathbf{V} ($\tilde{\mathbf{U}}$ e $\tilde{\mathbf{V}}$) consiste em normalizar as coordenadas de cada imagem, i.e., subtrair a cada coordenada a média da sua linha. Forma-se assim a matriz de observação normalizada $\tilde{\mathbf{W}}$, com o mesmo tamanho de \mathbf{W} ,

$$\tilde{\mathbf{W}} = \begin{bmatrix} \tilde{\mathbf{U}} \\ \tilde{\mathbf{V}} \end{bmatrix} \quad \text{Eq. 3.3}$$

Ao normalizar a matriz \mathbf{W} retira-se a translação do objecto ao longo da sequência de imagens, ficando o centro geométrico dos pontos característicos em $(0,0)$.

Forma

Define-se a matriz da forma \mathbf{S} que reúne as coordenadas dos pontos característicos no espaço tridimensional, tendo como origem o seu centróide. Tem o tamanho $3 \times P$.

$$\mathbf{S} = \begin{bmatrix} \mathbf{s}_1 & \cdots & \mathbf{s}_p \end{bmatrix} = \begin{bmatrix} x_1 & \cdots & x_p \\ y_1 & \cdots & y_p \\ z_1 & \cdots & z_p \end{bmatrix} \quad \text{Eq. 3.4}$$

Rotação

Define-se a matriz de rotação \mathbf{R} que representa as orientações da horizontal e vertical do objecto em cada imagem do vídeo. Reúne os 2 versores tridimensionais que definem os eixos de cada imagem bidimensional. Tem o tamanho $2F \times 3$.

$$\mathbf{R} = \begin{bmatrix} \mathbf{i}_1^T \\ \vdots \\ \mathbf{i}_F^T \\ \mathbf{j}_1^T \\ \vdots \\ \mathbf{j}_F^T \end{bmatrix} = \begin{bmatrix} i_{x1} & i_{y1} & i_{z1} \\ \vdots & \vdots & \vdots \\ i_{xF} & i_{yF} & i_{zF} \\ j_{x1} & j_{y1} & j_{z1} \\ \vdots & \vdots & \vdots \\ j_{xF} & j_{yF} & j_{zF} \end{bmatrix} \quad \text{Eq. 3.5}$$

A Fig. 3.1 ilustra os dois sistemas de referência do problema no modelo de projecção ortográfica. Em ortografia qualquer projecção \mathbf{k} é paralela ao produto externo de \mathbf{i} e \mathbf{j} , verificando-se,

$$\mathbf{k} = \mathbf{i} \times \mathbf{j} \quad \text{Eq. 3.6}$$

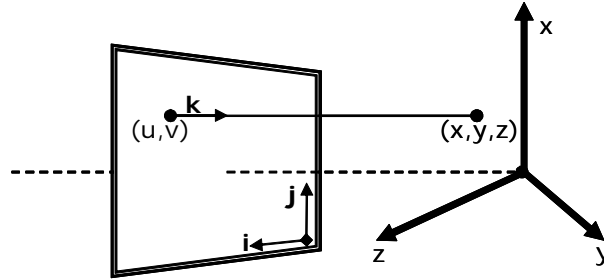


Fig. 3.1: Modelo de projecção ortográfica.
Em ortografia qualquer projecção é perpendicular ao plano da imagem.
Ilustram-se a relação dos dois sistemas: o plano da imagem e o espaço 3D.

Translação

Define-se o vector de translação \mathbf{T} que reúne as coordenadas do centro geométrico dos pontos característicos no plano bidimensional da imagem. Isso é obtido calculando a média de cada linha da matriz \mathbf{W} , isto é, calculando em cada imagem a média de cada coordenada u , v como na Eq. 3.7. Tem o tamanho $2F \times 1$.

$$T_i = \frac{1}{P} \sum_{j=1}^P w_{ij} \quad , \text{ com } i \in [1, 2F] \quad \text{Eq. 3.7}$$

Definições Equivalentes

É possível utilizar o mesmo formalismo organizando as matrizes \mathbf{W} , \mathbf{R} e \mathbf{T} de outra forma. As conclusões de Tomasi e Kanade (ver próxima secção) continuam válidas ao utilizar

$$\mathbf{W} = \begin{bmatrix} u_{11} & u_{12} & \cdots & u_{1P} \\ v_{11} & v_{12} & \cdots & v_{1P} \\ u_{21} & u_{22} & \cdots & u_{2P} \\ v_{21} & v_{22} & \cdots & v_{2P} \\ \vdots & \vdots & \ddots & \vdots \\ u_{F1} & u_{F2} & \cdots & u_{FP} \\ v_{F1} & v_{F2} & \cdots & v_{FP} \end{bmatrix} ; \quad \mathbf{R} = \begin{bmatrix} i_{x1} & i_{y1} & i_{z1} \\ j_{x1} & j_{y1} & j_{z1} \\ i_{x2} & i_{y2} & i_{z2} \\ j_{x2} & j_{y2} & j_{z2} \\ \vdots & \vdots & \vdots \\ i_{xF} & i_{yF} & i_{zF} \\ j_{xF} & j_{yF} & j_{zF} \end{bmatrix} ; \quad \mathbf{T} = \begin{bmatrix} t_{u1} \\ t_{v1} \\ t_{u2} \\ t_{v2} \\ \vdots \\ t_{uF} \\ t_{vF} \end{bmatrix}$$

Esta alteração permite aumentar a eficácia e simplicidade dos métodos de tratamento das oclusões descritos na secção 3.4.

3.3. Factorização Matricial

Uma das conclusões de Tomasi e Kanade [2] é a equação Eq. 3.8 quando se assume o modelo de projecção ortogonal.

$$\tilde{\mathbf{W}} = \mathbf{RS} \tag{Eq. 3.8}$$

Outra das conclusões é a igualdade Eq. 3.9, válida em ortografia,

$$\mathbf{W} = \mathbf{RS} + \mathbf{T} \times \mathbf{1}_{1 \times P} = [\mathbf{R} | \mathbf{T}] \begin{bmatrix} \mathbf{S} \\ \mathbf{1}_{1 \times P} \end{bmatrix} \tag{Eq. 3.9}$$

em que $\mathbf{1}_{1 \times P}$ é uma linha de P uns.

Por construção, Tomasi e Kanade concluem que $\tilde{\mathbf{W}}$ tem característica 3 e que \mathbf{W} tem característica 4. Assim, na presença de ruído esta será a primeira condição a aplicar: impor a característica. A operação de calcular uma aproximação à característica desejada denomina-se por factorização.

A segunda condição a impor vem por definição dos versores que constituem a matriz de rotação \mathbf{R} . Os vectores \mathbf{i} e \mathbf{j} são unitários e ortogonais entre si em cada imagem, verificando o sistema da Eq. 3.10. Resolver este sistema denomina-se por normalização.

$$\begin{cases} |\mathbf{i}_f| = |\mathbf{j}_f| = 1 \\ \mathbf{i}_f \cdot \mathbf{j}_f = \mathbf{i}_f^T \mathbf{j}_f = 0 \end{cases} \tag{Eq. 3.10}$$

O método de factorização tenta encontrar a melhor aproximação dentro das condições impostas. O problema de minimização da Eq. 3.11 é a solução ML ao utilizar a matriz de observação normalizada de característica 3. De notar que $\|\cdot\|_F$ é a norma de Frobenius.

$$\min_{(\mathbf{RS}) \in S_3} \|\tilde{\mathbf{W}} - \mathbf{RS}\|_F \tag{Eq. 3.11}$$

Ao utilizar a matriz de observação de característica 4 escreve-se,

$$\min_{(\mathbf{RS} + \mathbf{T}\mathbf{1}) \in S_4} \|\mathbf{W} - (\mathbf{RS} + \mathbf{T}\mathbf{1}_{1 \times P})\|_F \tag{Eq. 3.12}$$

No Anexo A descreve-se a implementação do método da factorização. Apresentam-se duas maneiras distintas de implementação: (i) o algoritmo como em [2], utilizando a característica 3; e (ii) o processo similar utilizando característica 4. Foi apenas implementada a factorização matricial de característica 3 pois apresenta maior resistência ao ruído. Isto porque calcula inicialmente a translação (\mathbf{T}) pela média das linhas de \mathbf{W} como na Eq. 3.7.

Ao utilizar a SVD no primeiro cálculo, a factorização com característica 4 torna-se mais sensível ao erro imposto pelo ruído.

Para desprezar o efeito da perspectiva é necessário que a profundidade relativa do objecto seja muito menor do que a sua distância ao plano da imagem. Deste modo o modelo de projecção ortogonal é uma boa aproximação da projecção perspectiva, ou seja, da realidade.

O modelo de projecção ortogonal não considera a translação ao longo do eixo óptico. O eixo óptico passa pelo centro da imagem e pelo centro do ponto, como ilustra a Fig. 3.1. Assim, é impossível estimar esta componente do movimento com este formalismo, bem como calcular a profundidade relativa do objecto e sua distância ao plano da imagem. Contudo, esta restrição aumenta a qualidade dos resultados da forma e dos outros 5 parâmetros do movimento.

A rigidez dos objectos, i.e., os pontos característicos caracterizarem uma forma inalterada no decorrer do vídeo, é um requisito do método da factorização como processo SFM. Com a estimação do movimento 2D apresentada, nomeadamente na selecção dos pontos característicos, várias situações tornam-se não rígidas. Uma luz/sombra forte é o melhor exemplo de uma não-rigidez pois resulta no seguimento de pontos característicos cujo movimento 2D não depende da forma do objecto. A rigidez não é uma característica local e tem de ser relacionada com o meio. Em [3] é estudado este problema.

De notar que o método da factorização funciona sem o *tracking* pois não necessita do movimento para conseguir a reconstrução da forma 3D. Assim, a continuidade temporal da sequência de imagens é apenas necessário para conseguir um bom seguimento, i.e., uma boa correspondência do mesmo ponto característico em duas imagens.

Uma boa reconstrução 3D exige um bom conhecimento do objecto e do seu movimento 2D no vídeo. Quanto maior o número de pontos característicos e de imagens melhor o resultado da minimização das Eq. 3.11, Eq. 3.12 e menor a sensibilidade ao ruído. O movimento câmara-objecto implica que partes do objecto se tornem visíveis e outras desapareçam, existindo assim informação acerca do objecto que é desconhecida. Na secção 3.4 descreve-se a utilização da factorização matricial com informação desconhecida.

3.4. Factorização com Dados Desconhecidos

3.4.1. Introdução

Ao processar o SFM pelo método da factorização de Tomassi e Kanade [2], i.e., ao resolver a Eq. 3.11, é necessário ter uma matriz de observação \mathbf{W} totalmente conhecida. Como visto na secção 2.3, isso é normalmente raro à saída do primeiro bloco da Fig. 1.1 sendo necessário resolver o problema SFM com informação desconhecida. Apresenta-se a solução de [7], que engloba a inicialização e dois métodos iterativos: o *Expectation-Maximization* (EM) e o *Two-Steps* (TS).

O problema das oclusões é resolvido dividindo o bloco B da Fig. 1.1 em dois, obtendo-se assim a Fig. 1.2. O funcionamento do último bloco da Fig. 1.2 – Método da Factorização – foi já descrito nas secções anteriores. O primeiro bloco – Tratamento de Oclusões – reduz o problema SFM com informação desconhecida ao problema SFM, cuja solução descrevemos na secção anterior. Isso é conseguido obtendo uma estimativa da matriz de observação \mathbf{W} completa.

A Fig. 3.2 ilustra a matriz \mathbf{W} com informação desconhecida (\mathbf{W}_0) que é gerada pelos métodos de seguimento. As áreas escuras referem-se a dados conhecidos e áreas brancas a dados desconhecidos. Para facilitar o tratamento organizam-se as colunas da matriz \mathbf{W}_0 por início e fim de seguimento.



Fig. 3.2: Matriz de observação \mathbf{W} com informação desconhecida (\mathbf{W}_0). Organizam-se as colunas da matriz \mathbf{W}_0 por início e fim de seguimento.

Os métodos que se apresentam minimizam o funcional que mede a distância da matriz de observação conhecida e incompleta \mathbf{W}_0 à matriz de característica deficiente estimada nas posições conhecidas $\tilde{\mathbf{W}}$,

$$\min_{\tilde{\mathbf{W}} \in S_4} \left\| (\mathbf{W}_0 - \tilde{\mathbf{W}}) \odot \mathbf{M} \right\|_F \quad \text{Eq. 3.13}$$

A matriz \mathbf{M} tem tamanho $2F \times P$ e verifica $\mathbf{M} = [m_{ij}]$ tal que

$$m_{ij} = \begin{cases} 1, & \text{se } w_{0ij} \text{ conhecido} \\ 0, & \text{se } w_{0ij} \text{ desconhecido} \end{cases}$$

e \odot é o produto ponto-a-ponto, também conhecido por produto de Hadamard.

No Anexo B descrevem-se o algoritmo da estimativa inicial e os dois métodos iterativos EM e TS. Na implementação destes algoritmos utilizou-se o código original dos autores [7], disponível em [19].

3.4.2. Análise dos métodos de estimação de informação desconhecida

Os métodos iterativos EM e TS e estimativa inicial podem ser aplicados para estimar dados desconhecidos de matrizes de qualquer característica. Neste trabalho apenas é relevante o seu funcionamento com característica 4 de modo a estimar uma matriz de observação \mathbf{W} completa.

Como se mostra em [7], utilizando a estimativa inicial proposta, os métodos usualmente convergem em menos de 10 iterações. Os resultados são afectados pelo ruído, pela quantidade de dados conhecidos e pela forma da “mancha” de dados conhecidos na matriz \mathbf{W}_0 (cor escura na Fig. 3.2). O tempo de processamento do método TS é maior do que o do método EM.

A utilização destes dois métodos de estimação é importante pois apresentam diferentes erros de estimação – resultado da Eq. 3.13 – em diferentes casos. Pode também dar-se o caso de um dos métodos não convergir.

4. Identificação de Reencontros

4.1. Introdução

Existe reencontro de um ponto característico quando após um período de seguimento e outro de oclusão se verifica outro período de seguimento, como ilustrado na Fig. 2.6. Os métodos de seguimento propostos no capítulo 2, tal como a generalidade dos métodos de seguimento usuais, não identificam reencontros de pontos característicos. Estes métodos consideram um ponto novo quando um reencontro ocorre. Pretende-se que o reencontro seja identificado para que os dois pontos característicos erradamente considerados no seguimento se tornem apenas um. Este é o problema abordado neste capítulo.

Uma boa reconstrução 3D exige um bom conhecimento do objecto no vídeo, i.e., ter informação relativa ao maior número de ângulos de vista do objecto no vídeo. Isto implica que, num vídeo considerado bom para a reconstrução 3D, alguns pontos têm vários períodos de seguimento, sendo habitual a ocorrência de reencontros. Além disso, um ponto característico pode deixar de ser seguido devido aos métodos de seguimento não serem totalmente fiáveis.

O estudo do reencontro de um ponto característico é importante pois otimiza o desempenho do SFM. Isto porque (i) melhora a estimação da informação desconhecida, feita no primeiro bloco da Fig. 1.2, e (ii) permite que o cálculo da estrutura 3D seja mais rigoroso. Ao invés de estimar dois períodos de oclusão de dois pontos característicos diferentes, cada um com um período de seguimento, torna-se apenas necessário tratar um ponto com dois ou mais períodos de seguimento. No segundo bloco da Fig. 1.2, a identificação dos reencontros permite que o cálculo da estrutura 3D seja mais rigorosa, assente em informação mais precisa. Principalmente na forma 3D resultante, a não identificação de um reencontro implica que sejam calculados dois pontos característicos muito próximos que na realidade o são mesmo. Para ilustrar a importância da identificação dos reencontros é realizada a experiência descrita na secção 5.5.

Deseja-se então encontrar uma matriz $\mathbf{W}_{0,R}$ que melhor caracterize os dados extraídos do vídeo (matriz \mathbf{W}_0). Assim, o Tratamento de Oclusões – segundo bloco da Fig. 1.2 – engloba duas acções: Tratamento dos Reencontros e Estimar a Informação Desconhecida, como ilustra a Fig. 4.1. A segunda acção foi já tratada na secção 3.4.

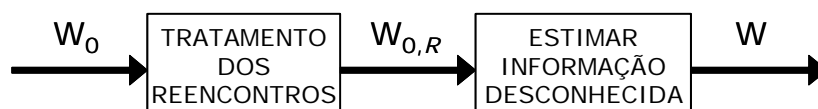


Fig. 4.1: Tratamento das Oclusões

Como ilustração considere-se um objecto a rodar. Existem pontos que inicialmente eram visíveis, ficaram oclusos e voltam a reaparecer. Neste exemplo as matrizes \mathbf{W}_0 e $\mathbf{W}_{0,R}$ têm o aspecto das observadas na Fig. 4.2, em que a cor mais clara identifica os reencontros. A Fig. 4.2(a) ilustra a matriz \mathbf{W}_0 que é calculada pelos métodos de seguimento, em que os reencontros levaram a que se considera-se dois pontos. A Fig. 4.2(b) ilustra a matriz $\mathbf{W}_{0,R}$

com os reencontros corrigidos. Com a identificação de reencontros tratada neste capítulo pretende-se transformar a matriz \mathbf{W}_0 da Fig. 4.2(a) na matriz $\mathbf{W}_{0,R}$ da Fig. 4.2(b).

Para ilustrar os pontos com vários períodos de seguimento, considera-se o 1º ponto reencontrado no 13º, o 2º reencontrado no 14º e o 3º reencontrado no 15º. O reencontro do 6º ponto com o 12º ilustra uma falha do método de seguimento, que por vezes não consegue corresponder o mesmo ponto em duas imagens consecutivas.

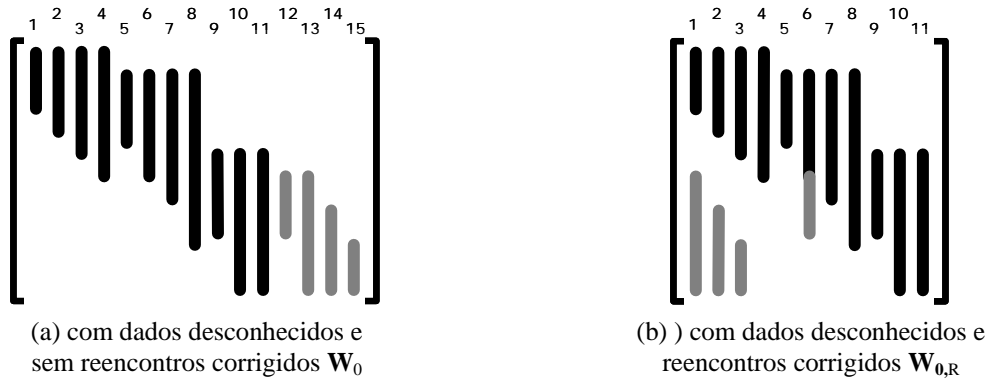


Fig. 4.2: Matriz de observação com informação desconhecida e reencontros (identificados pela cor clara)

Qualquer matriz de observação com reencontros, corrigidos ($\mathbf{W}_{0,R}$) ou não (\mathbf{W}_0), continua a ter característica 4. Isto porque se mantém o formalismo descrito no capítulo 3.

Trabalhos Relacionados

Existem outros métodos SFM que utilizam o conceito de reencontro de pontos característicos. Utilizam apenas algumas imagens com diferentes ângulos de vista do objecto, selecciona os pontos característicos em cada e estabelece os reencontros. Esses métodos, difíceis de resolver, são denominados por *matching* na literatura anglo-saxónica. Reparar que nesse caso não é feito o *tracking* dos pontos característicos numa sequência de imagens mas sim correspondidos em instantes dispersos no tempo – os pontos são reencontrados.

Identificar Reencontros Possíveis

O primeiro passo para tratamento dos reencontros é o de identificar os reencontros possíveis. Isto é feito por inspecção à matriz \mathbf{W}_0 , que colecciona as coordenadas de cada ponto característico em cada imagem. A Fig. 4.3 ilustra como identificar e criar um possível reencontro. As áreas escuras referem-se a dados conhecidos, sendo a cor mais clara os reencontros possíveis, e as áreas brancas a dados desconhecidos. Para facilitar o tratamento ordenam-se as colunas da matriz \mathbf{W}_0 por início e fim de seguimento. Para identificar um possível reencontro encontram-se colunas em que o início de seguimento seja pelo menos uma posição abaixo do fim de seguimento do ponto considerado, como se observa na Fig. 4.3(a). Cria-se assim uma lista dos reencontros possíveis com pares do tipo $(p_{1^\circ \text{seguimento}}, p_{2^\circ \text{seguimento}})$ com os índices das colunas do primeiro e segundo períodos de seguimento. Na Fig. 4.3(b) observa-se que ao juntar os dados conhecidos de duas colunas cria-se um reencontro.

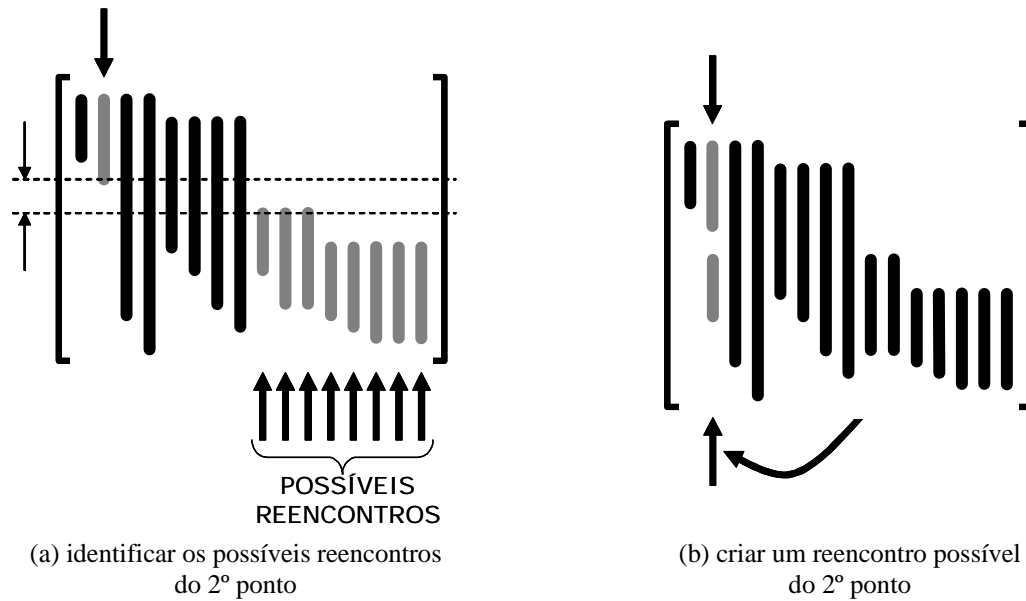


Fig. 4.3: Reencontros possíveis na matriz W_0

Para analisar os reencontros possíveis e concluir acerca da sua validade introduzem-se dois métodos: *Inspecção da Forma* e *Análise do Custo*. Estes são tratados de seguida.

4.2. Inspecção da Forma

4.2.1. Descrição do Método

Apresenta-se o método de identificação de reencontros *Inspecção da Forma*.

Da não identificação de um reencontro resulta que na forma 3D estimada existem dois pontos característicos muito próximos. Como concluído no capítulo 2, nos métodos de *tracking* propostos um ponto característico resulta em tantos pontos característicos quantos os seus períodos de seguimento. Sendo estes pontos característicos na realidade apenas um, se a reconstrução da forma 3D for eficiente estes têm a mesma posição ou estão muito próximos. O método *Inspecção da Forma* utiliza este facto para avaliar os reencontros possíveis.

A Fig. 4.4 ilustra o algoritmo, em que três blocos foram já descritos: Identificar Possíveis Reencontros, Estimar Dados Desconhecidos e Método da Factorização.

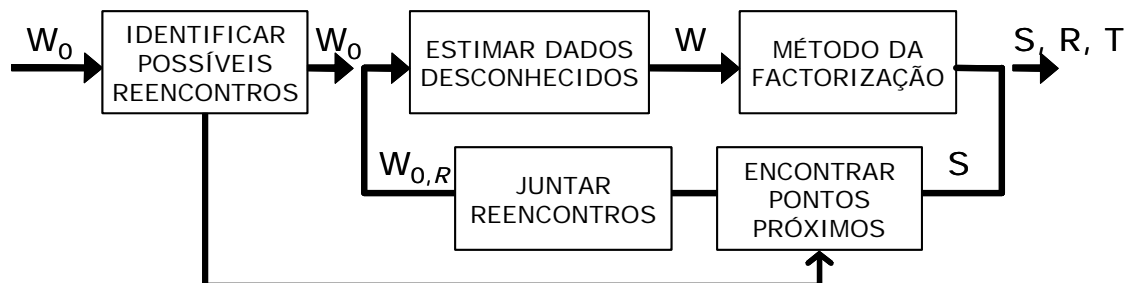


Fig. 4.4: Método da Inspecção da Forma

De notar que, na ausência de ruído, o ponto e os seus reencontros teriam as mesmas coordenadas após o cálculo do SFM. Isso geralmente não acontece devido ao ruído (ver secção 6.1), que influi negativamente nos resultados quer da estimação de matrizes completas

(ver secção 3.4) quer da factorização. Assim, o ponto e os seus reencontros aparecem na forma 3D com alguma distância.

O método Inspeção da Forma depende então da eficiência do processo SFM inicial para identificar correctamente os reencontros pela distância.

Distância de Reencontro

Define-se a distância de reencontro D como a distância máxima a que é considerado o reencontro. No bloco “Encontrar Pontos Próximos” da Fig. 4.4 é medida a distância d entre os pares de pontos característicos identificados no primeiro bloco como possíveis reencontros. Se esta distância $d(p_{1^{\circ} \text{seguimento}}, p_{2^{\circ} \text{seguimento}})$ for menor do que a distância de reencontro D , como na Eq. 4.1, é considerado o reencontro.

$$d(p_{1^{\circ} \text{seguimento}}, p_{2^{\circ} \text{seguimento}}) < D \quad \text{Eq. 4.1}$$

4.2.2. Cálculo da Distância de Reencontro

Pretende-se encontrar uma gama de valores válidos para a distância de decisão dos reencontros. Uma distância de reencontros grande beneficia a criação dos reencontros, enquanto que um valor baixo mantém o número de pontos extraídos do seguimento.

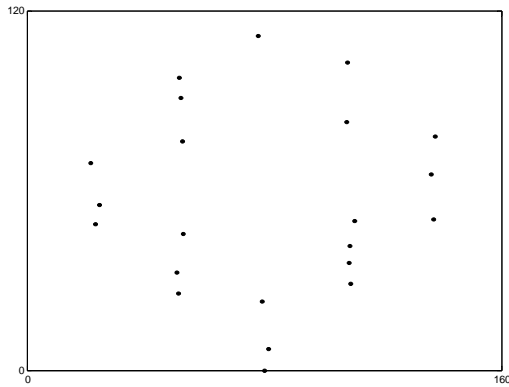
A escolha da distância de reencontro D não é fácil pois depende da relação entre a profundidade relativa do objecto e a sua distância ao plano da imagem. Como visto na secção 3.2, estes dois dados não são conhecidos devido à utilização do modelo de projecção ortogonal.

De notar que a eficiência do método da factorização [2] depende da boa aproximação do modelo de projecção ortogonal, o que implica desprezar o efeito da perspectiva. Para isso a profundidade relativa do objecto é muito menor do que a distância entre o plano da imagem e o objecto, como visto na secção 3.3.

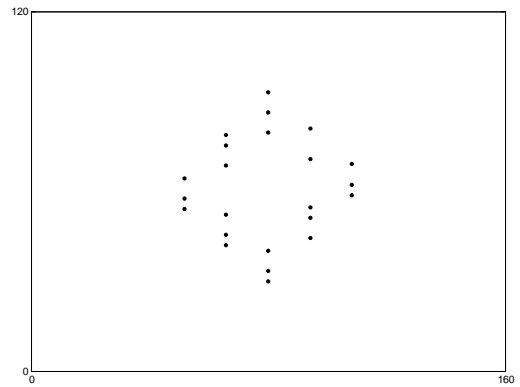
A Fig. 4.5 ilustra a dependência da distância de reencontro D da relação entre a profundidade relativa do objecto e a sua distância ao plano da imagem. Utilizou-se uma matriz de observação com dados artificiais, com 70.3% de dados conhecidos e 3 reencontros. De notar que na imagem de cada caso das Fig. 4.5(a1, a2) não são visíveis todos os pontos pois foram os apagados ao impor informação desconhecida.

Considera-se o mesmo objecto e o mesmo tamanho de imagem. No exemplo da Fig. 4.5 a profundidade relativa mantém-se e mostra que a distância de reencontro D é maior quando a distância entre o plano da imagem e o objecto é menor. De outro modo a distância entre quaisquer pontos característicos do objecto seria menor do que D , resultando em reencontros errados. Assim, o valor de D deve diminuir de (a) para (b) pois a distância entre o plano da imagem e o objecto é maior.

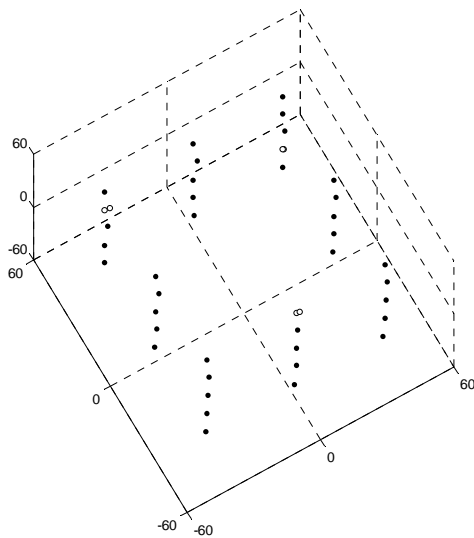
Deve então existir um valor de D diferente para cada reconstrução 3D, mas não existe a informação necessária para o calcular com os métodos propostos neste trabalho. Assim, o método Inspeção da Forma necessita da interacção do utilizador, que corrige o valor de D observando os resultados. A utilização de métodos que tornem o cálculo da distância de reencontro D possível permitirá aumentar a eficácia do método *Inspeção da Forma*.



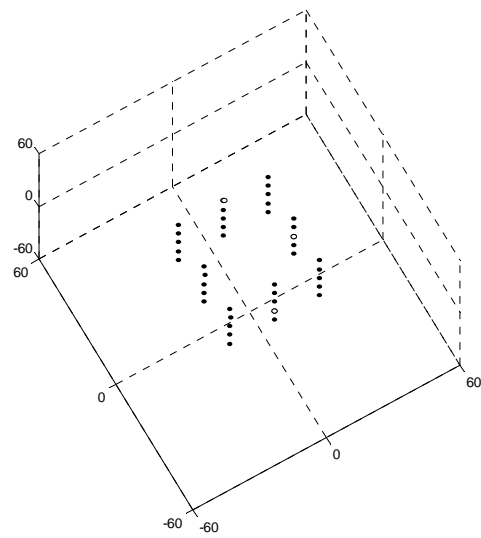
(a1) imagem do vídeo



(a2) imagem do vídeo



(b1) reconstrução 3D onde os círculos identificam os reencontros



(b2) reconstrução 3D onde os círculos identificam os reencontros

Fig. 4.5: Escolha da distância de decisão do reencontro. Esta distância depende da relação entre a profundidade relativa do objecto e a sua distância ao plano da imagem. Neste exemplo a distância entre o plano da imagem e o objecto é maior em (a) do que em (b).

Valor Utilizado na Implementação

O valor da distância de reencontro é escolhido atendendo ao método de selecção de pontos característicos proposto na secção 2.2. No último passo deste método são eliminados os pontos próximos entre si e apenas seleccionado o ponto característico considerado melhor desse conjunto. Para decidir que os pontos são próximos entre si é definida uma região quadrada R_2 com lado $L(R_2)$ que os envolve. É coerente que a distância de decisão utilizada em 2D para considerar pontos próximos entre si no plano da imagem seja a utilizada como distância de decisão em 3D. Assim, $D = L(R_2) = 21 \text{ pixels}$. De notar que este valor é escolhido qualitativamente, pelo observador, por observação do vídeo.

Devido à ambiguidade na definição da distância de decisão, interessa procurar um método de identificação de reencontros mais fiável. Na próxima secção apresenta-se o método *Análise de Custo*.

4.3. Análise do Custo

4.3.1. Descrição do Método

Apresenta-se o método de identificação de reencontros *Análise do Custo*.

A Fig. 4.6 ilustra o procedimento de Descrição de Comprimento Mínimo – *Minimum Description Length* (MDL) na literatura anglo-saxónica – que aplica o conceito de Verossimilhança Penalizada – *Penalized Likelihood* (PL) na literatura anglo-saxónica –, onde se inspira este método. Neste método, a avaliação dos reencontros é feita em função do seu custo. É relevante aplicar conhecimentos de MDL pois pretende-se a melhor representação do objecto no vídeo com o menor número de pontos.

Ao resolver o problema de estimar a informação desconhecida, assume-se que o número de pontos é conhecido. Isto não é o caso quando se sabe existir possíveis reencontros. A identificação de reencontros é assim um problema de selecção de modelos. Por modelo entende-se uma matriz $\mathbf{W}_{0,R}$ com oclusões, i.e., com dados desconhecidos, e com reencontros. Um modelo caracteriza-se pelo número de pontos e pelos reencontros criados. A Fig. 4.6 ilustra o processo de escolha do modelo, em que o conjunto de dados iniciais identifica a matriz \mathbf{W}_0 e cada modelo uma matriz $\mathbf{W}_{0,R}$.

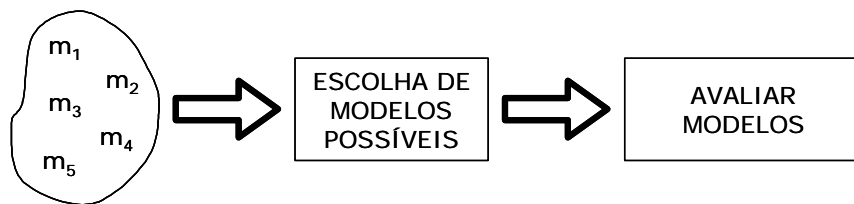


Fig. 4.6: Processo de escolha do modelo (extraído de [12])

O segundo bloco da Fig. 4.6 representa uma acção já descrita acima: identificar reencontros possíveis. As matrizes $\mathbf{W}_{0,R}$ possíveis são todos os reencontros mais as combinações de dois ou mais reencontros, existindo também o modelo sem reencontros ($\mathbf{W}_{0,R} = \mathbf{W}_0$). O conjunto dos modelos possíveis tem M elementos.

A Fig. 4.7 ilustra o método da *Análise do Custo* na identificação dos reencontros.

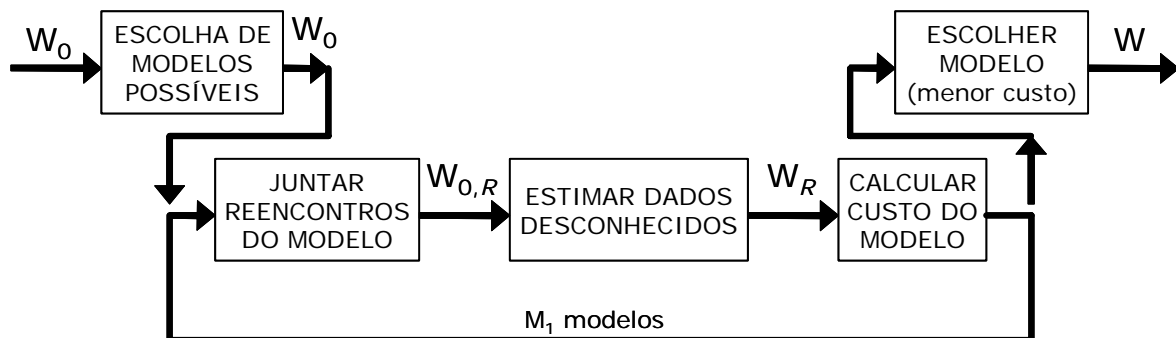


Fig. 4.7: Método da Análise do Custo

É necessário avaliar apenas os M_1 modelos que têm 1 reencontro criado e o modelo sem reencontros. De notar que $(M_1 + 1)$ é normalmente muito menor do que M . Depois de

concluir quais os reencontros a criar, é escolhido o modelo que os reúne dentro do conjunto dos M modelos possíveis.

O último bloco da Fig. 4.7 – Avaliar Modelos – decide quais reencontros criar. Escolhem-se os reencontros que apresentam um custo c menor do que o custo de não fazer reencontros c_0 , i.e., do que o custo do modelo sem reencontros ($\mathbf{W}_{0,R} = \mathbf{W}_0$).

Custo do Modelo

O princípio MDL [13, 14] conduz a uma soma de duas parcelas: uma que quantifica a semelhança do modelo (*likelihood* na literatura anglo-saxónica) ao sinal ou amostras a descrever; outra que quantifica o tamanho da descrição. Funções de custo que contemplam estes dois termos são habitualmente designadas por PL [9]. No Anexo A analisam-se a expressão MDL e várias expressões para o cálculo do custo. De seguida apresenta-se a expressão escolhida.

Assim, o custo de um modelo c pesa a aproximação aos dados conhecidos e a complexidade do modelo. É dado por

$$c = e + \alpha \cdot P \quad \text{Eq. 4.2}$$

em que e é o erro, por ponto, entre o modelo estimado e a observação, P o número de pontos e α o factor de custo.

O erro de estimação e é dado pela Eq. 4.3, em que $\tilde{\mathbf{W}}_{0,R}$ é a matriz completa estimada e $W_{conhecidos}$ é o número de coordenadas conhecidas. A matriz $\tilde{\mathbf{W}}_{0,R}$ resulta dos métodos de estimação de matrizes completas tratados na secção 3.4.

$$e = \frac{\|(\mathbf{W}_{0,R} - \tilde{\mathbf{W}}_{0,R}) \odot \mathbf{M}\|_F}{W_{conhecidos}} \quad \text{Eq. 4.3}$$

A medida de semelhança utilizada é $(-e)$, que é sempre negativa. É tanto maior quanto melhor for a aproximação aos dados conhecidos.

O factor de custo α define o peso da complexidade da solução, i.e., do número de pontos P que descrevem o objecto. A secção 4.3.3 trata como obter valores para o factor de custo.

Na próxima secção descreve-se a implementação do método *Análise do Custo*.

4.3.2. Implementação do método

A estimação da informação desconhecida no segundo bloco da Fig. 4.7 é feita com os algoritmos iterativos EM e TS descritos na secção 3.4.

Identificar Reencontros Possíveis

Encontra-se os reencontros possíveis como descrito na secção 4.1. Obtém-se assim uma lista com pares do tipo $(p_{1^\circ \text{seguinte}}, p_{2^\circ \text{seguinte}})$ com os índices das colunas do primeiro e segundo períodos de seguimento.

Primeiro Cálculo (custo do modelo sem reencontros)

Inicia-se o cálculo do custo com o modelo sem reencontros criados ($\mathbf{W}_{0,R} = \mathbf{W}_0$). Como não tem reencontros para juntar no primeiro bloco da Fig. 4.7, inicia-se o segundo bloco calculando-se a estimativa inicial como na secção 3.4.2. A estimativa inicial do modelo sem

reencontros $\mathbf{W}_0^{(0)}$ é guardada, bem como \mathbf{A} e \mathbf{B} , os respectivos espaços das linhas e colunas. De seguida executam-se os métodos EM e TS escolhendo-se o menor erro de estimação e_0 . O custo c_0 deste modelo é calculado como na Eq. 4.4, em que P_0 identifica o número de pontos, i.e., de colunas de \mathbf{W}_0 .

$$c_0 = e_0 + \alpha P_0 \quad \text{Eq. 4.4}$$

Cálculos Seguintes (custo dos modelos com um reencontro)

Cria-se o reencontro do modelo como na Fig. 4.3(b). No cálculo dos modelos com um reencontro são preparadas inicializações para os métodos iterativos EM e TS a partir da estimativa inicial do modelo sem reencontros $\mathbf{W}_0^{(0)}$. Para isso é retirada a coluna $p_{2^\circ \text{seguimento}}$ na matriz do espaço das colunas \mathbf{B} guardada. Desta maneira o algoritmo apenas necessita de calcular a estimativa inicial uma vez, com uma óbvia diminuição do custo computacional. O custo do reencontro $c(p_{1^\circ \text{seguimento}}, p_{2^\circ \text{seguimento}})$ é calculado por

$$c(p_{1^\circ \text{seguimento}}, p_{2^\circ \text{seguimento}}) = e(p_{1^\circ \text{seguimento}}, p_{2^\circ \text{seguimento}}) + \alpha(P_0 - 1) \quad \text{Eq. 4.5}$$

Escolher Reencontros a Corrigir

Os custos $c(p_{1^\circ \text{seguimento}}, p_{2^\circ \text{seguimento}})$ são comparados com o custo do modelo sem reencontros c_0 de modo a decidir a validade de cada reencontro. O reencontro é criado em \mathbf{W}_R se verificar

$$c(p_{1^\circ \text{seguimento}}, p_{2^\circ \text{seguimento}}) < c_0 \quad \text{Eq. 4.6}$$

Corrigir os Reencontros

Os pares de reencontros a criar são armazenados numa lista. Ordenam-se os índices $p_{1^\circ \text{seguimento}}$ desta lista por ordem decrescente. De notar que se tem sempre

$$p_{1^\circ \text{seguimento}} > p_{2^\circ \text{seguimento}}$$

devido a organizar-se as colunas da matriz \mathbf{W}_0 por início e fim de seguimento, como ilustrado na Fig. 3.2.

Inicia-se a correcção dos reencontros em \mathbf{W}_R pelo início da lista dos pares de reencontros a criar. Assegura-se assim que o método *Análise do Custo* consegue identificar e corrigir múltiplos reencontros, i.e., pontos característicos com vários períodos de seguimento.

4.3.3. Cálculo do Factor de Custo

Pretende-se encontrar uma gama de valores válidos para o factor de custo. Como visto acima, o factor de custo α define o peso do número de pontos P da solução em comparação com o erro de estimação e . Um factor de custo elevado beneficia a criação dos reencontros, enquanto que um valor baixo mantém o número de pontos extraídos do seguimento.

O factor de custo está dependente do método de estimação da informação desconhecida utilizado e da forma como é calculado o erro de estimação. Os métodos EM e TS têm normalmente erros de estimação muito próximos e por isso são tratados em conjunto. O erro de estimação utilizado é calculado como na Eq. 4.3 mas existem outras maneiras. Por exemplo, o quinto valor singular pode ser utilizado para quantificar o erro da aproximação à característica 4.

Com o fim de determinar o factor de custo realizaram-se experiências para analisar o efeito dos reencontros no erro da estimação da matriz de observação completa. A Fig. 4.8(a) ilustra as conclusões retiradas das experiências descritas no Anexo D. O erro sem reencontros e_0 foi já definido acima, e_{errado} é o menor erro dos reencontros errados e e_{certo} o maior erro dos reencontros certos.

A Fig. 4.8(b) mostra a transformação pretendida com a definição do custo de reencontro. O custo sem reencontros c_0 foi já definido acima e é dado pela Eq. 4.4, c_{errado} é o menor custo dos reencontros errados e c_{certo} o maior custo dos reencontros certos. Enquanto que pelo erro na Fig. 4.8(a) não é possível aferir a validade dos reencontros, pelo custo na Fig. 4.8(b) identificam-se quais os reencontros certos.

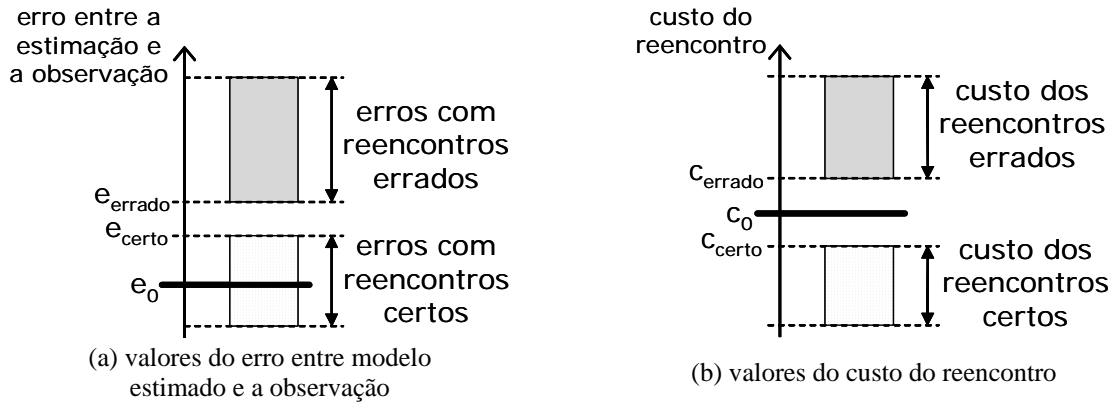


Fig. 4.8: Transformação pretendida com a utilização do custo de reencontro
Em (a) não é possível avaliar correctamente quais os reencontros a corrigir. Em (b) é possível a identificação de reencontros.

Como se pode observar na Fig. 4.8(a) o erro com reencontros certos é menor do que com reencontros errados, existindo um intervalo que os separa. Também se observa que o erro sem reencontros é menor do que com reencontros errados. O valor do erro com reencontros certos é próximo do erro sem reencontros. A Eq. 4.7 formaliza as conclusões do Anexo D, onde também se verifica a sua validade. Estas conclusões foram obtidas para uma gama pertinente de valores de desvio padrão do ruído, de percentagem de informação conhecida, de número de imagens, de número de pontos e de número de reencontros. Também no Anexo D é aferida a validade das relações

$$\begin{aligned} e_{errado} &> e_0 \\ e_{certo} &> e_{errado} \\ e_{certo} &\sim e_0 \end{aligned} \quad \text{Eq. 4.7}$$

Aplicando a definição do custo de um reencontro da Eq. 4.2 obtêm-se

$$\begin{aligned} c_{certo} &= e_{certo} + \alpha(P_0 - 1) \\ c_{errado} &= e_{errado} + \alpha(P_0 - 1) \end{aligned} \quad \text{Eq. 4.8}$$

Pretende-se calcular o factor de custo para que as relações da Eq. 4.9 sejam válidas.

$$c_{certo} < c_0 < c_{errado} \quad \text{Eq. 4.9}$$

A partir da Eq. 4.4, Eq. 4.8 e Eq. 4.9 obtêm-se as expressões Eq. 4.10 para o máximo e mínimo do factor de custo.

$$\alpha > e_{certo} - e_0$$

$$\alpha < e_{errado} - e_0$$
Eq. 4.10

Para determinar uma gama de valores possíveis para o factor de custo realizam-se experiências com várias matrizes de observação. É necessário conhecer quantas vezes a Eq. 4.9 se verifica aplicando valores dentro do intervalo da Eq. 4.10. Este é um dos objectivos do Anexo C, onde também se escolhe o factor de custo com melhores resultados nos dados reunidos.

A Fig. 4.9 mostra a evolução do número de testes em que se verifica a Eq. 4.9 em função do factor de custo. É possível também observar o andamento de outras relações entre os vários custos definidos acima.

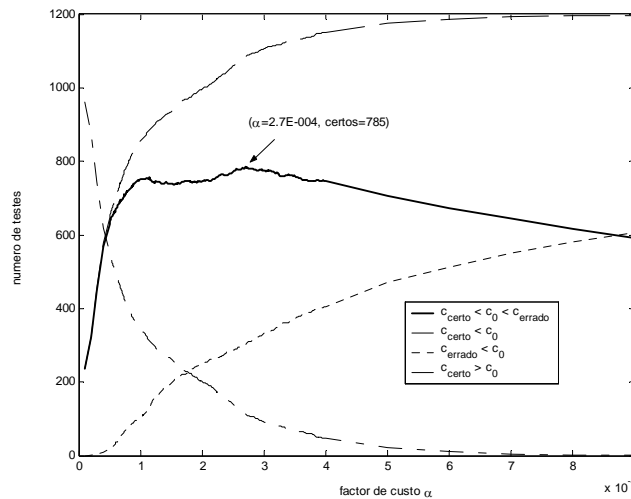


Fig. 4.9: Evolução do número de reencontros identificados em do factor de custo α . É escolhido o valor que maximiza o número de reencontros identificados.

As experiências foram realizadas com matrizes de observação de característica 4 gradadas a partir de dados aleatórios. Estes dados correspondem a coordenadas de pontos em imagens de tamanho $160 \times 120 \text{ pixels}$.

Pela Fig. 4.9 conclui-se que o intervalo

$$\alpha \in [0.1 \times 10^{-4}; 0.4 \times 10^{-4}]$$

apresenta bons resultados nas experiências efectuadas. De notar que esta análise de desempenho é feita com os dados reunidos no anexo D, em que se testa uma gama pertinente de valores de desvio padrão do ruído, de percentagem de informação conhecida, de número de imagens, de número de pontos e de número de reencontros.

Valor Utilizado na Implementação

É escolhido o valor $\alpha = 2.7 \times 10^{-4}$ pois apresenta melhores resultados. Este valor será utilizado nas próximas experiências em se aplicar o método *Análise do Custo*. Tendo sido feitos 2700 testes, o valor escolhido permitiu identificar todos os reencontros em 784 testes (29.0%). Destes 2700 a premissa necessária ao método *Análise do Custo* da Eq. 4.7 verificou-se 1152 vezes. É neste conjunto de testes que a transformação erro/custo se mostra relevante, mostrando-se válida em 68.1% com este factor de custo.

4.4. Análise dos Métodos de Identificação de Reencontros

O método *Inspecção da Forma* é dependente do desempenho do processo SFM que é feito inicialmente, bem como de uma distância de decisão. A definição do elemento de decisão – distância entre pontos característicos – não é fácil, dependendo fortemente de algo que é desconhecido, a distância do plano da imagem ao objecto.

O método *Análise de Custo* é um procedimento inspirado no princípio MDL. Procura-se escolher uma solução que minimize o erro de estimação dos dados desconhecidos e penalize soluções complexas, i.e., com mais pontos característicos.

Os dois métodos de identificação de reencontros dependem do desempenho dos métodos de estimação de informação desconhecida. Esta estimação depende do ruído, da percentagem de dados conhecidos e da forma da “mancha” de dados conhecidos na matriz \mathbf{W}_0 .

Os métodos *Inspecção da Forma* e *Análise do Custo* estão dependentes da definição de um parâmetro: a distância de reencontro e o factor de custo, respectivamente. Parâmetros com valores elevados beneficiam a criação dos reencontros, enquanto que valores baixos mantêm o número de pontos extraídos do seguimento. Enquanto no *Inspecção da Forma* a distância de reencontro é definida qualitativamente com a interacção do utilizador, no *Análise do Custo* é possível descobrir uma gama de valores com bons resultados.

É possível identificar pontos característicos com vários períodos de seguimento com os dois métodos.

Com o propósito de comparar os métodos *Inspecção da Forma* e *Análise de Custo* foram feitas várias experiências, descritas na secção 5.5. A sensibilidade ao ruído é estudada na secção 6.4.

5. Resultados Experimentais

5.1. Introdução

Nas secções seguintes expõem-se e analisam-se os resultados das experiências efectuadas. Nestas experiências assume-se quando necessário um valor de ruído realista. O estudo do efeito do ruído é realizado no capítulo 6.

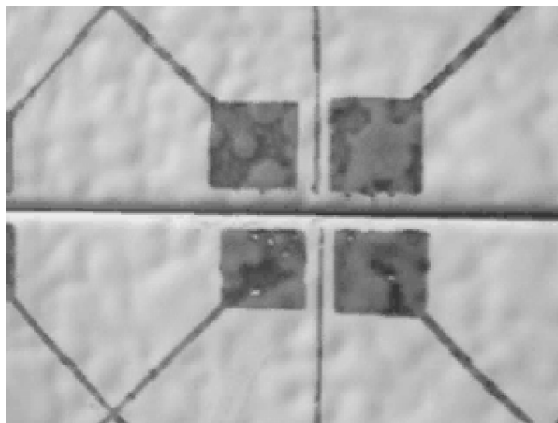
Na secção 5.2 analisa-se cada passo do algoritmo de *Seleção* proposto na secção 2.2. O método *Seguidor de Pontos Característicos* é analisado na secção 5.3. Para o estudo do bloco B da Fig. 1.1 nas secções 5.4, 5.5 e 5.6 são gerados dados artificiais de modo a ter controlo sobre as coordenadas dos pontos característicos e do tipo de ruído. A importância da identificação de reencontros no processo SFM é tratada na secção 5.5. Na secção 5.6 são analisados os métodos *Inspecção da Forma* e *Análise do Custo*. Na secção 5.6 apresenta-se uma experiência de modelação 3D a partir de vídeo, que ilustra o benefício da identificação de reencontros.

5.2. Seleção de Pontos Característicos

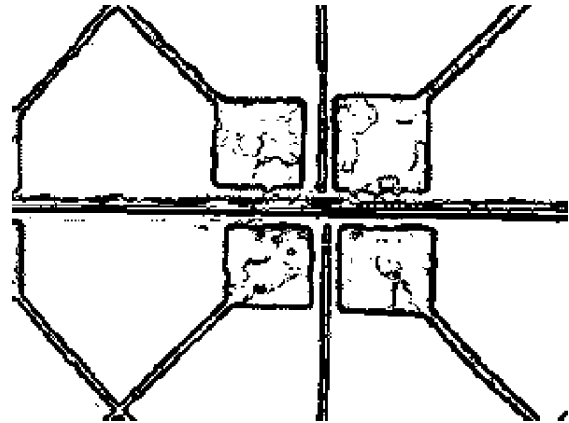
Nesta secção analisa-se a necessidade e eficácia de cada um dos três passos do processo de seleção proposto na secção 2.2, que está ilustrado na Fig. 2.3. Nestas experiências utiliza-se a imagem da Fig. 5.1(a), de tamanho 320×240 pixels.

Gradiente

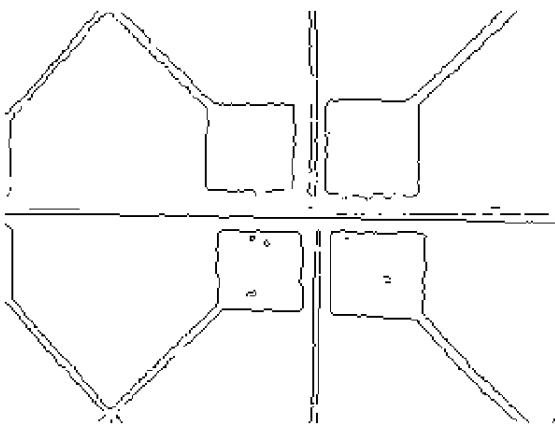
Este passo reduz os pontos a processar considerando apenas os pontos pertencentes aos contornos visíveis da imagem. Observa-se o resultado do passo Gradiente na seleção de pontos na imagem da Fig. 5.1(a). Na Fig. 5.1(b) são escolhidos os pontos correspondentes aos máximos relativos do gradiente do padrão de intensidades, enquanto que nas Fig. 5.1(c,d) define-se um valor mínimo para a norma do gradiente, respectivamente $\xi_1 = 10$ e $\xi_1 = 30$. Sabe-se que o máximo da norma do gradiente do padrão de intensidades desta imagem é $\max(\|\nabla \mathbf{I}\|) = 79.4$.



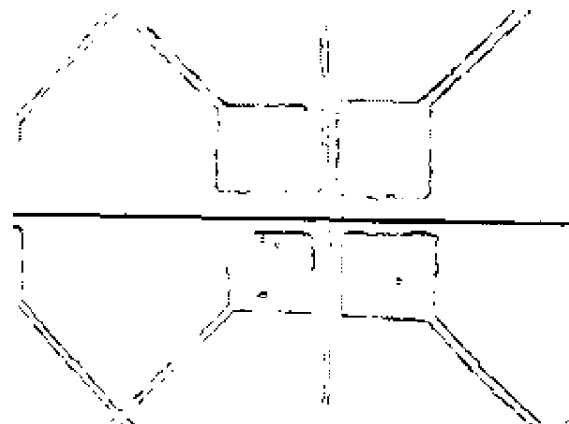
(a) imagem a processar



(c) definido um valor mínimo para a norma do gradiente ($\xi_1 = 10$)



(b) pelos máximos relativos do gradiente



(d) definido um valor mínimo para a norma do gradiente ($\xi_1 = 30$)

Fig. 5.1: Resultados do passo Gradiente do processo de *Seleção*.
 Observa-se a quantidade de pontos escolhidos por este passo na
 imagem (a) para vários valores do gradiente.

Verifica-se que o número de pontos escolhidos pelo passo Gradiente cresce quanto menor o *threshold* ξ_1 . A escolha de ξ_1 depende da imagem a tratar, logo é feita por observação da imagem e do resultado.

A Tab. 5.1 apresenta valores do tempo de processamento do algoritmo de seleção para os valores de ξ testados. É também possível comparar o número de pontos que o passo Gradiente escolhe e os que resistem à triagem do passo Abertura. Nestas experiências aplica-se o passo Abertura em regiões de lado $L(R_1) = 5 \text{ pixels}$.

Tab. 5.1: Resultados do passo Gradiente do processo de selecção

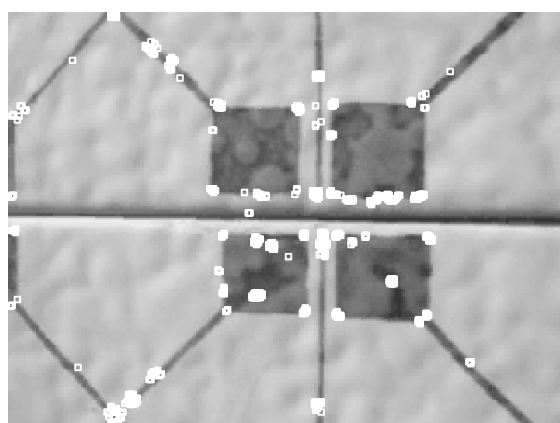
ξ_2	Método do passo Gradiente	Pontos escolhidos em Gradiente	Pontos escolhidos em Abertura	Tempo de Processamento [s]
8	máximos relativos	2465	223	2,4
	$\xi_1 = 30$	2167	144	2,2
	$\xi_1 = 20$	5499	484	3,8
	$\xi_1 = 10$	11128	1605	6,5
	sem passo Gradiente	76800	46932	221,1
2	máximos relativos	2465	17	2,5
	$\xi_1 = 30$	2167	6	2,6
	$\xi_1 = 20$	5499	38	3,6
	$\xi_1 = 10$	11128	166	6,2
	sem passo Gradiente	76800	9370	40,3

Observa-se que, além do número de pontos, também o tempo de processamento de todo o algoritmo de selecção cresce quanto menor o valor de *threshold* ξ_1 .

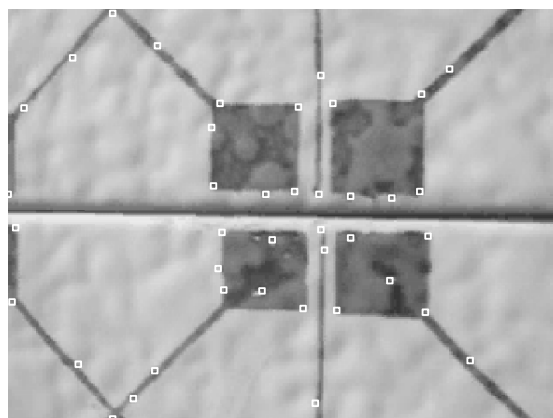
Abertura

Este é o passo que distingue os pontos característicos dos restantes e por isso é sempre utilizado. Selecciona o ponto central de regiões onde padrão de intensidades varia tanto verticalmente como horizontalmente. Pelas experiências sobre os passos Gradiente e Abertura, cujos resultados se resumem na Tab. 5.1, conclui-se que se deve utilizar um valor ξ_1 que permita contornos carregados como na Fig. 5.1(c) e um valor ξ_2 não muito pequeno que possibilite boas escolhas nestes contornos.

Observam-se na Fig. 5.2(a) os pontos seleccionados pelos passos Gradiente e Abertura. Foram utilizados os valores $\xi_1 = 20$ $\xi_2 = 8$ e $L(R_1) = 5\text{ pixels}$.



(a) pontos seleccionados pelos passos Gradiente ($\xi_1=20$) e Abertura ($\xi_2=8$, $L(R_1)=5\text{ pixels}$)



(b) processo Selecção completo (3 passos) Gradiente ($\xi_1=20$), Abertura ($\xi_2=8$, $L(R_1)=5\text{ pixels}$) Eliminar Pontos Próximos ($L(R_2)=21\text{ pixels}$)

Fig. 5.2: Resultados do processo de *Seleção* com e sem o passo Eliminar Pontos Próximos

Eliminar Pontos Próximos

No passo Eliminar Pontos Próximos são eliminados os pontos próximos entre si e apenas seleccionado o ponto característico considerado melhor desse conjunto. Para decidir que os pontos são próximos entre si é definida uma região que os envolve. A escolha do tamanho desta região depende do tamanho do objecto e da definição, i.e., número de *pixels*, do vídeo. Então a escolha do tamanho da região varia em cada situação, sendo feita por observação das imagens a tratar.

Os pontos escolhidos na Fig. 5.2(a) estão muito próximos entre si, dificultando o seguimento nas imagens seguintes. Na Fig. 5.2(b) observa-se o resultado final do processo de selecção proposto, aplicando o passo Eliminar Pontos Próximos numa região de lado $L(R_2) = 21 \text{ pixels}$.

As experiências efectuadas mostram que aplicando uma região pequena no passo Eliminar Pontos Próximos são seleccionados muitos pontos. Por serem demasiado próximos alguns destes pontos escolhidos são difíceis de seguir nas imagens seguintes e não são tão relevantes para a definição do objecto como outros por perto. Aplicando uma região grande significa escolher poucos pontos. Ao eliminarem-se muitos pontos corre-se o risco de se perder os que realmente interessam ao processo de seguimento. Por serem demasiado espaçados entre si, os pontos não definem correctamente o objecto.

Conclusão

O desempenho do algoritmo de Selecção proposto é observado na Fig. 5.2(b). Foram seleccionados $P = 39$ pontos característicos numa imagem de tamanho $320 \times 240 \text{ pixels}$, demorando 3,54 segundos.

Verifica-se que os pontos escolhidos fazem parte de cantos ou de texturas acentuadas. O passo Abertura não escolheu os pontos pertencentes à recta horizontal que percorre a imagem.

Verifica-se ainda que a definição das regiões e valores de treshold devem ser feitas por observação e experimentação no vídeo a tratar. Só assim é possível a selecção de bons pontos característicos que sejam passíveis de seguimento.

5.3. Seguimento de Pontos Característicos

Para seguimento dos pontos característicos numa série de imagens utiliza-se o método *Seguidor de Pontos Característicos* proposto no capítulo 2, que está ilustrado na Fig. 2.5.

Na Fig. 5.3 observa-se o desempenho deste método numa sequência de $F = 10$ imagens e $P = 23$ pontos. Mostram-se apenas alguns dos pontos seguidos para ilustrar os vários tipos de seguimento de um ponto característico. É visível na Fig. 5.3(a) os pontos escolhidos na primeira imagem e nas Fig. 5.3(c,d) as trajectórias destes na sequência de imagens, existindo também pontos cujo seguimento se iniciou na 6ª imagem. Os pontos seleccionados na 1ª imagem são identificados por um círculo, os seleccionados na 6ª imagem por um ponto e os pontos seleccionados na 10ª imagem por quadrados.

Na Fig. 5.3(b) observa-se a mancha de dados conhecidos da matriz de observação, de tamanho 20×23 .

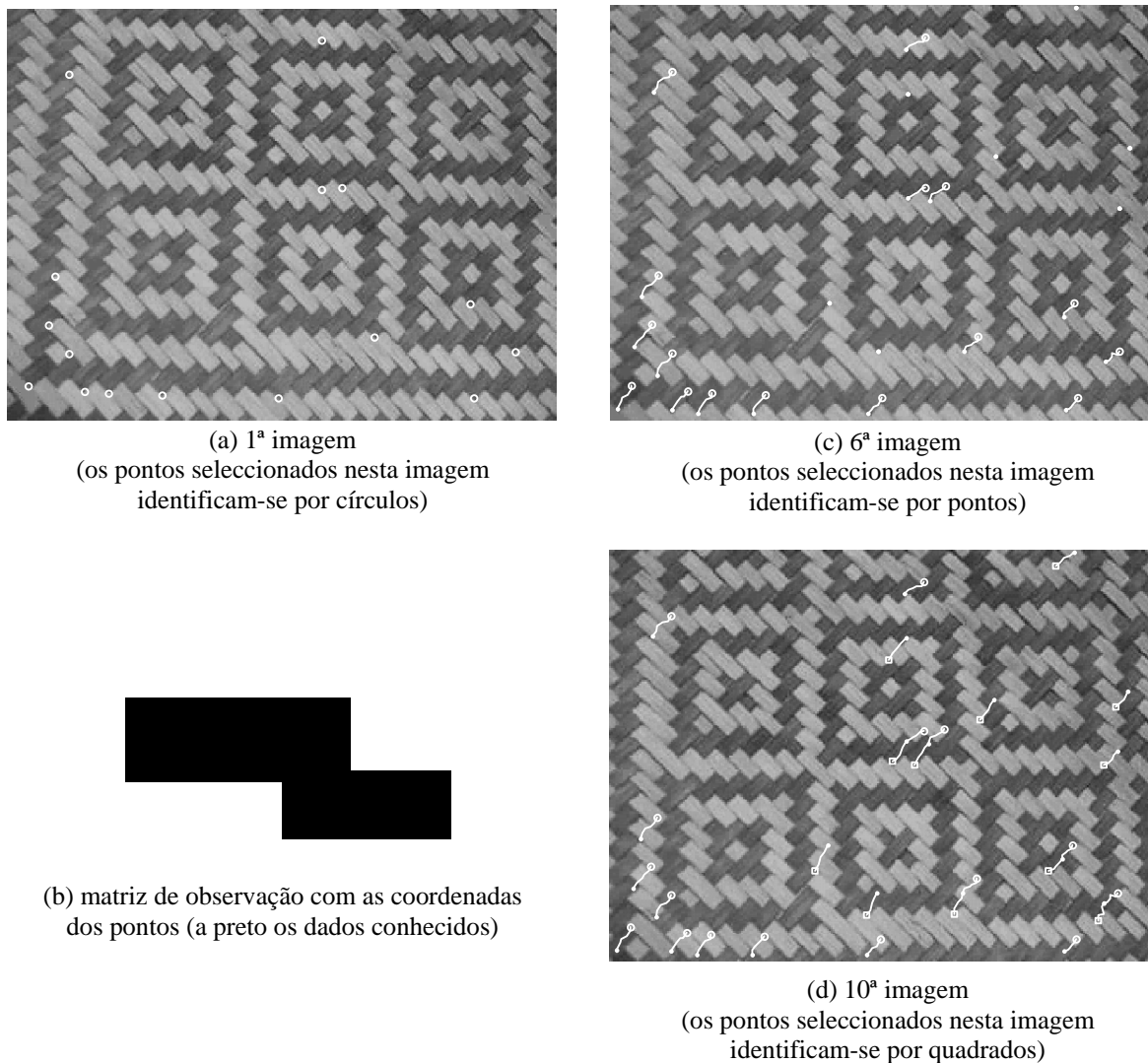


Fig. 5.3: Resultados do Seguidor de Pontos Característicos

Na experiência ilustrada na Fig. 5.3 é possível realçar as características dos vários tipos de seguimento de pontos característicos.

Seguido desde a 1ª à 10ª imagem

Estes pontos foram seguidos em toda a sequência de imagens. Demonstram o correcto desempenho do *Seguidor de Pontos Característicos*.

Seguido desde 1ª à 6ª imagem

Existem duas situações para isto acontecer: (i) o ponto desaparecer da imagem; ou (ii) o método falhar o seu seguimento.

Seguido desde a 6ª à 10ª imagem

Existem duas situações para isto acontecer: (i) o ponto aparecer pela primeira vez na imagem ou (ii) o método ter falhado o seu seguimento até esta imagem.

Ruído

Em alguns pontos a trajectória é uma linha suave. Noutros pontos, afectados pelo ruído, a linha da trajectória tem variações rápidas. Este ruído é devido a duas razões: (i) as imagens tem pouca qualidade e não permitem distinguir os pontos correctamente; e (ii) o seguimento não é perfeito.

Conclusão

Para que os pontos mal seguidos sejam poucos, calibra-se os parâmetros do método, de modo a que os pontos considerados importantes para a definição do objecto sejam bem seguidos no maior número de imagens possível. A definição das regiões e valores de threshold devem então ser feitas por observação e experimentação do vídeo a tratar, para obter o máximo de pontos seguidos que melhor definem o objecto.

5.4. Factorização Matricial

Nesta secção pretende-se exemplificar o desempenho do método da factorização como processo SFM com e sem informação desconhecida. Como descrito no capítulo 3, o algoritmo calcula uma nova coordenada dos pontos característicos para a transformação de 2D para 3D.

Nestas experiências utiliza-se um cubo de 40 pontos em 30 imagens de tamanho 160×120 pixels.

Dados Conhecidos

Na Fig. 5.4(a) observa-se uma imagem da matriz de observação \mathbf{W} totalmente conhecida e sem ruído. Na Fig. 5.4(b) observa-se a forma 3D resultante da reconstrução 3D.

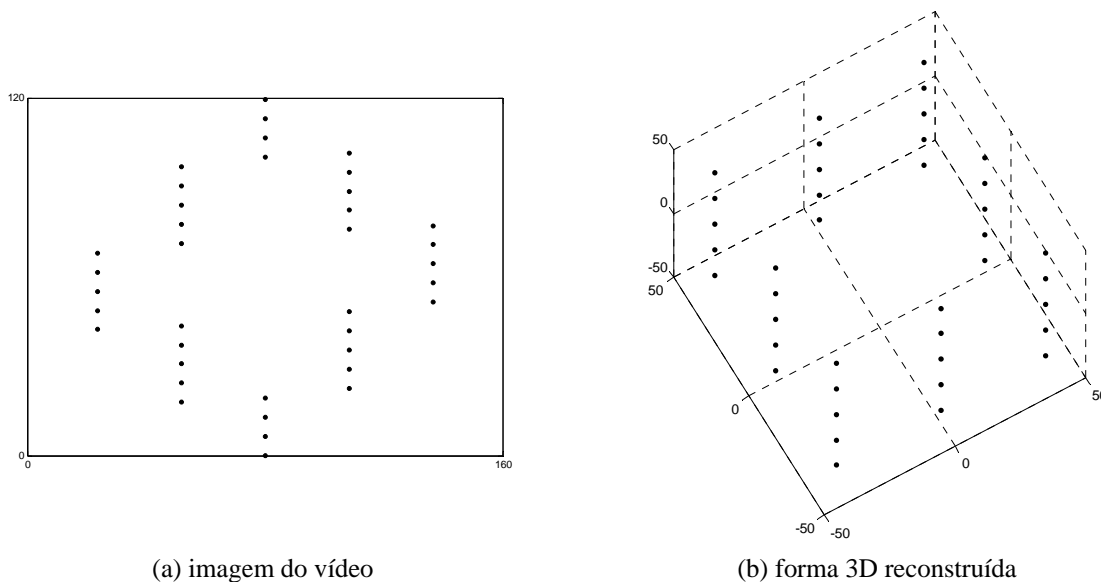


Fig. 5.4: Desempenho do método da factorização com dados conhecidos e sem ruído

Verifica-se que a reconstrução 3D com dados conhecidos e sem ruído é perfeita, i.e., sem erro, o que ilustra o correcto desempenho do método da factorização como processo SFM.

Dados Desconhecidos

Para resolver o problema SFM com informação desconhecida (oclusões) utilizam-se os métodos iterativos EM e TS [7]. Estes métodos calculam uma matriz de observação completa, estimando os dados desconhecidos.

A matriz de observação com dados desconhecidos \mathbf{W}_0 utilizada está ilustrada na Fig. 5.5(a) e resulta da introdução de oclusões na matriz \mathbf{W} da experiência anterior. São conhecidos 39.9% dos dados. A Fig. 5.5(b) mostra a evolução do erro entre a estimação e os dados conhecidos em cada iteração. Na Fig. 5.5(c) observa-se a forma 3D à saída do processo SFM.

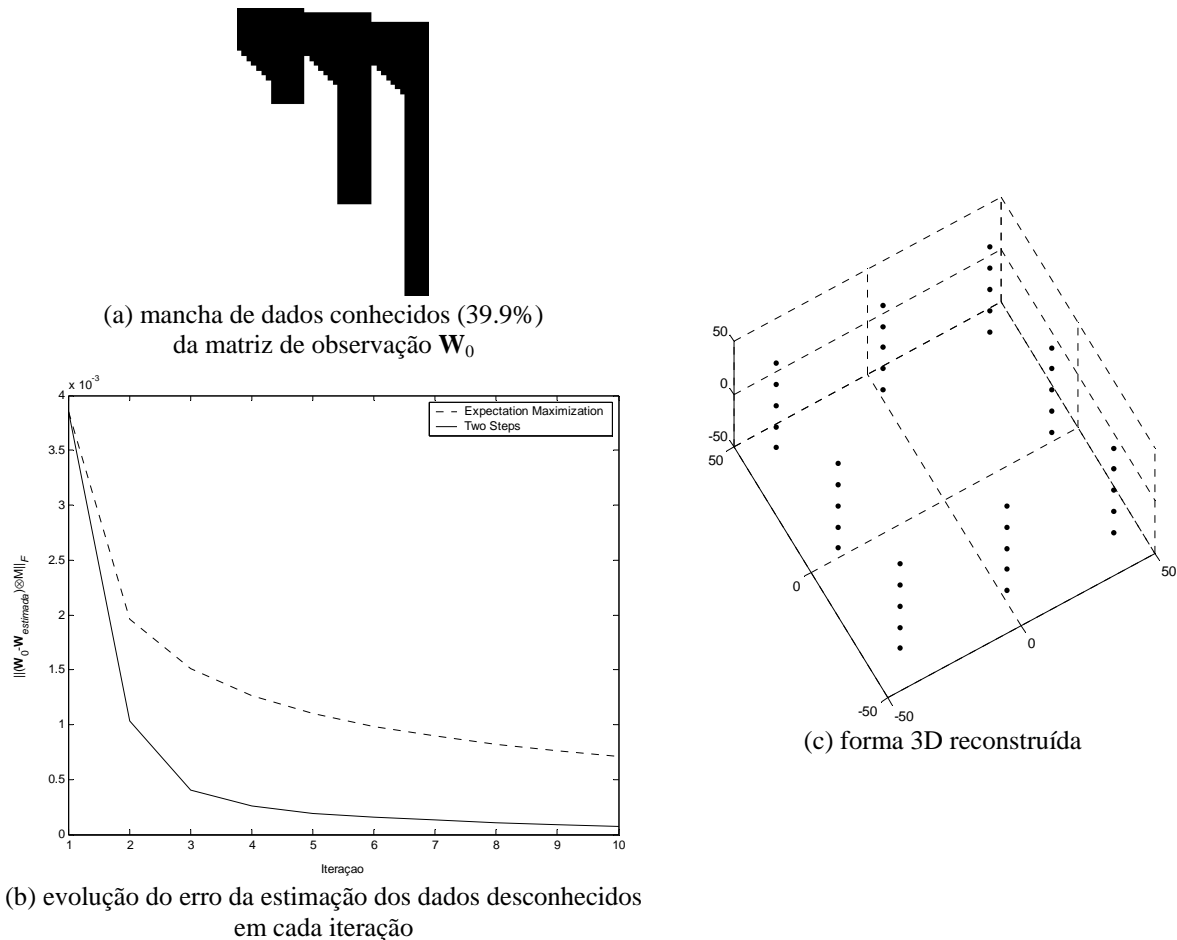


Fig. 5.5: Desempenho do método da factorização com informação desconhecida e sem ruído.

Como visto na secção 3.4, a cada iteração os métodos EM e TS minimizam a expressão do erro, visível na Fig. 5.5(b). A matriz \mathbf{M} é composta por 0's e 1's e identifica as coordenadas conhecidas.

Verifica-se o bom desempenho do processo SFM com informação desconhecida proposto. Neste caso sem ruído, a discrepância entre a entrada e a saída depende apenas desempenho dos métodos de estimação da informação desconhecida, EM e TS. Na ausência de ruído estes métodos apresentam um erro quase nulo. Com a introdução de ruído, estes métodos dependem também da quantidade de informação conhecida e da forma como os dados conhecidos se apresentam – mancha escura da Fig. 5.5(a).

Nesta experiência o erro por ponto entre a matriz de entrada \mathbf{W} (igual à da experiência anterior) e a matriz $\mathbf{W}_{estimada}$ estimada

$$\frac{\|\mathbf{W} - \mathbf{W}_{estimada}\|_F}{\sqrt{2F \cdot P}} = 2.5 \times 10^{-5} \text{ pixels}$$

A matriz de observação \mathbf{W}_{SFM} calculada com os resultados da reconstrução 3D – \mathbf{S}_{SFM} , \mathbf{R}_{SFM} e \mathbf{T}_{SFM} – é dada por

$$\mathbf{W}_{SFM} = \mathbf{R}_{SFM} \mathbf{S}_{SFM} + \mathbf{T}_{SFM} \times \mathbf{1}_{1 \times P} = [\mathbf{R}_{SFM} \mid \mathbf{T}_{SFM}] \begin{bmatrix} \mathbf{S}_{SFM} \\ \mathbf{1}_{1 \times P} \end{bmatrix} \quad \text{Eq. 4.11}$$

O erro por ponto entre a matriz de entrada \mathbf{W} e a matriz \mathbf{W}_{SFM} é

$$\frac{\|\mathbf{W} - \mathbf{W}_{SFM}\|_F}{\sqrt{2F \cdot P}} = 1.7 \times 10^{-5} \text{ pixels}$$

Mostra-se assim que o método da factorização melhora a estimação da matriz de observação com informação desconhecida \mathbf{W}_0 feita pelos métodos EM e TS. Na secção 6.2 é analisada esta resistência do método da factorização ao erro das coordenadas observadas (ruído).

5.5. Importância da Identificação de Reencontros

Pretende-se concluir acerca da importância da identificação de reencontros no processo SFM, nomeadamente a sua influência na estimação da informação desconhecida e na estimação da estrutura 3D. Para isso é realizada uma experiência onde se compara o resultado da reconstrução 3D nos casos com e sem reencontros. As conclusões apresentadas são gerais, i.e., válidas para todas as matrizes que permitam estimar correctamente a informação desconhecida pelos métodos descritos na secção 3.4. A experiência aqui descrita ilustra estas conclusões.

Na Fig. 5.6 observam-se as matrizes utilizadas. Geram-se matrizes relativas à forma do objecto (\mathbf{S}) e rotação (\mathbf{R}) e translação (\mathbf{T}) da câmara, criando-se de seguida uma matriz \mathbf{W} completa de característica 4. Nesta experiência utiliza-se um cubo de 40 pontos em 30 imagens de tamanho 160×120 pixels e é adicionado ruído com desvio padrão $\sigma = 3$ pixels. Impõem-se períodos de oclusão em vários pontos, formando assim a matriz sem reencontros $\mathbf{W}_{0,R}$ da Fig. 5.6(a). Introduzem-se novos pontos no final de $\mathbf{W}_{0,R}$ que são reencontros de pontos já seguidos, obtendo-se assim a matriz \mathbf{W}_0 da Fig. 5.6(b) com reencontros não corrigidos. A matriz $\mathbf{W}_{0,R}$ da Fig. 5.6(c) tem os reencontros corrigidos.

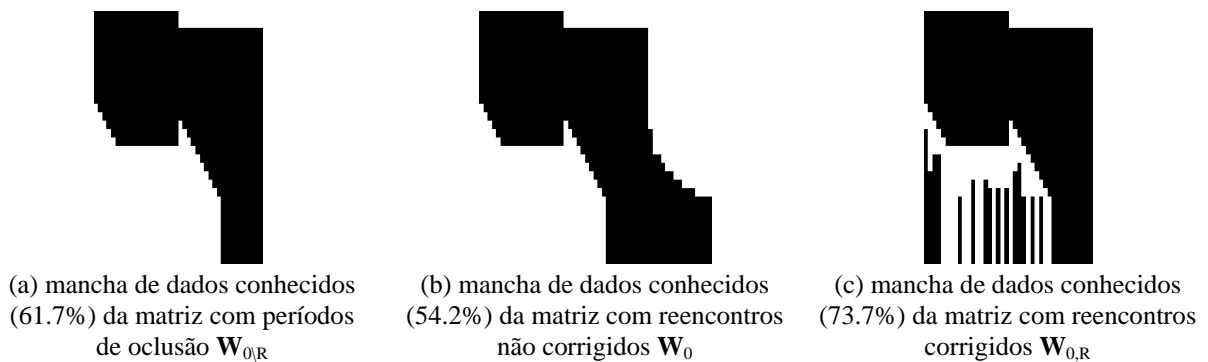


Fig. 5.6: Matrizes de observação utilizadas na experiência que ilustra a importância dos reencontros

Realiza-se o processo SFM com informação desconhecida com as três matrizes de observação descritas. Para isso são estimadas matrizes $\mathbf{W}_{estimada}$ completas pelos métodos EM e TS, antes de se aplicar o método de factorização.

Na Fig. 5.7 observa-se a discrepância dos resultados do processo SFM com as três matrizes de observação descritas. Esta discrepância é medida com o erro, por ponto, entre as matrizes \mathbf{W} , \mathbf{S} , \mathbf{R} e \mathbf{T} geradas à entrada e os resultados $\mathbf{W}_{estimada}$, \mathbf{W}_{SFM} , \mathbf{S}_{SFM} , \mathbf{R}_{SFM} e \mathbf{T}_{SFM} de cada caso. A matriz \mathbf{W}_{SFM} calculada como na Eq. 4.11 serve para comparação com a matriz de observação de entrada.

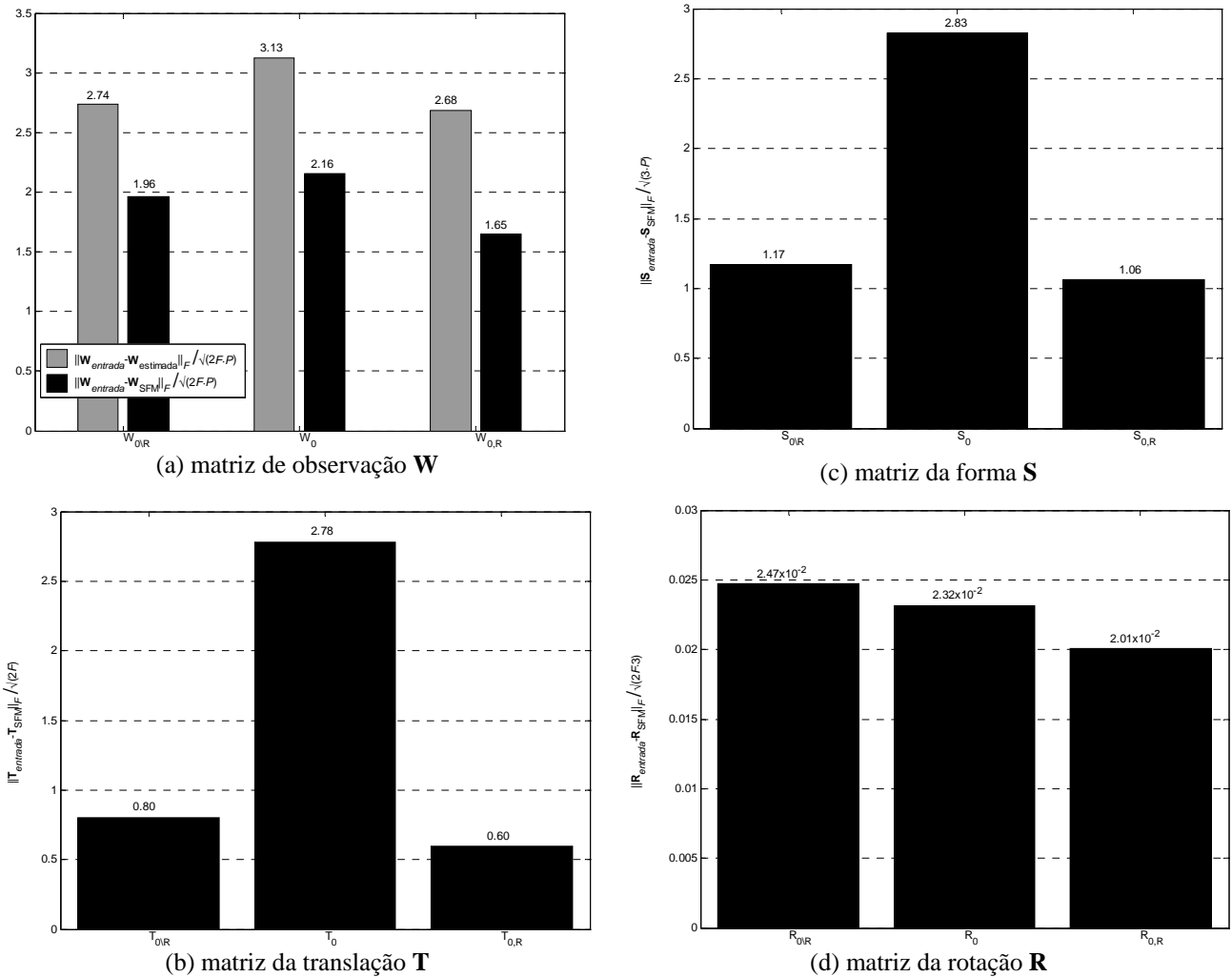


Fig. 5.7: Importância da identificação dos reencontros no processo SFM
 Comparação do erro, por ponto, dos resultados da estimação de matrizes completas e da estimação da estrutura 3D com e sem reencontros.

Observa-se que o facto de existirem reencontros não corrigidos dificulta quer a estimação de matrizes de observação completas quer a estimação da estrutura 3D. De notar que a análise do caso sem reencontros serve apenas para comparação, pois a ocorrência de reencontros é habitual, como visto na secção 4.1.

Como as colunas com reencontros têm normalmente poucos dados conhecidos, quer a estimação de dados quer a reconstrução 3D apresentam normalmente erro maior com reencontros não corrigidos (\mathbf{W}_0) do que sem reencontros ($\mathbf{W}_{0,R}$). Isto não acontece sempre, como observado na Fig. 5.7(d). Com reencontros, mesmo que não corrigidos, existe mais

informação, o que pode permitir que quer a estimação de dados quer a reconstrução 3D apresentem melhores resultados do que sem reencontros.

Comparando o caso sem reencontros $\mathbf{W}_{0,R}$ e o de reencontros corrigidos $\mathbf{W}_{0,R}$ conclui-se que os reencontros corrigidos permitem uma reconstrução 3D mais precisa, assente em informação mais precisa.

Verifica-se assim que a identificação dos reencontros é importante para o SFM na medida em que permite melhorar quer a estimação da informação desconhecida quer a estimação da estrutura 3D.

5.6. Métodos de Identificação de Reencontros

Analisa-se resultados de experiências com o propósito de comparar os dois métodos de identificação de reencontros: *Inspecção da Forma* e *Análise do Custo*. É exemplificado o desempenho, a sensibilidade à quantidade de informação conhecida, e quantificado o seu custo computacional de cada método. A sensibilidade ao ruído é estudada do capítulo 6.

5.6.1. Experiência Com Dados Artificiais

Para exemplificar o desempenho dos métodos de identificação de reencontros foi utilizada uma esfera de 81 pontos em 50 imagens de tamanho 160×120 pixels e é adicionado ruído com desvio padrão $\sigma = 1$ pixels. Foram criados 5 reencontros e a matriz de observação com reencontros não corrigidos \mathbf{W}_0 tem 71.4% de informação conhecida.

Na Fig. 5.8 observam-se a forma do objecto antes e depois da identificação de reencontros. Na Fig. 5.8(a) mostra-se a forma com reencontros não corrigidos, que estão identificados por círculos. Na Fig. 5.8(b) é visível a forma com todos os reencontros corrigidos.

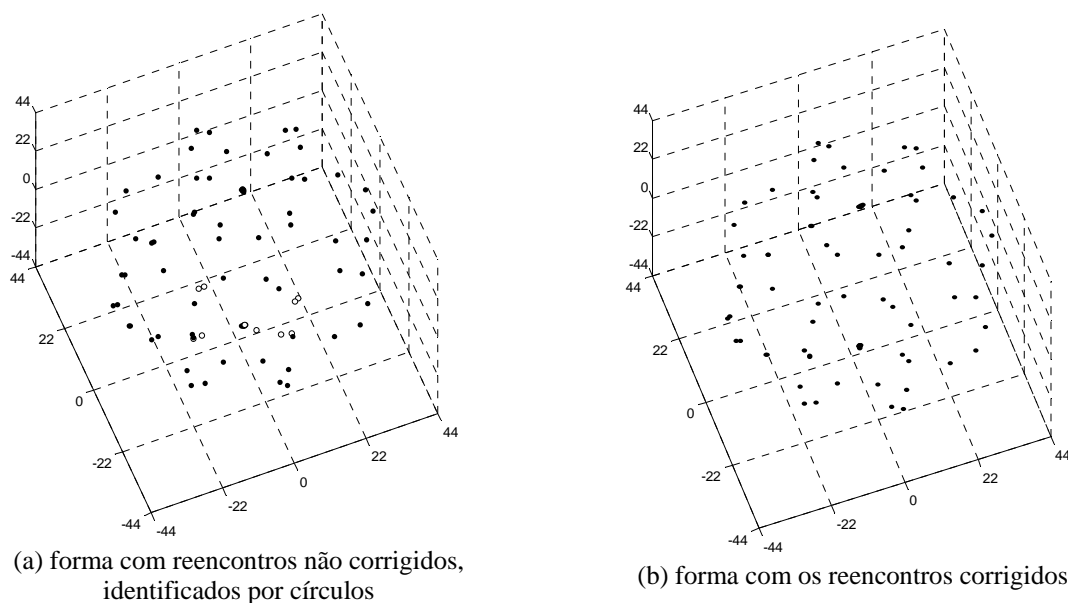


Fig. 5.8: Forma do objecto antes e depois da identificação de reencontros

O método *Análise do Custo* conseguiu encontrar todos os 5 reencontros, o que exemplifica a sua eficácia. Foi utilizado o factor de custo $\alpha = 2.7 \times 10^{-4}$.

O método *Inspecção da Forma* depende da interacção do utilizador para escolher a distância de reencontro correcta. Assim, é necessário experimentar vários valores de distância

de reencontro e analisar os resultados, nomeadamente a forma reconstruída. Na Tab. 5.2 resume-se o resultado da identificação de reencontros para várias distâncias de reencontro.

Tab. 5.2: Variação do número de reencontros descobertos com a distância de reencontro no método *Inspecção da Forma*

distância de reencontro D	número de reencontros descobertos certos	número de reencontros descobertos errados
21	3	2
9	4	1
3	5	0

De notar que estes resultados variam em cada vídeo.

5.6.2. Experiência Com Vídeo Real

Pretende-se ilustrar a sensibilidade dos métodos de identificação de reencontros propostos à quantidade de informação conhecida. Nesta experiência foi utilizado um vídeo real de uma pilha de discos em rotação.

Na Fig. 5.9 observa-se a trajectória dos 54 pontos característicos na sequência de 23 imagens. Nestes 54 pontos existem 18+18 (ponto+reencontro) situações de reencontro, correspondentes ao cantos de uma das faces do objecto.

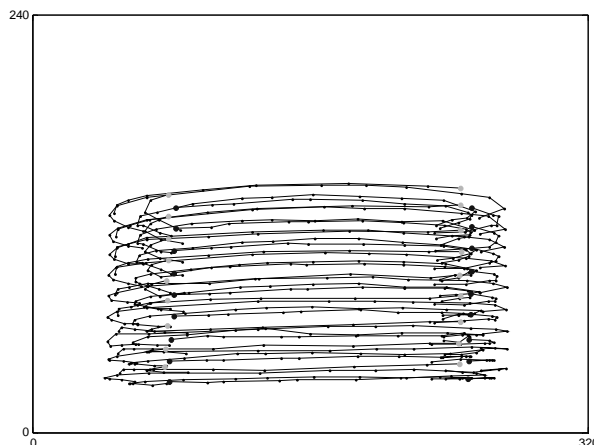
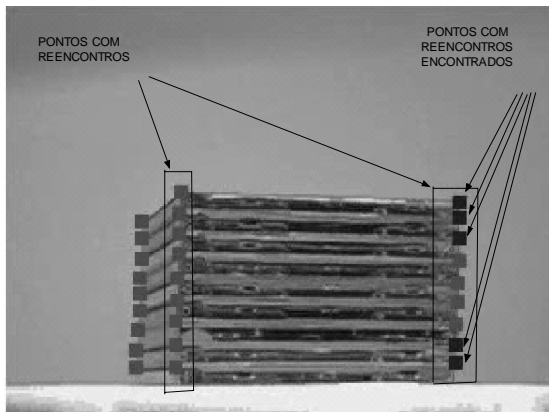


Fig. 5.9: Trajectórias dos pontos característicos no plano da imagem na experiência com a pilha de discos

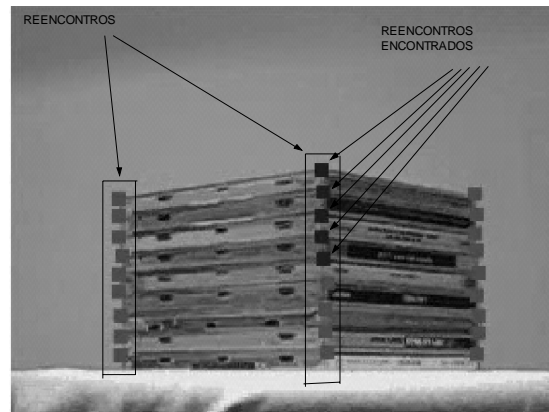
A matriz de observação sem reencontros \mathbf{W}_0 tem então tamanho $(2 \cdot 23) \times 54$ com 47.1% de dados conhecidos.

Como visto na anteriormente, o método *Inspecção da Forma* depende do desempenho do processo SFM inicial. Neste caso com pouca quantidade de informação, os métodos de estimação de matrizes completas não encontram uma boa aproximação dos dados desconhecidos, mesmo com 15 iterações. Assim, não é possível identificar os reencontros pela distância pois a forma 3D reconstruída é disforme.

O método *Análise do Custo* também é sensível à quantidade de informação. Consegue identificar 3 reencontros correctamente e 2 errados. Na Fig. 5.10(a) observam-se os pontos seleccionados na 2ª imagem, bem como os pontos com reencontros e os que o método assume com reencontros. A Fig. 5.10(b) apresenta os pontos na 21ª imagem, os reencontros existentes que deveriam ter sido identificados e os reencontros realmente encontrados pelo método *Análise do Custo*.



(a) 2ª imagem - pontos seleccionados identificados por quadrados; assinalam-se os pontos com reencontros e os encontrados pelo método *Análise do Custo*



(b) 21ª imagem - pontos seleccionados identificados por quadrados; assinalam-se os reencontros e os encontrados pelo método *Análise do Custo*

Fig. 5.10: Seguimento no vídeo, reencontros correctos e reencontros encontrados na experiência com a pilha de discos

Mesmo com os 3 reencontros correctos corrigidos, a forma 3D reconstruída continua bastante disforme, atingindo erros da ordem das centenas, não sendo por isso aqui apresentada.

Verificou-se que os métodos *Inspecção da Forma* e *Análise do Custo* são sensíveis à quantidade de informação. O método *Análise do Custo* consegue apresentar alguns resultados correctos, mas também introduz erros ao criar reencontros erradamente.

5.6.3. Custo Computacional

Com o propósito de comparar o custo computacional dos métodos *Inspecção da Forma* e *Análise do Custo*, foram realizadas experiências com matrizes de tamanho diferentes e medidos os tempos de processamento.

Foram geradas artificialmente matrizes de observação de imagens de tamanho 160×120 , com desvio padrão do ruído $\sigma = 1$ pixel, com 73% de informação conhecida e com 5 reencontros. A Tab. 5.3 expõe os resultados em função do número de pontos (sem reencontros) P , do número de imagens F e do número de reencontros possíveis M_1 . Na Fig. 5.11 observa-se a evolução do tempo de processamento em função de M_1 .

O tempo de processamento do método *Inspecção da Forma* depende do tamanho da matriz, sendo praticamente o tempo que demora a realizar a SFM sem identificar os reencontros.

O tempo de processamento do método *Análise do Custo* depende do tamanho da matriz de observação e do número de reencontros possível M_1 .

Tab. 5.3: Tempo de processamento dos métodos *Inspecção da Forma* e *Análise do Custo*

				Tempo de
--	--	--	--	----------

F	P	M_1	Processamento [s]	
			Análise do Custo	Inspeção da Forma
30	40	100	27.3	0.6
50	81	205	147.6	1.3
70	121	305	458.5	2.8
125	231	580	6056.6	23.6

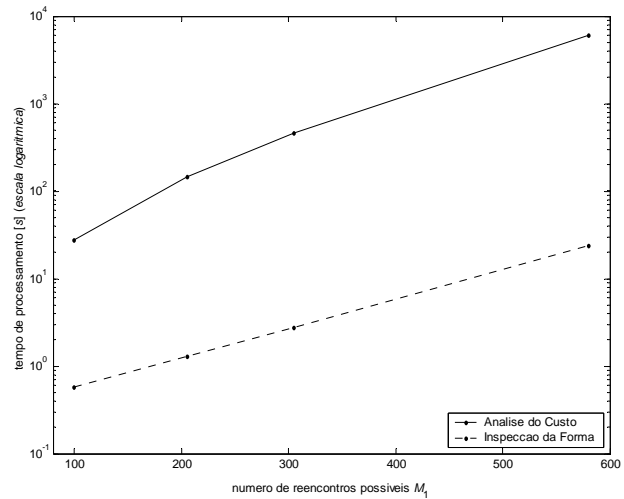


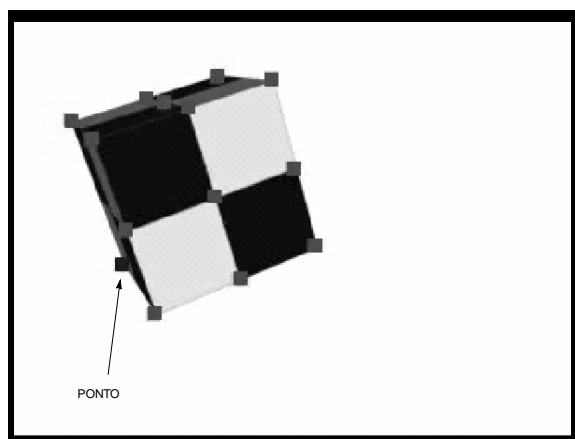
Fig. 5.11: Tempo de processamento dos métodos *Inspeção da Forma* e *Análise do Custo*

O tempo de processamento do método TS é maior do que o do método EM. É então possível reduzir o tempo de processamento dos métodos *Inspeção da Forma* e *Análise do Custo* ao utilizar apenas o método EM.

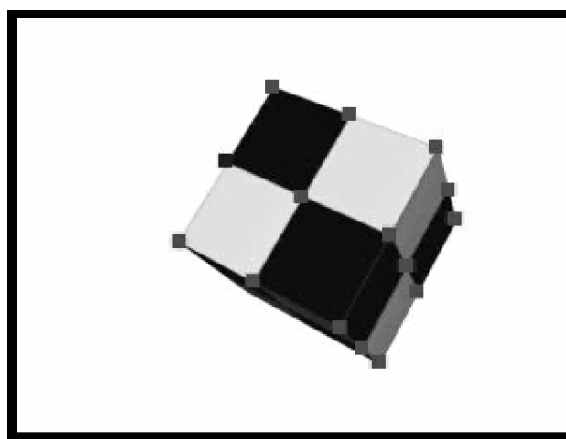
5.7. Modelação 3D a Partir de Vídeo

Para esta experiência gerou-se um vídeo por computador de modo a ter liberdade sobre o movimento do objecto. Utilizou-se um cubo numa sequência de 32 imagens, sendo possível identificar o objecto com 18 pontos característicos. Um destes pontos tem um reencontro.

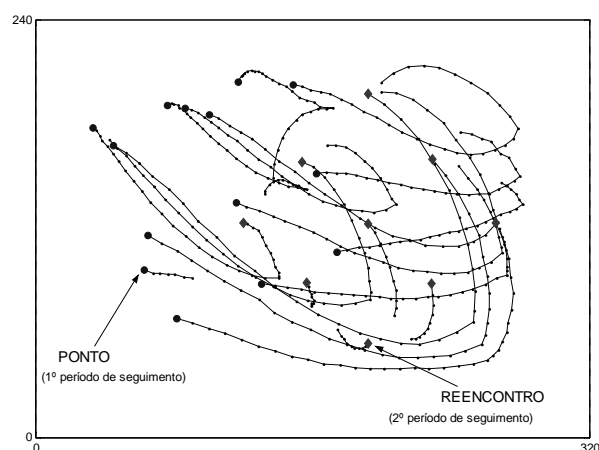
Na Fig. 5.12(a) observa-se os pontos escolhidos na 1ª imagem, na Fig. 5.12(c) os pontos da 15ª imagem e na Fig. 5.12(d) os da 26ª imagem. É identificado o ponto da Fig. 5.12(a) que tem um reencontro. Na Fig. 5.12(c) não se observa o ponto pois este está ocluído. Na Fig. 5.12(d) o ponto é novamente visível, i.e., foi reencontrado.



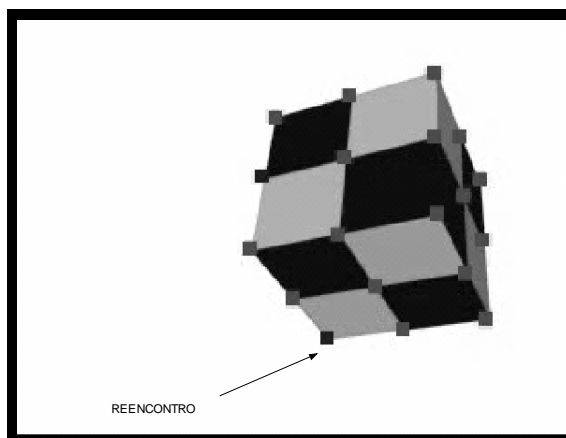
(a) 1ª imagem - pontos seleccionados identificados por quadrados; assinala-se o ponto que tem um reencontro



(c) 15ª imagem - pontos seleccionados identificados por quadrados; o ponto antes assinalado está ocluído



(b) trajectórias dos pontos característicos no plano da imagem



(d) 26ª imagem - pontos seleccionados identificados por quadrados; assinala-se o reencontro do ponto antes assinalado

Fig. 5.12: Seguimento no vídeo e reencontro na experiência com o cubo

Na Fig. 5.12(b) ilustram-se as trajectórias dos pontos característicos no plano da imagem. Estas trajectórias representam o movimento 2D extraído do vídeo, que resulta do *tracking*. Em particular, são visíveis os dois períodos de seguimento do ponto e do seu reencontro – ver Fig. 2.6 que ilustra a definição de reencontro.

Como ao realizar o seguimento não são identificados os reencontros, são considerados dois pontos que na realidade são o mesmo (ponto+reencontro). Assim, a matriz de observação que resulta do seguimento \mathbf{W}_0 , visível na Fig. 5.13(a), tem tamanho $(2 \cdot 32) \times 19$ com 77.5%

de dados conhecidos. Na Fig. 5.13(b) observa-se a matriz de observação com reencontros corrigidos $\mathbf{W}_{0,R}$, de tamanho $(2 \cdot 32) \times 18$ com 81.8% de dados conhecidos.

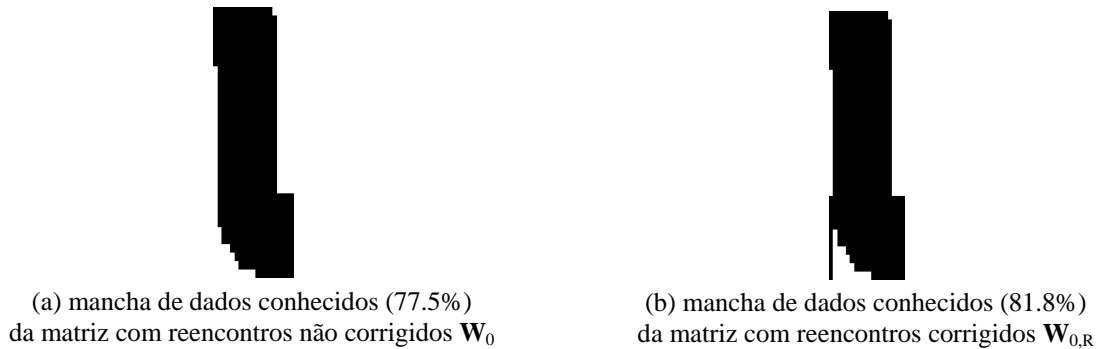


Fig. 5.13: Matrizes de observação da experiência com o cubo

Na Fig. 5.14 observam-se as formas 3D reconstruídas. A forma 3D, visível na Fig. 5.14(a), sem reencontros corrigidos é um cubo disforme dados. A forma 3D reconstruída a partir da informação com os reencontros corrigidos, visível na Fig. 5.14(b), representa a forma cúbica observada no vídeo.

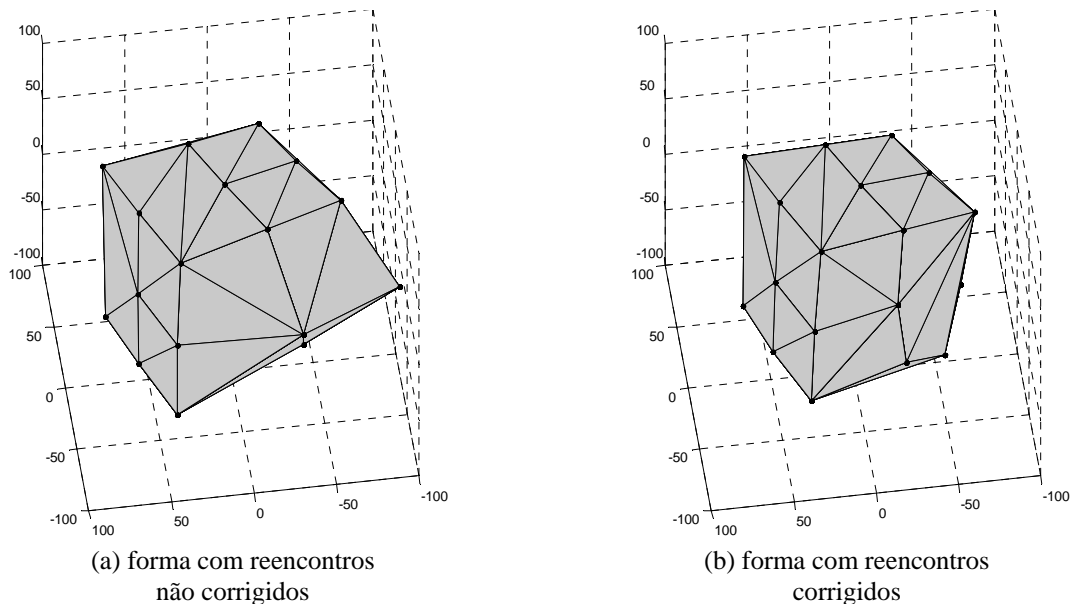


Fig. 5.14: Forma 3D reconstruída na experiência com o cubo

O método *Análise do Custo* consegue identificar o reencontro.

O método *Inspecção da Forma* não consegue identificar o reencontro. Isto porque a forma 3D calculada na primeira fase do algoritmo, visível na Fig. 5.14(a), é demasiado disforme para identificar o reencontro pela distância. Quaisquer valores de distância de reencontro resultam em erros.

Verifica-se a importância da identificação de reencontros, que neste exemplo permitiu uma reconstrução da forma 3D mais correcta. Exemplificou-se a eficiência do método *Análise do Custo* na identificação de reencontros com vídeo real. Verificou-se que o método *Inspecção da Forma* é dependente do desempenho do processo SFM inicial. Apresentou-se uma reconstrução 3D a partir de vídeo completa com identificação de reencontros.

6. Análise do Desempenho com Ruído

6.1. Introdução

O ruído presente na matriz de observação faz com que as coordenadas dos pontos característicos contêm erro. Este deve-se a três factores: a qualidade das imagens do vídeo, os algoritmos de seguimento e o modelo de projecção utilizado.

A qualidade das imagens do vídeo condiciona o desempenho dos algoritmos na selecção e seguimento dos pontos característicos no vídeo. Se as imagens apresentam o objecto distorcido, com “chuva” ou com “efeito de bloco” não é possível escolher pontos que caracterizem o objecto, nem encontra-los nas imagens posteriores. A definição do vídeo, i.e., o tamanho dos *pixels* condiciona a quantidade de informação por *pixel*, o que também condiciona a selecção dos pontos característicos.

Ao utilizar o modelo de projecção ortogonal são desprezados os movimentos ao longo do eixo óptico. Isso implica que no caso de existirem movimentos ao longo do eixo óptico a reconstrução 3D é influenciada. A Fig. 6.1 ilustra o modelo de projecção em perspectiva, semelhante à realidade. Ao comparar a Fig. 6.1 com a Fig. 3.1 é evidente a diferença das coordenadas no plano da imagem que os dois modelos de projecção permitem obter. Admite-se que esta diferença é desprezável e que os dois modelos de projecção são equivalentes considerando a profundidade relativa do objecto muito menor do que a sua distância ao plano da imagem.

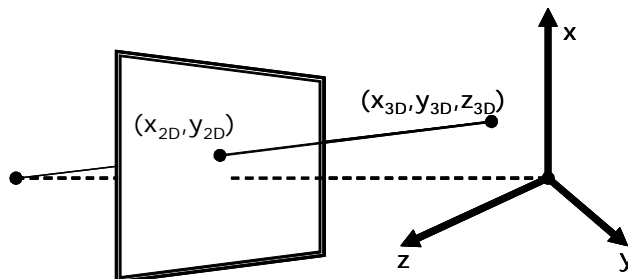


Fig. 6.1: Modelo de projecção em perspectiva

No estudo do ruído em SFM tem de se ter em conta que uma reconstrução 3D eficiente depende do número de imagens, se estas definem bem o objecto e se foi possível seguir os pontos característicos com sucesso. A imunidade do método da factorização ao ruído depende destes factores pois o efeito do ruído é combatido com um grande número de imagens com variedade de ângulos (matriz \mathbf{R}).

Para estudar o efeito do ruído utilizam-se matrizes de observação \mathbf{W} geradas artificialmente de modo a ter controlo sobre as coordenadas dos pontos característicos e do tipo de ruído. Assim, são criadas as matrizes da forma do objecto (\mathbf{S}) e da rotação (\mathbf{R}) e translação (\mathbf{T}) da câmara e calculada a matriz \mathbf{W} como em Eq. 3.9. É somado ruído gaussiano de média nula e variado o seu desvio padrão σ . De modo a obter resultados gerais as experiências são repetidas 100 vezes para cada valor de σ e calculada a média de cada resultado.

Nas secções seguintes analisa-se o efeito do ruído no método da factorização e nos métodos de identificação de reencontros propostos.

6.2. Factorização Matricial

Nesta secção pretende-se analisar a sensibilidade do método da factorização ao ruído, observando o seu efeito em cada saída. É relevante comparar as matrizes \mathbf{S} , \mathbf{R} e \mathbf{T} criadas artificialmente com as resultantes da reconstrução 3D. A análise da matriz \mathbf{W} não é tão relevante pois esta pertence apenas aos cálculos intermédios do processo SFM.

Nestas experiências utiliza-se a matriz \mathbf{W} usada na secção 5.4, onde se experimentou o método da factorização sem ruído. Considera-se um cubo de 40 pontos em 30 imagens de tamanho 160×120 *pixels*.

Resistência ao Ruído

Observa-se a resistência do método da factorização ao ruído na Fig. 6.2. Considera-se a mesma matriz de observação \mathbf{W} das experiências da secção 5.4, à qual se adiciona ruído com desvio padrão $\sigma = 4$ *pixels*. Na Fig. 6.2(a) observa-se uma das imagens com ruído e na Fig. 6.2(b) a respectiva reconstrução 3D.

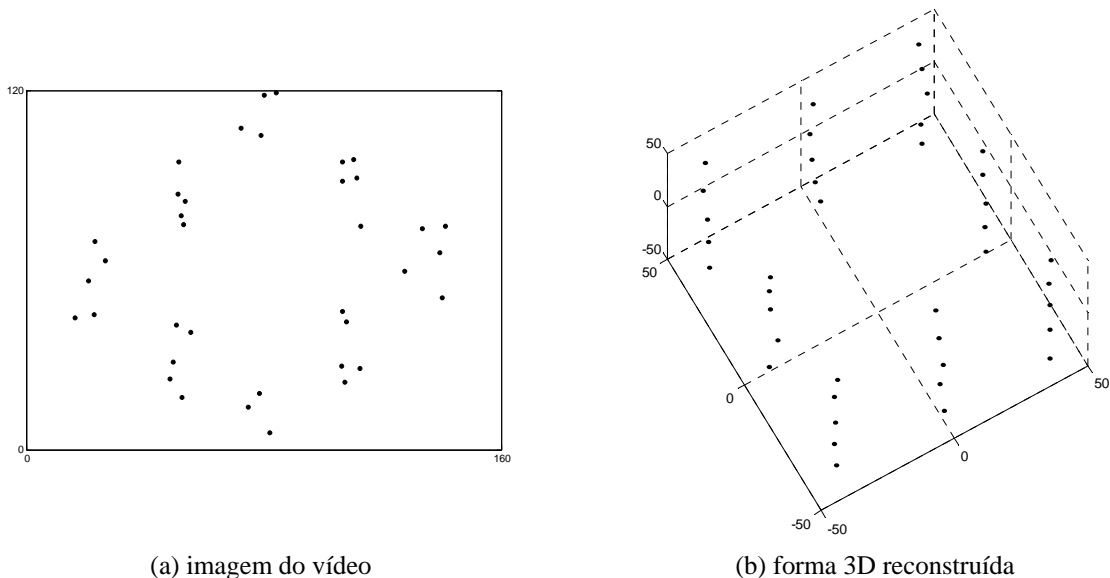
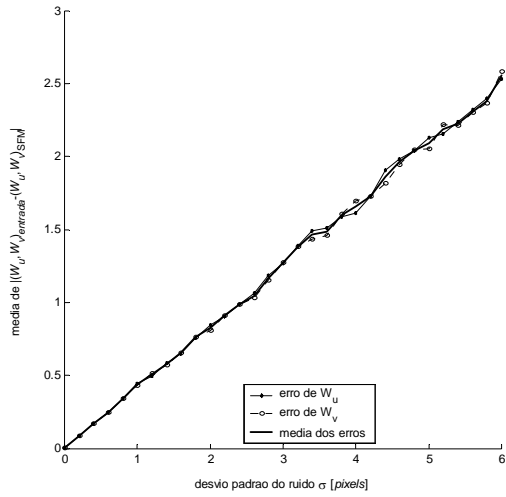


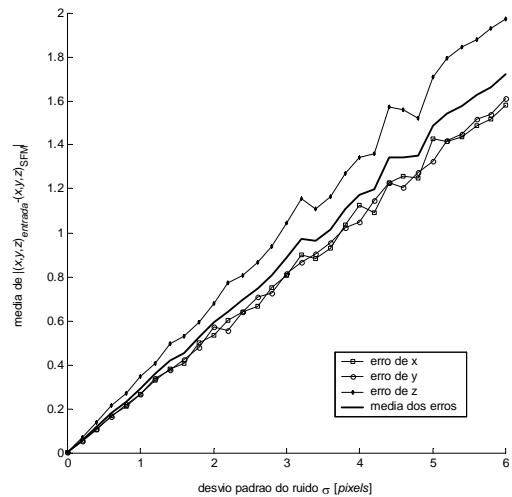
Fig. 6.2: Desempenho do método da Factorização com dados conhecidos e com ruído ($\sigma = 4$ *pixels*)

Verifica-se que o método da Factorização permite uma reconstrução da forma 3D bastante boa, mesmo com a introdução de ruído apreciável. Interessa saber como o ruído afecta cada saída da reconstrução 3D.

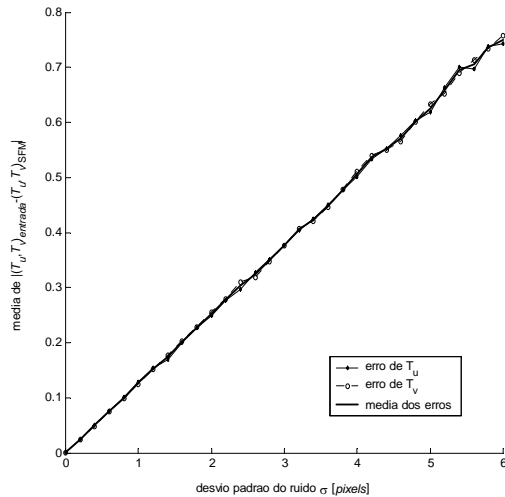
O efeito do ruído nas quatro matrizes \mathbf{W} , \mathbf{S} , \mathbf{R} e \mathbf{T} é observado na Fig. 6.3, onde se varia o desvio padrão do ruído σ . Gera-se a matriz de observação e adiciona-se ruído. Estima-se a forma do objecto (\mathbf{S}), rotação (\mathbf{R}) e translação (\mathbf{T}) da câmara pelo método da factorização. A matriz \mathbf{W}_{SFM} é calculada pela Eq. 4.11 a partir dos resultados. É calculada a diferença em cada coordenada de cada matriz, sendo depois feita a sua média.



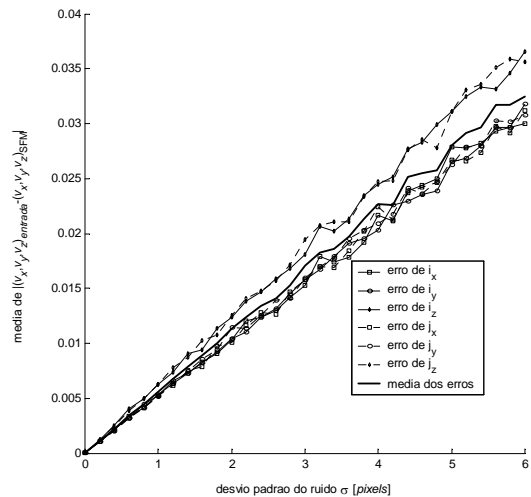
(a) matriz de observação \mathbf{W}



(c) matriz da forma \mathbf{S}



(b) matriz da translação \mathbf{T}



(d) matriz da rotação \mathbf{R}

Fig. 6.3: Efeito do ruído no processo *Structure From Motion*

Na Fig. 6.3 observa-se que cada coordenada de \mathbf{T} é afectada pelo ruído de igual forma, o que se esperava atendendo a que o ruído é de média nula e \mathbf{T} é composto pelas médias de cada linha de \mathbf{W} . Nas matrizes \mathbf{S} e \mathbf{R} observa-se que a terceira e nova coordenada calculada na reconstrução 3D é a mais afectada pelo ruído, nomeadamente porque foi calculada pelas outras duas coordenadas às quais se somou ruído.

Verifica-se que a existência de ruído implica haver erro na reconstrução 3D. Este erro é muito pequeno comparado com o nível de ruído. Demonstra-se assim a resistência do método da factorização a vários níveis de ruído.

Importância do Número de Imagens e do Número de Pontos

Observa-se na Fig. 6.4 a resistência do método da Factorização ao ruído com o aumento do número de imagens, nomeadamente na estimação da forma 3D. Foi utilizado um cubo com 40 pontos conhecidos e experimentados três casos com número de imagens F diferentes.

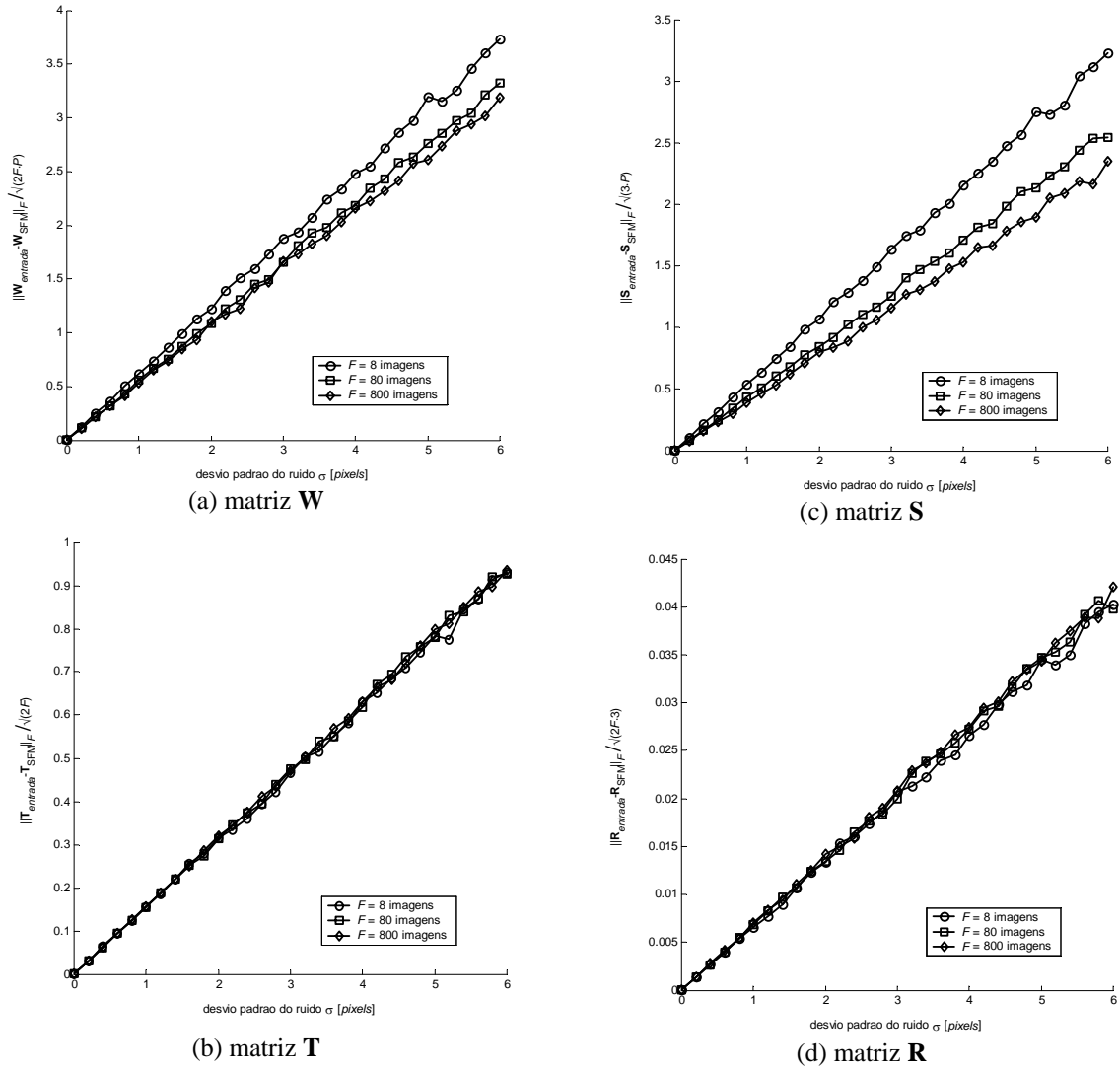


Fig. 6.4: Importância do número de imagens no combate ao ruído

Verifica-se que o erro por coordenada da forma (**S**) diminui com o aumento do número de imagens. Nas matrizes **T** e **R** não se observa alteração. Este efeito decorre da aplicação da factorização, em que um maior número de linhas permite uma melhor estimação do espaço das colunas. Como se viu no capítulo 3, a matriz da forma (**S**) é extraída do espaço das colunas da matriz de observação **W**.

O efeito similar seria observado ao aumentar o número de pontos. Um número de pontos maior permite uma melhor estimação do espaço das linhas. Assim, a rotação (**R**) e translação (**T**) da câmara apresentam menor erro porque são calculadas pelo espaço das linhas. O erro da forma (**S**) não é afectado pelo número de pontos.

De notar que o número de pontos característicos é sempre limitado aos existentes em cada objecto mas o número de imagens não: quanto mais se filmar melhor se pode estimar a sua forma 3D.

Dados Desconhecidos

Para estimação da informação desconhecida são utilizados os métodos EM, TS e estimativa inicial [7]. O efeito do ruído nestes algoritmos não é objecto de experiência neste trabalho. Em [16] conclui-se que o resultado varia com a introdução do ruído mas que o erro se mantém baixo para muitas situações com quantidade e forma de informação conhecida

pertinentes. Com ruído os métodos iterativos EM e TS apresentam uma convergência mais lenta e torna evidente a necessidade de uma boa estimativa inicial.

6.3. Identificação dos Reencontros

Pretende-se observar o efeito do ruído no resultado dos métodos de identificação de reencontros propostos: *Inspecção da Forma* e *Análise do Custo*. Para isso são contabilizados os reencontros descobertos para cada nível de ruído. Nesta experiência utilizou-se

Na Fig. 6.5 observa-se o número de reencontros descobertos e o número de reencontros errados que seriam corrigidos em função do desvio padrão do ruído. Nestas experiências foi utilizado um cubo de 45 pontos em 30 imagens de tamanho 160×120 pixels. Destes 45 pontos 5 eram reencontros não corrigidos. A matriz de observação tinha 68.9% de informação conhecida. Nos métodos *Inspecção da Forma* e *Análise do Custo* utilizam-se os parâmetros: distância de reencontro $D = 21$ pixels e factor de custo $\alpha = 2.7 \times 10^{-4}$, respectivamente.

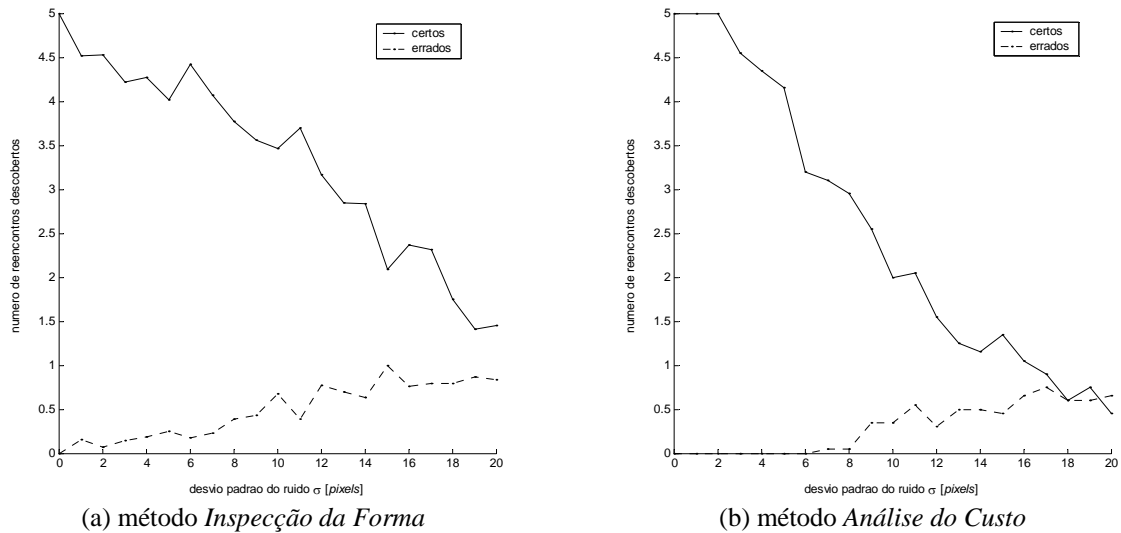


Fig. 6.5: Sensibilidade dos métodos *Inspecção da Forma* e *Análise do Custo* ao ruído

Observa-se que o método *Inspecção da Forma* é mais sensível ao ruído, considerando reencontros errados com um nível de ruído baixo. Como visto na secção 4.2.2., este erro é corrigido pelo utilizador. Ao reduzir a distância de reencontro D o algoritmo torna-se menos sensível ao ruído, tendo como contrapartida o menor número de reencontros descobertos.

O método *Análise do Custo* mostra-se eficiente para vários níveis de ruído, nomeadamente para $\sigma = 5 \sim 8$ pixels. A partir deste valor começam-se a considerar reencontros errados. Isto acontece porque as relações Eq. 4.7 entre o erro sem reencontros e o erro com reencontros não se mantêm. Estas relações são essenciais ao desempenho do método *Análise do Custo*. Como concluído no Anexo D, as relações entre estes erros dependem do ruído, da quantidade de informação conhecida, do número de imagens, do número de pontos e do número de reencontros.

De notar que os métodos estão dependentes da estimação de informação desconhecida. Como visto acima, na presença de ruído os métodos iterativos EM e TS apresentam uma convergência mais lenta. Assim, os métodos *Inspecção da Forma* e *Análise do Custo* podem apresentar melhores resultados aumentando o número de iterações dos métodos EM e TS.

7. Conclusões

7.1. Sumário

Neste trabalho implementou-se um processo SFM que utiliza o método da factorização matricial, proposto por Tomasi e Kanade em [2], e possibilita o tratamento de oclusões com identificação de reencontros, i.e., de partes do objecto que estiveram visíveis durante alguns instantes, ficando depois oclusas e que voltam a aparecer no vídeo. Mostrou-se que a identificação de reencontros é importante para o processo SFM, melhorando quer a estimação de matrizes de observação completas quer a reconstrução 3D. Assim, foi nesta a área que se propuseram soluções, nomeadamente dois métodos originais: *Inspecção da Forma* e *Análise do Custo*.

Implementou-se um algoritmo de selecção de pontos característicos numa imagem. Para estimar o movimento 2D do objecto na sequência de imagens foram estudados dois métodos de seguimento: *Seguidor de Intensidades* e *Seguidor de Pontos Característicos*. O *Seguidor de Pontos Característicos* e o algoritmo de selecção foram alvo de experiências que mostraram que a sua eficácia depende da interacção do utilizador na calibragem dos parâmetros para cada vídeo.

Foi ilustrada a eficácia do método da factorização como processo SFM, conseguindo recuperar eficientemente as informações relativas à forma do objecto e rotação e translação da câmara. Ficou demonstrada a sua robustez ao ruído. Mostrou-se que as estimativas das coordenadas da translação têm a mesma sensibilidade ao ruído. Na forma 3D e rotação 3D a terceira e nova coordenada estimada é mais sensível ao ruído do que as duas coordenadas relacionadas com a observação 2D. Verificou-se também que a forma 3D reconstruída é menos sensível ao erro com o aumento do número de imagens. Analogamente, as estimativas dos parâmetros que descrevem o movimento 3D tornam-se menos sensíveis ao ruído com o aumento do número de pontos característicos.

Quanto aos metodos propostos para identicao de reencontros, o método *Inspecção da Forma*, baseado nas distâncias entre as estimativas dos pontos característicos no espaço 3D, mostrou requerer interacção do utilizador. O método *Análise de Custo*, baseado na minimização de uma função de custo global, permite resolver eficientemente o problema SFM considerando os reencontros. Embora o seu desempenho dependa do nível de ruído e da quantidade de informação conhecida, mostrou funcionar adequadamente para uma gama pertinente de valores. No âmbito deste método foi analisada a expressão de custo a utilizar e foram feitas experiências para encontrar uma gama de valores válidos para o parâmetro que quantifica custo de cada ponto na descrição do modelo 3D.

A experiência realizada de reconstrução 3D a partir de video ilustra como o desempenho do processo SFM melhora quando se utilizam os métodos de identificação de reencontros propostos.

7.2. Direcções Futuras

Como se mostrou neste trabalho, é importante aprofundar o estudo sobre os reencontros, melhorando o desempenho dos métodos propostos neste trabalho ou adoptando outros que se mostrem mais eficazes.

A utilização de métodos que tornem possível o cálculo da distância de reencontro permitirá aumentar a eficácia do método *Inspecção da Forma*, dispensando assim da interacção do utilizador. Deverá então ser implementado um método que estime a distância entre o plano da imagem e o objecto. Métodos SFM que permitem estimar eficientemente esta distância utilizam o modelo de perspectiva, que não permite utilizar a factorização matricial.

No âmbito do método Análise do Custo existem duas questões que merecem uma análise aprofundada. Deve ser encontrada outra medida de semelhança mais eficiente que compare os vários modelos com os dados observados. Isto porque o erro utilizado entre a matriz completa estimada e a matriz observada é sensível ao nível de ruído e quantidade de informação conhecida, como se mostra neste trabalho, o que põe em causa a eficiência do método Análise do Custo. Deve também ser analisado o factor de custo, derivando-o analiticamente como descrito no MDL e noutros critérios de selecção de modelos.

Anexo A. Descrição do Método da Factorização

A.1. Introdução

Neste anexo descreve-se a implementação do método da factorização proposto por Tomasi e Kanade [2], que resolve o problema *structure from motion* (SFM). O método utiliza o modelo de projecção ortográfica e as igualdades

$$\mathbf{W} = \mathbf{RS} + \mathbf{T} \times \mathbf{1}_{1 \times P} = [\mathbf{R} | \mathbf{T}] \begin{bmatrix} \mathbf{S} \\ \mathbf{1}_{1 \times P} \end{bmatrix} \quad \text{Eq. A.1}$$

$$\tilde{\mathbf{W}} = \mathbf{RS}$$

em que $\mathbf{1}_{1 \times P}$ é uma linha de P uns. A partir da matriz de observação \mathbf{W} estimam-se as informações relativas à forma do objecto (\mathbf{S}) e rotação (\mathbf{R}) e translação (\mathbf{T}) da câmara. A matriz $\tilde{\mathbf{W}}$ é a matriz de observação normalizada.

Por construção, Tomasi e Kanade concluem que $\tilde{\mathbf{W}}$ tem característica 3 e que \mathbf{W} tem característica 4. O método de factorização tenta encontrar a melhor aproximação dentro das condições impostas. O problema de minimização da Eq. A.2 é a solução ML ao utilizar a matriz de observação normalizada de característica 3. De notar que $\|\cdot\|_F$ é a norma de Frobenius.

$$\min_{(\mathbf{RS}) \in S_3} \|\tilde{\mathbf{W}} - \mathbf{RS}\|_F \quad \text{Eq. A.2}$$

Ao utilizar a matriz de observação de característica 4 escreve-se,

$$\min_{(\mathbf{RS} + \mathbf{T}\mathbf{1}) \in S_4} \|\mathbf{W} - (\mathbf{RS} + \mathbf{T}\mathbf{1}_{1 \times P})\|_F \quad \text{Eq. A.3}$$

O método da factorização tem dois blocos distintos: a factorização, onde se impõe a característica desejada, e a normalização, onde se aplicam as restrições métricas. É ainda possível alinhar a 1ª imagem com a origem do espaço tridimensional, sendo este passo opcional.

Apresentam-se duas maneiras distintas de implementação. Na secção A.2 descreve-se o algoritmo como em [2], utilizando a característica 3. Na secção A.3 descreve-se o processo similar utilizando característica 4.

A.2. Matriz com Característica 3

Nesta secção descreve-se o método de factorização com característica 3. Este algoritmo foi o utilizado na implementação.

Factorização

Após calcular \mathbf{T} por

$$T_i = \frac{1}{P} \sum_{j=1}^P w_{ij} \quad , \text{ com } i \in [1, 2F] \quad \text{Eq. A.4}$$

calcula-se a matriz de observação normalizada $\tilde{\mathbf{W}}$ com,

$$\tilde{\mathbf{W}} = \mathbf{W} - \mathbf{T} \cdot \mathbf{1}_{1 \times P} \quad \text{Eq. A.5}$$

Aplicando a decomposição em valores singulares – *singular-value decomposition* (SVD) na literatura anglo-saxónica – à matriz $\tilde{\mathbf{W}}$ podem obter-se mais do que 3 valores singulares devido à presença de ruído.

$$\tilde{\mathbf{W}} = \mathbf{U}\mathbf{\Sigma}\mathbf{V} \quad \text{Eq. A.6}$$

As três matrizes \mathbf{U} , $\mathbf{\Sigma}$ e \mathbf{V} da Eq. A.6 são recortadas de modo a utilizar apenas os 3 maiores valores singulares, transformando-se nas \mathbf{U}' , $\mathbf{\Sigma}'$ e \mathbf{V}' , respectivamente com tamanhos $2F \times 3$, 3×3 e $3 \times P$.

A matriz $\hat{\mathbf{W}}$ dada por Eq. A.7 é a melhor estimativa da matriz de observação normalizada ideal. Reparar na notação utilizada para definir a redução da característica.

$$\hat{\mathbf{W}} = \tilde{\mathbf{W}} \downarrow S_3 = \mathbf{U}'\mathbf{\Sigma}'\mathbf{V}' \quad \text{Eq. A.7}$$

Definindo \mathbf{M} ($2F \times 3$) e \mathbf{N} ($3 \times P$) é possível escrever

$$\begin{aligned} \mathbf{M} &= \mathbf{U}'[\mathbf{\Sigma}']^{1/2} \\ \mathbf{N} &= [\mathbf{\Sigma}']^{1/2}\mathbf{V}' \end{aligned} \quad \rightarrow \quad \hat{\mathbf{W}} = \mathbf{M}\mathbf{N} \quad \text{Eq. A.8}$$

Normalização

Pretende-se normalizar \mathbf{M} e \mathbf{N} de modo a verificar as restrições métricas de \mathbf{i}_f e \mathbf{j}_f , que são vectores unitários ortogonais entre si. Verifica-se então

$$\begin{cases} |\mathbf{i}_f| = |\mathbf{j}_f| = 1 \\ \mathbf{i}_f \cdot \mathbf{j}_f = \mathbf{i}_f^T \mathbf{j}_f = 0 \end{cases} \quad \text{Eq. A.9}$$

Existe uma matriz \mathbf{Q} (3×3) tal que

$$\mathbf{R} = \mathbf{M}\mathbf{Q} \quad \text{Eq. A.10}$$

$$\mathbf{S} = \mathbf{Q}^{-1}\mathbf{N} \quad \text{Eq. A.11}$$

Pelas Eq. A.9 e Eq. A.10 sabemos que para cada imagem se verifica

$$\mathbf{i}_f^T = [i_{fx} \quad i_{fy} \quad i_{fz}] = \hat{\mathbf{i}}_f^T \mathbf{Q}$$

Deseja-se descobrir \mathbf{Q} que verifica o sistema

$$\hat{\mathbf{i}}_f^T \mathbf{Q} \mathbf{Q}^T \hat{\mathbf{i}}_f = 1 \quad \text{Eq. A.12}$$

$$\hat{\mathbf{j}}_f^T \mathbf{Q} \mathbf{Q}^T \hat{\mathbf{j}}_f = 1 \quad \text{Eq. A.13}$$

$$\hat{\mathbf{i}}_f^T \mathbf{Q} \mathbf{Q}^T \hat{\mathbf{j}}_f = 0 \quad \text{Eq. A.14}$$

Define-se $\mathbf{A} = \mathbf{Q}\mathbf{Q}^T$ (3×3). Por construção tem-se

$$\mathbf{A}^T = \mathbf{Q}\mathbf{Q}^T = \mathbf{A} \quad \text{Eq. A.15}$$

logo

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{12} & a_{22} & a_{23} \\ a_{13} & a_{23} & a_{33} \end{bmatrix}$$

As Eq. A.12, Eq. A.13 e Eq. A.14 são 3 conjuntos de F equações que definem um sistema de $3F$ equações e 6 incógnitas.

Para exemplificação da resolução utiliza-se uma linha do último conjunto de equações Eq. A.14. Note-se que

$$\hat{\mathbf{i}}_f^T \equiv [m_{f1} \quad m_{f2} \quad m_{f3}] \equiv [\hat{i}_{fx} \quad \hat{i}_{fy} \quad \hat{i}_{fz}]$$

Por simplificação não se utilizam os circunflexos nos cálculos seguintes. Para uma imagem qualquer – para um f qualquer – tem-se

$$\begin{aligned} & \begin{bmatrix} i_x & i_y & i_z \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{12} & a_{22} & a_{23} \\ a_{13} & a_{23} & a_{33} \end{bmatrix} \begin{bmatrix} j_x \\ j_y \\ j_z \end{bmatrix} = 0 \\ & \begin{bmatrix} i_x a_{11} + i_y a_{12} + i_z a_{13} & i_x a_{12} + i_y a_{22} + i_z a_{23} & i_x a_{13} + i_y a_{23} + i_z a_{33} \end{bmatrix} \begin{bmatrix} j_x \\ j_y \\ j_z \end{bmatrix} = 0 \\ & (i_x a_{11} + i_y a_{12} + i_z a_{13}) j_x + (i_x a_{12} + i_y a_{22} + i_z a_{23}) j_y + (i_x a_{13} + i_y a_{23} + i_z a_{33}) j_z = 0 \\ & \begin{bmatrix} i_x j_x & i_y j_x + i_x j_y & i_z j_x + i_x j_z & i_y j_y & i_z j_y + i_y j_z & i_z j_z \end{bmatrix} \begin{bmatrix} a_{11} \\ a_{12} \\ a_{13} \\ a_{22} \\ a_{23} \\ a_{33} \end{bmatrix} = 0 \end{aligned}$$

Assim, o sistema acima definido pelas Eq. A.12, Eq. A.13 e Eq. A.14 é equivalente a

$$\begin{bmatrix} (i_x i_x)_1 & (i_y i_x + i_x i_y)_1 & (i_z i_x + i_x i_z)_1 & (i_y i_y)_1 & (i_z i_y + i_y i_z)_1 & (i_z i_z)_1 \\ \vdots & & & & & \\ (i_x i_x)_F & (i_y i_x + i_x i_y)_F & (i_z i_x + i_x i_z)_F & (i_y i_y)_F & (i_z i_y + i_y i_z)_F & (i_z i_z)_F \\ (j_x j_x)_1 & (j_y j_x + j_x j_y)_1 & (j_z j_x + j_x j_z)_1 & (j_y j_y)_1 & (j_z j_y + j_y j_z)_1 & (j_z j_z)_1 \\ \vdots & & & & & \\ (j_x j_x)_F & (j_y j_x + j_x j_y)_F & (j_z j_x + j_x j_z)_F & (j_y j_y)_F & (j_z j_y + j_y j_z)_F & (j_z j_z)_F \\ (i_x j_x)_1 & (i_y j_x + i_x j_y)_1 & (i_z j_x + i_x j_z)_1 & (i_y j_y)_1 & (i_z j_y + i_y j_z)_1 & (i_z j_z)_1 \\ \vdots & & & & & \\ (i_x j_x)_F & (i_y j_x + i_x j_y)_F & (i_z j_x + i_x j_z)_F & (i_y j_y)_F & (i_z j_y + i_y j_z)_F & (i_z j_z)_F \end{bmatrix} \begin{bmatrix} a_{11} \\ a_{12} \\ a_{13} \\ a_{22} \\ a_{23} \\ a_{33} \end{bmatrix} = \begin{bmatrix} 1 \\ \vdots \\ 1 \\ \vdots \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

Atribuem-se nomes a estas matrizes

$$\mathbf{B} \cdot \mathbf{a} = \mathbf{G} \quad \text{Eq. A.16}$$

Existindo mais equações do que incógnitas, o sistema da Eq. A.16 é resolvido pela melhor aproximação aplicando o método dos mínimos quadrados – *Least Squares* (LS) na literatura anglo-saxónica,

$$\mathbf{a} = \mathbf{B}^\# \mathbf{G} \quad \text{Eq. A.17}$$

$$\mathbf{B}^\# = (\mathbf{B}^T \mathbf{B})^{-1} \mathbf{B}^T \quad \text{Eq. A.18}$$

A Eq. A.18 é a pseudo-inversa de Moore-Penrose.

A matriz procurada calcula-se factorizando a matriz \mathbf{A} , que por construção

$$\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{U}^T$$

Calcula-se assim

$$\mathbf{Q} = \mathbf{U}\mathbf{\Sigma}^{\frac{1}{2}}$$

Obtêm-se assim \mathbf{R} , \mathbf{S} .

Alinhar Objecto (opcional)

De modo a alinhar a 1ª imagem com a origem do espaço 3D, aplica-se uma rotação que transforme o referencial da 1ª imagem numa matriz identidade.

$$\begin{aligned} \mathbf{R}_{\text{alinhado}} &= \mathbf{R}\mathbf{R}_0 \\ \mathbf{S}_{\text{alinhado}} &= \mathbf{R}_0^T\mathbf{S} \end{aligned}, \text{ com } \mathbf{R}_0 = [\mathbf{i}_1 \quad \mathbf{j}_1 \quad \mathbf{k}_1] \quad \text{Eq. A.19}$$

Obtêm-se assim os versores do sistema de referência da primeira imagem

$$\mathbf{i}_1 = [1 \quad 0 \quad 0]^T, \quad \mathbf{j}_1 = [0 \quad 1 \quad 0]^T$$

alinhados com os eixos x e y do sistema de referência do espaço 3D.

A.3. Matriz com Característica 4

Nesta secção descreve-se o método de factorização com característica 4. Este algoritmo não foi implementado pois utilizou-se o método de factorização com característica 3 descrito acima.

Parte do método de factorização com característica 4 é similar ao método de factorização com característica 3. Não se descreve a operação de alinhar a 1ª imagem com a origem do espaço tridimensional por ser igual à descrita acima.

Factorização

Aplicando a SVD à matriz \mathbf{W} podem obter-se mais do que 4 valores singulares devido à presença de ruído.

$$\mathbf{W} = \mathbf{U}\mathbf{\Sigma}\mathbf{V} \quad \text{Eq. A.20}$$

As três matrizes \mathbf{U} , $\mathbf{\Sigma}$ e \mathbf{V} são recortadas de modo a utilizar apenas os 4 maiores valores singulares, transformando-se nas \mathbf{U}' , $\mathbf{\Sigma}'$ e \mathbf{V}' , respectivamente com tamanhos $2F \times 4$, 4×4 e $4 \times P$.

A matriz $\hat{\mathbf{W}}$ dada por Eq. A.21 é a melhor estimativa da matriz de observação ideal. Reparar na notação utilizada para definir a redução da característica.

$$\hat{\mathbf{W}} = \mathbf{W} \downarrow S_4 = \mathbf{U}'\mathbf{\Sigma}'\mathbf{V}' \quad \text{Eq. A.21}$$

Definem-se \mathbf{M} ($2F \times 4$) e \mathbf{N} ($4 \times P$) como na Eq. A.8.

Normalização

Pretende-se normalizar \mathbf{M} e \mathbf{N} de modo a verificar as restrições métricas de \mathbf{i}_f e \mathbf{j}_f da Eq. A.9.

Existe uma matriz \mathbf{Q} (4×4) tal que

$$[\mathbf{R} \mid \mathbf{T}] = \mathbf{M}\mathbf{Q} \quad \text{Eq. A.22}$$

$$\begin{bmatrix} \mathbf{S} \\ \mathbf{1} \end{bmatrix} = \mathbf{Q}^{-1}\mathbf{N} \quad \text{Eq. A.23}$$

Distingue-se em \mathbf{Q} a matriz \mathbf{Q}' (4×3) e o vector \mathbf{q}' (4×1) – Eq. A.24.

$$\mathbf{Q} = [\mathbf{Q}' | \mathbf{q}'] = \begin{bmatrix} q_{11} & q_{12} & q_{13} & q_{14} \\ q_{21} & q_{22} & q_{23} & q_{24} \\ q_{31} & q_{32} & q_{33} & q_{34} \\ q_{41} & q_{42} & q_{43} & q_{44} \end{bmatrix} \quad \text{Eq. A.24}$$

Atendendo a Eq. A.22 escrevem-se as Eq. A.25 e Eq. A.26

$$\mathbf{R} = \mathbf{M}\mathbf{Q}' \quad \text{Eq. A.25}$$

$$\mathbf{T} = \mathbf{M}\mathbf{q}' \quad \text{Eq. A.26}$$

Pelas Eq. A.25 e conhecendo a forma como os vectores se arrumam na matriz \mathbf{R} sabemos que

$$\mathbf{i}_f^T = [i_{fx} \quad i_{fy} \quad i_{fz}] = \hat{\mathbf{i}}_f^T \mathbf{Q}' = [m_{f1} \quad m_{f2} \quad m_{f3} \quad m_{f4}] \mathbf{Q}'$$

Deseja-se descobrir \mathbf{Q}' que verifica as Eq. A.12, Eq. A.13 e Eq. A.14.

Por um método similar ao utilizado para a característica 3 obtém-se \mathbf{Q}' .

Designam-se os elementos de \mathbf{Q}^{-1} por b_{ij} . Distingue-se em \mathbf{Q}^{-1} a matriz \mathbf{Q}'' (3×4) e o vector \mathbf{q}'' (1×4) – Eq. A.27.

$$\mathbf{Q}^{-1} = \begin{bmatrix} \mathbf{Q}'' \\ \mathbf{q}'' \end{bmatrix} = \begin{bmatrix} b_{11} & b_{12} & b_{31} & b_{14} \\ b_{21} & b_{22} & b_{23} & b_{24} \\ b_{31} & b_{32} & b_{33} & b_{34} \\ b_{41} & b_{42} & b_{43} & b_{44} \end{bmatrix} \quad \text{Eq. A.27}$$

Atendendo a Eq. A.23 escrevem-se as Eq. A.28 e Eq. A.29.

$$\mathbf{S} = \mathbf{Q}''\mathbf{N} \quad \text{Eq. A.28}$$

$$\mathbf{1} = \mathbf{q}''\mathbf{N} \quad \text{Eq. A.29}$$

Da Eq. A.29 encontra-se o sistema da Eq. A.30.

$$\mathbf{N}^T \mathbf{q}''^T = \mathbf{1}^T \quad \text{Eq. A.30}$$

Existindo mais equações do que incógnitas, o sistema da Eq. A.30 é resolvido pela melhor aproximação pelo LS.

$$\mathbf{q}''^T = \mathbf{N}^{\#} \mathbf{1}^T \quad \text{Eq. A.31}$$

$$\mathbf{N}^{\#} = (\mathbf{N}\mathbf{N}^T)^{-1} \mathbf{N} \quad \text{Eq. A.32}$$

Para encontrar os elementos de \mathbf{q}' utiliza-se os elementos de \mathbf{Q}' e \mathbf{q}'' já descobertos e a definição de inversa

$$b_{ij} = \frac{1}{|\mathbf{Q}|} \text{cof}(q_{ij})$$

em que $\text{cof}(q_{ij})$ é o co-factor do elemento correspondente da matriz \mathbf{Q} . Obtém-se assim \mathbf{q}' , logo \mathbf{Q} é conhecida.

Obtém-se assim \mathbf{R} , \mathbf{S} e \mathbf{T} .

Anexo B. Descrição da Estimação de Matrizes Completas

B.1. Introdução

Neste anexo resolve-se o problema de estimar matrizes completas a partir de matrizes com entradas desconhecidas, como a ilustrada na Fig. B.1.



Fig. B.1: Matriz de observação \mathbf{W} com informação desconhecida (\mathbf{W}_0)

Apresenta-se a solução de [7], que engloba a inicialização e dois métodos iterativos: o *Expectation-Maximization* (EM) e o *Two-Steps* (TS).

B.2. Estimativa Inicial

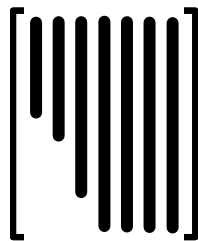
Qualquer matriz $\tilde{\mathbf{W}}$ pode ser escrita como na Eq. B.1, em que \mathbf{A} de tamanho $2F \times 4$ determina o espaço das colunas e \mathbf{B} de tamanho $4 \times P$ o espaço das linhas.

$$\tilde{\mathbf{W}} = \mathbf{AB} \in S_4 \quad \text{Eq. B.1}$$

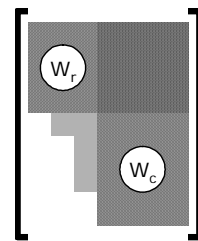
A estimativa inicial $\tilde{\mathbf{W}}^{(0)}$ é calculada por combinação do espaço das linhas com o espaço das colunas da porção conhecida de \mathbf{W}_0 . Para isso divide-se \mathbf{W}_0 em submatrizes mais pequenas.

Estimação do Subespaço

Considere-se uma matriz \mathbf{W}_0 como na Fig. B.2(a), em que algumas colunas e linhas são totalmente conhecidas. Na Fig. B.2(b) observa-se a mancha escura que identifica os dados conhecidos e as submatrizes \mathbf{W}_c e \mathbf{W}_r que reúnem as colunas e linhas totalmente conhecidas, respectivamente.



(a) algumas colunas e linhas são totalmente conhecidas



(b) destacam-se as submatrizes \mathbf{W}_r e \mathbf{W}_c

Fig. B.2: Matriz \mathbf{W}_0 e a estimação do subespaço

A partir de \mathbf{W}_c é possível estimar o espaço das colunas \mathbf{A} pela Eq. B.2 – recordar operação no passo da factorização na secção 3.3.2.

$$\mathbf{A} = \mathbf{W}_c \downarrow S_4 \quad \text{Eq. B.2}$$

A melhor aproximação do espaço das linhas \mathbf{B} é dada pela ML, aplicando o método LS que utiliza a pseudo-inversa de Moore-Penrose $\mathbf{A}_r^\#$ como definida em Eq. A.18. Esta matriz \mathbf{A}_r reúne as linhas de \mathbf{A} que correspondem às linhas de \mathbf{W}_r .

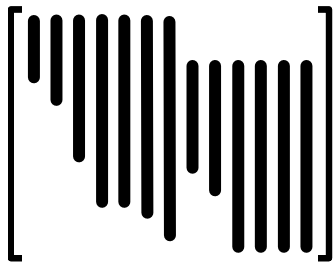
$$\mathbf{B} = \mathbf{A}_r^\# \mathbf{W}_r \quad \text{Eq. B.3}$$

As restrições resultam da utilização da característica 4, em que a submatriz \mathbf{W}_c tem de ter 4 colunas linearmente independentes e \mathbf{W}_r 4 linhas linearmente independentes.

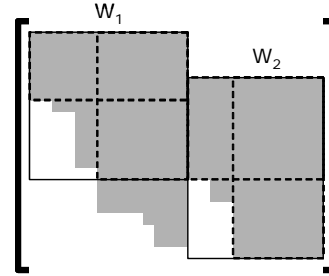
Combinação dos Subespaços

No caso geral a matriz \mathbf{W}_0 apresenta várias submatrizes como a definida na Fig. B.2(a). Exemplifica-se o algoritmo da estimativa inicial por combinação de espaços com uma matriz \mathbf{W}_0 como na Fig. B.3, composta por duas submatrizes que agrupam pontos com o mesmo início de seguimento. O processo de combinação de subespaços é repetido se a matriz \mathbf{W}_0 apresentar mais submatrizes do mesmo tipo.

Na Fig. B.3(b) encontram-se definidas as submatrizes \mathbf{W}_1 e \mathbf{W}_2 , tendo cada uma o mesmo aspecto que na Fig. B.2(b). A tracejado e para cada uma reúnem-se as colunas e linhas totalmente conhecidas.



(a) existem dois conjuntos de pontos com inícios de seguimento diferentes



(b) destacam-se as submatrizes \mathbf{W}_1 e \mathbf{W}_2

Fig. B.3: Matriz \mathbf{W}_0 e a combinação de subespaços

As submatrizes \mathbf{W}_1 e \mathbf{W}_2 devem ter pelo menos 4 colunas e 4 linhas sem dados desconhecidos. Cada submatriz é factorizada como na Eq. B.2 de modo a obter os respectivos espaços de colunas, \mathbf{A}_1 e \mathbf{A}_2 . Por definição de \mathbf{W}_1 e \mathbf{W}_2 , \mathbf{A}_1 e \mathbf{A}_2 tem linhas em comum que são reunidas em \mathbf{A}_{12} e \mathbf{A}_{21} , respectivamente. Como a matriz de observação \mathbf{W} procurada tem característica 4, \mathbf{A}_{12} e \mathbf{A}_{21} relacionam-se por uma transformação linear como em Eq. B.4, em que \mathbf{N} tem tamanho 4×4 .

$$\mathbf{A}_{12} \simeq \mathbf{A}_{21} \mathbf{N} \quad \text{Eq. B.4}$$

Na Eq. B.4 estima-se \mathbf{N} por LS, em que $\mathbf{A}_{21}^\#$ é a pseudo-inversa de Moore-Penrose.

$$\mathbf{N} = \mathbf{A}_{21}^\# \mathbf{A}_{12} \quad \text{Eq. B.5}$$

Obtém-se assim na Eq. B.6 a combinação dos espaços das colunas das submatrizes \mathbf{W}_1 e \mathbf{W}_2 , em que \mathbf{A}_{21} reúne as linhas de \mathbf{A}_2 não correspondidas em \mathbf{W}_1 .

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_1 \\ \mathbf{A}_{2|1}\mathbf{N} \end{bmatrix} \quad \text{Eq. B.6}$$

A combinação do espaço das linhas das submatrizes \mathbf{W}_1 e \mathbf{W}_2 é calculada como na Eq. B.3.

Assim se obtém a estimativa inicial da matriz $\tilde{\mathbf{W}}$,

$$\tilde{\mathbf{W}}^{(0)} = \mathbf{A}\mathbf{B} \quad \text{Eq. B.7}$$

B.3. Algoritmo Iterativo Expectation-Maximization

No algoritmo iterativo *Expectation-Maximization* (EM) a estimação dos dados desconhecidos da matriz de observação \mathbf{W}_0 é feita em dois passos alternados. Utiliza a estimativa inicial $\tilde{\mathbf{W}}^{(0)}$ acima descrita.

Passo Expectation

Para a iteração $\hat{\mathbf{W}}^{(k)}$ estimam-se os valores desconhecidos da matriz. Estes valores são igualados aos valores da iteração anterior $\tilde{\mathbf{W}}^{(k-1)}$,

$$\hat{\mathbf{W}}^{(k)} = \mathbf{W} \odot \mathbf{M} + \tilde{\mathbf{W}}^{(k-1)}[\mathbf{1} - \mathbf{M}] \quad \text{Eq. B.8}$$

Passo Maximization

Como visto no método da factorização, a estimativa ML de característica 4 obtém-se factorizando $\hat{\mathbf{W}}^{(k)}$, impondo a característica desejada,

$$\tilde{\mathbf{W}}^{(k)} = \hat{\mathbf{W}}^{(k)} \downarrow S_4 \quad \text{Eq. B.9}$$

B.4. Algoritmo Iterativo Two-Steps

O algoritmo iterativo *Two-Steps* (TS) aplica a definição da Eq. B.1 directamente no problema de minimização global do método da factorização,

$$\min_{\mathbf{A}, \mathbf{B}} \left\| (\mathbf{W}_0 - \mathbf{A}\mathbf{B}) \odot \mathbf{M} \right\|_F \quad \text{Eq. B.10}$$

O algoritmo é composto por dois passos que se executam alternadamente e não utiliza a computacionalmente pesada SVD. É inicializado pela matriz do espaço das colunas \mathbf{A} da estimativa inicial $\tilde{\mathbf{W}}^{(0)}$ da Eq. B.7 acima descrita.

Cálculo de B conhecendo A

Sendo \mathbf{A} conhecida, o problema de minimização da Eq. B.10 tem solução LS. esta solução é calculada para cada coluna \mathbf{b}_p de \mathbf{B} utilizando a pseudo-inversa,

$$\mathbf{b}_p = \left[\mathbf{A}^T (\mathbf{A} \odot \mathbf{M}_p) \right]^{-1} \mathbf{A}^T (\mathbf{w}_p \odot \mathbf{m}_p) \quad \text{Eq. B.11}$$

As letras minúsculas identificam vectores coluna das matrizes correspondentes com a letra maiúscula. Como mostra a Eq. B.12, a matriz \mathbf{M}_p tem tamanho $2F \times 4$ com 4 colunas iguais a \mathbf{m}_p . Este vector \mathbf{m}_p verifica os dados conhecidos de \mathbf{W}_0 e os correspondentes em \mathbf{A} .

$$\mathbf{M}_p = \mathbf{m}_p \mathbf{1}_{1 \times 4} \quad \text{Eq. B.12}$$

Este procedimento é repetido para as $p = 1, \dots, P$ colunas.

Cálculo de A conhecendo B

Sendo \mathbf{B} conhecida estima-se cada linha \mathbf{a}_f de \mathbf{A} de uma maneira similar. Na Eq. B.13 as letras minúsculas identificam vectores linha em vez de coluna e a matriz \mathbf{M}_f verifica a Eq. B.14.

$$\mathbf{a}_f = (\mathbf{w}_f \odot \mathbf{m}_f) \mathbf{B}^T \left[\mathbf{B}^T (\mathbf{B} \odot \mathbf{M}_f) \right]^{-1} \quad \text{Eq. B.13}$$

$$\mathbf{M}_f = \mathbf{1}_{1 \times 4} \mathbf{m}_f \quad \text{Eq. B.14}$$

Este procedimento é repetido para as $f = 1, \dots, 2F$ linhas.

Anexo C. Expressão do Custo de Reencontro

C.1. Introdução

Nesta secção analisa-se qual a expressão do custo do reencontro a utilizar.

Esta expressão será utilizada no método de identificação de reencontros Análise do Custo. Este método é inspirado no princípio de Descrição de Comprimento Mínimo - *minimum description length* (MDL) – que aplica o conceito de Verosimilhança Penalizada – *Penalized Likelihood* (PL).

Minimum Description Length

Um problema de selecção de modelos surge quando é necessário eleger entre vários modelos aquele que melhor descreve um conjunto de amostras. Em 1978, Rissanen apresenta o princípio MDL [Rissanen], que escolhe o modelo mais conciso que melhor descreve as amostras. Este princípio é usualmente aplicado para escolher o melhor código binário numa transmissão de dados mas também foi aplicado em várias outras situações. Existem por isso várias expressões para o cálculo MDL, sendo a mais usual

$$-L(D;\theta) + \frac{p}{2} \log_2 N \quad \text{Eq. C.1}$$

em que D são os dados, θ os parâmetros do modelo, p o número de parâmetros do modelo e N o número de símbolos do modelo.

Na primeira parcela da Eq. C.1 encontra-se $L(D;\theta)$, que quantifica a semelhança (*likelihood* na literatura anglo-saxónica) do modelo aos dados. Este é o termo que procurará um modelo que descreva bem os dados. A segunda parcela mede a complexidade do modelo e portanto procurará pelo modelo com menos símbolos.

A Eq. C.1 calcula o custo de um reencontro. Após calcular o custo de todos os modelos é escolhido o de menor custo.

De uma maneira geral escreve-se

$$-\left(\begin{array}{c} \text{semelhança} \\ \text{do modelo} \\ \text{com os dados} \end{array} \right) + (\text{factor}) \cdot \left(\begin{array}{c} \text{complexidade} \\ \text{do modelo} \end{array} \right) \quad \text{Eq. C.2}$$

Como o MDL calcula o custo de cada modelo, o factor da Eq. C.2 é designado por factor de custo. Funções de custo com estes dois termos são funções PL [9].

Reencontros

A Fig. D.1 mostra um exemplo das matrizes utilizadas. Cria-se uma matriz \mathbf{W} (a) de característica 4, à qual se impõe períodos de oclusão em vários pontos, formando assim a matriz sem reencontros $\mathbf{W}_{0,R}$ (b). Introduzem-se novos pontos no final de $\mathbf{W}_{0,R}$ que são reencontros de pontos já seguidos, obtendo-se a matriz \mathbf{W}_0 (c). Existe uma matriz $\mathbf{W}_{0,R}$ (d) que identifica os reencontros criados.

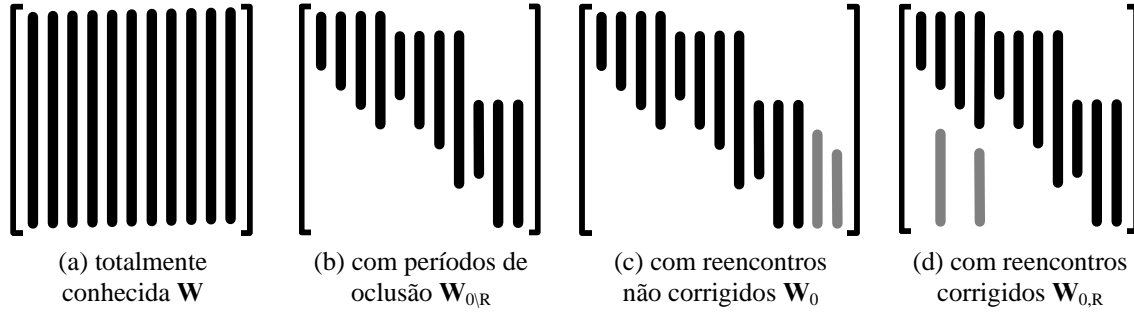


Fig. C.1: Exemplos das matrizes de observação utilizadas

Análise do Custo

O método Análise do Custo funciona com base na premissa que

$$\begin{aligned}
 e_{errado} &> e_0 \\
 e_{certo} &> e_{errado} \\
 e_{certo} &\sim e_0
 \end{aligned}
 \tag{Eq. C.3}$$

em que se tem o erro sem reencontros e_0 , o menor erro dos reencontros errados e_{errado} e o maior erro dos reencontros certos e_{certo} . As relações da Eq. C.3 foram testadas no Anexo D.

C.2. Cálculo do Custo do Reencontro

São analisadas várias hipóteses para o cálculo do custo dos reencontros. Para aferir a validade de cada expressão é necessário calcular uma gama de valores possíveis para o factor de custo. Para isso são utilizados valores resultantes de experiências descritas no Anexo B.

1ª Expressão

Decide-se a criação do reencontro comparando o erro e , por ponto, da aproximação aos dados conhecidos da matriz de observação $\mathbf{W}_{0,R}$, dado por

$$e = \frac{\|(\mathbf{W}_{0,R} - \tilde{\mathbf{W}}_{0,R}) \odot \mathbf{M}\|_F}{W_{conhecidos}}
 \tag{Eq. C.4}$$

em que $\tilde{\mathbf{W}}_{0,R}$ é a matriz completa estimada e $W_{conhecidos}$ é o número de coordenadas conhecidas. A matriz $\tilde{\mathbf{W}}_{0,R}$ resulta dos métodos de estimação da informação desconhecida tratados na secção 3.4. A matriz \mathbf{M} verifica $\mathbf{M} = [m_{ij}]$ tal que

$$m_{ij} = \begin{cases} 1, & \text{se } w_{0ij} \text{ conhecido} \\ 0, & \text{se } w_{0ij} \text{ desconhecido} \end{cases}$$

e \odot é o produto ponto-a-ponto, também conhecido por produto de Hadamard.

A expressão do custo c é então

$$c = e
 \tag{Eq. C.5}$$

Esta expressão tem sucesso em 41 testes de um total de 2700 (1.5%). Claramente que a decisão pelo erro não é eficiente para identificar reencontros.

2ª Expressão

A medida de semelhança é dada por $(-e)$. Esta medida de semelhança é sempre negativa. É tanto maior quanto melhor for a aproximação aos dados conhecidos.

A medida de complexidade é o número de pontos P do modelo.

A expressão do custo c é então

$$c = e + \alpha \cdot P \quad \text{Eq. C.6}$$

em que α é o factor de custo.

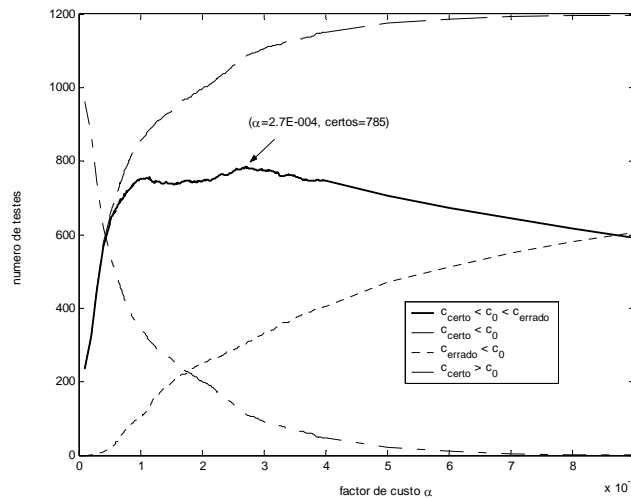


Fig. C.2: Estimar o factor de custo α – 2ª expressão

É escolhido o valor $\alpha = 2.7 \times 10^{-4}$ pois apresenta melhores resultados. Tendo sido feitos 2700 testes, o valor escolhido permitiu identificar todos os reencontros em 784 testes (29.0%). Destes 2700 a premissa necessária ao método *Análise do Custo* da Eq. C.3 verificou-se 1152 vezes. É neste conjunto de testes que a transformação erro/custo se mostra relevante, mostrando-se válida em 68.1% com este factor de custo.

3ª Expressão

A expressão do custo c é

$$c = e + \alpha \cdot \log_{10}(P) \quad \text{Eq. C.7}$$

A análise da complexidade com o logaritmo justifica-se pois está presente na Eq. C.1.

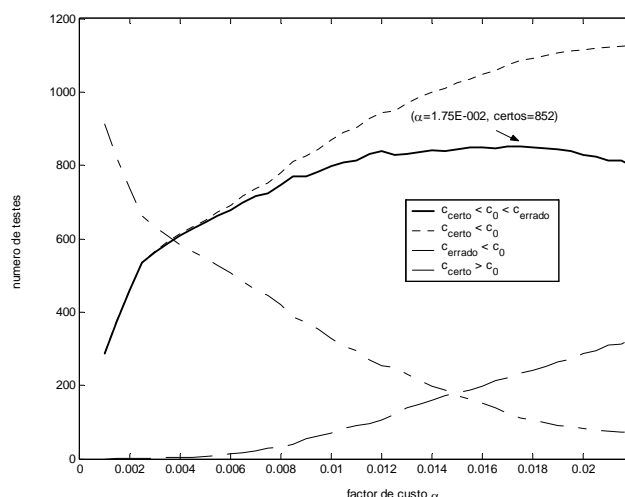


Fig. C.3: Estimar o factor de custo α – 3ª expressão

Mesmo que esta expressão pareça apresentar resultados bastante bons, a utilização de uma operação não linear como o logaritmo não permite um factor de custo constante. É possível aproximar a função logaritmo por vários troços lineares, mas para cada troço o factor de custo seria diferente.

É escolhido o valor $\alpha = 1.75 \times 10^{-2}$ pois apresenta melhores resultados.

Utilizando esta expressão com $\alpha = 1.75 \times 10^{-2}$ apenas nos dados com erros pequenos não é possível identificar nenhum dos reencontros.

4ª Expressão

A expressão do custo c é

$$c = e + \alpha \cdot e_0 \cdot P \quad \text{Eq. C.8}$$

Nesta hipótese a complexidade é dada por $(e_0 \cdot P)$. Isso introduz o efeito do erro e da semelhança na parcela da complexidade, e por isso esta expressão foge um pouco ao conceito da expressão MDL Eq. C.2.

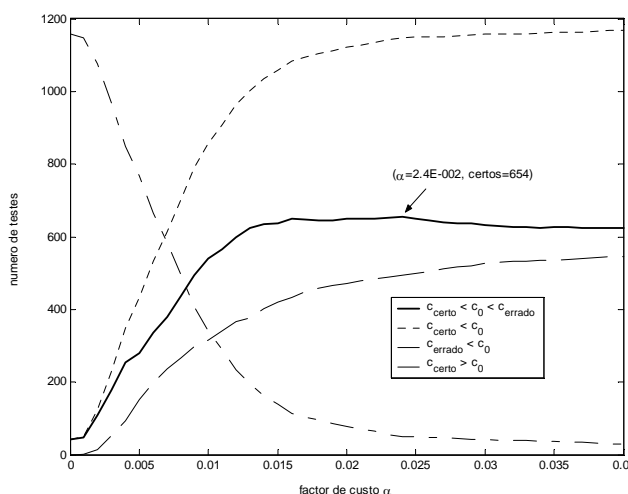


Fig. C.4: Estimar o factor de custo α – 4ª expressão

É escolhido o valor $\alpha = 2.4 \times 10^{-2}$ pois apresenta melhores resultados. Tendo sido feitos 2700 testes, o valor escolhido permitiu identificar todos os reencontros em 654 testes (24.2%).

Destes 2700 a premissa necessária ao método *Análise do Custo* da Eq. C.3 verificou-se 1152 vezes. É neste conjunto de testes que a transformação erro/custo se mostra relevante, mostrando-se válida em 56.8% com este factor de custo.

5ª Expressão

A expressão do custo c é

$$c = -\frac{1}{e} + \alpha \cdot P \quad \text{Eq. C.9}$$

Nesta hipótese utiliza-se $(1/e)$ como medida de semelhança.

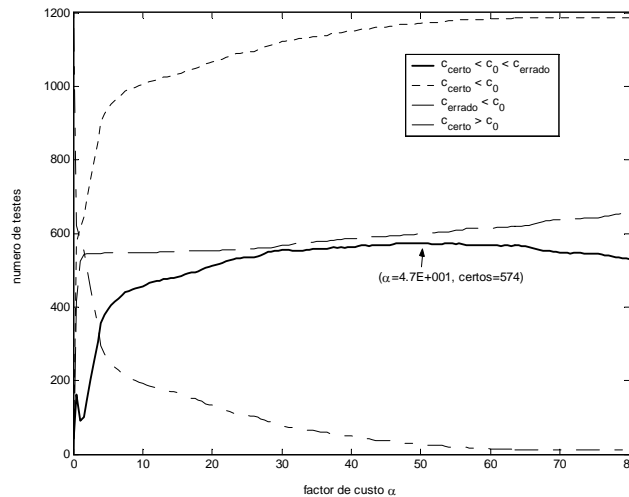


Fig. C.5: Estimar o factor de custo α – 5ª expressão

É escolhido o valor $\alpha = 47$ pois apresenta melhores resultados. Tendo sido feitos 2700 testes, o valor escolhido permitiu identificar todos os reencontros em 574 testes (21.3%). Destes 2700 a premissa necessária ao método *Análise do Custo* da Eq. C.3 verificou-se 1152 vezes. É neste conjunto de testes que a transformação erro/custo se mostra relevante, mostrando-se válida em 49.8% com este factor de custo.

C.3. Conclusões

Verifica-se que a 2ª expressão proposta é a que obtém melhores resultados. Além desta as 4ª e 5ª expressões também se mostram boas hipóteses. Foi ainda calculado um intervalo de valores para o factor de custo e escolhido um para futuras experiências.

De notar que esta análise de desempenho é feita nos dados reunidos e não é de modo algum geral. Apenas serviu para comparação entre as várias hipóteses.

Anexo D. Efeito dos Reencontros ao Estimar Matrizes Completas

D.1. Introdução

Pretende-se analisar os resultados da estimação da informação desconhecida quando se consideram reencontros.

Utilizam-se dados artificiais de modo a controlar o ruído, o número de pontos, a percentagem de informação conhecida, o número de reencontros e sua posição. Obtém-se o valor do erro de estimação dos métodos EM e TS [7] como na Eq. D.1 para cada reencontro possível de recriar. São comparados os resultados de criar reencontros certos e errados.

$$e = \left\| (\mathbf{W}_{0,R} - \mathbf{W}_R) \odot \mathbf{M} \right\|_F \quad \text{Eq. D.1}$$

A Fig. D.1 ilustra as matrizes utilizadas. Cria-se uma matriz \mathbf{W} (a) de característica 4, à qual se impõe períodos de oclusão em vários pontos, formando assim a matriz sem reencontros $\mathbf{W}_{0,R}$ (b). Introduzem-se novos pontos no final de $\mathbf{W}_{0,R}$ que são reencontros de pontos já seguidos, obtendo-se a matriz \mathbf{W}_0 (c). Existe uma matriz $\mathbf{W}_{0,R}$ (d) que identifica os reencontros criados.

A matriz \mathbf{M} tem tamanho $2F \times P$ e verifica $\mathbf{M} = [m_{ij}]$ tal que

$$m_{ij} = \begin{cases} 1, & \text{se } w_{0ij} \text{ conhecido} \\ 0, & \text{se } w_{0ij} \text{ desconhecido} \end{cases}$$

e \odot é o produto ponto-a-ponto, também conhecido por produto de Hadamard.

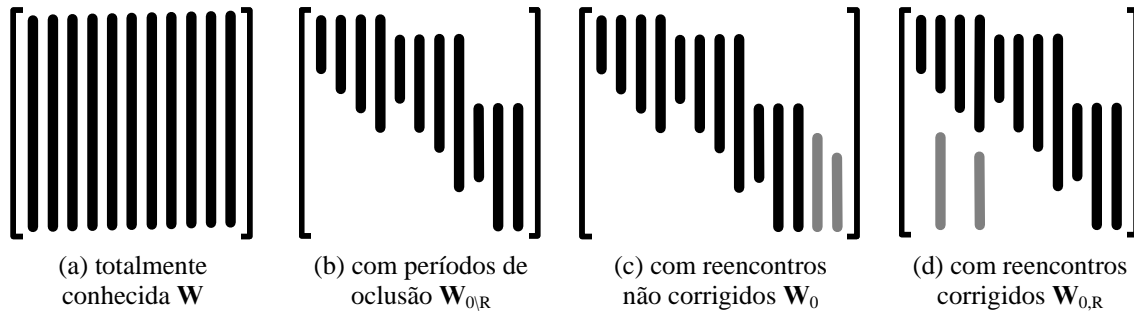


Fig. D.1: Exemplos das matrizes de observação utilizadas

Os reencontros possíveis são identificados como pares do tipo $(p_{1^\circ \text{seguimento}}, p_{2^\circ \text{seguimento}})$ com os índices das colunas do primeiro e segundo períodos de seguimento.

D.2. Evolução do Erro da Estimação

Interessa conhecer o andamento do erro da estimação da informação desconhecida. Foram realizadas várias experiências com este propósito. Dado que os resultados de cada uma são parecidos, descrevem-se apenas duas experiências para ilustrar as conclusões.

Em cada experiência observam-se os resultados numa tabela e num gráfico. Não interessa observar valores muito elevados mas sim os que permitam comparar o erro sem reencontros e_0 e o erro com reencontros $e(p_{1^\circ \text{seguimento}}, p_{2^\circ \text{seguimento}})$ certos e errados.

As experiências foram realizadas com matrizes de observação de característica 4 geradas a partir de dados aleatórios. Estes dados correspondem a coordenadas de pontos em imagens

de tamanho 160×120 pixels. Testou-se uma gama pertinente de valores de desvio padrão do ruído, de percentagem de informação conhecida, de número de imagens, de pontos e de reencontros.

Aplica-se ruído gaussiano de média nula variando o desvio padrão σ , que é calculado como uma percentagem do máximo dos dados da matriz \mathbf{W} .

1ª Experiência

Matriz \mathbf{W} com tamanho 40×33 com ruído de $\sigma = 0.0001 \times \max(\mathbf{W})$. Criam-se 2 reencontros, tendo \mathbf{W}_0 o tamanho 40×35 e 41,2% de informação conhecida.

Tab. D.1: Erros da estimação de informação desconhecida com reencontros – 1ª experiência

Colunas com reencontros		12	16		
		$P_{1^\circ \text{seguimento}}$	$P_{2^\circ \text{seguimento}}$	Número de Colunas P	Erro
Resultados sem juntar reencontros	sem reencontros	0	0	35	0.023708
	reencontros certos	16	35	34	0.023834
		12	34	34	0.02394
	reencontros errados	7	18	34	0.34085
		1	28	34	0.72392
		4	21	34	0.74429
		7	34	34	0.77173
		5	17	34	0.86179
8		21	34	0.89205	
Resultados juntando 1 reencontro	sem reencontros	0	0	34	0.023834
	reencontro certo	12	34	33	0.024032
	reencontros errados	7	18	33	0.34126
		4	21	33	0.7493
		7	34	33	0.77328
		1	28	33	0.80601
		5	17	33	0.86342
		8	21	33	0.90655
Resultados juntando 2 reencontro	sem reencontros	0	0	33	0.024032
	reencontros errados	7	18	32	0.35101
		4	21	32	0.77381
		1	28	32	0.8382
		5	17	32	0.8649
		8	21	32	0.9222
		5	24	32	0.94054

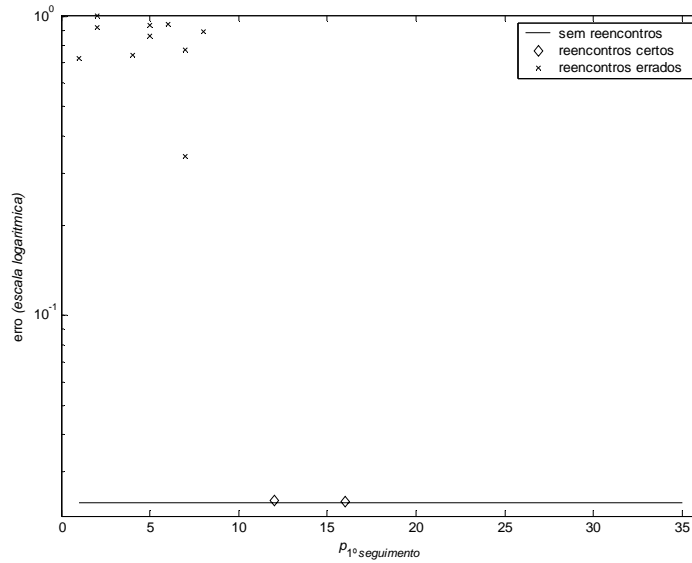


Fig. D.2: Erros da estimação de informação desconhecida com reencontros – 1ª experiência

2ª Experiência

Matriz \mathbf{W} com tamanho 70×40 com ruído de $\sigma = 0.01 \times \max(\mathbf{W})$. Criam-se 2 reencontros, tendo \mathbf{W}_0 o tamanho 70×42 e 47,7% de informação conhecida.

Tab. D.2: Erros da estimação de informação desconhecida com reencontros – 2ª experiência

Colunas com reencontros		7	10
-------------------------	--	---	----

		$p_{1^\circ \text{seguinto}}$	$p_{2^\circ \text{seguinto}}$	Número de Colunas P	Erro
Resultados sem juntar reencontros	sem reencontros	0	0	42	5.7852
	reencontros certos	10	41	41	5.7831
		7	42	41	5.7909
	reencontros errados	17	42	41	5.892
		4	42	41	5.9151
		15	42	41	5.9457
		8	29	41	6.0427
		7	27	41	6.0495
10	33	41	6.05		
Resultados juntando 1 reencontro	sem reencontros	0	0	41	5.7909
	reencontro certo	10	41	40	5.9238
	reencontros errados	10	33	40	6.0054
		8	29	40	6.0387
		13	32	40	6.0849
		10	38	40	6.1696
		2	41	40	6.2808
1	36	40	6.3078		
Resultados juntando 2 reencontro	sem reencontros	0	0	40	5.9238
	reencontros errados	8	29	39	6.0273
		13	32	39	6.1934
		1	36	39	6.4369
		12	29	39	6.5151
		6	29	39	6.5247
4	27	39	6.5286		

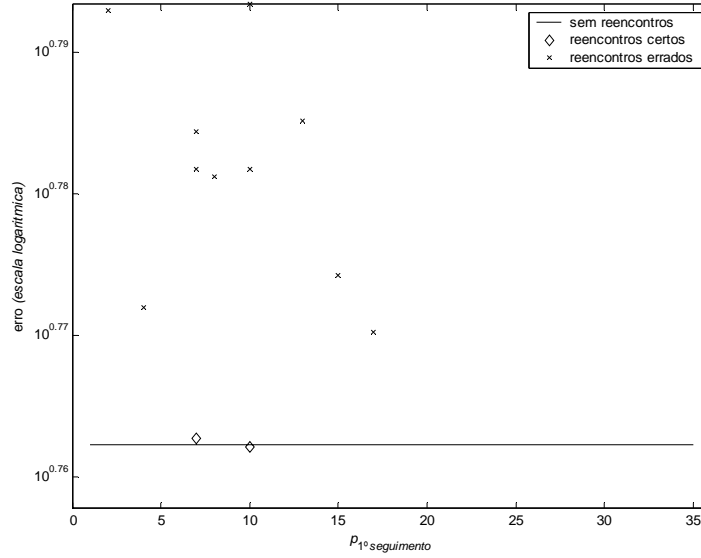


Fig. D.3: Erros da estimação de informação desconhecida com reencontros – 2ª experiência

Conclusões

Observa-se que o andamento do erro com os reencontros testados é igual em várias situações e que existem relações entre os vários erros. Define-se o erro sem reencontros e_0 , o erro de um reencontro certo $e_{reencontro}$ e o erro de um reencontro errado e_{outro} .

As relações que se observam nas tabelas e figuras são

$$\begin{aligned}
 e_{outro} &> e_0 \\
 e_{reencontro} &> e_{outro} \\
 e_{reencontro} &\sim e_0
 \end{aligned}
 \tag{Eq. D.2}$$

Também se observa que estas relações se mantêm ao corrigir apenas alguns dos reencontros.

Na próxima secção verifica-se a validade destas conclusões.

D.3. Relações Entre os Erros dos Reencontros Possíveis

Nesta secção pretende-se realizar experiências para se verificar a validade das conclusões retiradas das experiências anteriores. Interessa conhecer os parâmetros que influem e como influem.

Executam-se testes a várias matrizes, variando o desvio padrão do ruído, a percentagem de informação conhecida, o número de imagens, o número de pontos e o número de reencontros. Cria-se assim uma tabela com os valores dos erros relevantes: o erro sem reencontros e_0 , o menor erro dos reencontros errados e_{errado} e o maior erro dos reencontros certos e_{certo} .

A Tab. D.3 resume os dados reunidos, contabilizado o número de testes em que se verifica cada relação da Eq. D.2.

Tab. D.3: Número de testes em que as relações entre os vários erros se verificam

F [imagens]	P [pontos]	Desvio Padrão do Ruído [pixels]	Informação Conhecida Sem Reencontros [%]	Número de Reencontros [pontos]	Número Total de Testes	Número de testes em que:			
						Situações erradas		Situações certas	
						$e_{errado} < e_0$	$e_{certo} > e_{errado}$	$e_{certo} > e_0$	$e_{certo} < e_0$
21	21	0,01	57,4	5	100	0	0	85	15
21	21	0,10	57,4	5	100	0	11	82	7
21	21	0,10	57,4	6	100	0	8	87	5
21	21	0,10	57,4	7	100	0	17	83	0
21	21	0,10	57,4	8	100	0	9	91	0
21	21	1,00	57,4	5	100	1	77	21	2
21	21	1,00	57,4	7	100	0	88	12	0
21	21	1,00	57,4	9	100	0	86	14	0
21	21	1,00	57,4	11	100	0	87	13	0
21	21	3,00	57,4	5	100	1	94	5	1
21	21	5,00	57,4	5	100	2	96	2	2
21	21	7,00	57,4	5	100	0	99	1	0
21	21	0,01	68,0	5	100	0	0	95	5
21	21	0,10	68,0	5	100	0	2	93	5
21	21	1,00	68,0	5	100	1	43	56	1
21	21	1,00	68,0	7	100	1	60	39	1
21	21	1,00	68,0	9	100	0	63	37	0
21	21	1,00	68,0	11	100	0	62	38	0
21	21	3,00	68,0	5	100	2	94	4	2
21	21	5,00	68,0	5	100	2	95	2	3
21	21	7,00	68,0	5	100	0	95	5	0
21	41	1,00	68,0	5	100	0	46	53	1
21	41	1,00	68,0	7	100	0	53	47	0
21	41	1,00	68,0	9	100	1	69	30	1
21	41	1,00	68,0	11	100	0	75	25	0
21	61	1,00	68,0	5	100	0	46	53	1
21	61	1,00	68,0	7	100	0	62	38	0
					2700	11	1537	1111	52
						0,40%	56,9 %	41,1 %	1,9 %

Observa-se que as relações da Eq. D.2 se verificam para situações com pouco ruído. O aumento do número de reencontros dificulta a estimação dos dados desconhecidos, logo também dificulta que as relações aconteçam. Isso acontece porque a coluna de um reencontro tem normalmente poucos dados conhecidos, como ilustrado na Fig. D.1.

D.4. Conclusões

Observa-se que o erro com reencontros certos é menor do que com reencontros errados, existindo um intervalo que os separa. Também se observa que o erro sem reencontros é menor do que com reencontros errados. O valor do erro com reencontros certos é próximo do erro sem reencontros.

$$\begin{aligned}
 e_{errado} &> e_0 \\
 e_{certo} &> e_{errado} \\
 e_{certo} &\sim e_0
 \end{aligned}
 \tag{Eq. D.3}$$

A Fig. D.4 ilustra estas conclusões, que se mantém ao juntar cada reencontro certo.

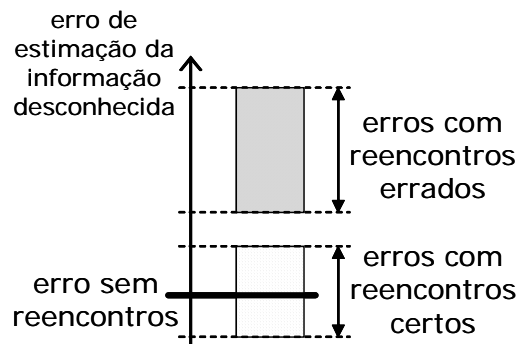


Fig. D.4: Erro na estimação de matrizes de observação completas com reencontros
 Valor sem reencontros e gamas de valores com reencontros errados e reencontros certos.

Estas conclusões são afectadas pelo ruído, pela percentagem de informação conhecida, pelo número de imagens, pelo número de pontos e pelo número de reencontros.

Anexo E. Referências

- [1] S. Ullman, “The interpretation of visual motion”, MIT Press, Cambridge MA, EUA, 1979
- [2] C. Tomasi e T. Kanade, “Shape and motion from image streams under orthography: a factorization method”, *International Journal of Computer Vision*, vol. 9, no. 2, 137-154, 1992
- [3] C. J. Poelman e T. Kanade, “A paraperspective factorization method for shape and motion recovery”, *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 19, no. 3, 1997
- [4] J. Costeira e T. Kanade, “A Multibody Factorization Method for Motion Analysis”, *International Journal of Computer Vision*, vol. 29, no. 3, 1998
- [5] D. Jacobs , “Linear Fitting with Missing Data for Structure-from-Motion”, *IEEE International Conference on Computer Vision Patern Recognition*, 1997
- [6] M. Maruyama e S. Kurumi, “Bidirectional Optimization for Reconstructing 3D Shape from an Image Sequence with Missing Data”, *IEEE International Conference on Image Processing*, Kobe, Japão, 1999
- [7] R. F. C. Guerreiro e P. M. Q. Aguiar, “3D structure from vídeo streams with pastialy overlapping images”, *IEEE International Conference on Image Processing*, EUA, 2001
- [8] P. M. Q. Aguiar e J. M. F. Moura, “Image Motion Estimation – Convergence and Error Analysis”, *IEEE International Conference on Image Processing*, Japão, 1999
- [9] P. J. Green, “Penalized Likelihood”, *Encyclopaedia of Statistical Sciencies*, Jonh Wiley and Sons, para publicação no volume de actualizações, 1996
- [10] D. A. Forsyth e J. Ponce, “Computer Vision – A Modern Aproach”, Prentice Hall, 2003
- [11] L. G. Shapiro e G. C. Stockman, “Computer Vision”, Prentice Hall, 2001
- [12] W. Niblack, “MDL Methods in Image Analysis and Computer Vision”, *Tutorial presented at IEEE International Conference on Computer Vision Patern Recognition*, New York, EUA, 1993
- [13] J. Rissanen, “Modeling by shortest data description”, *Automatica*, no. 14, 465-471, 1978
- [14] J. Rissanen, “Minimum-Description-Length Principle”, *Encyclopaedia of Statistical Sciencies*, Jonh Wiley and Sons, vol. 5, 523-527, 1985
- [15] J. Shi e C. Tomasi, “Good features to track”, *IEEE International Conference on Computer Vision Patern Recognition*, 1994
- [16] R. F. C. Guerreiro, “Modelação 3D a Partir de Vídeo”, Relatório de Trabalho de Final de Curso, Instituto Superior Técnico, Lisboa, Portugal, 2002
- [17] “The Computer Vision Homepage”, <http://www-2.cs.cmu.edu/~cil/vision.html>
- [18] “Minimum Description Length on the Web”, <http://www.mdl-research.org/>
- [19] <http://www.isr.ist.utl.pt/~aguiar/code.html>