



**UNIVERSIDADE TÉCNICA DE LISBOA**  
**INSTITUTO SUPERIOR TÉCNICO**

**DETECÇÃO E SEGUIMENTO DE ESTRUTURAS LINEARES EM**  
**IMAGENS DE FOTOGRAFIA AÉREA COM BASE EM**  
**MODELOS DE MARKOV/GIBBS**

JORGE AUGUSTO CASTRO NEVES BARBOSA  
(LICENCIADO)

Dissertação para obtenção do grau de Mestre em

ENGENHARIA ELECTROTÉCNICA E DE COMPUTADORES

Lisboa, Setembro de 1994

Tese realizada sob a orientação de

**Victor Alberto Neves Barroso**

Professor Associado do

Departamento de Engenharia Electrotécnica e de Computadores

Instituto Superior Técnico

## **DEDICATÓRIA**

Dedico esta tese à Nanda porque quanto mais não seja continua a ser um prazer poder partilhar com ela uma época e um espaço principalmente agora na companhia do nosso pequeno Carlos Renato a quem também dedico a tese

.

## AGRADECIMENTOS

Gostaria de agradecer muito sinceramente ao Prof. Dr. Victor Barroso, meu orientador científico, toda a amabilidade e disponibilidade que sempre manifestou bem como todas as preciosas ideias e sugestões que deu ao longo do desenvolvimento deste trabalho e sem as quais resultaria um trabalho de menor qualidade.

Também manifesto o meu agradecimento aos Drs. Carlos e Regina Correia pelo apoio precioso que facultaram. À minha mãe por todo o carinho e força de vontade que sempre me incutiu e que mais uma vez se mostrou essencial.

À Nanda gostaria de agradecer não só todo o encorajamento que me deu para a frequência deste mestrado mas também a compreensão, se calhar por vezes tão difícil de dar, que sempre mostrou e que quantas vezes não se traduziu numa sobrecarga de trabalho para eu ficar mais aliviado.

## RESUMO

Este trabalho aborda o problema da detecção de estruturas lineares (estradas) presentes em imagens de fotografia aérea. Usam-se métodos bayesianos baseados num modelo de Markov/Gibbs que incorpora conhecimento prévio sobre a imagem a processar sob a forma de uma representação conjunta dos campos de intensidade e de linha.

Uma vez que a qualidade dos resultados finais depende fortemente da modelização do campo de linha, especial atenção é dedicada a este problema. O modelo desenvolvido integra restrições à detecção de formas geométricas com baixa probabilidade de ocorrência, controlando-se assim a verosimilhança de que as estruturas detectadas correspondam a contornos dos objectos a isolar (estradas).

Consideram-se duas realizações alternativas do algoritmo de optimização: (i) sequencial, onde a detecção de linhas é executada sobre uma segmentação prévia; e (ii) integrada, onde a optimização é feita conjuntamente sobre os campos de intensidade e de linha. Verifica-se que a implementação sequencial tem um desempenho bastante superior ao da integrada.

O método bayesiano suportado pelo modelo proposto é comparado com técnicas clássicas baseadas em modelos heurísticos. Este estudo permite concluir que o recurso a metodologias estatísticas (como a bayesiana), desde que suportadas em modelos *a priori* estruturados de acordo com o problema em causa, conduz a soluções bastante mais eficientes. Dos resultados obtidos, conclui-se ainda que o modelo do campo de linha desenvolvido representa adequadamente o tipo de estruturas a detectar.

**Palavras Chave:** Detecção de Contornos, Campos Aleatórios de Markov, Distribuição de Gibbs, Campo de Intensidade, Campo de Linha

## ABSTRACT

This work addresses the problem of detecting linear structures (such as roads) in aerial photographic images. Bayesian methods based on a Markov/Gibbs model are used. The Markov/Gibbs model describes the prior knowledge as a joint representation of the intensity and line fields.

Since the quality of the solution is strongly dependent on the line field modeling, we focus our attention in this problem. The derived model integrates several constraints related to unlike geometric forms. By doing so, the likelihood that the detected structures correspond to the objects of interest is increased.

Two different implementations of the optimization algorithm are considered: (i) sequential, where the line detection and following is performed on a previously segmented image; and (ii) integrated, where the intensity and line fields are jointly optimized. The superior performance of the former approach is verified.

The bayesian approach supported on the model here proposed is compared with well known techniques based on heuristic models. This study shows that the statistical approach, provided that it relies on structured prior models in accordance with the specific problem under consideration, yields much more efficient solutions.

Based on the results obtained, it can be concluded that the line field model here developed is an adequate representation of the particular structures of interest.

**Keywords:** Edge Detection, Markov Random Fields, Gibbs Distribution, Intensity Field, Line Field

# INDÍCE

## 1. INTRODUÇÃO

1.1 Objectivos -----	1
1.2 Contribuições do Trabalho -----	2
1.3 Generalidades Sobre a Detecção de Estradas em Imagens de Fotografia Aérea -----	4
1.4 Técnicas para a Detecção de Estruturas Lineares -----	8
1.5 Organização do Trabalho -----	10

## 2. TÉCNICAS CLÁSSICAS DE DETECÇÃO DE ESTRUTURAS LINEARES

2.1 Introdução -----	12
2.2 Identificação de Estradas em Imagens Digitalizadas -----	13
2.3 Exemplo de Aplicação -----	18
2.3 Sumário -----	23

## 3. TÉCNICAS BAYESIANAS PARA DETECÇÃO DE ESTRUTURAS LINEARES

3.1 Introdução -----	25
3.2 Modelo de Markov/Gibbs -----	27
3.2.1 Modelo de Markov -----	30
3.2.2 Modelo de Gibbs -----	31
3.2.3 Equivalência entre os Modelos de Markov e de Gibbs -----	32
3.3 Especificação do Modelo <i>A Priori</i> para Detecção de Estruturas Lineares -----	34
3.3.1 Especificação da Função de Potencial Associada ao Campo de Intensidade -----	36
3.3.2 Especificação da Função de Potencial Associada ao Campo de Linha -----	37

3.4 Modelo dos Dados -----	43
3.5 Critério de Optimização -----	44
3.6 Sumário -----	45
<b>4. APLICAÇÃO A IMAGENS DE FOTOGRAFIA AÉREA</b>	
4.1 Introdução-----	47
4.2 Algoritmo de Optimização -----	48
4.3 Resultados da Aplicação a Imagens de Fotografia Aérea -----	49
4.3.1 Abordagem Sequencial -----	50
4.3.2 Abordagem Integrada -----	52
4.4 Análise de Resultados -----	55
4.5 Comparação com as Abordagens Clássicas -----	56
4.6 Sumário -----	57
<b>5. CONCLUSÕES</b>	
5.1 Introdução -----	60
5.2 Trabalho Desenvolvido -----	60
5.3 Comparação de Resultados Obtidos -----	62
<b>BIBLIOGRAFIA -----</b>	<b>65</b>
<b>APÊNDICE A -----</b>	<b>67</b>

# 1. INTRODUÇÃO

## 1.1 Objectivos

O objectivo do presente trabalho é a detecção de estruturas lineares em imagens de fotografia aérea digitalizadas com base em modelos estatísticos em que o conhecimento *a priori* é codificado por um modelo de Markov/Gibbs.

A detecção de estruturas lineares pode ser considerada como o primeiro nível do problema mais vasto da extracção automática de, por exemplo, redes viárias ou hidrográficas em fotografias aéreas. Esta extracção é feita normalmente em dois níveis de processamento. No primeiro nível faz-se a detecção e seguimento de todas as estruturas lineares presentes na imagem e, no segundo nível, a sua classificação, por exemplo, como estradas ou rios. Se considerarmos a morosidade das operações cartográficas tradicionais para a elaboração de cartas e mapas, a perspectiva de uma extracção automática deste tipo de redes é bastante interessante e pode enquadrar-se no conceito mais amplo de Cartografia Automática.

No presente trabalho, restringimo-nos ao primeiro nível de processamento referido, o da detecção de estruturas lineares. Para isso, desenvolvemos uma abordagem do tipo bayesiano baseada em modelos estocásticos de Markov/Gibbs que codificam o conhecimento *a priori*. No modelo desenvolvido integramos termos que constituem restrições à detecção de estruturas geométricas que não são típicas dos objectos que se pretende isolar. Como se demonstrará ao longo deste trabalho, esta abordagem estruturada mostra-se adequada à resolução do problema considerado.

No sentido de obter termos comparativos que permitam avaliar o desempenho dos métodos de processamento estatístico propostos, a tese considera também a aplicação de outro tipo de técnicas. Em particular, será apresentada uma síntese das técnicas mais usuais em processamento de imagem com vista à extracção e interpretação de estruturas lineares. Estas técnicas (aqui designadas por clássicas) usam heurísticas deduzidas das imagens a processar e são baseadas em modelos não estruturados, contrariamente ao modelo aleatório que será desenvolvido.

Da comparação dos resultados obtidos com estes dois tipos de abordagens, clássica e bayesiana, tornar-se-á evidente que os obtidos com a abordagem bayesiana baseada no modelo de Markov/Gibbs desenvolvido são, sem dúvida, bastante melhores do que os obtidos com a abordagem clássica.

## **1.2 Contribuições do Trabalho**

A principal contribuição deste trabalho decorre do desenvolvimento de uma abordagem bayesiana baseada em modelos de Markov/Gibbs ao problema da detecção de estruturas lineares em imagens digitalizadas. A eficiência da solução proposta constitui uma contribuição importante para a solução do problema mais complexo, e relevante do ponto de vista da aplicação em Cartografia Automática, da extracção automática de redes viárias (e/ou hidrográficas) a partir do processamento de imagens digitalizadas.

Considerando a grande flexibilidade da especificação dos campos de Markov/Gibbs, é forçoso que o modelo *a priori* seja o mais adequado tendo em vista o objectivo final do processamento. Neste trabalho, desenvolve-se um modelo que integra formalmente restrições geométricas impostas pela natureza dos contornos a detectar. Fica demonstrado, pela análise dos resultados obtidos, que é possível deste modo, e no nível mais baixo de

processamento, controlar de modo efectivo a verosimilhança de que as estruturas lineares detectadas correspondam a contornos de estradas.

Do estudo feito, torna-se claro que a qualidade dos resultados é sensível ao tipo de distribuição de probabilidade *a priori* assumida para representar o campo de intensidades da imagem. Evidencia-se que o modelo gaussiano não é o mais adequado quando, como é o caso, o objectivo consiste na detecção das estruturas lineares presentes na imagem. Ao contrário do alisamento (*smoothing*) do campo de intensidades forçado pelo modelo gaussiano, é possível, como se ilustra na tese, recorrer a outras distribuições de probabilidade que modelizam do modo mais eficiente as descontinuidades.

A abordagem usada possibilita considerar dois tipos de implementação do detector bayesiano de estruturas lineares. Mostra-se que a implementação sequencial (segmentação óptima seguida da estimação do campo de linhas) é preferível, apesar de custos computacionais adicionais, a uma implementação integrada.

Embora constituindo uma contribuição complementar ao corpo fundamental deste trabalho, a síntese de técnicas clássicas usadas na abordagem ao problema considerado permitiu identificar uma metodologia alternativa à proposta na tese e conducente a resultados referenciados como bons.

Os resultados da aplicação dos métodos bayesianos a imagens de fotografia aérea permitem concluir da validade e adequação do modelo desenvolvido à resolução do problema em questão, em particular, quando comparados com os obtidos por aplicação dos métodos heurísticos clássicos.

### 1.3 Generalidades Sobre a Detecção de Estradas em Imagens de Fotografia Aérea

Investigações sobre a detecção de estradas em fotografias aéreas, obtidas quer através de satélites, quer através de aviões, têm sido realizadas nas últimas duas décadas. Começaram por ser efectuadas sobre as primeiras imagens obtidas com o satélite *ERTSI*, actualmente designado por *Landsat*, sobre fotografias obtidas com aviões e, mais recentemente, sobre imagens obtidas com os satélites *SPOT* e os últimos satélites da série *Landsat*.

Os primeiros trabalhos, por assentarem em fotografias de baixa resolução, por exemplo fotografias do *ERTSI*, não permitiam a detecção de estradas relativamente pouco importantes mas tão somente de auto-estradas ou outras vias mais importantes. Convém não esquecer, como atrás referido, que as fotografias do *ERTSI* eram de baixa resolução, concretamente, cada pixel destas imagens correspondia a uma área de terreno de  $57 \times 79 \text{ m}^2$  o que, à partida, limitava as estruturas passíveis de serem detectadas e analisadas. À medida que se foram disponibilizando fotografias de maior resolução, passou-se a poder detectar e analisar outras estruturas mais pequenas tais como estradas secundárias e outras estruturas feitas pelo homem. Tal facto, deveu-se não só ao aparecimento de coberturas fotográficas realizadas a grandes escalas e um pouco por toda a parte, efectuadas para a elaboração de cartografia apoiada por fotogrametria, mas também pelo facto das próprias fotografias de satélite terem aumentado de resolução. Por exemplo, com o *Landsat 5* a resolução no terreno passou para os 30 metros e com o *SPOT* passou para os 20 metros no modo pancromático, e para os 10 metros no monocromático. Estes factos, aliados ao próprio desenvolvimento do processamento de imagem e à sua aplicação em muitas outras áreas científicas, para as quais foram desenvolvidos equipamentos e técnicas posteriormente adaptados, permitiram reorientar as investigações em curso e levar os investigadores a pensar em termos de uma "cartografia automática". Este conceito, surge assim, no âmbito da cartografia, como a possibilidade de, a partir de imagens digitalizadas da superfície terrestre, se poderem obter, sem intervenção da

pessoa humana, cartas e mapas. Considerando, por um lado as rápidas transformações que a superfície da Terra sofre constantemente devido à acção do homem, e por outro lado a morosidade das operações para levantamentos cartográficos, a ideia de uma cartografia automática surge como uma perspectiva de grande interesse pois pode tornar bastante mais rápida a obtenção de cartas e mapas actualizados.

Apesar destas perspectivas bastante promissoras, a sua concretização prática torna-se bastante difícil devido à variedade de casos existentes no terreno e que devem ser todos analisados.

O processamento digital de fotografias é desejável, porque a entrada e a saída de resultados é feita num formato digital. Como já referimos, estes dados podem provir de fotografias obtidas com aviões ou com satélites. Isto implica que os dados obtidos contenham mais ou menos informação pois, entre outras causas, devido à maior altitude a que são tiradas as fotografias de satélites, são geralmente de mais baixa resolução que as fotografias tiradas de avião. Estas fornecem maiores quantidades de informação e permitem uma melhor destrição entre os objectos a analisar. No caso da detecção de estradas, e devido ao facto de algumas delas serem relativamente estreitas, é necessário utilizar fotografias com uma resolução capaz de permitir a sua identificação. Imagens de satélite são consideradas de alta resolução se o tamanho do pixel de imagem corresponder a uma área de terreno menor ou igual a  $30 \text{ m}^2$ . A visibilidade de estradas em fotografias aéreas não depende somente da resolução espacial mas também da sua escala. A altitude e a distância focal da câmara definem a escala da imagem.

Daqui por diante, o termo imagem aérea referir-se-à sempre a imagens fotográficas obtidas a partir de aviões, e posteriormente digitalizadas.

Uma imagem digital é uma representação discreta de uma imagem contínua. A digitalização espacial de uma imagem resulta da amostragem da imagem original, seguida de uma quantificação dos níveis de cinzento.

Deste modo, a gama de intensidades é linearmente dividida num número fixo de níveis de cinzento. O computador representa uma imagem digital numa matriz de níveis de cinzento, possibilitando o seu processamento usando operadores aritméticos e lógicos. A finalidade deste processamento por computador poderá ser:

**Restauração:** Esta operação tem por objectivo corrigir distorções na imagem. Devido às características dos sensores, e aos movimentos dos satélites e aviões (e da Terra) durante o levantamento, podem surgir distorções geométricas e radiométricas;

**Realçamento:** Conjunto de técnicas que têm por objectivo o aperfeiçoamento e a distinção visual de características da imagem. Entre outras, estas técnicas constam da remoção de ruído, realçamento de contornos e manipulação de contrastes;

**Análise:** Técnicas para o reconhecimento e extracção de objectos relevantes na imagem como, por exemplo, estradas, prédios, aeroportos, etc;

**Representação:** A descrição e representação de resultados numa forma apropriada para posterior processamento em computador.

São numerosas as imagens aéreas da superfície da Terra que podem ser caracterizadas por uma composição de texturas modelizadas por processos aleatórios bi-dimensionais. Estas texturas podem representar diversos tipos de terrenos, água, desertos, arborizações, etc. A classificação e a segmentação dessas imagens em secções de textura conhecida é importante para a codificação de imagem e para a cartografia automática. Tanto na classificação como na segmentação é possível determinar uma função de probabilidade que se adeque aos dados da imagem. A detecção de objectos consiste em analisar uma dada zona da imagem e decidir sobre a presença de uma área cujas características diferem de algum modo do meio envolvente.

Neste trabalho, os objectos que nos interessa detectar e seguir são estruturas lineares, por exemplos contornos de estradas. As estradas caracterizam-se por quatro aspectos fundamentais: geometria, radiometria, contexto e função.

- Geometria:

Dimensão - as estradas são estruturas geralmente longas e estreitas. Em geral, uma auto-estrada ou uma estrada importante atravessa toda a imagem. Esta observação é pertinente, pois quando encontramos um objecto com estas características, trata-se geralmente de uma estrada;

Forma - o aspecto geral é o de um objecto muito alongado, de largura constante. Quando uma estrada importante tem uma curva, esta é bastante suave. A forma do objecto em questão é fundamental, pois pode permitir fazer a distinção entre estradas, rios, casas ou campos;

- Radiometria:

Constância - a radiometria da superfície de uma estrada mantém-se geralmente constante. Assim, a partir da superfície de uma estrada é geralmente possível seguir a estrada por propagação;

Contraste entre a radiometria da estrada e a das zonas envolventes - quando uma estrada atravessa uma determinada região, a sua radiometria é diferente da das árvores e campos adjacentes;

Textura - a textura de uma estrada é muito particular, na medida em que é regular e apresenta linhas longitudinais, que podem ser causadas por marcas no pavimento, filas de trânsito, árvores plantadas ao longo da estrada, etc;

- Contexto:

O contexto em que o objecto surge também pode permitir tirar conclusões. Por exemplo, se se conseguir detectar árvores ou casas alinhadas, é muito possível que estas estejam dispostas ao longo de uma estrada;

- Função:

A função de uma estrada permite-nos tirar conclusões como a seguinte: "Se encontrarmos um objecto que se assemelhe a uma estrada e conseguirmos encontrar um veículo, então é possível que se trate de uma estrada".

Estas são algumas das características usadas na implementação de métodos heurísticos para a detecção e identificação de estradas. No entanto, características do tipo geométrico e radiométrica podem também ser usadas formalmente e integrados em modelos matemáticos. Nesta tese, o modelo aleatório que descreve o conhecimento *a priori* sobre a imagem e as estruturas a detectar codifica de forma particular as características atrás referidas.

#### **1.4 Técnicas para a Detecção de Estruturas Lineares**

As técnicas utilizadas para a detecção de estruturas lineares em imagens de fotografia aérea, podem ser divididas em dois grupos: as técnicas clássicas, as quais se baseiam em métodos heurísticos, e as técnicas estruturadas que, combinando modelos lineares com métodos estatísticos, permitem o estabelecimento de modelos que integram as características das imagens como conhecimento *a priori*. Este trabalho dedica especial atenção a este último grupo de técnicas, e baseia-se na classe particular dos modelos de Markov/Gibbs.

As técnicas clássicas, usando modelos *a priori* genéricos baseados em observações experimentais, podem levar a falsas detecções ou a erros de interpretação. Isto resulta do facto de modelos genéricos que estão bem

adaptados a determinado tipo de imagens se revelarem totalmente inadequados para outros.

Com a utilização de técnicas estruturadas baseadas nos modelos de Markov/Gibbs, se bem que à custa de uma maior complexidade computacional, consegue-se obter um processamento mais eficiente, principalmente em imagens de alta resolução e com grande quantidade de informação. Para além disso, estes modelos são mais robustos relativamente à diversidade do tipo de imagens em que podem ser aplicados.

Deste modo, os métodos e algoritmos baseados em modelos não estruturados de conhecimento *a priori*, isto é, decorrentes de heurísticas fundamentadas em observações experimentais, quando comparados com os modelos estatísticos, conduzem a perdas de eficiência qualitativa dos respectivos desempenhos. Este facto é mais evidente em imagens de elevada resolução e com grande quantidade de informação. Torna-se assim necessário recorrer a modelos mais estruturados e complexos que integrem formalmente o conhecimento *a priori* das características das imagens, independentemente da respectiva resolução e conteúdo informativo. Os métodos e algoritmos de processamento baseados em modelos de Markov/Gibbs mostram-se mais adequados para o tipo de problemas aqui considerado. Deste modo, uma das perspectivas de orientação deste trabalho relaciona-se com o tipo de optimização adequado ao problema da detecção e seguimento de estruturas lineares baseada no modelo de Markov/Gibbs. Apesar dos algoritmos envolverem uma grande quantidade de cálculos, a sua estrutura é regular e pouco complexa, tornando acessível a sua implementação computacional.

## **1.5 Organização do Trabalho**

Para além deste primeiro capítulo, o trabalho está organizado como em seguida se descreve. No capítulo 2 faz-se uma descrição das técnicas clássicas de detecção de estruturas lineares mais usuais, as quais recorrem, em geral, a

algoritmos heurísticos. De seguida, e ainda nesse capítulo, descreve-se um algoritmo particular que implementa este tipo de técnicas. Comentam-se ainda os resultados obtidos com a sua aplicação a imagens de fotografia aérea. No capítulo 3, faz-se uma abordagem genérica ao problema da reconstrução de imagens com particular ênfase nos aspectos de modelização estocástica, detalhando-se de seguida a construção do modelo *a priori* de Markov/Gibbs com base nas restrições impostas pela aplicação específica considerada. Os modelos considerados assumem duas distribuições de probabilidade alternativas para representar o campo de intensidades. Define-se também um modelo simplificado dos dados, a partir do qual, e aplicando a teoria bayesiana, se constrói a função de distribuição de probabilidade *a posteriori*. Esta distribuição de probabilidade é usada para definir a funcional de custo (energia) cuja minimização permite obter uma estimativa de *maximum a posteriori*. No capítulo 4, o algoritmo usado na optimização desta funcional de custo é sumariamente descrito. Nesse capítulo, são ensaiadas várias estratégias de optimização da função de energia, testando-se também os dois modelos alternativos considerados para representar o campo de intensidades. Deste estudo, conclui-se, que a hipótese gaussiana associada ao campo de intensidades não é a mais adequada ao problema da detecção de contornos. Comentam-se os resultados obtidos pela aplicação do método a imagens de fotografia aérea, segundo as várias estratégias de optimização atrás referidas. Estes resultados são também comparados com os obtidos no capítulo 2. Verifica-se que o modelo desenvolvido permite a obtenção de excelentes resultados na detecção de estruturas lineares. Conclui-se deste modo que o modelo proposto é válido e bastante adequado ao problema considerado. Em particular, e por comparação directa, evidencia-se que a abordagem bayesiana proposta conduz a níveis de desempenho superiores aos obtidos usando técnicas heurísticas. Finalmente no capítulo 5 referem-se as principais conclusões do trabalho desenvolvido nesta tese.

## **2. TÉCNICAS CLÁSSICAS DE DETECÇÃO DE ESTRUTURAS LINEARES**

### **2.1 Introdução**

As estruturas lineares presentes numa imagem de fotografia aérea poderão corresponder, por exemplo, a redes de estradas ou redes hidrográficas. Em geral, a extracção automática destas redes passa por dois níveis de processamento. Num primeiro nível são detectadas todas as estruturas lineares presentes na imagem, sendo o segundo nível responsável por uma tarefa de classificação a partir da qual se isolam os objectos de interesse.

Neste capítulo aborda-se o problema referido numa perspectiva não estruturada, comum às metodologias mais usuais. Assim, na secção 2.2 apresenta-se uma síntese das técnicas clássicas usadas no problema da extracção automática de redes de estradas em imagens digitalizadas. Na secção 2.3 apresentam-se os resultados obtidos com imagens de fotografia aérea por aplicação de um algoritmo particular, referenciado como um dos que exibem melhor desempenho. Estes resultados serão usados mais tarde para efeitos de comparação com aqueles obtidos usando a abordagem estruturada, do tipo bayesiano, desenvolvida na tese. Finalmente, na secção 2.3 faz-se um resumo das principais conclusões deste capítulo.

## 2.2 Identificação de Estradas em Imagens Digitalizadas

Em geral, as técnicas clássicas usadas na resolução deste tipo de problema seguem uma metodologia que pode ser esquematizada como se mostra na figura 1.

Considerando, a título de exemplo, o problema da extracção automática de redes de estradas por processamento de imagens digitalizadas, podem identificar-se dois níveis de análise:

- Segmentação;
- Interpretação de Estradas.

No nível de Segmentação particiona-se a imagem em várias regiões de acordo com um critério de homogeneidade pré-definido. As técnicas de segmentação são diversas dependendo dos modelos *a priori* considerados. O nível de Interpretação de Estradas controla o nível de Segmentação. Em presença das estruturas lineares detectadas, são testadas algumas características típicas de uma estrada, tais como a presença de cruzamentos, o efeito de sombras, propriedades geométricas, etc., as quais, uma vez verificadas, permitem identificar as estruturas de interesse.

Na segmentação consideram-se normalmente duas fases:

Detecção;

Seguimento de Linhas.

Nos métodos mais usuais de segmentação, as fases de detecção e seguimento de linhas podem ser executadas sequencialmente ou simultaneamente por recurso a técnicas integradas.

# **ANÁLISE DA IMAGEM**

## **SEGMENTAÇÃO**

### **ABORDAGEM SEQUENCIAL**

#### **DETECÇÃO**

LIMIAR DE NÍVEL DE CINZENTO

DETECÇÃO DE DESCONTINUIDADES

DETECÇÃO DE LINHAS

#### **SEGUIMENTO DE LINHAS**

SEGUIDOR ESTRUTURAL DE LINHAS

"EDGE RELAXATION"

PROGRAMAÇÃO DINÂMICA

### **TÉCNICAS INTEGRADAS**

#### **DETECÇÃO DE PONTOS DE PARTIDA**

OPERADOR DE ANÁLISE DE PERFIL

#### **SEGUIMENTO DE LINHAS**

ANÁLISE DE PERFIL BASEADA NA CORRELAÇÃO

SEGUIDOR DE LINHAS BASEADO EM DESCONTINUIDADES

## **INTERPRETAÇÃO DE ESTRADAS**

COMBINAÇÃO DE CONHECIMENTOS

EXTRACÇÃO DE PORMENORES

CONTROLO

Fig. 1

Na fase de detecção a imagem é analisada pixel a pixel, tomando-se decisões para cada um deles. A detecção pode ser executada por comparação com um limiar de cinzento (*threshold*) [2], por determinação de descontinuidades (*edges*) entre os pixels [1] [13], ou ainda com base na verificação da correlação entre uma determinada máscara, correspondente a um tipo de segmento de linha, e uma dada vizinhança do pixel em análise [6] [10].

Quando se usam os limiares de níveis de cinzento, os pixels são seleccionados tendo em atenção o seu valor na escala de cinzento da imagem. Estas técnicas comparam o valor de cinzento de cada pixel com um valor de referência pré-definido, o valor de limiar, resultando uma imagem binária onde, idealmente, os pixels pertencentes à estrutura linear aparecem destacados. É frequente basear a escolha do valor de referência no histograma dos níveis de cinzento da imagem. Num histograma bimodal, o valor de referência é escolhido como a média dos dois valores extremos, implementando-se a solução de máxima verosimilhança do problema clássico de detecção binária em presença de ruído gaussiano aditivo. A eficiência desta solução é tanto maior quanto mais próxima do modelo binário subjacente estiver a imagem original. No entanto, em imagens de fotografia aérea, o modelo binário é evidentemente pouco realista, pelo que, em geral, os resultados obtidos com este tipo de técnicas não são satisfatórios. As falsas detecções são consequência do facto de se usarem medidas absolutas de intensidade de cada pixel em vez de se considerarem as diferenças relativas a um conjunto de pixels vizinhos. Este tipo de medidas de descontinuidades que usam estas diferenças relativas são usados nos chamados detectores de descontinuidades (*edge detectors*) e conduzem a soluções mais eficientes para o tipo de problemas em causa.

Outros operadores procuram segmentos de linha em pequenas vizinhanças de cada pixel e são designados por detectores de linhas. Por exemplo, pode fazer-se a convolução da vizinhança de determinado pixel com uma máscara, normalmente binária, que modeliza determinado tipo de

segmento de linha e, em face do resultado obtido, assume-se a existência ou não do segmento de linha correspondente a essa máscara.

Uma vez detectados os pixels pertencentes às fronteiras de estruturas lineares, procede-se ao seguimento dessas linhas. Para isso, os chamados operadores de seguimento de linhas ligam entre si os pixels candidatos a pertencerem a essa estrutura linear. Os operadores do tipo seguidor estrutural de linhas usam a informação estrutural da imagem binária para tomar as suas decisões sobre ligar ou remover pixels ainda não ligados e que são candidatos a pertencerem à estrutura linear [2] [10] [13]. Na técnica designada por *edge relaxation* utiliza-se informação sobre a vizinhança de descontinuidades para determinar se o pixel pertence à linha [1]. A programação dinâmica é outra técnica utilizada para se fazer o seguimento das linhas [6]. É uma técnica onde se pesquisa o caminho óptimo numa região de interesse. Uma rede é considerada como sendo um grafo com valores (custos) atribuídos aos arcos. Para cada caminho é calculado um custo e o objectivo é encontrar o caminho óptimo através da rede correspondente ao custo total mínimo. O caminho final é obtido seguindo o caminho de custo mínimo. Uma limitação desta técnica é que o ponto de partida e o de chegada devem ser indicados previamente.

As técnicas integradas combinam a detecção e o seguimento de linhas [12]. Estas técnicas normalmente constam de dois passos. No primeiro passo, começa-se por detectar pontos de partida da estrutura linear e de seguida comparam-se vizinhanças de pixels com perfis representativos de determinadas estruturas lineares, seleccionando-se deste modo os pontos candidatos a pertencerem à estrutura linear. O segundo passo consiste na predição do próximo ponto pertencente à estrutura linear por comparação com uma máscara cuja correlação com os pixels da imagem é testada. Alternativamente, este segundo passo pode consistir no seguimento da linha por análise das descontinuidades vizinhas. Comparativamente com o método anterior, este método tem a vantagem de usar diferenças relativas de intensidade de pixels vizinhos em vez de usar medidas absolutas de intensidade.

O segundo nível como já foi referido é o da Interpretação de Estradas. Este nível controla o precedente. O resultado da segmentação, que consiste na extracção de primitivas da estrada tais como intersecções, mudanças de material de que ela é revestida, etc., é combinado em elementos de estrada. Para isto, heurísticas do tipo *if-then-else* definidas a partir de um modelo estabelecido *a priori* para a estrada são verificadas. Deste modo são extraídos os pormenores considerados de interesse. Se o pormenor em análise, obtido através de uma determinada técnica de segmentação, conduzir a problemas de identificação, o nível de Interpretação encarrega-se de seleccionar uma ferramenta de segmentação alternativa. Este procedimento pode ser repetido até que se consiga uma identificação correcta. Neste sentido, pode afirmar-se que o nível de Interpretação controla o nível de Segmentação.

O estabelecimento dos modelos *a priori* usados nestas abordagens, baseia-se normalmente em certas propriedades das estradas definidas no contexto da imagem. Ao conjunto destas propriedades dá-se o nome de modelo genérico da estrada. Estes modelos podem ter por base propriedades visuais das estradas deduzidas a partir de propriedades físicas e geométricas como, por exemplo:

- Propriedades Espectrais: correspondentes aos materiais utilizados na construção das estradas (betão, asfalto, rochas, etc.);
- Forma: em geral, uma estrada toma a forma de uma fita estreita com raios de curvatura elevados e largura aproximadamente constante;
- Intensidade: o nível de intensidade dos pixeis pertencentes à estrada contrasta com o do terreno adjacente e é aproximadamente uniforme.

No entanto há que ter alguns cuidados na especificação dos modelos genéricos pois, por exemplo, algumas das propriedades atrás indicadas para estradas podem ser confundidas com propriedades de estruturas hidrográficas, dando origem a eventuais ambiguidades. Como alternativa, pode considerar-

se um modelo organizacional em que um processamento de alto nível é responsável pela validação de hipóteses sobre os objectos detectados por operadores de baixo nível. Deste modo, no processamento de alto nível determina-se a estratégia a adoptar no procedimento de detecção.

Outro dos problemas que poderá surgir no estabelecimento deste tipo de modelos prende-se com a sua aplicabilidade a diferentes tipos de imagens. Por exemplo, um mesmo modelo dificilmente será adequado para representar imagens de diferente resolução.

### **2.3 Exemplo de Aplicação**

Com o objectivo de comparar os desempenhos dos métodos clássicos descritos na secção anterior com os obtidos usando as técnicas baseadas em modelos estatísticos, que serão abordadas em detalhe nos capítulos seguintes, implementou-se sobre imagens de fotografia aérea digitalizadas o algoritmo proposto por Nevatia & Babu em [13]. Escolheu-se este algoritmo por ter servido de base para trabalhos de outros autores [1] [20] com níveis de desempenho considerados bons. Em relação à metodologia descrita na figura 1, neste algoritmo faz-se uma abordagem sequencial. Deste modo, procede-se primeiro à detecção de descontinuidades e depois ao seu seguimento.

A fase de detecção de descontinuidades é executada em dois passos. Começa-se por convolver os pixels vizinhos de cada um dos pontos da imagem com 6 diferentes máscaras 5x5, cada uma delas modelizando uma descontinuidade com uma determinada direcção (ver figura 2). De seguida, selecciona-se a direcção correspondente à máscara com a qual se obteve o valor máximo para as 6 convoluções calculadas. Deste modo, obtêm-se valores de amplitude e direcção das eventuais descontinuidades existentes em cada um dos pontos da imagem, os quais são guardados para posterior processamento.

É vulgar que as linhas detectadas em processamento de imagem tenham espessuras superiores a um pixel procedendo-se então a um adelgaçamento dessa linha de modo a reduzir a sua espessura a um pixel. Esta operação é designada por *thinning*. Por outro lado, a comparação de valores de cinzento com um dado limiar de cinzento é aqui designada por *thresholding*.

De seguida, e afim de se determinar as descontinuidades existentes, é feito em simultâneo, e com base nos valores de amplitude e de direcção obtidos no passo anterior, um *thinning* e *thresholding* da imagem. Esta selecção é feita de acordo com as seguintes regras:

- 1) A amplitude da descontinuidade desse pixel é maior do que as amplitudes das descontinuidades dos pixels vizinhos segundo uma direcção normal à dessa descontinuidade;
- 2) As direcções dos dois pixels vizinhos não diferem em mais do que uma unidade (30 graus) em relação ao pixel central;
- 3) A amplitude do pixel central excede um determinado valor de limiar;

Se as condições 1 e 2 forem verificadas em simultâneo, os dois pixels vizinhos são imediatamente desqualificados como candidatos a descontinuidades. Assim, das 8 descontinuidades vizinhas são escolhidas as que tenham uma direcção que não difira mais de 30 graus e que simultaneamente tenha um valor de amplitude superior ao de limiar.

-100	-100	0	100	100
-100	-100	0	100	100
-100	-100	0	100	100
-100	-100	0	100	100
-100	-100	0	100	100

a) 0 graus

-100	32	100	100	100
-100	-78	92	100	100
-100	-100	0	100	100
-100	-100	-92	78	100
-100	-100	-100	-32	100

b) 30 graus

100	100	100	100	100
-32	78	100	100	100
-100	-92	0	92	100
-100	-100	-100	78	32
-100	-100	-100	-100	-100

c) 60 graus

100	100	100	100	100
100	100	100	100	100
0	0	0	0	0
-100	-100	-100	-100	-100
-100	-100	-100	-100	-100

d) 90 graus

100	100	100	100	100
100	100	100	78	-32
100	92	0	-92	-100
32	-78	-100	-100	-100
-100	-100	-100	-100	-100

e) 120 graus

100	100	100	32	-100
100	100	0	-100	-100
100	100	0	-100	-100
100	-78	-100	-100	-100
-100	32	-100	100	-100

f) 150 graus

Fig. 2

A fase seguinte (seguimento) consiste em determinar as descontinuidades vizinhas a que cada uma das descontinuidades deve ser ligada. Para isso, começa-se por determinar os predecessores e sucessores de cada elemento usando-se a informação atrás obtida sobre as direcções dos 8 vizinhos. Só serão então consideradas as descontinuidades cujas direcções sejam iguais ou, quando muito, diferirem de 30 graus. Podem ocorrer 3 situações:

- Só um elemento é candidato a sucessor (predecessor);
- Dois elementos são candidatos a sucessores (predecessores) e no caso de se ter uma bifurcação para além de se ter um sucessor (predecessor) primário, ter-se-à também um sucessor (predecessor) secundário;
- Três elementos são sucessores (predecessores) candidatos. Neste caso, temos sempre uma bifurcação sendo a cadeia principal formada pela mais próxima das duas descontinuidades que tenham a mesma direcção. O outro elemento que tem a direcção diferente formará o outro ramo.

Finalmente procede-se ao traçado dos segmentos usando-se uma aproximação linear. O traçado dos segmentos é efectuado em três etapas:

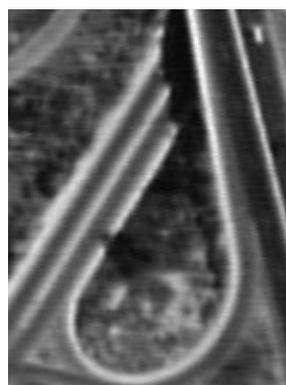
- Na primeira etapa, o traçado começa nas descontinuidades que não tenham predecessores. Se uma descontinuidade tiver 2 sucessores é escolhido o sucessor primário. O traçado deste segmento termina quando for encontrado uma descontinuidade que não tenha sucessor.
- Na segunda etapa, o traçado começa nos sucessores secundários das descontinuidades e prossegue como na etapa anterior.
- Na terceira etapa, o traçado começa nas descontinuidades que ainda não estejam contidas em nenhum traçado e prossegue como na primeira etapa.

Cada segmento determinado em qualquer uma destas 3 etapas, é aproximado por pequenos segmentos lineares segundo um método de aproximação linear descrito em [5].

Este algoritmo foi aplicado às imagens de fotografia aérea mostradas nas figuras 3 a) e b). Estas duas figuras referem-se a 2 imagens digitalizadas em 256 níveis de cinzento e tendo 190x150 pixels.



a) Imagem 1



b) Imagem 2

Fig. 4

Os resultados obtidos são visualizados nas figuras 4 e 5, respectivamente, para a imagem 1 e 2. Em 4 a) e 5 a) visualiza-se a saída dos blocos de *thresholding* e *thinning* e em 4 b) e 5 b) representam-se os resultados finais obtidos.



a)



b)

Fig. 4

a) Saída dos Blocos de Thinning e Thresholding

b) Resultado Final

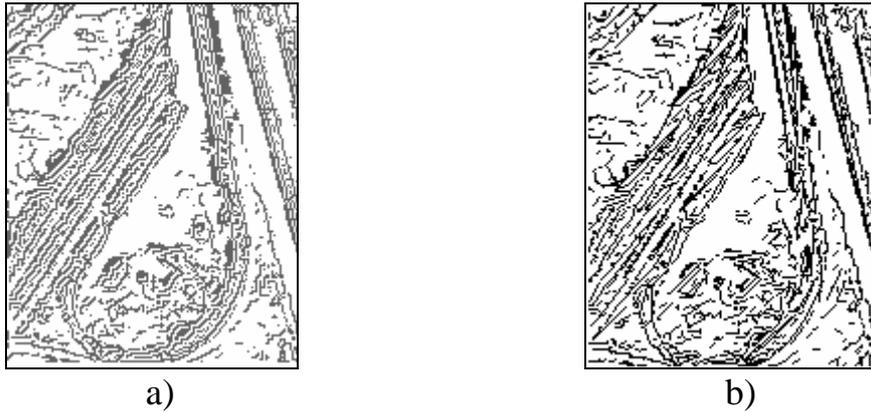


Fig. 5

- a) Saída dos Blocos de Thinning e Thresholding  
b) Resultado Final

A análise dos resultados permite constatar que o passo intermédio de *thresholding* e *thinning* apresenta um bom desempenho. As estruturas lineares representadas pela bifurcação da estrada e pelo nó da auto-estrada são detectadas. A fase final, de traçado dos segmentos revela-se extremamente sensível a falsas detecções provenientes da fase anterior de processamento, originando falsos segmentos e bifurcações não existentes. Dos resultados obtidos nos dois casos, conclui-se que a grande quantidade de informação disponível se torna responsável pela detecção de linhas não relevantes, tendo em atenção o objectivo pretendido. Isto resulta da não utilização de um modelo estruturado para a imagem o que origina falsas detecções de estruturas aproximadamente lineares.

### 2.3 Sumário

Neste capítulo foram descritas algumas das técnicas clássicas usadas em processamento de imagens de fotografia aérea com vista à extracção e interpretação de estruturas lineares. Foi ainda implementado um algoritmo baseado numa destas técnicas.

Como se pode constatar pelos resultados obtidos, uma das dificuldades inerentes aos métodos clássicos para detecção de estruturas lineares é que estes se mostram bastante sensíveis a falsas detecções. Isto deve-se principal-

mente a eventuais interacções entre as diferentes fases de processamento que podem levar a erros de interpretação nas fases subsequentes.

Outra das dificuldades inerentes a estes métodos é o estabelecimento dos modelos genéricos baseados em conhecimento *a priori* obtido a partir de observações experimentais.

Na aplicação considerada, tornou-se evidente a dificuldade do método implementado em lidar com a grande quantidade de informação presente nas fotografias aéreas de grande escala. Este facto traduz-se pelo elevado número de falsas detecções presentes nos resultados. Em contrapartida, pode afirmar-se que o método ensaiado envolve baixos custos computacionais.

Em alternativa às técnicas consideradas, pode-se pensar que abordagens que utilizam modelos estruturados da imagem poderão ser mais adequadas para o processamento de imagens de alta resolução e grande quantidade de informação. Os métodos baseados em modelos estatísticos do tipo Markov/Gibbs permitem estruturar o problema do processamento na medida em que integram formalmente o conhecimento prévio sobre as características das imagens a processar. Por outro lado, estes modelos não dependem da escala das imagens de fotografia aérea que aqui se consideram. Com base nestas premissas, é possível prever que este tipo de abordagem conduza a resultados mais eficientes em comparação com os que foram obtidos usando o algoritmo descrito neste capítulo.

### **3. TÉCNICAS BAYESIANAS PARA DETECÇÃO DE ESTRUTURAS LINEARES**

#### **3.1 Introdução**

Como referido anteriormente, na maior parte das arquitecturas de processamento de imagem é possível distinguir dois níveis de processamento.

No processamento de baixo nível, a representação de uma imagem é transformada utilizando operações tais como detecção de descontinuidades e segmentação de regiões. A partir de uma representação numérica, como uma matriz de intensidades dos pixels, é construída uma representação simbólica que usa um conjunto de primitivas da imagem espacialmente relacionadas, tais como descontinuidades e regiões.

No processamento de alto nível é usado o conhecimento do tipo da imagem para atribuir rótulos ou interpretações às primitivas atrás referidas de modo a construir uma descrição "do que está presente na imagem".

Considerado na sua globalidade, o problema particular da detecção de estradas pode também ser decomposto em dois níveis de processamento. Num primeiro nível são detectadas as diversas estruturas lineares presentes na imagem, seguindo-se um processo de classificação, em geral bastante complexo, a partir do qual se podem isolar os objectos de interesse.

Neste trabalho, aborda-se apenas o problema da detecção de estruturas lineares envolvendo, como procedimentos de baixo nível, a segmentação da imagem original e a construção das linhas que delimitam as estruturas candidatas a serem interpretadas como estradas. Estes procedimentos podem ser efectuados conjuntamente ou em passos separados. Em qualquer dos

casos, as técnicas aqui usadas baseiam-se em modelos que incorporam algum conhecimento prévio sobre características específicas dos objectos a isolar. Torna-se assim possível, mesmo no nível mais baixo de processamento, controlar a verosimilhança de que as estruturas lineares detectadas correspondam a estradas. Por outras palavras, o problema é estruturado de forma que, por recurso a métodos bayesianos bem conhecidos, seja maximizada a probabilidade de que as estruturas lineares detectadas correspondam aos objectos de interesse.

A estruturação acima referida passa pela construção de um modelo probabilístico que representa o conhecimento *a priori* sobre a imagem original, bem como pela formalização de hipóteses adicionais sobre as transformações conducentes à imagem disponível para processamento. Este baseia-se no critério de máximo *a posteriori*, isto é, conduz a uma solução que maximiza a probabilidade de ocorrência de uma imagem (incluída na classe representada pelo modelo *a priori*) condicionada nos dados.

Assume-se que o modelo *a priori* é determinado por um campo aleatório de Markov, MRF, cuja estrutura será detalhada mais adiante. Com generalidade, pode afirmar-se que o modelo markoviano descreve as propriedades locais do campo (imagem) sendo extremamente difícil especificá-lo de forma directa. No entanto, mediante a verificação de certas hipóteses de carácter técnico, é possível estabelecer a equivalência entre a descrição markoviana e a distribuição de Gibbs, GD. O modelo de Gibbs constitui uma representação global da imagem e apresenta a vantagem de poder ser especificado de forma mais simples. O modelo considerado representa simultaneamente o campo de intensidades da imagem e o seu dual designado por campo de linhas. Note-se que o campo de linhas não tem realização objectiva na imagem, ao contrário do campo de intensidades que é quantificado pelos níveis de cinzento de cada pixel. A concretização do campo de linhas, subjacente às discontinuidades presentes no campo de intensidades, pode ser fortemente condicionada se na especificação do modelo respectivo se incluírem termos de penalização de características geométricas

atípicas nas estruturas lineares que se pretende detectar. Torna-se assim claro que a modelização adequada do campo de linhas tem uma extrema importância na qualidade dos resultados.

Este capítulo está organizado do seguinte modo. Na secção 3.2 introduz-se de modo genérico o modelo de Markov para campos aleatórios  $2D$  bem como o seu equivalente especificado pela distribuição de Gibbs. A secção 3.3 constitui o corpo fundamental deste capítulo. Nela se especifica o modelo conjunto dos campos de intensidade e de linhas, dedicando especial atenção a este último. O modelo dos dados é apresentado na secção 3.4, e a construção da funcional de custo, distribuição da probabilidade *a posteriori*, é detalhada na secção 3.5. Finalmente, na secção 3.6, resumem-se os tópicos mais importantes discutidos neste capítulo.

### 3.2 Modelo de Markov/Gibbs

Suponhamos que uma imagem digitalizada é representada sobre um suporte dado pela matriz  $M$  de  $N_1 \times N_2$  pixels definindo uma grelha rectangular finita:

$$M = \{(i, j) : 1 \leq i \leq N_1, 1 \leq j \leq N_2\} \quad , \quad (1)$$

ou seja, a matriz  $M$  tem  $N_1$  linhas e  $N_2$  colunas, e o pixel  $(i, j)$  está situado na linha  $i$  e coluna  $j$ .

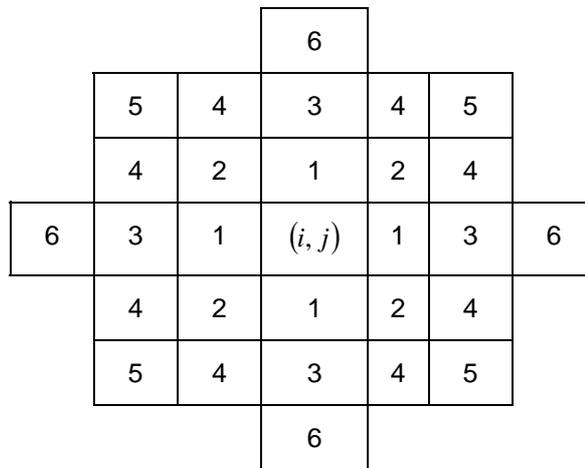
Considera-se que uma colecção de subconjuntos de  $M$ , eventualmente singular, descritos por

$$\eta = \{\eta_{i,j} : (i, j) \in M, \eta_{i,j} \subset M\} \quad (2)$$

constitui um sistema de vizinhanças em  $M$  se, e só se:

- a)  $(i, j) \notin \eta_{ij}$ , isto é, um pixel não pertence à sua própria vizinhança
- b) se  $(k, l) \in \eta_{ij} \Rightarrow (i, j) \in \eta_{kl}$ ,  $\forall (i, j) \in M$ , isto é, se um pixel pertence à vizinhança de outro, então esse outro também pertence à vizinhança do primeiro.

A extensão do suporte de interacção entre pixels vizinhos depende da ordem do sistema de vizinhanças. Deste modo, os sistemas de vizinhanças em relação ao pixel central podem ser ordenados segundo uma hierarquia sequencial. Assim, o sistema de vizinhanças de ordem zero é constituído só pelo pixel central, o de 1ª ordem,  $\eta^1$ , é constituído pelos 4 pixels que têm um lado comum com o pixel  $(i, j)$ , e o de segunda ordem,  $\eta^2$ , é constituído pelos 8 pixels vizinhos de  $(i, j)$  mas agora com um lado ou um canto comum com  $(i, j)$ .



$$\eta^m = \{\eta_{ij}^m\}$$

$$\eta_{ij}^m = \{k : k \leq m\}$$

Fig. 6  
Exemplos de Sistemas de Vizinhança

Com generalidade, podem considerar-se sistemas de vizinhanças de ordem superior, como se mostra na figura 6. No entanto, a complexidade do modelo cresce com a ordem do respectivo sistema de vizinhanças, o que se traduz por custos computacionais que podem ser inoportáveis.

Outro conceito importante a definir é o de *clique*. Dada uma imagem com suporte na matriz  $M$ , e um sistema de vizinhanças  $\eta$ , define-se *clique*  $c$  com base no par ordenado  $(M, \eta)$ , como sendo um conjunto de pixels, eventualmente singular, tais que:

$$(i, j) \neq (k, l), (i, j) \in c \wedge (k, l) \in c \Rightarrow (k, l) \in \eta_{i,j} . \quad (3)$$

De acordo com esta definição, e tendo em conta a definição de vizinhança, verifica-se que se dois pixels diferentes pertencerem à mesma *clique*, então um deles pertence à vizinhança do outro. Designa-se por  $C(M, \eta)$  ou simplesmente  $C$  o conjunto de todas as *cliques* pertencentes a  $(M, \eta)$ .

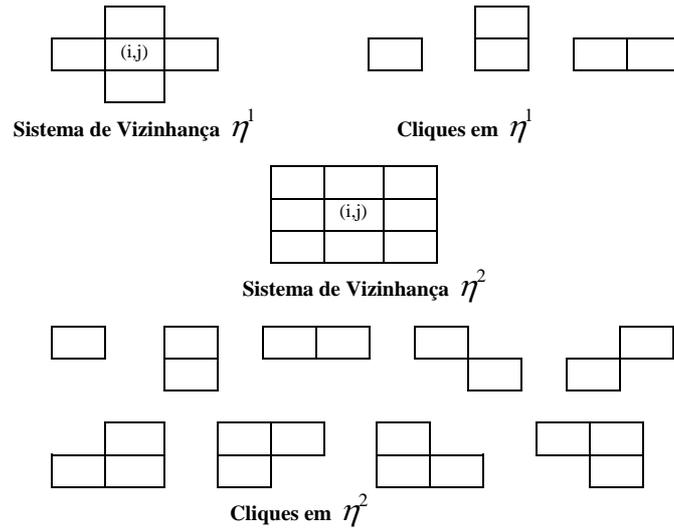


Fig. 7

Sistemas de Vizinhanças de 1ª e de 2ª Ordem e respectivas Cliques Associadas

Usando as definições anteriores, podem enumerar-se os diferentes tipos de *cliques* associadas a cada sistema de vizinhanças. Na figura 7 representam-se os sistemas de vizinhança de 1ª e de 2ª ordem,  $\eta^1$  e  $\eta^2$ , respectivamente, bem como os correspondentes tipos de *cliques*.

### 3.2.1 Modelo de Markov

Considerando um sistema de vizinhanças  $\eta$  definido sobre o suporte  $M$ , e sendo  $P(X = x)$  a probabilidade associada a uma dada realização da imagem, o campo aleatório  $X = \{X_{i,j}\}$  definido em  $M$  é um campo aleatório de Markov relativamente ao sistema de vizinhança  $\eta$  se, e só se,  $\forall \mathbf{a}_f \in M$ :

i)  $P(X = x) > 0, \forall x$

$$\text{ii) } \begin{aligned} &P(X_{ij} = x_{ij} | X_{kl} = x_{kl}, (k, l) \in M, (k, l) \neq (i, j)) = \\ &= P(X_{ij} = x_{ij} | X_{kl} = x_{kl}, (k, l) \in \eta_{ij}) \end{aligned} \quad (4)$$

A primeira condição é conhecida como condição de positividade e a segunda como condição de Markovianidade. Esta propriedade condicional que caracteriza um campo de Markov é chamada de característica local do campo aleatório. Esta característica local dos *MRF* tem uma razão de ser intuitiva: é natural que numa imagem o valor de um pixel não esteja relacionado com o valor dos pixels não pertencentes à sua vizinhança (figura 8). Isto é, existe uma forte relação entre um pixel e os pixels mais próximos sendo possível obter informação sobre um pixel recorrendo apenas à informação sobre os seus vizinhos mais próximos.

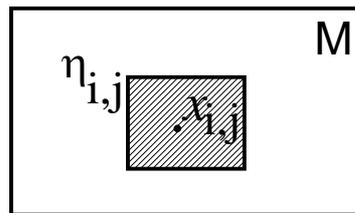


Fig. 8

Uma vez que as imagens têm normalmente uma estrutura regular com distribuições *a priori* homogêneas e isotrópicas, o conhecimento *a priori* sobre as dependências espaciais pode ser perfeitamente codificado num modelo de imagem construído com base em vizinhanças de dimensão suficientemente reduzida. Em termos práticos, e tendo em conta os custos computacionais, este facto é bastante importante do ponto de vista das aplicações práticas, justificando o recurso ao modelo de Markov em problemas de processamento de imagem.

Por outro lado, os modelos *MRF* são muito flexíveis e constituem uma ferramenta matemática poderosa se pensarmos em pequenos sistemas de vizinhança como, por exemplo, os de 1ª ou de 2ª ordem. Para vizinhanças de ordem superior, as dificuldades de caracterização da função de distribuição de

probabilidade  $P(X)$  são razoavelmente superadas se se fizer uso da equivalência entre os modelos de Markov e de Gibbs.

### 3.2.2 Modelo de Gibbs

Seja  $\eta$  um sistema de vizinhanças definido em  $M$ , um campo aleatório  $X = \{X_{i,j}\}$  definido em  $M$  é um campo aleatório de Gibbs, *GRF*, isto é, um campo aleatório com uma distribuição de Gibbs relativamente a  $\eta$ , se, e só se, a sua probabilidade conjunta tiver a forma:

$$P(X = x) = \frac{1}{z} e^{-\frac{1}{T} U(x)}, \quad (5)$$

em que

$$U(X) = \sum_{c \in \mathcal{C}} V_c(x). \quad (6)$$

$U(X)$  é normalmente designado por função de energia e  $V_c(x)$  é o potencial associado com a *clique*  $c$ . A função

$$z = \sum_x e^{-\frac{1}{T} U(x)},$$

designada por função de partição, é simplesmente uma constante de normalização que garante que a probabilidade total de todas as configurações possíveis da imagem seja igual a 1.  $T$  é um parâmetro designado por *temperatura* do sistema, o qual controla o alisamento das distribuições.

A interpretação física da distribuição de Gibbs expressa em (5) é a de que quanto menor for o valor da energia  $U(X)$  correspondente à realização  $X$ , maior é a probabilidade dessa realização, ou seja, aos estados mais prováveis corresponde uma energia menor.

As funções de potencial,  $V_c(x)$ , são funções totalmente arbitrárias associadas a cada *clique*. No entanto, cada função de potencial relaciona entre si os pixels pertencentes a uma determinada *clique* e o seu valor não depende de pixels exteriores a essa *clique*. Tendo em conta as definições (5) e (6), torna-se óbvio que a escolha das funções de potencial  $V_c(x)$  permitem definir uma grande variedade de distribuições de Gibbs.

### 3.2.3 Equivalência entre os Modelos de Markov e de Gibbs

Da discussão anterior é possível retirar dois factos importantes: (i) o modelo de Markov é de grande flexibilidade pois permite descrever as características locais do campo sendo, no entanto, muito difícil especificar de modo coerente as distribuições de probabilidade condicionais que o caracterizam; (ii) o modelo de Gibbs pode ser especificado de modo mais simples e descreve a distribuição conjunta do campo.

A equivalência entre os dois modelos constitui um resultado de grande importância, pois permite fazer uso da simplicidade de especificação da distribuição de Gibbs para, a partir desta, se obterem as distribuições condicionais definidas em (4) e que caracterizam completamente um campo de Markov. Formalmente, este facto traduz-se pelo seguinte teorema [4]:

"Seja  $\eta$  um sistema de vizinhanças definido sobre uma matriz finita  $M$ . Um campo aleatório  $X = \{X_{i,j}\}$  é um campo aleatório de Markov relativamente a  $\eta$  se, e só se, a sua distribuição conjunta for uma distribuição de Gibbs, com *cliques* associadas a  $\eta$ ".

Nestas condições, as distribuições condicionais (4) vêm dadas por

$$\begin{aligned}
P(X_{ij} = x_{ij} | X_{kl} = x_{kl}, (k, l) \in \eta_{ij}) &= \\
&= \frac{e^{-\frac{1}{T} \sum_{c \in C} V_c(x)}}{\sum_{x_{ij}} e^{-\frac{1}{T} \sum_{c \in C} V_c(x)}} \quad (7)
\end{aligned}$$

onde a soma no denominador é feita sobre todos os valores possíveis de  $x_{ij}$ .

Assim, através da utilização de modelos de Markov/Gibbs, podem relacionar-se as probabilidades locais com as probabilidades globais conjuntas e para além disso, a especificação de campos de Markov torna-se conceptualmente mais simples, pois reduz-se a uma especificação de potenciais.

Usando a equivalência *MRF/GD* consegue-se assim que:

- a distribuição conjunta do campo aleatório seja obtida facilmente usando a distribuição de Gibbs ((5) e (6));
- as características locais sejam obtidas facilmente a partir da distribuição conjunta, definindo única e consistentemente um campo aleatório de Markov (7);

Na secção seguinte, o modelo de Markov/Gibbs a usar na detecção de estruturas lineares será detalhado, definindo-se, em particular, as funções de potencial que especificam a distribuição de Gibbs.

### 3.3 Especificação do Modelo *A Priori* para Detecção de Estruturas Lineares

Habitualmente as técnicas utilizadas para deduzir valores estatísticos médios dos *MRF*, consistem na obtenção de funções explícitas dos dados e parâmetros do modelo. Os cálculos analíticos necessários para a obtenção destas expressões são, normalmente, bastante pesados e é costume

considerarem-se algumas aproximações. O problema aqui consiste em aproximar uma superfície a uma grelha de posições  $2D$  regular a partir de dados esparsos e com ruído.

Como vimos pela equivalência  $MRF/GD$ , qualquer campo aleatório de Markov pode ser descrito por uma distribuição de probabilidade de Gibbs, obtendo-se uma função de energia (6) que pode ser calculada a partir das contribuições locais de cada posição  $(i, j)$  da grelha. Uma das maiores potencialidades dos modelos  $MRF$  é o de se poder lidar directamente com descontinuidades introduzindo um campo binário, o chamado campo de linha [8].

O campo de linha pode ser definido sobre uma grelha dual, representando explicitamente a presença ou ausência de descontinuidades. Assim, a grelha dual é conjugada da grelha do campo de intensidades, no sentido em que a cada ponto da grelha do campo de intensidades correspondem dois pontos na grelha dual: um ponto correspondendo a uma linha vertical e outro correspondendo a uma linha horizontal (ver figura 9). Define-se assim um processo de linha horizontal  $h_{i,j}$ , o qual liga a posição  $(i, j)$  com a posição  $(i-1, j)$ , e o processo de linha vertical  $v_{i,j}$  que liga a posição  $(i, j)$  à posição  $(i, j-1)$ .

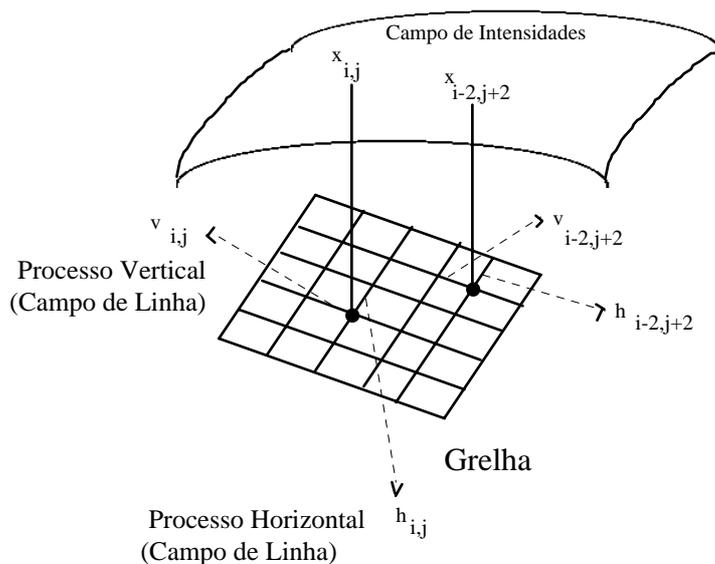


Fig. 9

Podemos então considerar três *MRF* acoplados:  $x$  representando o campo de intensidade e  $h$  e  $v$  representando, respectivamente, os processos de linha horizontal e vertical.

Deste modo, na especificação das funções de potencial em (6) tem que se ter em conta a interacção entre os campos de intensidade e de linha. Sendo assim, define-se

$$V_c(x_s) = VI(x_s) + VL(x_s) \quad , \quad (8)$$

onde  $VI(x_s)$  e  $VL(x_s)$  designam as funções de potencial associadas aos campos de intensidade e de linha, respectivamente.

### 3.3.1 Especificação da Função de Potencial Associada ao Campo de Intensidade

O termo  $VI(x_s)$  em (8) refere-se à função de potencial associada ao campo de intensidade, a qual, para um sistema de vizinhança de 1ª ordem, toma a forma

$$VI(x_s) = \beta_1 f(x_{i,j} - x_{i-1,j}) + \beta_2 f(x_{i,j} - x_{i,j-1}) + \beta_3 f(x_{i,j} - x_{i+1,j}) + \beta_4 f(x_{i,j} - x_{i,j+1}) \quad (9)$$

onde os parâmetros  $\beta_1, \beta_2, \beta_3, \beta_4$  constituem pesos relativos que ponderam a influência de pixels pertencentes à vizinhança  $\eta^1$ . Tendo por objectivo estudar

a influência de modelos estatísticos diversos, em termos da eficiência de detecção de linhas em imagens de fotografia aérea, consideraram-se em (9) as seguintes funções:

• *Quadrática*:  $f(\cdot) = (\cdot)^2$

e

• *Recíproca*:  $f(\cdot) = -\frac{1}{1+|\cdot|}$

Estas funções correspondem a assumir distribuições de probabilidade *a priori* diferentes. A função quadrática (modelo gaussiano) provoca um alisamento entre regiões de grande contraste enquanto que a função recíproca [9] realça as discontinuidades entre essas duas regiões. Sendo assim, é de esperar que o modelo baseado na função recíproca conduza a processos de detecção de contornos mais eficientes.

### **3.3.2 Especificação da Função de Potencial Associada ao Campo de Linha**

O termo  $VL(x_s)$  em (8), refere-se ao campo de linha. Neste trabalho foi dedicada especial atenção à especificação de  $VL(x_s)$  no sentido de modelizar correctamente certo tipo de características geométricas das linhas que se pretendem detectar, mantendo, no entanto, a complexidade do modelo em níveis aceitáveis. Atendendo a estas considerações, a função  $VL(x_s)$  que modeliza o campo de linha é dada por:

$$\begin{aligned}
VL(x_s) = & \alpha \left( \Delta_{i,j}^h (1 - h_{i,j}) + \Delta_{i,j}^v (1 - v_{i,j}) \right) + & (10) \\
& \gamma_1 h_{i,j} H_{i,j} + \gamma_2 v_{i,j} V_{i,j} + \\
& \xi_1 HV1_{i,j} + \xi_2 HV2_{i,j} + \\
& \xi_3 HV3_{i,j} + \xi_4 HV4_{i,j} + \\
& \zeta_1 PH_{i,j} + \zeta_2 PV_{i,j}
\end{aligned}$$

O primeiro termo, pesado pelo parâmetro de alisamento  $\alpha$ , favorece a detecção de descontinuidades do campo de intensidade. Os parâmetros  $h_{i,j}$  e  $v_{i,j}$  referem-se aos valores a estimar do processo de linha horizontal e vertical, podendo tomar os valores 1 ou 0. As funções quadráticas representadas por  $\Delta_{i,j}^h$  e  $\Delta_{i,j}^v$  são valores do gradiente de intensidade em cada direcção:

$$\begin{aligned}
\Delta_{i,j}^h &= (x_{i,j} - x_{i-1,j})^2 \\
\Delta_{i,j}^v &= (x_{i,j} - x_{i,j-1})^2
\end{aligned}$$

Estas funções dão uma medida das amplitudes das descontinuidades horizontais e verticais (ver figura 10). Note-se que o favorecimento da detecção de descontinuidades ( $h_{i,j} = 1$  e/ou  $v_{i,j} = 1$ ) se evidencia pelo facto de, nestes casos, o 1º termo de (10) se anular, contribuindo para a minimização da energia.

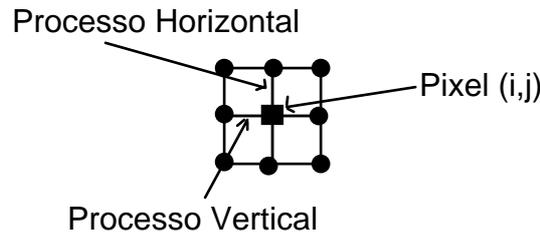


Fig. 10

Os restantes termos de (10) são termos de penalização. Recorde-se que o objectivo principal consiste na detecção das linhas que delimitam estradas, pelo que estes termos de penalização actuam sobre características geométricas

atípicas neste tipo de estruturas. As características geométricas a penalizar em ambos os processos de linha horizontal e vertical são as seguintes: (i) descontinuidades das linhas, (ii) cantos, e (iii) linhas paralelas muito próximas (1 a 2 pixels). No primeiro caso, a razão é óbvia, uma vez que as descontinuidades são em geral consequência de artefactos como, por exemplo, sombras provocadas por árvores, postes, etc. A penalização da formação de cantos deve-se ao facto de em geral, as estradas apresentarem raios de curvatura elevados. Embora possam existir em cruzamentos ou pela presença de viadutos, os cantos são formas geométricas que, neste contexto, têm baixas probabilidades de ocorrência. Finalmente, com a penalização de linhas paralelas muito próximas pretende-se obter uma maior definição dos contornos.

### 3.3.2.1 Penalização da Formação de Descontinuidades de Linhas

Os factores

$$H_{i,j} = |h_{i-1,j} - h_{i+1,j}|$$

$$V_{i,j} = |v_{i,j-1} - v_{i,j+1}|$$

penalizam descontinuidades das linhas horizontais e verticais.

Para melhor se perceber a influência destes factores no valor da função de energia consideremos o esquema representado na figura 11 referente à grelha dual do campo de linha.

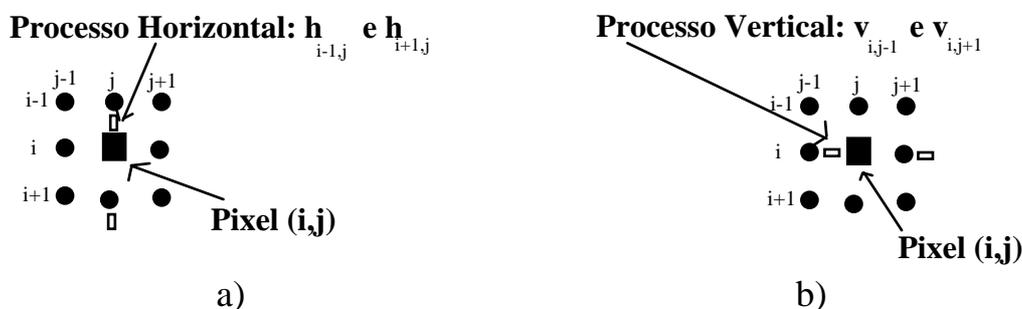


Fig. 11

Processos de Linha: a)Horizontal e b) Vertical

Por exemplo, para  $H_{i,j}$ , caso haja continuidade para "trás" e para a "frente" ( $h_{i-1,j} = h_{i+1,j} = 1$ ),  $H_{i,j}$  toma o valor 0, e a tendência será no sentido de  $h_{i,j}$  tomar o valor 1. Caso contrário (um dos  $h$  é zero),  $H_{i,j}$  toma o valor 1, e o valor de  $h_{i,j}$  fica condicionado por  $\Delta_{i,j}^h$  e pelo peso relativo dos parâmetros  $\alpha$  e  $\gamma_1$ . Considerações análogas podem ser feitas relativamente a  $V_{i,j}$ . Na figura 12, representam-se as configurações horizontais e verticais a penalizar e a favorecer.

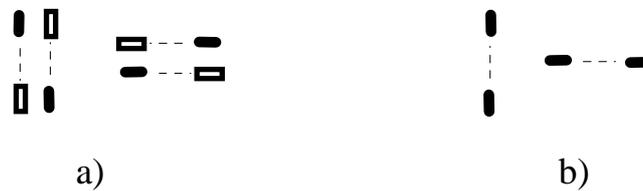


Fig. 12  
Configurações: a) A Penalizar e b) A Favorecer

### 3.3.2.2 Penalização da Formação de Cantos

Os factores  $HV1_{i,j}, K, HV4_{i,j}$  influenciam de um modo semelhante o valor da função de energia, penalizando o aparecimento de estruturas circulares (cantos), sendo definidos por:

$$HV1_{i,j} = h_{i-1,j} \cdot v_{i,j}$$

$$HV2_{i,j} = h_{i,j} \cdot v_{i,j}$$

$$HV3_{i,j} = h_{i,j} \cdot v_{i,j-1}$$

$$HV4_{i,j} = h_{i-1,j} \cdot v_{i,j-1} \cdot$$

É de notar que estes termos só aparecem caso haja evidência da formação de cantos. Para se compreender o modo como estes factores influenciam o valor da função de energia, tomemos como exemplo o factor  $HV1_{i,j}$  e considere-se o esquema representado na figura 13-a).

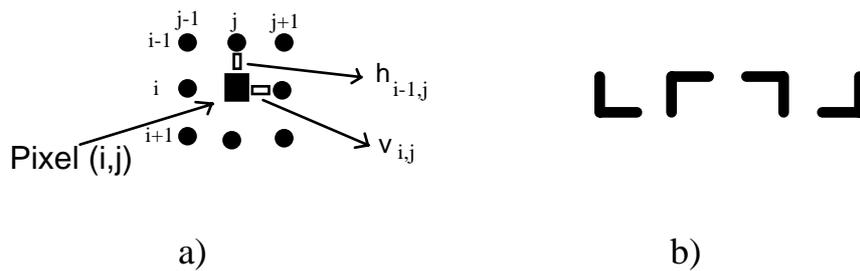


Fig. 13

- a) Esquema na grelha dual de um canto
- b) Configurações dos vários cantos a penalizar

Caso  $h_{i-1,j}$  seja zero, o termo  $HV1_{i,j}$  anula-se não influenciando consequentemente o valor da função de energia. Caso contrário, se  $h_{i-1,j}$  for 1, o valor a estimar para  $v_{i,j}$  dependerá do balanço entre os vários termos onde  $v_{i,j}$  apareça bem como dos pesos relativos dos parâmetros com ele relacionados. Assim, caso nesta posição,  $(i-1, j)$ , o valor obtido para o processo horizontal tenha sido 1, haverá tendência a considerar para  $v_{i,j}$  o valor 0 de modo a baixar o valor da função de energia e deste modo penalizar-se uma eventual formação de um canto. Caso o valor obtido para o processo horizontal em  $(i-1, j)$  tenha sido 0,  $HV1_{i,j}$  anula-se ficando então a estimativa para o valor de  $v_{i,j}$  dependente somente dos outros factores. Considerações análogas podem ser feitas em relação aos restantes factores  $HV2_{i,j}, HV3_{i,j}, HV4_{i,j}$  de penalização da formação de cantos, conseguindo-se deste modo inibir uma sua fácil formação. Em 13-b) representam-se as configurações de cantos a penalizar.

### 3.3.2.3 Penalização da Formação de Linhas Paralelas Contíguas

Por sua vez os factores

$$PH_{i,j} = h_{i,j-1} \cdot h_{i,j} \cdot h_{i,j+1}$$

$$PV_{i,j} = v_{i-1,j} \cdot v_{i,j} \cdot v_{i+1,j}$$

têm uma influência similar no valor da função de energia, penalizando o aparecimento de linhas contiguamente paralelas (horizontais e verticais), e também aparecem na expressão da função de energia apenas quando há evidência da formação destas linhas.

Consideremos o esquema representado na figura 14-a) referente à grelha dual do processo de linha horizontal.

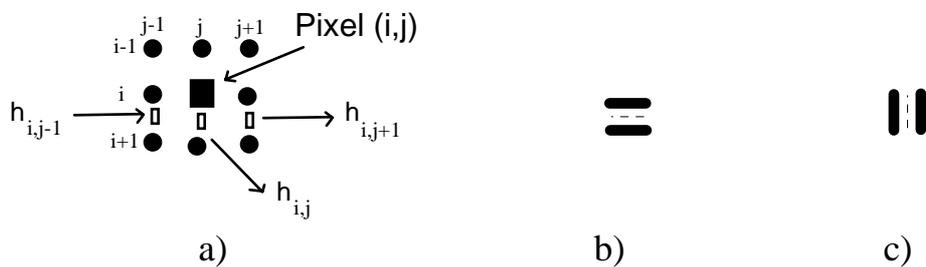


Fig. 14

- a) Esquema na grelha dual de uma linha paralela contígua
- b) e c) Configurações das linhas paralelas a penalizar

Repare-se que caso os valores do campo de linha relativos ao processo horizontal obtidos para os pixels vizinhos de  $(i, j)$  sejam simultaneamente 1, isto é, caso haja processo de linha horizontal à direita e à esquerda, o que corresponderia a formar linhas verticalmente paralelas, a tendência será no sentido de  $h_{i,j}$  tomar o valor 0, dependendo do balanço entre os vários termos dependentes de  $h_{i,j}$ . Deste modo, pode-se penalizar-se o aparecimento destas linhas paralelas. Considerações análogas podem ser feitas relativamente a  $PV_{i,j}$ . Na figura 14-b) e c) representam-se as configurações a penalizar.

### 3.3.2.4 Parâmetros do Modelo

Em (9), os parâmetros  $\beta_1, \beta_2, \beta_3, \beta_4$  constituem pesos relativos que ponderam a influência de pixels pertencentes à vizinhança  $\eta^l$ . Por sua vez em (10)  $\alpha$  é um parâmetro de alisamento,  $\gamma_1$  e  $\gamma_2$  são factores de penalização da formação de linhas horizontais e verticais, e os parâmetros  $\xi_1, \xi_2, \xi_3, \xi_4$  e  $\zeta_1, \zeta_2$  são factores de ponderação que devem ser escolhidos no sentido de penalizar mais fortemente a detecção de cantos e paralelas.

Normalmente os valores destes parâmetros são escolhidos arbitrariamente e ajustados experimentalmente.

## 3.4 Modelo dos Dados

Em geral, a imagem observada é do tipo

$$Y = \Phi(X(I, L)) \Theta R$$

em que  $I$  representa o campo de intensidades e  $L$  o campo de linha.  $\Phi$  é uma transformação sobre a imagem original que traduz as perturbações introduzidas na imagem original devidas a problemas, por exemplo, de focagem das lentes, do filme utilizado, de calibração dos sensores ou mesmo devido a movimentos da Terra ou dos captores aquando da tomada das fotografias.  $\Theta$  é um operador, multiplicativo ou aditivo, sobre o campo de ruído  $R$ .

Neste trabalho, pelo facto de não conhecermos as perturbações atrás referidas de modo a poder modelizá-las correctamente, consideramos apenas o caso particular em que  $\Phi$  é uma identidade e  $\Theta$  um operador de soma. Por

outro lado, uma eventual consideração destas perturbações iria também aumentar a complexidade do modelo o que poderia traduzir-se por um aumento significativo da carga computacional envolvida. Assim, preferiu-se investir na complexidade do modelo *a priori* em detrimento de uma especificação mais detalhada do modelo de observações. A qualidade dos resultados obtidos, como se verá no capítulo 4, dá justificação a esta abordagem.

Assim, o modelo dos dados considerado é traduzido pela expressão

$$Y = X + R,$$

onde  $R$ , estatisticamente independente de  $X$ , é suposto ser um campo de ruído branco gaussiano, com realizações independentes em cada pixel, de média nula e variância  $\sigma^2$ .

### 3.5 Critério de Optimização

Nestas condições, pelo teorema de Bayes e usando as distribuições de  $X$  e do ruído  $R$ , a probabilidade de  $X$  condicionada nos dados  $Y$  vem

$$P_{X|Y}(x|y) = \frac{P_{Y|X}(y|x)P_X(x)}{P_Y(y)}. \quad (11)$$

Como  $P_{Y|X}(y|x) = P_R(y-x)$  e

$$P_R(r) = \prod_s \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{r_s^2}{2\sigma^2}} = c_0 e^{-\sum_s \frac{r_s^2}{2\sigma^2}}$$

resulta

$$P_{Y|X}(y|x) = \prod_s \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(y_s-x_s)^2}{2\sigma^2}} = c_0 e^{-\sum_s \frac{(y_s-x_s)^2}{2\sigma^2}}. \quad (12)$$

De (5), (11) e (12) obtém-se

$$P_{X|Y}(x | y) = \frac{e^{-\frac{1}{T} \sum_c V_c(x)} \cdot c_0 e^{-\frac{1}{2\sigma^2} \sum_c (y_s - x_s)^2}}{Z} \cdot P(y) \quad , \quad (8)$$

onde  $V_c(x)$  é a função de potencial associada à clique  $c$  e especificada na secção 3.3.

De acordo com o critério de *maximum a posteriori*, a expressão (13) traduz a funcional de custo a maximizar. Note-se que maximizar esta funcional é equivalente a minimizar a função de energia

$$U(x) = \frac{1}{T} \sum_c V_c(x) + \frac{1}{2\sigma^2} \sum_c (y_s - x_s)^2 \quad (14)$$

Neste ponto, o problema a resolver formula-se do seguinte modo: dada a imagem  $Y = \{y_{ij}\}$  calcular o estado  $X = \{x_{ij}\}$  que minimiza a energia (14). Este problema será abordado no capítulo 4.

### 3.6 Sumário

Abordou-se neste capítulo o problema da aplicação do modelo de Markov/Gibbs à detecção de estruturas lineares em imagens de fotografia aérea. Esta detecção corresponde a um primeiro nível do problema mais amplo da extracção de redes viárias ou hidrográficas em imagens de fotografia aérea. Este nível de processamento envolve procedimentos de baixo nível tais como segmentação da imagem original e construção das linhas que permitam delimitar as estruturas passíveis de serem classificadas posteriormente, por exemplo, como estradas.

Tendo-se estabelecido este objectivo da detecção de estruturas lineares, assumiu-se que o modelo *a priori* era determinado por um campo aleatório de Markov. Dado ser extremamente difícil a especificação de forma directa do modelo de Markov, considerou-se a equivalência entre a descrição

markoviana e o modelo de Gibbs. Este modelo constitui uma representação global da imagem e apresenta a vantagem de poder ser especificado de forma mais simples. O modelo de Markov/Gibbs considerado, e cuja estrutura se detalhou, representa simultaneamente o campo de intensidades da imagem e o seu dual designado por campo de linhas a cuja especificação foi dedicada especial atenção. A concretização do campo de linhas, subjacente às descontinuidades presentes no campo de intensidades foi condicionada especificando-se no modelo termos que permitem favorecer continuidades de linhas e termos que permitem penalizar o aparecimento de determinadas estruturas tais como cantos e linhas contiguamente paralelas. Torna-se assim possível, mesmo neste primeiro nível de processamento, controlar a verosimilhança de que as estruturas lineares detectadas correspondam, por exemplo, a estradas.

Especificou-se também um modelo simplificado dos dados em que, sobreposto à imagem original, se considera haver um termo de ruído branco aditivo.

A partir deste modelo dos dados e aplicando a teoria bayesiana, especificou-se, com base no modelo de Markov/Gibbs, uma função de distribuição de probabilidade *a posteriori* a partir da qual se obteve uma funcional de custo (energia) cuja minimização conduz a uma estimativa de *maximum a posteriori*.

No capítulo seguinte, recorre-se a um algoritmo de optimização já conhecido para proceder ao cálculo da estimativa do campo de linha em imagens de fotografia aérea, usando o modelo *a priori* especificado neste capítulo. Recorde-se que este modelo formaliza restrições adequadas ao tipo de linhas (neste caso, contornos de estradas) que se pretende detectar.

## **4. APLICAÇÃO A IMAGENS DE FOTOGRAFIA AÉREA**

### **4.1 Introdução**

Neste capítulo iremos abordar a aplicação do modelo descrito na secção anterior a imagens de fotografia aérea digitalizadas.

A funcional de custo (energia) que foi construída a partir do modelo integrado do campo de intensidades e de linha deve ser otimizada de modo a poder reconstruir-se o campo de linhas (tendo em atenção as restrições impostas pelo modelo) com base na segmentação óptima do campo de intensidades. Nesta optimização, e com a finalidade de se testar desempenhos, foram seguidas duas estratégias de optimização do campo de intensidades baseadas em modelos probabilísticos diferentes: uma baseada numa função designada por quadrática e a outra baseada numa função designada por recíproca. A optimização do campo de intensidades e de linhas pode ser feita segundo uma abordagem sequencial, optimização factorizada, ou segundo uma abordagem integrada, optimização conjunta, tendo-se também testado estas duas alternativas.

Será sumariamente referido o algoritmo de optimização utilizado e as razões que levaram à sua utilização. Os vários passos do algoritmo são descritos no Apêndice A. De seguida mostram-se resultados da aplicação em imagens de fotografia aérea digitalizadas e analisam-se esses resultados.

## 4.2 Algoritmo de Optimizaçã

A optimizaçã da funçã de energia foi feita com base no algoritmo *Highest Confidence First (HCF)*, descrito em [3]) o qual, quando comparado com outros algoritmos, mostra ser computacionalmente mais eficiente. Para alã disso, verifica as seguintes propriedades:

- É um algoritmo determinístico que efectua cálculos locais;
- Com base num esquema global de minimizaçã da energia toma decisões sobre qual o pixel a visitar em cada iteraçã;
- A ordem de visita é determinada por forma a que se garanta um decréscimo máxmo de energia, acelerando-se de modo controlado o processo de optimizaçã;
- Este controlo baseia-se numa medida de estabilidade dos estados condicionada pela actualizaçã de um determinado pixel;
- A configuraçã final obtida com o *HCF* corresponde, em geral, a um mímimo local de energia.

O algoritmo iterativo termina quando todas as estabilidades associadas aos pixels forem positivas, isto é, quando não for mais possível baixar a energia.

O algoritmo aqui considerado goza ainda das seguintes propriedades genéricas:

- Eficiêcia: Um método iterativo não deve ser apenas determinístico mas deve tentar otimizar cada passo, de modo a garantir maior rapidez, sem prejudicar a qualidade da estimativa;

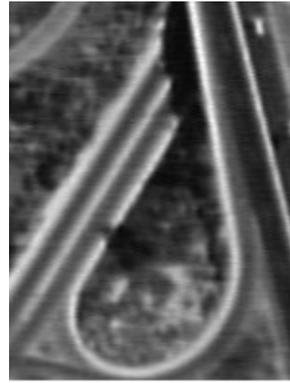
- Robustez: As estimativas não deverão ser afectadas pelas características de larga escala dos campos de Markov escolhidos.

### 4.3 Resultados da Aplicação a Imagens de Fotografia Aérea

A optimização dos campos de intensidade e de linha pode ser feita em conjunto ou em separado, fazendo-se neste último caso primeiro a optimização do campo de intensidades e de seguida a do campo de linha. Dado o objectivo do trabalho não ser a reconstrução da imagem, a optimização do campo de intensidade é feita no sentido de realçar os contrastes entre regiões. Considera-se a redução da imagem original a 2 ou a 4 níveis de cinzento.

Nos testes realizados, usaram-se as mesmas imagens de fotografia aérea digitalizadas em 256 níveis de cinzento já utilizadas no capítulo 2 e aqui representadas nas figuras 15-a) e b). Para os parâmetros de modelização do campo de linha consideraram-se os valores  $\sigma = 0.3$ ,  $\alpha = 0.5$ ,  $\xi_1 = \xi_2 = \xi_3 = \xi_4 = 0.5$ ,  $\gamma_1 = \gamma_2 = 0.5$ ,  $\zeta_1 = \zeta_2 = 0.5$ , e para o campo de intensidades  $\beta_1 = \beta_2 = \beta_3 = \beta_4 = 1$ .

Para o processamento destas imagens, foi utilizada uma *workstation* DEC ALPHA 3400 com processador AXP RISC de 64 Bits, 92 MFLOPS e 266 MIPS.

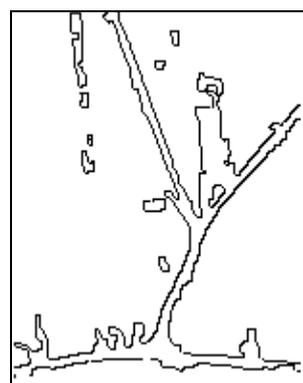
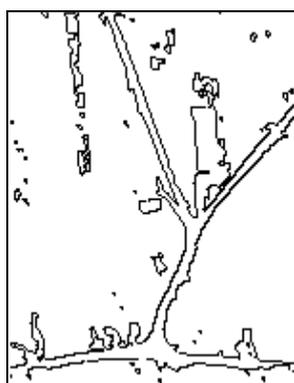


a) Imagem 1      b) Imagem 2

Fig. 15  
Imagens Originais

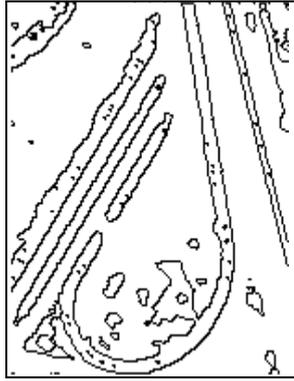
### 4.3.1 Abordagem Sequencial

Nas figuras 16 a 19, representam-se os campos de linha detectados após processamento factorizado dos campos de intensidade e de linhas. Nestas figuras, a) refere-se ao processamento das imagens utilizando-se a função recíproca no campo de intensidade enquanto que em b) se representam as imagens resultantes desse processamento com base na função quadrática. As imagens representadas nas figuras (16 e 17) e (18 e 19) resultam da optimização factorizada dos campos de intensidade e de linha com redução a 2 e a 4 níveis, respectivamente.

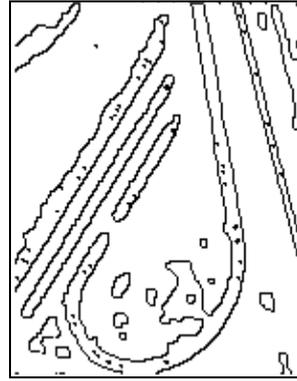


a) Recíproca      b) Quadrática

Fig. 16  
Imagem 1 Pré-Segmentada a 2 Níveis



a) Recíproca



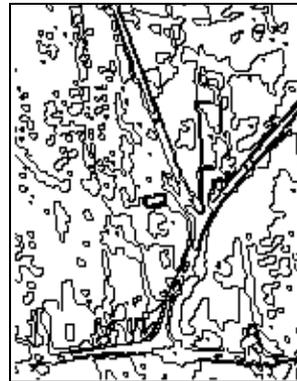
b) Quadrática

Fig. 17

Imagem 2 Pré-Segmentada a 2 Níveis



a) Recíproca



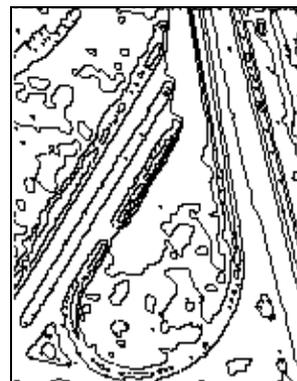
b) Quadrática

Fig. 18

Imagem 1 Pré-Segmentada a 4 Níveis



a) Recíproca



b) Quadrática

Fig. 19

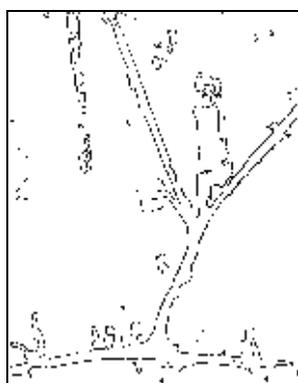
Imagem 2 Pré-Segmentada a 4 Níveis

Estes resultados permitem constatar que:

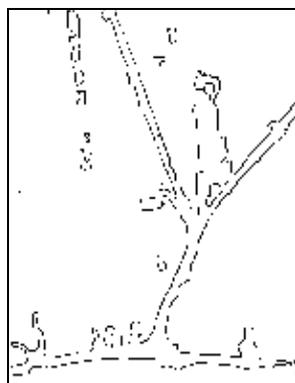
- A utilização da função recíproca no campo de intensidades, como se observa em a) de qualquer destas figuras, e como seria de esperar da discussão apresentada no capítulo 3, conduz a uma detecção de contornos mais eficiente do que se se utilizar a função quadrática;
- Com a redução a 4 níveis, muito mais detalhe é representado se bem que o tempo de processamento (ver tabela 1) aumente ligeiramente.

### 4.3.2 Abordagem Integrada

As figuras 20 a 23 mostram o resultado do processamento com optimização conjunta dos campos de intensidades e de linha. A disposição das várias figuras coincide com a atrás descrita.



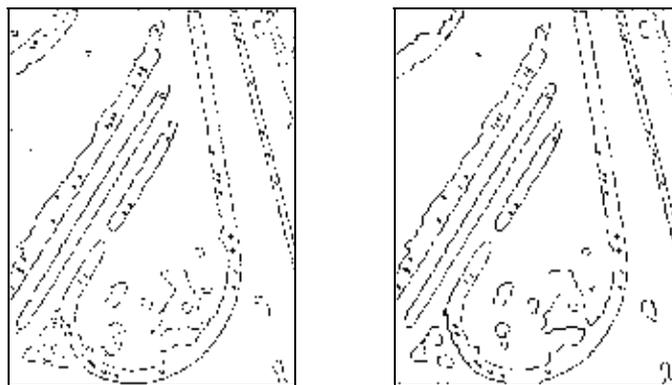
a) Recíproca



b) Quadrática

Fig. 20

Imagem 1 - Optimização Conjunta com Redução a 2 Níveis

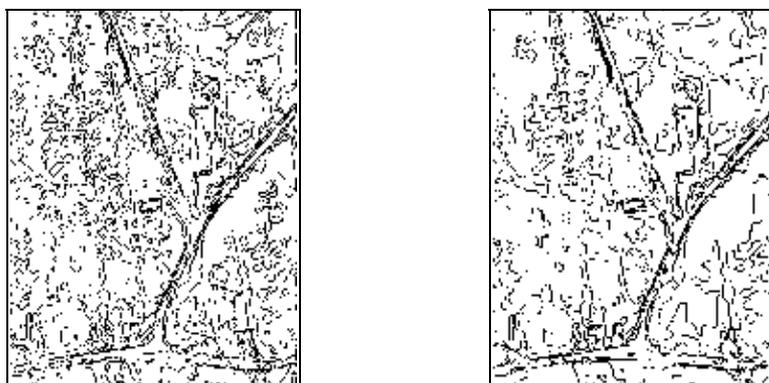


a) Recíproca

b) Quadrática

Fig. 21

Imagem 2 - Otimização Conjunta com Redução a 2 Níveis

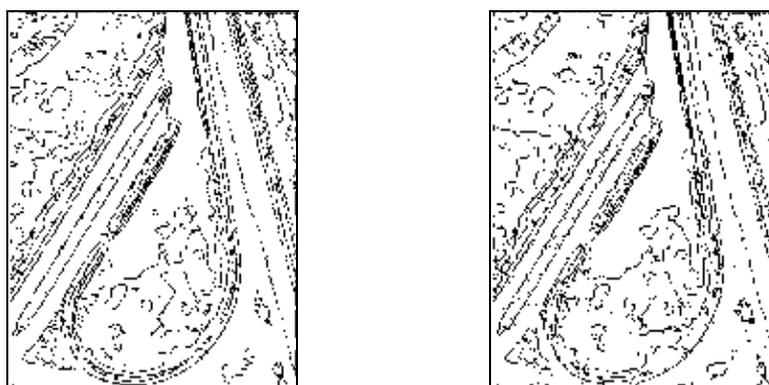


a) Recíproca

b) Quadrática

Fig. 22

Imagem 1 - Otimização Conjunta com Redução a 4 Níveis



a) Recíproca

b) Quadrática

Fig. 23

Imagem 2 - Otimização Conjunta com Redução a 4 Níveis

- Como se constata, comparando com as figuras anteriores, se bem que os tempos de processamento sejam menores, os resultados obtidos com esta optimização conjunta são piores do que os obtidos com a optimização factorizada. Isto deve-se ao facto de o mínimo local obtido com esta abordagem integrada ser pior do que o obtido usando a abordagem sequencial.

- A utilização desta abordagem integrada conduz a tempos de processamento menores do que os obtidos com a optimização factorizada. Isto era previsível se atendermos ao facto de que a optimização factorizada corresponde a dois passos de optimização: primeiro optimiza-se o campo de intensidades e de seguida optimiza-se o campo de linhas.

	a)	b)
Fig. 16	1:46:57	1:38:15
Fig. 17	1:57:46	1:47:09
Fig. 18	2:08:08	1:58:52
Fig. 19	2:07:12	1:58:20
Fig. 20	1:26:51	1:22:11
Fig. 21	1:52:11	1:45:07
Fig. 22	1:37:05	1:31:01
Fig. 23	1:53:14	1:51:36

**Tabela 1**

#### 4.4 Análise de Resultados

Da análise destas figuras, pode concluir-se que relativamente ao modelo ensaiado:

i) Verifica-se que com o modelo recíproco se obtêm melhores resultados do ponto de vista da continuidade das linhas detectadas;

ii) A utilização da optimização conjunta dos campos de intensidade e de linha, apesar de ter tempos de processamento menores, não conduz a resultados tão bons como os obtidos com a optimização factorizada;

iii) Como se pode constatar das figuras 16 a 19, o contraste da imagem original é importante para os resultados obtidos. Para a imagem 1, que é bastante mais contrastada do que a 2, as linhas mais relevantes são facilmente detectadas quando a imagem é reduzida a 2 níveis enquanto que para a imagem 2 há linhas relevantes que só são detectadas se a imagem estiver reduzida a 4 níveis.

Assim, para imagens bem contrastadas e considerando o objectivo deste trabalho, a optimização do campo de intensidades com redução da imagem a 2 níveis de cinzento, para além de ser mais rápida, permite obter melhores resultados, comparativamente à optimização com redução da imagem a 4 níveis. Note-se que neste caso, embora se consiga um acréscimo no detalhe com que as estruturas de interesse são detectadas, surgem muitos outros contornos de objectos irrelevantes do ponto de vista do problema em consideração. Isto significa que a abordagem aqui considerada, embora (e como se verá de seguida) apresente desempenhos muito superiores aos conseguidos com as técnicas clássicas, não dispensa um nível superior de processamento. Este nível de interpretação poderá, por seu lado, recorrer à informação contida nos campos de linha, obtidos por redução dos campos de intensidade a 2 e 4 níveis, para resolver ambiguidades e eliminar objectos que não sejam de interesse.

#### 4.5 Comparação com as Abordagens Clássicas

Comparativamente aos resultados obtidos com os da implementação feita no capítulo 2, os resultados agora obtidos são bastante melhores. As figuras 24 e 25 representam os resultados obtidos, respectivamente, para as imagens originais 1 e 2, usando o método clássico implementado no capítulo 2 (figuras 24 e 25 b)) e com o método bayesiano descrito no capítulo 3 (figuras 24 e 25 c)). Neste caso, estes resultados referem-se à otimização factorizada com utilização da função recíproca no campo de intensidades e redução a 2 níveis de cinzento.

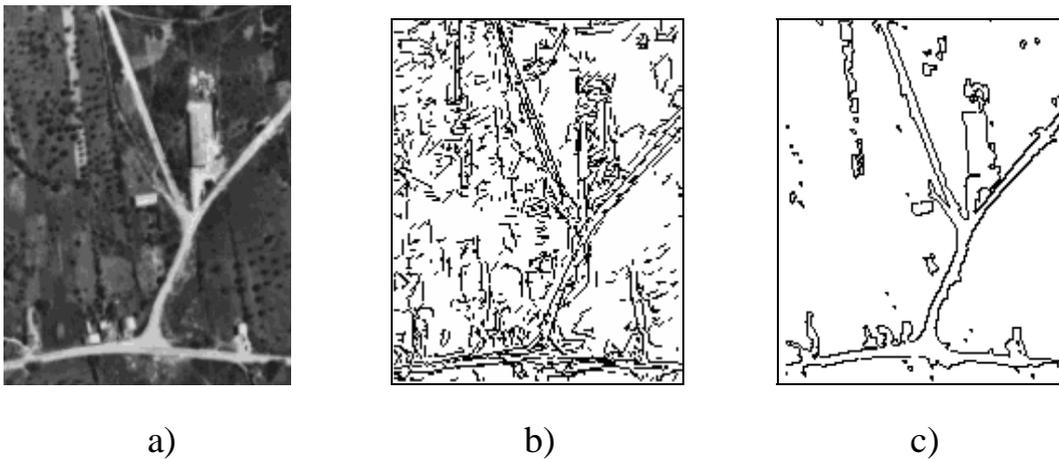


Fig. 24

Comparação de Resultados

a) Imagem Original 1

b) Abordagem Clássica

c) Abordagem Bayesiana

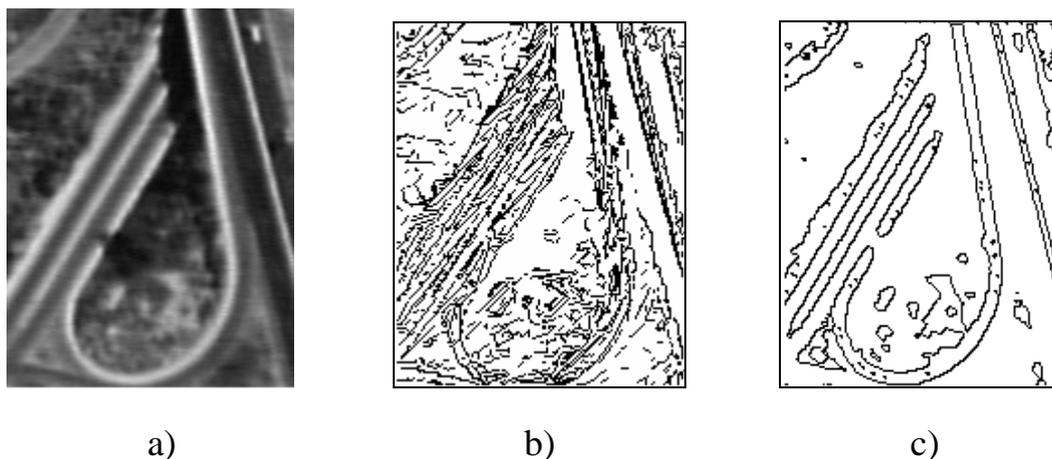


Fig. 25

Comparação de Resultados Obtidos

a) Imagem Original 2    b) Abordagem Clássica    c) Abordagem Bayesiana

Da análise destas figuras observa-se que os resultados obtidos com o emprego das técnicas bayesianas desenvolvidas no capítulo 3 são bastante superiores aos obtidos com a utilização do método clássico descrito na capítulo 2. Verifica-se que a abordagem bayesiana, suportada no modelo *a priori* especificado na tese, é mais eficiente na medida em que, para além de conduzir a um melhor definição dos contornos das estruturas de interesse, é responsável por um número muito menor de falsas detecções. Sublinhe-se que estes factos resultam da adequação do modelo do campo de linha desenvolvido na tese, o qual formaliza as restrições associadas ao tipo de estruturas lineares que se pretende detectar.

Pode ainda concluir-se que, para os casos em que o modelo escolhido é o mais adequado, o peso computacional envolvido é compensado pelo desempenho superior do método considerado.

#### 4.6 Sumário

O modelo descrito no capítulo 3 foi aqui aplicado a imagens de fotografia aérea digitalizadas com o objectivo de se detectarem contornos de estradas. Esta abordagem combina a imagem digitalizada (dados) com

conhecimento *a priori* representado por um Campo Aleatório de Markov (*Markov Random Field, MRF*).

Para a optimização da função de energia, utilizou-se o algoritmo *Highest Confidence First (HCF)* [3] o qual permite obter um mínimo local para a função de energia.

O estudo feito neste capítulo teve como objectivo fundamental avaliar o desempenho da abordagem bayesiana para a detecção dos contornos de estradas em imagens de fotografia aérea. Consideraram-se dois tipos de modelos *a priori* do campo de intensidades (aqui designados por modelos quadrático e recíproco) e duas abordagens possíveis para efectuar a optimização dos campos de intensidade e de linha. Os resultados obtidos foram ainda comparados com aqueles decorrentes da aplicação do algoritmo descrito no capítulo 2 e baseado em métodos não bayesianos.

Em síntese, o estudo feito pode ser esquematizado do seguinte modo:

1. Desempenho do método bayesiano baseado em modelos de Markov/Gibbs:
  - 1.a- Estudo de sensibilidade relativamente à distribuição de probabilidade *a priori* do campo de intensidades;
  - 1.b- Estudo comparativo de duas implementações possíveis de processamento:
    - abordagem sequencial (segmentação óptima seguida da estimação do campo de linha);
    - abordagem integrada (optimização conjunta);

2. Comparação de desempenhos entre técnicas bayesianas baseadas em modelos de Markov/Gibbs e técnicas clássicas baseadas em modelos heurísticos.

As conclusões retiradas deste estudo foram as seguintes:

- Os resultados obtidos com a utilização da função recíproca no campo de intensidade foram melhores do que os obtidos com a utilização da função quadrática;
- Fazendo-se a otimização factorizada obtêm-se resultados bastante melhores do que os obtidos com a otimização conjunta;
- O modelo estruturado utilizado baseado no modelo de Markov/Gibbs revelou-se muito melhor adequado ao processamento destas imagens do que o modelo não estruturado baseado em heurísticas usado no método clássico.

Como conclusão genérica final pode afirmar-se que os resultados obtidos com a utilização de técnicas bayesianas baseadas no modelo de Markov/Gibbs, embora envolvendo custos computacionais consideráveis, apresentam uma qualidade muito superior aos obtidos com algoritmos de processamento baseados em abordagens clássicas do tipo heurístico.

## 5. CONCLUSÕES

### 5.1 Introdução

No presente trabalho abordou-se o problema da detecção de estruturas lineares em imagens de fotografia aérea. A abordagem proposta é do tipo bayesiano, baseada em modelos de Markov/Gibbs e constitui o primeiro nível de processamento associado ao problema mais lato da extracção automática de redes viárias.

Definiu-se, como objectivo central da tese, o estabelecimento de um modelo aleatório de Markov/Gibbs que incorpora de modo adequado o conhecimento *a priori* sobre a imagem a processar. A validade do modelo foi verificada experimentalmente, evidenciando-se o seu bom desempenho na detecção de contornos de estradas.

### 5.2 Trabalho Desenvolvido

Foi feita uma análise de alguns métodos clássicos para a identificação de estradas em imagens digitalizadas tendo-se também procedido à implementação de um algoritmo baseado numa destas técnicas de tipo heurístico.

De seguida abordou-se o problema em questão à luz da teoria bayesiana. Após uma descrição genérica do problema da modelização estocástica, detalhou-se a construção do modelo *a priori* incluído na classe de modelos de Markov/Gibbs. A partir de um modelo simplificado dos dados determinou-se a função de distribuição *a posteriori* a partir da qual se obteve uma funcional

de custo (energia) cuja minimização permitiu obter uma solução de *maximum a posteriori*.

Como já foi referido, foi dedicada especial atenção à especificação do modelo *a priori*. Dada a dificuldade em especificar de forma directa o modelo de Markov, considerou-se a equivalência entre este modelo e o modelo de Gibbs. O modelo de Markov/Gibbs considerado representa simultaneamente o campo de intensidades e o seu dual designado por campo de linhas. A modelização do campo de linhas, subjacente às descontinuidades presentes no campo de intensidades, permite condicionar o tipo de estruturas a detectar pela definição de termos que favorecem a continuidades de linhas e de termos que penalizam o aparecimento de determinadas estruturas geométricas tais como cantos e linhas contiguamente paralelas. Tornou-se assim possível, mesmo neste primeiro nível de processamento, controlar a verosimilhança de que as estruturas lineares detectadas correspondam, de facto, a estradas. Definiu-se também um modelo simplificado dos dados em que, sobreposto à imagem original, se considera existir um termo de ruído branco aditivo. A partir deste modelo dos dados, e aplicando a teoria bayesiana, especificou-se com base no modelo de Markov/Gibbs uma função de distribuição de probabilidade *a posteriori* a partir da qual se obteve uma funcional de custo (energia) cuja minimização conduz a uma estimativa de *maximum a posteriori*.

Para otimizar a função de energia assim obtida, usou-se o algoritmo *HCF* cuja descrição se detalha no Apêndice A.

Com o objectivo de avaliar a adequação do modelo e o desempenho da metodologia bayesiana proposta, procedeu-se ao conjunto de testes que seguidamente se detalham.

- 1- Na modelização do campo de intensidades foram ensaiados dois tipos diferentes de medida local de descontinuidades, um designado por

função quadrática e correspondendo a um modelo gaussiano e outro designado por função recíproca;

2- Outra comparação de desempenho testada prende-se com o modo como a optimização é efectuada tendo-se considerado:

- Optimização conjunta do campo de intensidades e do campo de linhas;
- Optimização factorizada do campo de intensidades e do campo de linha;

Recorde-se que, no âmbito deste trabalho, por optimização do campo de intensidade se entende segmentação óptima por redução da imagem original a 2 e a 4 níveis de cinzento.

3- Comparação dos resultados decorrentes do método proposto e do algoritmo heurístico implementado.

### **5.3 Comparação de Resultados Obtidos**

Como se constatou pelos resultados obtidos no capítulo 2, os métodos clássicos de detecção de estruturas lineares conduzem geralmente a um número elevado de falsas detecções, particularmente em imagens de elevada resolução como são as de fotografia aérea. Outra dificuldade prende-se com o estabelecimento dos modelos genéricos em que se baseiam as heurísticas de processamento. No exemplo de aplicação implementado no capítulo 2, tornou-se evidente a dificuldade do método considerado em lidar com a grande quantidade de informação presente nas fotografias aéreas de grande escala. Este facto traduziu-se pelo elevado número de falsas detecções presentes nos resultados. Em contrapartida, pode afirmar-se que o método ensaiado envolve baixos custos computacionais.

Os métodos baseados em modelos estatísticos do tipo Markov/Gibbs permitem estruturar o problema do processamento na medida em que integram formalmente o conhecimento prévio sobre as características das imagens a processar. Por outro lado, estes modelos não dependem da escala das imagens de fotografia aérea aqui consideradas. Com base nestas premissas, foi possível fundamentar a opção por este tipo de modelos tendo em vista a obtenção de resultados mais eficientes em comparação com os que foram obtidos usando o algoritmo heurístico descrito no capítulo 2.

Os resultados obtidos permitiram confirmar este pressuposto. Constatou-se que a solução bayesiana baseada no modelo de Markov/Gibbs conduz à detecção de estruturas lineares melhor definidas do que as obtidas com base na solução clássica. Deste modo, o peso computacional acrescido na implementação bayesiana é sobejamente compensado pela superioridade dos resultados obtidos.

Relativamente às estratégias alternativas consideradas na optimização da funcional desenvolvida com base no modelo de Markov/Gibbs, pode-se considerar que:

- 1- Os resultados obtidos com a optimização factorizada são bastante melhores do que os obtidos com a optimização conjunta, apesar de no primeiro caso ser necessário mais tempo de processamento. Isto deve-se ao facto da optimização factorizada conduzir a um mínimo local de energia melhor do que o obtido com a optimização conjunta;
- 2- Utilizando-se o modelo recíproco na optimização do campo de intensidades obtêm-se melhores resultados do que com o modelo gaussiano, embora os tempos de processamento sejam ligeiramente maiores;

Tendo em vista o objectivo proposto no trabalho, os melhores resultados, do ponto de vista da detecção e continuidade das linhas, são obtidos pela

optimização sequencial com redução do campo de intensidades a 2 níveis e usando o modelo recíproco.

Como conclusão final, pode afirmar-se que o modelo de representação conjunta dos campos de intensidade e de linha especificado nesta tese é adequado ao problema da detecção de estruturas lineares em imagens de fotografia aérea com elevada resolução. Isto resulta do facto de o modelo do campo de linha integrar restrições geométricas típicas das estruturas de interesse. Ficou igualmente demonstrado que a formalização do problema em causa segundo uma perspectiva bayesiana baseada em modelos de Markov/Gibbs conduz a resultados de qualidade superior, quando comparados com os decorrentes da aplicação de técnicas heurísticas clássicas.

## BIBLIOGRAFIA

- [1] Z. Aviad, P. D. Carnine. *Road Finding for Road-Network Extraction*, IEEE Proceedings on Computer Vision and Pattern Recognition, Ann Arbor, Michigan, 1988
- [2] R. Bajcsy, M. Tavakoli. *Computer Recognition of Roads from Satellite Pictures*, IEEE Trans. on SMC, Vol. SMC-6 n° 9, 1976
- [3] P. B. Chou, C. M. Brown. *The Theory and Practice of Bayesian Image Labeling*, International Journal of Computer Vision, 4, 1990
- [4] H. Derin. *The Use of Gibbs Distributions In Image Processing*, in *Communications and Networks, A Survey of Recent Advances*, Ed. I. Blake & H. Poor, Springer-Verlag, Nova Iorque, 1986
- [5] R. O. Duda, P. E. Hart. *Pattern classification and Scene Analysis*, Wiley , Nova Iorque, 1973
- [6] M. Fischler, J. Tenenbaum, H. Wolf, *Detection of Roads and Linear Structures in Low-Resolution Aerial Imagery Using a Multisource Knowledge Integration Techniques*, Computer Graphics and Image Processing, N° 15, 1981
- [7] D. Geiger, F. Girosi. *Parallel and Deterministic Algorithms from MRF's: Surface Reconstruction*, IEEE Trans. on PAMI, Vol. 13, n° 5, Maio 1991
- [8] S. Geman, D. Geman. *Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images*, IEEE PAMI, 6, Novembro 1984
- [9] D. Geman, C. Reynolds. *Constrained Restoration and the Recovery of Discontinuities*, IEEE Trans. on PAMI, Vol 14, N° 3, Março 1992

- [10] A. Huertas, W. Cole, R. Nevatia. *Detecting Runways in Aerial Images*, Readings in Computer Vision, Morgan Kaufmann Publishers, California
- [11] S. Lakshmanan, H. Derin. *Simultaneous Parameter Estimation and Segmentation of Gibbs Random Fields Using Simulated Annealing*, IEEE Trans. on PAMI, Vol. 11, n° 8, Agosto 1989
- [12] D. McKeown, J. Denlinger. *Cooperative Methods for Road Tracking in Aerial Imagery*, IEEE Proceedings on Computer Vision and Pattern Recognition, Ann Arbor, Michigan, 1988
- [13] R. Nevatia, K. R. Babu. *Linear Feature Extraction and Description*, Computer Graphics and Image Processing, 1980
- [14] M. L. Zhu, P. S. Yeh. *Automatic Road Network Detection on Aerial Photographs*, IEEE Proceedings on Computer Vision and Pattern Recognition, 1986

## APÊNDICE A - Descrição do Algoritmo de Optimização

### A.1 Estimação *HCF*

O algoritmo utilizado na optimização da função de energia foi o algoritmo *Highest Confidence First, HCF*, proposto em [3].

Antes de se fazer a descrição do algoritmo, é necessário definir alguns conceitos para melhor se compreender o seu funcionamento.

Seja  $L = \{l_1, l_2, \dots, l_r\}$  um conjunto mutuamente exclusivo e exaustivo de todos os estados possíveis de cada pixel da imagem e seja  $\Omega$  o correspondente espaço de configurações possíveis. Considerando apenas o campo de intensidades quantificado em 256 níveis de cinzento, ter-se-á  $L = \{0, 1, 2, \dots, 255\}$ . Seja  $\bar{L}$ , o conjunto aumentado de estados, uma extensão do conjunto  $L$ , e que resulta da união do conjunto  $L$  com um novo estado  $l_0$ ,  $\bar{L} = L \cup \{l_0\}$ . Este estado  $l_0$  é um estado neutro. Seja  $\bar{\Omega} = \{w = (w_1, \dots, w_r) \mid w_s \in \bar{L}, \forall_{s \in C}\}$  o espaço aumentado de todas as configurações possíveis onde  $w$  é uma dada configuração. A ideia básica do *HCF* é construir uma sequência de configurações  $w^0, w^1, \dots, w^K$  de  $\bar{\Omega}$ , a partir da configuração inicial  $w = (l_0, \dots, l_0)$  e em que a configuração final obtida,  $w^f \in \Omega$ , corresponda a uma energia  $U_0(w^f)$  que é um mínimo local relativamente a  $\Omega$ .

Em cada passo da iteração, considera-se que um pixel está *committed* se o seu estado for um dos estados que compõem o conjunto  $L$ , e que estará *uncommitted* se o seu estado for  $l_0$ . Inicialmente, todos os pixels estão no estado neutro e, à medida que a iteração prossegue, vão passar a estar *committed*, diminuindo sucessivamente o número de pixels no estado  $l_0$ . No

final da iteração, não deverá haver nenhum pixel *uncommitted* pois, como veremos, a este estado corresponde uma estabilidade negativa. Um pixel após passar a *committed* não poderá voltar a *uncommitted*, mas poderá mudar para outro estado, caso resulte a um valor de energia menor.

## A.2 Determinação da Energia Local

Como referido, o problema a resolver é o seguinte: dada a imagem  $Y = \{y_{ij}\}$ , calcular o estado  $X = \{x_{ij}\}$  que minimize a energia. Este estado é obtido por maximização da probabilidade *a posteriori* expressa em (13) ou, o que é o mesmo, por minimização da função de energia (14).

Considerando o estado  $l \in L$ , relativamente a um pixel  $s \in M$  e a uma dada realização  $w' \in \bar{\Omega}$ , pode-se rescrever (14) em termos de energia local aumentada da seguinte forma [3]

$$U_s(l) = \sum_{c \in C_s} V_c(w') + \frac{(y_s - l)^2}{2\sigma^2}. \quad (9)$$

em que  $w' \in \bar{\Omega}$  é a configuração idêntica à realização  $w$  em todos os pixels excepto se  $w'_s = l$ . Quanto à função de potencial  $V_c$ , para cada  $r \in c$ , valerá 0 se  $l = l_0$ . Esta medida permite quantificar a qualidade de uma estimativa relativamente à configuração instantânea dos seus vizinhos. Repare-se que só os pixels que já fizeram o seu *commitment* é que contribuem para esta medida, pois os pixels que estão no estado neutro não são considerados na determinação da energia uma vez que o valor das funções de potencial que lhes corresponderá será zero.

No entanto, como se consideram os pixels pertencentes à *clique* do sistema de vizinhanças, quando um pixel faz o seu *commitment*, leva em linha de conta os vizinhos que já estão *committed*. Deste modo, se todos os vizinhos de determinado pixel ainda estiverem *uncommitted*, a parcela

correspondente ao somatório das funções de potencial anular-se-á pelo que (15) reduzir-se-á a:

$$U_s(l) = \frac{(y_s - l)^2}{2\sigma^2} \quad (16)$$

ou seja, neste caso, o valor da energia é obtido por uma estimativa de máxima verosimilhança pois é obtida apenas à custa dos dados.

### A.3 Determinação das Estabilidades

Em cada passo da construção só o pixel menos estável é passível de ver o seu estado alterado ou fazer o seu *commitment* e deste modo a ordem de visita deverá ser criteriosamente escolhida afim de se garantir a qualidade da estimativa final,  $w^f$ .

A estabilidade  $G_s$  de um dado pixel  $s$ , relativamente à configuração instantânea  $w$ , é definida então como:

- $G_s(w) = \min_{k \in L, k \neq w} \Delta U_s(k, w)$ , se  $w \in L$  (17)

- $G_s(w) = - \min_{j, k \in L, k \neq j} \Delta U_s(k, j)$ , se  $w = l_0$  (18)

em que  $j \in L$  e:

$$U_s(j) = \min_{k \in L} U_s(k)$$

$$\Delta U_s(j, k) = U_s(j) - U_s(k) \wedge U_s(j) \neq U_s(k)$$

Deste modo, a medida da estabilidade constitui uma combinação da medida da evidência observável e do conhecimento *a priori* sobre o estado actual e os outros estados alternativos. Um valor negativo de  $G$  indica que se poderá obter um estado mais estável (menor energia) caso o pixel faça um outro *commitment*. Para se calcular as estabilidades iniciais, começa-se por colocar todos os estados dos pixeis no estado de *uncommitment*, isto é,

igualam-se todos os estados a  $l_0$ . De seguida, calculam-se as estabilidades de cada um dos pixels usando as relações anteriores. No início, nenhum pixel tem vizinhos activos, pelo que o cálculo da energia usa apenas a estimativa de máxima verosimilhança, já que as funções de potencial que relacionam um pixel com os pixels vizinhos são todas nulas.

Assim, no início da iteração, uma vez que nenhum pixel tem os vizinhos activos e considerando as relações (16), (17) e (18), teremos:

$$U_s(j) = \min_{l \in L} U_s(l) = \min_{l \in L} \frac{(y_s - l)^2}{2\sigma^2} = 0$$

Como  $y_s$  é o valor lido na imagem fonte, temos que  $U_s(j) = 0$ , valor obtido quando  $l = y_s$ .

A estabilidade inicial determina-se então da seguinte forma:

$$\begin{aligned} G_s(w) &= - \min_{k \in L, k \neq j} \Delta U_s(k, j) & (19) \\ &= - \min_{k \in L, k \neq j} (U_s(k) - U_s(j)) \\ &= - \min_{k \in L, k \neq j} (U_s(k) - 0) \\ &= - \min_{k \in L, k \neq j} (U_s(k)) \\ &= - \min_{k \in L, k \neq j} \left( \frac{(y_s - k)^2}{2\sigma^2} \right) \end{aligned}$$

Sendo  $k \neq j$ , o valor que minimiza o quadrado desta expressão é  $k = y_s \pm 1$  dado que  $j = y_s$  pois  $U_s(j) = 0$  quando  $l = y_s$ , vindo então:

$$G_s = - \frac{1}{2\sigma^2}$$

Deste modo, as estabilidades iniciais são todas iguais e valem  $\left(-\frac{1}{2\sigma^2}\right)$ , o que é compreensível dado que todos os pixels usam apenas a informação local disponível.

Uma vez desencadeado o processo, esta homogeneidade das estabilidades desaparece rapidamente pois o *commitment* de um pixel vai provocar alterações nas estabilidades dos pixels vizinhos, dado que as funções de probabilidade deixarão de ser nulas. Como o módulo de  $G$  indica a quantidade de energia que se ganha ou se perde por haver uma mudança de estado, se  $G$  for negativo, existe um estado de menor energia que deverá então ser determinado. Deste modo, o pixel a ser visitado deverá ser aquele para o qual  $G$  tenha o valor negativo mais baixo para que deste modo se obtenha o maior ganho de estabilidade possível e então a ordem de visita deverá ser permanentemente actualizada para se respeitar esta condição.

O algoritmo termina quando todas as estabilidades associadas aos pixels forem positivas, isto é, quando não for mais possível baixar a energia.

#### **A.4 Implementação do Algoritmo**

Em termos de desenvolvimento do algoritmo HCF podemos considerar o seguinte esquema simplificado dos passos necessários para implementar o que foi referido anteriormente:

### ALGORITMO HCF

```
w = (l0, ..., lo)
Topo = Pilha(w)
Enquanto(GTopo < 0)
{
    Topo = Pilha(w)
    Alterar _ Estado(ws)
    Actualizar _ G(Gs)
    Ajustar _ Pilha(s)
    Para _ Cada(r ∈ ηs) {
        Actualizar _ G(Gr)
        Ajustar _ Pilha(r)
    }
}
return(w)
```

Começa-se por colocar no estado de *uncommitted* todos os pixeis, o que é feito na primeira instrução.

De seguida, estabelece-se uma ordem de visita, que no início pode ser qualquer já que todos os elementos estão *uncommitted*, e que é indicada pela ordem dos elementos de uma pilha. A segunda instrução,

$$Topo = Pilha(w)$$

destina-se a inicializar a variável *Topo* com o primeiro elemento da pilha.

Enquanto não houver estabilidades negativas, o que é garantido pela condição de controlo do ciclo *Enquanto*( ), vão-se seleccionando sucessivamente os pixeis (*Topo = Pilha(w)*) correspondentes ao valor da estabilidade mais negativa e utilizando-se a sua informação calculam-se as novas estabilidades.

Esta selecção é feita escolhendo-se entre todos os estados possíveis pertencentes ao conjunto  $L$ , e de acordo com o critério expresso nas condições (17) e (18), o estado que corresponde à energia mínima. Esta escolha é assegurada pela instrução

$$\textit{Alterar\_Estado}(w_s)$$

Como este novo estado pode ter alterado a estabilidade associada a esse pixel, há que recalculá-la, o que é feito através da instrução

$$\textit{Actualizar\_G}(G_s)$$

Eventualmente, este novo cálculo das estabilidades pode fazer com que a pilha referente aos valores da estabilidade deixe de estar ordenada crescentemente de modo a apresentar no topo da pilha o valor mais negativo tendo-se então de reordenar essa pilha,

$$\textit{Ajustar\_Pilha}(s)$$

Como a alteração do estado de um pixel pode interferir na estabilidade dos pixels vizinhos, tem-se que percorrer todos os pixels vizinhos e recalcular as suas estabilidades, o que é feito através do ciclo

$$\begin{array}{l} \textit{Para\_Cada}(r \in \eta_s) \{ \\ \quad \textit{Actualizar\_G}(G_r) \\ \quad \textit{Ajustar\_Pilha}(r) \\ \} \end{array}$$

Garante-se assim que os estados de todos os pixels vizinhos são actualizados. O índice  $r$  denota um pixel vizinho de  $s$  pertencente à *clique*  $c$ . Este ciclo compreende apenas duas instruções, as quais são equivalentes às instruções já descritas para o pixel  $s$ , isto é, determina-se a nova estabilidade para o pixel vizinho entrando em conta com o novo estado do pixel  $s$  reordenando-se de seguida a lista de estabilidades.

Ao efectuarmos uma mudança de estado aumentamos a estabilidade e diminuimos a energia. A iteração termina quando todas as estabilidades forem positivas o que corresponde a um estado global de equilíbrio estável ou seja corresponde a ter-se atingido uma configuração a que corresponde um mínimo de energia.