

ANTÓNIO MANUEL RIBEIRO DE SOUSA

Localização Automática de Objectos em Sequências de Imagens

Dissertação de Mestrado em Informática

Dissertação submetida à Universidade do Minho, para a obtenção do grau de Mestre,
elaborada sob a orientação do Professor Doutor João Luís Ferreira Sobral

Universidade do Minho
Escola de Engenharia
Departamento de Informática

Braga, 2005

É autorizada a reprodução integral desta dissertação, apenas para efeitos de investigação, mediante declaração escrita do interessado, que a tal se compromete.

Resumo

Palavras-chave: localização de objectos, detecção de sinais, seguimento de objectos, assinaturas de contorno, espaço cor *HSI*, correlação.

A detecção e seguimento de objectos tem uma grande variedade de aplicações em visão por computador. Embora tenha sido alvo de anos de investigação, continua a ser um tópico em aberto. Continua a ser ainda hoje um grande desafio a obtenção de uma abordagem que inclua simultaneamente flexibilidade e precisão, principalmente quando se trata de ambiente aberto.

O objectivo desta dissertação é o desenvolvimento de uma metodologia que permita a localização de objectos genéricos e uma outra de localização de objectos conhecidos (sinais de trânsito), em sequências de imagens em ambiente aberto, sendo, nesta última, efectuado também o seu reconhecimento.

No caso da primeira metodologia o objectivo proposto é concretizado com a indicação do objecto de interesse, através da sua selecção, numa primeira imagem, sendo o seu seguimento efectuado, numa primeira fase, recorrendo a uma aproximação grosseira à posição do objecto, utilizando informação de cor (característica interna), seguida de uma aproximação refinada, utilizando informação de forma (característica externa). No caso da segunda metodologia, a localização (detecção e seguimento) do objecto é realizada com base na informação de cor, através da segmentação de cor (azul e vermelha) no espaço cor *HSI*, e na forma, através das assinaturas de contorno. Finalmente é utilizada uma base de dados constituída pelas imagens dos objectos que se pretende reconhecer para identificar o objecto.

Para determinar a viabilidade das metodologias propostas, foram efectuados vários testes dos quais se obtiveram, para a metodologia de localização de um objecto genérico, resultados aceitáveis, tendo em conta, por um lado, a não utilização de informação específica sobre o objecto, e por outro lado a complexidade contida nas sequências de imagens testadas, obtidas de ambiente aberto.

A segunda metodologia, que corresponde à localização automática de objectos, obteve bons resultados, apesar dos testes terem sido direccionados para a sinalização rodoviária e restringida à localização de quatro formas e duas cores em concreto. A metodologia foi submetida, tal como no caso anterior, a cenas em ambiente aberto, mais concretamente 172 imagens, das quais se observaram 238 sinais de trânsito em condições de serem localizados, e dos quais resultaram 90,3% detectados correctamente por cor e forma e destes 82,8% foram reconhecidos correctamente, apesar do algoritmo utilizado nesta fase de reconhecimento ter sido aplicado apenas como abordagem inicial. Os resultados obtidos das metodologias desenvolvidas são encorajadores e um forte incentivo para continuar a apostar no seu melhoramento.

Abstract

Keywords: object localization, traffic sign detection, object tracking, contour signatures, HSI color space, correlation.

Object detection and tracking has a wide range of applications in computer vision. Although it has been studied for many years, it remains an open research problem. A flexible and accurate approach is still a great challenge today, specially in outdoor environments.

The objective of this thesis is the development of a methodology able to track generic objects and another able to localize known objects (traffic signs) and their recognition, in outdoor environment image sequences.

The proposed objective concerning the first methodology is achieved by selecting the object of interest in a first frame, and the tracking performed, in a first step, by a coarse approach to the object position, using color information (internal feature), followed by a refined approach, using shape information (external feature). In the second methodology, the object localization (detection and tracking) is based on color information, through color segmentation (blue and red) in HSI color space, and shape, through contour signatures. Object identification is performed using a database filled with the objects images to recognize.

Several tests were performed to determine the proposed methodologies effectiveness, obtaining acceptable results in the generic object localization methodology, taking into account, on one hand, the non utilization of any specific information about the object, and the other hand, the tested outdoor environment image sequences complexity.

The second methodology, corresponding to the automatic object localization, obtained good results, although the tests were directed to traffic signs and restricted to four shapes and two colors. The methodology was submitted, as in the previous case, to outdoor environment scenes, more specifically 172 images, from which 238 localizable traffic signs were spotted. In this test 90.3% color and shape were correctly detected and from these 82.8% were correctly recognized, although the algorithm used in this recognition phase is only an initial approach. The developed methodologies results are encouraging and a strong incentive for future improvements.

Agradecimentos*

Gostaria de agradecer a todos que contribuíram para este trabalho, de forma directa ou indirecta, através de apoio, sugestões ou de informação disponibilizada.

Um agradecimento especial ao Prof. Doutor João Luís Ferreira Sobral pelo seu trabalho de supervisão, através de críticas e sugestões fornecidas.

* Este trabalho foi parcialmente suportado pelo projecto SIGMA (Sistema de Georeferenciação Móvel Assistido por imagem), financiado pela Agência de Inovação.

Aos meus pais e irmão

Índice

1. Introdução	1
1.1 Objectivos do trabalho	2
1.2 Estrutura do relatório	2
2. Metodologias para detecção e seguimento de objectos	5
2.1 Detecção baseada em características	5
2.1.1 Abordagens baseadas na forma	6
2.1.2 Abordagens baseadas na cor	6
2.2 Detecção baseada em modelos	7
2.2.1 Correspondência com modelo fixo	8
2.2.2 Correspondência com modelo deformável	8
2.3 Detecção com base no movimento	9
2.3.1 Diferença entre imagens	9
2.3.2 Sistemas de energia global	10
2.4 Seguimento utilizando informação de movimento	10
2.4.1 Abordagens baseadas na forma	11
2.4.1.1 Contornos activos deformáveis (<i>Snakes</i>)	11
2.4.1.2 Contornos activos geodésicos	12
2.4.2 Abordagens baseadas na cor	13
2.4.3 Utilização informação movimento disponível em ficheiros vídeo	14
3. Módulo de PI do sistema SIGMA	15
3.1 Enquadramento	15
3.2 Localização assistida	16
3.2.1 Detecção	16
3.2.1.1 Imagem de energia	17
3.2.1.2 Extracção de características	18
3.2.2 Seguimento	18
3.2.2.1 Translação	19
3.2.2.2 Escala	20
3.2.2.3 Rotação	26
3.3 Localização automática	27

3.3.1 Detecção e segmentação de regiões por cor.....	27
3.3.1.1 Segmentação com base na cor	28
3.3.1.2 Etiquetagem	33
3.3.1.3 Filtragem.....	34
3.3.1.4 Normalização	36
3.3.2 Detecção de formas	36
3.3.2.1 Circular	38
3.3.2.2 Triangular	39
3.3.2.3 Quadrangular	40
3.3.2.4 Triangular Invertida	41
3.3.2.5 Classificador	42
3.3.2.6 Exemplo da classificação de formas.....	42
3.3.3 Seguimento.....	45
3.3.4 Reconhecimento	46
4. Resultados e discussão.....	49
4.1 Sistema assistido	49
4.1.1 Escolha do detector de contornos.....	49
4.1.1.1 Sobel3x3	49
4.1.1.2 LVIF.....	50
4.1.1.3 Comparação Sobel3x3 / LVIF	51
4.1.2 Detecção e seguimento.....	52
4.1.3 Localização de objectos - exemplos.....	54
4.2 Sistema automático	56
4.2.1 Detecção e segmentação com base na cor.....	56
4.2.2 Classificador de forma	63
4.2.2.1 Escolha do classificador de forma	63
4.2.2.2 Sensibilidade à rotação	72
4.2.3 Reconhecimento	74
4.2.4 Seguimento.....	75
5. Conclusão.....	77
6. Referências.....	81

Lista de Figuras

Figura 1 – Selecção do objecto de interesse pelo utilizador	16
Figura 2 – Objecto de interesse (a) imagem <i>RGB</i> (b) imagem de energia	17
Figura 3 – Imagem (a) <i>N</i> original (b) objecto seleccionado na imagem <i>N</i> (c) área de pesquisa do objecto na imagem <i>N+1</i>	19
Figura 4 – Objecto sujeito a translação e alteração de escala entre imagens consecutivas	20
Figura 5 – Compensação do efeito de translação do objecto	20
Figura 6 – Comparação do objecto na imagem <i>N</i> e <i>N+1</i>	21
Figura 7 – Contorno do objecto (a) na imagem <i>N</i> (b) obtida na imagem <i>N+1</i> (c) pretendida na imagem <i>N+1</i>	22
Figura 8 – Assinaturas de contorno do objecto	22
Figura 9 – Aproximação das assinaturas de contorno do objecto	23
Figura 10 – Aproximação final das assinaturas de contorno do objecto	24
Figura 11 – Suavização da assinatura de contorno	25
Figura 12 – Aproximação final suavizada da assinatura de contorno do objecto	25
Figura 13 – Árvore de detecção	27
Figura 14 – Espaço cor HSI num modelo duplo cone	29
Figura 15 – Conversão (a) <i>RGB</i> para as componentes (b) <i>H</i> , (c) <i>S</i> e (d) <i>I</i>	31
Figura 16 – Componente (a) saturação (b) cromática	32
Figura 17 – Segmentação (a) <i>Hue/Saturação</i> (b) <i>Hue/Cromatismo</i>	33
Figura 18 – Regiões etiquetadas de uma imagem	34
Figura 19 – Etiquetagem com vizinhança 4 de um pixel	34
Figura 20 – Exemplo de uma imagem (a) original e (b) ilustração da filtragem	35
Figura 21 – Resultado da filtragem	36
Figura 22 – Regiões normalizadas	36
Figura 23 – Exemplo de assinaturas resultantes de contornos	37
Figura 24 – Detecção do contorno circular	39
Figura 25 – Assinatura de contorno circular	39
Figura 26 – Detecção do contorno triangular	39

Figura 27 – Assinatura de contorno triangular.....	40
Figura 28 – Detecção do contorno quadrangular	40
Figura 29 – Assinatura de contorno quadrangular	41
Figura 30 – Detecção do contorno triangular invertido	41
Figura 31 – Assinatura de contorno triangular invertido	41
Figura 32 – Exemplo de detecção cor e forma.....	43
Figura 33 – Assinaturas de contorno das regiões detectadas	43
Figura 34 – Assinaturas de contorno das formas pré-definidas	44
Figura 35 – Seguimento de objectos numa sequência de imagens.....	45
Figura 36 – Distribuição de pixels numa máscara 3x3	50
Figura 37 – Aplicação à imagem (a) original (b) o detector de contornos <i>Sobel3x3</i> (c) e o <i>LVIF</i>	51
Figura 38 – Seguimento de um objecto numa sequência (imagens 1, 10 e 20)	52
Figura 39 – Detecção e seguimento de um objecto entre a imagem (a) N e (b)-(c) N+1 de uma sequência.....	52
Figura 40 – Sequência de operações executadas na detecção e seguimento assistido	53
Figura 41 – Localização de um sinal quadrangular numa sequência (imagens 1, 21 e 28)	54
Figura 42 – Localização de um automóvel numa sequência (imagens 1, 17 e 58).....	54
Figura 43 – Localização de um sinal circular numa sequência (imagens 1, 10 e 15).....	55
Figura 44 – Localização de um sinal quadrangular numa sequência (imagens 1, 8 e 15)	55
Figura 45 – Sinais unidos a outro sinal	58
Figura 46 – Gráfico relação altura/largura dos sinais	58
Figura 47 – Distinção entre sinais normais e unidos.....	59
Figura 48 – Resultado da separação dos sinais unidos.....	59
Figura 49 – Sinais unidos ao fundo não detectados	60
Figura 50 – Sinais partidos parcialmente/não detectados	61
Figura 51 – Sinais partidos parcialmente/não detectados recuperados	62
Figura 52 – Resultados obtidos para o 1º Classificador.....	65
Figura 53 – Resultados obtidos para o 2º Classificador.....	66
Figura 54 – Resultados obtidos para o 3º Classificador.....	66
Figura 55 – Resultados para o método complementar do 1º Classificador.....	68
Figura 56 – Resultados para o método complementar do 2º Classificador.....	68

Figura 57 – Resultados para o método complementar do 3º Classificador	69
Figura 58 – Exemplos de erros de classificação por forma.....	71
Figura 59 – Assinaturas de duas formas pré-definidas e da região com forma errada.....	71
Figura 60 – Resultado da rotação da forma <i>Triangular</i>	73
Figura 61 – Resultado da rotação da forma <i>Triangular Invertida</i>	73
Figura 62 – Resultado da rotação da forma <i>Quadrangular</i>	73
Figura 63 – Base de dados de sinais para reconhecimento	74
Figura 64 – Sinais com ideogramas semelhantes.....	75
Figura 65 – Sinais com grafismo variável.....	75
Figura 66 – Reconhecimento de sinais com sombras.....	75
Figura 67 – Detecção de um sinal de trânsito (a)-(k) numa sequência de imagens	76

Lista de Tabelas

Tabela 1 – Algoritmo de conversão <i>RGB</i> para <i>HSI</i>	30
Tabela 2 – Valores de <i>Hue</i> e Saturação para a segmentação de cor.....	31
Tabela 3 – Valores de <i>Hue</i> e Cromatismo utilizados na segmentação de cor	32
Tabela 4 – Medidas obtidas do classificador de forma	44
Tabela 5 – Formas dos sinais analisados.....	56
Tabela 6 – Resultado filtro de dimensão.....	57
Tabela 7 – Resultado da detecção inicial por cor.....	57
Tabela 8 – Resultado final da detecção inicial por cor	62
Tabela 9 – Sinais analisados pelo classificador de forma	63
Tabela 10 – Valores obtidos para os classificadores.....	67
Tabela 11 – Falsos positivos e negativos obtidos para os classificadores.....	67
Tabela 12 – Resumo dos valores do classificador.....	69
Tabela 13 – Resultados obtidos da detecção e classificação por forma	70
Tabela 14 – Medidas obtidas do classificador de forma	72
Tabela 15 – Resultados do reconhecimento dos sinais	74

1. Introdução

Os vídeos são na realidade sequências de imagens, cada uma das quais normalmente designada de quadro (*frame*), exibidas numa frequência suficientemente rápida para que o olho humano não se aperceba da cadência de imagens. Este efeito deve-se também ao facto do conteúdo de duas quaisquer imagens consecutivas ser normalmente muito semelhante, permitindo a aplicação de metodologias de localização de objectos com base em técnicas de processamento de imagem, baseadas nessa semelhança.

O conteúdo visual de uma imagem pode ser modulado como uma hierarquia de abstracções. No primeiro nível estão os pixels com informação de cor ou brilho. Processando esta informação podem ser obtidas características tais como contornos, cantos, linhas, curvas, e regiões de cor. Uma camada de abstracção mais elevada pode combinar e interpretar estas características como objectos e respectivos atributos. No nível mais elevado estão os conceitos de nível humano envolvendo um ou mais objectos e relações entre eles, ou seja, a compreensão da imagem.

A detecção de objectos em vídeos envolve a verificação da presença de um determinado objecto em sequências de imagens e possivelmente a determinação da posição exacta. O seguimento de objectos consiste na monitorização das alterações sofridas por um determinado objecto durante uma sequência de vídeo, incluindo a sua presença, posição, tamanho, forma, etc. Isto é conseguido resolvendo o problema da correspondência entre imagens, ou seja, o problema de corresponder a região alvo em imagens sucessivas de uma sequência de imagens obtidas com intervalos de tempo reduzidos. A detecção e seguimento estão intimamente relacionados porque o seguimento requer a detecção inicial do objecto. A detecção e seguimento de um qualquer objecto corresponde à sua localização numa ou em várias imagens.

A detecção e seguimento de objectos tem uma grande variedade de aplicações em visão por computador tal como a compressão de vídeo, vídeo vigilância, controlo baseado em visão, interfaces humano – computador, imagiologia médica, auxílio na geo-referenciação e robótica. Além disso, fornece informação necessária em tarefas de visão de nível mais elevado, tais como reconstrução 3D e representação 3D. Desempenha também um papel importante em bases de dados vídeo por exemplo na sua indexação e extracção baseada no conteúdo.

Embora tenha sido alvo de anos de investigação, a detecção e seguimento de objectos continua a ser um tópico em aberto. Continua a ser ainda hoje um grande desafio a obtenção de uma abordagem que inclua simultaneamente flexibilidade e precisão, principalmente quando se trata de ambiente aberto. O nível de dificuldade do problema depende em grande medida da forma como é definido o objecto a ser detectado e seguido. Se apenas um número reduzido de características visuais, tal como uma cor específica,

forem utilizadas como representação de um objecto, torna-se relativamente mais fácil a identificação de todos os pixels com a mesma cor como o objecto. No outro extremo, a face de uma pessoa específica, com detalhes que variam com a postura e a iluminação, torna-se complexa a sua correcta detecção, reconhecimento e seguimento. Um desafio importante surge da variabilidade da imagem vídeo porque os objectos vídeo se encontram geralmente em movimento. À medida que um objecto se move através do campo visual de uma camera, as imagens do objecto podem sofrer alterações significativas. Esta variabilidade surge de três fontes principais: variação na postura ou deformações do alvo, variações na iluminação, e oclusão parcial ou total do alvo [Hager98].

1.1 Objectivos do trabalho

Este trabalho pretende desenvolver um módulo de processamento de imagem que inclua metodologias que permitam a localização de objectos genéricos e objectos conhecidos (sinais de trânsito) em sequências de imagens em ambiente aberto sendo, na localização de objectos conhecidos, efectuado também o seu reconhecimento.

Este objectivo é concretizado de duas formas: primeiro, através do desenvolvimento de uma metodologia que permita uma localização assistida (mais genérica), através da detecção e seguimento de um objecto genérico definido manualmente pelo utilizador numa primeira imagem; segundo, o desenvolvimento de uma metodologia que permita a localização automática (mais específica), através da detecção, seguimento e reconhecimento de objectos pré-definidos no sistema.

Relativamente ao módulo de processamento de imagem, apostou-se pelo desenvolvimento de um protótipo que permitisse realizar um trabalho experimental, possibilitando a validação dos conceitos e das metodologias propostas assim como aferir a sua viabilidade através do estudo de casos concretos em testes efectuados com imagens ou sequências de imagens em ambiente aberto.

1.2 Estrutura do relatório

No próximo capítulo, são analisados sistemas e metodologias comuns para detecção e seguimento de objectos genéricos quer eles sejam deformáveis ou não deformáveis. É dada particular ênfase à forma como se resolvem problemas recorrentes neste tipo de sistemas, tais como problemas de iluminação, oclusões, elementos de distração da imagem (*clutter*), transformações de escala, rotação, e translação.

O terceiro capítulo, inclui o enquadramento do trabalho efectuado no sistema para o qual foi desenvolvido. São explicadas detalhadamente as metodologias e algoritmos implementados no módulo de PI (processamento de imagem), quer na localização assistida

quer na localização automática que inclui também a abordagem utilizada no reconhecimento de objectos.

O quarto capítulo, é dedicado à apresentação dos resultados obtidos da aplicação das metodologias implementadas a casos reais, assim como a explicação de determinadas opções tomadas no decurso do seu desenvolvimento resultantes de experimentação efectuada. A inclusão, neste capítulo, dessas explicações permite complementá-las com resultados obtidos utilizando exemplos práticos.

No quinto e último capítulo, são apresentadas as conclusões finais sobre o trabalho desenvolvido, as suas potencialidades, possíveis melhorias a efectuar e apontadas algumas sugestões como trabalho futuro.

2. Metodologias para detecção e seguimento de objectos

A detecção de objectos em vídeos envolve a verificação da presença de um determinado objecto em sequências de imagens e possivelmente a determinação da posição exacta. O seguimento de objectos consiste na monitorização das alterações sofridas por um determinado objecto durante uma sequência de vídeo, incluindo a sua presença, posição, tamanho, forma, etc.

Existem duas fontes de informação numa sequência de imagens que podem ser utilizadas na detecção e seguimento de objectos: imagens isoladas, de onde se podem extrair características visuais como a cor, textura e forma, e diferença entre imagens obtendo-se informação de movimento. A combinação de análise das características visuais das imagens isoladas e de informação de movimento temporal, resultante da diferença entre imagens, normalmente resulta em abordagens mais eficientes. Uma das estratégias mais frequentemente utilizada neste tipo de abordagem começa pela segmentação da imagem em regiões, baseada na cor e textura, seguida da união de regiões com vectores de movimento semelhantes.

Têm sido propostas uma grande quantidade de abordagens na literatura para a detecção e seguimento de objectos. Grande parte delas faz uso de múltiplas técnicas, existindo também combinações e intersecções entre diferentes métodos. Este facto torna complexa a tarefa de obtenção de uma classificação uniforme das abordagens existentes.

Existem abordagens em que a detecção de objectos é efectuada com base apenas em algumas das suas características como seja a cor ou a forma. Existem outras que se baseiam no modelo do objecto, efectuando a sua correspondência, as quais podem ou não lidar com problemas de deformação dos objectos. Noutros casos, a detecção é efectuada com base no movimento do objecto. O seguimento, por outro lado, é efectuado com base na informação de movimento da forma, da cor, ou informação de movimento dos objectos, disponível em ficheiros vídeo.

Nas subsecções seguintes, serão analisadas separadamente abordagens propostas tendo em conta estas diferentes vertentes de investigação.

2.1 Detecção baseada em características

Neste tipo de abordagem são extraídas uma ou mais características dos objectos de interesse, sendo estes modulados com base nessas características. A detecção e

reconhecimento dos objectos podem ser transformados num problema de correspondência de grafos. Na detecção de objectos baseada em características, é importante a normalização dessas características. As imagens podem ter de ser transformadas para lidarem com alterações na iluminação (espaço de cor diferente), tamanho e orientação (transformações de escala e rotação). As características mais comuns utilizadas neste tipo de abordagem são as de forma e cor.

2.1.1 Abordagens baseadas na forma

Estas abordagens baseiam-se na característica de forma dos objectos, em que a sua forma é representada por um conjunto discreto de pontos do seu contorno. O processo é normalmente iniciado com a segmentação dos objectos de interesse nas imagens. A sua forma pode ser encontrada por algoritmos de detecção e seguimento de contornos. A detecção e caracterização por forma dos objectos torna-se mais difícil em cenas complexas, onde existem problemas de rotação e escala, variações do fundo ou objectos com oclusões e sombras [Falou94] [Pentl94].

As características de forma do objecto são obtidas em [Grigo03] tendo em conta a distribuição espacial dos seus pontos de contorno. Essa distribuição é definida através da escolha dos n pontos que melhor possam caracterizar a sua forma, por exemplo cantos, sendo de seguida calculado o conjunto das distâncias entre cada um dos pontos seleccionados e todos os outros. A detecção da forma numa nova imagem é iniciada, nos testes efectuados a sinais de trânsito, com o cálculo da transformada *wavelet* de 2º nível. Seguidamente são eliminados em cada banda resultante da *wavelet* os contornos de menor amplitude e efectuada a sua sobreposição. Na imagem resultante é calculado, para cada ponto, o conjunto das suas distâncias a todos os outros. Finalmente é efectuada a correspondência entre os dois conjuntos de conjuntos de distâncias que minimiza a sua diferença. Esta abordagem é invariante à rotação e translação uma vez que apenas são determinadas distâncias, mas é dependente da escala, a qual pode ser tornada independente bastando para isso, por exemplo, dividir todas as distâncias de cada conjunto pela distância que une os dois pontos mais afastados. Como todas as distâncias variam em proporção com o tamanho, as distâncias normalizadas mantêm-se constantes.

2.1.2 Abordagens baseadas na cor

Ao contrário de outras características de imagem (p.ex.: forma), a cor é menos sensível a alterações do ponto de visualização e computacionalmente mais simples a sua obtenção. Embora a cor nem sempre seja apropriada como único meio para detectar e seguir objectos, os baixos custos computacionais dos algoritmos propostos tornam a cor uma característica desejável a explorar sempre que tal seja apropriado [Vinod97].

Em [Grove98] é proposto um algoritmo para detectar e seguir veículos ou peões em tempo real utilizando uma técnica baseada no histograma de cor. Este algoritmo requer

duas condições ao fundo da sequência de imagens: que se mantenha constante ao longo do tempo e a existência de um conjunto de imagens representativas desse fundo. Com base nessas imagens é criado um Modelo de Mistura Gaussiana (*Gaussian Mixture Model*) para descrever a sua distribuição de cor com base nas componentes *Hue* e *Saturação* do espaço de cor *HSI*. Este modelo de cor é utilizado para segmentar as imagens da sequência utilizando um nível de limiar e determinar assim a presença de objectos na imagem. Seguidamente são utilizados operadores morfológicos para eliminar ruído e unir regiões resultantes da segmentação, das quais são apenas extraídas as correspondentes a potenciais objectos tendo como critério a sua área. Nas imagens seguintes, as cores dos pixels não pertencentes ao fundo são utilizadas para determinar um histograma de cor que se aproxime da distribuição dos pixels de cada objecto através da correlação. A oclusão dos objectos foi resolvida utilizando um *buffer* de oclusão. A aplicabilidade deste algoritmo é no entanto limitada a situações em que os objectos tenham cores claramente diferentes do fundo da imagem.

[Fiegu97] efectua o seguimento de faces humanas em tempo real usando informação de cor. São definidas regiões com tamanho fixo e posições relativas dentro dos limites do objecto a ser seguido. Cada região é caracterizada por um vector de cor calculado por sub-amostragem dos pixels pertencentes à região, o qual representa a cor média dos pixels dessa região. Este método de seguimento simples é baseado no seguimento de regiões com cor semelhante de imagem para imagem. A semelhança de cor é obtida, para cada região com centro em (x, y) , do cálculo da razão entre o seu vector de cor e o de cada uma das M regiões vizinhas. Cada uma dessas regiões tem centro em $(x + \Delta x_m, y + \Delta y_m)$, em que $\Delta x_m, \Delta y_m$ representa o deslocamento da região M em relação à região a ser seguida. A correspondência entre as regiões do objecto na imagem antiga e na nova imagem é obtida da semelhança máxima de cor utilizando uma abordagem designada por *maximum likelihood approach*. O processo de oclusão foi tratado através da sua modulação explícita, ou seja, sempre que existe oclusão de um objecto é efectuada uma previsão da sua localização. Este objectivo foi alcançado utilizando um modelo designado por *hypothesis-tree model*.

2.2 Detecção baseada em modelos

Quando existir e for adequada a utilização de um modelo (*template*) que descreva um objecto específico, a detecção de objectos transforma-se num processo de correspondência (*matching*) de características entre o modelo e a sequência de imagens em análise. A detecção desses objectos implica a correspondência de todo o modelo, exigindo geralmente um carga computacional elevada. A qualidade da correspondência depende não só do grau de detalhe fornecido pelo modelo do objecto mas também da sua capacidade para lidar com problemas como a translação, rotação, escala, oclusões e deformações do objecto. A correspondência de modelos (*template matching*) de objectos pode ser dividida

basicamente em correspondência com modelo fixo e correspondência com modelo deformável.

2.2.1 Correspondência com modelo fixo

Um modelo fixo (imagem de referência) é normalmente localizado numa outra imagem (imagem de pesquisa) correlacionando o modelo com a imagem de pesquisa, dentro de um determinado limite de deslocamento. A posição do modelo, na imagem de pesquisa, é encontrada quando a correlação entre as imagens é máxima. Os modelos fixos são úteis quando a forma do objecto não se altera em relação ao ângulo de visão da camera. Além dos problemas referidos anteriormente, existem também, neste tipo de modelos, problemas com o seu brilho médio. Dado que o cálculo da correlação envolvendo toda a imagem, tem um custo computacional elevado é necessária a utilização de algoritmos com custo mais reduzido, sempre que esteja envolvido o processamento de imagens em tempo real. Uma das abordagens mais conhecidas é a utilização de pirâmides de resolução que permite reduzir consideravelmente o tamanho das imagens envolvidas, sendo a pesquisa do modelo efectuada da menor para a maior resolução.

[Kratt94] propõe a correlação com determinados pontos seleccionados do modelo em vez de todo o modelo. São determinadas todas as combinações possíveis de M pontos para o modelo e seleccionadas aleatoriamente N combinações. Seguidamente é efectuada a correlação entre os pontos pertencentes a cada uma das N combinações e escolhida aquela cuja correlação entre pontos é mais elevada. O passo seguinte é efectuar recursivamente a correlação dos pontos da combinação com cada um dos outros pontos do modelo sendo consecutivamente ordenados pelo resultado obtido. Os M pontos com melhor correlação são então utilizados no processo de correspondência.

2.2.2 Correspondência com modelo deformável

As abordagens de correspondência de modelos deformáveis [Sclar01] [Zhong00] [Schoe01] (*deformable template matching*) são mais adequadas a casos onde os objectos se modificam devido apenas a deformações rígidas (i.é. transformações de translação, rotação, ou escala) e não rígidas, ou seja, alterações na estrutura dos objectos. Os modelos deformáveis são mais apelativos em tarefas de seguimento devido à natureza deformável dos objectos na maioria dos vídeos, causada frequentemente por alterações da distância entre os objectos e a camera.

Em [He02] o objecto é modulado por características locais e globais. As características locais são obtidas da selecção de pontos característicos com base na sua energia, resultantes da aplicação da transformada *wavelet* de Gabor (*Gabor Wavelet Transform*). As características globais são representadas por áreas triangulares, obtidas da malha 2D constituída pelas posições espaciais dos pontos característicos. Esta abordagem tira partido das características das *wavelets* de Gabor, as quais são muito localizadas tanto

no domínio do espaço como das frequências, e da invariância da malha 2D a transformações do plano. A detecção do objecto na imagem seguinte é iniciada com o deslocamento da malha 2D na vizinhança de cada ponto característico e encontrado o máximo local. O processo utilizado na pesquisa do máximo local é um algoritmo designado por secção dourada (*Golden Section Algorithm*) o qual é normalmente utilizado para determinar o máximo de uma função unimodal. O objecto estará localizado na nova imagem quando forem encontrados os máximos locais que maximizam a semelhança entre características. Esta metodologia foi testada no seguimento de faces.

Apesar da flexibilidade deste tipo de abordagem, estas exigem no entanto que o fundo da imagem seja uniforme. A aplicação mais importante das técnicas de correspondência de modelo deformável é na detecção de movimento de objectos em imagens vídeo a qual será analisada na secção seguinte.

2.3 Detecção com base no movimento

Neste tipo de abordagens a detecção de objectos é baseada na informação de movimento por eles produzida, devido a alterações espaciais e temporais. A detecção de objectos com base no movimento, ou detecção de movimento, tem muita importância na detecção e seguimento de objectos em vídeo. Na última década foi concentrado neste problema um grande esforço de investigação. Comparada com a detecção de objectos sem movimento, a detecção de movimento, por um lado, torna mais complexo o problema da sua detecção acrescentando requisitos de alteração temporal do objecto, por outro lado, também fornece outra fonte de informação para a detecção e seguimento, ou seja, a informação de movimento.

Têm sido propostos vários algoritmos de detecção de movimento, os quais podem ser divididos entre os que utilizam para esse fim a diferença entre imagens e os que utilizam sistemas de energia global.

2.3.1 Diferença entre imagens

Em casos em que a camera se encontra estática, a detecção de movimento é frequentemente baseada apenas na diferença entre imagens. Essa diferença pode ser obtida utilizando abordagens mais simples como a diferença ao nível de pixel ou ao nível de bloco de pixels, aumentando a tolerância ao ruído, sendo efectuada a binarização da imagem resultante utilizando um valor pré definido. Em abordagens mais sofisticadas a diferença entre imagens pode ser modulada como uma mistura de distribuições *Gaussianas* ou *Laplacianas*. Em muitos casos reais a aplicação apenas da diferença entre imagens não é suficiente devido à existência de movimento com origem no movimento da camera. Este problema pode no entanto ser resolvido calculando o movimento da camera e criando uma nova sequência compensada.

A presença de movimento provoca normalmente o aparecimento de três tipos de alteração entre imagens:

- fundo estático destapado: região da imagem onde parte do objecto esteve na imagem anterior e já não está;
- fundo estático tapado: região da imagem onde parte do objecto está na imagem actual e não estava na anterior;
- sobreposição de duas sucessivas projecções do objecto: zona da imagem em que existe sobreposição de parte do objecto da imagem anterior e actual.

Este último tipo de regiões é de difícil obtenção quando a superfície do objecto é muito uniforme, sendo necessária a aplicação de técnicas complementares que permitam a extracção de informação específica acerca da localização exacta dos objectos em movimento. Uma dessas técnicas consiste na aplicação de máscaras de detecção de alterações temporais como os detectores de alterações espacio-temporais de *Gabor*. [Parag99].

A abordagem referida anteriormente, para a detecção de movimento, baseia-se na diferença entre imagens consecutivas (*inter-frame difference*). Existem no entanto outras abordagens, também por vezes utilizadas, em que a diferença é efectuada entre a imagem actual e a de referência (*reference difference*), sendo a imagem de referência constituída apenas pelo fundo. No caso de não ser disponibilizada imagem de referência, abordagens mais complexas, efectuem a construção da mesma com informação estática recolhida ao longo da sequência.

2.3.2 Sistemas de energia global

Neste tipo de abordagem o problema de detecção de movimento é formulado minimizando uma função objectivo global, sendo normalmente implementada utilizando algoritmos de relaxamento estocásticos (*Mean-Field*, *Simulated Annealing*) ou determinísticos (*Iterated Conditional Modes*, *Highest Confidence First*). Os *Markov Random Fields* [Parag99] espaciais têm sido amplamente utilizados nesse sentido e a detecção de movimento passou a ser considerada como um problema de cálculo estatístico. Embora esse cálculo seja muito poderoso, requer também normalmente elevados recursos computacionais.

2.4 Seguimento utilizando informação de movimento

Este tipo de abordagens utiliza métodos cuja função é efectuar o agrupamento de informação visual semelhante ao longo do tempo. O seguimento dos objectos é efectuado com base em medições sobre as suas propriedades geométricas, ou seja, de forma, ou sobre

as propriedades da sua região, i.é. cor. As abordagens existentes podem então ser divididas com base no tipo de propriedades que utiliza [Parag00]:

- abordagens baseadas na forma: utilizam a informação fornecida pelos contornos do objecto;
- abordagens baseadas na cor: utilizam a informação fornecida por toda a região do objecto tal como, a textura e propriedades baseadas no movimento.

2.4.1 Abordagens baseadas na forma

Este tipo de abordagens baseiam-se na informação fornecida pela forma do objecto. Têm sido amplamente adoptadas no seguimento de objectos porque a informação obtida não depende do tipo de movimento, nem da forma do objecto. Estes modelos podem utilizar abordagens de minimização da energia ou de geometria dos objectos, evoluindo de uma curva inicial sob a influência de energias “externas” e sendo ao mesmo tempo restringidos por energias “internas”. Os modelos utilizados podem ser divididos entre os que necessitam de parametrização (contornos activos deformáveis) e os que não necessitam (contornos activos geodésicos).

2.4.1.1 Contornos activos deformáveis (*Snakes*)

O seguimento de formas utilizando contornos deformáveis (*deformable active contours*), mais conhecidos por *snakes*, foi originalmente introduzido por Terzopoulos [Kass87]. As *snakes* são modelos deformáveis baseados na minimização de energia ao longo de uma curva. Essa curva, deforma-se com o objectivo de minimizar as energias “internas” e “externas” ao longo do seu contorno. A energia interna provoca a suavização da curva, enquanto que a energia externa transporta a curva em direcção à forma do objecto na imagem [Golde01]. As *snakes* são designadas por contornos activos porque se adaptam às alterações de forma do objecto sobre o qual são definidos inicialmente. Esta adaptação ou convergência é efectuada deslocando o contorno com base na influência da intensidade da imagem (energias externas), obtidas do seu gradiente, e submetidas a certas restrições de deformação internas obtidas da própria curva (energias internas) [Peter99]. Uma das vantagens do modelo *snake* ser induzido por energias ou forças é o facto de poder incluir mais facilmente nas suas equações de movimento as dinâmicas, derivadas de formas que variam com o tempo.

As *snakes* necessitam de parametrização sendo habitualmente efectuada utilizando *B-splines* [Curwe92] e o espaço de soluções é restringido a uma forma pré-definida. Estes métodos exigem a definição do contorno inicial próximo à forma real do objecto. Em [LuoH00] é no entanto proposto ultrapassar o problema de inicialização do contorno através da utilização, não só da informação de forma do objecto, como da informação de cor da região, na obtenção das forças externas da *snake*. O contorno resultante da inicialização converge iterativamente em direcção à solução de uma equação diferencial

parcial (*PDE - Partial Differential Equation*) [Araki98] [Lin97], ou seja, o contorno inicial é deformado, com base na forma do objecto, utilizando *PDE's*.

Esta convergência, pode no entanto não ser efectuada com precisão, devido à reduzida capacidade dos modelos de contornos activos em distinguir objectos ou fundos com características complexas. Esta reduzida capacidade, deve-se ao facto da deformação dos modelos (convergência do contorno) ser efectuada principalmente com base em características locais de imagem ao longo da forma do objecto. Em [Matsu99] é proposta uma abordagem para reduzir o efeito deste problema. O primeiro passo consiste em definir as curvas iniciais, uma dentro do objecto e outra no seu exterior, envolvendo-o. Seguidamente, as curvas iniciais são seccionadas automaticamente sendo gerado para cada uma um novo contorno inicial. Posteriormente é determinada, para cada uma, a semelhança de cor do seu interior, com base na sua distribuição. Tendo como referência essa semelhança é efectuado entre eles uma competição pela região. Finalmente a forma do objecto é encontrada passando a ser definida por um conjunto de contornos activos.

Normalmente neste tipo de modelos, a função objectivo, ou seja, a convergência entre o contorno definido e a forma real do objecto, é determinada com base na informação de intensidade. Em [Kosch02] é apresentada uma extensão a estes modelos, utilizando informação de cor *RGB* na fase de convergência do contorno definido, com a forma do objecto.

2.4.1.2 Contornos activos geodésicos

Os contornos activos geodésicos (*geodesic active contours*) são modelos dinâmicos, introduzidos por Casseles [Casse95] como uma alternativa geométrica às *snakes* [Kass87]. Dada uma curva inicial, a evolução dessa curva é definida por medidas geométricas externas e internas à curva, designadas por fluxo geométrico (vectores de velocidade). A aproximação à forma do objecto é obtida minimizando a função que define esse fluxo, normalmente através de *PDE's*. O modelo derivado do fluxo geométrico é designado por modelo geométrico, mais conhecido por modelo de contorno activo geodésico. Devido à natureza geométrica do modelo este não necessita de parametrização. Apesar destes modelos terem algumas vantagens em relação aos modelos de contornos activos deformáveis, uma das desvantagens é a sua não linearidade, resultando em implementações ineficientes. Em [Golde01] foi proposta uma nova implementação numérica do modelo de contorno activo geodésico que tenta aumentar a eficiência deste modelo.

Em [Casse96], é proposta uma abordagem em três passos, a qual começa por detectar os contornos dos objectos a serem seguidos. De seguida é calculada uma estimativa do vector de velocidade ao longo dos contornos detectados. Seguidamente, é definida uma *PDE* para deslocar os contornos em direcção às formas dos objectos em movimento. Estes contornos são posteriormente utilizados como aproximações iniciais aos contornos na imagem seguinte e o processo repete-se. Mais recentemente, foi proposto em [Bertal99] um algoritmo para o seguimento de objectos baseado em duas *PDE's*. Uma das *PDE's* efectua a deformação da primeira imagem em relação à segunda. A outra *PDE*

efectua a deformação das curvas de interesse da primeira imagem. O seguimento é efectuado projectando as velocidades da primeira equação na segunda. Neste algoritmo são utilizadas as imagens e curvas de interesse dos objectos a seguir.

2.4.2 Abordagens baseadas na cor

Estas abordagens utilizam técnicas de cálculo do movimento das regiões dos objectos com base na cor/textura da sua região. Neste caso, o cálculo da velocidade do alvo é baseada na correspondência, em diferentes instantes de tempo, da sua região. Esta operação consome normalmente muito tempo, devido à necessidade de uma correspondência ponto a ponto em toda a região. A aceleração dessa operação pode ser conseguida através da utilização de modelos de movimento paramétricos, que descrevam o movimento do alvo com um conjunto limitado de parâmetros. Em casos em que os objectos estejam sujeitos a deformações, a utilização destes modelos introduz uma maior dificuldade no seguimento da forma real do objecto. A sua utilização, aumenta no entanto o desempenho do processo de seguimento, uma vez que é explorada a informação fornecida por toda a região do objecto [Parag00]. O movimento aparente de texturas com brilho elevado numa sequência de imagens é normalmente designado por fluxo óptico (*optical flow*) ou fluxo de movimento (*motion flow*) sendo este um dos métodos mais utilizados nesta categoria.

Os algoritmos de fluxo óptico calculam o vector de deslocamento de cada ponto do campo óptico relativo ao observador. O fluxo óptico representa o vector de velocidade de cada pixel na imagem. Uma das formas mais básicas de efectuar o seu cálculo consiste na utilização dos mínimos quadrados (*least squares*) [Yan03]. As várias metodologias existentes para o seu cálculo podem classificar-se com base no tipo de informação que utilizam, ou seja, informação de gradiente [Pless00], correlação e energia [Yan03]. Os métodos mais tradicionais de cálculo do fluxo óptico são aplicados apenas em imagens com brilho constante e homogéneas. Parte deles têm dificuldade em conseguir calcular correctamente o fluxo em situações de descontinuidade do movimento, regiões com oclusão e descontinuidades na intensidade da imagem [Yan03]. Este tipo de métodos exige alguma carga computacional sendo no entanto atractiva a sua utilização na detecção e seguimento de objectos em vídeos com o fundo em movimento ou obtidos de uma camera em movimento.

Um dos problemas centrais em aplicações de vídeo vigilância aérea consiste na detecção e seguimento de pequenos objectos, com movimento independente, em imagens com ruído. Em [Pless00], é proposta a introdução do fluxo óptico como forma de resolver esse problema. O processo é iniciado com a obtenção do modelo de movimento do fundo da imagem, para estabilizar o plano de fundo, utilizando o fluxo óptico, com base no seu gradiente. De seguida, é detectado o movimento independente, como sendo fluxo residual ou fluxo na direcção do gradiente da imagem, que não é detectado pelo movimento do plano de fundo. Isto é efectuado “subtraindo” os valores obtidos para o fluxo óptico na nova imagem, com os obtidos para o modelo de movimento do fundo. Embora tenha um

custo de computação ligeiramente mais elevado, esta medida tem um significado geométrico mais directo do que a obtida da subtracção do fundo numa imagem estabilizada.

2.4.3 Utilização informação movimento disponível em ficheiros vídeo

As abordagens anteriores têm de extrair informação de movimento dos valores dos pixels pertencentes às imagens dos vídeos. Nestes casos, existe a necessidade da aplicação de técnicas mais ou menos complexas de análise e processamento ao nível do pixel. Técnicas essas que poderão exigir, em alguns casos, carga computacional elevada. Na verdade, já existe informação de movimento em vídeos codificados em formatos vídeo standard e globalmente utilizados. Por exemplo, o *MPEG* adoptou uma técnica designada de cálculo/compensação de movimento, a qual efectua o cálculo dos deslocamentos de blocos de pixels, com correspondência em imagens consecutivas, como vectores de movimento, codificando-os em ficheiros *MPEG* standard. Em muitos casos, os valores do vector de movimento reflectem com fidelidade o movimento de objectos na cena, especialmente objectos com texturas definidas. [Khan01] [Jones99] em vez de terem um módulo separado para extrair informação de movimento, utilizaram directamente estes valores do vector de movimento, ou uma versão suavizada [Ahmad03]. No entanto, a utilização da informação do vector de movimento, poderá reduzir a precisão na determinação da fronteira do objecto, uma vez que os vectores de movimento são associados apenas com blocos de pixels, cuja dimensão é definida no *MPEG*. Este facto poderá, no entanto, não resultar num grande problema, em aplicações cujo objectivo do seguimento, não seja o de determinar a correspondência exacta, para cada imagem, mas sim o de determinar, de forma global, o movimento de toda uma região alvo durante uma sequência de imagens. A relativa facilidade na aquisição de informação de movimento e os algoritmos de seguimento altamente eficientes ao nível de bloco tornam este tipo de abordagem atractiva.

3. Módulo de PI do sistema SIGMA

O módulo de PI (Processamento de Imagem) do sistema SIGMA é constituído basicamente por duas metodologias, uma para a localização assistida outra para a localização automática de objectos de interesse, incluindo nesta última também o seu reconhecimento, em sequências de imagens.

Neste capítulo, será efectuado um enquadramento do trabalho desenvolvido nesta dissertação, no sistema SIGMA, assim como a análise detalhada dos procedimentos utilizados em cada uma das metodologias implementadas no módulo de PI.

3.1 Enquadramento

O SIGMA (Sistema de Georeferenciação Móvel Assistido por imagem) é um projecto de investigação em consórcio, parcialmente financiado pela ADI (Agência De Inovação), no âmbito do Programa Operacional Sociedade da Informação – POSI.

Este projecto pretende melhorar o processo de aquisição de informação geo-referenciada utilizando um Sistema Móvel Multi-sensor (SMM), sendo o objectivo deste tipo de sistemas a aquisição de grande quantidade de informação necessária para a criação e actualização de bases de dados para Sistemas de Informação Geográfica (SIG).

A GEONAV, empresa pertencente ao consórcio, dispõe já de um SMM designado por (GISMAP) para aquisição de informação geo-referenciada, que integra medidas *GPS* com medidas obtidas a partir de uma unidade inercial. A experiência adquirida com a sua utilização permitiu concluir que a integração com sistemas de aquisição de imagem e a automatização do processo de análise e tratamento de dados aumentaria significativamente as potencialidades e campos de aplicação do sistema.

A inovação que se pretende introduzir prende-se com a implementação da capacidade de aquisição, a partir de um veículo em movimento, de grandes volumes de informação usando sensores digitais de imagem simultaneamente com informação de *GPS*, com a posterior identificação e geo-referenciação automática de objectos nas imagens, e correspondente armazenamento em base de dados SIG.

O módulo de processamento de imagem do sistema SIGMA, será baseado nas metodologias desenvolvidas nesta dissertação que incluem a localização assistida (i.é., com intervenção do utilizador) e automática de objectos de interesse e que serão abordadas de seguida.

3.2 Localização assistida

A metodologia desenvolvida baseia-se em duas fases: uma aproximação grosseira à posição do objecto, utilizando informação de cor, seguida de uma aproximação refinada, utilizando informação de forma, tendo como base a filosofia dos contornos activos deformáveis (*snakes*).

O termo *localização assistida* ou *semi automática*, no contexto do trabalho desenvolvido, significa apenas que existe uma intervenção por parte do utilizador do sistema na fase inicial de localização de um determinado objecto, definindo-o na imagem inicial de uma sequência de imagens. A partir dessa definição são extraídas características de forma e cor desse objecto que permitirão a partir desse momento o seu seguimento automático nas imagens seguintes da sequência.

Nesta secção são apresentados os algoritmos implementados no sistema assistido de detecção e seguimento de objectos em sequências de imagens.

3.2.1 Detecção

A detecção é efectuada inicialmente através da indicação do objecto a ser seguido por parte do utilizador do sistema. A partir desse momento a localização é efectuada automaticamente nas imagens seguintes, utilizando para isso informação de cor do objecto (imagem *RGB*) e a imagem de energia do mesmo, que permitirá extrair a informação de forma utilizando as assinaturas de contorno.

A fase de detecção inicial é a única que necessita da intervenção do utilizador do sistema. A definição do objecto de interesse é efectuada manualmente numa determinada imagem da sequência envolvendo-o com um rectângulo. Esta definição indica ao sistema a zona da imagem onde ele se encontra (Figura 1). Esta intervenção permite flexibilizar o sistema no seguimento de objectos de interesse, uma vez que não são impostas qualquer tipo de restrições ao tipo de objecto seleccionado. Uma maior flexibilidade do sistema implica, no entanto, também como consequência uma maior probabilidade de erro no processo de seguimento.



Figura 1 – Selecção do objecto de interesse pelo utilizador

3.2.1.1 Imagem de energia

O algoritmo utilizado pelo sistema na obtenção de informação de forma do objecto baseia-se no cálculo da variância local da imagem σ_t^2 (imagem de energia) utilizando a imagem em níveis de cinzento do objecto. A caixa definida pelo utilizador para seleccionar o objecto de interesse, em geral, inclui duas partes: o objecto e o fundo. Normalmente é utilizada a segmentação, com base em características de cor ou textura, para efectuar essa separação. No entanto, essa segmentação introduziria, erros no contorno do objecto que não são desejáveis uma vez que o sistema implementado se baseia em parte na informação de forma do objecto.

Inicialmente é aplicado à imagem *RGB* um filtro de mediana, que tem como objectivo eliminar possível ruído caracterizado por picos isolados de energia, porque o algoritmo utilizado pelo sistema nessa detecção de contornos se baseia no cálculo da variância local σ_t^2 como medida da componente de alta frequência [Won02] num dado pixel t :

$$\sigma_t^2 = \frac{1}{|\zeta_t|} \sum_{s \in \zeta_t} (Z_s - \bar{Z}_t)^2,$$

Onde ζ_t é um conjunto de pixels incluindo o pixel t e os seus pixels vizinhos (conectividade 8), $|\zeta_t|$ é o número de pixels de ζ_t , Z_s é o nível de cinzento do pixel s e \bar{Z}_t representa a média dos níveis de cinzento de ζ_t .

A *LVIF* (*Local Variance Image Field*), ou seja, a imagem da variância local, é resultado da obtenção das variâncias locais para todos os pixels, quantificando-as e limitando-as a valores inteiros de 0 a 255 (níveis de cinzento), a qual representa a quantidade de energia da componente de alta frequência em cada pixel, ou seja, as zonas de maior transição de tons (i.é., contornos) (Figura 2). A *LVIF* $Y = \{Y_t = y_t, t \in \Omega\}$ para todos os pixels do conjunto Ω , é dada por:

$$y_t = \begin{cases} \langle \sigma_t^2 \rangle, & \text{se } \langle \sigma_t^2 \rangle \leq 255 \\ 255, & \text{se } \langle \sigma_t^2 \rangle > 255 \end{cases}$$

onde $\langle \sigma_t^2 \rangle$ assume o valor inteiro mais próximo de σ_t^2 . À medida que y_t aumenta são esperados na vizinhança transições de t níveis de cinzento da imagem original mais elevadas. Esses valores sugerem a aproximação a zonas de contorno do objecto. A atribuição de $y_t = 255$ quando se verifica $\langle \sigma_t^2 \rangle > 255$ pretende realçar essas zonas de contorno (Figura 2).



Figura 2 – Objecto de interesse (a) imagem *RGB* (b) imagem de energia

3.2.1.2 Extracção de características

Após a definição por parte do utilizador do objecto de interesse e sempre que este é detectado numa nova imagem, são extraídas características de cor e forma do objecto para utilização na fase de seguimento.

A característica de cor extraída é representada pela imagem *RGB* do objecto de interesse. Esta imagem *RGB* é actualizada sempre que o objecto é detectado numa nova imagem da sequência.

A característica de forma é extraída determinando a assinatura de contorno do objecto de interesse. A assinatura é obtida com base na imagem de energia já referida anteriormente, percorrendo-a e identificando as zonas com maior transição de tom escuro/claro (contorno). Esta abordagem será explicada com maior detalhe na próxima secção. Informação mais detalhada sobre assinaturas de contorno será abordada na secção 3.3.2 Detecção de forma. Ao longo de uma sequência de imagens o objecto poderá sofrer alterações na sua forma, nomeadamente, devido ao efeito de escala. Estas alterações são compensadas determinando a assinatura de contorno do objecto sempre que este seja detectado numa nova imagem.

A definição inicial, por parte do utilizador, do objecto de interesse deve ser efectuada numa zona com o fundo homogéneo. Esta limitação da abordagem deve-se à necessidade da obtenção correcta da forma do objecto.

3.2.2 Seguimento

O movimento de um objecto numa sequência de imagens é representado por três transformações base, as quais devem ser levadas em conta no desenvolvimento de um sistema que pretenda efectuar o seu seguimento: a translação, a mudança de escala e a rotação sofrida pelo objecto. O sistema desenvolvido não lida no entanto com deformações, transformações de perspectiva ou oclusões.

No sistema desenvolvido a cada imagem da sequência analisada está sempre associada:

- a localização $I_{xy} = (I_x, I_y)$ do objecto na imagem N;
- uma imagem *RGB* extraída a partir de I_{xy} com tamanho $d_{xy} = (d_x, d_y)$, centro em $C_{xy} = (I_x + d_x / 2, I_y + d_y / 2)$ e que representa o objecto na imagem N (Figura 3b);
- uma imagem *RGB* (janela de pesquisa) com centro em C_{xy} e tamanho $d1_{xy} = (d_x + \Delta x, d_y + \Delta y)$ sendo $\Delta x = \Delta y = const$ um parâmetro do sistema, na imagem N+1 (Figura 3c).

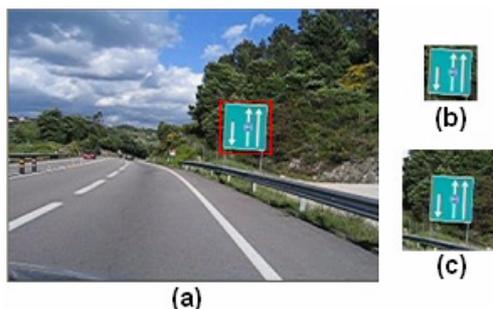


Figura 3 – Imagem (a) N original (b) objecto seleccionado na imagem N (c) área de pesquisa do objecto na imagem N+1

Sempre que uma nova imagem da sequência é analisada são actualizados os três parâmetros anteriores. De seguida será abordado o processo implementado tendo como tópicos as três transformações referidas: translação, escala e rotação. A metodologia utilizada pressupõe que as alterações sofridas pelo objecto, devidas às transformações referidas, sejam reduzidas entre imagens consecutivas de uma sequência.

3.2.2.1 Translação

O deslocamento de um objecto ao longo de uma sequência de imagens é traduzido por uma translação T_{xy} . No sistema implementado o deslocamento que o objecto sofre da imagem N da sequência para a imagem N+1 é calculado com base na comparação da informação de cor (característica interna) do objecto (imagem *RGB*) na imagem N da sequência com a informação de cor da janela de pesquisa (imagem *RGB*) na vizinhança do objecto na imagem N+1 da sequência, ou seja, é efectuada uma correspondência de modelos (*template matching*) entre estas duas imagens *RGB* (Figura 3) e devolvida a nova localização I_{xy} do objecto na imagem N+1. Tendo em conta que a correspondência de modelos envolve um número elevado de operações, optou-se por implementar uma pirâmide de resolução. A pirâmide é obtida da média dos pixels em coluna (duas a duas) seguida da média em linha (duas a duas) resultando numa imagem com $\frac{1}{4}$ da dimensão original. A pirâmide permite, apesar de diminuir a resolução da imagem, reduzir consideravelmente o tempo de processamento necessário, sem comprometer a sua eficiência.

Entre a posição do objecto na imagem N e a posição na imagem N+1 obtém-se a translação sofrida pelo objecto de uma imagem para a outra $T_{xy} = I_{xy} - I_{xy}$. É assim obtida uma aproximação grosseira à posição do objecto. Os pontos azuis da Figura 4 representam o centro do objecto sujeito a translação entre imagens consecutivas. A Figura 5 representa a compensação do efeito de translação.

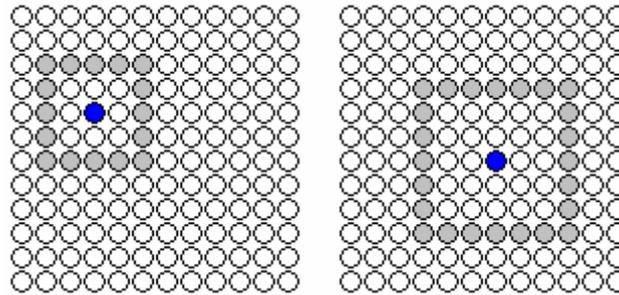


Figura 4 – Objecto sujeito a translação e alteração de escala entre imagens consecutivas

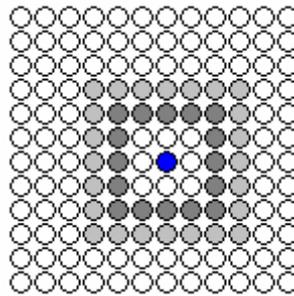


Figura 5 – Compensação do efeito de translação do objecto

No processo de seguimento o sistema deve ser também tolerante a alterações de escala (Figura 5). Seguidamente será abordada a forma como o sistema lida com este tipo de deformações.

3.2.2.2 Escala

A mudança de escala de um objecto, ao longo de uma sequência de imagens, é provocada por uma alteração da sua distância ao ponto de visualização. O sistema implementado quantifica essa alteração comparando a forma do objecto (característica externa) na imagem N com a sua forma na imagem N+1 (Figura 5), efectuando assim uma aproximação refinada à forma do objecto, tendo como base a filosofia dos contornos activos deformáveis (*snakes*).

A metodologia de refinamento baseia-se em 5 fases:

- 1ª fase – cálculo do contorno do objecto na imagem antiga;
- 2ª fase – cálculo do contorno do objecto na nova imagem;
- 3ª fase – convergência entre as duas curvas;
- 4ª fase – suavização da curva obtida;
- 5ª fase – actualização da informação do objecto;

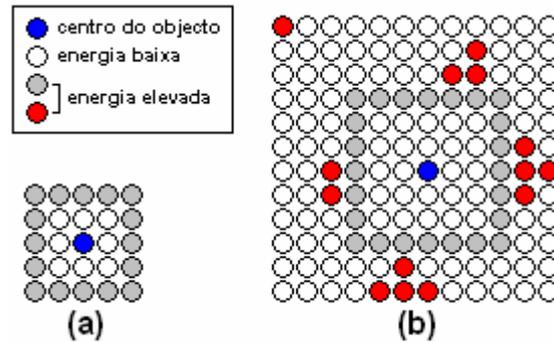


Figura 6 – Comparação do objecto na imagem N e N+1

Considere-se para ilustração dos passos do algoritmo:

- a Figura 6a) como sendo a imagem de energia obtida para o objecto na imagem N de uma sequência (os pontos a cinzento indicam a forma do objecto);
- a Figura 6b) como sendo a imagem de energia obtida para a janela de pesquisa do objecto na imagem N+1 da mesma sequência (os pontos a cinzento indicam a forma do objecto a determinar, pontos a vermelho indicam ruído);
- compensação prévia do efeito de translação (Figura 5).

1ª Fase: A forma do objecto é obtida do cálculo dos n pontos da sua assinatura de contorno, denominada $ACOI$, com um intervalo de θ° entre si.

$$n = \sqrt{d_x^2 + d_y^2} \text{ e } \theta = \frac{360^\circ}{n}, \text{ sendo } d_{xy} = (d_x, d_y) \text{ a dimensão da imagem de energia}$$

que representa o objecto. O número de pontos da assinatura de contorno e o ângulo entre eles varia com a dimensão da imagem que representa o objecto.

A cada ponto da assinatura de contorno está associada uma recta que une o centro do objecto C_{xy} (ponto a azul na Figura 7a)) ao ponto mais distante da imagem e que faz um ângulo θ com o eixo dos xx (Figura 7). Posteriormente é percorrida cada uma dessas rectas na direcção do centro C_{xy} , desde o ponto mais distante. É considerado ponto de contorno, em cada uma delas, aquele que possuir maior transição de tom escuro/claro.

2ª Fase: Esta fase é iniciada com o cálculo aproximado do centro $C_{1,xy}$ do objecto na imagem N+1 (ponto a azul na Figura 7b)). $C_{1,xy} = (I1_x + d_x / 2, I1_y + d_y / 2)$, sendo $I1_x, I1_y$ a nova localização do objecto, obtida da translação. Este cálculo pressupõe que a variação de tamanho do objecto entre imagens é pequena.

Procede-se de seguida à determinação grosseira (contorno partido + fundo) da assinatura de contorno do objecto na imagem N+1 (Figura 7b)). Essa assinatura,

denominada de *ACO2*, é obtida de forma idêntica à da fase anterior, para o mesmo número de pontos e ângulo entre eles. A aplicação do algoritmo é efectuada neste caso à imagem de energia resultante da janela de pesquisa do objecto com centro em $C1_{xy}$.

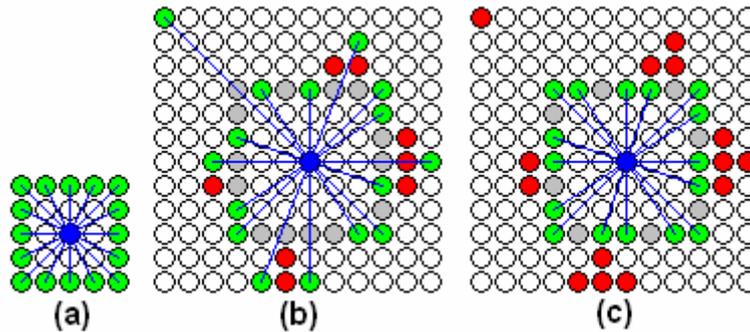


Figura 7 – Contorno do objecto (a) na imagem N (b) obtida na imagem N+1 (c) pretendida na imagem N+1

3ª Fase: As assinaturas de contorno *ACO1* e *ACO2*, são representadas pela distância de cada um dos seus pontos ao centro do objecto. A Figura 8 representa cada uma dessas distâncias, obtidas do exemplo da Figura 7, assim como as distâncias que se pretendem determinar, representadas pela Figura 7c). Na Figura 8 o deslocamento em y entre assinaturas resulta da alteração de escala.

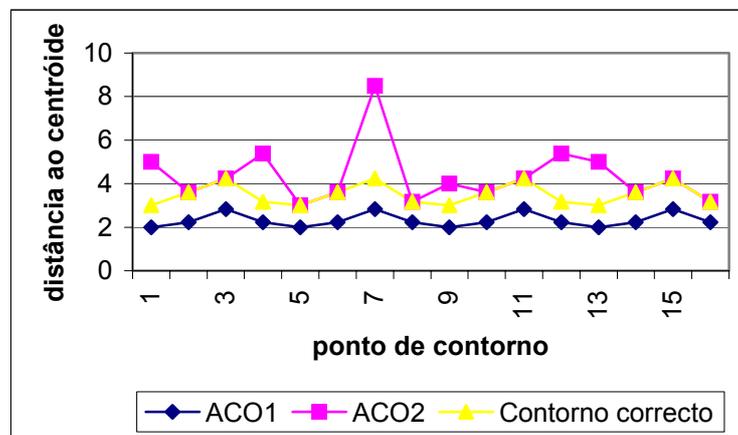


Figura 8 – Assinaturas de contorno do objecto

A comparação entre as curvas *ACO1* e *ACO2* é baseada nos valores absolutos das diferenças ponto a ponto. $\Delta d_i = |x_i - y_i|$, sendo i um ponto de contorno, x_i o valor da *ACO1* no ponto i e y_i o valor da *ACO2* no ponto i . A semelhança entre as duas curvas é obtida da combinação de duas medidas:

- Média aritmética dos valores de Δd_i .

$$\bar{\mu} = \frac{1}{n} \sum_{i=1}^n \Delta d_i, \quad n \text{ é o número de pontos da assinatura}$$

- Desvio padrão dos valores de Δd_i .

$$\sigma = \sqrt{\frac{1}{n} \sum_{i=1}^n (\Delta d_i - \bar{\mu})^2}$$

A média aritmética dos valores de Δd_i quantifica a variação média de escala (distância entre curvas) do objecto entre duas imagens consecutivas. O desvio padrão dos valores de Δd_i quantifica o rigor da aproximação entre as duas curvas. Em condições ideais este valor seria nulo.

A aproximação da curva *ACO2* à *ACO1* é efectuada combinando as duas medidas anteriores. Para cada Δd_i é verificado se cumpre a condição $\forall \Delta d_i, i \in [0, n], \Delta d_i < \bar{\mu} + \sigma$. Caso não seja cumprida esta condição, o valor correspondente da *ACO2* é igualado ao seu vizinho imediatamente anterior ($y_i = y_{i-1}$) que cumpre essa condição. Esta igualdade parte do pressuposto que num objecto real o valor y_{i-1} da assinatura de contorno é próximo do seu vizinho y_i . Este requisito é cumprido se for utilizado um número adequado de pontos de contorno. A combinação das duas medidas referidas é utilizada como forma de validar se a curva *ACO2* possui o mesmo comportamento que a curva *ACO1*.

A Figura 9 ilustra como exemplo a aproximação efectuada ao ponto assinalado da assinatura de contorno que não cumpre a condição imposta anteriormente.

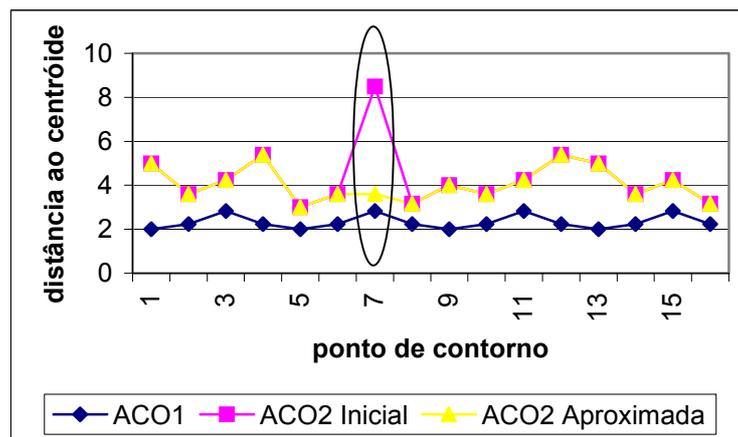


Figura 9 – Aproximação das assinaturas de contorno do objecto

Esta fase é executada recursivamente m vezes até que seja atingido o grau de aproximação definido entre as duas curvas. Esse grau de aproximação é obtido quando o

desvio padrão dos valores de Δd_i verificar a condição em que $\sigma_m = \sigma_{m-1}$, i.é., não existam pontos com $\Delta d_i \geq \bar{\mu} + \sigma$. Existe uma segunda condição, definida por $\sigma_m \leq k$ constante, que determina o valor abaixo do qual a aproximação é considerada suficiente. O valor para k foi obtido empiricamente ($k = 0,5$) testando casos reais.

Caso esta condição não seja cumprida, que não é o caso do exemplo apresentado, a janela de pesquisa é reduzida de um determinado valor ($\Delta x, \Delta y$). Este passo é utilizado essencialmente em casos em que o ruído de fundo nos limites da imagem da janela de pesquisa é muito elevado e tenha assim consequentemente uma grande influência na obtenção da *ACO2*. Seguidamente é obtida a nova *ACO2* e efectuada nova aproximação à *ACO1* repetindo-se recursivamente todo o processo a partir da 2ª Fase até que seja verificada a condição $\sigma_m \leq k$.

A Figura 10 mostra a fase final de aproximação da *ACO2* à *ACO1* após o cumprimento das condições $\sigma_m = \sigma_{m-1}$ e $\sigma_m \leq k$.

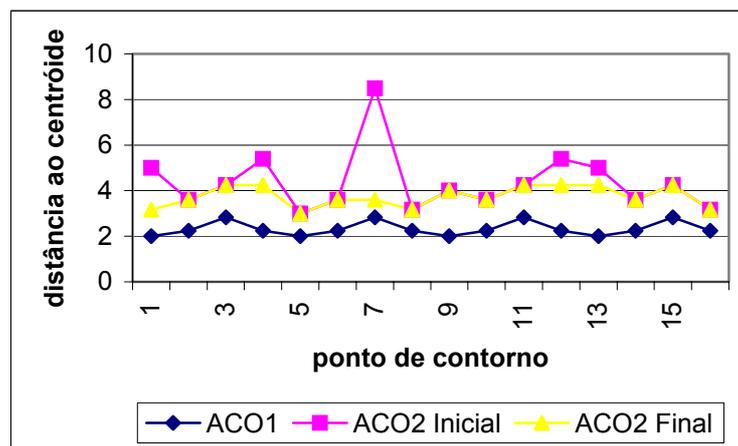


Figura 10 – Aproximação final das assinaturas de contorno do objecto

4ª Fase: A *ACO2* é submetida nesta fase a uma suavização com o objectivo de a tornar mais uniforme eliminando possíveis pontos isolados de ruído. O processo é iniciado de forma semelhante ao da 3ª Fase, com o cálculo dos valores absolutos das diferenças, neste caso, entre cada valor y_i da *ACO2* e o valor vizinho y_{i+1} , ou seja, $\Delta d_i = |y_i - y_{i+1}|$ em que y_i é o valor da *ACO2* no ponto i e y_{i+1} é o valor da *ACO2* no ponto $i+1$.

Os pontos da *ACO2* são comparados entre si utilizando o mesmo método da 3ª Fase, ou seja, a média aritmética dos valores de Δd_i ($\bar{\mu}$) e o desvio padrão dos valores de Δd_i (σ). Tal como nessa fase é analisado cada valor de Δd_i e verificado se cumpre a condição $\forall \Delta d_i, i \in [0, n], \Delta d_i < \bar{\mu} + \sigma$. Caso não seja cumprida esta condição, o valor y_i correspondente da *ACO2* é igualado à média dos seus vizinhos ($y_i = (y_{i-1} + y_{i+1})/2$). Este cálculo parte também do pressuposto que num objecto real quaisquer valores y_{i-1}, y_i e y_{i+1} pertencentes à assinatura de contorno possuem valores próximos.

Esta fase é executada recursivamente m vezes até que seja atingida a suavização requerida da $ACO2$ (Figura 11). Essa suavização verifica-se quando o desvio padrão dos valores de Δd_i for mínimo, ou seja, assim que se verifique a condição em que $\sigma_m = \sigma_{m-1}$.

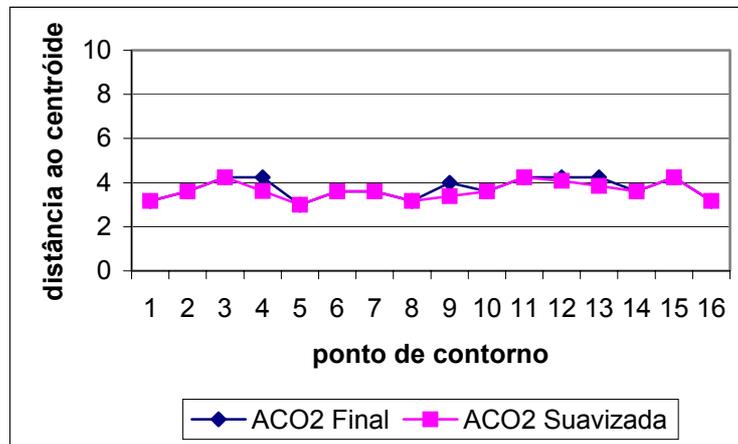


Figura 11 – Suavização da assinatura de contorno

A Figura 12 compara a $ACO1$ com a evolução da $ACO2$ da fase inicial para a fase final (suavização) assim como a $ACO2$ real na janela de pesquisa. De notar que a diferença resultante, ou seja, erro de aproximação, ponto a ponto, entre a $ACO2$ real e a $ACO2$ final obtida ($ACO2$ Suavizada) é inferior a 1 pixel.

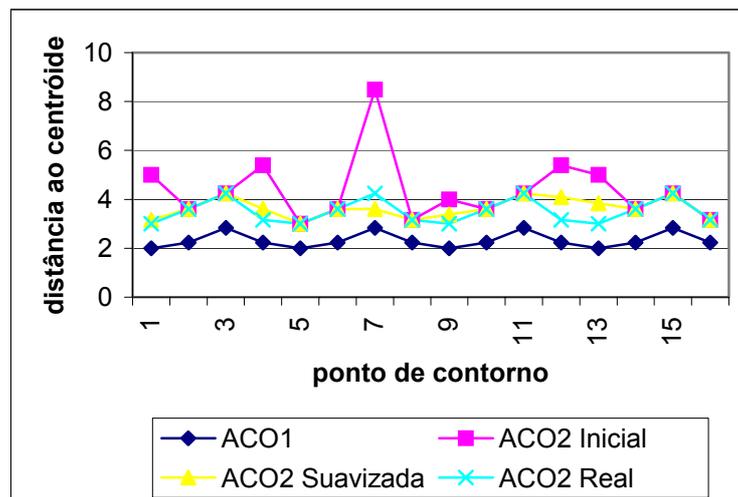


Figura 12 – Aproximação final suavizada da assinatura de contorno do objecto

5ª Fase: Finalmente nesta fase, e após determinação da $ACO2$ Suavizada que representa uma aproximação à $ACO2$ Real, é determinada a nova dimensão do objecto devida à alteração de escala. Este cálculo pressupõe que a variação de tamanho do objecto entre imagens é pequena. A cada ponto da $ACO2$ Suavizada corresponde uma localização

na imagem em coordenadas cartesianas, das quais se obtêm o menor e o maior valor em x (x_{\min} , x_{\max}), assim como o menor e o maior valor em y (y_{\min} , y_{\max}). Daqui se obtêm a dimensão (d_x, d_y) do objecto na nova imagem, ou seja, $d_x = (x_{\max} - x_{\min})$ e $d_y = (y_{\max} - y_{\min})$. O sistema está em condições de actualizar os três parâmetros utilizados no seguimento com os novos valores:

- a localização $I_{xy} = (I_x, I_y)$ do objecto na imagem N+1;
- a imagem *RGB* extraída a partir de I_{xy} com tamanho $d_{xy} = (d_x, d_y)$, centro em $C_{xy} = (I_x + d_x / 2, I_y + d_y / 2)$ e que representa o objecto na imagem N+1;
- a imagem *RGB* (janela de pesquisa) com centro em C_{xy} e tamanho $d_{l_{xy}} = (d_x + \Delta x, d_y + \Delta y)$ sendo $\Delta x = \Delta y = const$ um parâmetro do sistema, na imagem N+2.

Na informação obtida é incluída a actualização às características internas (imagem *RGB*) e externas (forma – assinatura de contorno) do objecto seguido para compensar as transformações por ele sofridas devido às alterações de escala. Nesta fase é verificado se algum dos pontos de contorno do objecto na imagem N+1 está localizado nos limites da imagem. Caso se verifique esta situação o processo de seguimento é imediatamente parado, uma vez que o objecto se encontra a sair da imagem. Caso contrário é de novo repetido todo o processo, iniciando-se com o cálculo da translação sofrida pelo objecto entre a imagem N+1 e N+2.

3.2.2.3 Rotação

O sistema pressupõe que a rotação sofrida pelo objecto a ser seguido é reduzida entre a imagem N e N+1 de uma sequência, não influenciando o processo de seguimento. No entanto essa rotação mesmo sendo reduzida poderia tornar-se crítica se fosse continuada ao longo da sequência, podendo resultar numa rotação final considerável. No sistema implementado essa situação não se verifica uma vez que, tendo em conta o que já foi dito na secção anterior, mais especificamente na 5ª Fase, as características internas e externas do objecto sobre o qual é efectuado o seguimento são actualizadas a cada nova imagem, sendo deste modo implicitamente efectuada a compensação à rotação sofrida pelo objecto entre as imagens N e N+1 da sequência.

3.3 Localização automática

A metodologia implementada no sistema, para a localização automática de objectos de interesse numa sequência de imagens, está dividida em quatro partes:

- **detecção e segmentação de regiões por cor:** São seleccionadas regiões com base na sua cor (azul e/ou vermelha), dimensão e localização na imagem;
- **detecção de formas:** São obtidas as regiões que possuam uma das formas pré-definidas no sistema - circular, triangular, quadrangular e triangular invertida;
- **seguimento:** É efectuado o seguimento dos objectos detectados e mantido o seu estado global actualizado;
- **reconhecimento:** O reconhecimento dos objectos é efectuado comparando-os com objectos pré-definidos numa base de dados, onde os objectos estão organizados segundo a sua forma.

A localização automática pressupõe por parte do sistema autonomia no processo de localização, ou seja, é dispensada a intervenção do utilizador. Dado que o sistema desenvolvido se destina à detecção de sinais de trânsito, procedeu-se à análise de uma vasta base de dados de sinais e chegou-se à conclusão que as características mais discriminantes entre os sinais analisados são a sua forma e a sua cor. Foram então introduzidas no sistema as características de forma e de cor mais comuns em sinais de trânsito (Figura 13), dividindo assim o problema da detecção em duas partes, ou seja, a detecção da sua cor, i.é., característica interna e da sua forma, i.é., característica externa.

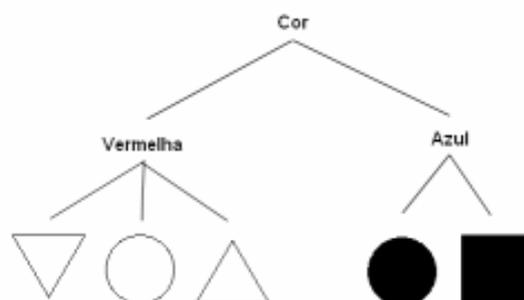


Figura 13 – Árvore de detecção

3.3.1 Detecção e segmentação de regiões por cor

A detecção das regiões de interesse é iniciada com a segmentação da imagem baseada na cor. As regiões que cumprirem determinados requisitos são normalizadas para posterior detecção por forma, a qual será abordada na secção 3.3.2. A detecção automática de um objecto numa imagem implica um conhecimento prévio por parte do sistema de um

número suficiente de características que o distingua não só do resto da imagem analisada como de outros possíveis objectos. Essa distinção é iniciada, nesta metodologia, com base na sua cor (azul e/ou vermelha), dimensão e localização na imagem.

3.3.1.1 Segmentação com base na cor

A identificação de zonas da imagem, de cor vermelha e/ou azul é efectuada através da segmentação da imagem em regiões. A característica comum entre os pixels de cada uma é a sua semelhança de cor. Esta decisão de considerar ou não determinadas zonas da imagem como sendo das duas cores desejadas é de extrema importância uma vez que influenciará decisivamente todo o processo subsequente. Por essa razão uma boa segmentação da imagem é essencial para o êxito do sistema. A análise de cor é essencial neste caso porque os sinais de trânsito são desenhados utilizando cores que pretendem chamar a atenção para a sua informação e que portanto se destacam da zona envolvente.

A imagem no espaço cor *RGB* (*R*= Vermelho, *G*= Verde, *B*= Azul) não é uma boa opção para servir de base à segmentação quando trata de imagens em ambiente aberto. Por um lado, devido a questões ambientais, não existe controlo sobre os níveis de iluminação do local. Por outro lado as cores dos sinais degradam-se devido ao seu envelhecimento. Um outro factor que contribui para esta degradação são as características intrínsecas das componentes *RGB*. Estas possuem uma correlação muito forte entre si, ou seja, a alteração de uma componente tende a influenciar as outras. Como consequência desta forte correlação, variações na intensidade de luz ambiente torna complexa a utilização do espaço cor *RGB*.

Apesar da existência de outros espaços de cor: *Nrgb*, *L*a*b*, etc. o espaço cor *HSI* (*Hue*, *Saturation*, *Intensity*) (Figura 14) é o utilizado na detecção/classificação de cor. Este espaço de cor representa diferentes partes de informação em cada componente [Arnou96], [Hibi96], [Picci96], além de ser mais próxima à forma como os seres humanos percebem as cores. Mais importante ainda, concentra melhor, relativamente a outros modelos, a variabilidade das condições de iluminação na sua componente *I*, permitindo a obtenção da componente *H* não correlacionada com a *I* [Arnou96].

O espaço cor *RGB* representa a quantidade de vermelho, verde e azul que existe em cada pixel de uma imagem. Não quantifica questões como a iluminação ou a pureza da cor de cada pixel que é precisamente o que o espaço cor *HSI* é capaz de fazer. Este espaço cor, também conhecido como *HSL*, está dividido em três componentes [Carro94]:

- *H (Hue)*: representa a característica de cor. Tem a propriedade de ser relativamente imune a sombras causadas pela fonte de luz. Esta característica tem uma representação angular (sendo de 0 a 255 depois da normalização);
- *S (Saturation)*: mede o grau de pureza da componente *H*;
- *I (Intensity)*: representa o nível de cinzento médio. Intensidade de cor do pixel ou a sua luminosidade.

Apesar de existirem diversas variantes da transformação *RGB* para *HSI* [Carro94], [Welch91], [Ledle90], a posição de um qualquer ponto no espaço *HSI* é modulada em termos de coordenadas cilíndricas, isto é, o triplo (h, s, i), representado na Figura 14 [Singh03], corresponde às coordenadas cilíndricas (θ, r, z). A componente H é representada pelo ângulo, variando entre 0° e 360° . A componente S corresponde ao raio, variando entre 0 e 1. I varia ao longo do eixo dos zz sendo 0 preto e 1 branco.

Quando $S=0$, a cor tem um valor de cinzento com intensidade 1. Quando $S=1$, a cor encontra-se na fronteira da base do cone superior. Quanto maior for o valor de saturação (S), mais afastada está a cor do branco/cinzento/preto (dependendo da intensidade (I)). O ajuste da componente H varia a cor de vermelho aos 0° , passando por verde aos 120° , azul aos 240° , e de novo ao vermelho aos 360° . Quando $I=0$, a cor é preto e portanto a componente H é indefinida. Quando $S=0$, a cor é um valor de cinzento e também neste caso H é indefinido. Ajustando a componente I , pode tornar-se uma cor mais clara ou mais escura. Mantendo-se $S=1$ e ajustando I , são criadas sombras dessa cor [Singh03].

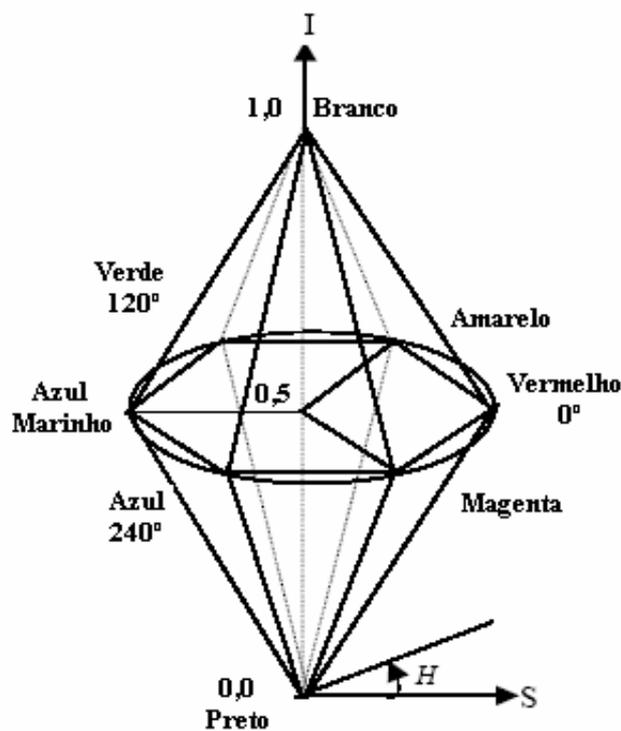


Figura 14 – Espaço cor HSI num modelo duplo cone

Na Tabela 1 é apresentado o algoritmo de conversão *RGB-HSI* utilizado no sistema desenvolvido o qual retorna para as três componentes valores entre 0,0-1,0 os quais são posteriormente normalizados para valores entre 0-255. Este algoritmo foi obtido de um programa comercial de conversão entre espaços de cor [Logic04].

```

var_R = ( R / 255 )           // valores RGB = de 0 a 255
var_G = ( G / 255 )
var_B = ( B / 255 )

var_Min = min( var_R, var_G, var_B ) //valor Min. de RGB
var_Max = max( var_R, var_G, var_B ) //valor Máx. de RGB
del_Max = var_Max - var_Min         //valor delta de RGB

I = ( var_Max + var_Min ) / 2

Se ( del_Max == 0 )               //Pixel escala de cinzentos...
{
    H = 0
    S = 0
}
Senão                               //Pixel com cor...
{
    Se ( I < 0.5 ) S = del_Max / ( var_Max + var_Min )
    Senão          S = del_Max / ( 2 - var_Max - var_Min )

    del_R = ( ( ( var_Max - var_R ) / 6 ) + ( del_Max / 2 ) ) / del_Max
    del_G = ( ( ( var_Max - var_G ) / 6 ) + ( del_Max / 2 ) ) / del_Max
    del_B = ( ( ( var_Max - var_B ) / 6 ) + ( del_Max / 2 ) ) / del_Max

    Se      ( var_R == var_Max ) H = del_B - del_G
    Senão Se ( var_G == var_Max ) H = ( 1 / 3 ) + del_R - del_B
    Senão Se ( var_B == var_Max ) H = ( 2 / 3 ) + del_G - del_R

    Se ( H < 0 ) ; H += 1
    Se ( H > 1 ) ; H -= 1
}

```

Tabela 1 – Algoritmo de conversão RGB para HSI

Na Figura 15 é apresentada uma imagem no espaço cor RGB e as suas respectivas componentes no espaço cor HSI. A componente de intensidade de cor é semelhante à que se obteria da conversão da imagem RGB em níveis de cinzento (luminosidade).

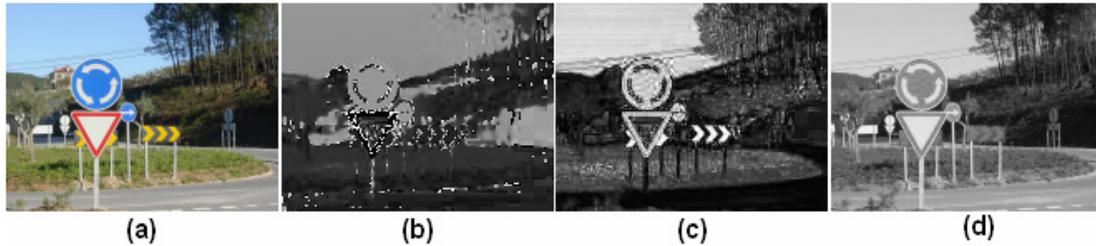


Figura 15 – Conversão (a) *RGB* para as componentes (b)*H*, (c)*S* e (d)*I*

Os valores apresentados na Tabela 2 foram obtidos com base nos resultados da análise de cor de vários sinais de trânsito em imagens reais com diferentes tipos de iluminação originados por diferentes estados de tempo, segmentados por um operador humano e dos quais foram guardados os valores de *Hue* e Saturação [Escal04]. De lá podemos verificar que são considerados pixels vermelhos todos aqueles que possuem um *Hue* na gama 0-15 ou 200-255 e que possuem uma saturação igual ou superior a 32. De igual modo são considerados pixels azuis todos aqueles que têm um valor de *Hue* entre 135-200 e saturação mínima de 84.

	Hue (H)		Saturação (S)
Sinais Vermelhos	[0 ... 15]	[200 ... 255]	>= 32
Sinais Azuis	[135 ... 200]		>= 84

Tabela 2 – Valores de *Hue* e Saturação para a segmentação de cor

As restrições impostas à gama de valores ilustradas na Tabela 2 são suficientes quando a análise de cor é efectuada apenas nos sinais, ou seja, estes se encontram isolados do fundo. Surge no entanto alguma dificuldade na segmentação quando zonas da imagem externas aos próprios sinais se encontram ligadas a estes podendo provocar uma segmentação incorrecta: claras (ex: céu) ou escuras (ex: estrada), que é o caso mais recorrente em imagens reais. Tal como referido anteriormente, o valor de saturação quantifica o afastamento da cor de cada pixel ao branco/cinza/preto (dependendo da intensidade (*I*)), mas é também ao mesmo tempo muito sensível ao ruído para valores baixos (preto) exemplificado na Figura 16. Por esse facto a componente de saturação é substituída, na segmentação, pela componente cromática da imagem *RGB*, ou seja, a diferença ao equivalente nível de cinza, mantendo-se válidos para esta os valores determinados para a componente de saturação (Tabela 3). A componente cromática possui valores baixos para cores *RGB* próximas aos níveis de cinza, aumentando esses valores para cores *RGB* mais distantes dos níveis de cinza. Esta componente pode considerar-se, por esse motivo, como uma componente de penalização.



Figura 16 – Componente (a) saturação (b) cromática

	Hue (H)		Cromatismo
Sinais Vermelhos	[0 ... 15]	[200 ... 255]	≥ 32
Sinais Azuis	[135 ... 200]		≥ 84

Tabela 3 – Valores de *Hue* e Cromatismo utilizados na segmentação de cor

Assim sendo a componente cromática da Figura 16 é calculada da seguinte forma [Escal04]:

1. É obtida uma imagem de níveis de cinzento das componentes RGB da imagem original, vermelho $R(x, y)$, verde $G(x, y)$ e azul $B(x, y)$, seguindo a fórmula utilizada no cálculo do sinal de luminância $L(x, y)$ de uma câmara RGB:

$$L(x, y) = 0,3 * R(x, y) + 0,6 * G(x, y) + 0,1 * B(x, y)$$

2. Os erros normalizados para cada componente são:

$$\ell_R = \frac{|L(x, y) - R(x, y)|}{L(x, y)}$$

$$\ell_G = \frac{|L(x, y) - G(x, y)|}{L(x, y)}$$

$$\ell_B = \frac{|L(x, y) - B(x, y)|}{L(x, y)}$$

As componentes calculadas são designadas de “erros” porque quantificam a distância da cor analisada à cor cinzenta. Os erros serão próximos de zero se a cor for cinzenta.

3. A componente cromática $A(x, y)$ representa o erro máximo:

$$A(x, y) = \max \{l_R, l_G, l_B\}$$

Finalmente, após a determinação das componentes referidas, é efectuada a segmentação da imagem RGB , em que o sistema considera como pixels de interesse todos aqueles que satisfazem as condições impostas na Tabela 3, atribuindo-lhes o valor máximo possível de nível de cinzento (255) e rejeitando todos os outros, atribuindo-lhes o valor mínimo possível (0). A Figura 17 compara o resultado da segmentação utilizando as componentes *Hue* e Saturação com o resultado da segmentação aplicada no sistema que utiliza as componentes *Hue* e Cromatismo.



Figura 17 – Segmentação (a) *Hue*/Saturação (b) *Hue*/Cromatismo

3.3.1.2 Etiquetagem

Após a fase de segmentação da imagem é efectuada a contagem das regiões daí resultantes. Para isso é efectuada a atribuição de um valor único aos pixels que constituem cada uma delas, à qual se dá o nome de etiquetagem. A Figura 18 apresenta o resultado de uma etiquetagem na qual foi atribuído um valor distinto a cada uma das regiões.

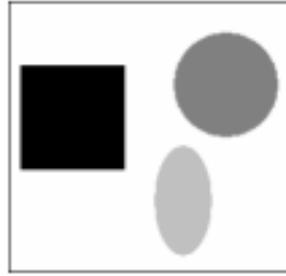


Figura 18 – Regiões etiquetadas de uma imagem

A imagem é percorrida sendo atribuído ao primeiro pixel encontrado um determinado valor (etiqueta) que será o mesmo para o pixel seguinte se este se encontrar numa vizinhança de 4 pixels relativamente ao anterior (Figura 19), senão é-lhe atribuído um novo valor e assim sucessivamente. Tomando como exemplo a Figura 19, o pixel central (azul) só pertencerá à região em causa se pelo menos um dos quatro pixels verdes fizer parte da mesma (vizinhança 4).

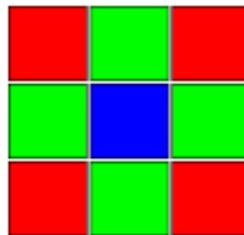


Figura 19 – Etiquetagem com vizinhança 4 de um pixel

3.3.1.3 Filtragem

A filtragem está dividida em quatro fases e consiste na separação e/ou eliminação de regiões, obtidas no processo de segmentação de cor, que não possuem determinados requisitos, tal como uma área mínima. A distinção entre regiões é um passo que permite posteriormente uma análise mais pormenorizada a cada uma delas e determinar a sua validação ou não como região de interesse. Parte das regiões etiquetadas não terão possibilidade de serem validadas. Para evitar sobrecarga de processamento, na análise dessas regiões, é realizada previamente uma filtragem dividida em 4 fases (Figura 20).

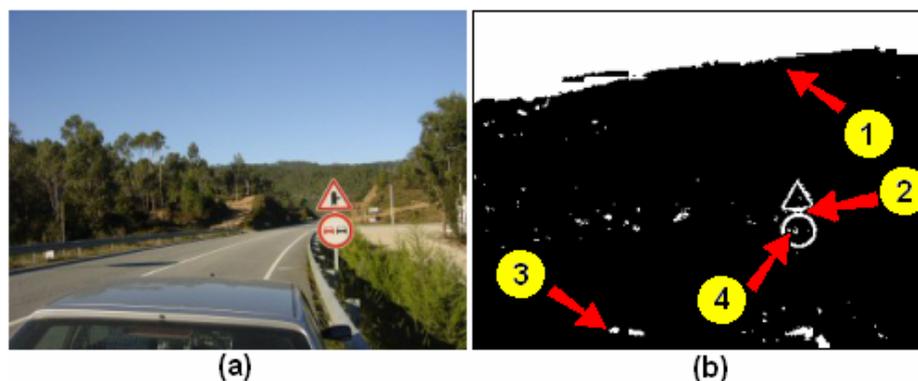


Figura 20 – Exemplo de uma imagem (a) original e (b) ilustração da filtragem

1ª Fase: São eliminadas as regiões que se encontram encostadas aos cantos da imagem:

- representam apenas ruído;
- são objectos que estão a entrar na área de visualização e que portanto serão com certeza detectados mais à frente na sequência;
- são objectos que estão a sair da área de visualização e nesse caso já terão sido detectados em imagens anteriores da sequência.

2ª Fase: Dado que o sistema lida com sinais de trânsito e tendo-se verificado por vezes a existência nas vias rodoviárias a presença de dois sinais num mesmo suporte, para evitar possíveis uniões de regiões relativas a esses sinais, procede-se à sua divisão desde que a relação *Altura/Largura* seja superior a um determinado valor.

3ª Fase: São eliminadas as regiões que não possuam :

- Área mínima;
- Altura e largura mínimas;

Os valores da 3ª fase foram obtidos empiricamente através de experimentação e análise visual a imagens com sinais de trânsito. Com esta imposição pretende-se eliminar regiões que não possuem informação mínima de cor (característica interna) e forma (característica externa) necessária no passo seguinte da detecção.

4ª Fase: Nesta última fase de filtragem são eliminadas as regiões que estão contidas totalmente dentro de outras, tentando-se evitar desta forma referências duplicadas da mesma região.

A Figura 21 apresenta o resultado da filtragem aplicada à imagem da Figura 20.

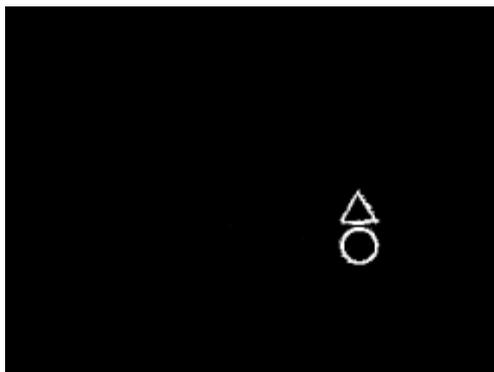


Figura 21 – Resultado da filtragem

3.3.1.4 Normalização

Para facilitar a comparação, reconhecimento e classificação de formas, as regiões são redimensionadas (normalizadas) em imagens de uma determinada altura e largura pré-definidas (obtidas experimentalmente 64x64 pixels), mantendo obviamente a sua relação altura/largura. No canto superior esquerdo da Figura 22 estão representadas as regiões normalizadas obtidas da filtragem.



Figura 22 – Regiões normalizadas

3.3.2 Detecção de formas

O sistema desenvolvido é capaz de detectar e reconhecer 4 formas pré-definidas de sinais de trânsito: circular, triangular, quadrangular e triangular invertida (Figura 13). A validação das formas de interesse é obtida com um classificador seleccionado após testes experimentais que será apresentado na secção 4.2.2.1. A informação encontra-se introduzida no sistema na forma de assinatura de contorno (Figura 23). Esta representação é obtida através de uma função 1-D que indica a distância do *centróide* ao contorno em função do ângulo $r(\theta)$ [Gonza02].

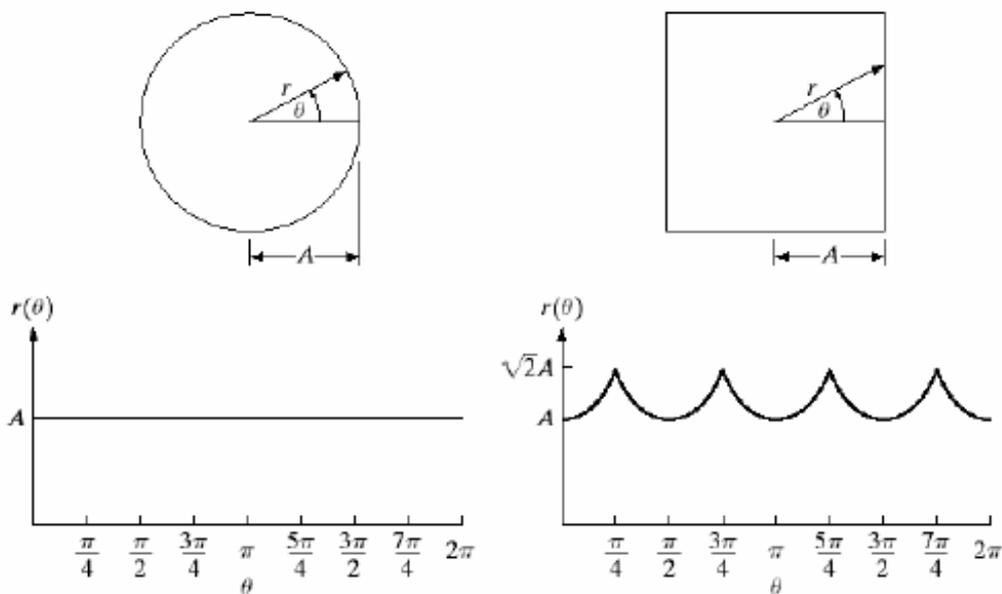


Figura 23 – Exemplo de assinaturas resultantes de contornos

O resultado obtido deste tipo de abordagem é independente da posição do objecto de interesse, mas depende da rotação e escala (Figura 23). O efeito da rotação num objecto traduz-se no deslocamento segundo o eixo dos xx da sua assinatura de contorno. No caso da transformação de escala traduz-se no deslocamento da mesma segundo o eixo dos yy .

A compensação da rotação implicaria por parte do sistema a capacidade de quantificar a rotação sofrida, aumentando a carga computacional. Uma vez que os objectos analisados pelo sistema são rígidos e normalmente colocados a aproximadamente 0° em relação ao eixo dos yy (sinais de trânsito) o sistema pressupõe que a rotação, mesmo que exista, seja reduzida e portanto não afecta a análise da forma utilizando esta metodologia, não sendo por esse motivo necessária qualquer tipo de compensação por esse facto. No caso da escala é utilizada a normalização dos valores de $r(\theta)$, ou seja, redimensionando as regiões detectadas a um determinado valor de altura e largura já referidos anteriormente.

As assinaturas de contorno constituem, com as devidas alterações já mencionadas, uma boa opção na detecção de formas pré-definidas como as utilizadas neste sistema, uma vez que exigem uma carga computacional reduzida no processo de detecção, sendo este facto determinante em casos em que existe uma grande quantidade de regiões a analisar e em que são exigidos requisitos temporais ao sistema [Peixo02]. Nos casos em que as formas a detectar têm concavidades as assinaturas de contorno não são adequadas. As assinaturas de contorno requerem para cada ângulo θ apenas um valor de $r(\theta)$ (Figura 23). Não é o caso neste sistema, que apenas analisa formas sem concavidades (sinais de trânsito).

As assinaturas de contorno utilizadas foram extraídas de sinais com formas ideais, digitalizados de documentação oficial de uma escola de condução, após a devida normalização. Os modelos dos sinais possuem as seguintes características:

- distância fixa;
- perpendicular ao eixo;
- localizado no centro da imagem.

De seguida serão apresentadas as assinaturas de contorno correspondentes às formas pré-definidas que o sistema utiliza como modelo e é capaz de detectar, extraídas dos sinais normalizados. Não são utilizados no sistema os 360 valores possíveis de obter da assinatura de contorno (Figura 23) mas sim uma amostragem desses pontos. Essa opção permite uma diminuição do número de pontos a analisar mantendo ao mesmo tempo informação de forma suficiente para a análise pretendida. Após vários testes de extracção de assinaturas de contornos chegou-se à conclusão que utilizando a regra seguinte para o número de amostragens é suficiente para a detecção dos objectos por forma:

$$n = \sqrt{x^2 + y^2} \quad \text{e} \quad \Delta\theta = \frac{360^\circ}{n}$$

Em que n representa o número de amostragens, $\Delta\theta$ o ângulo de intervalo entre amostragens e x, y as dimensões da imagem.

Nos exemplos apresentados de seguida, para cada uma das formas pré-definidas, foram utilizadas imagens, após normalização, de dimensão 64 x 64 pixels (Figura 24), que corresponde a $n=90$ pontos de assinatura de contorno (Figura 25) e a um intervalo de amostragem $\Delta\theta = 4^\circ$. Os valores obtidos foram arredondados para o inteiro mais próximo.

A assinatura de contorno (distância ao *centróide*) de cada uma das formas apresentadas de seguida, ao contrário da Figura 23, não são constantes uma vez que foram obtidas directamente de formas digitalizadas e que portanto contêm ruído próprio da representação digital (discreta) do sinal. Essa situação no entanto não representa qualquer tipo de problema ao sistema porque as assinaturas de contorno das formas pré-definidas são suficientemente distintas entre si, como se verá posteriormente.

3.3.2.1 Circular

A assinatura de contorno da forma circular (Figura 24) é semelhante ao da Figura 23 uma vez que se trata da mesma forma. Possui também uma distância ao centróide aproximadamente constante a qual caracteriza as formas circulares.

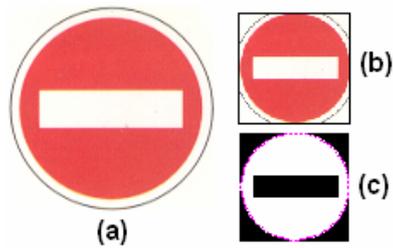


Figura 24 – Detecção do contorno circular

(a) sinal modelo, (b) sinal normalizado e (c) sinal com contorno

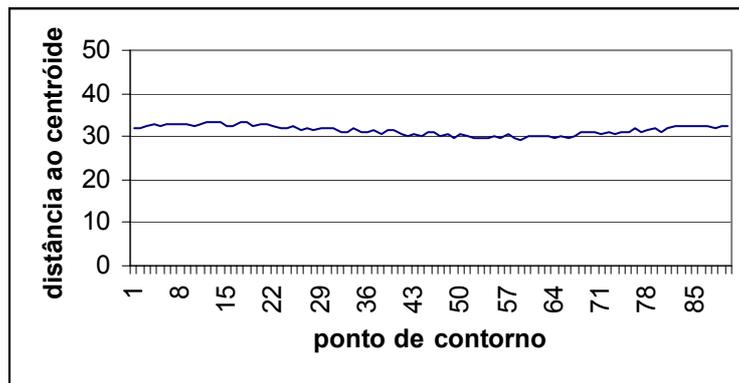


Figura 25 – Assinatura de contorno circular

3.3.2.2 Triangular

A assinatura de contorno triangular tem como principais características, facilmente observáveis da Figura 27:

- três pontos com distância máxima ao *centróide*, que se destacam de todos os outros, que correspondem aos cantos do triângulo (Figura 26);
- três pontos de distância mínima que correspondem aos pontos médios entre os cantos do triângulo.

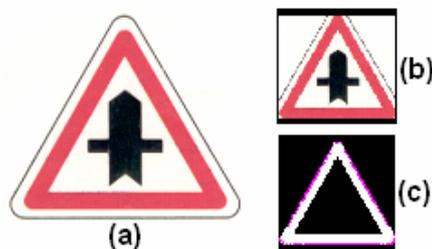


Figura 26 – Detecção do contorno triangular

sinal modelo, (b) sinal normalizado e (c) sinal com contorno

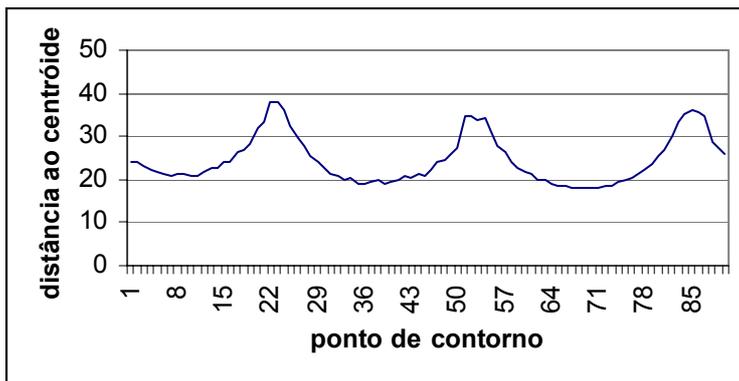


Figura 27 – Assinatura de contorno triangular

3.3.2.3 Quadrangular

A assinatura de contorno da forma quadrangular (Figura 29) tem um aspecto semelhante ao da forma triangular com a diferença que neste caso existem:

- quatro pontos bem destacados na assinatura, que representam os quatro cantos da forma (Figura 28);
- quatro pontos, que podem ser considerados os mínimos da assinatura, que representam os pontos médios entre o cantos da forma quadrangular.

Esta assinatura de contorno é semelhante ao exemplo, para um caso ideal, da Figura 23.

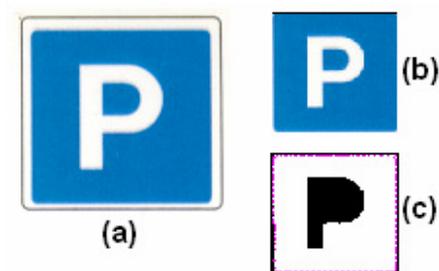


Figura 28 – Detecção do contorno quadrangular
sinal modelo, (b) sinal normalizado e (c) sinal com contorno

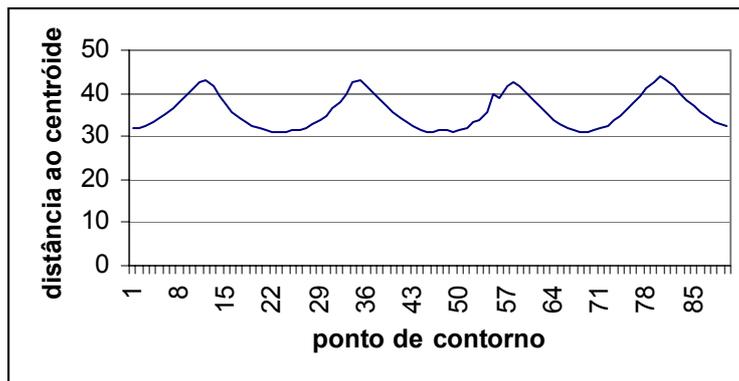


Figura 29 – Assinatura de contorno quadrangular

3.3.2.4 Triangular Invertida

A forma triangular invertida (Figura 30) é semelhante à forma triangular a menos de uma rotação de 180° , com origem no centro da forma. Em termos de assinatura de contorno este facto reflecte-se num desfasamento de 180° (Figura 31), i.é. 45 pontos, em relação à da forma triangular (Figura 27).

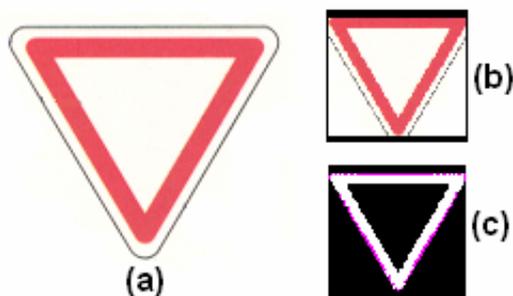
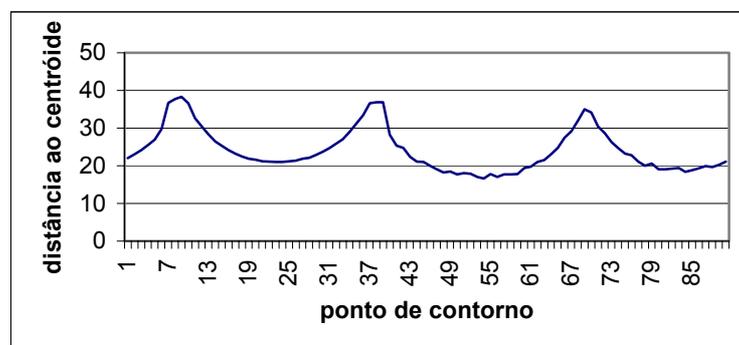
Figura 30 – Detecção do contorno triangular invertido
sinal modelo, (b) sinal normalizado e (c) sinal com contorno

Figura 31 – Assinatura de contorno triangular invertido

3.3.2.5 Classificador

Para cada uma das regiões resultantes do processo de detecção de cor é determinada a assinatura de contorno correspondente, do modo referido anteriormente. De seguida esta é comparada com as assinaturas das formas pré-definidas e decidido se a região analisada é válida (i.é. sinal com uma das formas pretendidas) ou não (i.é. sinal com outro tipo de forma/ruído). Tal é efectuado através de um procedimento que percorre a assinatura de cada uma das regiões e calcula o desvio médio (σ) da sua assinatura (forma) a cada uma das assinaturas pré-definidas.

$$\sigma = \frac{1}{n} \sum_{i=1}^n |x_i - y_i - \mu|$$

Sendo,

$$\mu = \frac{1}{n} \sum_{i=1}^n |x_i - y_i|: \text{ a média aritmética dos valores absolutos das diferenças entre}$$

assinaturas ponto a ponto;

n : número de pontos da assinatura;

x_i : o valor da assinatura da região em análise no ponto i ;

y_i : o valor da assinatura de uma das formas pré-definidas no ponto i ;

Sendo quatro o número de formas pré-definidas, obtêm-se quatro desvios médios para cada região, ou seja, $\sigma_{região} = \{\sigma_{circular}, \sigma_{triangular}, \sigma_{quadrangular}, \sigma_{triang_inv}\}$. Destes é obtido o menor valor, $\sigma_{min} = \min(\sigma_{região})$ e associada à região a forma correspondente. Finalmente é verificado se $\sigma_{min} \leq k$ constante, em que o valor k , obtido experimentalmente, representa o desvio médio máximo admitido pelo sistema para que a região analisada seja considerada suficientemente próxima da forma associada.

3.3.2.6 Exemplo da classificação de formas

Para finalizar a parte relativa à detecção é apresentado um exemplo prático da aplicação da metodologia implementada no sistema. À imagem original é aplicada a segmentação, com base na detecção de cor, e determinada a assinatura de contorno de cada uma das regiões candidatas a sinal (Figura 32).

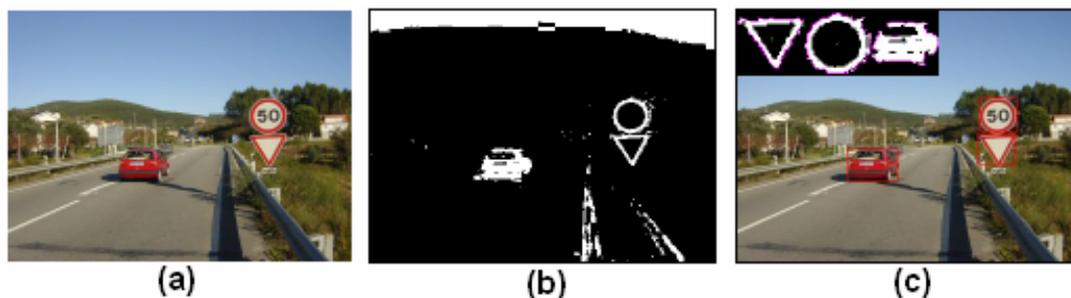


Figura 32 – Exemplo de detecção cor e forma
 (a) imagem original, (b) segmentação e (c) detecção

Neste exemplo, são detectados apenas 3 objectos (Figura 32c)) que cumprem todos os requisitos impostos pelo sistema na fase de detecção por cor. Dois deles são efectivamente sinais de trânsito, sendo o terceiro um automóvel. Seguidamente estes três objectos são submetidos à detecção por forma.

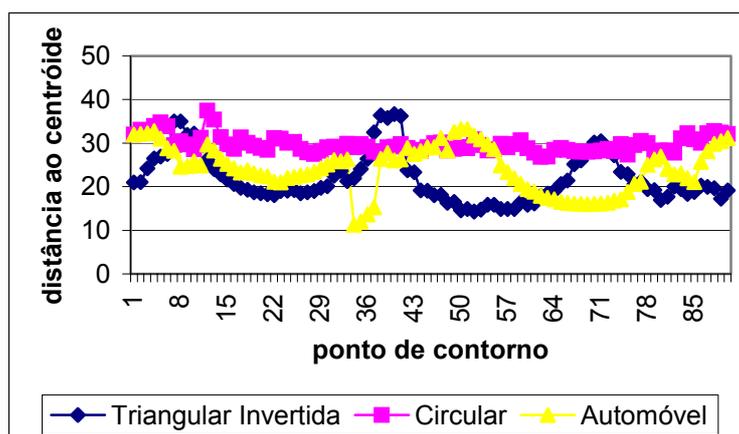


Figura 33 – Assinaturas de contorno das regiões detectadas

As assinaturas de contorno resultantes do processo anterior (Figura 33) são comparadas com as assinaturas de contorno das formas pré-definidas (Figura 34) existentes no sistema, utilizando o classificador de forma, definindo assim, se a forma da região analisada é ou não uma das pré-definidas e qual.

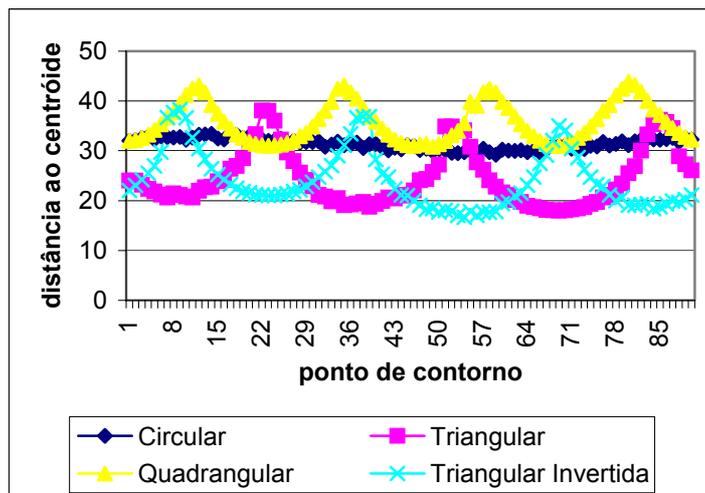


Figura 34 – Assinaturas de contorno das formas pré-definidas

Aplicando os passos e a notação da secção anterior a este exemplo, obtêm-se os valores da Tabela 4. Para cada região analisada é quantificada a proximidade da sua forma a cada uma das formas pré-definidas, utilizando o classificador de forma. De seguida é identificado para cada região o valor mínimo de entre os quatro calculados, ou seja, a forma mais próxima. Finalmente é verificado se a região é suficientemente próxima, menor ou igual que um valor k constante, da forma pré-definida que lhe foi associada. O valor $k \leq 2,43$ foi obtido experimentalmente para o classificador, o qual será analisado na secção 4.2.2.1 (Tabela 12).

Neste caso concreto, a região1, que representa o sinal triangular invertido, foi classificada correctamente, sendo associada à forma pré-definida triangular invertida. A região2, que representa o sinal circular, foi também classificada correctamente, sendo associada à forma circular. Finalmente, a região3, que representa o automóvel, foi rejeitada pelo sistema, uma vez que não é suficientemente próxima a nenhuma das formas pré-definidas.

	$\sigma_{circular}$	$\sigma_{triangular}$	$\sigma_{quadrangular}$	σ_{triang_inv}
$\sigma_{região1}$	3,65	4,76	5,50	1,49
$\sigma_{região2}$	1,02	2,86	3,68	2,74
$\sigma_{região3}$	4,13	2,97	5,62	4,89

Tabela 4 – Medidas obtidas do classificador de forma

3.3.3 Seguimento

O seguimento é efectuado com base na detecção individual dos objectos em cada imagem, exigindo por esse motivo, a sua localização em todas as imagens.

Terminada a fase de detecção dos objectos de interesse é iniciada a fase do seu seguimento. Para que esta fase seja realizada com sucesso é necessário ter em conta alguns aspectos importantes. O sistema deverá nesta fase ser capaz de identificar:

- o mesmo objecto de interesse em imagens consecutivas (Figura 35 objecto 2);
- objectos de interesse que “entram” na área de visualização (Figura 35 objecto 3);
- objectos de interesse que “saem” da área de visualização (Figura 35 objecto 1).

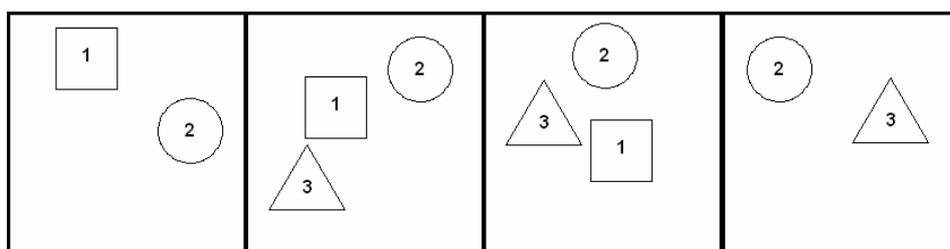


Figura 35 – Seguimento de objectos numa sequência de imagens

Tendo em vista o cumprimento destes requisitos o sistema implementado possui um “estado global actual” que consiste numa estrutura onde estão registados todos os sinais de trânsito detectados numa dada imagem n da sequência, tendo como parâmetros a sua localização (x, y) e a sua forma, sendo de seguida os dados obtidos comparados com os registados na imagem $n-1$ e actualizados.

1º Requisito: Um determinado sinal é validado pelo sistema como sendo o mesmo da imagem $n-1$ se a diferença entre as suas localizações for inferior a um determinado valor e forem da mesma forma. É efectuada uma actualização à localização (x, y) do sinal. Esta correspondência permite fazer o seguimento de cada um dos sinais ao longo da sequência de imagens.

2º Requisito: Sempre que exista um sinal registado na estrutura do sistema, respeitante à imagem n , que não tem correspondência na imagem $n-1$, é criado um novo registo no sistema para este sinal. Isto é resultado da entrada de um novo sinal na sequência de imagens.

3º Requisito: No caso de se verificar que um determinado sinal na imagem $n-1$ não possui correspondência em nenhum sinal na imagem n o registo desse sinal é retirado da

estrutura do sistema. Neste caso é resultado da saída de um sinal anteriormente detectado pelo sistema e que já não se encontra na imagem actual.

Tendo em conta o que foi dito pode concluir-se que o seguimento não é mais do que a união de duas partes. Uma primeira parte de detecção, em que se procuram identificar os objectos de interesse em cada uma das imagens da sequência e uma outra parte em que a esses objectos é atribuída uma identidade própria com o objectivo de os identificar como sendo os mesmos objectos ou não (distinguir), nas imagens seguintes da sequência.

3.3.4 Reconhecimento

O sistema desenvolvido utiliza para o reconhecimento o método da correlação normalizada de níveis de cinzento. O processo de reconhecimento de um dado objecto exige, tal como na fase da sua detecção automática, conhecimento prévio por parte do sistema das características que o distinguem dos demais. Existem no entanto dificuldades recorrentes que surgem no processo de reconhecimento que devem ser consideradas, acrescidas de outras quando esse reconhecimento é efectuado em sinais de trânsito:

- elevado número de sinais;
- ideogramas ligeiramente/consideravelmente diferentes do mesmo sinal [Pacli99];
- oclusão total/parcial de sinais;
- variação de iluminação;
- diferenças de escala.

Este último problema é facilmente ultrapassado uma vez que após a detecção o sinal é normalizado para uma determinada dimensão, utilizando uma sub-amostragem/replicação dos pixels. Uma solução frequentemente utilizada para resolver problemas de oclusão é a utilização de redes neuronais treinadas com sinais parcialmente visíveis [LuoR92], justificando-se por isso uma análise mais profunda a este tipo de solução. Esta abordagem contém alguns problemas na fase de treino nomeadamente de iluminação uma vez que a luz influencia a distribuição de níveis de cinzento no interior do sinal, ainda que seja utilizada uma binarização prévia do sinal o problema não é resolvido porque o limiar de binarização varia muito de uma imagem para outra. Os dois primeiros problemas apresentados são os mais difíceis de resolver. Em vários artigos os modelos de sinal utilizados são um número muito restrito de todos os sinais de trânsito [Aoyag96]. Se o sistema for utilizado num levantamento da sinalização rodoviária ou mesmo utilizado em vários países a base de dados de treino da rede neuronal seria enorme.

Tendo em conta todas estas considerações o sistema desenvolvido utiliza para o reconhecimento o método da correlação normalizada de níveis de cinzento, o qual tem sido frequentemente utilizado ao longo dos anos [Escal04][Miura02], [Betke95].

$$C = \frac{\sum_{x,y} (I - \bar{I})(T - \bar{T})}{\sqrt{\sum_{x,y} (I - \bar{I})^2 \sum_{x,y} (T - \bar{T})^2}}, \text{ onde } \bar{I} = \sum \frac{I}{N} \text{ e } \bar{T} = \sum \frac{T}{N}$$

T e I são a informação do sinal modelo e da região em análise, respectivamente, e N é o número de pixels das imagens.

A base de dados de sinais está dividida em classes segundo a sua forma, uma vez que os sinais detectados se encontram também assim divididos. Existem portanto quatro classes de sinais: circulares, triangulares, quadrangulares e triangulares invertidos já referidos anteriormente. Esta divisão tem por objectivo diminuir o número de candidatos da base de dados a utilizar na correlação, diminuindo o tempo de reconhecimento.

Para evitar erros de correlação provocados por pequenas translações do sinal em análise em relação ao sinal modelo são provocados deslocamentos de ± 2 pixels em x e em y, sendo o valor de correlação considerado, para cada candidato, o máximo resultante destas 25 correlações. Finalmente o candidato escolhido é aquele que obtiver um valor de correlação mais elevado.

4. Resultados e discussão

Este capítulo apresenta os resultados obtidos com as metodologias implementadas, nomeadamente na localização assistida, automática e no reconhecimento dos objectos, aplicados a casos reais, acompanhados de alguns comentários e análises às opções tomadas no desenvolvimento dessas metodologias.

4.1 Sistema assistido

O funcionamento do sistema no modo assistido requer a intervenção do utilizador apenas uma vez, na definição do objecto, indicando a sua localização na imagem da sequência a partir da qual se pretende efectuar o seu seguimento automático.

Os testes efectuados nesta secção estão organizados numa primeira parte sobre a fundamentação das opções tomadas na implementação do algoritmo de detecção, assim como da análise pormenorizada dos passos executados pelo sistema no seguimento de um objecto real entre a imagem N e $N+1$ de uma sequência de imagens, tendo em conta as duas fases de aproximação ao objecto: grosseira e refinada. Numa segunda parte é testado o comportamento do sistema em várias sequências de imagens reais.

4.1.1 Escolha do detector de contornos

A escolha do detector de contornos tornou-se crítica para o processo que envolve a localização assistida, devido à sua decisiva contribuição na definição da forma do objecto e consequentemente no seu seguimento. Por esse motivo foram testados dois métodos de detecção de contornos: *Sobel3x3* e a *LVIF (Local Variance Image Field)*.

4.1.1.1 Sobel3x3

A inclusão deste detector de contornos nos testes preliminares deveu-se ao facto de ser um dos mais utilizados para este fim assim como o reduzido tempo de processamento exigido.

Considere-se a distribuição de pixels em torno do pixel $[i, j]$ ilustrada na Figura 36.

a_0	a_1	a_2
a_7	$[i, j]$	a_3
a_6	a_5	a_4

Figura 36 – Distribuição de pixels numa máscara 3x3

No operador *Sobel* a magnitude do gradiente é dada por:

$$S_m = \sqrt{S_x^2 + S_y^2}$$

onde as derivadas parciais são dadas por:

$$S_x = (a_2 + ca_3 + a_4) - (a_0 + ca_7 + a_6)$$

$$S_y = (a_0 + ca_1 + a_2) - (a_6 + ca_5 + a_4)$$

com a constante $c = 2$.

Tal como outros operadores de gradiente, S_x e S_y podem ser implementados usando máscaras de convolução:

$$S_x = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} \quad S_y = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix}$$

4.1.1.2 LVIF

Este segundo método foi proposto como parte de um algoritmo de segmentação automática de objectos para imagens onde apenas o objecto de interesse se encontra bem focado encontrando-se as regiões do fundo desfocadas e esbatidas [Won02].

No contexto desse algoritmo o cálculo da *LVIF* surge como forma de auxiliar essa separação através do realce que faz às zonas de contorno, surgindo portanto nesse algoritmo basicamente como detector de contornos, que é o que se pretende para este sistema.

A explicação deste método foi já efectuada na secção 3.2.1.1.

4.1.1.3 Comparação Sobel3x3 / LVIF

Foram efectuados vários testes aos dois detectores tendo sido seleccionada a imagem da Figura 37 propositadamente pelas suas características (ruído de fundo), como forma de comparação entre os dois métodos.

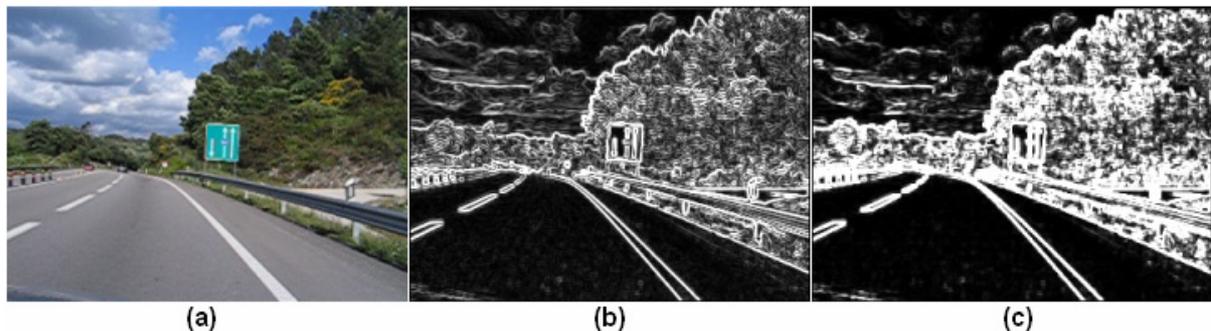


Figura 37 – Aplicação à imagem (a) original (b) o detector de contornos *Sobel3x3* (c) e o *LVIF*

Analisando o resultado obtido na Figura 37 poderia escolher-se o *Sobel3x3* como o detector de contornos mais adequado para o sistema, talvez porque neste caso visualmente se distinga melhor o contorno do objecto de interesse (sinal de trânsito) do fundo que o envolve (ruído).

No entanto há que ter em conta o facto do sistema implementado necessitar da informação de forma do objecto, obtida do seu contorno utilizando as assinaturas de contorno, como meio de o identificar. Este requisito exige ao sistema a utilização de um detector de contornos com grande sensibilidade às transições de cor, de forma a maximizar a energia da zona de contorno do objecto. Comparando agora o resultado dos dois detectores, tendo em conta esta informação, verifica-se que o detector de contornos *LVIF* é o que possui maior energia na zona de contorno. Quanto ao ruído de fundo, de facto o *LVIF* sendo mais sensível às transições de cor coloca-o em maior evidência do que o *Sobel3x3*. Este no entanto não é um factor crítico para o sistema, dado que possui informação prévia sobre a forma do objecto na imagem N-1 da sequência, o que lhe permite distinguir qual a zona de contorno com a forma definida.

A confirmação prática desta ideia foi obtida com testes efectuados em sequências de imagens (Figura 38), comparando o resultado obtido no seguimento de objectos, utilizando como método de detecção de contornos o *Sobel3x3* (Figura 38a, b) e c)) com o obtido utilizando o *LVIF* (Figura 38d, e) e f)). Este último consegue seguir com mais precisão o objecto definido.

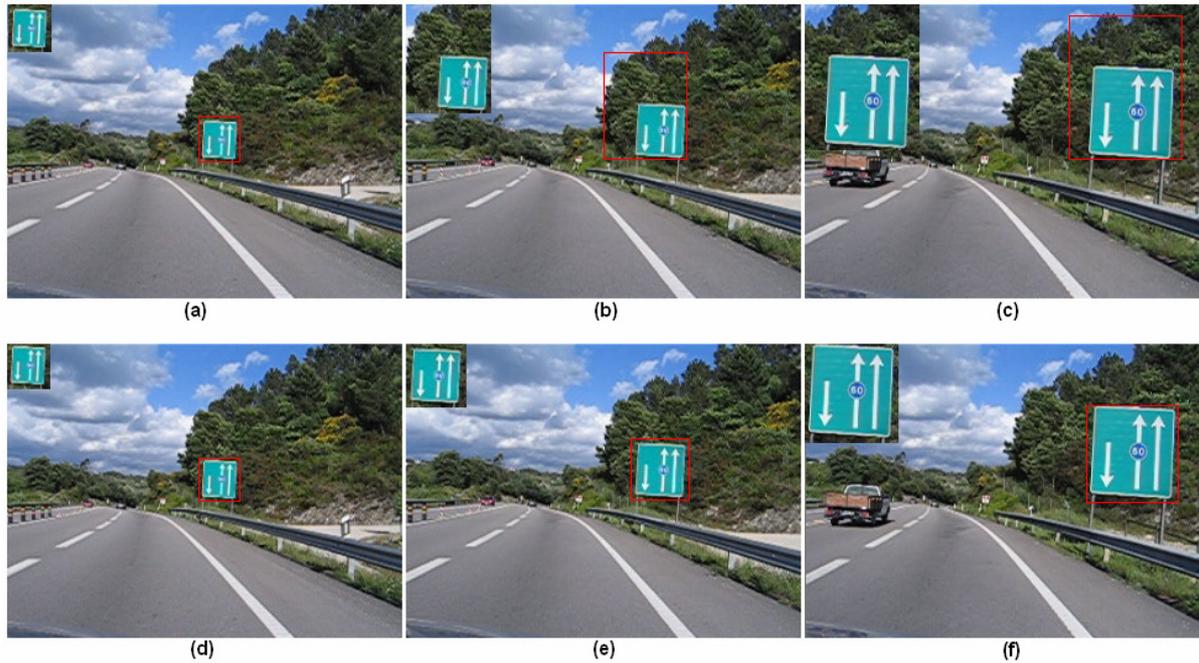


Figura 38 – Seguimento de um objecto numa sequência (imagens 1, 10 e 20)

A Figura 38 evidencia a dificuldade do algoritmo em efectuar o seguimento do objecto utilizando o método de detecção de contornos *Sobel3x3*. Esta dificuldade resulta da menor capacidade de maximização da energia do contorno do objecto por parte do *Sobel3x3* em relação ao *LVIF*.

4.1.2 Detecção e seguimento

O exemplo da Figura 39 serve de base ao esquema da Figura 40, o qual apresenta as operações que o sistema executa no seguimento de um objecto entre a imagem N e $N+1$ de uma sequência de imagens.

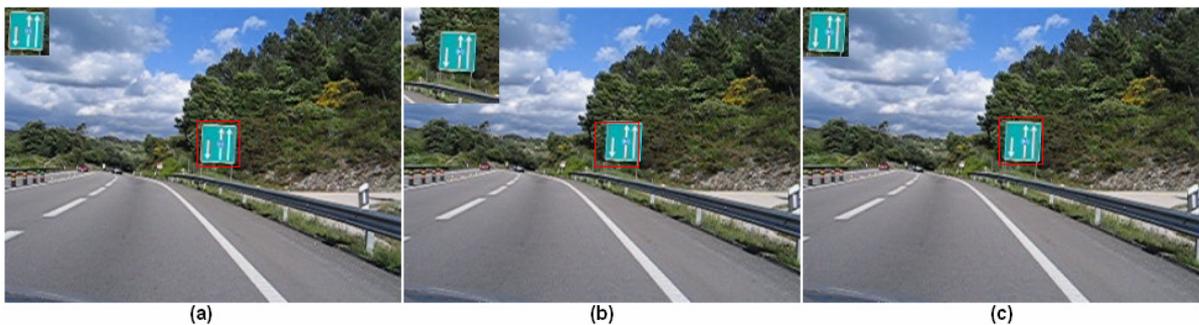


Figura 39 – Detecção e seguimento de um objecto entre a imagem (a) N e (b)-(c) $N+1$ de uma sequência

As operações executadas pelo sistema são (Figura 40):

1. definição do objecto para detecção/seguimento, por parte do utilizador, na imagem N da sequência (Figura 39a));
2. obtenção automática da janela de pesquisa na imagem N+1, com base na localização do objecto na imagem N (Figura 39b));
3. determinação do deslocamento do objecto entre as imagens através da correspondência de modelos (*template matching*), em que é efectuada uma aproximação grosseira à posição do objecto;
4. determinação da imagem de energia correspondente à janela de pesquisa;
5. determinação da imagem de energia do objecto;
6. cálculo do contorno do objecto;
7. aproximação inicial do contorno na janela de pesquisa;
8. aproximação final (recursiva) ao contorno do objecto na imagem N+1 tendo como comparação o obtido na imagem N;
9. suavização do contorno obtido;
10. actualização da informação do objecto obtida na imagem N com a informação obtida na imagem N+1 (Figura 39c)).

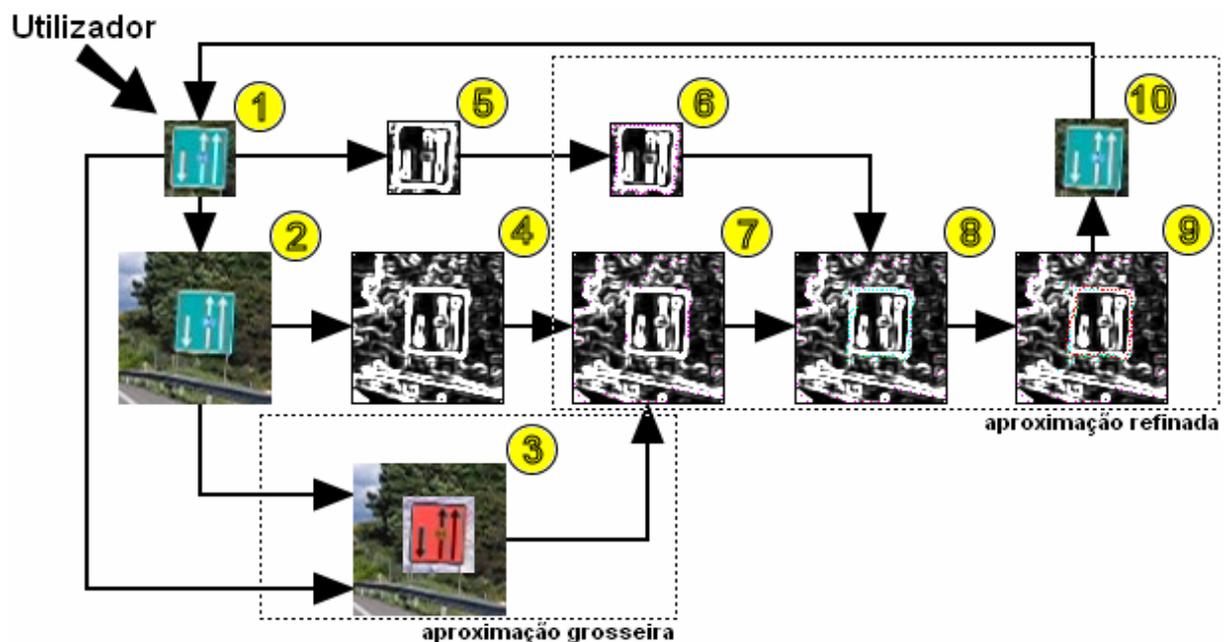


Figura 40 – Sequência de operações executadas na detecção e seguimento assistido

4.1.3 Localização de objectos - exemplos

Nesta secção são apresentados 4 exemplos, seleccionados dos vários testes efectuados ao sistema utilizando esta metodologia. São apontadas também as potencialidades e limitações deste tipo de abordagem. O objecto a ser localizado ao longo de cada uma das sequências, encontra-se definido na primeira imagem pelo utilizador.

No exemplo da Figura 41, é definido um sinal de trânsito com a forma quadrangular. Este possui cores contrastantes com o fundo e contornos da forma bem definidos, sendo localizado correctamente ao longo da sequência de imagens, apesar de existir uma significativa alteração de escala do objecto.



Figura 41 – Localização de um sinal quadrangular numa sequência (imagens 1, 21 e 28)

A Figura 42 apresenta a localização de um objecto com forma irregular, mais precisamente de um automóvel. Tal como no caso anterior, existe uma alteração de escala do automóvel significativa ao longo da sequência. Apesar disso a localização é efectuada de forma correcta, existindo também neste caso um bom contraste entre as cores do automóvel e o fundo.



Figura 42 – Localização de um automóvel numa sequência (imagens 1, 17 e 58)

No caso da Figura 43 é efectuada a localização, com sucesso, de um sinal de trânsito de forma circular. Apesar de existir, na parte final da sequência, uma diminuição do contraste entre o sinal e o fundo, este é ainda suficiente para não comprometer a sua localização com sucesso.



Figura 43 – Localização de um sinal circular numa sequência (imagens 1, 10 e 15)

Finalmente, na Figura 44 procedeu-se à localização de um sinal de trânsito de forma quadrangular. Neste caso, a parte escura do sinal é muito semelhante ao fundo, tendo implicações na fase de aproximação grosseira à posição do objecto nas imagens seguintes da sequência. Existem também, pelo mesmo motivo e pela união do sinal à barreira de protecção, incorrecções na fase de aproximação refinada à sua forma, não permitindo a compensação correcta do efeito de escala. A acumulação destes erros de aproximação ao objecto, resultam no insucesso da sua localização ao longo da sequência de imagens. Uma possível solução para este problema seria a utilização de informação de movimento do sinal ao longo da sequência.



Figura 44 – Localização de um sinal quadrangular numa sequência (imagens 1, 8 e 15)

4.2 Sistema automático

Os testes efectuados ao sistema implementado e já detalhadamente descrito, podem dividir-se em três partes:

- **detecção:** Foram analisadas imagens de cenas de tráfego rodoviário com a presença de sinais de trânsito das formas (Circular, Triangular, Quadrangular e Triangular Invertida) e cores (Azul e/ou Vermelho) pré-definidas para testar as capacidades do sistema na sua detecção;
- **reconhecimento:** Após a detecção dos sinais de trânsito e com o auxílio de uma base de dados de sinais é efectuado o seu reconhecimento;
- **seguimento:** Foram testadas algumas sequências de imagens com a presença de sinais de trânsito com as mesmas características pré-definidas e efectuado o seu seguimento.

4.2.1 Detecção e segmentação com base na cor

A primeira parte dos testes consistiram na análise de 172 imagens de 320x240 pixels com sinais de trânsito, obtidas com uma câmara fotográfica digital. A informação relativa à sinalização contida nessas imagens encontra-se resumida na Tabela 5.

Formas	Quantidade (#)	Quantidade (%)
Circular	143	55,43%
Triangular	47	18,22%
Quadrangular	35	13,57%
Triangular Invertida	33	12,79%
	258	Total sinais

Tabela 5 – Formas dos sinais analisados

Os 5 passos abordados de seguida têm por objectivo ilustrar os problemas, assim como as soluções encontradas, no processo de classificação das regiões por cor. Esse processo é efectuado isolando em cada imagem apenas as regiões que cumprem determinados requisitos.

1ºPasso: Após a segmentação e etiquetagem de cada uma das imagens surgiu a necessidade de escolher as regiões daí resultantes candidatas a objecto de interesse, ou seja, sinal da(s) cor(es) desejada(s). Para isso procedeu-se à imposição de alguns limites abaixo dos quais, apesar de uma dada região ser efectivamente um sinal ou de conter um sinal, não seja possível a obtenção de informação mínima na imagem, necessária à detecção e/ou ao reconhecimento. Assim, determinou-se a rejeição de regiões com:

- Área (número de pixels) inferior a 64 pixels;

- Dimensão ($\Delta x, \Delta y$): $\Delta x < 16 \vee \Delta y < 16$ pixels.

Além da eliminação das regiões já referidas são também eliminadas outras regiões, as quais foram já explicadas na secção 3.3.1.3.

Da imposição destas restrições ao total de sinais inicial (Tabela 5) obteve-se:

Total de sinais	258
Dimensão < 16X16 pixels	17
Área < 64 pixels (Área mínima)	3
Total rejeitados pelo sistema	20
Total de sinais	238

Tabela 6 – Resultado filtro de dimensão

2ºPasso: Efectuou-se a análise dos 238 sinais filtrados no passo anterior, com o algoritmo implementado para a detecção inicial por cor. Procedeu-se à separação das regiões detectadas que continham realmente sinais (mesmo com ruído) das que não continham ou continham apenas em parte. Desta separação resultaram 3 tipos de problemas: sinais unidos a outro sinal, sinais unidos ao fundo e sinais com contorno partido. A Tabela 7 apresenta a quantidade de casos verificados para cada um dos problemas referidos.

Total de sinais	238
Unidos a outro sinal	14
Unidos ao fundo	3
Partidos	15
Parcialmente/Não detectados por cor	32

Tabela 7 – Resultado da detecção inicial por cor

3ºPasso: Por forma a tornar o algoritmo capaz de detectar sinais unidos a outro sinal (Tabela 7), procedeu-se ao cálculo da relação Altura/Largura para todas as regiões detectadas que correspondiam a sinais e separou-se as regiões que continham apenas um sinal das que continham dois (máximo existente nas imagens testadas Figura 45), o resultado obtido encontra-se no gráfico da Figura 46.

Apesar da igualdade entre $\sigma_{normais}$ e σ_{unidos} , optou-se pela sua inclusão no cálculo de μ_f , por se pretender caracterizar a dispersão da classe (Figura 47).

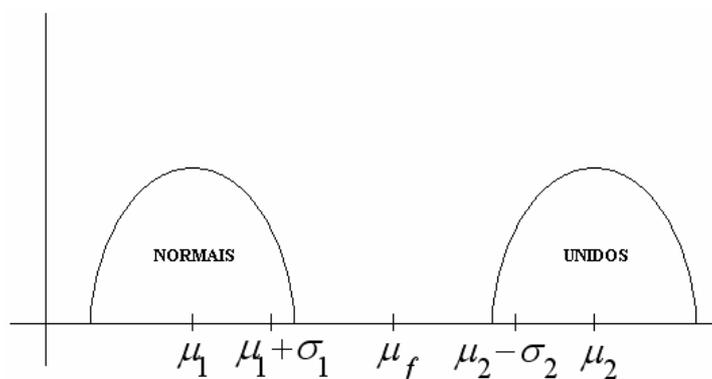


Figura 47 – Distinção entre sinais normais e unidos

Sempre que são detectadas pelo sistema regiões em que o valor da sua relação Altura/Largura é superior a $\mu_f = 1,46$ procede-se à sua divisão a meio em altura, resultando em duas regiões tratadas a partir desse momento independentemente uma da outra. A Figura 48 ilustra o resultado da aplicação deste passo às imagens da Figura 45, o qual não conseguiu separar correctamente os sinais apenas em 1 caso. Uma vez que o erro se verificou apenas em 1 caso, optou-se por adoptar este processo de separação sem introdução de refinamentos. No entanto uma possível solução seria a divisão das regiões na zona de contacto mais estreita.

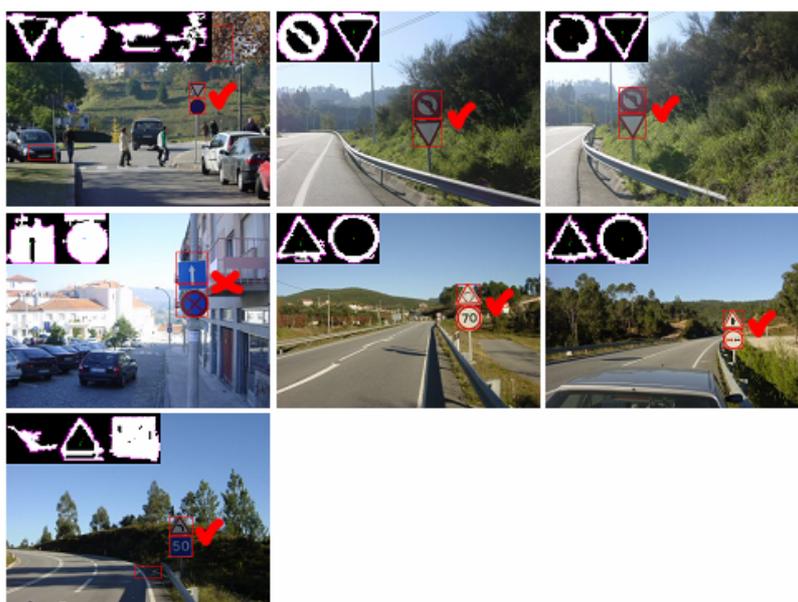


Figura 48 – Resultado da separação dos sinais unidos

4ºPasso: Os sinais unidos ao fundo (Tabela 7) são consequência de uma má separação na fase de segmentação (Figura 49) sendo envolvidos/ligados pelo fundo e tornando-os por esse motivo não detectados pelo sistema. Uma vez que é uma situação pouco frequente optou-se por não tratar esta situação.

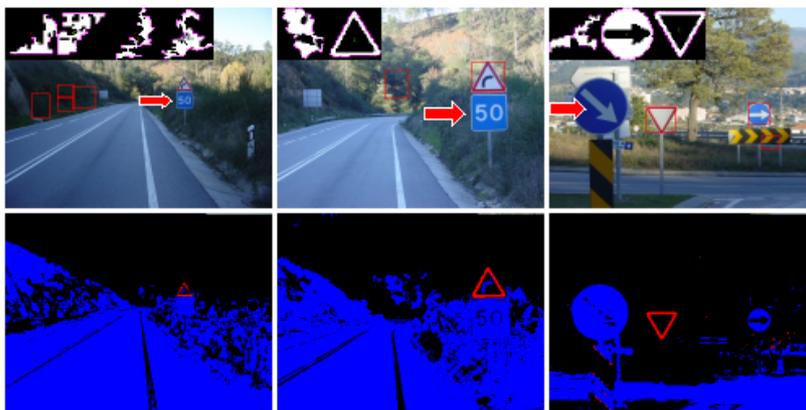


Figura 49 – Sinais unidos ao fundo não detectados

5ºPasso: A recuperação de sinais partidos (Tabela 7) é efectuada com base na repetição do processo descrito anteriormente, até ao 3ºPasso inclusive, com a diferença que neste caso é aplicada à imagem obtida da fase de segmentação de cor uma dilatação seguida de uma erosão, unindo assim pequenas regiões muito próximas que eventualmente corresponderão, após a sua ligação, a sinais de trânsito. Este passo serve apenas como complemento ao processo de detecção normal, ou seja, só serão considerados válidos objectos de interesse detectados por este processo, aqueles que não tiverem sido já detectados pelo processo normal (evitando detecções repetidas de objectos) ou tendo sido detectados foram rejeitados pelo sistema (mal detectados, possivelmente partidos).

Os sinais partidos sinalizados por uma seta na Figura 50 foram eliminados pelo sistema por duas possíveis razões:

1. não possuíam os requisitos mínimos para serem considerados para análise (não identificados no topo das imagens, por exemplo: sinais assinalados nas imagens da 2ª linha da Figura 50);
2. simplesmente foram rejeitados (*Classificador de forma* - seleccionados nas imagens mas partidos, por exemplo: sinais assinalados nas imagens da 1ª linha da Figura 50).

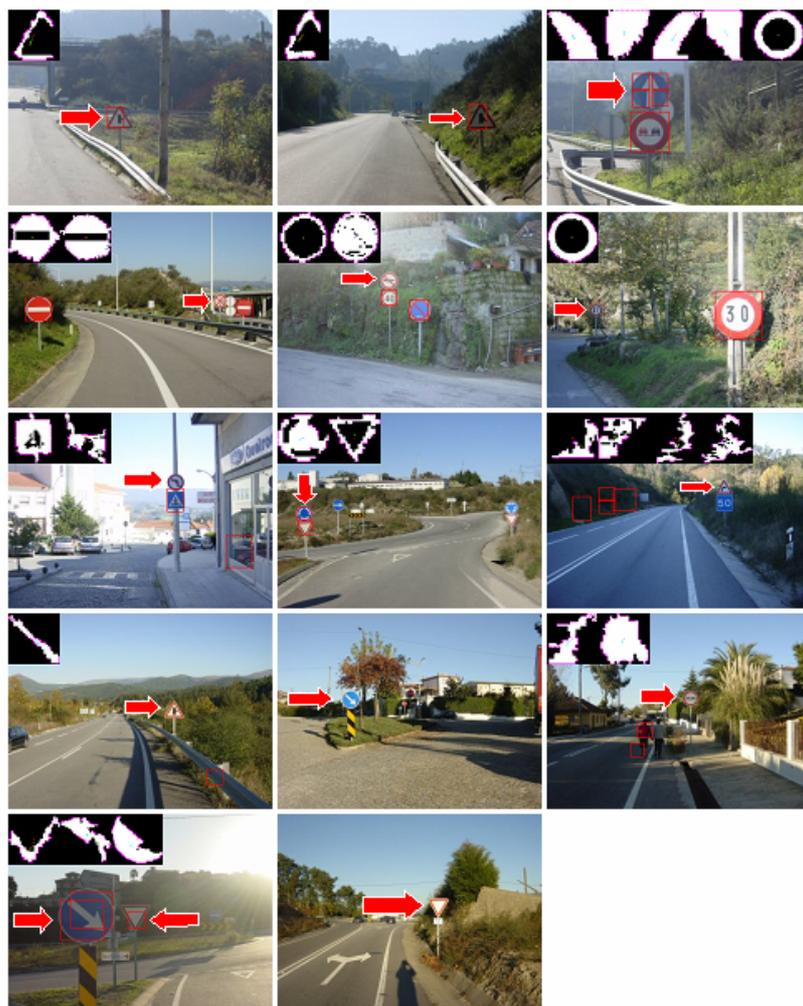


Figura 50 – Sinais partidos parcialmente/não detectados

Dos 15 sinais partidos da Tabela 7 e sinalizados na Figura 50 foram recuperados 13 (identificados na Figura 51), utilizando este passo complementar.

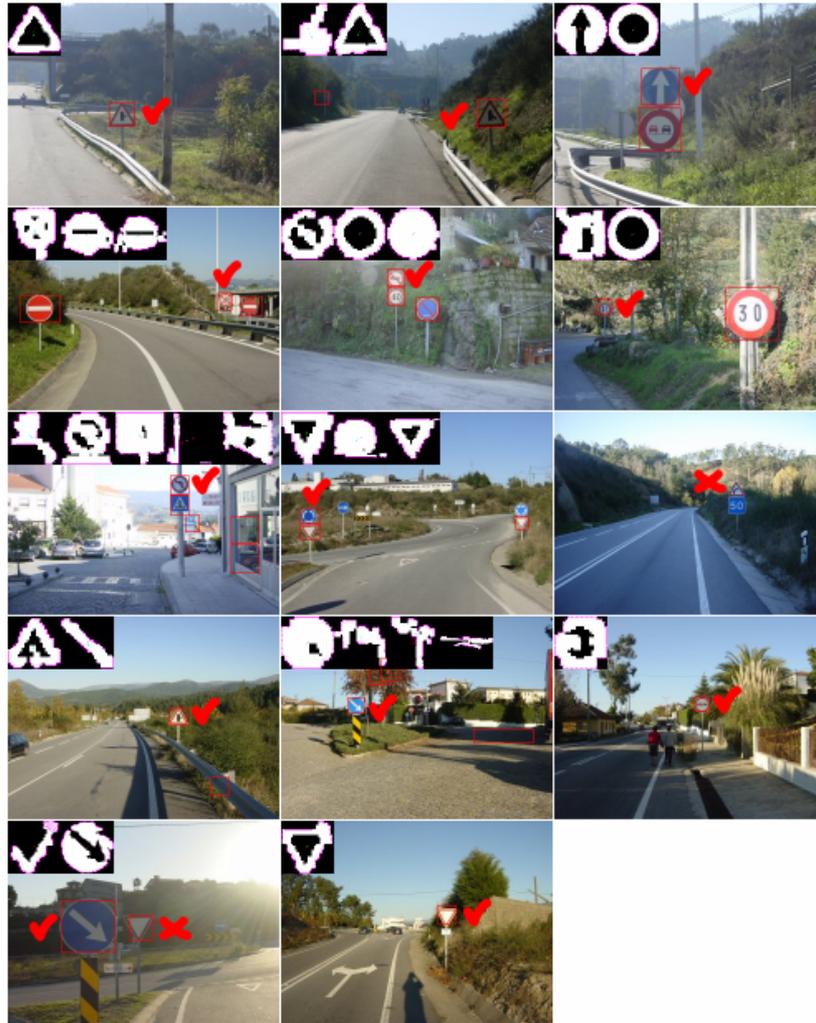


Figura 51 – Sinais partidos parcialmente/não detectados recuperados

Os resultados finais da detecção inicial por cor, tendo como termo de comparação a Tabela 7 estão resumidos na Tabela 8. Dos 14 sinais inicialmente unidos a outro sinal apenas 2 foram mal separados. Mantêm-se os mesmos 3 sinais inicialmente unidos ao fundo. Finalmente, dos 15 sinais inicialmente partidos apenas 2 não foram recuperados.

Não Recuperados	
Unidos a outro sinal	2
Unidos ao fundo	3
Partidos	2
Parcialmente/Não detectados por cor	7
Total de sinais	238
Total detectados por cor (#)	231
Total detectados por cor (%)	97,06%

Tabela 8 – Resultado final da detecção inicial por cor

4.2.2 Classificador de forma

A classificação das regiões por forma, obtidas da detecção por cor, tem por objectivo não só validar a sua forma como sendo uma das quatro pré-definidas do sistema mas também eliminar possíveis sinais que, apesar de estarem contidos em determinada região, estão mal segmentados e que portanto não corresponderão a nenhuma das formas procuradas. É por esse motivo que na detecção por cor apenas houve a preocupação de identificar as regiões detectadas que podem conter sinais, apesar de poder existir ruído nas mesmas. Tendo em conta este facto verifica-se que para além das 231 regiões com sinais (Tabela 8) passam ainda para esta fase os *parcialmente detectados*, nomeadamente os *unidos a outro sinal* (Figura 48) e um dos *partidos* (Figura 51). Passam assim para esta fase 234 regiões com sinais (ver Tabela 9).

Total detectados por cor	231
Unidos a outro sinal	2
Partidos	1
Total analisados pelo classificador de forma	234

Tabela 9 – Sinais analisados pelo classificador de forma

4.2.2.1 Escolha do classificador de forma

O classificador de forma foi escolhido de entre 3 classificadores que têm em comum a quantificação do grau de semelhança entre duas formas, representadas pelas suas assinaturas de contorno. Nesta fase surgiu a necessidade de encontrar um procedimento que permitisse extrair de entre todas as regiões detectadas pelo sistema aquelas que são das formas pré-definidas. O primeiro passo foi a extracção da assinatura de contorno de cada uma das regiões normalizadas em imagens de 64x64 pixels. A amostragem da assinatura de contorno é efectuada com $\Delta\theta = \frac{360}{\sqrt{64^2 + 64^2}} = 4^\circ$ cuja equação foi já referida anteriormente.

Após a obtenção das assinaturas de contorno de cada uma das regiões, assim como, previamente, de cada uma das formas pré-definidas, foram testados três classificadores. A formulação dos classificadores foi obtida de testes experimentais. Têm como principal característica o facto de medirem a semelhança entre formas, neste caso entre a forma de uma dada região e as formas pré-definidas, para o mesmo número de pontos. Os três classificadores testados foram:

(1) **Coefficiente de Correlação (CC):**

$$\rho_{x,y} = \frac{Cov(X,Y)}{\sigma_x \cdot \sigma_y} \quad \text{onde} \quad -1 \leq \rho_{x,y} \leq 1 \quad \text{e} \quad Cov(X,Y) = \frac{1}{n} \sum_{i=1}^n (x_i - \mu_x)(y_i - \mu_y)$$

sendo:

n número de pontos do contorno (90 pontos);

x_i a distância do ponto i de contorno da região ao seu centro;

y_i a distância do ponto i de contorno da forma pré-definida ao seu centro;

μ_x, μ_y as médias das distâncias;

$$\mu_x = \frac{1}{n} \sum_{i=1}^n x_i, \quad \mu_y = \frac{1}{n} \sum_{i=1}^n y_i$$

σ_x, σ_y os desvios padrão;

$$\sigma_x = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \mu_x)^2}, \quad \sigma_y = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \mu_y)^2}$$

Os valores obtidos do classificador, que se situam entre $-1 \leq \rho_{x,y} \leq 1$, são normalizados para valores entre $-100\% \leq \rho_{xy} \leq +100\%$. Um valor para a correlação próximo do mínimo (-100%), indica que as duas formas analisadas são muito diferentes. Por outro lado um valor para a correlação próximo do máximo (+100%), significa que as duas formas analisadas são muito semelhantes. Nos resultados apresentados com este classificador os valores são sempre positivos devido ao facto do classificador devolver ao sistema apenas o coeficiente de correlação mais elevado, dos quatro obtidos da comparação com as formas pré-definidas, indicando assim ao sistema a forma mais semelhante.

(2) **Média aritmética dos valores absolutos das diferenças entre contornos (MédiaDif):**

$$\mu = \frac{1}{n} \sum_{i=1}^n |x_i - y_i|$$

(3) **Média aritmética dos desvios absolutos dos pontos de contorno a partir da sua média (Desvio Médio):**

$$\sigma = \frac{1}{n} \sum_{i=1}^n |x_i - y_i| - \mu$$

Tendo como objectivo determinar de entre os três classificadores aquele que melhor distinguia as regiões que são sinais daquelas que não o são procedeu-se à comparação da assinatura de contorno de cada uma delas com as das formas pré-definidas, utilizando para cada classificador a fórmula respectiva. Daqui se obtém, para cada classificador, quatro valores, um de cada forma pré-definida, dos quais apenas é retornado aquele que maximiza a semelhança entre formas. No caso do primeiro classificador é retornado o valor mais elevado encontrado, i.é. correlação mais elevada, no caso dos outros dois classificadores é retornado o valor mais baixo encontrado (menor diferença). De seguida, procedeu-se à separação dos valores retornados pelos classificadores, correspondentes a regiões que continham sinais, dos valores daquelas que não continham. Finalmente, são omitidos os valores dos classificadores associados a regiões com sinais mal detectados, isto para evitar a introdução de ruído na análise dos classificadores. Os resultados obtidos são apresentados na Figura 52, Figura 53 e Figura 54.

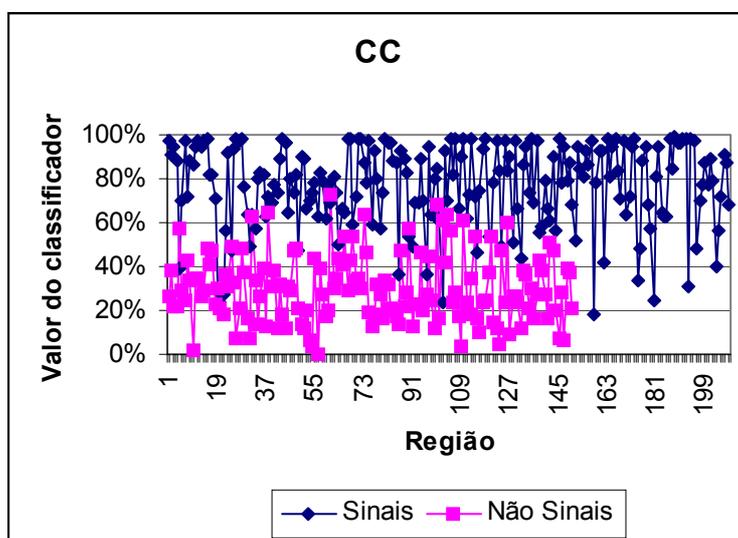


Figura 52 – Resultados obtidos para o 1º Classificador

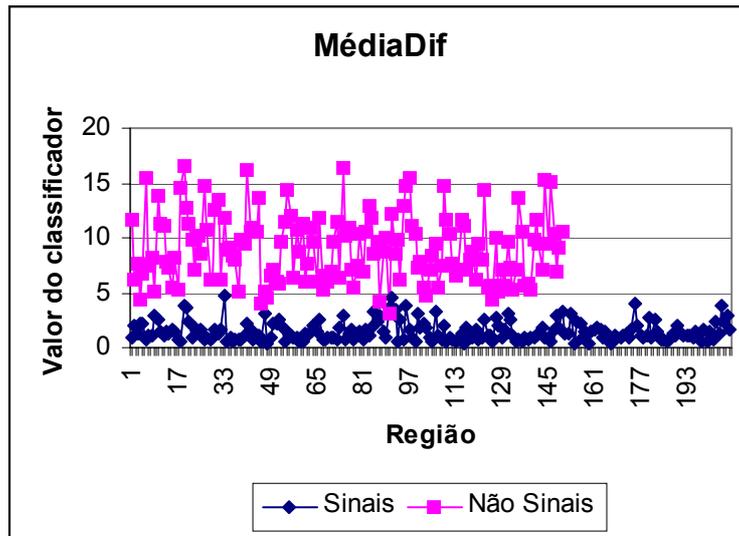


Figura 53 – Resultados obtidos para o 2º Classificador

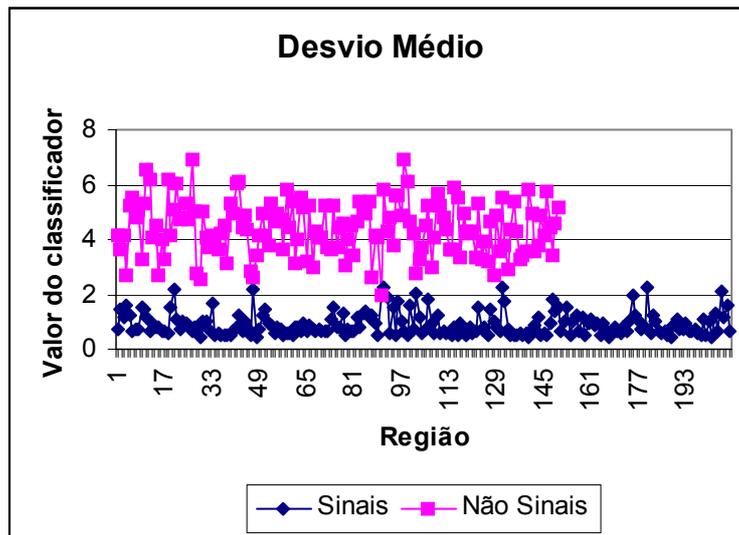


Figura 54 – Resultados obtidos para o 3º Classificador

Analisando os gráficos da Figura 52, Figura 53 e Figura 54, aparentemente o 3º classificador é o que melhor separa a classe dos sinais da classe dos não sinais, não sendo no entanto evidente qual o valor a atribuir ao classificador. Por esse motivo, procedeu-se à determinação desse valor para cada classificador, utilizando duas metodologias:

1. tendo em conta apenas a concentração dos valores de cada uma das classes.

$\mu_1 = (\mu_{\text{sinais}} + \mu_{\text{nãosinais}}) / 2$, sendo $\mu = \frac{1}{n} \sum_{i=1}^n x_i$ a média dos n valores de cada classe e x_i um dado valor na posição i ;

2. tendo em conta a concentração e a dispersão dos valores (Figura 47) de cada uma das classes. $\mu_2 = ((\mu_{\text{sinais}} + \sigma_{\text{sinais}}) + (\mu_{\text{nãosiniais}} - \sigma_{\text{nãosiniais}}))/2$, sendo

$$\sigma = \frac{1}{n} \sum_{i=1}^n |x_i - \mu|$$

o desvio médio dos n valores de cada classe. No caso do

$$\text{classificador } CC \mu_2 = ((\mu_{\text{sinais}} - \sigma_{\text{sinais}}) + (\mu_{\text{nãosiniais}} + \sigma_{\text{nãosiniais}}))/2.$$

Os valores obtidos para cada um dos classificadores, aplicando as duas metodologias anteriores, são apresentados na Tabela 10.

Parâmetros	Classificadores					
	CC		MédiaDif		Desvio Médio	
	Sinais	Não Sinais	Sinais	Não Sinais	Sinais	Não Sinais
Média	76,78%	30,14%	1,46	9,02	0,91	4,41
Desvio Médio	15,57%	12,64%	0,69	2,56	0,33	0,80
Metodologia 1	53,46%		5,24		2,66	
Metodologia 2	51,99%		4,30		2,43	

Tabela 10 – Valores obtidos para os classificadores

Seguidamente, procedeu-se à aplicação dos valores obtidos na Tabela 10, para cada um dos classificadores, aos valores dos resultados obtidos para sinais e não sinais, apresentados na Figura 52, Figura 53 e Figura 54 respectivamente, e determinados os falsos positivos (FP) e falsos negativos (FN) resultantes (Tabela 11).

Parâmetros	Classificadores					
	CC		MédiaDif		Desvio Médio	
	FP	FN	FP	FN	FP	FN
Metodologia 1	16	25	14	0	4	0
Metodologia 2	16	24	3	2	1	0

Tabela 11 – Falsos positivos e negativos obtidos para os classificadores

Finalmente, analisando a Tabela 10 e a Tabela 11 em conjunto, pode concluir-se que o classificador *Desvio Médio* (3º Classificador), utilizando o valor 2,43 obtido com a metodologia 2, é o que minimiza a obtenção de FP e FN , sendo portanto o classificador que apresenta menor erro de separação entre sinais e não sinais.

Apesar de ser este o classificador utilizado genericamente no sistema o valor encontrado de 2,43 é utilizado apenas na distinção entre sinais e não sinais detectados pelo método normal. No caso do método complementar de detecção (aplicação de uma dilatação e erosão à imagem segmentada) claramente terá que ser encontrado um outro valor (mais restritivo) na distinção entre sinais e não sinais. Foi utilizado o mesmo procedimento para

o método complementar e obtidos os valores apresentados na Figura 55, Figura 56 e Figura 57.

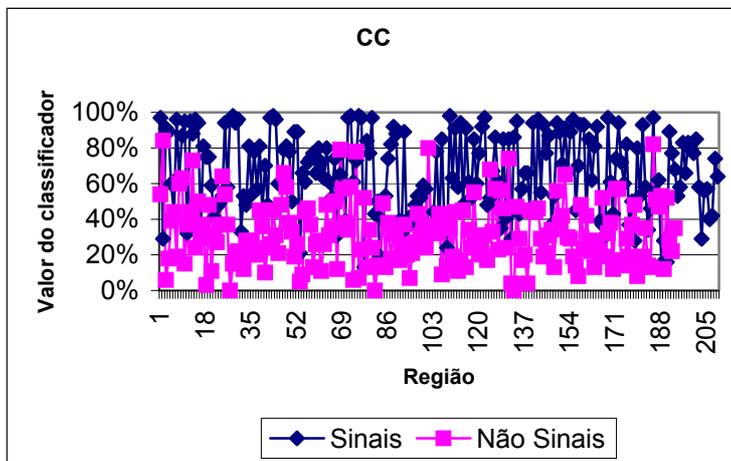


Figura 55 – Resultados para o método complementar do 1º Classificador

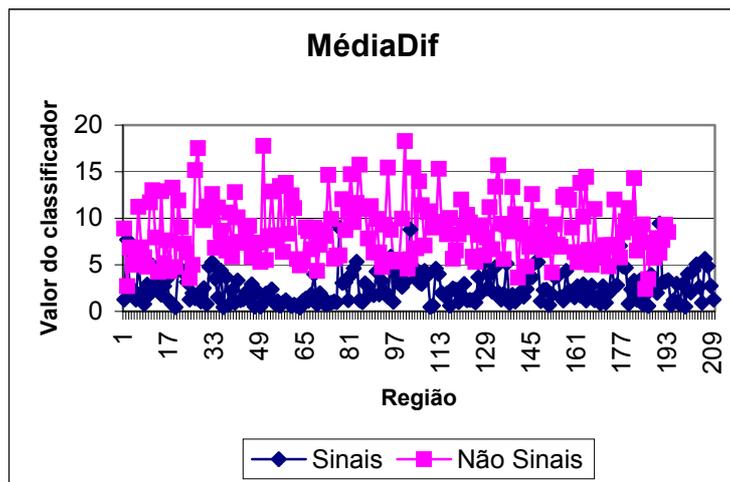


Figura 56 – Resultados para o método complementar do 2º Classificador

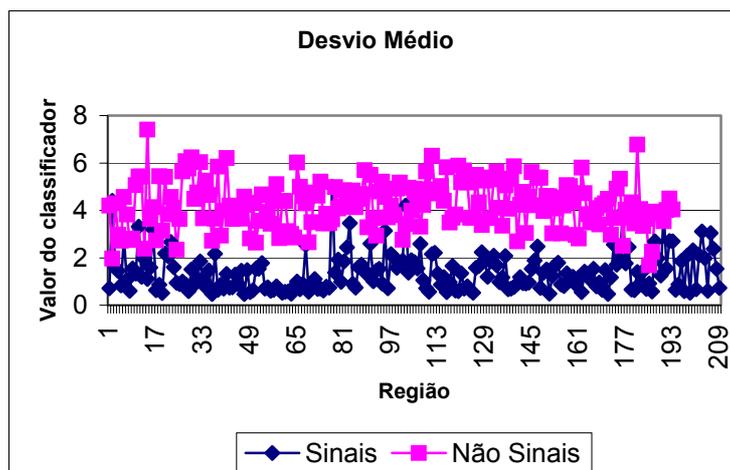


Figura 57 – Resultados para o método complementar do 3º Classificador

O valor que distingue as regiões que são sinais das formas desejadas das que não são para o método complementar não pode ser determinado da mesma forma que o foi para o método normal. Uma das razões tem a ver com o facto de neste caso as regiões correspondentes a sinais completos (não partidos) possivelmente já detectadas se encontrarem agora deformadas ou ligadas a outras regiões (devido à dilatação - erosão), justificação em parte para a menor separação das classes verificada nos resultados dos gráficos da Figura 55, Figura 56 e Figura 57. Apesar disso é relativamente fácil pela observação dos mesmos que continua a ser o 3º classificador (*desvio médio*) que melhores resultados apresenta na distinção entre sinais e não sinais. O valor obtido para o método complementar teve então de ser determinado empiricamente tendo em atenção a necessidade de ser um valor suficientemente baixo em relação ao utilizado no método normal (mais restritivo) de forma a evitar a introdução de erros de detecção no sistema e ao mesmo tempo que fosse suficientemente alto para detectar (recuperar) o maior número possível de sinais não detectados pelo método normal. Tendo em conta estas considerações e o gráfico da Figura 57 chegou-se ao valor 2,20. A Tabela 12 resume os valores utilizados no sistema para o classificador de forma escolhido (*desvio médio*).

Método	Sinal	Não sinal
Normal	$\leq 2,43$	$> 2,43$
Complementar	$\leq 2,20$	$> 2,20$

Tabela 12 – Resumo dos valores do classificador

Para cada região analisada é quantificada a proximidade da sua forma a cada uma das formas pré-definidas (Circular, Triangular, Quadrangular e Triangular Invertida), utilizando o classificador *desvio médio*. De seguida é identificado para cada região, o valor mínimo de entre os quatro calculados, ou seja, a forma mais próxima. Finalmente é verificado, utilizando os valores da Tabela 12, se a região é suficientemente próxima da forma pré-definida que lhe foi associada. Em caso afirmativo, a região é classificada como

senal, com a forma associada. Caso contrário a região não é validada como sinal e é simplesmente rejeitada pelo classificador de forma.

Aplicando ao sistema implementado os limites definidos na Tabela 12 procedeu-se uma vez mais à análise das regiões resultantes da segmentação (Tabela 9) dos quais 3 apresentavam já problemas (2 unidos a outro sinal e 1 partido). Os resultados obtidos dessa análise estão resumidos na Tabela 13.

Total analisados pelo classificador de forma	234
<i>Eliminados pelo classificador</i>	12
<i>Forma atribuída errada</i>	7
Classificados por forma correctamente(#)	215
Classificados por forma correctamente(%)	91,88%
Falsos positivos	3
Deteção e classificação por forma correcta	90,34%

Tabela 13 – Resultados obtidos da deteção e classificação por forma

Todos os sinais *eliminados pelo classificador* apresentavam alterações devidas, umas vezes, a contornos incompletos, noutros casos, a contornos com ruído de fundo colado (Figura 58a) e b)). Os sinais com *forma atribuída errada* apresentavam algumas deformações muito particulares as quais conjugadas levaram a que o resultado do *desvio médio* fosse menor (mais próximo) da forma que lhe foi atribuída do que à que efectivamente pertence, causado pela segmentação defeituosa (Figura 58c)). Dos 3 sinais da Tabela 9 que já apresentavam problemas, um dos *unidos a outro sinal* assim como o *partido* foram *eliminados pelo classificador*, ao outro *unido a outro sinal* foi atribuída uma *forma errada*. Finalmente os *falsos positivos* foram obtidos porque cumpriram todos os requisitos exigidos pelo sistema e apresentavam semelhanças evidentes com as formas pré-definidas procuradas. Um dos *falsos positivos* resulta da aplicação do *método normal* (Tabela 11), sendo os outros dois resultado da aplicação do *método complementar de deteção*. No caso da Figura 58d) à região sinalizada foi atribuída a forma *Triangular*.

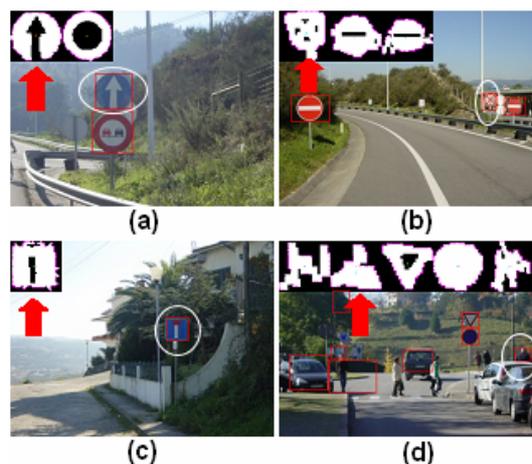


Figura 58 – Exemplos de erros de classificação por forma

O exemplo da Figura 58c) mostra um sinal de forma *Quadrangular* ao qual foi atribuída a forma *Circular*. A forma do sinal encontra-se representada pela sua assinatura de contorno no gráfico da Figura 59, bem como as assinaturas das formas pré-definidas: *Circular* e *Quadrangular*.

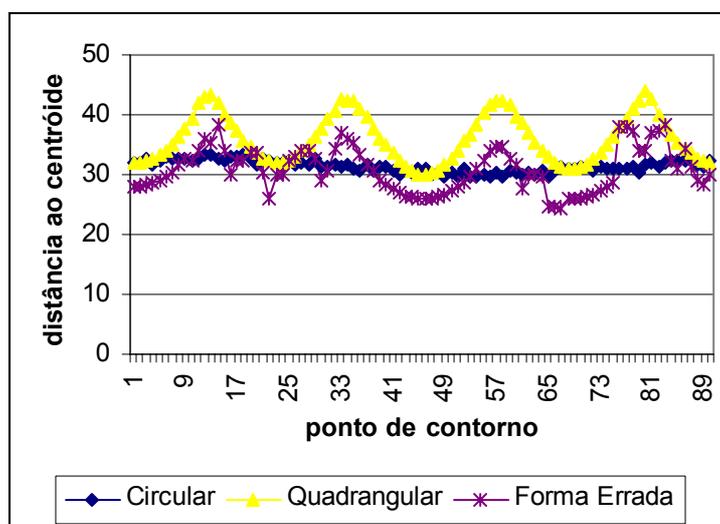


Figura 59 – Assinaturas de duas formas pré-definidas e da região com forma errada

Analisando o sinal de forma *Quadrangular* da Figura 58c) e a sua assinatura de contorno no gráfico da Figura 59 pode concluir-se que devido a má segmentação do sinal a sua normalização produziu um efeito de escala no objecto. Essas alterações produziram uma alteração do seu contorno e alteração na posição da sua assinatura de contorno relativamente à assinatura da forma *Quadrangular* pré-definida. Essa alteração de posição traduz-se, no gráfico da Figura 59, num deslocamento no eixo dos yy em direcção à posição da assinatura de contorno da forma *Circular* e num conseqüente afastamento da posição da assinatura da forma *Quadrangular*. Daqui resultou uma alteração do valor do

classificador, aumento do *desvio médio*, em relação à forma *Quadrangular* e uma diminuição do *desvio médio*, em relação à forma *Circular*, originando assim a atribuição da forma errada ao sinal (Tabela 14).

	$\sigma_{circular}$	$\sigma_{quadrangular}$
σ_{forma_errada}	1,55	2,05

Tabela 14 – Medidas obtidas do classificador de forma

A forma de sinal *Quadrangular* é de facto a mais susceptível a confusão com a forma de sinal *Circular* por parte do classificador de forma. Esta confirmação foi obtida dos testes experimentais realizados, resumidos na Tabela 13, de onde se concluiu que os 7 sinais com *forma atribuída errada* eram de facto sinais de forma *Quadrangular* mas que o classificador de forma atribuiu a forma *Circular*. Uma possível solução para este problema, seria a utilização de um segundo classificador de forma, sempre que resultasse do primeiro classificador (desvio médio) a validação de mais de uma forma pré-definida. Este caso verifica-se no exemplo da Tabela 14 o qual possui duas formas pré-definidas dentro dos limites do classificador (2,43).

4.2.2.2 Sensibilidade à rotação

Com o objectivo de quantificar a influência da rotação do objecto de interesse, efectuou-se um teste no qual foram utilizadas imagens *RGB* de três das formas pré-definidas (*Triangular*, *Triangular Invertida* e *Quadrangular*) às quais foram introduzidas rotações entre $[-20^\circ, \dots, +20^\circ]$, com saltos de 1° , e determinado o $\Delta\theta$ máximo admitido pelo sistema para cada uma dessas formas, sem que se verifiquem erros na classificação do sinal por forma.

O sistema, tal como já mencionado anteriormente, pressupõe que os sinais a detectar se encontrem posicionados a aproximadamente 0° , ou seja, $\theta_r = 0^\circ \pm \Delta\theta$ em relação ao eixo dos *yy*. Numa situação real essa rotação $\Delta\theta$ poderá ter mais ou menos influência na detecção do sinal por forma dependendo do seu valor.

Os resultados obtidos nos testes efectuados encontram-se apresentados na Figura 60, Figura 61 e Figura 62. A cada ângulo testado o sistema devolveu o valor do classificador de forma e a forma atribuída: classe correcta (pontos a azul), classe errada (pontos a vermelho). A partir do valor 2,43 (ver Tabela 12) o classificador (eixo dos *yy*) considera a forma inválida (barras vermelhas no eixo dos *xx*), abaixo desse valor a forma é validada (barra a verde no eixo dos *xx*).

Nestes resultados constata-se que o classificador admite rotações para a forma:

- *triangular* entre $[-11^\circ, \dots, +12^\circ]$;
- *triangular Invertida* entre $[-12^\circ, \dots, +12^\circ]$;

- quadrangular entre $[-6^\circ, \dots, +5^\circ]$.

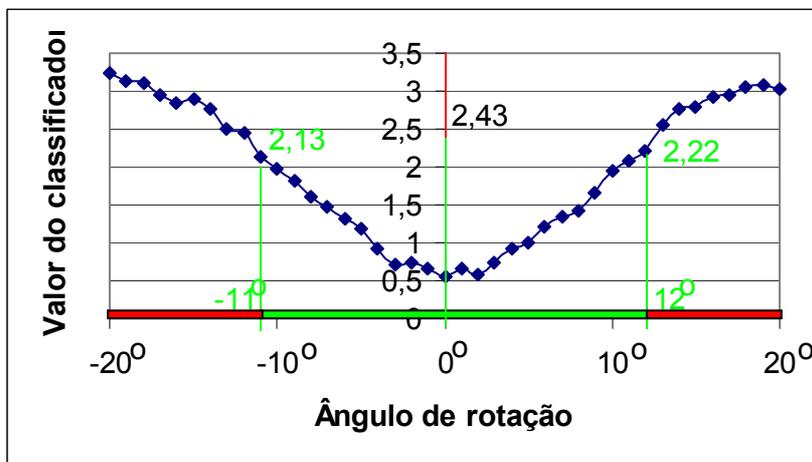


Figura 60 – Resultado da rotação da forma *Triangular*

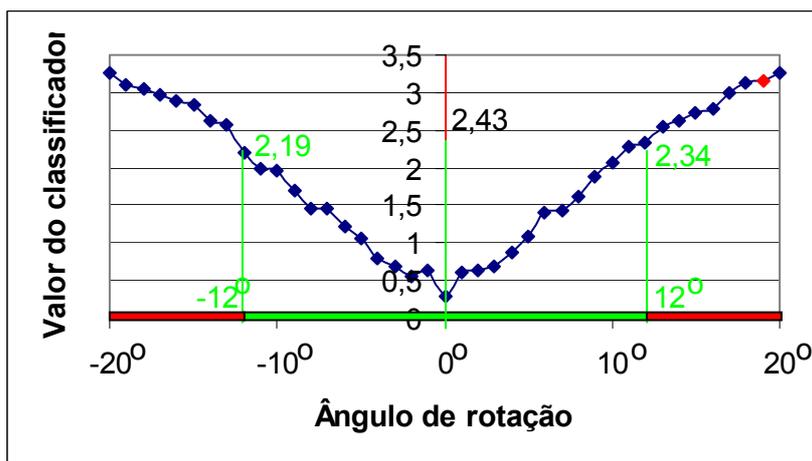


Figura 61 – Resultado da rotação da forma *Triangular Invertida*

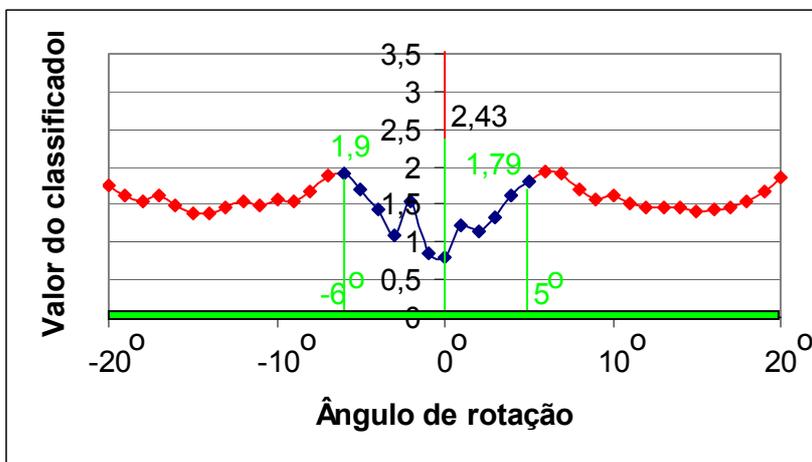


Figura 62 – Resultado da rotação da forma *Quadrangular*

4.2.3 Reconhecimento

O sistema possui para a fase de reconhecimento uma base de dados de sinais divididos pela sua forma (Circular, Triangular, Quadrangular e Triangular Invertida) com as cores vermelha e/ou azul. A informação existente na base de dados é constituída por 38 imagens em níveis de cinzento dos sinais de trânsito passíveis de serem encontrados nas 172 imagens que servem de teste ao sistema, normalizados para 64x64 pixels (Figura 63).



Figura 63 – Base de dados de sinais para reconhecimento

Após a classificação das regiões por forma, procedeu-se ao seu reconhecimento. Esse reconhecimento é efectuado utilizando a correlação normalizada de níveis de cinzento com cada um dos sinais da base de dados com a mesma forma da região analisada. Para minimizar erros na correlação, devidos a pequenas translações em x e/ou y resultantes da fase de detecção, introduziu-se uma pequena tolerância no deslocamento, de +/- 2 pixels em x e y. Daqui resultam 25 correlações, sendo aceite aquela que obtiver melhor resultado. Dos testes efectuados obtiveram-se os resultados apresentados na Tabela 15:

Classificados por forma correctamente	215
Mal reconhecidos	37
Reconhecidos correctamente	178
Taxa de erro	17,21%
Taxa de acerto	82,79%

Tabela 15 – Resultados do reconhecimento dos sinais

A taxa de erro obtida deve-se em parte aos seguintes factores:

- Existência de sinais com ideogramas muito semelhantes que a partir de determinada distância se torna muito difícil a sua distinção (Figura 64);



Figura 64 – Sinais com ideogramas semelhantes

- Existência de sinais com o mesmo significado mas com grafismo diferente (Figura 65).



Figura 65 – Sinais com grafismo variável

Apesar dos factores apresentados limitarem em parte o desempenho do sistema no reconhecimento dos sinais de trânsito, este mostra no entanto grande robustez quando na presença de sinais sujeitos a variações na iluminação que incide sobre eles (Figura 66).



Figura 66 – Reconhecimento de sinais com sombras

4.2.4 Seguimento

Um dos testes efectuados à metodologia de seguimento encontra-se apresentado na Figura 67. Esta sequência é constituída por 11 imagens de 320x240 pixels processada pelo sistema implementado. Na sequência, aparece o único sinal existente e detectado em cada

imagem. No canto superior esquerdo de cada imagem encontra-se representada a imagem normalizada em 64x64 pixels correspondente à detecção.

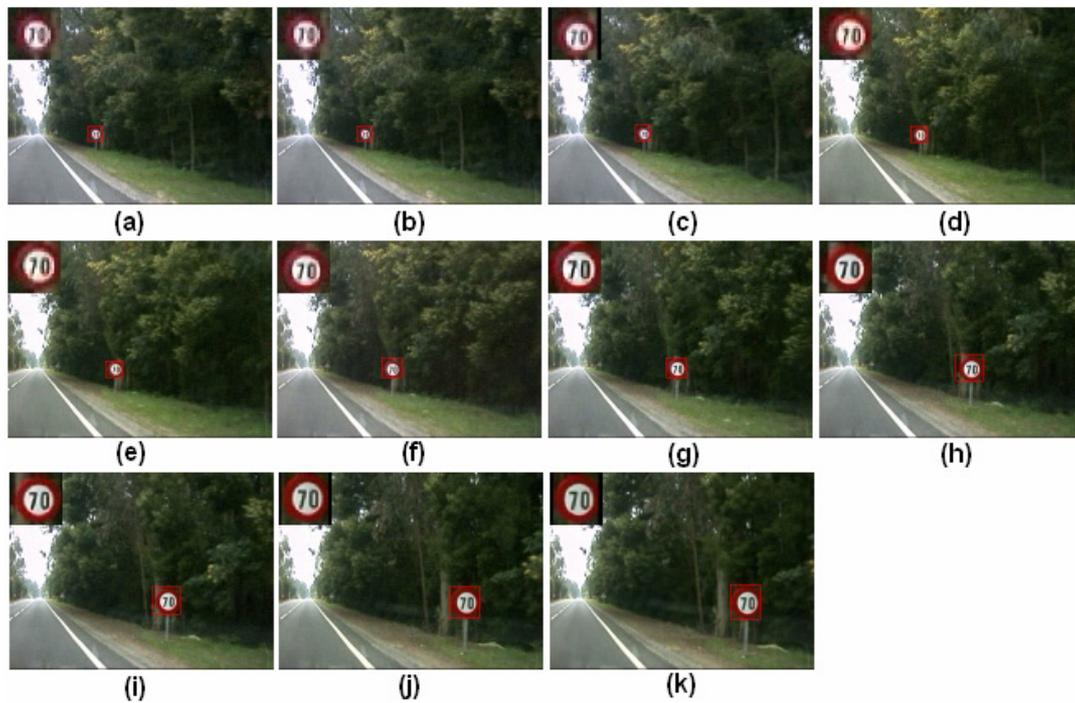


Figura 67 – Detecção de um sinal de trânsito (a)-(k) numa sequência de imagens

5. Conclusão

O trabalho desenvolvido nesta dissertação teve como factor principal de motivação o facto de se tratar do desenvolvimento de um módulo de processamento de imagem com aplicação prática tornando-o parte integrante de um sistema mais global. No entanto este factor de aplicabilidade prática limitou de alguma forma o leque de opções e abordagens passíveis de serem adoptadas devido quer a restrições temporais de desenvolvimento quer à necessidade de cumprir certas necessidades do sistema para o qual se destina.

Um dos requisitos do SIGMA prendia-se com a necessidade de existir uma metodologia inserida no módulo de visão que efectuasse a localização semi automática de um qualquer objecto, após a sua definição por parte do utilizador numa primeira imagem. Esse requisito foi cumprido com o desenvolvimento de uma metodologia que efectua uma aproximação grosseira à posição do objecto com base na sua cor seguida de uma aproximação refinada com base na forma.

Pode existir no entanto alguma dificuldade na localização exacta do contorno do objecto na aproximação refinada, nomeadamente quando existe uma semelhança grande entre a coloração do objecto e o fundo da imagem que o envolve, originando problemas de segmentação da imagem. Este problema pode no entanto ser minimizado se futuramente se associar neste passo, à informação de forma também informação de cor do objecto, possivelmente utilizando outro espaço de cor que não o *RGB*.

O sistema SIGMA necessitava também de uma metodologia capaz de localizar automaticamente numa sequência de imagens objectos de interesse cujas características eram conhecidas previamente pelo sistema. Tendo em conta que a actividade principal do sistema SIGMA se prende com o levantamento da sinalização rodoviária optou-se por seleccionar as formas e as cores mais comuns de sinais de trânsito como as características de interesse para utilização na localização de objectos.

Assim sendo, a metodologia desenvolvida localiza os objectos de interesse, que neste caso são sinais de trânsito, nas sequências de imagens com base na sua informação de cor, e de forma. É de referir no entanto que esta limitação é imposta apenas como forma de validar os resultados. A flexibilidade da abordagem permite incluir qualquer tipo de informação de forma e cor que se desejar. Este facto poderá apenas ter como inconveniente a diminuição do desempenho do sistema em termos de velocidade de análise à medida que for aumentando a quantidade de objectos a reconhecer.

O levantamento da sinalização rodoviária só está no entanto completa se associada à informação de posição (localização) do sinal for também efectuada a sua identificação (reconhecimento). Para esse efeito foi criada uma base de dados, com imagens *RGB* normalizadas, dos sinais a identificar separados em classes pela sua forma e comparados

com o sinal alvo. Este procedimento exige alguma qualidade das imagens, tanto dos sinais da base de dados como dos sinais alvo.

Apesar de neste trabalho e nesta fase de desenvolvimento do módulo de processamento de imagem do sistema SIGMA não ter havido a preocupação nem a necessidade de cumprir requisitos de tempo real, houve no entanto essa preocupação na escolha das metodologias, permitindo assim com mais facilidade no futuro, com as devidas optimizações de processamento, o cumprimento desses requisitos. Além de que a crescente capacidade de processamento disponibilizada pelos computadores torna possível o processamento, em tempo real, de sequências de imagens, aplicando algoritmos cada vez mais complexos. Esta evolução permite a automatização de mais tarefas, mantendo ou melhorando o grau de fiabilidade do processamento.

Para demonstrar a viabilidade das metodologias propostas, foram efectuados vários testes dos quais se obtiveram, para a metodologia de localização de um objecto genérico, resultados aceitáveis, tendo em conta, por um lado, a escassez de informação do objecto utilizada pelo sistema e, por outro lado, a complexidade contida nas sequências de imagens, obtidas de ambiente aberto, às quais a metodologia foi submetida.

A segunda metodologia, que corresponde à localização automática de objectos, obteve bons resultados, apesar dos testes terem sido direccionados para a sinalização rodoviária e restringida à localização de quatro formas e duas cores em concreto. A metodologia foi submetida, tal como no caso anterior, a cenas em ambiente aberto, mais concretamente 172 imagens, das quais se observaram 238 sinais de trânsito em condições de serem localizados, e dos quais resultaram 90,3% detectados correctamente por cor e forma e destes 82,8% foram reconhecidos correctamente, apesar do algoritmo utilizado nesta fase de reconhecimento ter sido aplicado apenas como abordagem inicial. Esta metodologia foi também submetida a sequências de imagens em ambiente aberto de sinalização rodoviária, tendo obtido também aqui resultados animadores. Os resultados obtidos das metodologias desenvolvidas são encorajadores e um forte incentivo para continuar a apostar no seu melhoramento.

Trabalho futuro

Com o trabalho desenvolvido no âmbito da presente dissertação o módulo encontra-se em condições de ser explorado, através de testes em campo, já integrado no sistema SIGMA, assim como paralelamente a optimização e extensão dos algoritmos implementados.

Dos testes efectuados verificou-se que grande parte dos erros detectados estavam associados directa ou indirectamente a problemas de segmentação. Isto derivado do facto de ser utilizada apenas informação de cor e forma do objecto. Dado que a localização dos objectos é efectuada em sequências de imagens seria interessante futuramente proceder-se à inclusão também de informação de movimento na segmentação utilizando para isso por exemplo técnicas de cálculo de fluxo óptico ou da diferença entre imagens e comparar os resultados obtidos com os obtidos neste trabalho.

Apesar do sistema SIGMA efectuar o levantamento em vídeo e registo, maioritariamente de sinalização rodoviária, e por esse motivo ter sido direccionado o desenvolvimento das metodologias assim como os testes efectuados ao sistema para esta área, existem vários outros tipos de objectos também relevantes nesse levantamento, os quais são localizados neste momento de forma manual. Seria interessante portanto, no futuro, inserir no módulo a possibilidade de efectuar o levantamento (localização) automático desses objectos como sejam: tampas de saneamento, caixas automáticas multibanco ou mesmo postes de iluminação pública. No caso das tampas de saneamento essa localização poderia concretizar-se utilizando um algoritmo especializado na detecção de formas elípticas numa determinada região da imagem que correspondesse à zona de estrada, ou então utilizar a metodologia já implementada neste módulo, juntando a forma elíptica às quatro formas já pré-definidas e testar também assim o desempenho do sistema.

6. Referências

- [Ahmad03] A. Ahmad, D. Chen, S. Lee. *Robust Object Detection Using Cascade Filter in MPEG Videos*. Proc. 5th IEEE Int. Symposium on Multimedia Software Engineering (ISMSE'03), Taichung, Taiwan, Dezembro 2003.
- [Aoyag96] Y. Aoyagi, T. Asakura. *A Study on Traffic Sign Recognition in Scene Image Using Genetic Algorithms and Neural Networks*. Proc. 22nd IEEE Int. Conf. Industrial Electronics, Control, and Instrumentation, Taipei, Taiwan, Agosto 1996.
- [Araki98] S. Araki, T. Matsuoka, H. Takemura, N. Yokoya. *Real-time Tracking of Multiple Moving Objects in Moving Camera Image Sequences Using Robust Statistics*. Proc. 14th IAPR Int. Conf. on Pattern Recognition (14ICPR), Brisbane, Australia, Agosto 1998.
- [Arnou96] P. Arnoul, M. Viala, J. Guerin, M. Mergy. *Traffic signs localisation for highways inventory from a video camera on board a moving collection van*. Proc. '96 IEEE Intelligent Vehicles Symp., Tokyo, Japan, Setembro 1996.
- [Bertal99] M. Bertalmío, G. Sapiro, G. Randall. *Morphing Active Contours*. IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 22, no. 7, pp. 733-737, Julho 2000.
- [Betke95] M. Betke, N. Makris. *Fast Object Recognition in Noisy Images Using Simulated Annealing*. Proc. 5th IEEE Int. Conf. Computer Vision, Cambridge, USA, Junho 1995.
- [Carro94] T. Carron, P. Lambert. *Color edge detector using jointly hue, saturation and intensity*. Proc. 1st IEEE Int. Conf. on Image Processing, Austin, USA, Novembro 1994.
- [Casse95] V. Casseles, R. Kimmel, and G. Sapiro. *Geodesic active contours*. Proc. '95 Int. Conf. Computer Vision, Boston, USA, Junho 1995.
- [Casse96] V. Casseles, B. Coll. *Snakes in Movement*. SIAM J. Numerical Analysis, vol. 33, pp. 2.445-2.456, 1996.
- [Curwe92] R. Curwen, A. Blake. *Dynamic Contours: Real-Time Active Splines*. *Active Vision*, A. Blake and A. Yuille, eds., pp. 39-58. MIT Press, 1992.
- [Escal04] A. Escalera, J. Armingol, J. Pastor, F. Rodríguez. *Visual Sign Information Extraction and Identification by Deformable Models for Intelligent Vehicles*. IEEE Transactions on Intelligent Transportation Systems, vol. 5, no. 2, pp. 57-68, Junho 2004.
- [Falou94] C. Faloutsos, R. Barber, M. Flickner, J. Hafner, W. Niblack, D. Petkovic, W. Equitz. *Efficient and Effective Querying by Image Content*. J. Intelligent Information Systems, vol. 3, no. 1, pp. 231-262, 1994.

- [Fiegu97] P. Fieguth, D. Terzopoulos. *Color-Based Tracking of Heads and Other Mobile Objects at Video Frame Rates*. Proc. '97 Conference on Computer Vision and Pattern Recognition (CVPR '97), San Juan, Puerto Rico, Junho 1997.
- [Golde01] R. Goldenberg, R. Kimmel, E. Rivlin, M. Rudzsky. *Fast Geodesic Active Contours*. IEEE Trans. on Image Processing, vol. 10, no. 10, pp. 1467-1475, Outubro 2001.
- [Gonza02] R. Gonzalez, R. Woods. *Digital Image Processing*, 2nd edition, Prentice Hall, 2002.
- [Grigo03] C. Grigorescu, N. Petkov. *Distance Sets for Shape Filters and Shape Recognition*. IEEE Transactions on Image Processing, vol. 12, no. 10, pp. 1274-1286, Outubro 2003.
- [Grove98] T. Grove, K. Baker, T. TAN. *Colour Based Object Tracking*. In 14th International Conference on Pattern Recognition (ICPR98), Abril 1998.
- [Hager98] G. Hager, P. Belhumeur. *Efficient Region Tracking With Parametric Models of Geometry and Illumination*. IEEE Trans. on Pattern Analysis and Machine Intelligence, vol. 20, no. 10, pp. 1025-1039, Outubro 1998.
- [He02] C. He, Y. Zheng, S. Ahalt. *Object Tracking Using the Gabor Wavelet Transform and the Golden Section Algorithm*, IEEE Trans. on Multimedia, vol. 4, no. 4, pp. 528-538, Dezembro 2002.
- [Hibi96] T. Hibi. *Vision based extraction and recognition of road sign region from natural color image, by using HSL and coordinates transformation*. Proc. 29th Int. Symp. Automotive Technology and Automation, Florence, Italy, Junho 1996.
- [Jolly96] M. Jolly, S. Lakshmanan, A. Jain. *Vehicle Segmentation and Classification using Deformable Templates*. IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 18, no. 3, pp. 293-308, Março 1996.
- [Jones99] R. Jones, D. DeMenthon, D. Doermann. *Building mosaics from video using MPEG motion vectors*. Proc. 7th ACM International Conference on Multimedia, Orlando, USA, Julho 1999.
- [Kass87] M. Kass, A. Witkin, D. Terzopoulos. *Snakes: Active Contour Models*. Int'l J. Computer Vision, vol. 1, no. 4, pp. 321-331, 1987.
- [Khan01] J. Khan, Z. Guo, W. Oh. *Motion based Object Tracking in MPEG-2 Stream for Perceptual Region Discriminating Rate Transcoding*. Proc. 9th ACM International Conference on Multimedia, Ottawa, Canada, 2001.
- [Kosch02] A. Koschan, S. Kang, J. Paik, B. Abidi, M. Abidi. *Video object tracking based on extended active shape models with color information*, Proc. 1st European Conf. Color in Graphics, Imaging, Vision, University of Poitiers, France, Abril 2002.
- [Kratt94] W. Krattenthaler, K. Mayer, M. Zeiller. *Point Correlation: A Reduced-Cost Template Matching Technique*. Proc. 1st International Conference on Image Processing, Austin, USA, Novembro 1994.

- [Ledle90] R. Ledley, M. Buas, T. Golab. *Fundamentals of True-Color Image Processing*. Proc. 10th IEEE Int. Conf. on Pattern Recognition, Atlantic City, USA, Junho 1990.
- [Lin97] Y. Lin, Y. Chang. *Tracking Deformable Objects with the Active Contour Model*. Proc. '97 IEEE Int. Conf. on Multimedia Computing and Systems, Ottawa, Canada, Junho 1997.
- [Logic04] *Logicol S.r.l. Software EasyRGB-PC*. Algoritmo de conversão de cor RGB-HSL, <http://www.easyrgb.com/math.html>, 2004.
- [LuoH00] H. Luo, Q. Lu, R. Acharya, R. Gaborski. *Robust Snake Model*. IEEE Conf. on Computer Vision and Pattern Recognition, vol. 1, pp. 452-457, Hilton Head Island, USA, Junho 2000.
- [LuoR92] R. Luo, H. Potlapalli, D. Hislop. *Neural network based landmark recognition for robot navigation*. Proc. '92 IEEE Int. Conf. Industrial Electronics, Control, Instrumentation, Automation, Power Electronics Motion Control, San Diego, USA, Novembro 1992.
- [Matsu99] Y. Matsuzawa, T. Abe. *Region Extraction Using Competition of Multiple Active Contour Models*. Proc. '99 IEEE Int. Conf. on Image Processing, Kobe, Japan, Outubro 1999.
- [Metax93] D. Metaxas, D. Terzopoulos. *Shape and Nonrigid Motion Estimation through Physics-Based Synthesis*. IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 15, no. 6, pp. 580-591, Junho 1993.
- [Miura02] J. Miura, T. Kanda, S. Nakatani, Y. Shirai. *An active vision system for on-line traffic sign recognition*. IEICE Trans. Inform. Syst., vol. E85-D, no. 11, pp. 1784-1792, 2002.
- [Pacli99] P. Paclik, J. Novovicova, P. Pudil, P. Somol. *Road Sign Classification using the Laplace Kernel Classifier*. Proc. 11th Scandinavian Conf. on Image Analysis, Kangerlussuaq, Greenland, 1999.
- [Parag99] N. Paragios, G. Tziritas. *Adaptive detection and localization of moving objects in image sequences*. Signal Processing: Image Comm., vol. 14, no. 4, pp. 277-296, Setembro 1999.
- [Parag00] N. Paragios, R. Deriche. *Geodesic Active Contours and Level Sets for the Detection and Tracking of Moving Objects*. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 22, no. 3, pp. 266-280, Março 2000.
- [Peixo02] P. Peixoto, J. Gonçalves, H. Araújo. *Real-Time Gesture Recognition System Based on Contour Signatures*. Proc. 16th IEEE Int. Conf. on Pattern Recognition, Agosto 2002.
- [Pentl94] A. Pentland, R. Picard, S. Sclaroff. *Photobook: Tools for Content-Based Manipulation of Image Databases. Storage and Retrieval of Image and Video Databases II*, no. 2185-05, pp. 34-47, San Jose, USA, Fevereiro 1994.

- [Peter99] N. Peterfreund. *Robust Tracking of Position and Velocity With Kalman Snakes*. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 21, no. 6, pp. 564-569, Junho 1999.
- [Picci96] G. Piccioli, E. Micheli, P. Parodia, M. Campani. *Robust method for road sign detection and recognition*. Image Vis. Comput., vol.14, no. 3, pp. 209-223, Abril 1996.
- [Pless00] R. Pless, T. Brodsky, Y. Aloimonos. *Detecting Independent Motion: The Statistics of Temporal Continuity*. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 22, no. 8, pp. 768-773, Agosto 2000.
- [Schoe01] T. Schoepflin, V. Chalana, D. Haynor, Y. Kim. *Video Object Tracking With a Sequential Hierarchy of Template Deformations*, IEEE Transactions on Circuits and Systems for Video Technology, vol. 11, no. 11, pp. 1171-1182, Novembro 2001.
- [Sclar01] S. Sclaroff, L. Liu. *Deformable Shape Detection and Description via Model-Based Region Grouping*. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 23, no. 5, pp. 475-489, Maio 2001.
- [Singh03] S. Singh, D. Chauhan, M. Vatsa, R. Singh. *A Robust Skin Color Based Face Detection Algorithm*. Tamkang Journal of Science and Engineering, vol. 6, no. 4, pp. 227-234, 2003.
- [Vinod97] V. Vinod, H. Murase. *Video Shot Analysis using Efficient Multiple Object Tracking*. Proc. '97 IEEE Int. Conf. on Multimedia Computing and Systems, Ottawa, Canada, Junho 1997.
- [Welch91] E. Welch, R. Moorhead, J. Owens. *Image processing using the HSI color space*. Proc. '91 IEEE Southeastcon, Williamsburg, USA, Abril 1991.
- [Won02] C. Won, K. Pyun, R. Gray. *Automatic Object Segmentation in Images with Low Depth of Field*. Proc. '02 IEEE Int. Conf. on Image Processing, Junho 2002.
- [Yan03] H. Yan, T. Tjahjadi. *Optical Flow Estimation and Segmentation through Surface Fitting and Robust Statistics*. IEEE '03 Int. Conf. on Systems, Man and Cybernetics, Outubro 2003.
- [Zhong00] Y. Zhong, A. Jain, M. Jolly. *Object Tracking Using Deformable Templates*. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 22, no. 5, pp. 544-549, Maio 2000.