



Pedro Miguel Ferreira Gomes

*Licenciado em Ciências de
Engenharia Electrotécnica e de Computadores*

Sistema de Detecção de Fogo e Detecção e Seguimento de Objectos

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

Orientador: José António Barata de Oliveira,
Professor Auxiliar, FCT-UNL

Co-orientador: Pedro Figueiredo Santana,
Professor Auxiliar, ISCTE-IUL

Júri:

Presidente: Prof. Doutor Luís Filipe dos Santos Gomes
Vogais: Prof. Doutor José Manuel Matos Ribeiro da Fonseca
Prof. Doutor José António Barata de Oliveira
Prof. Doutor Pedro Figueiredo Santana

Copyright

Sistema de Detecção de Fogo e Detecção e Seguimento de Objectos

A Faculdade de Ciências e Tecnologia e a Universidade Nova de Lisboa têm o direito, perpétuo e sem limites geográficos, de arquivar e publicar esta dissertação através de exemplares impressos reproduzidos em papel ou de forma digital, ou por qualquer outro meio conhecido ou que venha a ser inventado, e de a divulgar através de repositórios científicos e de admitir a sua cópia e distribuição com objectivos educacionais ou de investigação, não comerciais, desde que seja dado crédito ao autor e editor.

Dedico esta dissertação à memória da minha Mãe.

Agradecimentos

A elaboração de uma dissertação e a obtenção do respectivo grau de mestre é um trabalho bastante exigente. No entanto, ao longo destes últimos seis anos estive rodeado de colegas e amigos que me auxiliaram na conclusão deste processo. Quero por isso expressar o meu agradecimento a todos, e em especial aos que seguidamente menciono.

Ao Prof. José Barata o meu agradecimento pelo suporte prestado e pela oportunidade de participar num projecto ambicioso no domínio da robótica. Ao meu co-orientador Prof. Pedro Santana uma palavra de especial apreço pelo seu apoio, motivação e, essencialmente, uma partilha de experiência e conhecimento científico. Um agradecimento pela sempre boa disposição da equipa de investigação, Gonçalo Cândido, Eduardo Pinto, Magno Guedes, Pedro Deusdado, Giovanni di Orio, Francisco Marques e André Caves.

Agradeço também a todos os colegas da Faculdade de Ciências e Tecnologia da Universidade Nova de Lisboa que ao longo destes anos se tornaram amigos, destacando Bruno Valente, Carlos Carvalhas, Fábio Júlio, Fábio Alves, Flávio Dinis, João Santos, Miguel Marques, Ricardo Mendonça e Tiago Xavier. Um agradecimento especial ao colega de laboratório e amigo, Ricardo Mendonça, pelo *feedback* e apoio prestados ao longo dos últimos dois anos. Desejo-lhe os melhores sucessos pessoais e profissionais.

Um grande apreço e gratidão à minha família e em especial aos meus pais, pelo apoio, motivação, suporte e carinho prestado em todas as etapas da minha vida. Uma palavra especial ao meu irmão, João Gomes, pelo incentivo e força transmitida.

Um agradecimento singular à minha namorada, Raquel Custódio, pela paciência, amizade e especial carinho demonstrado ao longo destes anos.

A todos os que mencionei e aos que injustamente não referi, o meu muito obrigado.

Abstract

This dissertation presents a system for early fire detection and object tracking in live video sequences obtained from fixed outdoor surveillance cameras. Focus is given to the challenges related to the actual deployment of the vision system. Namely, background subtraction, which is a key to determine which regions of the environment belong to the foreground, is performed in a windowed way for improved accuracy. To reduce the computational cost, an attentive mechanism is employed to focus a computationally expensive frequency analysis of potential fire regions. To promptly discriminate fire regions from fire colored moving objects, a new colour-based model of fire's appearance and a new Wavelet-based model of fire's frequency signature are proposed. Besides that, to reduce the false alarms on fire detection of moving object with fire-colored appearance, an innovative solution to integrate the results of the two algorithms is proposed. Namely, the movement of the tracked object on the environment is analyzed. In addition, camera-world mapping is approximated according to a GPS-based learning calibration process to generate geo-located alarms, and to estimate the object height in the image plane. Experimental results demonstrate the ability of the proposed model to robustly detect fires and track moving objects, even in the presence of severe occlusions. Concretely, an average success rate of 92.7 % to detect fire and 92.8 % to tracking objects at a processing frequency of 10 Hz shows the applicability of the model to real-life applications.

Keywords: fire detection, *Wavelet* transform, tracking, particle filters

Resumo

Esta dissertação apresenta um sistema para uma rápida detecção de fogo e detecção e seguimento de objectos em tempo real, que recorre a sequências de imagens obtidas a partir de câmaras de vigilância fixas em ambientes exteriores. Os desafios relacionados com o desenvolvimento do sistema de visão constituem o foco desta dissertação. Nomeadamente, a subtracção do plano de fundo, que é uma chave para determinar quais as regiões do ambiente que pertencem ao primeiro plano, é realizada em forma de grelha para melhorar a precisão. Para reduzir o custo computacional, um mecanismo de atenção é utilizado para focar uma análise da frequência nas potenciais regiões de fogo, a qual é computacionalmente pesada. Para uma rápida diferenciação entre regiões com fogo e objectos em movimento com cores semelhantes às das chamas, um modelo de cor baseado na aparência do fogo e um modelo de análise da assinatura do fogo na frequência através da transformada Wavelet são propostos. Para reduzir os falsos alarmes na detecção de fogo provenientes de objectos em movimento com aparência semelhante ao fogo, uma solução inovadora de integração dos resultados dos dois algoritmos é proposta. Para tal, o movimento dos objectos a serem seguidos é analisado. Para além disso, um mapeamento do plano da imagem é aproximado através de um processo de calibração baseado na aprendizagem de coordenadas *GPS*, permitindo gerar alarmes geo-localizados e estimar a altura dos objectos no plano da imagem. Resultados experimentais demonstram a capacidade do modelo proposto para uma detecção robusta de fogo e seguimento de objectos, mesmo na presença de oclusões. Concretamente, uma taxa média de sucesso de 92.7 % para detectar fogo e 92.8 % para seguimento de objectos a uma taxa de processamento de 10 Hz mostra a aplicabilidade do modelo para aplicações em tempo real.

Palavras-Chave: detecção de fogo, transformada *wavelet*, seguimento de objectos, filtro de partículas

Conteúdo

Lista de Figuras	xv
Lista de Tabelas	xix
1 Introdução	1
1.1 Motivação e Visão Geral	1
1.2 Objectivos	2
1.3 Solução Proposta	2
1.4 Estrutura	3
2 Revisão da Literatura	5
2.1 Detecção de Fogo	5
2.2 Detecção e Seguimento de Objectos	8
2.2.1 Detecção de Regiões de Movimento	8
2.2.2 Seguimento de Objectos	11
2.2.2.1 Oclusões Entre Objectos	13
2.2.2.2 Múltiplas Câmaras	13
3 Metodologia Proposta	15
3.1 Detecção de Regiões de Movimento	16
3.2 Algoritmo de Detecção de Fogo	18
3.2.1 Detecção de Regiões de Píxeis Com Base na Cor do Fogo	18
3.2.1.1 Métodos de Classificação dos Píxeis Com Base na Cor	19
3.2.1.2 Análise do Desempenho dos Métodos de Classificação dos Píxeis com Base na Cor	22
3.2.2 Detecção de Texturas Dinâmicas	25
3.2.2.1 Segmentação das Regiões de Fogo	26

3.2.3	Análise no Domínio da Frequência	29
3.2.3.1	Transformada <i>Wavelet</i> Discreta no Espaço da Frequência Temporal	30
3.2.3.2	Transformada <i>Wavelet</i> Discreta no Espaço da Frequência Espacial	34
3.3	Algoritmo de Detecção e Seguimento de Objectos	36
3.3.1	Subtracção do Plano de Fundo	37
3.3.2	Detecção de Objectos e Remoção de Sombras	40
3.3.3	Seguimento de Múltiplos Objectos	43
3.3.3.1	Filtro de Partículas	43
3.3.3.2	Estado do seguimento	48
3.3.3.3	Classificação dos Tipos de Objectos	51
3.4	Calibração do Sistema	52
3.5	Integração dos Algoritmos e Geração de Alarmes	54
3.5.1	Integração dos Algoritmos	54
3.5.2	Geração de Alarmes	56
4	Resultados Experimentais	59
4.1	Instalação Experimental	59
4.2	Resultados da Detecção de Fogo	60
4.3	Resultados da Detecção e Seguimento de Objectos	62
5	Conclusões e Trabalhos Futuros	65
5.1	Conclusões	65
5.2	Trabalhos Futuros	66
5.3	Disseminação	66
	Bibliografia	69

Lista de Figuras

2.1	Estrutura típica dos algoritmos de detecção de fogo baseados em sequências de imagens.	5
2.2	Visão global dos espaços de cores <i>HSI</i> , <i>CIE L*a*b</i> e <i>YCbCr</i>	6
2.3	Elementos distractivos na detecção de fogo com base na aparência do fogo	6
2.4	Métodos de análise das características dinâmicas conhecidas do fogo no domínio da frequência	8
2.5	Estrutura típica de um algoritmo de detecção e seguimento de objectos	8
2.6	Visão global dos métodos de detecção de regiões de movimento	10
2.7	Exemplos de aplicação da técnica de fluxo óptico	10
2.8	Exemplo do sistema em funcionamento, no qual são detectadas as regiões do corpo humano	11
3.1	Diagrama do sistema proposto	16
3.2	Detecção de regiões de movimento através do método para estimação do plano de fundo	17
3.3	Variação do limiar de diferenciação do método para estimação do plano de fundo	18
3.4	Classificação dos píxeis com base na cor através do espaço de cores <i>RGB</i>	20
3.5	Classificação dos píxeis com base na cor através do espaço de cores <i>YCbCr</i>	21
3.6	Classificação dos píxeis com base na cor através do espaço de cores <i>HSI</i>	22
3.7	Segmentação manual das regiões de chamas	23
3.8	Resultados obtidos para a detecção de fogo com base em 5 modelos de cor	24
3.9	Resultados obtidos para a detecção de fogo com base em 5 modelos de cor com elementos distractivos no ambiente	25
3.10	Detecção de texturas dinâmicas	26
3.11	Processo de remoção de contornos	27
3.12	Localização das regiões de fogo	27

3.13	Intersecção entre as caixas delimitadoras de duas regiões detectados	28
3.14	Estrutura básica de um banco de filtros com dois níveis de transformação	31
3.15	Sinais nos domínios do tempo e frequência para análise da transformada <i>wavelet</i> discreta no espaço da frequência temporal	31
3.16	Resultado da transformada <i>wavelet</i> discreta no espaço da frequência temporal	32
3.17	Sub-imagens de decomposição resultantes da <i>DWT</i> bidimensional com três níveis de transformação	34
3.18	Resultado da análise da transformada <i>wavelet</i> discreta no espaço da frequência espacial	35
3.19	Forma de onda de um pixel ao longo do tempo	37
3.20	Visão global do modelo do plano de fundo	38
3.21	Deteccção de regiões de movimento através da subtracção do plano de fundo	39
3.22	Aprendizagem do modelo do plano de fundo	39
3.23	Visão global da detecccção de objectos	41
3.24	Visão global do método de remoção de sombras	42
3.25	Resultado do método de remoção de sombras	43
3.26	Etapas de execução de um filtro de partículas	44
3.27	Representação das 4 dimensões das partículas	46
3.28	Máscaras de movimento utilizadas para criação e actualização dos filtros de partículas	47
3.29	Localização dos objectos a serem seguidos pelo sistema	47
3.30	Resultado do algoritmo de detecccção e seguimento de objectos	48
3.31	Intersecção de objectos em movimento a serem seguidos pelo sistema	49
3.32	Oclusão entre um objecto em movimento e um objecto fixo no ambiente	50
3.33	Objectos em movimento a abandonar o campo de visão da câmara	51
3.34	Classificação do tipo de objecto a ser seguido pelo sistema	51
3.35	Calibração do sistema	53
3.36	Definição dos polígonos de calibração do sistema	54
3.37	Resultado da integração dos resultados dos algoritmos de detecccção de fogo, detecccção e seguimento de objectos	55
3.38	Caminho percorrido por objectos seguidos pelo sistema	57
4.1	Exemplos de imagens do bancos de dados com a saída do algoritmo de detecccção de fogo sobreposta	60
4.2	Exemplos de classificação e respectivas avaliações em termos da presença e ausência de fogo	61

4.3	Exemplos de imagens do bancos de dados com a saída do algoritmo de detecção e seguimento de objectos sobreposta	62
4.4	Exemplos de categorização para obtenção dos resultados experimentais do algoritmo de detecção e seguimento de objectos	63
4.5	Resultado do método de estimação do plano de fundo quando um objecto está longe da câmara.	64

Lista de Tabelas

3.1	Resultados experimentais da detecção de fogo através do método de classificação dos píxeis com base na cor (<i>MCC</i>)	23
3.2	Correspondência entre as coordenadas <i>GPS</i> e as coordenadas do objecto no plano da imagem	53
4.1	Resultados experimentais da detecção de fogo	61
4.2	Resultados experimentais da detecção e seguimento de objectos	63

Lista de Notações

ω	Valor máximo para a comparação de histogramas de cores
α	Constante de definição da velocidade de actualização das novas observações
β	Parâmetro de análise do sinal de saída da transformada <i>wavelet</i> discreta no espaço da frequência espacial
χ_i	Partícula i
γ	Valor médio do parâmetro β
κ	Intensidade média dos píxeis numa região
λ	Distância entre as arestas das caixas delimitadoras de duas regiões detectadas
λ_1	Constante de sintonização do filtro temporal recursivo
λ_2	Constante de sintonização do filtro temporal recursivo
B_n	Modelo do plano de fundo na imagem n
F_{sp}	Imagem que representa a análise do sinal de saída do banco de filtros da transformada <i>wavelet</i> discreta no espaço de frequência temporal
H_{calib}	Imagem de calibração da altura esperada dos objectos
I_s	Imagem actual
I_s^n	Imagem do sistema na imagem n
I_{BS}	Máscara resultante do método de subtracção do plano de fundo
I_f	Máscara com as caixas delimitadoras das regiões de fogo detectadas
I_{HBE}	Máscara resultante do método para estimação do plano de fundo
I_{PC}	Máscara resultante do método de classificação dos píxeis com base na cor
I_{PF}	Máscara com as caixas delimitadoras dos objectos a serem seguidos
I_{RC}	Máscara resultante do processo de remoção de contornos
I_{SR}	Máscara resultante do método de remoção de sombras
I_{TD}	Máscara resultante do método de detecção de texturas dinâmicas
M_{PF}	Máscara de inicialização e actualização do filtro de partículas

T_n	Limiar de diferenciação na imagem n
x_{hh}	Sub-imagem alto-alto resultante da transformada <i>wavelet</i> discreta no espaço da frequência espacial
x_{hl}	Sub-imagem alto-baixo resultante da transformada <i>wavelet</i> discreta no espaço da frequência espacial
x_{lh}	Sub-imagem baixo-alto resultante da transformada <i>wavelet</i> discreta no espaço da frequência espacial
ν	Constante de sensibilidade do método de remoção de sombras
ρ	Número de píxeis em movimento na região do objecto
σ	Valor máximo dos picos do sinal de saída do banco de filtros da transformada <i>wavelet</i> discreta no espaço de frequência temporal
τ	Constante de sensibilidade da função densidade de probabilidade
τ_H	Constante de sensibilidade do método de remoção de sombras
τ_S	Constante de sensibilidade do método de remoção de sombras
θ	Parâmetro de conversão entre espaços de cores
v	Constante de sensibilidade do método de remoção de sombras
ε	Constante de sensibilidade do método de remoção de sombras
φ	Constante de sensibilidade de actualização de histogramas de cores
ϑ	Número de píxeis em movimento na região de cada partícula
ξ_1	Valor mínimo do parâmetro β
ξ_2	Valor máximo do parâmetro β
ζ	Constante de calibração da altura do objecto
A_n	Escalas de aproximação do sinal de saída do banco de filtros
B	Componente azul
b	Componente azul normalizado
C_f	Conjunto de píxeis considerados fogo
C_b	Componente crominância azul
$C_{b_{mean}}$	Valor médio da componente crominância azul
C_r	Componente crominância vermelho
$C_{r_{mean}}$	Valor médio da componente crominância vermelho
D_n	Escalas de detalhe do sinal de saída do banco de filtros
d_{battch}	Métrica de distância de Bhattacharyya

<i>DWT</i>	Transformada <i>wavelet</i> discreta
<i>dwt_{out}</i>	Sinal de saída do banco de filtros
<i>G</i>	Componente verde
<i>g</i>	Componente verde normalizado
<i>g_f</i>	Função de resposta ao impulso de um filtro passa-baixo
<i>GPS</i>	Sistema de posicionamento global
<i>H</i>	Componente matiz
<i>h₁</i>	Valor mínimo da componente matiz
<i>h₂</i>	Valor máximo da componente matiz
<i>H_b</i>	Banda de alta frequência resultante da transformada <i>wavelet</i> discreta no espaço da frequência espacial
<i>h_f</i>	Função de resposta ao impulso de um filtro passa-alto
<i>h_i</i>	Altura da caixa delimitadora do objecto <i>i</i>
<i>h_r^[n]</i>	Histograma de referência do objecto <i>n</i>
<i>h_{mp}^[n]</i>	Histograma de aparência da média das melhores partículas do objecto <i>n</i>
<i>h_{obj}</i>	Altura do objecto
<i>HBE</i>	Método de estimação do plano de fundo
<i>HOG</i>	Histograma de gradientes orientados
<i>hp_i</i>	Histograma de cor da partícula <i>i</i>
<i>HSI</i>	Espaço de cores
<i>HSV</i>	Espaço de cores
<i>HY</i>	Modelo de cor
<i>HYR</i>	Modelo de cor
<i>I</i>	Componente Intensidade
<i>i₁</i>	Valor mínimo da componente luminosidade
<i>i₂</i>	Valor máximo da componente luminosidade
<i>int_{pf}</i>	Distância máxima entre as arestas das caixas delimitadoras de dois objectos para se considerar que estes se interseccionam
<i>L_b</i>	Banda de baixa frequência
<i>M</i>	Conjunto de píxeis em movimento
<i>MCC</i>	Métrica de classificação de imagens binárias
<i>mov_{rpf}</i>	Número de localizações do objecto a ser seguido para representar o seu movimento

N_{dwt}	Número total de píxeis classificados como fogo
N_{initBS}	Sequência de imagens para aprendizagem do modelo do plano de fundo
N_{picos}	Número total de picos do sinal de saída do banco de filtros da transformada <i>wavelet</i> discreta no espaço de frequência temporal
NC_{dwt}	Número de imagens recolhidas para análise da transformada <i>wavelet</i> discreta
R	Componente vermelho
r	Componente vermelho normalizado
r_1	Ruído das partículas relacionado com largura do objecto
r_2	Ruído das partículas relacionado com altura do objecto
r_o	Ruído das partículas relacionado com a posição no plano da imagem do objecto
R_T	Valor global da componente vermelho
R_{pf}	Ruído das partículas para cada iteração do filtro de partículas
RGB	Espaço de cores
S	Componente saturação
s_1	Valor mínimo da componente saturação
s_2	Valor máximo da componente saturação
$STFT$	Transformada de <i>Fourier</i> de curto espaço de tempo
thr_{BS}	Número de píxeis em movimento numa região da máscara de subtracção do plano de fundo
thr_{int}	Limite inferior do rácio de píxeis classificados como fogo pela regra baseada na intensidade dos píxeis
thr_{mf}	Distância Euclidiana mínima para se considerar que um objecto não se encontra estático no ambiente
thr_{zc}	Limite inferior do número acumulado de passagens por zero do sinal de saída do banco de filtros
$update_{BS}$	Sequência de imagens para actualização do método de subtracção do plano de fundo
V	Componente valor
w_i	Largura da caixa delimitadora do objecto i
$w_i^{[n]}$	Peso da partícula i para o objecto n
WT	Transformada <i>wavelet</i>
x_f	Sinal original no domínio do tempo
x_i	Posição x da caixa delimitadora do objecto i
X_r	Incerteza atribuída ao movimento do objecto em relação à coordenada x da imagem

x_{obj}	Posição do objecto em relação à coordenada x da imagem
XML	Formato para criação de documentos
Y	Componente luminosidade
Y_1	Limite superior da curva de correlação entre componente azul e saturação
Y_2	Limite inferior da curva de correlação entre componente azul e saturação
y_i	Posição y da caixa delimitadora do objecto i
Y_r	Incerteza atribuída ao movimento do objecto em relação à coordenada y da imagem
Y_{alto}	Saída do filtro passa alto
Y_{baixo}	Saída do filtro passa baixo
Y_{mean}	Valor médio da componente luminosidade
y_{obj}	Posição do objecto em relação à coordenada y da imagem
y_{obs}	Probabilidade de observação
$YCbCr$	Espaço de cores

Capítulo 1

Introdução

1.1 Motivação e Visão Geral

A segurança das pessoas e dos seus bens é um tema de grande preocupação para a sociedade. Com recurso a sistemas de vídeo vigilância é possível obter índices de segurança e supervisão mais elevados, visto que estes permitem o controlo e monitorização de uma determinada área de forma permanente e exaustiva. Estes sistemas são tipicamente monitorizados por um conjunto de pessoas com recurso a monitores que exibem as imagens das câmaras em tempo real. Consoante o cenário em causa, é de extrema dificuldade para os operadores destes sistemas conseguirem detectar todos os eventos que ocorrem em tempo real.

Recorrendo a algoritmos de visão computacional é possível automatizar aplicações de vídeo vigilância, processando em tempo real os dados provenientes dos referidos sistemas, de forma a detectar situações anómalas e seguir objectos em movimento. Assim é possível gerar automaticamente diversos tipos de alarmes de modo a assistir os operadores do sistema.

A referida automatização permite, entre outros, detectar múltiplos focos de incêndio. Novas técnicas e estudos realizados recentemente demonstram que a detecção de fogo através de uma sequência de imagens pode ser uma alternativa fiável ou um complemento às técnicas existentes, tais como os dispositivos automáticos de detecção de calor, chama, fumo ou luz. No entanto, em ambientes não estruturados não é possível recorrer aos referidos dispositivos visto que estes apresentam diversos problemas, como sejam uma elevada taxa de falsos alarmes e uma necessidade de proximidade à chama.

Para além da detecção de fogo, a crescente necessidade de automatizar computacionalmente processos de vigilância permitiu o desenvolvimento de algoritmos com capacidade para seguir objectos em movimento (regiões de píxeis) ao longo de uma sequência de imagens. O seguimento de objectos isolados que permanecem no campo de visão da câmara é uma tarefa de elevada dificuldade e está dependente da complexidade dos objectos a seguir e do ambiente no qual estão inseridos. No entanto, as principais dificuldades para o sistema surgem quando é necessário lidar com oclusões temporárias entre objectos em movimento ou com objectos fixos no ambiente (e.g,

obstáculos), com a formação de grupos de objectos ou com o desaparecimento parcial dos mesmos. Nestes casos, não é possível seguir isoladamente os objectos e o sistema deverá ser capaz de os reconhecer quando estas abandonarem o estado de oclusão.

O volume de trabalhos publicados na área da segurança e vigilância automática e a evolução dos algoritmos de visão computacional conduz a sistemas cada vez mais complexos e sofisticados. No entanto, é possível projectar e implementar sistemas que incluam métodos mais eficientes e robustos na detecção de focos de incêndio e seguimento de múltiplos objectos.

1.2 Objectivos

O objectivo desta dissertação é projectar e implementar um sistema para mitigar algumas das limitações dos algoritmos de detecção de fogo e seguimento de objectos presentes na literatura. Como tal, é proposto um sistema que tem a capacidade de detectar múltiplos focos de incêndio e, em simultâneo, detectar e seguir objectos em tempo real, mesmo que estes entrem em oclusão com objectos em movimento ou com objectos presentes no ambiente no qual o sistema de vídeo vigilância está instalado. De referir que o sistema recorre a uma câmara fixa no ambiente para obter uma sequência de imagens.

Com a introdução do algoritmo de detecção de fogo no sistema proposto, pretende-se detectar múltiplos focos de incêndio de forma eficiente e robusta. Recorrendo a métodos que analisam as características dinâmicas conhecidas do fogo no domínio da frequência, é possível discriminar entre um fogo e um objecto em movimento com cores semelhantes às das chamas.

Com base no algoritmo de detecção e seguimento de objectos é possível seguir objectos em movimento em ambientes exteriores e com variação das condições dinâmicas do ambiente (e.g nuvens, luminosidade e sombras) no qual o sistema de vídeo vigilância está instalado, recorrendo a técnicas de extracção de regiões de movimento e a modelos estocásticos para executar o seguimento dos mesmos.

1.3 Solução Proposta

O sistema proposto nesta dissertação introduz um conjunto de inovações em relação aos algoritmos presentes na literatura, destacando-se:

- **Algoritmo de detecção de fogo:** com base no estudo da detecção de regiões de píxeis com cor de fogo, nesta dissertação é proposto um modelo de cor denominado *HY*. Para a discriminação entre uma região de fogo e um objecto em movimento com cores semelhantes às das chamas, o algoritmo recorre a métodos de análise das características dinâmicas conhecidas do fogo no domínio da frequência (Töreyn et al., 2006). Apesar de apresentar bons resultados, estes métodos são computacionalmente pesados. Com vista a colmatar esta limitação, é proposto a utilização de um filtro temporal recursivo para focar o algoritmo na

detecção de texturas dinâmicas. Caso este detecte uma constante presença de uma região de píxeis considerados fogo, a análise no domínio da frequência na referida região é aplicada. Foi também proposto uma extensão ao método de análise das texturas dinâmicas no domínio da frequência (Töreyn et al., 2006), aplicando uma transformada *wavelet* discreta com três níveis de transformação, de modo a tornar a detecção de fogo mais robusta.

- **Algoritmo de detecção e seguimento de objectos:** para segmentar as regiões de movimento na imagem actual são utilizados os métodos de subtracção do plano de fundo (Kim et al., 2005) e de estimação do plano de fundo (Collins et al., 2000). O método proposto por Kim et al. (2005) baseia-se na aprendizagem de um único modelo do plano de fundo, tornando o sistema demasiado sensível a alterações de luminosidade. Nesta dissertação é proposto uma extensão a este método, onde uma grelha regular é sobreposta à imagem actual, associando a cada posição da grelha um modelo do plano de fundo. Assim, é possível que cada modelo aprenda o plano de fundo independentemente do movimento existente noutra região da imagem.

Com determinadas condições de luminosidade, o resultado da detecção de movimento inclui as sombras dos objectos. Para que estas não afectem o processo de seguimento, surge a necessidade de as remover. Este processo é efectuada através de uma extensão ao método proposto por Cucchiara et al. (2003), onde apenas são removidas as sombras na região inferior dos objectos detectados.

O processo de seguimento de objectos recorre a filtros de partículas que se baseiam nas características de aparência dos objectos para estimar a posição dos mesmos. A forma de adicionar ruído aleatório às partículas foi alterado, de modo a actualizar as 4 dimensões do ruído das partículas com base nas últimas localizações do objecto em movimento e assim, na iteração seguinte do filtro, direccionar as partículas para a localização esperada do objecto.

Apesar de o algoritmo de detecção de fogo incluir métodos suficientemente robustos para diferenciar um fogo de um objecto em movimento com cores semelhantes às das chamas, poderão existir situações nas quais objectos com movimentos repetitivos afectem a robustez do algoritmo. Neste sentido, o sistema proposto nesta dissertação inclui uma solução inovadora ao integrar o resultado dos dois algoritmos, de forma a reduzir a taxa de falsos alarmes na detecção de fogo proveniente dos referidos objectos.

Para a validação do sistema proposto, os resultados experimentais foram obtidos a partir de um banco de dados de 15 vídeos. Os resultados mostram que o sistema detecta regiões de chamas em 92.7% das imagens testadas, detecta e segue objectos em movimento em 95.87% e 92.77% das imagens testadas, respectivamente.

1.4 Estrutura

Esta dissertação está organizada da seguinte forma:

Capítulo 2 Análise da literatura relacionada a detecção de fogo e detecção seguimento de múltiplos objectos;

Capítulo 3 Descrição do sistema proposto para detecção de fogo e detecção e seguimento de objectos em tempo real;

Capítulo 4 Apresentação da configuração experimental e os resultados obtidos a partir de um de banco de dados com 15 vídeos;

Capítulo 5 Apresentação de um conjunto de conclusões, principais contribuições da dissertação e oportunidades de investigação na área.

Capítulo 2

Revisão da Literatura

O sistema proposto nesta dissertação é baseado na integração dos algoritmos de detecção de fogo e detecção e seguimento de objectos em tempo real. Neste capítulo é revista a literatura relacionada com os referidos algoritmos, detalhada nas secções 2.1 e 2.2, respectivamente.

2.1 Detecção de Fogo

A análise da literatura permite concluir que a detecção de fogo através de algoritmos de visão computacional engloba três passos fundamentais (ver figura 2.1): (1) detecção de regiões de movimento; (2) detecção de regiões baseado na aparência e (3) análise das características dinâmicas conhecidas do fogo. Os dois primeiros passos permitem segmentar regiões com cores semelhantes às chamas que apresentam movimento. O terceiro passo é necessário para diminuir a taxa de falsos alarmes proveniente dos objectos em movimento com aparência semelhante ao fogo.

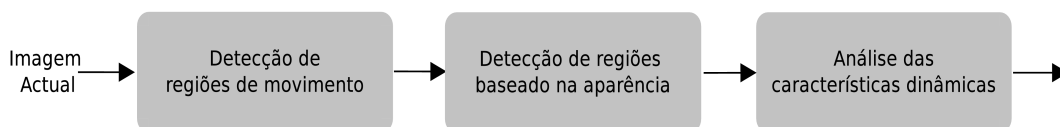


Figura 2.1: Estrutura típica dos algoritmos de detecção de fogo baseados em sequências de imagens.

Uma aproximação à detecção de fogo através de algoritmos de visão computacional consiste em classificar os píxeis da imagem de acordo com um modelo de aparência baseado no fogo (Chen et al., 2010). Esta aproximação, que se baseia na classificação dos píxeis com base na informação da cor, tem um papel fundamental nos algoritmos de detecção de fogo, pois os restantes métodos implementados estão dependentes da capacidade desta técnica para segmentar regiões de píxeis com cores semelhantes às das chamas.

Na literatura existe um conjunto de métodos de classificação dos píxeis com base na cor, fundamentados por regras predefinidas. O modelo proposto por Phillips III et al. (2002) utiliza a informação do espaço de cores *RGB* e a variação temporal dos píxeis para a detecção do fogo. Para tal, uma

tabela treinada é utilizada para extrair as regiões com cores semelhantes às das chamas. Chen et al. (2010) utiliza um algoritmo de filtragem dos píxeis, aplicando um conjunto de regras no espaço de cores *RGB*. Por outro lado, Töreyn et al. (2006) usa uma distribuição pré-determinada de cores no espaço de cores *RGB* para executar uma segmentação dos píxeis com base na cor.

De modo a especificar as intensidades de cor e brilho separadamente, são introduzidos outros espaços de cores no método de classificação dos píxeis com base na cor (ver figura 2.2). Os modelos propostos por Celik and Demirel (2009) e Celik (2010) utilizam os espaços de cores *YCbCr* e *CIE L*a*b* para construir um modelo genérico de segmentação dos píxeis através de um conjunto de regras. Por sua vez, Horng et al. (2005) utiliza uma transformação para o espaço de cores *HSI* e aplica o método de separação de cores para segmentar regiões com aparência semelhante ao fogo, removendo posteriormente os píxeis saturados e com pouca intensidade que não contenham chamas, tais como fumo e regiões com reflexões.

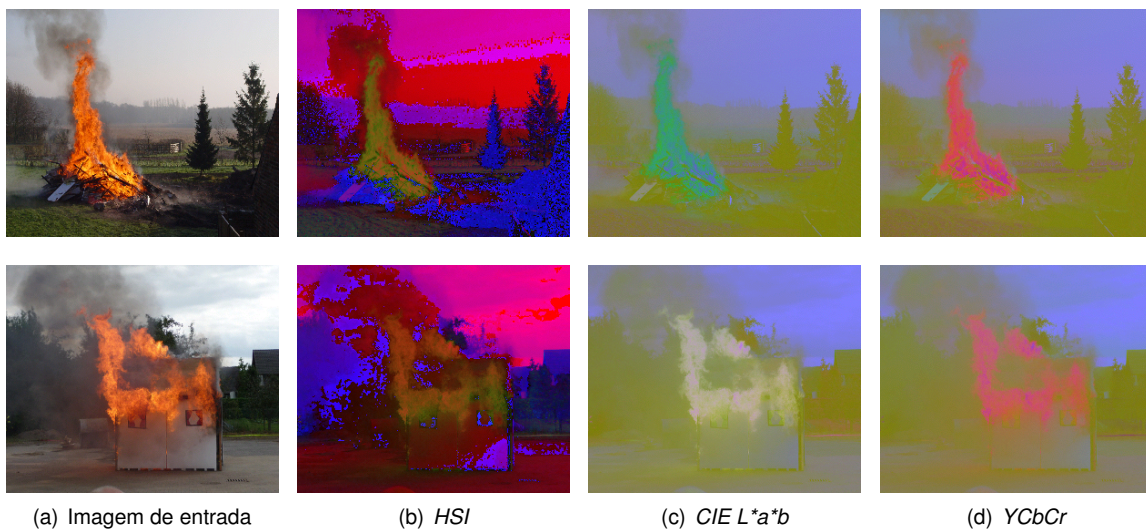


Figura 2.2: Visão global dos espaços de cores *HSI*, *CIE L*a*b* e *YCbCr*. (a) Imagem original no espaço de cores *RGB*; imagem convertida para o espaço de cores: *HSI* (Horng et al., 2005) (b), *CIE L*a*b* (Celik, 2010) (c) e *YCbCr* (Celik and Demirel, 2009) (d).

Embora os métodos de classificação dos píxeis com base na cor apresentem uma elevada robustez, estes estão sujeitos a um conjunto elevado de distractivos, isto é, objectos estáticos ou em movimento que apresentam cores semelhantes às das chamas (ver figura 2.3).



Figura 2.3: Elementos distractivos na detecção de fogo com base na aparência do fogo: areias (a), árvores (b), relvado (c) e armários (d).

Os métodos propostos por Celik and Demirel (2009), Horng et al. (2005), Chen et al. (2003) e Healey et al. (1993) baseiam-se apenas na análise das características de aparência do fogo para executar a detecção do mesmo. Estes classificam uma região como fogo quando esta apresenta

um número de píxeis com aparência semelhante ao fogo acima de um determinado limiar. Neste sentido, estão sujeitos a uma elevada taxa de falsos alarmes, pois qualquer objecto que apresente cores semelhantes às das chamas é classificado como fogo.

O fogo apresenta-se como um objecto estático com forma e textura dinâmicas. Neste sentido, é possível reduzir a taxa de falsos alarmes na detecção de fogo, eliminando regiões de píxeis que representam objectos estáticos presentes no ambiente. Os métodos propostos por Phillips III et al. (2002), Celik (2010), Celik et al. (2007, 2006) e Ko et al. (2009) incorporam a técnica de diferenciação temporal de imagens consecutivas para detectar regiões de movimento. Assim, com base nesta técnica e na detecção de regiões de píxeis através da aparência do fogo, apenas as regiões de movimento são classificados com base na informação de cor. Uma análise detalhada à detecção de regiões de movimento é apresentada na secção 2.2.1.

Uma das diversas características do fogo é a oscilação dos píxeis nas regiões circundantes às chamas (Detriche and Lanore, 1980; Drysdale, 2011; Hammis et al., 1992). Para aumentar a fiabilidade da detecção de fogo e diminuir a taxa de falsos alarmes provenientes de objectos em movimento com cores semelhantes às das chamas, podem ser aplicados métodos de análise das características dinâmicas conhecidas do fogo nos domínios do tempo e frequência.

Para classificar uma região como fogo, os algoritmos propostos por Chen et al. (2004, 2010) incluem métodos de análise das características dinâmicas conhecidas do fogo no domínio do tempo, propondo a análise da variação da forma das chamas e da variação de brilho dos píxeis, com base no número de píxeis considerados fogo pela técnica de classificação dos píxeis com base na cor que apresentam movimento.

A análise das características dinâmicas conhecidas do fogo no domínio da frequência acrescentam uma elevada robustez aos algoritmos de detecção de fogo, pois permitem analisar: (1) a oscilação dos píxeis na zona circundante às chamas (Töreyn et al., 2006; Toreyn et al., 2005; Chen et al., 2010, 2004), (2) a variação da textura das chamas (Töreyn et al., 2006; Toreyn et al., 2005) e (3) a variação da forma das chamas (Liu and Ahuja, 2004). É importante referir que os algoritmos de detecção de fogo que incluem métodos de análise no domínio da frequência apresentam uma taxa de falsos alarmes inferior na detecção de fogo em relação aos restantes algoritmos.

O algoritmo proposto por Töreyn et al. (2006) analisa as referidas características do fogo através da transformada *wavelet* discreta (*DWT*) no espaço da frequência temporal e espacial, de modo a avaliar o comportamento oscilatório dos píxeis nas zonas circundantes às chamas e a variação de textura das chamas nas regiões classificadas como fogo, respectivamente. Apesar de apresentar bons resultados para o conjunto de dados testado, os métodos implementados são pesados computacionalmente e o modelo de cores utilizado não segmenta convenientemente as chamas.

No método proposto por Liu and Ahuja (2004), os contornos das chamas são representados através dos coeficientes de *Fourier*. Para a discriminação entre uma região de fogo e um objecto em movimento com cores semelhantes às das chamas é analisada a variação destes coeficientes ao longo do tempo. No entanto, este método é limitado a chamas de grande dimensão.

Na figura 2.4 estão apresentados exemplos de análise de uma região com chamas no domínio da frequência, através dos métodos propostos por Töreyn et al. (2006) e Liu and Ahuja (2004).

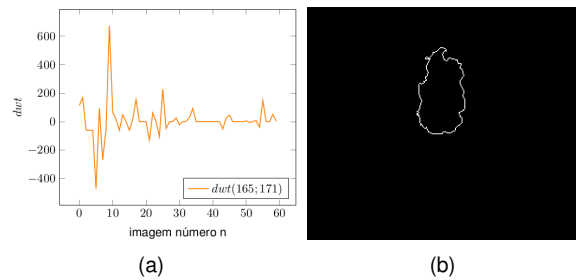


Figura 2.4: Métodos de análise das características dinâmicas conhecidas do fogo no domínio da frequência: frequência de oscilação de um pixel na região circundante às chamas (Töreyn et al., 2006) (a); contorno das chamas (Liu and Ahuja, 2004) (b).

2.2 Detecção e Seguimento de Objectos

Enquanto o objecto a ser seguido pelo sistema se encontra na perspectiva da câmara, este pode estar ocluído (parcialmente ou totalmente) por outros objectos ou obstáculos. Neste sentido, um algoritmo de detecção e seguimento de objectos deve ter a capacidade de seguir objectos e conseguir distingui-los após uma ou mais oclusões, mesmo com alterações da luminosidade e movimentos de outros objectos.

A análise da literatura permite concluir que o seguimento de objectos através de algoritmos de visão computacional inclui três passos fundamentais: segmentação, detecção e seguimento de objectos (ver figura 2.5). Os processos de segmentação e detecção são utilizados para extrair regiões de movimento e detectar objectos, através de técnicas de subtracção do plano de fundo ou diferenciação temporal. Para o processo de seguimento de objectos são utilizadas filtros estocásticos que garantem o seguimento de um determinado objecto, mesmo em situações de oclusões entre objectos ou obstáculos.



Figura 2.5: Estrutura típica de um algoritmo de detecção e seguimento de objectos.

De seguida, são analisadas um conjunto de técnicas de detecção regiões de movimento e técnicas de seguimento de objectos, descritas nas secções 2.2.1 e 2.2.2, respectivamente.

2.2.1 Detecção de Regiões de Movimento

O objectivo das técnicas de detecção de movimento é isolar um objecto ou parte deste a partir de uma imagem, sendo esta operação fundamental em aplicações de vídeo vigilância. A detecção de movimento através de sequências de imagens é um problema de difícil resolução, pois existem alterações no campo visual da câmara. Em ambientes exteriores, estas alterações estão relacionadas com objectos em movimento presentes no ambiente (e.g folhas das árvores e arbustos) e alterações dinâmicas da luminosidade, tais como nuvens e objectos que alteram a entrada de luz no ambiente.

Existem três aproximações principais para a detecção de objectos em movimento (Moeslund et al., 2006): (1) técnica de diferenciação temporal; (2) técnica de subtracção do plano de fundo e (3) técnica de fluxo óptico.

A técnica de diferenciação temporal consiste na subtracção directa de imagens consecutivas. Esta técnica é ideal para ambientes com grandes variações dinâmicas, apesar não extrair convenientemente as regiões de movimento e não detectar movimento quando os objectos se encontram a uma determinada distância da câmara. O método proposto por Lipton et al. (1998) detecta regiões de movimento através desta técnica e utiliza a análise de contornos para extrair as regiões de movimento. Por outro lado, Collins et al. (2000) propõe um método de detecção de regiões de movimento através de duas operações, diferenciação de três imagens consecutivas e subtracção do plano de fundo.

Para compensar os efeitos da alteração da luminosidade e dos movimentos repetitivos dos objectos num ambiente exterior, a técnica de subtracção do plano de fundo ajusta um modelo de séries temporais para cada pixel ou grupo de píxeis, sendo capaz de lidar com flutuações temporais, embora seja computacionalmente complexo (Toyama et al., 1999). Esta técnica baseia-se na representação do plano de fundo através de um modelo, onde as regiões de píxeis em movimento são determinadas através da comparação entre este e a imagem actual.

Entende-se como plano de fundo as regiões do ambiente que estão fixas, tais como estradas e edifícios. Por seu lado, qualquer elemento que se mova no ambiente da câmara ou um elemento que esteja estacionário durante um curto período de tempo (e.g. automóveis estacionados) são considerados elementos do primeiro plano. É possível que diversos elementos do plano de fundo possam apresentar movimento, tais como árvores ou ervas.

Os métodos propostos por Stauffer and Grimson (1999), Grimson et al. (1998), KaewTraKulPong and Bowden (2001), Friedman and Russell (1997) e Butler et al. (2003) utilizam uma mistura de distribuições gaussianas para construir um modelo do plano de fundo com base numa sequência de imagens. Para cada pixel da nova imagem, o seu valor é comparado com o modelo e, caso esteja dentro de um intervalo dado pelo desvio padrão, os parâmetros da gaussiana são ajustados. Caso contrário, a distribuição gaussiana que modela o pixel é substituída.

Javed et al. (2002) propõe a utilização da informação do gradiente no modelo do plano de fundo para reduzir o erro de classificação da informação de cor e para lidar com as mudanças de luminosidade. Com a alteração das condições de luminosidade, a informação de cor de cada pixel é alterada subitamente e segundo Javed et al. (2002), o gradiente do plano de fundo mantém-se estável durante a ocorrência dessas alterações.

Zivkovic and van der Heijden (2006) propuseram um modelo de plano de fundo que adapta automaticamente o número de distribuições gaussianas usadas para modelar um determinado pixel. No método proposto por Ridder et al. (1995), cada pixel é modelado através de um filtro de *Kalman*, com o objectivo de compensar a variação de iluminação no campo visual da câmara. Por outro lado, Kim et al. (2005) utiliza um vector de *codebooks* para construir um modelo que representa o plano de fundo, após a observação de uma sequência de imagens. Uma descrição detalhada deste método é apresentada na secção 3.3.1.

De referir que a escolha do espaço de cores é determinante para os métodos de subtracção do plano de fundo. O espaço de cores *RGB* permite medir a distorção do brilho e da cor. No entanto, para métodos que têm a capacidade de lidar com flutuações dinâmicas do ambiente, os espaços de cores *YCbCr* e *HSV* são mais adequados, pois permitem separar a cor e a intensidade de um determinado pixel. Para além do espaço de cores *RGB*, foram utilizadas outras representações para extrair o plano de fundo. Concretamente, Wren et al. (1997) propôs a utilização do espaço de cores *YUV* e Cucchiara et al. (2003) o espaço de cores *HSV*.

A figura 2.6 apresenta dois exemplos das técnicas de diferenciação temporal e subtracção do plano de fundo, aplicadas a imagens provenientes de sistemas de vídeo vigilância instalados em ambientes exteriores.

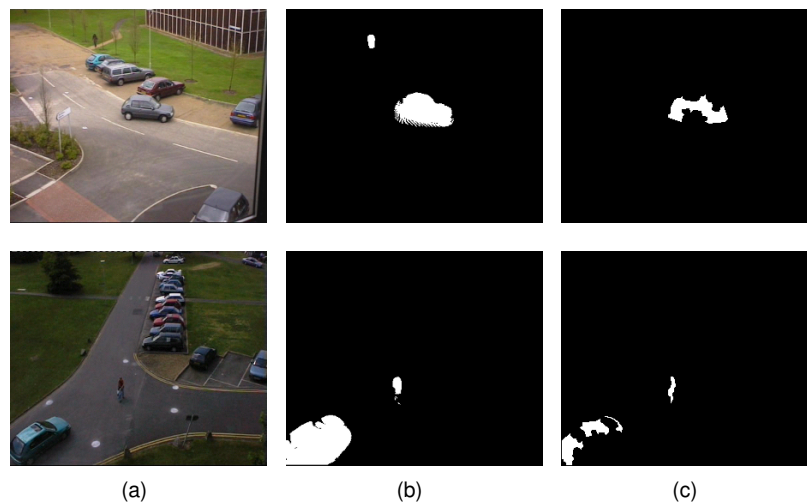


Figura 2.6: Visão global dos métodos de detecção de regiões de movimento. (a) imagem original; máscaras resultante da técnica de subtracção do plano de fundo (b) e diferenciação temporal (c).

Outra das formas de detectar objectos em movimento é utilizar a técnica de fluxo óptico. Esta técnica é definida como o deslocamento aparente dos padrões de brilho da imagem, obtendo-se para cada pixel da imagem um vector que representa a direcção e a intensidade do movimento (Horng and Schunck, 1981; Barron et al., 1992; Roberts et al., 2009). O fluxo óptico apresenta uma excelente segmentação das regiões em movimento, embora seja computacionalmente pesado, tornando-se assim inadequada para o sistema proposto. A figura 2.7 expõe dois exemplos de aplicação da técnica de fluxo óptico.

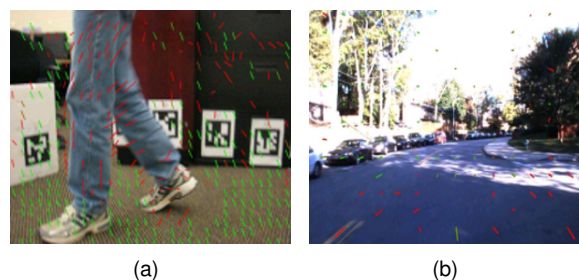


Figura 2.7: Exemplos de aplicação da técnica de fluxo óptico para ambientes: interiores (a) e exteriores (b). Os vectores de fluxo a vermelho indicam as regiões onde existe movimento na imagem actual (adaptado de Roberts et al. (2009)).

2.2.2 Seguimento de Objectos

O processo de seguimento de objectos pode ser dividido em duas aproximações: (1) a cada nova imagem, detectar todos os objectos presentes no ambiente e comparar com a lista de objectos da imagem anterior; (2) detectar e seguir objectos através da extracção de uma ou várias características que os descrevem unicamente, e recorrer a técnicas para executar o seguimento dos mesmos com base nas referidas características: algoritmo *mean shift* e *CAMShift*, filtro de *Kalman* e filtro de partículas.

Assim, ao utilizar a segunda aproximação para o seguimento de objectos, é necessário extrair pelo menos uma das características dos objectos:

- Características geométricas - posição e tamanho do objecto no plano da câmara, através da caixa delimitadora (*bounding box*) do objecto (Matsumura et al., 2002).
- Características de contorno - a forma do objecto é extraída através da segmentação das regiões de movimento (Hariatoglu et al., 1998; Haritaoglu et al., 2000).
- Características de cor e textura - recorrendo a modelos de aparência (Chien et al., 2006; Kang et al., 2003; Hariatoglu et al., 1998; Lu and Tan, 2001) o processo de seguimento de objectos torna-se substancialmente mais robusto e complexo computacionalmente.

Para o processo de seguimento é necessário adequar as características dos objectos ao ambiente, isto é, ter em consideração os parâmetros de aquisição da câmara (imagens a cor ou na escala de cinzentos, resolução da imagem e o campo visual da câmara), bem como o tipo de ambiente no qual o sistema de vídeo vigilância está instalado (ambientes interiores ou exteriores, dia ou noite).

A grande maioria dos trabalhos publicados têm por base a técnica de subtracção do plano de fundo para extrair regiões de movimento, e a partir destas executar o seguimento de objectos. Berclaz et al. (2006) executa o seguimento de objectos através do processamento das trajectórias individuais de cada objecto ao longo do tempo. Por outro lado, Hariatoglu et al. (1998) e Wren et al. (1997) utilizam uma análise da forma (cabeça, tronco, mãos e pés) dos objectos para o seguimento dos mesmos, criando modelos de aparência para distinguir objectos em situações de oclusão. A figura 2.8 mostra um exemplo do sistema proposto por (Hariatoglu et al., 1998), no qual são detectadas as regiões do corpo humano.

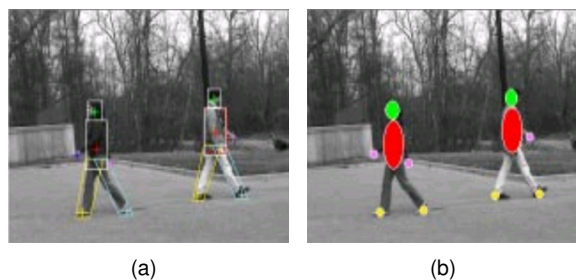


Figura 2.8: Exemplo do sistema proposto por Hariatoglu et al. (1998) em funcionamento: delimitação dos objectos em movimento (a) e ilustração das partes do corpo humano detectadas (b).

Haritaoglu et al. (1999) propôs um sistema de detecção de objectos através da extracção de regiões de movimento, executando posteriormente o seguimento dos objectos através do histograma de aparência. Simultaneamente, combina esta informação com a localização da cabeça das pessoas dentro de uma determinada região.

O método proposto por Zhao and Nevatia (2004) utiliza um modelo da forma da cabeça humana para segmentar regiões de movimento e seguir as pessoas. O algoritmo apresenta a particularidade da câmara poder estar instalada a poucos metros de altura e apontada para baixo. Kang et al. (2003) também aplica a detecção da cabeça após a detecção de movimento com base na técnica de subtracção do plano de fundo, seguindo a pessoa através da comparação da aparência da cabeça. Por outro lado, os métodos propostos por Wang et al. (2002) e Matsumura et al. (2002) aplicam a detecção de regiões de movimento para extrair regiões de pele na imagem actual, construindo um modelo da cor da pele para seguir objectos.

Os sistemas de vídeo vigilância instalados em ambientes exteriores estão sujeitos a grandes variações do ambiente. Neste sentido, as técnicas de segmentação de regiões de movimento poderão apresentar uma pobre segmentação dos objectos em movimento, fazendo com que os sistemas que dependem das características da cabeça e pele de um humano para executar o seguimento dos mesmos possam ser pouco robustos. A forma de solucionar este problema é recorrer a características que apresentem uma maior informação acerca dos objectos (cor e a textura), conjugadas com técnicas que adicionam ao sistema a capacidade de previsão. Existem duas abordagens principais para prever a posição de um determinado objecto na próxima imagem: (1) filtro de *Kalman* e (2) filtro de partículas.

O conceito base de um filtro de *kalman* (Thacker and Lacey, 2006; Kalman, 1960; Black et al., 2005) consiste na estimação recursiva de parâmetros através de um conjunto de premissas e na presença de ruído. Neste sentido, é possível maximizar a probabilidade à posteriori mantendo pouca informação transacta, actualizando iterativamente o modelo. Apesar de ser uma técnica robusta, o filtro de *kalman* é incapaz de lidar com distribuições não-gaussianas, tornando-se inadequado para sistemas de vídeo vigilância.

Para lidar com movimentos não lineares é necessário recorrer a outros filtros estocásticos mais avançados, tais como o filtro de partículas. Este tem a capacidade de estimar um conjunto de parâmetros (Arulampalam et al., 2002), permitindo assim seguir objectos previamente detectados. Este filtro consiste em utilizar um conjunto de amostras (denominadas partículas) ponderadas para aproximar uma determinada função densidade de probabilidade, estimando assim a posição do objecto a ser seguido. Uma descrição do filtro de partículas é apresentada na secção 3.3.3.1.

Para aplicações de vídeo vigilância automática, é recorrente a utilização do filtro de partículas para seguir objectos em movimento. O método proposto por Okuma et al. (2004) utiliza a combinação de um classificador de padrões humanos e filtros de partículas orientados aos modelos aprendidos para o seguimento de pessoas. Yang et al. (2005) propõe a combinação da informação de cor e silhueta do objecto para estimar a posição do mesmo através de filtros de partículas. Por outro lado, Nummiaro et al. (2003) utiliza a informação de cor e a técnica de filtro de partículas para seguir diferentes tipos de objectos.

2.2.2.1 Oclusões Entre Objectos

As oclusões entre objectos no mundo real são frequentes e estão associadas a dois tipos: (1) entre um objecto em movimento e objectos fixos no ambiente (e.g árvore e armário); (2) entre múltiplos objectos que se encontram a ser seguidos.

Uma das formas de lidar com oclusões entre objectos é utilizar a capacidade de previsão dos filtros de partículas com base na aparência dos objectos. Outra das formas é recorrer a múltiplas câmaras com campos de visão sobrepostos. Os métodos propostos por Khan et al. (2001) e Khan and Shah (2003), exploram a relação espacial conhecida entre os diversos campos de visão das câmaras para facilitar a correspondência entre os objectos. De forma semelhante, Kim and Davis (2006) e Khan and Shah (2006) executam uma associação de dados entre pontos de vista e exploram combinações e restrições de homografia, através da extracção dos eixos principais (verticais) de um ser humano e das características associadas com a silhueta, respectivamente.

Existem outras formas de lidar com oclusões. O método proposto por Lu and Tan (2001) antecipa a oclusão entre objectos em movimento, verificando a cada instante se a mínima caixa delimitadora ocupada pelos objectos poderá entrar em intersecção. Por outro lado, Koller et al. (1994) estima a posição de profundidade dos objectos em seguimento, classificando-os com base na sua posição vertical e combina essa informação com as posições esperadas, determinadas por um filtro de *Kalman*.

2.2.2.2 Múltiplas Câmaras

Um dos temas em ascensão na área de detecção e seguimento de objectos é a utilização de múltiplas câmaras com vistas disjuntas. O método proposto por Javed et al. (2003) estabelece uma correspondência entre objectos quando estes se movimentam entre câmaras nesta configuração, através de uma aprendizagem da topologia das câmaras e dos caminhos percorridos pelos objectos.

Os métodos propostos por Zhu et al. (2011) e Kettner and Zabih (1999) associam as trajectórias dos objectos em movimento sobre as múltiplas câmaras sem sobreposição do campo de visão para encontrar correspondência entre estes. Chen et al. (2008) propôs um método para aprendizagem de relações temporais, espaciais e de aparência entre câmaras para monitorização contínua de longa duração. Através de uma observação estatística, o método proposto por Makris et al. (2004) determina automaticamente a topologia das câmaras com vistas disjuntas. Para além de calcular automaticamente a topologia das câmaras, Black et al. (2005) estima o tempo de transição entre os diversos pontos disjuntos.

Este tipo de sistemas estão sujeitos a demasiadas falhas, visto que um determinado objecto pode mudar de aparência na transição entre as múltiplas câmaras e mudar o seu padrão de comportamento, tornando o reconhecimento de padrões altamente falível.

Capítulo 3

Metodologia Proposta

Este capítulo descreve a metodologia proposta para a detecção de múltiplos focos de incêndio e, simultaneamente, detecção e seguimento de múltiplos objectos em tempo real, em ambientes exteriores e com alterações dinâmicas das condições de luminosidade inerentes ao local onde o sistema de vídeo vigilância está instalado.

O sistema proposto é composto pelos algoritmos de detecção de fogo e detecção e seguimento de objectos, descritos nas secções 3.2 e 3.3, respectivamente. Para a detecção de focos de incêndio são analisados a informação da cor das regiões de píxeis em movimento (descrito na secção 3.1) e as características dinâmicas conhecidas do fogo no domínio da frequência. Os métodos implementados para a detecção e seguimento de objectos incluem técnicas de segmentação de regiões de movimento, filtros estocásticos para o seguimento de objectos e classificadores treinados para a discriminação entre um objecto genérico e um humano.

Com base no algoritmo de detecção de fogo é possível diferenciar uma região de fogo de um objecto em movimento com cores semelhantes às das chamas. No entanto, objectos com movimentos repetitivos podem afectar a robustez dessa diferenciação. Neste sentido, o sistema proposto nesta dissertação integra o resultado dos dois algoritmos (ver Secção 3.5) de modo a reduzir a taxa de falsos alarmes na detecção de fogo, eliminando os alarmes provenientes dos referidos objectos.

Consoante o tipo de evento que ocorre no campo de visão da câmara, o sistema reporta alarmes com as coordenadas do evento no sistema de posicionamento global (*GPS*). Esta capacidade deve-se à calibração do sistema, detalhada na secção 3.4.

A figura 3.1 mostra o processamento do sistema proposto para cada imagem. Em primeiro lugar, as regiões de píxeis em movimento são extraídas a cada nova imagem. Seguidamente, o algoritmo de detecção de fogo é processado, onde os píxeis em movimento são classificados a partir de um modelo baseado na cor do fogo, e o conjunto de píxeis classificados como fogo são utilizados para a detecção de texturas dinâmicas, através de um filtro temporal recursivo proposto nesta dissertação. As regiões detectadas, isto é, com assinatura temporal semelhante às das chamas, são analisadas no domínio da frequência, avaliando a oscilação dos píxeis e as texturas dinâmicas com base na transformada *wavelet* discreta no espaço de frequência temporal e espacial, respectivamente.

De seguida, o algoritmo de detecção e seguimento de pessoas é executado, onde o plano de fundo é removido para extrair as regiões de movimento com base na história dos píxeis. Caso uma região de movimento seja detectada (objecto), é aplicado o método de remoção de sombras para extrair os píxeis considerados sombra. Quando um determinado objecto for consecutivamente detectado e apresentar movimento, este é seguido através de um filtro de partículas, que estima a sua posição através do modelo de cor que o define. Por fim, o objecto é definido como humano através de um classificador baseado em histogramas de gradientes orientados (*HOG*).

O último passo do sistema é a integração dos resultados dos referidos algoritmos. Esta integração é essencial para verificar se um resultado positivo da análise *wavelet* do algoritmo de detecção de fogo é causado por um objecto em movimento. Por fim, é gerado um conjunto de alarmes (fogo, movimento e seguimento) consoante o tipo de evento a ocorrer no ambiente.

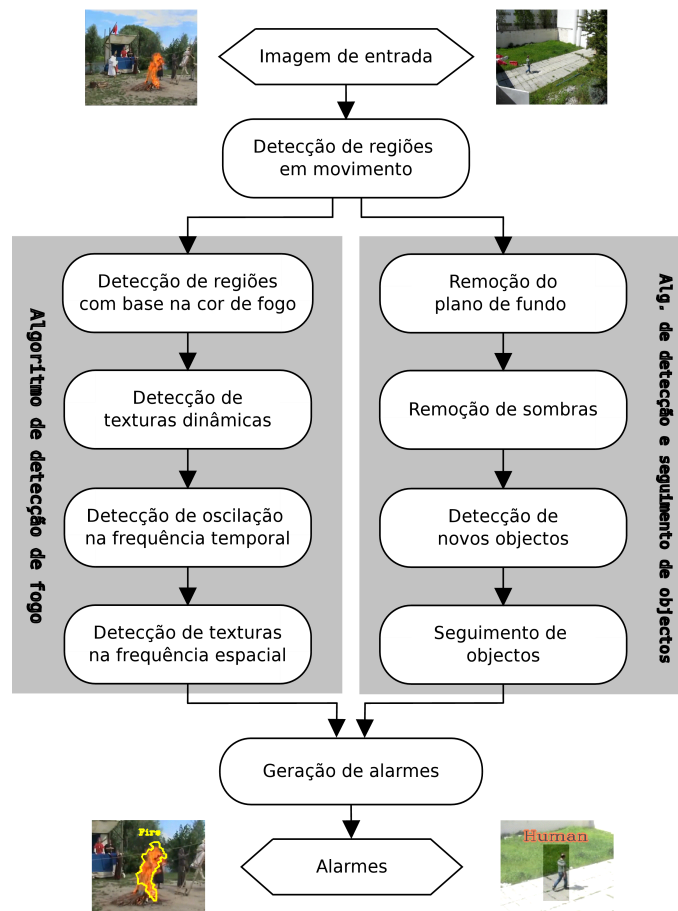


Figura 3.1: Diagrama de processamento do sistema proposto para a detecção de fogo e detecção e seguimento de objectos para cada nova imagem.

3.1 Detecção de Regiões de Movimento

O primeiro passo de uma aplicação de vídeo vigilância é a extracção de regiões de movimento. Com base no método para a estimação do plano de fundo proposto por Collins et al. (2000), *HBE*, é possível segmentar regiões de píxeis em movimento através de sequências de imagens. Embora este método seja baseado em duas operações, com vista a diminuir o custo computacional através

da substituição da subtracção adaptativa do plano de fundo pelo algoritmo de componentes ligados, nesta dissertação apenas é considerada a diferenciação de três imagens consecutivas. Assim, a informação de um pixel é considerada em movimento se o valor da sua intensidade variar significativamente entre a imagem actual e as duas imagens anteriores, ou seja,

$$M = \{(x, y) : (|\mathbf{I}_s^n(x, y) - \mathbf{I}_s^{n-1}(x, y)| > \mathbf{T}_n(x, y)) \wedge (|\mathbf{I}_s^n(x, y) - \mathbf{I}_s^{n-2}(x, y)| > \mathbf{T}_n(x, y))\}, \quad (3.1)$$

onde \mathbf{I}_s^n representa a intensidade do pixel na posição (x, y) na imagem n . O parâmetro \mathbf{T}_n representa o limiar de diferenciação que descreve uma mudança significativa na intensidade do pixel na posição (x, y) em n .

O modelo de plano de fundo \mathbf{B}_n e o limiar de diferenciação \mathbf{T}_n são actualizados ao longo do tempo através das expressões:

$$\mathbf{B}_n(x, y) = \begin{cases} \alpha \mathbf{B}_{n-1}(x, y) + (1 - \alpha) \mathbf{I}_s^{n-1}(x, y), & \text{se } (x, y) \notin M, \\ \mathbf{B}_{n-1}(x, y), & \text{se } (x, y) \in M. \end{cases} \quad (3.2)$$

$$\mathbf{T}_n(x, y) = \begin{cases} \alpha \mathbf{T}_{n-1}(x, y) + (1 - \alpha)(5|\mathbf{I}_s^{n-1}(x, y) - \mathbf{B}_n(x, y)|), & \text{se } (x, y) \notin M, \\ \mathbf{T}_n(x, y), & \text{se } (x, y) \in M, \end{cases} \quad (3.3)$$

onde α representa uma constante que especifica a velocidade pela qual as novas observações se sobrepõem às observações anteriores, definida no intervalo $[0, 1]$. Quando se pretende extrair apenas as regiões de movimento da imagem actual, sem ter em consideração o movimento das mesmas em instantes anteriores, o parâmetro α é definido como 1. No instante inicial, \mathbf{B}_0 é actualizado com a primeira imagem do sistema e \mathbf{T}_0 é definido como um valor superior a zero (ver em baixo).

Na figura 3.2 estão expostos resultados do método quando aplicado em ambientes exteriores, contendo chamas e pessoas em movimento.

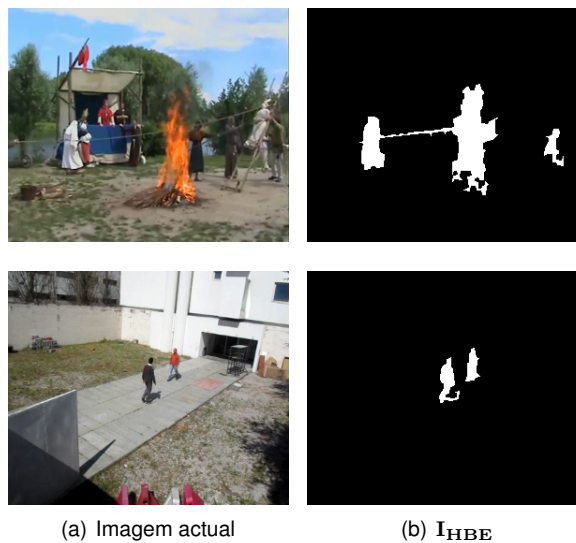


Figura 3.2: Detecção de regiões de movimento através do método para estimação do plano de fundo. (a) Imagem actual; (b) máscara resultante do método HBE , \mathbf{I}_{HBE} .

É importante notar que o valor inicial do limiar de diferenciação T_0 necessita de reflexão, pois quanto mais elevado for o seu valor, menos sensível se torna o método *HBE* na detecção de regiões de movimento. A figura 3.3 ilustra três exemplos do resultado do método para a estimação do plano de fundo para diferentes valores do limiar de diferenciação T_0 .

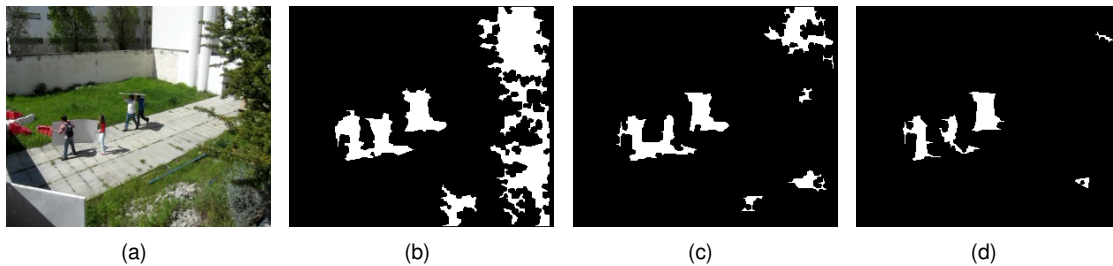


Figura 3.3: (a) Imagem actual; máscara resultante do método *HBE* para diferentes valores do limiar de diferenciação T_0 no ambiente da figura (a), definido como 5 (b), 10 (c) e 15 (d).

Na figura 3.3(b) o limiar T_0 foi definido como 5, apresentando assim uma maior segmentação das regiões de movimento. No entanto este limiar é demasiado baixo, visto que a mínima alteração de movimento na imagem actual é detectada, acrescentando demasiado ruído. No caso da figura 3.3(d), o limiar de diferenciação foi definido como 15 e assim não são apresentadas as regiões de movimento provenientes das folhas das árvores (lado direito da imagem). No entanto, este limiar é demasiado elevado para se segmentar o movimento de objectos de pequena dimensão. Neste sentido, após uma parametrização do limiar de diferenciação, este foi definido como 10. A figura 3.3(c) apresenta o resultado do método *HBE* para esta parametrização, onde o movimento proveniente das folhas das árvores não é significativo, ao mesmo tempo que segmenta o movimento das pessoas (incluindo as suas sombras).

3.2 Algoritmo de Detecção de Fogo

Esta secção descreve os métodos implementados para a detecção de fogo e o trabalho realizado no âmbito da dissertação para aumentar a robustez e eficiência da referida detecção.

3.2.1 Detecção de Regiões de Píxeis Com Base na Cor do Fogo

As cores das chamas são normalmente avermelhadas, tornando-se azuis à medida que a temperatura do fogo aumenta ou amarelas quando a temperatura do fogo diminui (Drysdale, 2011). Cada imagem digital a cores é composta por três planos de cores: vermelho (R), verde (G) e azul (B). A combinação de planos de cores *RGB* permite representar uma determinada cor em componentes digitais, sendo este espaço de cores um dos mais usados para processamento digital de imagens.

A detecção de regiões com base na cor do fogo é executada através da técnica de classificação dos píxeis com base na cor, que se baseia num conjunto pré-definido de regras para classificar um pixel como fogo. Visto que geralmente o fogo apresenta-se como um objecto estático com forma e textura dinâmicas, esta técnica só é aplicada apenas às regiões de píxeis em movimento, provenientes do método *HBE*, de modo a tornar o sistema mais robusto e eficiente.

Uma das contribuições desta dissertação é estudar e analisar a eficiência de alguns dos métodos de classificação dos píxeis com base na cor presentes na literatura. Neste sentido, seguidamente são descritos os métodos propostos por Celik and Demirel (2009), Horng et al. (2005) e Chen et al. (2010), que recorrem a diferentes métricas e espaços de cores para segmentar regiões com chamas. De referir que o objectivo de um espaço de cores é facilitar a representação das cores de uma imagem através de um padrão, isto é, um modelo matemático para formalizar a descrição de cores.

3.2.1.1 Métodos de Classificação dos Píxeis Com Base na Cor

O primeiro método de classificação dos píxeis com base na cor descrito é proposto por Chen et al. (2010), que se baseia na filtragem dos píxeis através de um conjunto de regras que descrevem as características da cor das chamas no espaço de cores *RGB*, ou seja,

$$\text{Regra 1: } R \geq R_T \quad \wedge \quad R \geq G \geq B \quad (3.4)$$

$$\text{Regra 2: } Y_2 \leq S \leq Y_1, \quad (3.5)$$

onde R , G e B representam os componentes vermelho, verde e azul no espaço de cores *RGB* de um determinado pixel da imagem. O parâmetro R_T (definido como 120) é o valor global da componente vermelha. A segunda regra é utilizada para restringir a saturação da cor, eliminando assim as regiões com elevada iluminação. As curvas Y_1 e Y_2 representam os limites superior e inferior da correlação entre a componente azul e a saturação, dado pela expressão:

$$Y_1 = \begin{cases} 100 - 0.48B, & \text{se } B \geq 117, \\ 79.27 - 0.311B, & \text{se } B < 117. \end{cases} \quad (3.6)$$

$$Y_2 = -2.0147 + 90.59435e^{-B/77.6027}, \quad (3.7)$$

onde o valor de saturação (S) de um determinado pixel é obtido através da conversão entre o espaço de cores *RGB* e *HSV*, dado pela expressão:

$$H = \arccos \frac{\frac{1}{2}((R - G) + (R - B))}{\sqrt{((R - G)^2 + (R - B)(G - B))}} \quad (3.8)$$

$$S = 1 - 3 \frac{\min(R, G, B)}{R + G + B} \quad (3.9)$$

$$V = \frac{1}{3}(R + G + B), \quad (3.10)$$

onde H representa a componente matiz, definida como a cor dominante. A componente saturação mede a proporção de cor e a componente valor (V) está relacionado com a luminosidade da cor.

Através da discriminação de cores, o método proposto por Chen et al. (2010) cria uma máscara

com as regiões de píxeis em movimento (pertencentes ao conjunto M) que validam as duas regras previamente descritas (ver figura 3.4).

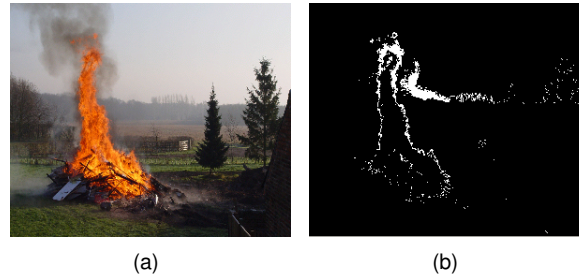


Figura 3.4: (b) Resultado do método de classificação dos píxeis com base na cor proposto por Chen et al. (2010) aplicado à figura (a).

Outro dos métodos analisados nesta dissertação foi proposto por Celik and Demirel (2009). Este utiliza o espaço de cores $YCbCr$ para eliminar o efeito da luminosidade, construindo um modelo genérico de classificação dos píxeis com base na cor através de um conjunto pré-definido de regras. O espaço de cores $YCbCr$ é composto por três componentes: Y representa a luminosidade; Cb e Cr são dois valores de cor formados através da subtração da componente luminosidade com as componentes azul e vermelho do espaço de cores RGB , denominados cromaticidade azul e cromaticidade vermelho, respectivamente. A conversão entre o espaço de cores RGB e $YCbCr$ é dada por:

$$\begin{bmatrix} Y \\ Cb \\ Cr \end{bmatrix} = \begin{bmatrix} 0.2568 & 0.5041 & 0.0979 \\ -0.1482 & -0.2910 & 0.4392 \\ 0.4392 & -0.3678 & -0.0714 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} + \begin{bmatrix} 16 \\ 128 \\ 128 \end{bmatrix}, \quad (3.11)$$

onde Y representa a componente luminosidade contida no intervalo $[16, 235]$. Os parâmetros Cb e Cr representam as componentes cromaticidade azul e cromaticidade vermelho, respectivamente, contidas no intervalo $[16, 240]$.

O conjunto de regras definidas por Celik and Demirel (2009) para segmentar regiões de chamas com base na informação da cor dos píxeis são dadas por:

$$\text{Regra 1: } Y > Cb \quad \wedge \quad Cr > Cb \quad (3.12)$$

$$\text{Regra 2: } Y > Y_{mean} \quad \wedge \quad Cb > Cb_{mean} \quad \wedge \quad Cr > Cr_{mean}, \quad (3.13)$$

onde Y , Cb e Cr são as componentes luminosidade, cromaticidade azul e cromaticidade vermelho de um determinado pixel da imagem, respectivamente. Os parâmetros Y_{mean} , Cb_{mean} e Cr_{mean} representam o valor médio das três componentes do espaço de cores $YCbCr$.

De forma semelhante ao método descrito anteriormente, o método proposto por Celik and Demirel (2009) cria uma máscara com as regiões de píxeis em movimento que validam as duas regras previamente descritas. Na figura 3.5 estão retratados os componentes Y , Cb , Cr , o resultado da conversão para o espaço de cores $YCbCr$ e o resultado do método de classificação dos píxeis com base na cor recorrendo ao referido espaço de cores (Celik and Demirel, 2009).

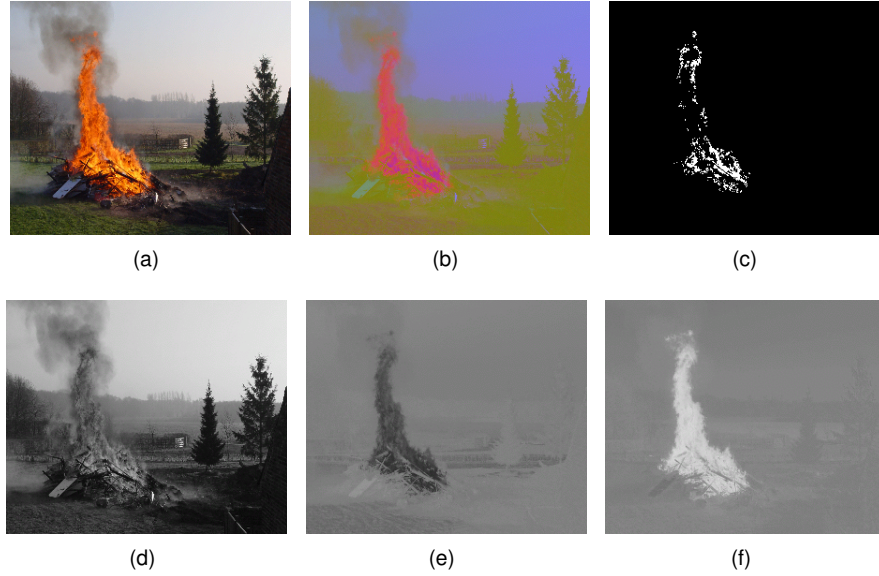


Figura 3.5: Classificação dos píxeis com base na cor através do espaço de cores $YCbCr$. (a) Imagem original; (b) imagem convertida para o espaço de cores $YCbCr$; (c) resultado do método de classificação dos píxeis com base na cor proposto por Celik and Demirel (2009); componentes: luminosidade (d), crominância vermelho (e) e crominância azul (f) do espaço de cores $YCbCr$ da figura (a).

O último método em análise foi proposto por Horng et al. (2005), que utiliza o espaço de cores HSI e aplica o método de separação de cores para segmentar regiões com cores semelhantes às das chamas. O espaço de cores HSI é formado por três componentes: (1) o matiz que está relacionado com o comprimento espectral da onda que produz a percepção da cor; (2) a componente saturação que está relacionada com a pureza da cor, isto é, a quantidade de cor em relação à cor cinza média; (3) a componente intensidade (I) que está associada ao brilho da cor (Vassili et al., 2003). A conversão entre o espaços de cores RGB e HSI é dada pelas expressões:

$$I = \frac{1}{3}(r + g + b) \quad (3.14)$$

$$S = 1 - \frac{3}{(r + g + b)}[\min(r + g + b)] \quad (3.15)$$

$$H = \begin{cases} \theta, & \text{se } b \leq g, \\ 360 - \theta, & \text{se } b > g. \end{cases} \quad (3.16)$$

$$\theta = \cos^{-1} \left\{ \frac{\frac{1}{2}[(r - g) + (r - b)]}{[(r - g)^2 + (r - b)(g - b)]^{\frac{1}{2}}} \right\}, \quad (3.17)$$

onde os valores das componentes r , g e b do espaço de cores RGB estão normalizados no intervalo $[0, 1]$, através das expressões:

$$r = \frac{R}{R + G + B}, \quad g = \frac{G}{R + G + B}, \quad b = \frac{B}{R + G + B}. \quad (3.18)$$

Através de uma análise empírica, Horng et al. (2005) efectuou uma análise às características de cor das chamas de acordo com o espaço de cores *HSI*, aplicando posteriormente o método de separação de cores para segmentar regiões com aparência semelhante ao fogo, construindo assim um conjunto de píxeis considerados chamas,

$$C_f([h_1, h_2], [s_1, s_2], [i_1, i_2]) = \{(H, S, I) | h_1 \leq H \leq h_2, s_1 \leq S \leq s_2 \wedge i_1 \leq I \leq i_2\}, \quad (3.19)$$

onde h_1, h_2, s_1, s_2, i_1 e i_2 são os valores máximos e mínimos das três componentes do espaço de cores *HSI*, expressos através de um estudo detalhado das características das chamas no referido espaço de cores (Horng et al., 2005). Os parâmetros H, S e I são os componentes matiz, saturação e intensidade do espaço de cores *HSI*, respectivamente.

Com base na informação da cor dos píxeis da imagem actual, o método proposto por Horng et al. (2005) cria uma máscara com as regiões de píxeis em movimento que estejam contidos no conjunto C_f . As componentes matiz, saturação e intensidade, a conversão para o espaço de cores *HSI* e o resultado do método de classificação dos píxeis com base na cor proposto por Horng et al. (2005) estão expostos na figura 3.6.

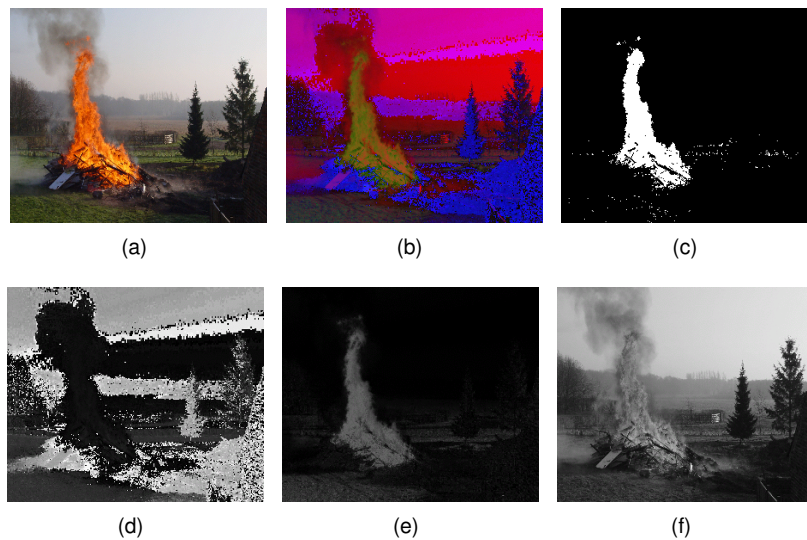


Figura 3.6: (a) Imagem original; (b) imagem convertida para o espaço de cores *HSI*; (c) resultado do método de classificação dos píxeis com base na cor proposto por Horng et al. (2005); componentes: matiz (d), saturação (e) e intensidade (f) do espaço de cores *HSI* da figura (a).

3.2.1.2 Análise do Desempenho dos Métodos de Classificação dos Píxeis com Base na Cor

Como referido anteriormente, a capacidade de extrair regiões de píxeis em movimento com cores semelhantes às das chamas é essencial para a detecção de fogo. Neste sentido, é necessário analisar e fundamentar qual dos métodos expostos (Horng et al., 2005; Celik and Demirel, 2009; Chen et al., 2010) se deve incorporar no algoritmo de detecção de fogo proposto nesta dissertação. De modo a simplificar a análise do desempenho dos métodos descritos anteriormente, estes serão referidos através da nomenclatura do espaço de cores (*HSI, YCbCr* e *RGB*) associados aos métodos expostos, respectivamente.

Considerando que os sistemas de vídeo vigilância estão presentes na grande maioria dos ambientes urbanos, é também verificado o comportamento dos referidos classificadores consoante o tipo de ambiente no qual o sistema de vídeo vigilância está instalado. Para tal, foram utilizadas 217 imagens retiradas de sítios na Internet, contendo chamas decorrentes de situações quotidianas, divididas em 4 categorias: interior, noite, rural e urbano. Com base nas imagens, foi executada uma segmentação manual para criar máscaras que contêm apenas as regiões com chamas. A figura 3.7 retrata exemplos das referidas categorias e a máscara utilizada para a análise do desempenho dos métodos de classificação dos píxeis com base na cor.

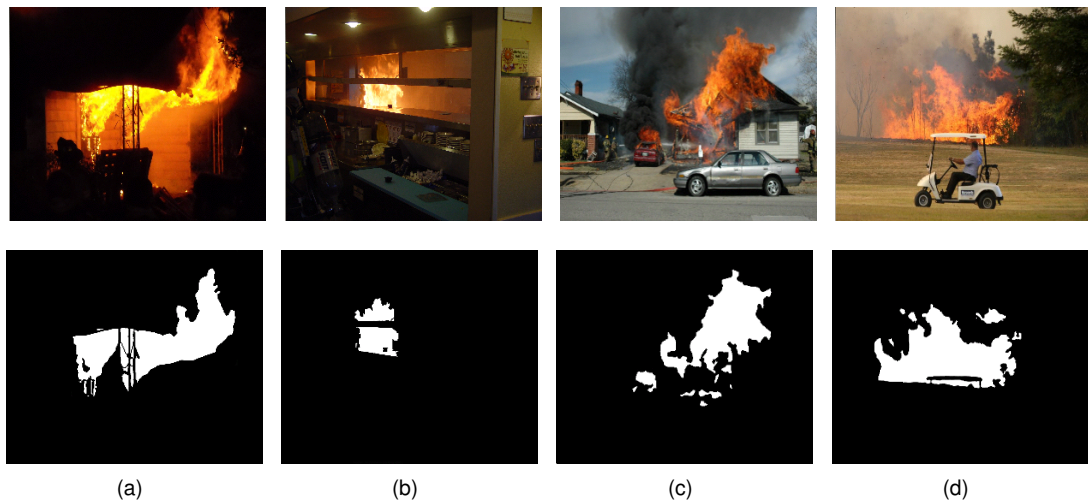


Figura 3.7: Segmentação manual das regiões de chamas. A primeira linha mostra as imagens originais e a segunda as máscaras resultantes da segmentação manual das chamas, relativas às categorias noite (a), interior (b), urbano (c) e rural (d).

Para além da análise da eficiência dos métodos de classificação dos píxeis com base na cor baseados nos espaços de cores *HSI*, *YCbCr* e *RGB*, são também analisadas duas das suas possíveis combinações, *HY* e *HYR*. Os píxeis destas combinações são classificados como fogo através da acumulação do resultado dos métodos que se baseiam nos espaços de cores *HSI* e *YCbCr* para a primeira combinação e nos espaços de cores *HSI*, *YCbCr* e *RGB* para a segunda combinação.

A tabela 3.1 apresenta os resultados obtidos com os três métodos de classificação dos píxeis previamente descritos e com os modelos de cor propostos nesta dissertação, *HYR* e *HY*. A métrica utilizada foi a *Matthews Correlation Coefficients (MCC)*, que permite a classificação binária de imagens, tendo em consideração verdadeiros e falsos positivos, e os verdadeiros e falsos negativos. De referir que um *MCC* de 1 representa uma classificação perfeita.

Tabela 3.1: Resultados experimentais da detecção de fogo através do método de classificação dos píxeis com base na cor (*MCC*)

Categorias	# Imagens	<i>HSI</i>	<i>YCbCr</i>	<i>RGB</i>	<i>HYR</i>	<i>HY</i>
Urbano	124	0.78	0.51	0.19	0.81	0.83
Rural	73	0.73	0.57	0.17	0.77	0.79
Interior	20	0.36	0.47	0.12	0.43	0.46
Noite	21	0.60	0.61	0.01	0.64	0.64
Todas	217	0.74	0.54	0.15	0.78	0.79

Os sombreados presentes na tabela 3.1 evidenciam qual o melhor método de classificação dos píxeis em função do tipo de imagem. Os resultados obtidos demonstram que o modelo de cor proposto nesta dissertação, *HY*, é o que melhor segmenta as regiões de chamas para o banco de imagens em estudo. Este modelo apresenta o melhor valor de *MCC* para todas as imagens e para praticamente todas as categorias, com excepção das imagens de interior, embora a diferença para o espaço de cores *YCbCr* seja pouco significativa. Por seu lado, o espaço de cores *RGB* é o que pior segmenta as regiões com chamas, apresentando em alguns casos uma segmentação nula.

Globalmente, os resultados experimentais obtidos através da classificação dos píxeis com base na cor realçam a incapacidade do espaço de cores *RGB* em detectar regiões de fogo e a complementaridade na detecção de fogo dos espaços de cores *HSI* e *YCbCr*. Na figura 3.8 encontram-se alguns exemplos dos resultados dos métodos implementados para a detecção de regiões de píxeis com base na cor do fogo.

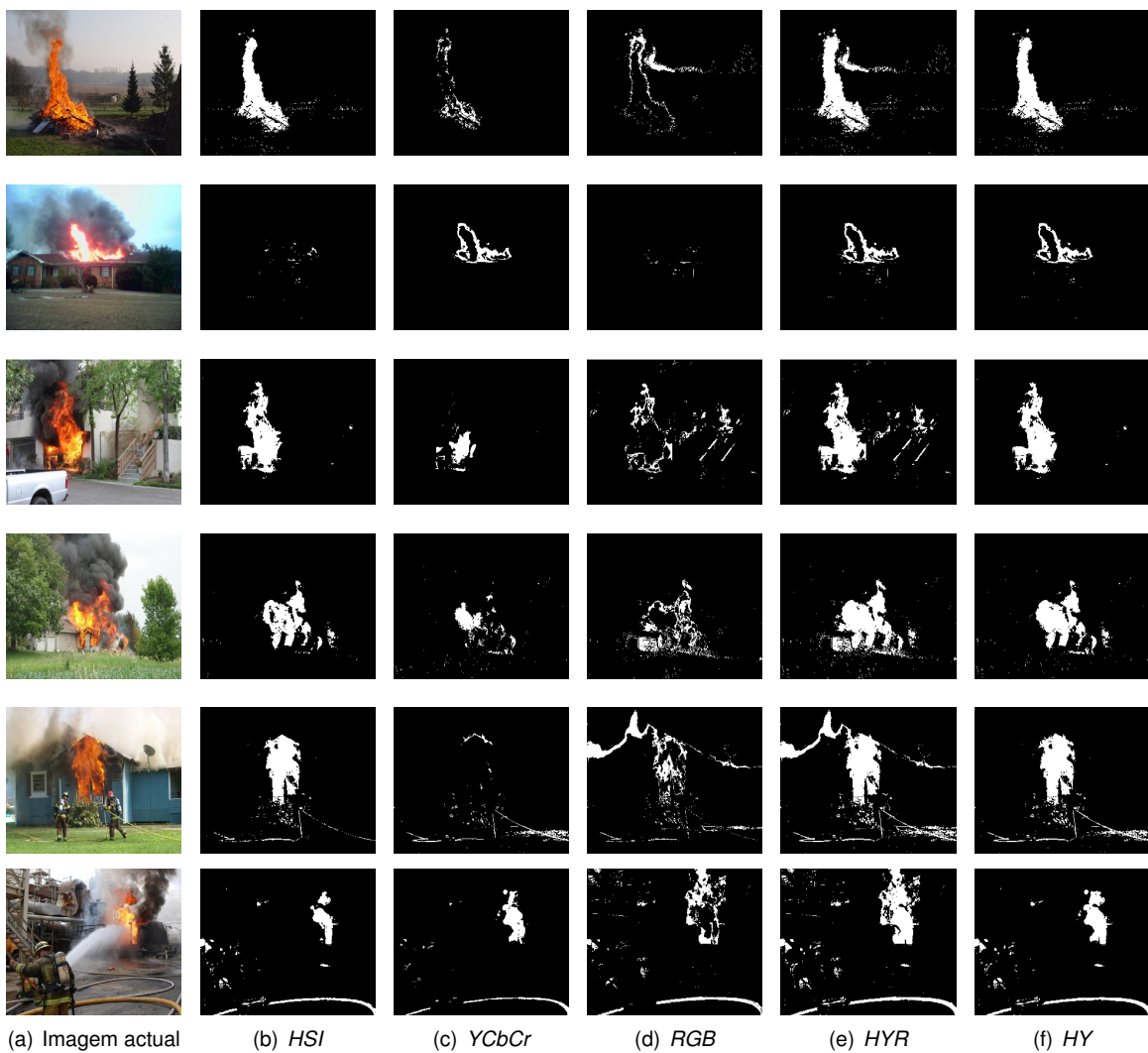


Figura 3.8: Resultados obtidos para a detecção de fogo em 6 imagens (uma por linha) através da informação da cor dos píxeis com base em 5 modelos de cor. Os píxeis a branco (b)-(f) correspondem aos píxeis classificados como fogo. Nos 6 casos expostos, a configuração *HY* é a que apresenta os melhores resultados, pois explora a complementaridade dos métodos baseados nos espaços de cores *HSI* e *YCbCr*.

Como referido anteriormente, as imagens provenientes de sistemas de vídeo vigilância instalados em ambientes exteriores contêm normalmente um conjunto de distractivos indesejáveis para a

detecção de fogo, os quais apresentam uma aparência semelhante ao fogo. A figura 3.9 apresenta exemplos destes elementos e as máscaras resultantes da classificação dos píxeis com base na cor.

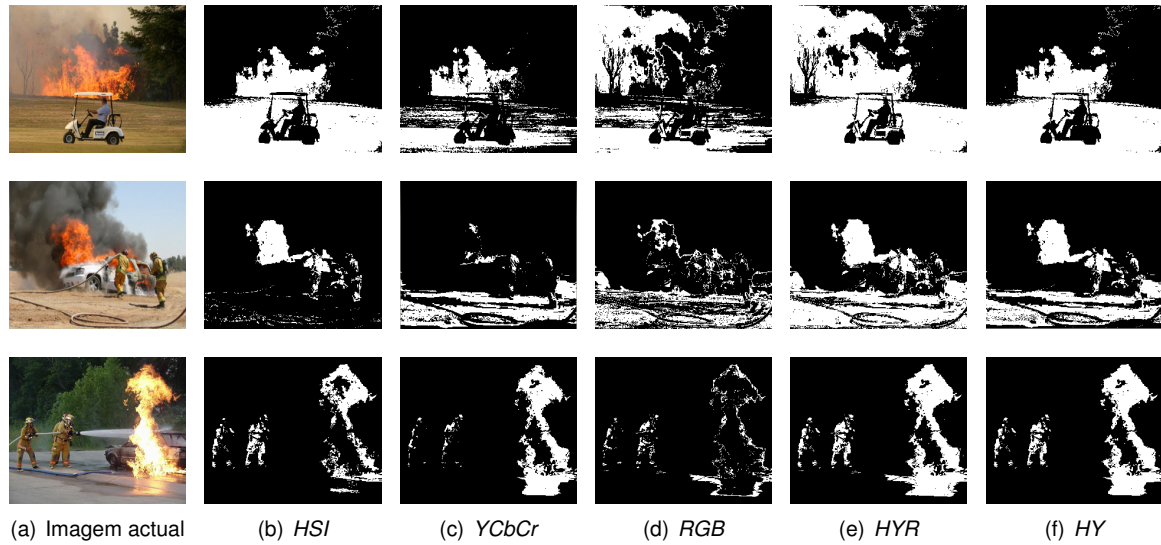


Figura 3.9: Resultados obtidos para a detecção de fogo em três imagens (uma por linha) através da informação da cor dos píxeis com base em 5 modelos de cor com elementos distractivos fixos no ambiente. As chamas e os elementos distractivos são classificados como fogo (píxeis brancos nas imagens (b)-(f)), embora o algoritmo só classifique píxeis que apresentem movimento. Na segunda linha, os bombeiros são igualmente classificados como fogo, o que pode resultar num falso alarme na detecção de fogo.

Independentemente do modelo de cor utilizado para classificar os píxeis com base na cor, verifica-se que praticamente todos os elementos distractivos presentes nas imagens são classificados como fogo. No entanto, como apenas as regiões em movimento (ver Secção 3.1) são classificadas através da informação da cor, são filtrados e eliminados os elementos distractivos estáticos presentes na imagem. Apesar da filtragem, a figura 3.9 mostra que os objectos em movimento com cores semelhantes às das chamas são segmentados como fogo, podendo assim gerar um falso alarme na sua detecção. Para resolver este problema, as regiões de movimento classificadas como fogo são filtradas para a detecção de texturas dinâmicas e, posteriormente, são analisadas no domínio da frequência.

3.2.2 Detecção de Texturas Dinâmicas

A análise no domínio da frequência (ver Secção 3.2.3) é pesada computacionalmente. Para mitigar este problema, esta secção apresenta o método de detecção de texturas dinâmicas, proposto nesta dissertação. Este representa a história dos píxeis segmentados como fogo pelo método de classificação dos píxeis com base na cor ao longo do tempo (ver figura 3.10).

O método tem como objectivo detectar a presença de texturas dinâmicas através de um filtro temporal recursivo. Este é alimentado através da imagem binária resultante da detecção de regiões com cor de fogo através do modelo de cor *HY* proposto nesta dissertação, I_{PC} , criada através do processo descrito na secção 3.2.1:

$$I_{TD}(n) \leftarrow I_{TD}(n-1) + \lambda_1 \cdot I_{PC}(n) - \lambda_2 \cdot (1 - I_{PC}(n)), \quad (3.20)$$

onde n é o índice da imagem, λ_1 e λ_2 são constantes definidas empiricamente. O resultado do processo de filtragem, I_{TD} , é truncado entre $[0, 255]$ para evitar a saturação da imagem e $I_{TD}(0)$ é definida como zero. Com a inclusão destas constantes, garante-se que I_{TD} cresce ao ritmo λ_1 quando existem regiões de píxeis em movimento e decresce ao ritmo λ_2 caso contrário. Deste modo, com $\lambda_1 = 3$ e $\lambda_2 = 1$, o filtro é sintonizado para a frequência típica das texturas dinâmicas das chamas.

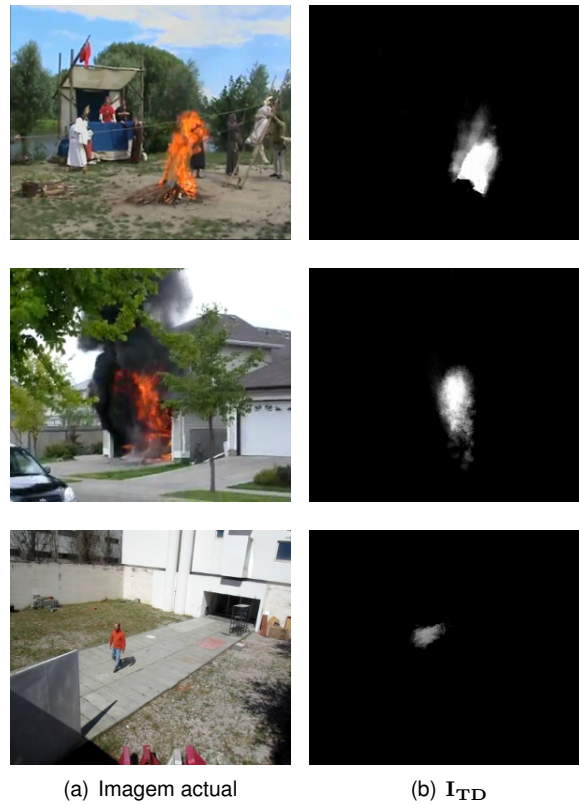


Figura 3.10: Detecção de texturas dinâmicas. a) Imagem actual; (b) máscara resultante do método de detecção de texturas dinâmicas, I_{TD} .

A figura 3.10 mostra que um objecto em movimento com uma aparência semelhante ao fogo pode apresentar uma textura dinâmica, gerando assim um falso alarme na detecção de fogo. Assim, caso o método reporte a presença de uma textura numa determinada região, isto é, quando a intensidade média dos píxeis nessa região no resultado do processo de filtragem for superior ao limiar κ (parametrizado para 12), esta é analisada no domínio da frequência (ver Secção 3.2.3).

3.2.2.1 Segmentação das Regiões de Fogo

O algoritmo de detecção de fogo proposto nesta dissertação tem por objectivo detectar e seguir múltiplos focos de incêndio. Para tal, é necessário agrupar regiões de píxeis considerados fogo, eliminar os contornos das mesmas e executar a correspondência entre estas regiões ao longo do tempo.

Um ser humano em movimento pode apresentar um regime de oscilação dos píxeis (em volta do tronco e membros) semelhante ao das chamas. Após a detecção de texturas dinâmicas, os píxeis

circundantes aos objectos detectados, os quais poderão apresentar uma elevada taxa de oscilação, são removidos. O objectivo deste processo é diminuir a taxa de falsos alarmes na detecção de fogo dos métodos de análise no domínio da frequência.

O resultado da remoção de contornos, I_{RC} , é dado pela máscara de classificação dos píxeis com base na cor, na qual os contornos (sequência de píxeis ligados entre si) detectados através do algoritmo de componentes ligados são desenhados a preto, eliminando assim os píxeis que apresentam uma maior oscilação. A figura 3.11 apresenta 2 exemplos de detecção e remoção dos contornos detectados.

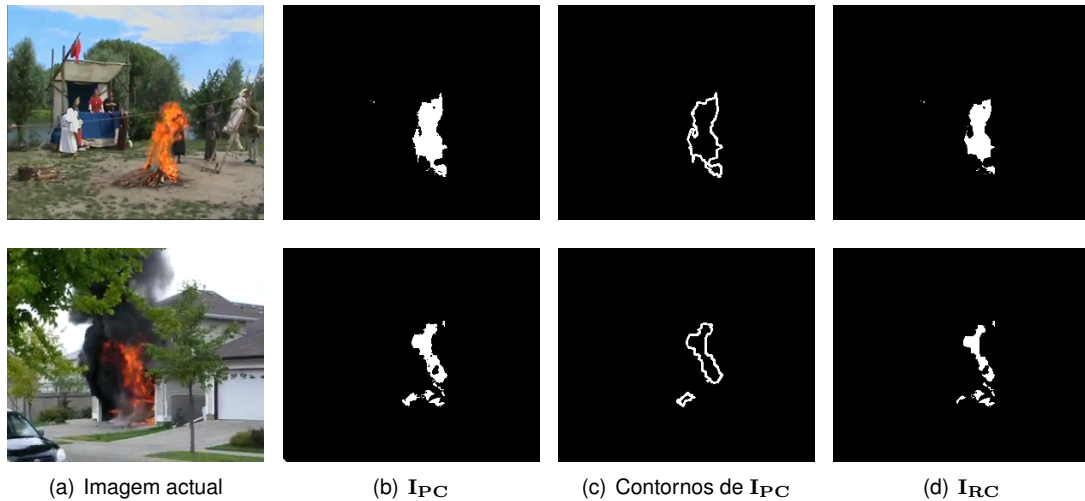


Figura 3.11: Processo de remoção de contornos. (b) máscara resultante do método de classificação dos píxeis com base na cor; (c) contornos das chamas extraídos a partir do algoritmo de componentes ligados com base na figura (b); (d) máscara resultante do processo de remoção dos contornos, onde os píxeis que apresentam uma elevada oscilação foram removidos.

A segmentação das regiões de chamas através do método de classificação dos píxeis com base na cor não é uniforme ao longo do tempo, isto é, a cada iteração do algoritmo de detecção de fogo uma determinada região de chamas poderá apresentar uma forma diferente e variar a sua localização. Assim, para cada região de píxeis detectada é extraída a sua caixa delimitadora e é verificado se esta se encontra em intersecção com o conjunto de regiões previamente detectadas (ver figura 3.12). Caso as caixas delimitadoras de duas regiões se intersectem, estas são fundidas e é efectuada uma correspondência entre os seus parâmetros de identificação. Caso não seja detectada nenhuma intersecção, uma nova região é criada.

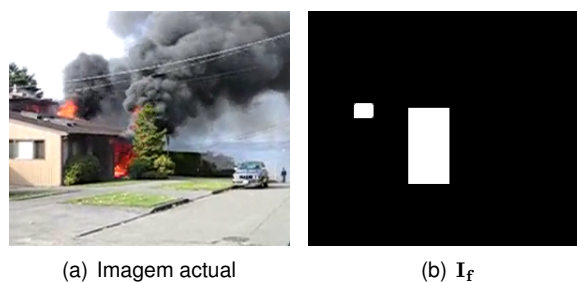


Figura 3.12: Localização das regiões de fogo. (a) Imagem original; (b) máscara com as caixas delimitadoras das regiões de fogo detectadas, I_f , marcadas a branco se foram classificadas como fogo ou em tons de cinza, caso os métodos de análise no domínio da frequência estejam a ser aplicados.

É importante mencionar que se a correspondência entre regiões de píxeis não fosse executada, a cada iteração do algoritmo de detecção de fogo seriam criados novas regiões e assim aplicadas várias análises no domínio da frequência sobre o mesmo objecto. Desta forma, reduz-se significativamente o custo computacional. As regiões de fogo detectados são identificadas através dos seguintes parâmetros:

- Identificador da região através de um valor numérico único.
- Posição actual da caixa delimitadora da região, definida pela sua posição na imagem, a sua altura e largura (x_i, y_i, w_i, h_i) .
- A análise no domínio da frequência de cada região é executada sobre a localização inicial da sua caixa delimitadora. Assim, é necessário guardar a primeira caixa delimitadora.
- Marcas com informação se os métodos de análise no domínio da frequência estão a ser aplicados e o resultado dos mesmos.
- A cada nova detecção é registado o número da imagem actual. Este parâmetro é necessário para: (1) eliminar uma região do sistema caso não seja detectado nas últimas 50 imagens; (2) verificar o número de imagens recolhidas para executar a análise das características dinâmicas conhecidas do fogo no domínio da frequência.

As caixas delimitadoras de duas regiões detectadas, i e j , encontram-se em intersecção se:

$$(|x_i - x_j| < \lambda(w_i + w_j)) \quad \wedge \quad (|y_i - y_j| < \lambda(h_i + h_j)) \quad (3.21)$$

onde os parâmetros x , y , w e h representam a posição na imagem, a largura e altura das caixas delimitadoras das regiões i e j . O parâmetro λ está relacionado com a distância entre as arestas das caixas delimitadoras (ver figura 3.13).

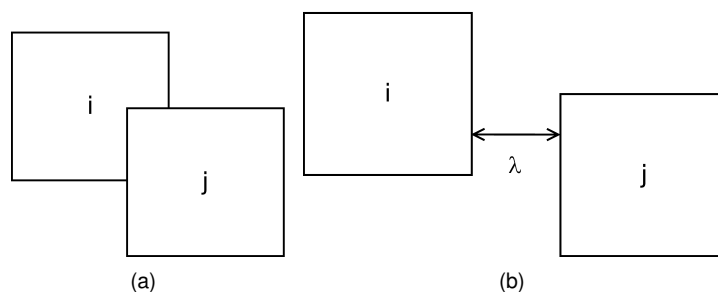


Figura 3.13: (a) Para verificar se caixas delimitadoras de duas regiões detectadas se intersectam, λ é definido como 1; (b) para verificar se estão a uma determinada distância, λ é definido com o valor da distância.

O algoritmo 1 sumariza o método de detecção de texturas dinâmicas e a subsequente segmentação das regiões de fogo, descritos nas secções 3.2.2 e 3.2.2.1, respectivamente.

Algoritmo 1 detectarObjectos

```

1: Entrada: conjunto de objectos detectados ( $N_{obj}$ ), máscara com regiões de píxeis com aparência
   semelhante ao fogo ( $I_{PC}$ ), máscara de filtragem temporal dos píxeis ( $I_{TD}$ ), máscara de remoção
   de contornos ( $I_{RC}$ )
2: Saída: conjunto actualizado de objectos detectados ( $N_{obj}$ )
3: Dados: sequência de píxeis ligados ( $C$ ), caixa delimitadora ( $rect$ ), intensidade média ( $med$ ),  $\kappa$  é
   uma constante definida empiricamente.
4:
5: detecção de texturas dinâmicas através da equação 3.20 com  $I_{TD}$  e  $I_{PC}$ 
6:
7: //remover a periferia das regiões de chamas
8: obter contornos  $C$  de  $I_{RC}$ 
9: foreach  $c \in C$  do
10:   pintar os píxeis  $p$  do contorno  $c$  a preto,  $\forall p \in c : I_{RC}(p) = 0$ 
11: end for
12:
13: //detecção de objectos e actualização dos parâmetros
14: foreach  $c \in C$  do
15:   encontrar caixa delimitadora do contorno  $c$ ,  $rect(c)$ 
16:   calcular intensidade média de  $I_{TD}$  delimitada por  $rect(c)$ ,  $med$ 
17:   if  $med > \kappa$  then
18:     foreach  $n \in N_{obj}$  do
19:       obter objecto  $n$  que se intersecta com  $rect(c)$  usando a equação 3.21 com  $\lambda = 1$ 
20:       actualizar objecto  $n$  (ver Secção 3.2.2.1)
21:     end for
22:     criar objecto  $n$  caso não existam objectos ou intersecções,  $N_{obj} \leftarrow N_{obj} \cup \{n\}$ 
23:   end if
24: end for
25: return  $N_{obj}$ 

```

3.2.3 Análise no Domínio da Frequência

Os algoritmos de detecção de fogo que se baseiam apenas nos métodos de detecção de regiões de píxeis com base na cor do fogo não têm a capacidade para distinguir uma região de chamas de um objecto em movimento com aparência semelhante ao fogo. No sentido de adicionar ao algoritmo proposto essa capacidade, são propostos métodos que analisam as características dinâmicas conhecidas do fogo no domínio da frequência através da transformada *wavelet* discreta.

Para diversas aplicações e, em particular os algoritmos de visão computacional, que necessitam de aceder ao conteúdo do sinal, o formato bruto do sinal torna-se inadequado, pois os conteúdos mais distintos de cada sinal estão embutidos na representação do sinal do domínio da frequência. Para tal, são aplicadas transformações ao sinal para aceder à informação nesse domínio.

Existem diversas transformações para analisar o sinal no domínio da frequência, destacando-se: (1) a transformada de *Fourier*, que fornece o conteúdo espectral do sinal, mas sem qualquer informação temporal; (2) transformada *wavelet* (*WT*), que fornece a informação de tempo e frequência simultaneamente, sendo adequada para extrair as características de sinais não estacionários (Rioul and Vetterli, 1991). A transformada *wavelet* discreta foi desenvolvida como uma alternativa à transformada de *Fourier* de curto espaço de tempo (*STFT*), que oferece resolução de tempo uniforme para todas as frequências, enquanto a *DWT* oferece diferentes resoluções para altas e baixas frequências (Tzanetakis et al., 2001).

A transformada *wavelet* discreta é um caso particular da *WT*, fornecendo uma representação compacta do sinal nos domínios do tempo e frequência discreto, sendo computacionalmente eficiente (Tzanetakis et al., 2001). Um caso particular de análise dos sinais com base na *DWT*, usado para analisar as características dinâmicas conhecidas do fogo no domínio da frequência, é implementado através de um banco de filtros passa alto e passa baixo (Mallat, 1989), filtrando o sinal original no domínio do tempo através das expressões:

$$Y_{alto}[k] = \sum_n x_f[n]g_f[2k - n] \quad (3.22)$$

$$Y_{baixo}[k] = \sum_n x_f[n]h_f[2k - n] \quad (3.23)$$

onde Y_{alto} e Y_{baixo} representam as saídas dos filtros passa alto e passa baixo após o processo de sub-amostragem no instante k , respectivamente. Os parâmetros g_f e h_f são as funções de resposta ao impulso de um filtro passa-baixo e passo-alto, respectivamente, cada um com metade da banda do sinal original. O parâmetro x_f representa o sinal original no domínio do tempo.

É importante mencionar que uma determinada região da imagem é classificada como fogo após ser validada pelos métodos de análise da *DWT* no espaço da frequência temporal (ver Secção 3.2.3.1) e espacial (ver Secção 3.2.3.2). Caso os dois métodos classifiquem uma região como fogo, os resultados dos algoritmos de detecção de fogo e detecção e seguimento de objectos são integrados de modo a eliminar os falsos alarmes provenientes de objectos em movimento com aparência semelhante ao fogo (ver Secção 3.5).

3.2.3.1 Transformada *Wavelet* Discreta no Espaço da Frequência Temporal

Uma das características dinâmicas do fogo é oscilar a uma frequência entre 1 e 10 Hz (Detriche and Lanore, 1980; Drysdale, 2011; Hammis et al., 1992), ou seja, a informação de cor de um determinado pixel na área circundante ao fogo pode desaparecer e aparecer várias vezes por segundo.

O quarto passo do algoritmo de detecção de fogo é uma extensão ao método proposto por Töreyn et al. (2006), onde a métrica de decisão e o modo como o método é aplicado são alterados. O objectivo deste método é analisar a história de oscilação dos píxeis considerados fogo pelo método de detecção de texturas dinâmicas no domínio da frequência, sendo aplicada a cada região com assinatura temporal do fogo uma análise da transformada *wavelet* discreta no espaço da frequência temporal.

Ao analisar o comportamento oscilatório do fogo, a sua detecção torna-se mais robusta, pois as regiões de píxeis segmentadas como movimento e classificadas como fogo poderão não evidenciar um regime de oscilação suficientemente alto para ser confundido com as chamas, tais como folhas das árvores, ervas, carros e pessoas. Por outro lado, com a inclusão do método proposto nesta dissertação em detrimento do método de Töreyn et al. (2006), o algoritmo de detecção de fogo torna-se mais eficiente, visto que este só é executado após a detecção de texturas dinâmicas.

Com vista a verificar o comportamento oscilatório das chamas através da *DWT* unidimensional, um banco de filtros de uma dimensão com dois níveis de transformação, onde o número de cascatas determina o nível de transformação, foi implementado. A figura 3.14 representa a estrutura básica do referido banco, composto por filtros passa alto e passa baixo com coeficientes $\{-0.25, 0.5, -0.25\}$ e $\{0.25, 0.5, 0.25\}$, respectivamente.

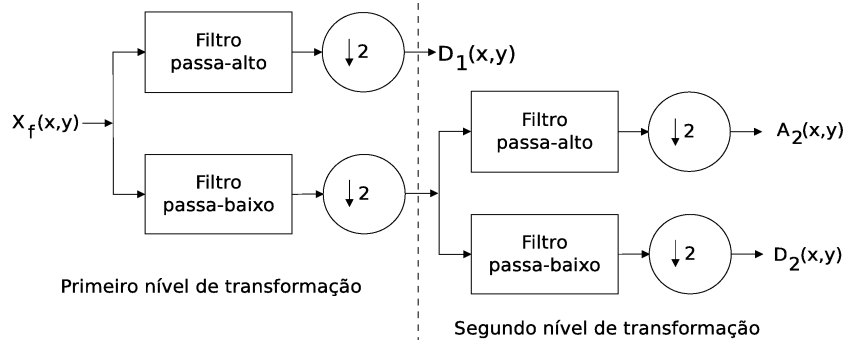


Figura 3.14: Estrutura básica de um banco de filtros com dois níveis de transformação, utilizado para análise da transformada *wavelet* discreta. As saídas do processo de transformação são dadas pelas escalas de aproximação A_n e de detalhe D_n .

O sinal de entrada x_f corresponde à componente vermelha das regiões de píxeis com cor de fogo (ver Secção 3.2.1). Assim, são recolhidas NC_{dwt} (definido como 60) imagens consecutivas dessa região, efectuando uma análise de cada pixel ao longo das imagens através da *DWT* no espaço da frequência temporal. É importante mencionar que as imagens recolhidas são criadas com o tamanho da primeira caixa delimitadora da região em análise.

Na figura 3.15 encontram-se expostos o sinal de entrada x_f no domínio do tempo (figura 3.15(a)), e os sinais de saída do banco de filtros unidimensional com dois níveis de transformação, dwt_{out} , dado pela concatenação da escalas D_1 e A_2 (figura 3.15(b)). De referir que na saída do banco de filtros ocorre o processo sub-amostragem, que reduz a taxa de amostragem dos sinais e, por motivos de apresentação, foi imposto o processo de sobre-amostragem que acrescenta zeros nas posições eliminadas pelo processo de sub-amostragem.

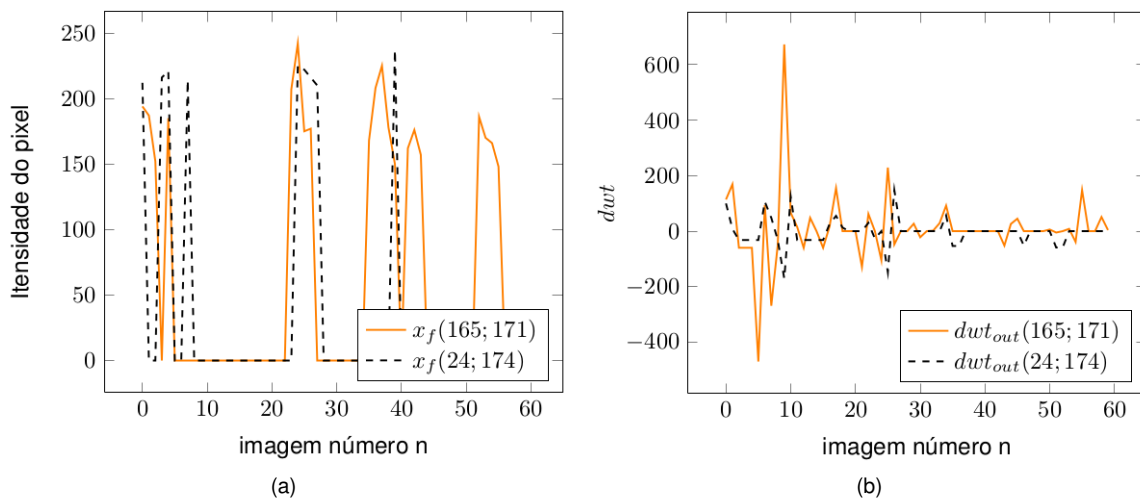


Figura 3.15: Sinais nos domínios do tempo (a) e de saída do banco de filtros de 2 níveis de transformação (b). Os píxeis nas posições (165, 171) e (24, 174) representam uma região de chamas e uma região com um objecto em movimento com cores semelhantes às das chamas, respectivamente.

Se um determinado pixel é considerado fogo pelo método de classificação dos píxeis com base na cor, então o seu sinal ao longo do tempo apresenta diversos picos, representando a transição entre o plano de fundo e as chamas (sinal $x_f(165, 171)$ da figura 3.15(a)), e a transição entre o plano de fundo e um objecto em movimento com cores semelhantes às das chamas (sinal $x_f(24, 174)$ da figura 3.15(a)). O pixel $x_f(165, 171)$ faz parte das chamas nas imagens $n = 1, 2, 4, 6, 22$ a $26, 34$ a $40, 41$ a 43 e 51 a 56 , tornando-se parte da imagem de fundo nas restantes imagens.

Na figura 3.15(b) estão expostos os sinais de saída do banco de filtros com dois níveis de transformação. Os sinais $dwt_{out}(165, 171)$ e $dwt_{out}(24, 174)$ representam a transformada *wavelet* discreta no espaço da frequência temporal de um pixel de fogo e de um pixel representativo de um objecto em movimento com aparência semelhante ao fogo, respectivamente.

Com base nos sinais de saída do banco de filtros, foram definidas regras para discriminar entre dois tipos de píxeis: um fogo ou um objecto em movimento com aparência semelhante ao fogo. Assim, para cada pixel da região detectada é analisado a intensidade e variação dos picos do correspondente sinal de saída do banco de filtros. Esta análise não é considerada no modelo original proposto por (Töreyn et al., 2006), onde apenas é tomado em consideração o número de passagens por zero. Com a inclusão deste processo pretende-se negligenciar a forte contribuição de alguns píxeis, isto é, evitar que os píxeis com elevada oscilação tenham uma ponderação mais elevada, de modo a não saturar da métrica de decisão.

Com base na métrica de decisão baseada na intensidade dos píxeis, cada pixel é classificado como fogo se o correspondente sinal de saída do banco de filtros satisfazer duas condições:

$$dwt_{out}(x, y) > \sigma \quad \wedge \quad N_{picos}(x, y) \leq 3, \quad (3.24)$$

onde o parâmetro N_{picos} representa o número total de picos acima do limiar σ na posição (x, y) . A figura 3.16 expõe dois exemplos do resultado deste processo, F_{sp} .

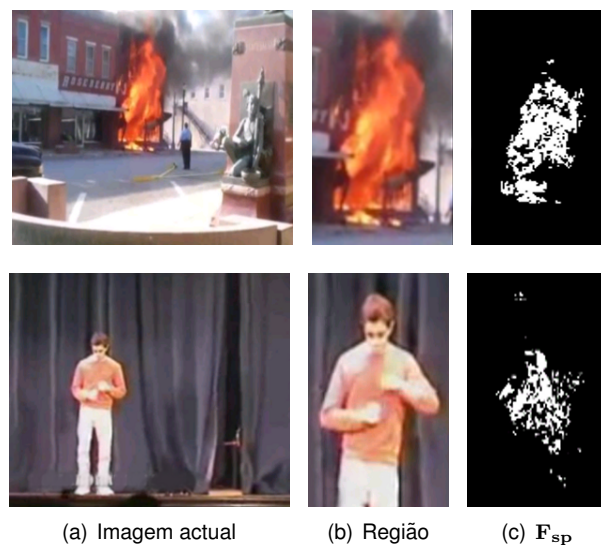


Figura 3.16: (a) Imagem actual; (b) região detectada com cores semelhantes às das chamas para análise da *DWT*; (c) máscara com os píxeis classificados como fogo pela métrica de decisão baseada na intensidade dos píxeis, os quais apresentam até três picos acima da amplitude σ no sinal de saída do banco de filtros.

Como referido, para não saturar a métrica de decisão e acrescentar uma maior robustez na detecção de fogo, são analisadas a intensidade e a variação dos picos do sinal de saída do banco de filtros. Neste sentido, uma região é considerada fogo se satisfazer duas condições: (1) o rácio dos píxeis classificados como fogo pela regra baseada na intensidade dos píxeis (ver equação 3.24) necessita de ser superior ao limiar thr_{int} (definido como 0.15); (2) o número acumulado de passagens por zero do sinal de saída do banco de filtros tem de ser superior ao limiar thr_{zc} , definido como três vezes a área da região em análise.

O método descrito para análise da transformada *wavelet* discreta no espaço da frequência temporal é sumariado no algoritmo 2.

Algoritmo 2 dwtTemporal

```

1: Entrada: conjunto de imagens recolhidas para aplicar a  $DWT (IR_{dwt})$ , caixa delimitadora do
   objecto ( $rect$ )
2: Saída: marca com o resultado da análise da  $DWT$  unidimensional ( $alarme$ )
3: Data:  $S_{in}$  e  $dwt_{out}$  são vectores auxiliares,  $thr_{int}$ ,  $thr_{zc}$  e  $NC_{dwt}$  são constantes definidas em-
   piricamente.
4:
5:  $F_{sp} \leftarrow \emptyset$ ,  $alarme \leftarrow falso$ 
6: for  $y = 0$  até altura  $rect$  do
7:   for  $x = 0$  até largura  $rect$  do
8:     //obter o sinal no domínio do tempo
9:      $V_{med} \leftarrow 0$ 
10:    for  $i = 0$  até  $NC_{dwt}$  do
11:       $S_{in}(i) \leftarrow IR_{dwt}(i, x, y)$ 
12:       $V_{med} \leftarrow V_{med} + S_{in}(i)$ 
13:    end for
14:
15:    //verificar se o valor médio do sinal é superior a zero para calcular a  $DWT$  unidimensional
16:    if  $V_{med} > 0$  then
17:      remover o valor médio do sinal,  $\forall s \in S_{in}, S_{in}(s) = S_{in}(s) - V_T$ ,  $V_T = \frac{V_{med}}{NC_{dwt}}$ 
18:      calcular a  $DWT$  unidimensional com base em  $S_{in}$ ,  $dwt_{out}$ 
19:      actualizar  $F_{sp}$  usando a equação 3.24 e  $dwt_{out}$ 
20:    end if
21:  end for
22: end for
23: if região  $rect$  valida as métricas de decisão de  $thr_{int}$  e  $thr_{zc}$  then
24:    $alarme \leftarrow verdadeiro$ 
25: end if
26: eliminar  $F_{sp}$ 
27: return  $alarme$ 

```

3.2.3.2 Transformada *Wavelet* Discreta no Espaço da Frequência Espacial

Para além de analisar a oscilação dos píxeis nas zonas circundantes às chamas através da *DWT* no espaço da frequência temporal, para reduzir ainda mais a probabilidade de gerar um falso alarme, a textura das chamas no domínio da frequência também é analisada, através da *DWT* no espaço da frequência espacial.

Töreysin et al. (2006) propôs um método de detecção da variação espacial dos píxeis, aplicando a *DWT* no espaço da frequência espacial através de um banco de filtro de duas dimensões com um nível de transformação. Contudo, para uma análise mais detalhada da textura dos píxeis considerados fogo e conseqüentemente uma maior robustez na detecção de fogo, um banco de filtro bidimensional com três níveis de transformação é considerado nesta dissertação. Para além disso, em vez de se aplicar um limiar para a energia dos píxeis numa imagem para diferenciar entre um fogo e um objecto em movimento com aparência semelhante ao fogo, o método proposto tem em consideração o valor máximo, mínimo e médio da variação de energia dos píxeis na sequência de imagens recolhidas para análise da *DWT*.

Através da concatenação de estruturas *DWT* unidimensional, obtidas a partir da estrutura definida na figura 3.14, uma nova estrutura para a análise da *DWT* bidimensional com três níveis de transformação pode ser conseguida. O objectivo da concatenação é aumentar a resolução da frequência e os coeficientes de aproximação decompostos com filtros passa alto e passa baixo. Na figura 3.17 estão expostas as sub-imagens de decomposição de uma imagem, quando uma *DWT* bidimensional com três níveis de transformação é aplicada. As sub-imagens correspondem a diferentes níveis de resolução e orientações (Antonini and Barlaud, 1992).

L_b^3	$L_b^3 H_b^3$	$L_b^2 H_b^2$	$L_b^1 H_b^1$
$H_b^3 L_b^3$	$H_b^3 H_b^3$		
$H_b^2 L_b^2$	$H_b^2 H_b^2$		
$H_b^1 L_b^1$		$H_b^1 H_b^1$	

Figura 3.17: Sub-imagens de decomposição resultantes da *DWT* bidimensional com três níveis de transformação (1, 2 e 3). As sub-imagens a cinzento são utilizadas para a discriminação entre um região com chamas e uma região com um objecto em movimento com aparência semelhante ao fogo. H_b e L_b representam as bandas de alta e baixa frequência, respectivamente.

Através do sinal de saída do banco de filtros com três níveis de transformação e com base nas NC_{dwt} imagens recolhidas, é possível analisar a energia dos píxeis que apresentam uma aparência semelhante ao fogo:

$$\beta(x, y) = \frac{1}{N_{dwt}} \sum_{x,y} |\mathbf{x}_{lh}[x, y]|^2 + |\mathbf{x}_{hl}[x, y]|^2 + |\mathbf{x}_{hh}[x, y]|^2, \quad (3.25)$$

onde x_{lh} , x_{hl} e x_{hh} são as sub-imagens resultantes da *DWT* bidimensional na posição (x, y) , denominadas sub-imagens baixo-alto, alto-baixo e alto-alto, respectivamente, ilustradas nas figuras 3.17 (a cinzento) e 3.18(d). O parâmetro N_{dwt} representa o número de píxeis considerados fogo pelo método de classificação dos píxeis com base na cor na região em análise.

O algoritmo de detecção de fogo classifica uma região como fogo se os valores mínimo, médio e máximo do parâmetro β se encontram acima dos limiares ξ_1 , ξ_2 e γ (parametrizadas respectivamente para 1, 0.01 e 50) ao longo das 60 imagens recolhidas para a análise da *DWT*. A figura 3.18 ilustra dois exemplos de aplicação da *DWT* bidimensional. De referir que o contraste das sub-imagens foi ajustado para uma melhor observação.

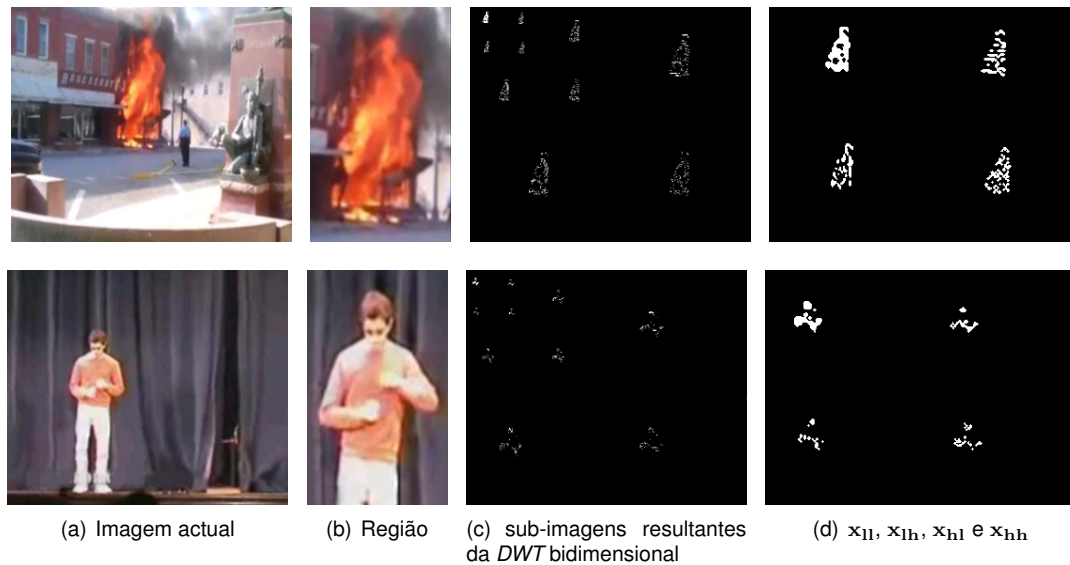


Figura 3.18: (a) imagem actual; (b) região detectada com aparência semelhante ao fogo para análise da *DWT* no espaço da frequência espacial; (c) sub-imagens resultantes da aplicação da *DWT* bidimensional com três níveis de transformação; (d) sub-imagens de decomposição x_{ll} , x_{lh} , x_{hl} e x_{hh} .

Os algoritmos 3 e 4 sumarizam o método de análise da transformada *wavelet* discreta no espaço da frequência espacial e o algoritmo de detecção de fogo, respectivamente.

Algoritmo 3 *dwtEspacial*

- 1: **Entrada:** conjunto de imagens recolhidas para aplicar a *DWT* (IR_{dwt})
 - 2: **Saída:** marca com o resultado da análise da *DWT* bidimensional (*alarme*)
 - 3: **Data:** dwt_{out} é um vector auxiliar, NC_{dwt} e β são constantes definidas empiricamente.
 - 4:
 - 5: $alarme \leftarrow falso$
 - 6: **for** $x = 0$ até NC_{dwt} **do**
 - 7: calcular *DWT* bidimensional com base em $IR_{dwt}(x)$, dwt_{out}
 - 8: usar equação 3.25 para analisar o sinal dwt_{out}
 - 9: **end for**
 - 10: **if** dwt_{out} valida a métrica de decisão do parâmetro β **then**
 - 11: $alarme \leftarrow verdadeiro$
 - 12: **end if**
 - 13: **return** $alarme$
-

Algoritmo 4 detectarFogo

```

1: Entrada: conjunto de objectos detectados ( $N_{obj}$ ), conjunto de objectos considerados fogo ( $N_f$ ),
   imagem actual ( $I_s$ ), máscara com regiões de movimento ( $I_{HBE}$ ), máscara com regiões de píxeis
   com cores semelhantes às das chamas ( $I_{PC}$ ), máscara de filtragem de texturas dinâmicas ( $I_{TD}$ ),
   máscara de remoção de contornos ( $I_{RC}$ )
2: Saída: conjunto actualizado de objectos considerados fogo ( $N_f$ )
3: Data: primeira caixa delimitadora do objecto ( $rect$ ), conjunto de imagens recolhidas para aplicar
   a  $DWT$  ( $IR_{dwt}$ ),  $NC_{dwt}$  é uma constante definida empiricamente.
4:
5: determinar as regiões de píxeis com cor de fogo usando  $I_s$  e  $I_{HBE}$  (ver Secção 3.2.1)
6: detectar regiões de píxeis,  $N_{obj} \leftarrow \mathbf{detectarObjectos}(N_{obj}, I_{PC}, I_{TD}, I_{RC})$  (ver Secção 3.2.2.1)
7:
8: //análise das características dinâmicas conhecidas do fogo no domínio da frequência
9: foreach  $n \in N_{obj}$  do
10:  //verificar se já foram recolhidas as  $NC_{dwt}$  para aplicar a  $DWT$ 
11:  if  $|IR_{dwt}(n)| < NC_{dwt}$  then
12:    obter a componente vermelho de  $I_s$ 
13:    definir região de interesse,  $ROI(I_s, rect(n))$ 
14:    guardar imagem,  $IR_{dwt}(n) \leftarrow IR_{dwt}(n) \cup \{I_s\}$ 
15:  else
16:    //validar regiões de chamas
17:    aplicar  $DWT$  unidimensional,  $alarme \leftarrow \mathbf{dwtTemporal}(IR_{dwt}(n), rect(n))$  (ver
    Secção 3.2.3.1)
18:    if  $alarme = verdadeiro$  then
19:      aplicar  $DWT$  bidimensional,  $alarme \leftarrow \mathbf{dwtEspacial}(IR_{dwt}(n))$  (ver Secção 3.2.3.2)
20:      if  $alarme = verdadeiro$  then
21:         $N_f \leftarrow N_f \cup \{n\}$ 
22:      end if
23:    end if
24:    eliminar  $IR_{dwt}(n)$ 
25:  end if
26: end for
27: return  $N_f$ 

```

3.3 Algoritmo de Detecção e Seguimento de Objectos

O processo de detectar e seguir objectos em tempo real é fundamental para sistemas de vídeo vigilância. É importante notar que o seguimento de objectos apresenta um grande desafio para o sistema, devido à presença de ruído, alterações dinâmicas do ambiente e oclusões entre objectos.

Para além de seguir múltiplos objectos presentes no campo de visão da câmara e gerar os respectivos alarmes, os resultados deste processo e da detecção de regiões com chamas são integrados (ver Secção 3.5), de modo a eliminar os falsos alarmes na detecção de fogo provenientes de objectos em movimento com aparência semelhante ao fogo.

O algoritmo proposto nesta dissertação para a detecção e seguimento de objectos é constituído por um conjunto de métodos. Em primeiro lugar, são extraídas regiões de movimento através de uma extensão ao método proposto por Kim et al. (2005) e do método para estimação do plano de fundo (Collins et al., 2000). Com base nestas regiões são removidas as sombras dos objectos em movimento através de uma extensão ao método proposto por Cucchiara et al. (2003). Posteriormente, os objectos em movimento são detectados de forma a gerir os seus parâmetros de identificação. Caso um objecto seja detectado um determinado número de vezes, este é seguido através de um filtro de partículas, que estima a posição do objecto com base no modelo aparência do mesmo. Por fim, o objecto é definido como humano através de um classificador proposto por (Dadal and Triggs, 2005). As próximas secções detalham este processo.

3.3.1 Subtracção do Plano de Fundo

A capacidade de extrair objectos em movimento através de uma sequência de imagens é o primeiro passo normalmente efectuado pelos algoritmos de seguimento de objectos. No sistema proposto, a técnica utilizada para este fim é a subtracção do plano de fundo, que tem como objectivo apresentar as diferenças entre a imagem actual e o modelo do plano de fundo com base na história dos píxeis, identificando assim regiões de movimento.

O método de subtracção do plano de fundo proposto por Kim et al. (2005) constrói um modelo do plano de fundo, representando estados deste através de *codebooks*, após a observação de uma sequência de imagens. Para cada pixel, é criado um *codebook* e através de uma métrica de variação de cor e brilho, um pixel da imagem actual é classificado como pertencente ao plano de fundo.

Na prática, um *codebook* é constituído por caixas que delimitam valores de intensidade de um pixel (ver figura 3.19). Inicialmente, a caixa é criada para um determinado valor, aumentando lentamente a sua área de modo a cobrir a variação de intensidade do pixel ao longo do tempo. Caso o valor do pixel varie significativamente, é criado um novo *codebook*. De referir que as caixas cobrem as três dimensões (canais) de cada pixel da imagem no espaço de cores *RGB*.

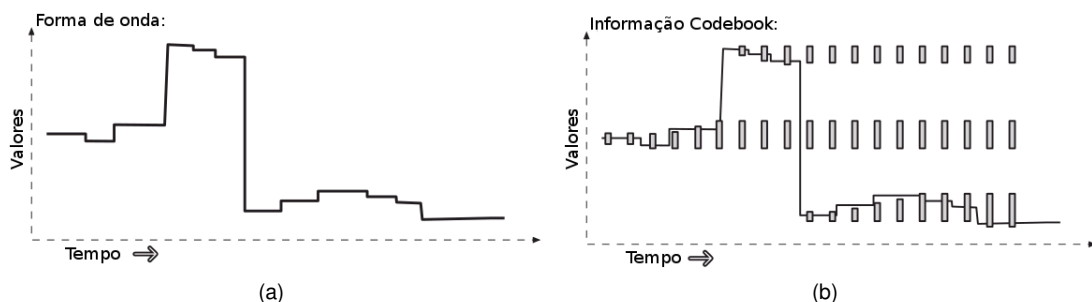


Figura 3.19: Forma de onda de um pixel ao longo do tempo (a), sobreposta com a variação dos codebooks (b) (adaptado de Bradski and kaehler (2008)).

Através de uma fase inicial de aprendizagem, o método proposto por Kim et al. (2005) aprende um modelo do plano de fundo. Através de novas observações e com base na história dos píxeis, este é iterativamente actualizado. No entanto, quando existe movimento na imagem actual o modelo não é aprendido, tornando-o inadequado para ambientes com variações de luminosidade e variações das condições dinâmicas, tais como sombras e movimentos das folhas das árvores.

Para colmatar este problema, esta dissertação propõe duas extensões ao método proposto por Kim et al. (2005), onde uma grelha regular é sobreposta à imagem actual (ver figura 3.20), associando a cada posição da grelha um modelo do plano de fundo. Para além disso, o processo de aprendizagem foi alterado, de modo a actualizar cada modelo independentemente das condições existentes nos restantes. Assim, cada modelo é actualizado caso valide as seguintes condições:

1. Não existem objectos a serem seguidos pelo sistema na região do modelo. Com base na máscara de localização dos objectos a serem seguidos, I_{PF} (ver Secção 3.3.3), verifica-se se existem píxeis a branco na região do modelo.
2. Não existem mais que thr_{BS} píxeis em movimento na região do modelo (parametrizado para 50). Esta constante é obtida através do somatório do número de píxeis em movimento provenientes do método *HBE* (ver Secção 3.1) ao longo das últimas 50 imagens. Com a inclusão desta condição, pretende-se que o método não aprenda objectos em movimento presentes no ambiente que não consegue detectar.
3. No instante da actualização, o método de subtracção do plano de fundo necessita de conter um número de píxeis em movimento superior a thr_{BS} .

De modo a reduzir o custo computacional associado a este método, o processo de actualização só é executado após uma sequência de $update_{BS}$ imagens (definido como 50) ou caso exista uma grande variação de luminosidade.

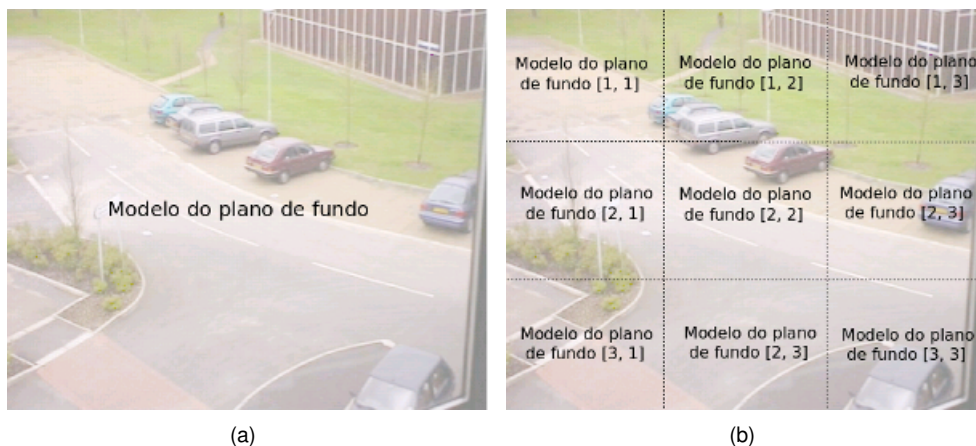


Figura 3.20: Visão global do modelo do plano de fundo. (a) Modelo do plano de fundo do método proposto por Kim et al. (2005); (b) disposição dos modelos do plano de fundo sobre uma grelha regular, de acordo com o proposto nesta dissertação.

Outro dos problemas do método proposto por (Kim et al., 2005) está relacionado com o tipo de movimento dos objectos, isto é, caso um objecto fique imóvel no ambiente da câmara durante um determinado número de imagens consecutivas, este será aprendido pelo modelo do plano de fundo e incorporado no mesmo. Caso este se mova, será novamente considerado como objecto do primeiro plano. No entanto, na região onde o objecto foi considerado como pertencente ao plano de fundo, o método não apresentará uma segmentação satisfatória de objectos com aparência semelhante a este. Para além de corrigir o problema associado à variação das condições do ambiente, as extensões propostas nesta dissertação permitem mitigar este tipo de problemas.

Como referido anteriormente, a aprendizagem inicial deste método é necessária para construir o modelo do plano de fundo. Durante esta fase, assume-se que não existem objectos em movimento, de modo a aprender apenas os elementos estáticos presentes no ambiente. Nesta dissertação foi considerada uma sequência de N_{initBS} imagens (parametrizado para 250) para a referida aprendizagem. O resultado deste método aplicado a ambientes exteriores está representado na figura 3.21.

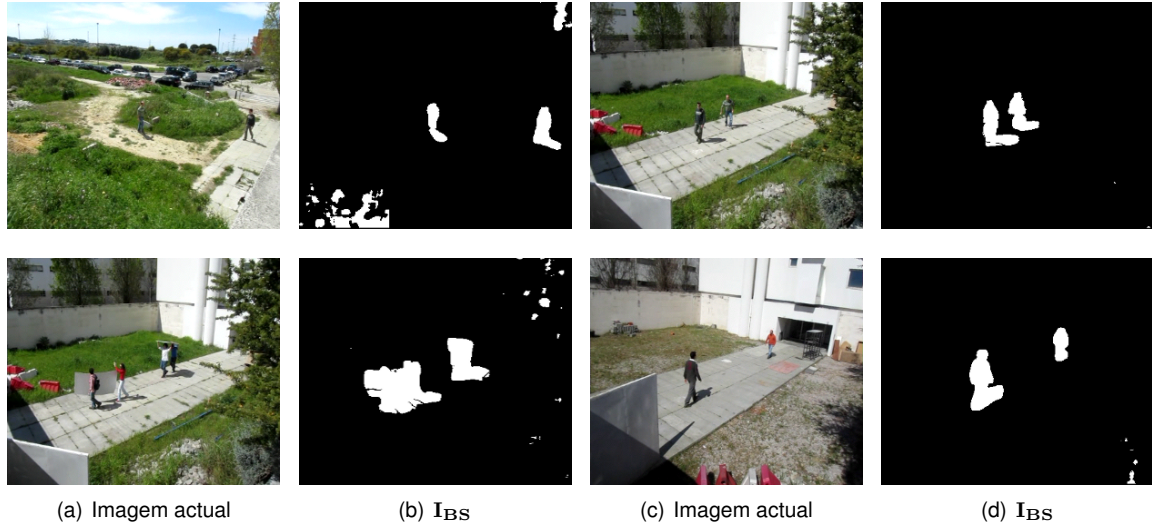


Figura 3.21: Detecção de regiões de movimento através da subtracção do plano de fundo. (a) e (c) Imagem actual; (b) e (d) máscaras resultante do método de subtracção do plano de fundo proposto, I_{BS} .

É importante notar que as alterações das condições de luminosidade na imagem têm um efeito catastrófico no método de subtracção do plano de fundo, visto que existe uma grande variação nos *codebooks* de cada pixel. A figura 3.22 apresenta 2 exemplos de aprendizagem do plano de fundo num ambiente exterior.

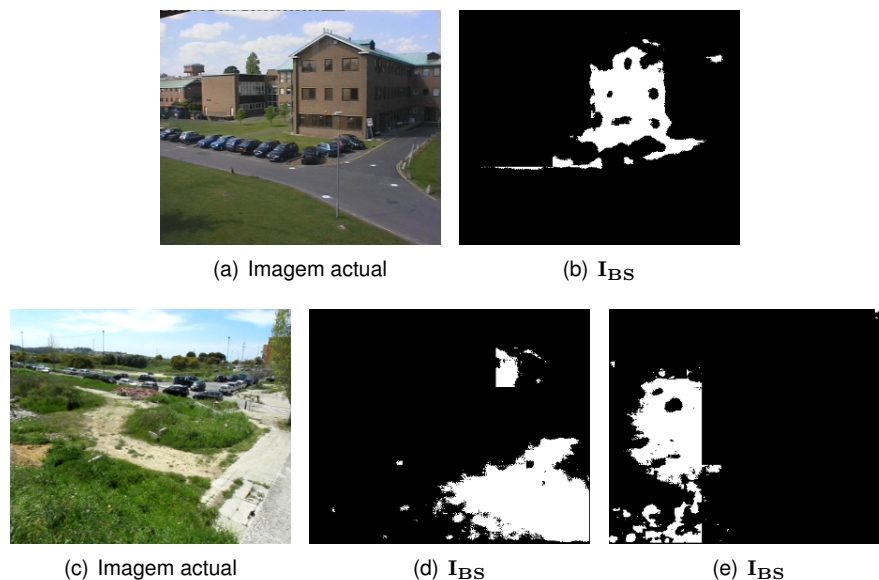


Figura 3.22: (a) e (c) Imagem actual; (b) situação onde a variação das condições de luminosidade numa determinada região afectam o método de subtracção do plano de fundo, levando-o a segmentar o prédio como objecto em movimento. Com base no método de aprendizagem proposto, o prédio é posteriormente considerado como objecto do plano de fundo, visto que cumpre as condições acima referidas. (d) e (e) sequência de imagens de uma situação de variação da exposição da câmara, isto é, quantidade de luz absorvida pelo sensor. Verifica-se que cada modelo é actualizado, independentemente do movimento existente nas restantes regiões.

O algoritmo 5 sumariza o processo de actualização dos modelos do plano de fundo proposto nesta dissertação.

Algoritmo 5 actualizarBS

```

1: Entrada: conjunto de codebooks ( $C_{CB}$ ), número da imagem actual ( $N_{frame}$ ), imagem filtrada ( $I_{blur}$ ), máscara de diferenciação temporal ( $I_{HBE}$ ), máscara com as caixas delimitadoras dos objectos seguidos ( $I_{PF}$ ), máscara de subtracção do plano de fundo ( $I_{BS}$ ), vector auxiliar (temp)
2: Saída: conjunto actualizado de codebooks ( $C_{CB}$ )
3: Data: caixa delimitadora (rect),  $update_{BS}$  e  $thr_{BS}$  são constantes definidas empiricamente.
4:
5:  $model \leftarrow 0$ ,  $w \leftarrow \frac{1}{3} \cdot largura(I_{blur})$ ,  $h \leftarrow \frac{1}{3} \cdot altura(I_{blur})$ 
6: for  $y = 0$  até 3 do
7:   for  $x = 0$  até 3 do
8:     definir região de interesse,  $ROI(I_{blur}, rect(x.w, y.h, w, h))$ ;
9:     definir região de interesse,  $ROI(I_{BS}, rect(x.w, y.h, w, h))$ ;
10:    segmentar píxeis do primeiro plano com base em  $C_{CB}(model)$ ,  $I_{blur}$  e  $I_{BS}$ 
11:
12:    //actualizar modelo do plano de fundo
13:    obter número de píxeis brancos,  $wp$ , de  $I_{HBE}$  na posição  $rect(x.w, y.h, w, h)$ ;
14:    obter número de píxeis brancos,  $wp_1$ , de  $I_{PF}$  na posição  $rect(x.w, y.h, w, h)$ ;
15:    obter número de píxeis brancos,  $wp_2$ , de  $I_{BS}$ 
16:     $temp(model) \leftarrow temp(model) + wp$ ;
17:    if  $Modulo(N_{frame}, update_{BS}) = 0$  then {//modelo actualizado de 50 em 50 imagens}
18:      if  $wp_1 = 0$  e  $wp_2 \geq thr_{BS}$  e  $temp(model) \leq thr_{BS}$  then
19:        actualizar o modelo de plano de fundo,  $C_{CB}(model)$  baseado em  $I_{blur}$ 
20:        limpar a entrada de codebooks com base em  $C_{CB}(model)$ 
21:      end if
22:       $temp(model) \leftarrow 0$ 
23:    end if
24:     $model \leftarrow model + 1$ 
25:    remover regiões de interesse de  $I_{blur}$  e  $I_{BS}$ 
26:  end for
27: end for
28: return ( $C_{CB}$ )

```

3.3.2 Detecção de Objectos e Remoção de Sombras

O algoritmo de detecção e seguimento de objectos proposto tem como objectivo seguir múltiplos objectos em movimento. Para tal, é necessário agrupar regiões de píxeis em movimento (objectos) e executar a correspondência entre estas regiões ao longo do tempo. Esta operação, denominada detecção de objectos, é utilizada para analisar o movimento dos objectos e extrair as suas características para que estes possam ser seguidos pelo sistema.

A detecção de objectos é baseada na máscara de subtracção do plano de fundo, na qual são aplicadas operações morfológicas de erosão e dilatação para limpar a imagem. Estas são responsáveis

por eliminar regiões isoladas ou detalhes irrelevantes que contenham poucos píxeis e para preenchimento de pequenos buracos. Com base nesta máscara, o algoritmo de componentes ligados é executado para extrair os objectos, identificados através dos parâmetros:

- Identificador do objecto através de um valor numérico único.
- Posição actual da caixa delimitadora do objecto, definida pela sua posição na imagem, altura e largura. É através da localização destas caixas que se identificam diferentes objectos.
- Número de imagens que o objecto foi detectado. Se este for detectado 20 vezes, é seguido pelo sistema (ver Secção 3.3.3).
- A cada detecção do objecto é registado o número da imagem actual. Assim, um objecto é eliminado do sistema caso não seja detectado nas últimas 100 imagens.
- Marcas com informação se o objecto já se encontra a ser seguido pelo sistema e se encontra em regiões interditas (ver Secção 3.4).

A figura 3.23 ilustra dois exemplos de detecção de objectos, onde os rectângulos a branco representam as suas caixas delimitadoras.

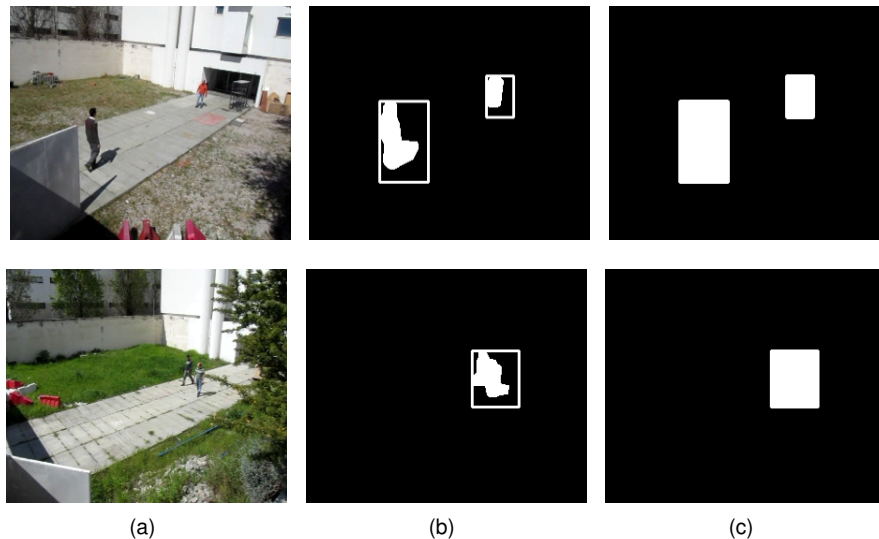


Figura 3.23: Visão global da detecção de objectos. (a) Imagem actual; (b) máscara de subtracção do plano de fundo com as caixas delimitadoras dos objectos detectados sobrepostas; (c) máscara com as caixas delimitadoras dos objectos detectados. Na segunda linha, dois objectos em oclusão são detectados através de uma única região. Apesar disso, caso as pessoas forem detectadas 20 vezes, são seguidas pelo sistema. Caso estas se afastem o suficiente para serem detectadas duas regiões, é criado um novo objecto sobre a região detectada que não esteja a ser seguida.

Os objectos em movimento variam a sua posição ao longo do tempo. Assim, a cada iteração do algoritmo é analisada a variação da localização das caixas delimitadoras dos objectos detectados, e verificado se existem intersecções entre estas e as dos objectos previamente detectados. Caso exista intersecção, são actualizados os parâmetros de identificação do objecto. Caso contrário, um novo objecto é criado se validar as seguintes condições:

1. Apresente uma altura dentro do intervalo estimado pela calibração do sistema (ver Secção 3.4),

$$\mathbf{H}_{\text{calib}}(x, y) - \mathbf{H}_{\text{calib}}(x, y) \cdot \zeta \leq h_{\text{obj}} \leq \mathbf{H}_{\text{calib}}(x, y) + \mathbf{H}_{\text{calib}}(x, y) \cdot \zeta, \quad (3.26)$$

onde H_{calib} representa a calibração da altura esperada dos objectos para cada posição (x, y) . Os parâmetros h_{obj} e ζ representam a altura do objecto detectado e a constante de ajuste do intervalo da altura esperada do objecto (parametrizado para $2/5$), respectivamente. Com a inclusão desta condição no sistema proposto, as regiões de movimento provenientes das folhas das árvores, sombras ou alterações de luminosidade não são detectadas nem processadas, tornando o sistema mais robusto e eficiente.

2. Não esteja em intersecção com regiões onde não se pretende detectar objectos (ver Secção 3.4).
3. Se as caixas delimitadoras do objecto detectado e de um objecto a ser seguido pelo sistema não se encontrem em intersecção.

Os objectos detectados contêm as sombras provenientes do seu movimento. De modo a que estas não afectem o processo de seguimento, que é baseado nas características de aparência dos objectos, surge a necessidade de as remover. Para tal, é proposto uma extensão ao método de Cucchiara et al. (2003), no qual a aparência do pixel do primeiro plano e do plano de fundo na mesma posição são comparados no espaço de cores *HSV*. Assim, para cada pixel pertencente à região do objecto detectado, este classifica os píxeis como sombra com base em três condições:

$$\varepsilon \leq \frac{\mathbf{I}_s(x, y)}{\mathbf{I}_{\text{BS}}(x, y)} \leq \nu \wedge |\mathbf{I}_s(x, y)S - \mathbf{I}_{\text{BS}}(x, y)S| \leq \tau_S \wedge H \leq \tau_H, \quad (3.27)$$

onde \mathbf{I}_s e \mathbf{I}_{BS} representam os valores dos píxeis na posição (x, y) na imagem actual e na máscara de subtracção do plano de fundo, respectivamente. Segundo Cucchiara et al. (2003), as componentes H , S e V do espaço de cores *HSV* encontram-se dentro de determinados limites e assim, as constantes ε , ν , τ_H e τ_S foram parametrizadas nesta dissertação para 0.5, 1, 20 e 50, respectivamente, com ε e ν contidos no intervalo $[0, 1]$. A máscara resultante deste processo, \mathbf{I}_{SR} , contém a região do objecto sem os píxeis considerados sombra.

O limite inferior ε é usado para definir o valor máximo do efeito de escurecimento das sombras no plano de fundo e é proporcional à intensidade da fonte de luz. O limite superior ν impede que o método identifique sombras em regiões que o fundo escureceu muito pouco em relação ao efeito esperado pela presença de sombras (Cucchiara et al., 2003).

O método proposto por Cucchiara et al. (2003) é aplicado na totalidade da imagem actual, tornando-o demasiado pesado computacionalmente. Em alternativa, nesta dissertação é removida a sombra da região inferior do objecto, mais concretamente um terço da sua altura (ver figura 3.24).



Figura 3.24: (a) Região do objecto em movimento; (b) região utilizada para remoção de sombras no método proposto.

Na figura 3.25 estão expostos dois exemplos dos resultados dos métodos de subtracção do plano de fundo e de remoção de sombras, propostos nesta dissertação.

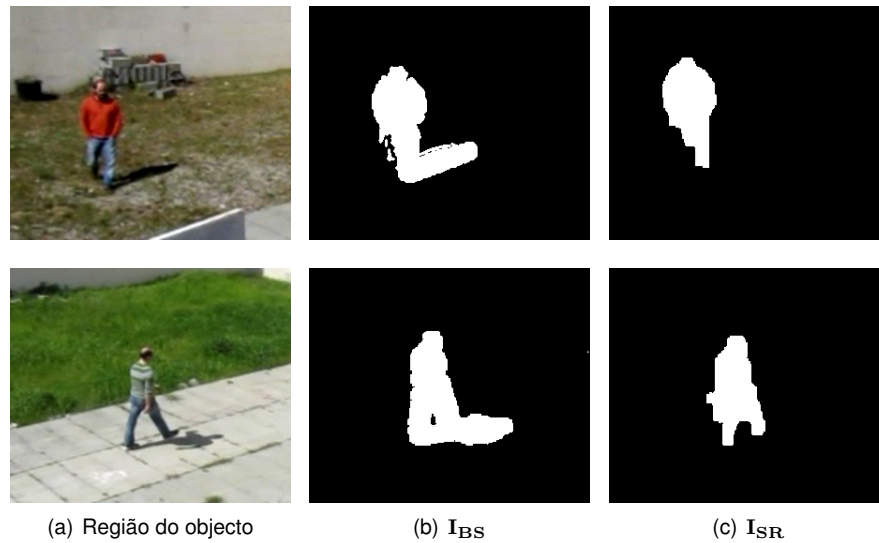


Figura 3.25: Resultado do método de remoção de sombras proposto. (a) Região do objecto detectado. Máscaras resultante do método de subtracção do plano de fundo (b) e do método de remoção de sombras (c), propostos nesta dissertação. Analisando as imagens I_{BS} e I_{SR} , verifica-se que o método proposto remove correctamente as sombras dos objectos e, para além disso, a informação relacionada com a sua aparência que é eliminada é irrelevante.

3.3.3 Seguimento de Múltiplos Objectos

O processo de seguimento de objectos com recurso a imagens de sistemas de vídeo vigilância em ambientes exteriores necessita de ser robusto o suficiente para lidar com oclusões entre objectos, alterações na direcção do movimento dos mesmos e variação das condições de luminosidade. Para tal, o algoritmo proposto recorre a filtros de partículas que têm a capacidade de lidar com distribuições não-gaussianas presentes neste tipo de ambientes, estimando assim a posição dos objectos com base na sua aparência. Esta secção descreve detalhadamente este processo.

3.3.3.1 Filtro de Partículas

O filtro de partículas, baseado no método de Monte Carlo (Doucet et al., 2001), permite estimar iterativamente uma determinada característica de um objecto. Com base no conjunto de partículas que representa o estado actual $\chi_i(m)$, nas partículas da última iteração $\chi_i(m-1)$ e através de uma observação $y_{obs}(m)$, é atribuído um peso $w_i^{[m]}$ a cada partícula. Este processo é efectuado através do modelo da dinâmica que simula a função densidade de probabilidade, no qual é aplicado ruído aleatório para espalhar as partículas em torno da posição do objecto e assim, representar a incerteza do resultado da simulação.

O sistema proposto recorre a filtros de partículas baseados na amostragem e reamostragem por importância das partículas, definindo a função densidade de probabilidade com base na aparência do objecto. Neste género de filtros, o processo de reamostragem das partículas com base nos

seus pesos é fundamental, pois as partículas com o peso mais baixo, isto é, as mais afastadas das regiões com maior probabilidade, são eliminadas e substituídas por novas partículas concentradas nas regiões de alta probabilidade. Caso este processo não fosse executado quando o número efectivo de partículas fosse, como considerado nesta dissertação, inferior a metade do número total de partículas, as regiões com elevada probabilidade teriam algumas partículas com pesos elevados e as restantes regiões partículas com pesos próximos de zero.

A figura 3.26 descreve globalmente o processo de iteração de um filtro de partículas e o processo de reamostragem das partículas de menor peso.

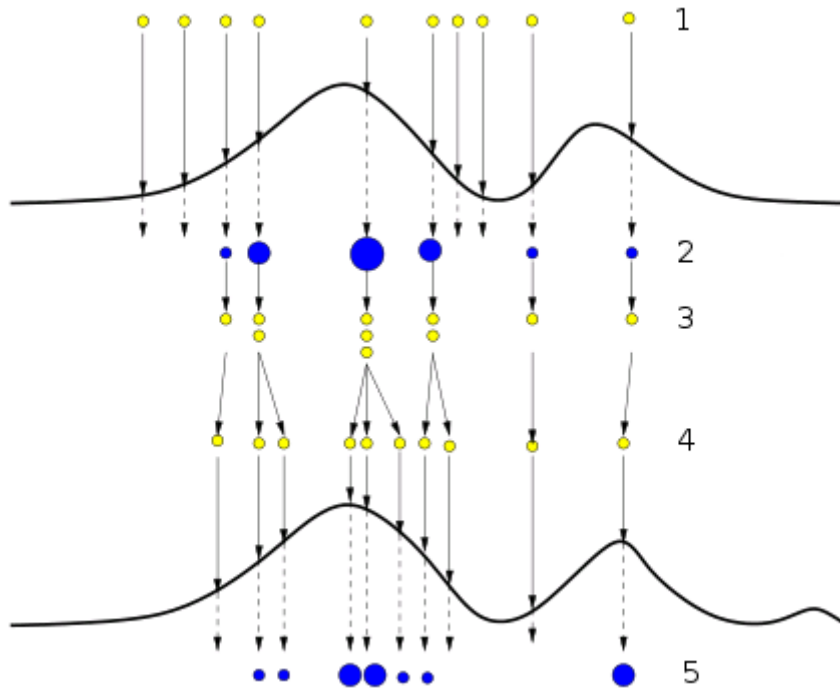


Figura 3.26: Etapas de execução de um filtro de partículas: (1) partículas com peso uniforme no instante $t = 0$; (2) para cada partícula, o seu peso é actualizado utilizando a informação no instante $t - 1$; (3) processo de reamostragem - as partículas com menor peso são eliminadas e as partículas com maior peso dão origem a mais partículas, para as regiões com maior probabilidade apresentarem uma maior quantidade de partículas; (4) partículas com peso uniforme no instante $t = t + 1$, derivadas do processo de reamostragem; (5) por fim, é executado o processo de amostragem das partículas (adaptado de Van Der Merwe et al. (2001)).

O objectivo do sistema proposto é seguir múltiplos objectos em movimento, mesmo que estes fiquem oclusos por outros objectos ou mudem de direcção. Para tal, é necessário ajustar os parâmetros do filtro de partículas:

1. O filtro é inicializado sobre um determinado objecto quando este tiver sido detectado 20 vezes. Assim, mesmo que um objecto tenha sido incorrectamente detectado, este necessita de ser iterativamente detectado para ser seguido pelo sistema.
2. O filtro é inicializado se as máscaras resultantes do método de subtracção do plano de fundo e do método *HBE* (ver Secção 3.1) contiverem movimento na região do objecto no instante da inicialização, e o objecto não esteja em intersecção com regiões interditas (ver Secção 3.4).

Devido ao facto de o método de subtracção do plano de fundo ajustar modelos de séries temporais a grupos de píxeis, faz com que possam existir regiões que não apresentem movimento

mas ainda estejam no primeiro plano. Assim, com a inclusão do método de estimação do plano de fundo no processo de inicialização de filtros de partículas, garante-se que este apenas é inicializado numa determinada região quando existe efectivamente movimento na mesma, pois o método *HBE* é baseado na diferenciação temporal de imagens.

3. O processo de actualização iterativo do modelo do filtro de partículas, que depende das máscaras de movimento (ver em baixo), é actualizado se o número de píxeis em movimento na região do objecto for superior a ρ (parametrizado para 50).
4. O processo de aplicar ruído aleatório às partículas foi alterado, de modo a adicionar ruído uniforme aditivo com base nas últimas localizações do objecto.

No filtro proposto, as partículas são representadas por 4 dimensões: posição na imagem, largura e altura do objecto detectado. Para além disso, a actualização iterativa do ruído das partículas é efectuado com base no movimento do objecto, espalhando assim as partículas para uma região próxima do objecto na iteração seguinte do algoritmo. Neste sentido, é adicionado ruído uniforme aditivo às partículas, R_{pf} , dado pela expressão:

$$R_{pf} = \begin{bmatrix} -r_0 + X_r & r_0 + X_r & -r_1 + Y_r & r_1 + Y_r \\ -r_2 & r_2 & -r_2 & r_2 \end{bmatrix}, \quad (3.28)$$

onde r_0 , r_1 e r_2 representam o ruído atribuído às 4 dimensões das partículas, isto é, posição na imagem, largura e altura, respectivamente, sendo parametrizados para 4, 3 e 2. Estes parâmetros são dinamicamente alterados em função do estado do objecto a ser seguido (ver Secção 3.3.3.2), isto é, se o objecto se encontra em intersecção com outro objecto em movimento, os parâmetros do ruído são alterados de modo a espalhar as partículas em torno dos objectos, sendo definidos como 8, 4 e 1.

Os parâmetros X_r e Y_r representam o movimento do objecto pelo ambiente, dado por:

$$X_r = \frac{x_{obj}(i) - x_{obj}(i - mov_{rpf})}{mov_{rpf}} \quad \wedge \quad Y_r = \frac{y_{obj}(i) - y_{obj}(i - mov_{rpf})}{mov_{rpf}}, \quad (3.29)$$

onde mov_{rpf} indica o número de posições do objecto necessárias para representar o seu movimento (definido como 4). O parâmetro x_{obj} e y_{obj} representam a localização do objecto na imagem no instante i , dado pela média das partículas com melhor pontuação (ver figura 3.27).

Nos métodos que recorrem a estes filtros, é necessário ter em consideração o número de partículas. Quanto menor for o seu número, menor é a complexidade do sistema e mais rápido se torna. Por outro lado, ao usar mais partículas, o processo de estimação com base nas características do objecto torna-se mais robusto. Nesta dissertação foram consideradas 100 partículas.

Na figura 3.27 estão expostos exemplos de um objecto a ser seguido com recurso a partículas de 4 dimensões. Os rectângulos a preto representam as caixas delimitadoras de cada partícula e os rectângulos a cor a média das partículas com melhor pontuação.

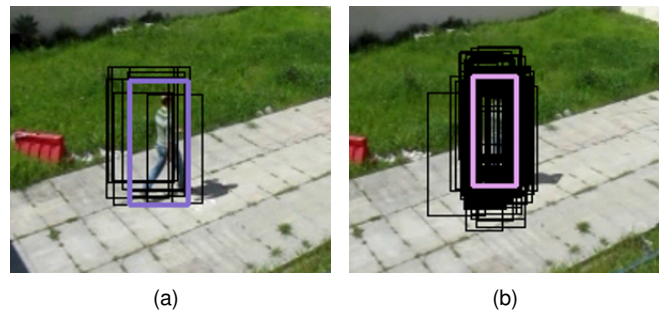


Figura 3.27: Filtro de partículas de 4 dimensões que recorre a 10 (b) e 100 (c) partículas para estimar a posição do objecto. Nos dois casos, o rectângulo que representa a posição do objecto na iteração actual está centrado com este, embora o filtro que recorre a 100 partículas o rectângulo está mais centrado com o objecto.

Através do modelo da dinâmica para atribuição dos pesos das partículas é possível estimar iterativamente a posição do objecto com base nas suas características e, independentemente do tipo de objecto que se deseja seguir, é necessário definir o modelo que o descreve. O filtro de partículas proposto baseia-se na aparência de cor dos objectos, de modo a tornar o sistema mais robusto a variações de forma e movimento do objecto. Para tal, é utilizada a comparação de histogramas de cor para a atribuição dos pesos das partículas, recorrendo à métrica de distância de *Bhattacharyya* no espaço de cores *RGB*, com um total de 25 *bins* por canal. Esta atribuição é dada pela comparação entre o histograma de cada partícula e o histograma de referência do objecto, e pelo número de píxeis do primeiro plano na região da partícula,

$$w_i^{[n]} \leftarrow \tau \cdot d_{battch}(hp_i, h_r^{[n]}) + (1 - \tau)(1 - \vartheta), \quad (3.30)$$

onde $w_i^{[n]}$ representa o peso de cada partícula para o objecto n e ϑ o número de píxeis em movimento na região de cada partícula, proveniente da máscara de movimento considerada nesta dissertação, \mathbf{M}_{PF} (ver em baixo). O parâmetro d_{battch} representa a métrica de distância de *Bhattacharyya*, com $\sum_i h(i) = 1$. Os histogramas de cada partícula e de referência do objecto n são representados por hp_i e $h_r^{[n]}$, respectivamente. Nesta dissertação, a constante τ foi definida como 0.98 para dar uma maior ponderação à aparência dos objectos.

O histograma de cores é realizado com base numa máscara com regiões de píxeis em movimento. Nesta dissertação, é proposto uma máscara de movimento (ver figura 3.28) para a criação e actualização do filtro de partículas, \mathbf{M}_{PF} , dada pela multiplicação das máscaras de subtracção do plano de fundo e de remoção de sombras. Com a inclusão desta máscara, pretende-se que os métodos de detecção de movimento definam equitativamente a aparência do objecto.

Para a comparação de histogramas é necessário ter em consideração o número de píxeis brancos na máscara \mathbf{M}_{PF} , isto é, se esta tem informação suficiente para discriminar a aparência do objecto a ser seguido. Assim, se o número de píxeis do primeiro plano na região do objecto for superior a ρ na máscara \mathbf{M}_{PF} , o histograma de cores é actualizado com base em \mathbf{M}_{PF} . Caso contrário, é verificado se a máscara de subtracção do plano de fundo cumpre a referida condição, para que o modelo do filtro seja actualizado com base na máscara \mathbf{I}_{BS} . Caso nenhuma das máscaras apresente movimento na região do objecto, este é considerado como ocluso (ver Secção 3.3.3.2).

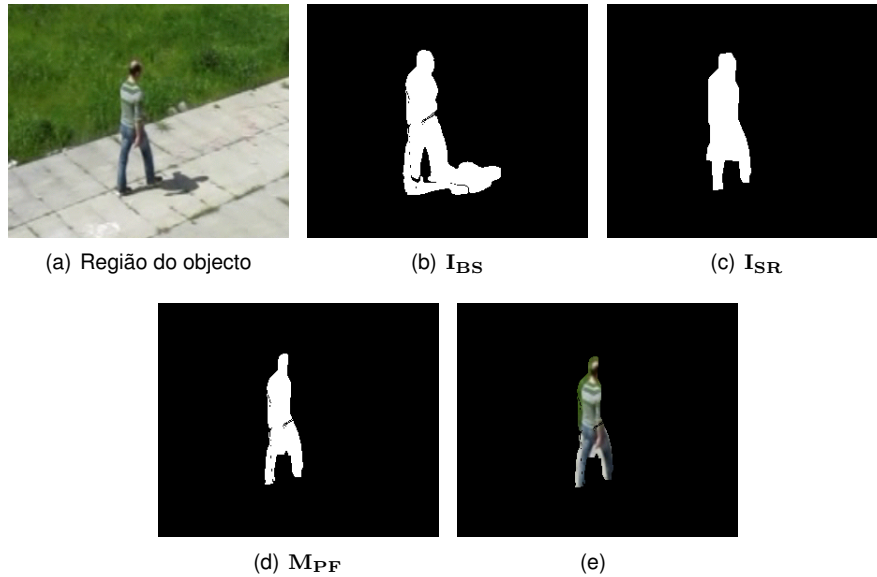


Figura 3.28: Máscaras utilizadas para criação e actualização do filtro de partículas proposto. (a) Região do objecto detectado; máscaras resultantes dos métodos de subtracção do plano de fundo (b) e remoção de sombras (c) para a região do objecto; (d) máscara M_{PF} , dada pela multiplicação de I_{SR} e I_{BS} ; (e) região do objecto filtrada pela máscara M_{PF} .

A cada iteração do filtro, é possível determinar a posição actual do objecto em relação à imagem anterior, recorrendo à média das partículas com o peso mais elevado. Com base na referida média, é iterativamente actualizado o histograma de referência do objecto em movimento, $h_r^{[n]}$, com base no histograma de aparência da média das melhores partículas, $h_{mp}^{[n]}$:

$$h_r^{[n]} \leftarrow \varphi \cdot h_r^{[n]} + (1 - \varphi) \cdot h_{mp}^{[n]}, \quad (3.31)$$

onde φ é uma constante definida empiricamente. De modo a adaptar o sistema a variações dinâmicas da aparência do objecto em movimento, φ é definido como 0,3.

O sistema proposto tem a capacidade de seguir múltiplos objectos em movimento. Para tal, a cada iteração do algoritmo é verificado se as caixas delimitadoras dos objectos a serem seguidos, dadas pela média da partículas com melhor pontuação, se encontram em intersecção com regiões interditas ou com fogo (ver Secção 3.5). Caso não estejam em intersecção, são actualizados os parâmetros de identificação dos objectos e a posição actual do mesmo (ver figura 3.29).

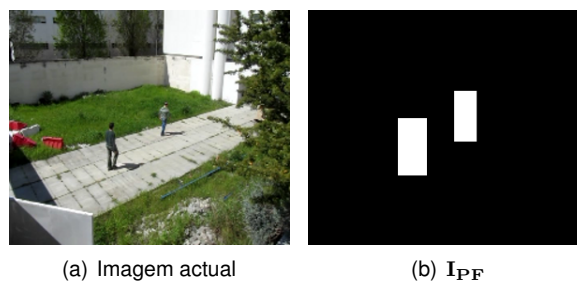


Figura 3.29: Localização dos objectos a serem seguidos pelo sistema. (a) Imagem actual; (b) máscara com as caixas delimitadoras dos objectos a serem seguidos pelo sistema, I_{PF} .

A figura 3.30 apresenta exemplos do resultado do algoritmo de detecção e seguimento de objectos proposto nesta dissertação, onde apenas as caixas delimitadoras dos objectos a serem seguidos não se encontram esbatidas.

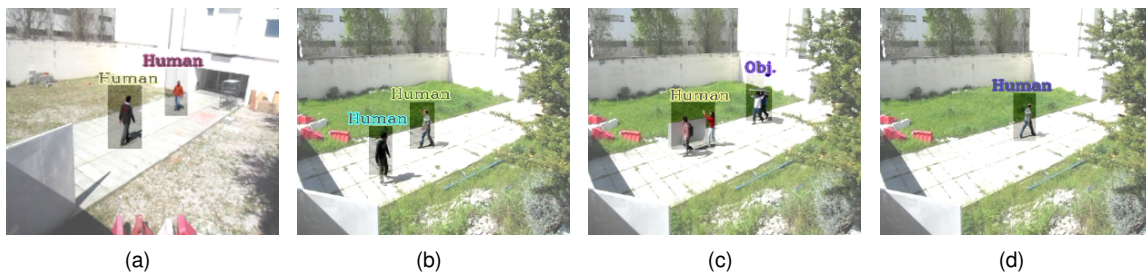


Figura 3.30: Resultado do algoritmo de detecção e seguimento de objectos proposto. As marcas *Human* e *Obj.* representam o tipo de objecto que o sistema está a seguir (ver Secção 3.3.3.3).

3.3.3.2 Estado do seguimento

Um dos maiores desafios dos métodos de seguimento de objectos é a capacidade de lidar com oclusões entre múltiplos objectos. Estas oclusões podem ser dinâmicas ou estáticas, isto é, entre múltiplos objectos que se encontram a ser seguidos ou entre um objecto em movimento e objectos fixos no ambiente (e.g árvore e armário).

Para mitigar o problema relacionado com as oclusões dinâmicas, após cada iteração do filtro de partículas a posição das caixas delimitadoras dos objectos a serem seguidos é verificada. O sistema considera que dois objectos a serem seguidos estão em intersecção se as suas caixas delimitadoras se intersectam ou estão relativamente próximas (ver Secção 3.2.2.1). Nestas situações, o objecto a ser seguido é considerado como estando em intersecção e é dinamicamente alterada a parametrização do filtro: (1) a função de pontuação do peso das partículas baseia-se apenas na comparação de histogramas de cor; (2) a actualização iterativa do histograma não é executada; (3) os parâmetros das 4 dimensões do ruído são alterados para espalhar as partículas de forma mais abrangente em torno dos objectos (ver Secção 3.3.3.1).

No entanto, caso os objectos em movimento se mantenham em intersecção mais que 150 imagens consecutivas, o filtro de partículas do objecto em intersecção que foi criado há menos tempo é eliminado do sistema, de modo a evitar a sobreposição de objectos a serem seguidos. De referir que a contabilização do número de intersecções entre objectos seguidos por filtros de partículas só é considerada quando a distância entre as caixas delimitadoras dos objectos for inferior ao limiar int_{pf} (parametrizado para 4).

Na figura 3.31 estão expostas três situações de intersecção entre objectos em movimento. A primeira e terceira linha desta figura ilustram duas situações típicas onde o sistema segue os objectos correctamente, mesmo quando estes se movem em direcções opostas e se intersectam durante um período de tempo suficientemente longo para as suas aparências mudarem. É importante notar que a aparência das pessoas é semelhante, visto que ambas estão vestidas de verde e, para além disso, o ambiente é também predominantemente verde. Através da cor identificadora de cada objecto, é possível observar que os objectos foram correctamente distinguidos após a oclusão.



Figura 3.31: Intersecção de objectos em movimento a serem seguidos pelo sistema em três situações (uma por linha). (a) Situação inicial do seguimento; (b) após algumas imagens, os objectos encontram-se em intersecção e durante esta fase, o algoritmo proposto utiliza apenas a comparação de histogramas de cores para distinguir os objectos; (c) após a intersecção, os objectos apresentam a mesma cor identificadora, significando que o sistema distinguiu correctamente os objectos.

Relativamente às oclusões entre objectos em movimento seguidos pelo sistema e objectos estáticos presentes no ambiente, existem duas aproximações para solucionar este problema:

- Aumentar o ruído aleatório das partículas em torno da última posição do objecto ocluído, seguindo-o novamente quando este abandonar o estado de oclusão. No entanto, com este processo as partículas tendem a convergir para qualquer objecto que esteja próximo dessa região, mesmo que o peso de cada partícula seja baixo.
- Categorizar o objecto como ocluído e não executar nenhum processo de seguimento ao mesmo. Caso seja detectado um novo objecto, verificar se este cumpre um conjunto de condições para que o sistema possa considerar que o novo objecto detectado é o objecto que se encontrava ocluído.

Nesta dissertação, as oclusões com objectos estáticos são considerados através da segunda aproximação (ver figura 3.32). Neste tipo de oclusões, não existem regiões de píxeis em movimento na região do objecto a ser seguido. Visto que a atribuição dos pesos das partículas depende do número de píxeis em movimento, esta tem de ser dinamicamente alterada durante este tipo de oclusões, ou seja, a constante τ é definida como 1 (ver equação 3.30). Com base nesta aproximação, o sistema proposto considera que o objecto detectado e o ocluído são o mesmo caso estes validem as seguintes condições:

1. Os objectos apresentam uma aparência semelhante, isto é, uma comparação de histogramas de cores inferior a ω (parametrizado para 0.8).
2. As posições das caixas delimitadoras dos objectos estão a uma distância Euclidiana máxima (definido como 80). Esta condição garante que os objectos não se encontram em regiões da imagem demasiado afastadas.
3. O objecto seja detectado pela primeira vez após a oclusão, garantindo assim que a atribuição do objecto ocluído não é feita a um objecto que apareceu no campo de visão da câmara antes da oclusão.

Se as condições indicadas anteriormente forem validadas, todos os parâmetros de identificação do objecto ocluído são adicionados ao novo objecto e é eliminado o objecto ocluído. Poderão existir circunstâncias nas quais este método falhe, isto é, considerar que o objecto detectado é o objecto ocluído, quando na realidade não é. Embora exista uma má atribuição dos objectos, quando o objecto sair do estado de oclusão é novamente detectado e seguido pelo sistema. De referir que, caso não exista nenhuma correspondência entre objectos após uma sequência de 300 imagens, o filtro de partículas do objecto ocluído é eliminado do sistema.

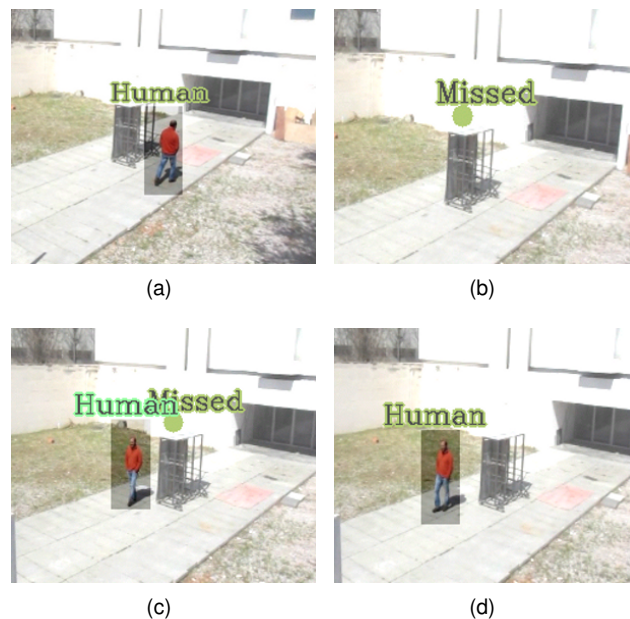


Figura 3.32: Oclusão entre um objecto em movimento e um objecto fixo no ambiente. (a) Situação inicial do seguimento. No local onde o objecto ficou ocluído por um obstáculo, é marcado um aviso de *Missed* (b). Após uma sequência de imagens, um objecto é detectado e seguido pelo sistema (c). O objecto a ser seguido cumpre as condições acima indicadas, logo este é considerado o objecto ocluído (d).

Para além de lidar explicitamente com oclusões entre objectos, o sistema proposto também detecta se um determinado objecto vai abandonar o campo de visão da câmara. De referir que este processo está subjacente à calibração do sistema (ver Secção 3.4), visto que é necessário definir as regiões pelas quais os objectos poderão sair do ambiente. Para tal, a cada iteração do algoritmo é verificado se as caixas delimitadoras dos objectos a serem seguidos intersectam as referidas regiões. Caso existe intersecção, é marcado um aviso na imagem (ver figura 3.33) e o objecto não é processado pelo sistema a partir desse instante.

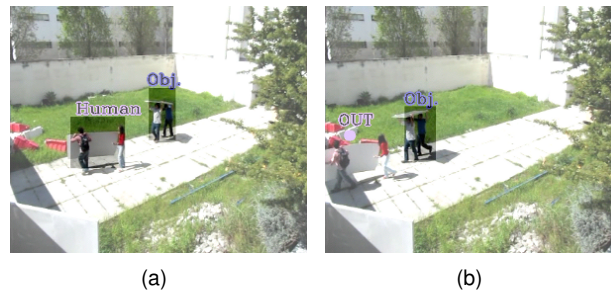


Figura 3.33: Objectos em movimento a abandonar o campo de visão da câmara. (a) Situação inicial do seguimento; (b) após uma sequência de imagens e quando um objecto abandona o campo de visão da câmara, é marcado um aviso de *out* na imagem.

3.3.3.3 Classificação dos Tipos de Objectos

O reconhecimento de padrões humanos numa sequência de imagens é um enorme desafio para algoritmos de visão computacional, devido à constante variação de aparência e forma dos objectos. Uma das tarefas de maior dificuldade é encontrar características robustas para discriminar de forma clara um ser humano.

Dadal and Triggs (2005) propôs um método de detecção de humanos, tendo por base a ideia que a aparência de um objecto é caracterizada pela distribuição da intensidade de gradientes e pela direcção dos contornos. Este utiliza blocos de 16×16 píxeis e histogramas de gradientes orientados para executar a detecção de humanos, recorrendo posteriormente a uma *support vector machine* para a classificação dos objectos.

Com base neste método, os objectos a serem seguidos pelo sistema são classificados como humano ou como objecto genérico (ver figura 3.34).



Figura 3.34: Classificação do tipo de objecto a ser seguido pelo sistema: humano (a) e objecto genérico (b).

Apesar de apresentar bons resultados, o método proposto por Dadal and Triggs (2005) é pesado computacionalmente. Com o objectivo de minimizar esse efeito, o método é apenas executado para cada objecto a ser seguido pelo sistema quando: (1) não tiver sido considerado humano em iterações anteriores do algoritmo; (2) não estiver em oclusão com nenhum objecto em movimento ou objecto fixo no ambiente e (3) não estiver em intersecção com regiões com chamadas. Para além disso, o método só é executado após uma sequência de 20 imagens.

Quando um determinado objecto a ser seguido se move pelo ambiente, este poderá ficar ocluído por regiões com chamadas, por objectos em movimento ou fixos no ambiente. Durante a ocorrência dessas oclusões, não é possível definir com clareza o tipo de objecto a ser seguido, logo este é reclassificado como objecto genérico.

O algoritmo 6 sumariza o algoritmo de detecção e seguimento de objectos proposto nesta dissertação.

Algoritmo 6 seguimento

```

1: Entrada: conjunto de objectos a serem seguidos ( $N_{seg}$ ), máscara de subtração do plano de
   fundo ( $I_{BS}$ ), máscara de remoção de sombras ( $I_{SR}$ ), máscara de calibração ( $I_{calib}$ )
2: Saída: conjunto actualizado de objectos a serem seguidos ( $N_{seg}$ )
3: Data: caixa delimitadora ( $rect$ ), número de píxeis brancos ( $wp$ ),  $\rho$  é uma constante definida
   empiricamente.
4:
5: foreach  $n \in N_{seg}$  do
6:   if  $ocluo(n) = falso$  e  $out(n) = falso$  then
7:     //verificar se objecto está em intersecção com regiões laterais (ver Secção 3.4)
8:     if  $rect(n)$  intersecta  $I_{calib}$  then { //usar equação 3.21 com  $\lambda = 1$  }
9:       objecto abandonou o campo de visão da câmara,  $out(n) \leftarrow verdadeiro$ 
10:    else
11:      //criar máscara de actualização do filtro de partículas,  $M_{PF}$  (ver figura 3.28)
12:      //verificar se  $I_{SR}$  ou  $I_{BS}$  têm movimento na região do objecto para criar  $M_{PF}$ 
13:      if  $wp(I_{SR}) \geq \rho$  then
14:        criar máscara de movimento  $M_{PF}$ , baseado em  $I_{SR}$  e  $I_{BS}$ 
15:        iterar filtro de partículas (ver Secção 3.3.3.1)
16:      else if  $wp(I_{BS}) \geq \rho$  then
17:        criar máscara de movimento  $M_{PF}$ , baseado em  $I_{BS}$ 
18:        iterar filtro de partículas
19:      else { //caso não exista movimento na região do objecto, objecto é considerado ocluso }
20:        objecto entrou no estado de oclusão,  $ocluo(n) \leftarrow verdadeiro$ ;
21:      end if
22:    end if
23:  else if  $ocluo(n) = verdadeiro$  then
24:    marcar aviso que objecto está ocluso e executar método de correspondência de objectos
25:  else
26:    marcar aviso que objecto abandonou o campo de visão da câmara
27:  end if
28: end for

```

3.4 Calibração do Sistema

O sistema proposto nesta dissertação gera alarmes geo-localizados consoante o tipo de evento a ocorrer no campo de visão da câmara. Para além disso, através de uma calibração é possível definir regiões onde não se pretende detectar objectos. Esta secção descreve o processo de adicionar a localização *GPS* aos alarmes do sistema e a calibração da detecção de objectos.

Existem algumas possibilidades para obter um mapa com as coordenadas *GPS* para todas as posições da imagem. Uma das possibilidades seria conhecer previamente as coordenadas *GPS* da câmara e através de um conjunto de suposições (e.g terreno planar) sobre o ambiente, executar um processo de calibração. Dado que nesta dissertação se pretende um sistema que modele qualquer tipo de ambiente, optou-se por efectuar um mapeamento entre as posições no plano da imagem e as coordenadas *GPS*. Para tal, o operador do sistema tem de percorrer o ambiente do campo visual da câmara e com auxílio de um dispositivo reportar as coordenadas *GPS*. Enquanto este se movimenta, esta coordenada é armazenada juntamente com a caixa delimitadora associada ao operador. Com base nos dados obtidos, uma correspondência entre as coordenadas é criada (ver tabela 3.2).

Tabela 3.2: Correspondência entre as coordenadas *GPS* e as coordenadas do objecto no plano da imagem

Objecto		GPS	
x	y	latitude	longitude
50	75	38.393589	9.121753
120	170	38.393610	9.121745
250	98	38.393628	9.121761
...

Esta calibração permite criar uma base de conhecimento, e através desta e da posição na imagem de um determinado evento é possível reportar alarmes geo-localizados. Este processo é efectuado com base nos três vizinhos mais próximos da posição do evento, isto é, para cada posição (x, y) da imagem, a sua coordenada *GPS* é dada pela média das três coordenadas *GPS* que se encontram a uma menor distância Euclidiana dessa posição.

Para além de serem utilizados para reportar alarmes geo-localizados, os dados provenientes da calibração *GPS* são interpolados de modo a obter uma calibração da altura esperada dos objectos para cada posição da imagem. O resultado deste processo, H_{calib} , é dado pela média das alturas das três coordenadas *GPS* que se encontram a uma menor distância Euclidiana de uma determinada posição. O objectivo desta interpolação é tornar o sistema mais robusto e substancialmente mais eficiente na detecção de objectos, visto que estes só são processados se a sua altura estiver dentro do intervalo esperado.

A figura 3.35 apresenta um exemplo típico de calibração do sistema num ambiente exterior.

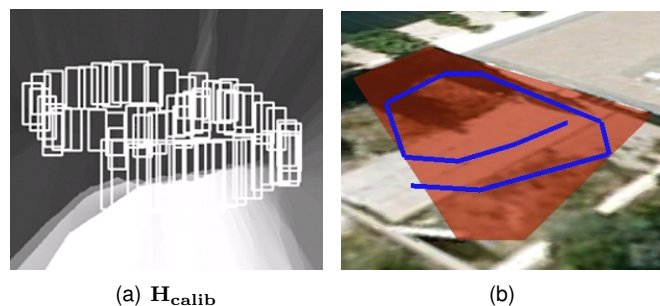


Figura 3.35: (a) Máscara de interpolação da altura esperada de um objecto para cada posição da imagem, H_{calib} , representada através do nível de brilho (valores escalados para melhor legibilidade). Os rectângulos representam as caixas delimitadoras do operador durante a fase de calibração; (b) conjunto de dados *GPS* sobrepostos à imagem de satélite. A região a vermelho representa o campo de visão da câmara da figura 3.36(a).

Para além da calibração *GPS*, são definidos os polígonos onde não se pretende detectar objectos em movimento (rectângulo central da figura 3.36(a)) e os polígonos de calibração laterais. Os primeiros são definidos quando não se pretende detectar regiões de movimento numa determinada região da imagem. Os polígonos laterais são necessários para distinguir objectos quando estes se intersectam nas regiões laterais do plano da imagem.



Figura 3.36: Definição dos polígonos de calibração do sistema. (a) Polígonos de calibração lateral e polígono onde não se pretende detectar objectos (rectângulo central), sobrepostos à imagem original; (b) o polígono de calibração lateral direito encontra-se sobreposto aos arbustos, no sentido de evitar que o movimento repetitivo das folhas afecte a detecção de objectos.

3.5 Integração dos Algoritmos e Geração de Alarmes

Esta secção descreve o processo de integração dos resultados dos algoritmos propostos e a geração de alarmes consoante o tipo de evento a ocorrer no campo de visão da câmara.

3.5.1 Integração dos Algoritmos

O algoritmo de detecção de fogo classifica uma região como fogo consoante o resultado da conjugação dos métodos implementados: movimento e informação da cor dos píxeis, detecção de texturas dinâmicas e análise no domínio da frequência. Apesar do algoritmo apresentar resultados robustos (ver Capítulo 4), poderão existir situações nas quais objectos com movimentos repetitivos e com cores semelhantes às das chamas gerem um falso alarme na detecção de fogo. Para contornar este problema, recorrente neste género de algoritmos, é proposto uma integração inovadora dos resultados da detecção de fogo e seguimento de objectos.

Com esta integração pretende-se que pelo facto de objectos estarem a ser seguidos, nenhum falso alarme na detecção de fogo é gerado. O sistema proposto classifica uma determinada região como fogo, caso esta valide as seguintes regras de integração:

1. Caso as caixas delimitadoras da região de fogo e dos objectos seguidos pelo sistema não estejam em intersecção. Este processo é baseado em I_f e I_{PF} (ver figura 3.37).
2. Caso exista intersecção entre a região de fogo e um objecto a ser seguido, deduz-se que: (1) a região de fogo apresentou movimento suficiente para ser detectada e seguida pelo sistema ou (2) um objecto em movimento entrou em intersecção com a região de chamas. Neste sentido, é analisado o movimento do objecto e uma região é considerada fogo se: (1) o movimento do objecto desde que foi detectado pela primeira vez até à posição actual, em termos de distância Euclidiana, for inferior ao limiar thr_{mf} (parametrizado para 50); (2) o objecto em movimento não tiver sido detectado pela primeira vez antes da região de fogo. O parâmetro thr_{mf} representa a distância Euclidiana mínima que um objecto a ser seguido tem de percorrer para se considerar que não se encontra estático no ambiente.

Caso as condições acima indicadas sejam validadas, o sistema proposto reporta um alarme de fogo. No entanto, poderão existir filtros de partículas na região classificada como fogo, sendo eliminados após 50 imagens consecutivas.

Na figura 3.37 estão expostos exemplos do resultado do sistema proposto para detecção de fogo e detecção e seguimento de objectos, aplicado num ambiente exterior com regiões de fogo e pessoas em movimento.

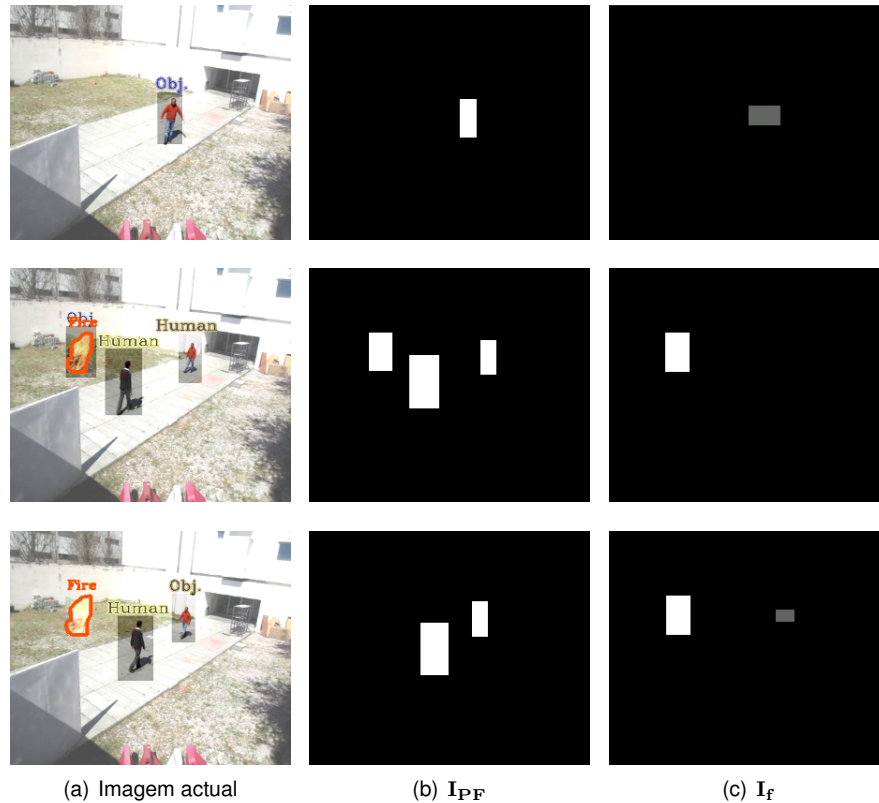


Figura 3.37: (a) Resultado do sistema proposto para detecção de fogo, detecção e seguimento de objectos com base numa sequência de imagens; (b) máscara com a localização dos objectos a serem seguidos, I_{PF} ; (c) máscara com a localização das regiões de chamas detectadas, I_f , marcadas a branco se foram classificadas como fogo ou em tons de cinza, caso os métodos de análise no domínio da frequência estejam a ser aplicados.

Na primeira linha da figura 3.37 está representada uma situação onde um objecto em movimento, com cores semelhantes às das chamas, simula movimentos corporais repetitivos no sentido de induzir um falso alarme na detecção de fogo. Visto que o objecto em movimento está a ser analisado pelo algoritmo de detecção de fogo, a região do objecto encontra-se marcada a cinza na máscara I_f . De notar que apesar de o objecto ter sido considerado humano, como este se encontra em intersecção com uma região que poderá conter chamas, este é reclassificado como objecto genérico (ver secção 3.3.3.3).

Na segunda linha da figura 3.37 verifica-se que o objecto referido anteriormente foi classificado como humano e, visto que não apresenta uma assinatura no domínio da frequência semelhante ao fogo, não foi classificado como fogo. Para além disso, caso o algoritmo de detecção de fogo reporte um falso alarme, o sistema proposto não classifica a região como fogo, pois neste caso não valida as regras de integração.

A região que contém um objecto em movimento com cores semelhantes às das chamas é no-

vamente analisada para verificar a existência de fogo (ver terceira linha da figura 3.37). É também possível observar que na região do objecto classificado como humano não é executada uma análise por parte do algoritmo de detecção de fogo, e o filtro de partículas em intersecção com a região classificada como fogo foi eliminado.

3.5.2 Geração de Alarmes

Com base no resultado do seguimento de objectos e da detecção de focos de incêndios, o sistema proposto reporta alarmes geo-localizados em função do tipo de evento a ocorrer no plano da imagem. Os alarmes são classificados consoante o tipo de evento:

- Movimento - relacionado com a detecção de objectos em movimento através do algoritmo de detecção e seguimento de objectos. Este é reportado quando um determinado objecto detectado: (1) não se encontra em intersecção com polígonos de calibração ou com regiões de chamas; (2) na posição actual do objecto, existem regiões de píxeis em movimento extraídas através do método de estimação do plano de fundo (ver Secção 3.1) .
- Seguimento - relacionado com o seguimento de objectos. Este alarme é reportado quando o objecto a ser seguido não se encontra em intersecção com polígonos de calibração lateral e não esteja ocluído por objectos fixos no ambiente.
- Fogo - após a integração dos resultados dos algoritmos de detecção de fogo e seguimento de objectos, é reportado um alarme de fogo caso a região tenha sido confirmada como fogo.

Um dos objectivos do sistema proposto é assistir o operador que controla e monitoriza uma determinada área, da forma mais eficiente e robusta possível. Para tal, é executada uma comunicação entre o sistema e o operador através da arquitectura cliente-servidor, onde o sistema funciona como prestador de serviços, reportando alarmes no formato *Extensible Markup Language (XML)* com os parâmetros: (1) número identificador do objecto; (2) data e hora de ocorrência do evento; (3) coordenadas *GPS* do evento; (4) tipo de alarme. No caso dos alarmes de seguimento, é adicionado o parâmetro de descrição do tipo de objecto (humano ou objecto genérico).

Para além dos referidos alarmes, o operador tem acesso às imagens com os resultados do sistema. Para tornar a sua consulta mais eficiente, é proposto nesta dissertação uma imagem com o caminho percorrido pelos objectos no campo de visão da câmara (ver figura 3.38). Esta é iterativamente actualizada com base num vector que guarda as localizações dos objectos a serem seguidos:

- Caso um determinado objecto não tenha sido classificado como humano, o seu caminho percorrido é marcado através de uma circunferência sem preenchimento.
- Assim que um objecto é classificado como humano, é marcada uma circunferência de maior dimensão e, a partir desse instante, as circunferências com o caminho percorrido pelo objecto passam a conter preenchimento.
- As localizações do objecto são marcadas de acordo com a sua cor identificadora.

- Quando um objecto entra em oclusão com um obstáculo, não é marcada nenhuma informação. No entanto, quando é efectuada uma correspondência entre um objecto detectado e o objecto ocluído, é marcado um M (de *missed*) e um rectângulo (proporcional ao tamanho do objecto) na região da correspondência. Na figura 3.38(f) é possível observar que após o objecto ficar ocluído pelo armário, visível através do fim da trajectória do objecto no lado direito da imagem, foi criado um novo objecto (a verde), e com base neste que foi executada a correspondência entre os objectos.

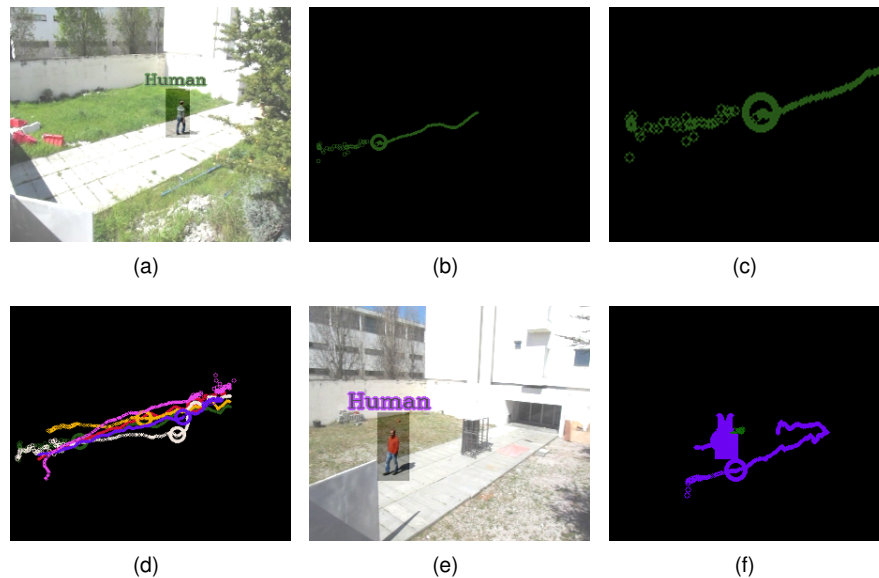


Figura 3.38: Caminho percorrido por objectos seguidos pelo sistema no campo de visão da câmara. (a) Estado inicial do seguimento; (b) caminho percorrido pelo objecto no ambiente da figura (a); (c) aproximação da figura (b) para uma análise detalhada às circunferências com e sem preenchimento, e à circunferência de maior dimensão, identificando o objecto como humano a partir dessa localização; (d) caminho percorrido por vários objectos no ambiente da figura (a); (e) objecto a ser seguido pelo sistema após a oclusão; (f) caminhos percorridos pelo objecto antes de ficar ocluído (a roxo) e pelo novo objecto detectado (a verde), e o rectângulo que representa a correspondência entre estes.

Capítulo 4

Resultados Experimentais

Este capítulo apresenta a instalação experimental e os resultados experimentais obtidos com o sistema proposto. Para uma validação mais eficiente e uma análise mais detalhada dos resultados obtidos, este capítulo encontra-se dividido em três secções: a secção 4.1 expõe a instalação experimental, e as secções 4.2 e 4.3 apresentam os resultados experimentais do algoritmo de detecção de fogo e algoritmo de detecção e seguimento de objectos, respectivamente.

4.1 Instalação Experimental

O sistema proposto nesta dissertação foi implementado na linguagem de programação C++ e testado num processador Intel Pentium(R) Dual-Core P7800 a 2.53GHz com 4Gb de memória RAM, correndo o sistema operativo Linux Ubuntu 10.10 (*Maverick Meerkat*).

Para a implementação do sistema proposto foi necessário recorrer a três bibliotecas: (1) biblioteca *OpenCV 2.3*¹ para a execução de algoritmos de visão computacional de baixo nível (Bradski and Kaehler, 2008); (2) biblioteca *wavelet1d*² para execução da transformada *wavelet* discreta; (3) biblioteca *Xerces-C++*³ para análise e validação de documentos *XML*.

De um modo geral, a validação experimental de um algoritmo de detecção de fogo é uma tarefa de elevada dificuldade, devido essencialmente a dois factores: (1) insuficiente número de bancos de dados disponíveis e (2) a dificuldade de gerar um fogo controlado para a ser gravado. Neste sentido, a validação da eficiência do algoritmo de detecção de fogo proposto foi executada com base num conjunto de 12 vídeos encontrados em sítios da Internet. Os vídeos englobam um total de 21992 imagens com uma resolução de 300×250 , gravados a uma taxa de 30 imagens por segundo e com a câmara em movimento.

Relativamente ao algoritmo de detecção e seguimento de objectos, um banco de dados com 3

¹*OpenCV 2.3*: <http://opencv.willowgarage.com/wiki/>

²*wavelet1d*: <http://code.google.com/p/wavelet1d>

³*Xerces-C++*: <http://xerces.apache.org/xerces-c>

vídeos a cores, abrangendo um total de 17247 imagens com uma resolução de 640×480 gravados a uma taxa de 30 imagens por segundo, foram obtidos com recurso a uma câmara fixa no ambiente.

Para além do banco de dados referido anteriormente, para validar experimentalmente a integração dos algoritmos de detecção de fogo e seguimento de pessoas são necessários vídeos que incluam simultaneamente pessoas em movimento e regiões com chamas. Visto não existir nenhum banco de dados disponível com as referidas especificações e como gerar um fogo representa um grande obstáculo, não foram obtidos resultados da integração dos algoritmos.

4.2 Resultados da Detecção de Fogo

Para a validação do algoritmo de detecção de fogo foi aplicado um conjunto de vídeos gravados em diferentes ambientes, com diversas condições de luminosidade e com situações típicas quotidianas de pessoas e objectos em movimento com aparência semelhante ao fogo. A figura 4.1 ilustra exemplos de imagens do bancos de dados, com a saída do algoritmo sobreposta.

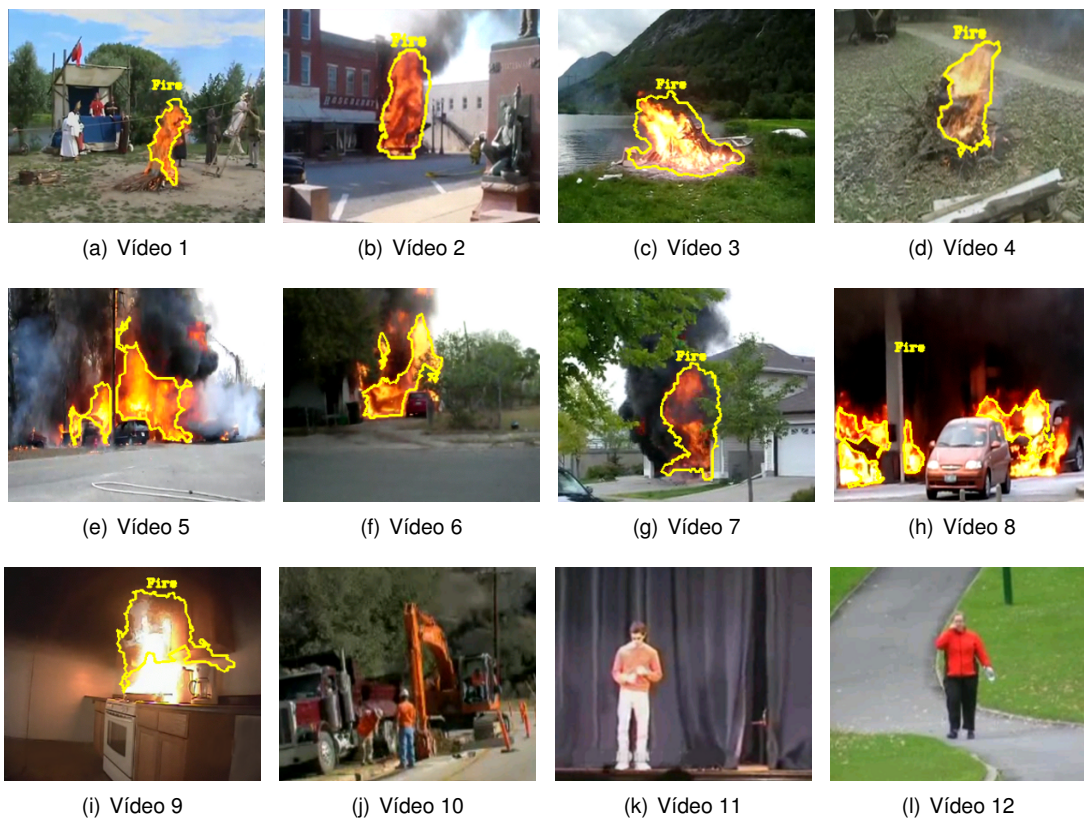


Figura 4.1: Exemplos de imagens do bancos de dados utilizado para validar o algoritmo de detecção de fogo, com o resultado deste representado pelos contornos amarelos. Os vídeos 10, 11 e 12 incluem situações com elementos dissuasores na detecção de fogo, isto é, objectos em movimento com aparência semelhante ao fogo.

O resultado do algoritmo proposto é representado pelos contornos das chamas. Analisando a figura 4.1, verifica-se que o algoritmo detecta as regiões de fogo quando estas estão presentes na imagem, apresentando uma segmentação das chamas superior a 90% do seu tamanho. É importante verificar que no banco de dados estão incluídas situações com chamas de pequena dimensão (ver vídeo 8) e com reflexos das mesmas no ambiente (ver vídeo 9). Para além disso, como mostram

as imagens dos vídeos 10, 11 e 12, o algoritmo não detecta regiões de fogo que poderão gerar falsos alarmes, tais como pessoas em movimento ou máquinas em manobras com cores semelhantes às das chamas. De notar ainda que os resultados obtidos não incluem a integração com o algoritmo de detecção e seguimento de objectos.

A tabela 4.1 enaltece o esforço efectuado nesta dissertação para a componente de detecção de fogo apresentar resultados robustos e fiáveis, sob quaisquer condições e ambientes. A taxa de sucesso reportada representa a percentagem de êxito do algoritmo, isto é, a proporção de imagens do vídeo onde as regiões de fogo, quando presentes, são correctamente detectadas.

Tabela 4.1: Resultados experimentais da detecção de fogo

Vídeos	Número de imagens analisadas	Número de imagens com chamas	Sucesso (%)
1	3123	2786	89.2
2	3904	3830	98.1
3	812	718	88.4
4	1423	1347	94.7
5	2422	2271	93.8
6	848	698	82.3
7	1073	922	85.9
8	3410	3212	94.2
9	2950	2536	86.0
10	265	0	100.0
11	1438	0	100.0
12	324	0	100.0

Globalmente, o algoritmo proposto apresenta uma detecção média de regiões de fogo de 92.7%. Para além disso, a taxa de 100% de sucesso, em 3 dos vídeos onde estão presentes objectos em movimento com aparência semelhante ao fogo, mostra claramente a robustez e fiabilidade do algoritmo.

Por fim, é importante mencionar a métrica utilizada para obtenção dos resultados experimentais na detecção de fogo. Visto que a forma das chamas varia iterativamente, a avaliação da presença ou ausência de fogo é um enorme desafio. Para tal, assume-se uma detecção positiva de uma região de fogo caso as chamas nessa região tenham sido segmentadas em pelo menos 90% do seu tamanho (ver figura 4.2(a)).



Figura 4.2: Exemplos de classificação e respectivas avaliações em termos da presença e ausência de fogo, com uma detecção positiva da região de fogo com praticamente 100% de segmentação das chamas (a) e uma detecção incorrecta da região de fogo (b).

4.3 Resultados da Detecção e Seguimento de Objectos

Para validar o algoritmo de detecção e seguimento de múltiplos objectos, foram utilizados 3 vídeos que incluem situações típicas existentes num ambiente exterior. Visto que se pretende um sistema capaz de lidar com oclusões estáticas e dinâmicas entre objectos, o banco de dados incorpora situações com pessoas a movimentarem-se em grupo e a transportar objectos. A figura 4.3 apresenta exemplos de imagens extraídas do banco de dados, com o resultado do algoritmo sobreposto à imagem original. Para facilitar a visualização, apenas as caixas delimitadoras dos objectos seguidos pelo sistema não se encontram esbatidas.

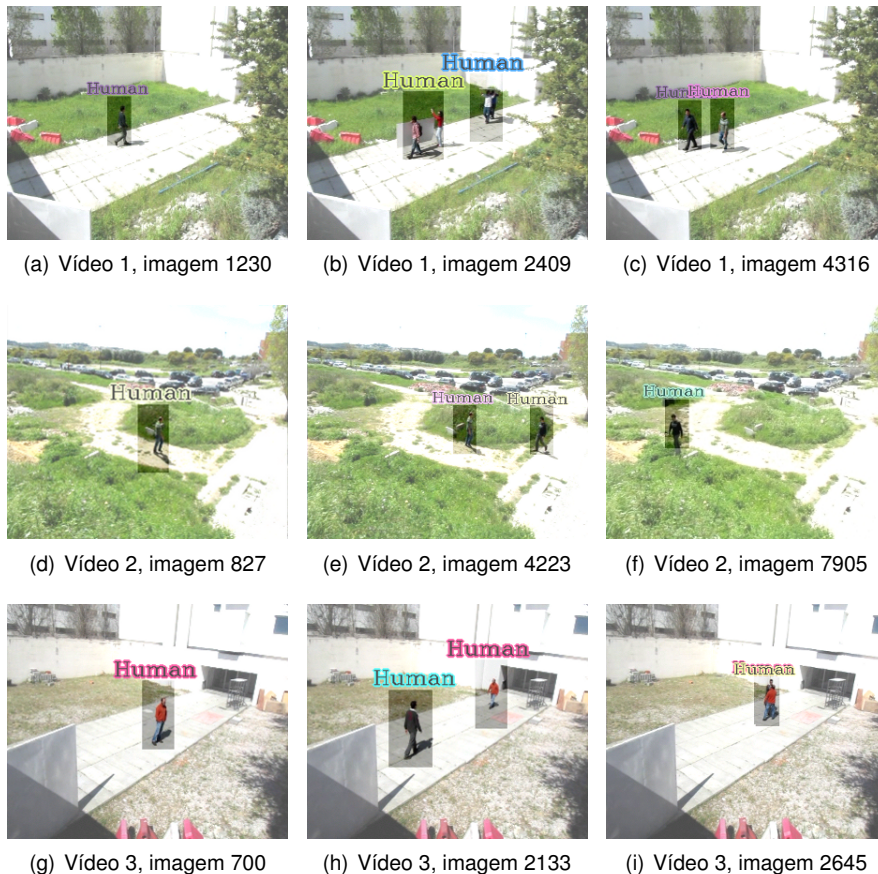


Figura 4.3: Exemplos de imagens dos bancos de dados utilizado para validar o algoritmo de detecção e seguimento de objectos com a saída do algoritmo sobreposta em três situações (uma por linha), onde apenas as caixas delimitadoras dos objectos seguidos pelo sistema não se encontram esbatidas.

Analisando as imagens do banco de dados (ver figura 4.3), verifica-se que o algoritmo proposto segue correctamente as pessoas em movimento, mesmo que estas apresentem uma aparência muito semelhante ao ambiente ou que se encontrem próximas. É interessante observar, através da figura 4.3(b), que as 4 pessoas em movimento são seguidas apenas por dois filtros de partículas, pois estas encontram-se ligadas por um objecto. No entanto, quando se separarem, o sistema vai executar o seguimento das mesmas. De mencionar que não são seguidos pelo sistema quaisquer objectos provenientes do movimento das folhas das árvores (ver vídeo 1), das ervas (ver vídeo 2) ou das sombras das pessoas. Para além das imagens apresentadas, outros exemplos de imagens do banco de dados são ilustrados nas figuras 3.30, 3.31, 3.32, 3.33, 3.37 e 3.38.

A tabela 4.2 expõe os resultados quantitativos do algoritmo de detecção e seguimento de objec-

tos. Para a obtenção destes resultados foram consideradas duas métricas, baseadas na detecção e no seguimento dos objectos. Estas métricas, dadas pela percentagem de objectos detectados e seguidos, representam o número de imagens que os objectos, quando presentes na imagem, foram correctamente detectados e seguidos pelo sistema, respectivamente.

Para além destas métricas, são propostas duas formas de validar a eficiência do algoritmo (ver figura 4.4). Para tal, são contabilizados o número de vezes que os objectos se encontram visíveis e que se encontram em intersecção com polígonos de calibração. A primeira contabilização representa os casos em que os objectos em movimento (pessoas, animais ou automóveis) se encontram no campo visual câmara, a uma distância máxima de 50 metros da mesma e não estão em sobreposição as regiões de calibração. A segunda, descreve o número de vezes que os objectos detectados se encontram em intersecção com as regiões de calibração.



Figura 4.4: Exemplos de categorização para obtenção dos resultados experimentais do algoritmo de detecção e seguimento de objectos, com os polígonos de calibração sobrepostos à imagem original. (a) Pessoa em movimento em intersecção com um polígono de calibração lateral; (b) pessoa em movimento a pelo menos 50 metros de distância da câmara. Nestas situações, para obtenção das métricas de detecção e seguimento não é incrementado o número de vezes que o objecto se encontra visível e, na situação da figura (a), é incrementado o número de intersecções com os polígonos de calibração. De mencionar que os polígonos de calibração não afectam o algoritmo de detecção de fogo.

Tabela 4.2: Resultados experimentais da detecção e seguimento de objectos

Vídeos	Número de imagens	Número de objectos visíveis	Número de objectos polígonos	Objectos detectados (%)	Objectos seguidos (%)
1	5549	3576	1480	96.31	88.98
2	8245	3387	2623	91.73	89.70
3	3453	3668	104	99.56	99.64
Média	5749	3544	1402	95.87	92.77

Analisando os resultados presentes na tabela 4.2, é possível concluir que a forma como os objectos se movimentam no campo de visão da câmara tem uma grande influência para a sua detecção e seguimento. Caso o vídeo seja constituído por vários objectos a entrarem e saírem do campo de visão da câmara, existe um aumento do número de intersecções entre os objectos e os polígonos de calibração laterais, levando a uma diminuição da taxa de detecção de objectos. Esta situação verifica-se nos vídeos 1 e 2. No caso do vídeo 3, onde o número de intersecções entre os objectos em movimento e os polígonos de calibração é claramente menor, o algoritmo proposto apresenta uma taxa de detecção e seguimento de objectos claramente superior.

Apesar de o método de subtração do plano de fundo proposto ser robusto, poderão existir situações nas quais o movimento repetitivo das folhas das árvores afectam a detecção de objectos. Neste sentido, é possível aumentar o tamanho dos polígonos de calibração laterais de modo a não

afectarem essa detecção (ver lado direito da figura 4.4(a)). No entanto, analisando os resultados experimentais do vídeo 1, verifica-se que este apresenta uma pior taxa de detecção e seguimento de objectos em relação ao terceiro vídeo. Este facto está relacionado com a forma como os objectos se movimentam no ambiente.

O vídeo 2 apresenta a menor taxa de detecção de objectos. Este facto deve-se à influência da distância dos objectos em relação à posição da câmara, isto é, a partir de uma determinada distância os métodos de subtração do plano de fundo e de estimação do plano de fundo (ver Secção 3.1) não têm capacidade para segmentar as regiões de movimento (ver figura 4.5).

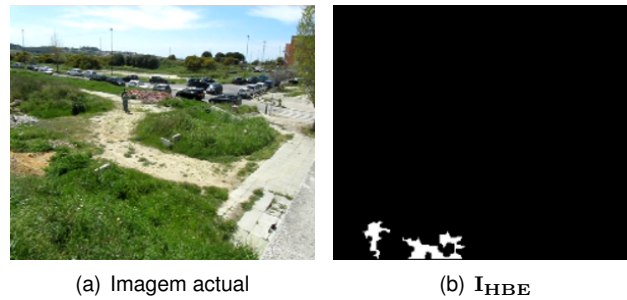


Figura 4.5: (a) Imagem actual; (b) máscara resultante do método de estimação do plano de fundo, quando um objecto em movimento se encontra a mais de 50 metros da câmara.

No que diz respeito à taxa de seguimento de objectos, esta é sempre inferior à taxa de detecção visto que o sistema impõe que os objectos sejam detectados um número pré-determinado de vezes para poderem ser seguidos.

É importante mencionar que quando os objectos em movimento não se encontram em intersecção com regiões de calibração e estão presentes na imagem, estes são correctamente detectados e seguidos pelo sistema. Para além disso, não foram detectados nem seguidos objectos com regiões de movimento provenientes das folhas das árvores, ervas ou as sombras dos objectos. Globalmente, o algoritmo de detecção e seguimento de objectos apresenta uma taxa média de detecção e seguimento de objectos de aproximadamente 96% e 93%, respectivamente.

Capítulo 5

Conclusões e Trabalhos Futuros

Este capítulo apresenta um conjunto de conclusões a partir dos resultados experimentais obtidos e expõe algumas linhas de investigação e desenvolvimentos futuros no tema da dissertação.

5.1 Conclusões

Nesta dissertação desenvolveu-se um sistema para a detecção de fogo e detecção e seguimento de objectos, a partir de sequências de imagens provenientes de câmaras fixas em ambientes exteriores. Com vista a tornar o sistema o mais robusto possível, os resultados dos dois algoritmos são integrados de forma inovadora para evitar falsos alarmes na detecção de fogo. Assim, visto que o fogo se apresenta como um objecto estático no ambiente, o movimento dos objectos seguidos pelo sistema é analisado.

A componente de detecção de fogo segmenta regiões de movimento com cores semelhante às das chamas através de métodos de diferenciação temporal de imagens e de classificação dos píxeis baseado no modelo de cor proposto nesta dissertação, *HY*. Com vista a diferenciar uma região com chamas de um objecto em movimento com aparência semelhante ao fogo, um método de análise na frequência foi incluído. Por ser demasiado pesado computacionalmente, propôs-se focar a sua aplicação às regiões da imagem que exibem texturas dinâmicas. Com base num filtro temporal recursivo, este processo analisa a história dos píxeis considerados fogo.

Com o intuito de seguir múltiplos objectos em movimento, a seguinte sequência de processamento foi proposta. Em primeiro lugar, as regiões de píxeis em movimento são iterativamente segmentadas através da subtracção do plano de fundo. Com vista a tornar o sistema mais robusto a variações de luminosidade, uma grelha regular é sobreposta à imagem actual, associando a cada posição um modelo do plano de fundo. Para além disso, o processo de aprendizagem foi alterado de modo a actualizar cada modelo independentemente das condições existentes nos restantes. Posteriormente, a sombra dos objectos que se destacam do plano de fundo é removida de modo a que esta não afecte o processo de seguimento. No sentido de reduzir o custo computacional, propôs-se remover a sombra da região inferior do objecto. Caso os objectos sejam consecutiva-

mente detectados, estes são seguidos através de filtros de partículas que se baseiam na aparência dos objectos para estimar a sua posição. Para uma estimação computacionalmente mais eficiente, uma actualização do modelo de aparência do objecto através de uma máscara que se baseia na ponderação equitativa dos métodos de detecção de movimento é proposta.

Nesta dissertação propôs-se ainda uma forma original de reportar alarmes (fogo, detecção e seguimento de objectos) geo-localizados com base nos eventos que ocorrem no ambiente. Para tal, as coordenadas do plano da imagem são mapeadas através de um processo de calibração. Neste processo, as coordenadas *GPS* reportadas pelo operador no campo de visão da câmara são associadas à posição da sua caixa delimitadora no plano da imagem e é efectuado uma interpolação dos dados. Para além disso, este processo permite ainda mapear a altura esperada dos objectos para cada posição da imagem. Assim, reduz-se o custo computacional e evita-se criar objectos provenientes do movimento das folhas das árvores, sombras ou alterações da luminosidade.

Para validar o sistema proposto recorreu-se a um conjunto de 15 vídeos com 39239 imagens, que incluem diferentes situações quotidianas num conjunto variado de ambientes. Os resultados obtidos mostram que o sistema proposto apresenta uma taxa de sucesso de 92.7% de detecção de regiões de fogo e uma taxa de sucesso de 92.8% de seguimento de objectos que se encontram no campo de visão da câmara e não estão em intersecção com regiões interditas.

5.2 Trabalhos Futuros

Esta dissertação contribui para o aumento da fiabilidade, eficiência e robustez de sistemas de vídeo vigilância que recorrem a algoritmos de visão computacional para executar a monitorização de uma determinada área, sendo capazes de detectar múltiplos focos de incêndio e seguir objectos. Apesar dos resultados alcançados, existem ainda diversas linhas de trabalho futuro que podem ser prosseguidas. Como extensão ao sistema proposto, destaco entre outros, a integração do sistema numa plataforma móvel, implementando para tal mecanismos de compensação do movimento da câmara. Seria também interessante definir o tipo de objecto que o sistema está a seguir, isto é, encontrar características robustas para discriminar os diferentes objectos presentes no ambiente. Poder-se-á ainda incluir processos de monitorização da actividade humana, os quais se baseiam na estimação do movimento e do comportamento das pessoas no ambiente da câmara. Outra linha de trabalho a seguir no futuro seria a melhoria dos modelos de aparência dos objectos. Esta abordagem tornaria o sistema mais robusto para ambientes nocturnos ou para uma arquitectura suportada na utilização de múltiplas câmaras, onde poderá existir uma grande variação das condições de luminosidade. Para além disso, seria interessante mitigar o problema relacionado com a detecção de objectos em movimento quando estes se encontram a uma distância considerável da câmara.

5.3 Disseminação

O sistema proposto nesta dissertação encontra-se parcialmente descrito na seguinte publicação:

- Santana, P., Gomes, P., Barata, J. A vision-based system for early fire detection. To appears in Proceedings of the International Conference on Systems, Man, and Cybernetics (SMC), 2012.

O sistema foi desenvolvido no contexto do projecto DVA - sistema avançado de vigilância baseado em agentes, em parceria com a empresa Holos e co-financiado pelo Quadro de Referência Estratégico Nacional (QREN).

Bibliografia

- Antonini, M. and Barlaud, M. (1992). Image coding using wavelet transform. *IEEE Transactions on Image Processing*, 1(2):205–220.
- Arulampalam, M., Maskell, S., Gordon, N., and Clapp, T. (2002). A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking. *IEEE Transactions on Signal Processing*, 50(2):174–188.
- Barron, J., Fleet, D., and Beauchemin, S. (1992). Performance of optical flow techniques. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 236–242.
- Berclaz, J., Fleuret, F., and Fua, P. (2006). Robust people tracking with global trajectory optimization. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 17–22.
- Black, J., Makris, D., and Ellis, T. (2005). Validation of blind region learning and tracking. In *Proceedings of the second IEEE International Workshop on Visual Surveillance and Performance Evaluation of Tracking and Surveillance*, pages 9–16.
- Bradski, G. and kaehler, A. (2008). *Learning OpenCV: computer vision with the OpenCV library*. O'Reilly Media.
- Butler, D., Sridharan, S., and Bove, V. (2003). Real-time adaptive background segmentation. In *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, volume 3, pages 349–352.
- Celik, T. (2010). Fast and efficient method for fire detection using image processing. *ETRI Journal*, 32(6).
- Celik, T. and Demirel, H. (2009). Fire detection in video sequences using a generic color model. *Fire Safety Journal*, 44(2):147–158.
- Celik, T., Demirel, H., and Ozkaramanli, H. (2006). Automatic fire detection in video sequences. In *Proceedings of the 14th European Signal Processing Conference (EUSIPCO)*.
- Celik, T., Demirel, H., Ozkaramanli, H., and Uyguroglu, M. (2007). Fire detection using statistical color model in video sequences. *Journal of Visual Communication and Image Representation*, 18(2):176–185.
- Chen, J., He, Y., and Wang, J. (2010). Multi-feature fusion based fast video flame detection. *Building and Environment*, 45(5):1113–1122.

- Chen, T., Kao, C., and Chang, S. (2003). An intelligent real-time fire-detection method based on video processing. In *Proceedings of the IEEE 37th Annual International Carnahan Conference on Security Technology*, pages 104–111.
- Chen, T., Wu, P., and Chiou, Y. (2004). An early fire-detection method based on image processing. In *Proceedings of the IEEE International Conference on Image Processing (ICIP)*, volume 3, pages 1707–1710.
- Chen, W., Lai, C., Hung, P., and Chen, C. (2008). An adaptive learning method for target tracking across multiple cameras. In *Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1–8.
- Chien, S., Chan, W., Cherng, D., and Chang, J. (2006). Human object tracking algorithm with human color structure descriptor for video surveillance systems. In *Proceedings of the IEEE International Conference on Multimedia And Expo*, pages 2097–2100.
- Collins, R., Lipton, A., Kanade, T., Fujiyoshi, H., Duggins, D., Tsin, Y., Tolliver, D., Enomoto, N., Hasegawa, O., Burt, P., and Wixson, L. (2000). A system for video surveillance and monitoring. Technical report, Robotics Institute, Pittsburgh, PA.
- Cucchiara, R., Grana, C., Piccardi, M., and Prati, A. (2003). Detecting moving objects, ghosts, and shadows in video streams. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(10):1337–1342.
- Dadal, N. and Triggs, B. (2005). Histograms of oriented gradients for human detection. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, volume 1, pages 886–893.
- Detriche, P. and Lanore, J. (1980). An acoustic study of pulsation characteristics of fires. *Fire Technology*, 16(3):204–211.
- Doucet, A., Freitas, N., and Gordon, N. (2001). *Sequential Monte Carlo methods in practice*. Springer Verlag.
- Drysdale, D. (2011). *An introduction to fire dynamics*. Wiley.
- Friedman, N. and Russell, S. (1997). Image segmentation in video sequences: A probabilistic approach. In *Proceedings of the 13th Conference Uncertainty in Artificial Intelligence (UAI)*, pages 175–181.
- Grimson, W., Stauffer, C., Romano, R., and Lee, L. (1998). Using adaptive tracking to classify and monitor activities in a site. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 22–29.
- Hammis, A., Yang, J., and Kashiwagi, T. (1992). An experimental investigation of the pulsation frequency of flames. *Symposium (International) on Combustion*, 24(1):1695–1702.
- Hariatoglu, I., Harwood, D., and Davis, L. (1998). W4: Who? When? Where? What? - A real time system for detecting and tracking people. In *Proceedings of the 3rd IEEE International Conference on Face and Gesture Recognition*, pages 222–227.
- Haritaoglu, D., Harwood, D., and Davis, L. (1999). Hydra: multiple people detection and tracking using silhouettes. In *Proceedings of the IEEE International Conference on Image Analysis and Processing*, pages 280–285.

- Haritaoglu, D., Harwood, D., and Davis, L. (2000). An appearance-based body model for multiple people tracking. In *Proceedings of the 15th IEEE International Conference on Pattern Recognition*, volume 4, pages 184–187.
- Healey, G., Slater, D., Lin, T., Drda, B., and Geodeke, A. (1993). A system for real-time fire detection. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 605–606.
- Horng, W., Peng, J., and Chen, C. (2005). A new image-based real-time flame detection method using color analysis. In *Proceedings of the IEEE International Conference on Networking, Sensing and Control*, pages 100–105.
- Horng, W. and Schunck, B. (1981). Determining optical flow. *Artificial Intelligence*, 17(1-3):185–203.
- Javed, O., Rasheed, Z., Shafique, K., and Shah, M. (2003). Tracking across multiple cameras with disjoint views. In *Proceedings of the 9th IEEE International Conference on Computer Vision*, volume 2, pages 952–957.
- Javed, O., Shafique, K., and Shah, M. (2002). A hierarchical approach to robust background subtraction using color and gradient information. In *Proceedings of the IEEE Workshop on Motion and Video Computing*, pages 22–27.
- KaewTraKulPong, P. and Bowden, R. (2001). An improved adaptive background mixture models for real-time tracking with shadow detection. In *Proceedings of the 2nd European Workshop on Advanced Video Based Surveillance Systems*, volume 25, pages 1–5.
- Kalman, R. (1960). A new approach to linear filtering and prediction problems. *Journal of Basic Engineering*, 82(1):35–45.
- Kang, S., Hwang, B., and Lee, S. (2003). Multiple people tracking based on temporal color feature. *International Journal of Pattern Recognition and Artificial Intelligence*, 17(6):931–949.
- Kettner, V. and Zabih, R. (1999). Bayesian multi-camera surveillance. In *Proceedings of the IEEE International Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, volume 2.
- Khan, S., Javed, O., and Shah, M. (2001). Tracking in uncalibrated cameras with overlapping field of view. In *Proceedings of the 2nd IEEE Workshop on Performance Evaluation of Tracking and Surveillance*, volume 50.
- Khan, S. and Shah, M. (2003). Consistent Labeling of tracked objects in multiple cameras with overlapping fields of view. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(10):1355–1360.
- Khan, S. and Shah, M. (2006). A multi-view approach to tracking people in crowded scenes using a planar homography constraint. In *Proceedings of the European Conference on Computer Vision (ECCV)*, pages 133–146.
- Kim, K., Chalidabhongse, T., Harwood, D., and Davis, L. (2005). Real-time foreground-background segmentation using codebook model. *Real-Time Imaging*, 11(3):172–185.
- Kim, K. and Davis, L. (2006). Multi-Camera tracking and segmentation of occluded people on ground plane using search-guided particle filter. In *Proceedings of the 9th European Conference on Computer Vision (ECCV)*, pages 98–109.

- Ko, B. C., Cheong, K., and Nam, J. (2009). Fire detection based on vision sensor and support vector machines. *Fire Safety Journal*, 44(3):322–329.
- Koller, D., Weber, J., and Malik, J. (1994). Robust multiple car tracking with occlusion reasoning. In *Proceedings of the European Conference on Computer Vision (ECCV)*, pages 186–196.
- Lipton, A., Fujiyoshi, H., and Patil, R. (1998). Moving target classification and tracking from real-time video. In *Proceedings of the 4th IEEE Workshop Applications of Computer Vision*, pages 8–14.
- Liu, C. and Ahuja, N. (2004). Vision based fire detection. In *Proceedings of the 17th IEEE International Conference on Pattern Recognition (ICPR)*, volume 4, pages 134–137.
- Lu, W. and Tan, Y. (2001). A color histogram based people tracking system. In *Proceedings of the IEEE International Symposium on Circuits and Systems*, volume 2, pages 137–140.
- Makris, D., Ellis, T., and Black, J. (2004). Bridging the Gaps between Cameras. In *Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition (CVPR)*, volume 2, pages 205–210.
- Mallat, S. (1989). A theory for multiresolution signal decomposition: the wavelet representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(7):674–693.
- Matsumura, A., Iwai, Y., and Yachida, M. (2002). Tracking people by using color information from omnidirectional images. In *Proceedings of the 41st SICE Annual Conference*, volume 3, pages 1772–1777.
- Moeslund, T., Hilton, A., and Kruger, V. (2006). A survey of advances in vision-based human motion capture and analysis. *Computer Vision and Image Understanding*, 104(2):90–126.
- Nummiaro, K., Koller-Meier, E., and Van Gool, L. (2003). An adaptive color-based particle filter. volume 21, pages 99–110. Elsevier.
- Okuma, K., Taleghani, A., Freitas, N., Little, J., and Lowe, D. (2004). A boosted particle filter: Multi-target detection and tracking. In *Proceedings of the 8th European Conference on Computer Vision (ECCV)*, pages 28–39.
- Phillips III, W., Shah, M., and Vitoria Lobo, N. (2002). Flame recognition in video. *Pattern Recognition Letters*, 23(1-3):319–327.
- Ridder, C., Munkelt, O., and Kirchner, H. (1995). Adaptive background estimation and foreground detection using kalman filter. In *Proceedings of the International Conference on Recent Advances in Mechatronics*, pages 193–199. Citeseer.
- Rioul, O. and Vetterli, M. (1991). Wavelets and signal processing. *IEEE Signal Processing Magazine*, 8(4):14–38.
- Roberts, R., Potthast, C., and Dellaert, F. (2009). Learning general optical flow subspaces for ego-motion estimation and detection of motion anomalies. In *Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 57–64.
- Stauffer, C. and Grimson, W. (1999). Adaptive background mixture models for real-time tracking. In *Proceedings of the IEEE Computer Society on Computer Vision and Pattern Recognition (CVPR)*, volume 2.

- Thacker, N. and Lacey, A. (2006). Tutorial: the likelihood interpretation of the Kalman filter. *TINA Memos: Advanced Applied Statistics*, 2.
- Töreyn, B., Dedeoğlu, Y., and Cetin, A. (2005). Flame detection in video using hidden Markov models. In *Proceedings of the IEEE International Conference on Image Processing (ICIP)*, volume 2, pages 1230–1233.
- Töreyn, B., Dedeoğlu, Y., Güdükbay, U., and Cetin, A. (2006). Computer vision based method for real-time fire and flame detection. *Pattern Recognition Letters*, 27(1):49–58.
- Toyama, K., Krumm, J., Brumitt, B., and Meyers, B. (1999). Wallflower: Principles and practice of background maintenance. In *Proceedings of the 7th IEEE International Conference on Computer Vision*, volume 1, pages 255–261.
- Tzanetakis, G., Essl, G., and Cook, P. (2001). Audio analysis using the discrete wavelet transform. In *Proceedings of the Conference in Acoustics and Music Theory Applications*.
- Van Der Merwe, R., Doucet, A., and Freitas, N. (2001). The unscented particle filter. *Advances in Neural Information Processing Systems*.
- Vassili, V., Sazonov, V., and Andreeva, A. (2003). A survey on pixel-based skin color detection techniques. In *Proceedings of the 13th International Conference on Computer Graphics and Vision*, pages 85–92.
- Wang, L., Hu, W., and Tan, T. (2002). Face tracking using motion-guided dynamic template matching. In *Proceedings of the 5th Asian Conference on Computer Vision (ACCV)*, volume 2, pages 448–453.
- Wren, C., Azarbayejani, A., Darrel, T., and Pentland, A. (1997). Pfunder: real-time tracking of the human body. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 19(7):780–785.
- Yang, C., Duraiswami, L., and Davis, L. (2005). Fast multiple object tracking via a hierarchical particle filter. In *Proceedings of the 10th IEEE International Conference on Computer Vision*, volume 1, pages 212–219.
- Zhao, T. and Nevatia, R. (2004). Tracking multiple humans in complex situations. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26(9):1208–1221.
- Zhu, X., Wang, J., Fu, W., Lu, H., and Fang, Y. (2011). Global trajectory construction across multi-cameras via graph matching. In *Proceedings of the 6th IEEE International Conference on Image and Graphics (ICIG)*, pages 801–806.
- Zivkovic, Z. and van der Heijden, F. (2006). Efficient adaptive density estimation per image pixel for the task of background subtraction. *Pattern Recognition Letters*, 27(7):773–780.