

**Detecção e seguimento de objectos em imagens
termográficas: análise experimental
de modelos de descrição**

Tânia Zhao Zhu



Mestrado em Engenharia Electrotécnica e de Computadores
Área de Especialização de Telecomunicações
Departamento de Engenharia Electrotécnica
Instituto Superior de Engenharia do Porto

2011

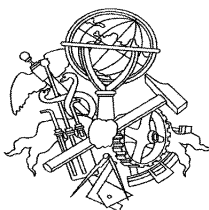
Este relatório satisfaz, parcialmente, os requisitos que constam da Ficha de Disciplina de Tese/Dissertação, do 2º ano, do Mestrado em Engenharia Electrotécnica e de Computadores

Candidata: Tânia Zhao Zhu, N° 1060424, 1060424@isep.ipp.pt

Orientação científica: Paula Maria Marques Moura Gomes Viana, pmv@isep.ipp.pt

Empresa: INESC Porto

Supervisão: Pedro Carvalho, pedro.carvalho@inescporto.pt



Mestrado em Engenharia Electrotécnica e de Computadores

Área de Especialização de Telecomunicações

Departamento de Engenharia Electrotécnica

Instituto Superior de Engenharia do Porto

20 de Julho de 2011

Agradecimentos

Em primeiro lugar, quero dirigir os meus agradecimentos ao meu supervisor do INESC Porto (Instituto de Engenharia de Sistemas e Computadores do Porto), Eng.º Pedro Carvalho, e à minha orientadora científica, a Prof. Doutora Paula Maria Marques Moura Gomes Viana. A ambos, o meu obrigado pela orientação, ensinamentos, disponibilidade em acompanhar o trabalho desenvolvido e oportunidade conferida em realizar um estágio curricular.

Gostaria de deixar uma palavra de apreço ao Eng.º Lucian Ciobanu e aos restantes colaboradores do INESC Porto que tão bem me receberam e acompanharam ao longo do estágio, tornando a minha integração fácil e rápida.

Agradeço, profundamente, a dois bons colegas, Paulo André da Silva Pereira e Telmo Afonso Vilar Gonçalves Oliveira pela extrema paciência e a possibilidade de partilha de conhecimentos.

Finalmente, à minha família e pessoas mais próximas, um “Muito Obrigado” pelo apoio que me foi conferido.

Resumo

A instalação de sistemas de videovigilância, no interior ou exterior, em locais como aeroportos, centros comerciais, escritórios, edifícios estatais, bases militares ou casas privadas tem o intuito de auxiliar na tarefa de monitorização do local contra eventuais intrusos. Com estes sistemas é possível realizar a detecção e o seguimento das pessoas que se encontram no ambiente local, tornando a monitorização mais eficiente.

Neste contexto, as imagens típicas (imagem natural e imagem infravermelha) são utilizadas para extrair informação dos objectos detectados e que irão ser seguidos. Contudo, as imagens convencionais são afectadas por condições ambientais adversas como o nível de luminosidade existente no local (luzes muito fortes ou escuridão total), a presença de chuva, de nevoeiro ou de fumo que dificultam a tarefa de monitorização das pessoas. Deste modo, tornou-se necessário realizar estudos e apresentar soluções que aumentem a eficácia dos sistemas de videovigilância quando sujeitos a condições ambientais adversas, ou seja, em ambientes não controlados, sendo uma das soluções a utilização de imagens termográficas nos sistemas de videovigilância.

Neste documento são apresentadas algumas das características das câmaras e imagens termográficas, assim como uma caracterização de cenários de vigilância. Em seguida, são apresentados resultados provenientes de um algoritmo que permite realizar a segmentação de pessoas utilizando imagens termográficas. O maior foco desta dissertação foi na análise dos modelos de descrição (Histograma de Cor, HOG, SIFT, SURF) para determinar o desempenho dos modelos em três casos: distinguir entre uma pessoa e um carro; distinguir entre duas pessoas distintas e determinar que é a mesma pessoa ao longo de uma sequência.

De uma forma sucinta pretendeu-se, com este estudo, contribuir para uma melhoria dos algoritmos de detecção e seguimento de objectos em sequências de vídeo de imagens termográficas. No final, através de uma análise dos resultados provenientes dos modelos de descrição, serão retiradas conclusões que servirão de indicação sobre qual o modelo que melhor permite discriminar entre objectos nas imagens termográficas.

Palavras – Chaves

Imagens termográficas, Segmentação, Histograma de Cor, HOG, SIFT, SURF

Abstract

This report presents the work accomplished for the Thesis/Dissertation module of the Masters Degree in Electrical and Computer Engineering – within the Telecommunications area of expertise.

Currently, automatic monitoring in video surveillance systems in environments such as airports, shopping malls, government buildings, office buildings, and private home is done through the use of detection and object tracking techniques.

Natural images and near-infrared images are mainly accessed through video surveillance in order to extract information on the object detected and subsequently being tracking. However, due to variations in environmental conditions within surveillance scenarios, severe drawbacks are exhibited when used for night-time surveillance and/or in scenes with harsh environmental conditions such as strong light, total darkness, smoke, rain and fog. Therefore, it became more and more important to present a solution that could overcome those disadvantages. A possible solution is to make use of thermal images.

This dissertation aims to analyze descriptors models such as Color Histograms, HOG, SIFT and SURF, to conclude if they are able or not to be used to distinguish between an object representing a non-person and a person and between two different persons due to their similarity. In addition, a study of a set of scenarios with harsh environmental conditions and also results of a segmentation algorithm are presented.

In short, the entire study intends to contribute for a better performance of video object detection and tracking algorithms. At the end, through the analysis of the set of results from the descriptors models, conclusions are drawn in order to indicate which of the models can better distinguish the detected objects in thermal images.

Keywords

Thermal Images, Segmentation, Color Histogram, HOG, SIFT, SURF

Índice

AGRADECIMENTOS	I
RESUMO	III
ABSTRACT	V
ÍNDICE DE FIGURAS	IX
ÍNDICE DE GRÁFICOS	XIII
ÍNDICE DE TABELAS	XV
ACRÓNIMOS	XVII
1. INTRODUÇÃO	1
1.1 CONTEXTO	1
1.2 OBJECTIVOS	5
1.3 CONCEITOS BÁSICOS	5
1.4 ESTRUTURA DO RELATÓRIO	7
2. TERMOGRAFIA.....	9
2.1 CÂMARAS TERMOGRÁFICAS.....	9
2.2 CAPTURA DE IMAGENS TERMOGRÁFICAS	11
3. ÁREAS DE APLICAÇÃO E TÉCNICAS DE TRATAMENTO DE IMAGENS TERMOGRÁFICAS	15
3.1 SEGURANÇA RODOVIÁRIA	16
3.2 VIGILÂNCIA COM RECURSO A UMA CÂMARA FIXA.....	20
3.3 VIGILÂNCIA COM RECURSO A UMA CÂMARA MÓVEL.....	22
4. CARACTERIZAÇÃO DE CENÁRIOS E IMPACTO DA VARIAÇÃO LUMINOSA	27
4.1 DESCRIÇÃO E ANÁLISE DE CENÁRIOS.....	27
4.2 ANÁLISE DO EFEITO DE VARIAÇÕES ABRUPTAS DE ILUMINAÇÃO	30
5. SEGMENTAÇÃO/DETECÇÃO DE OBJECTOS.....	37
5.1 SEGMENTAÇÃO BASEADA EM THRESHOLD.....	37
5.2 AVALIAÇÃO DOS ALGORITMOS DE SEGMENTAÇÃO	41

6. DESCRIÇÃO E SEGUIMENTO	47
6.1 DESCRITOR: HISTOGRAMA DE COR	47
6.2 DESCRITOR: HISTOGRAMA DE GRADIENTES (HOG)	56
6.3 DESCRITOR: SIFT	61
6.4 DESCRITOR: SURF	67
6.5 CONCLUSÕES DECORRENTES DA ANÁLISE EFECTUADA	72
7. CONCLUSÃO E DESENVOLVIMENTOS FUTUROS.....	75
7.1 CONCLUSÃO	75
7.2 DESENVOLVIMENTOS FUTUROS.....	77
REFERÊNCIAS DOCUMENTAIS	79

Índice de Figuras

Figura 1.2 - 1 Espectro electromagnético	3
Figura 1.2 - 2 Imagem à esquerda: Filtro de corte infravermelho numa câmara; Ao centro: Posição do filtro durante o dia e posição do filtro durante a noite; À direita: Exemplo de uma fonte de luz artificial.....	3
Figura 1.2 - 3 Esquerda: Imagem Natural; Direita: Imagem Infravermelha	4
Figura 1.4 - 1 Ilustração dos conceitos básicos associados à detecção e seguimento.....	6
Figura 2.1 - 1 Espectro Electromagnético	9
Figura 2.1 - 2 Reflexão, absorção e transmissão de energia.....	10
Figura 2.1 - 3 Câmara termográfica: Imagem resultante em escala cinza e pseudo-cores.....	11
Figura 2.2 - 1 Paleta de cores: Imagem da esquerda em Gray; Imagem ao centro em Iron; Imagem da direita em Rainbow	12
Figura 2.2 - 2 Escala ajustável: Detalhe observado	12
Figura 2.2 - 3 Imagem termográfica com uma escala fixa	13
Figura 3.1 - 1 Instalação da câmara termográfica num carro	16
Figura 3.1 - 2 Tratamento de artefactos escuros, [OMA10]	17
Figura 3.1 - 3 Esquerda: Imagem dividida em células de 4x8; Direita: Resultado do descritor HOG, [OMA10] ...	19
Figura 3.2 - 1 Da esquerda para a direita: Imagem termográfica capturada; Imagem normalizada; Eliminação de zonas incandescentes; Imagem binária, [FER10]	20
Figura 3.2 - 2 a) Região de interesse seleccionada b) Histograma aplicado c) Resultado da inferência, [FER10]	21
Figura 3.2 - 3 a) Sub-regiões obtidas na sub-etapa de delimitação vertical; b) Imagem resultante após aplicação do <i>threshold</i> ; c) Resultados obtidos, [FER10]	21
Figura 3.3 - 1 Da esquerda para a direita: Imagem no instante $t-1$; Imagem no instante t ; Posição actual e posição anterior; aplicação do <i>threshold</i> para eliminar a posição fantasma, [FER10A]	23
Figura 3.3 - 2 Resultados obtidos com um método de subtracção de imagens	23
Figura 3.3 - 3 Resultados obtidos com um método de <i>optical flow</i> de Lucas-Kanade.....	24
Figura 3.3 - 4 <i>Elliptic contour model</i> e divisão do contorno elíptico, [TRE06].....	24
Figura 4.1 - 1 Recriação de acontecimentos	28
Figura 4.1 - 2 Esquerda: corredor comum; Ao centro: corredor com vidros; Esquerda: garagem	29

Figura 4.2 - 1 Comparação entre as imagens da câmara dia e noite e termográfica em situações de variação da luminosidade.....	31
Figura 4.2 - 2 Planta do local da gravação.....	32
Figura 4.2 - 3 Ilustração do porquê da ocorrência do erro	36
Figura 5.1 - 1 Fluxograma do código de segmentação.....	37
Figura 5.1 - 2 a) Imagem termográfica com escala ajustável; b) Aplicação do <i>threshold</i> = 60; c) Imagem termográfica com escala fixa; d) Aplicação do <i>threshold</i> = 40.....	38
Figura 5.1 - 3 a) <i>Threshold</i> da imagem com escala ajustável; Operações morfológicas na imagem com escala ajustável;	39
Figura 5.1 - 4 Da coluna esquerda para a direita: Resultado da aplicação do <i>threshold</i> ; Resultado da aplicação das operações morfológicas; Resultado obtido com os critérios de selecção.....	40
Figura 5.1 - 5 Da esquerda para a direita: Imagem original; Máscara; Resultado obtido com a operação AND..	41
Figura 5.1 - 6 Da esquerda para a direita: Imagem original; Imagem gerada com a ferramenta de Bounding Box; Imagem gerada com a ferramenta de Berkeley.....	42
Figura 5.1 - 7 Imagem10 - Da esquerda para a direita: Imagem de referência de segmentação; Resultado do ISegmentation; Resultado do OpenCv; Resultado do <i>threshold</i> ; Resultado da combinação do OpenCv com o <i>threshold</i>	45
Figura 5.1 - 8 Imagem67 - Da esquerda para a direita: Imagem de referência de segmentação; Resultado do ISegmentation; Resultado do OpenCv; Resultado do <i>threshold</i> ; Resultado da combinação do OpenCv com o <i>threshold</i>	45
Figura 6.1 - 1 Comparação entre objectos diferentes na mesma imagem	48
Figura 6.1 - 2 Comparação entre o mesmo objecto, mas em imagens diferentes	49
Figura 6.1 - 3 Comparação de o objecto A com o objecto B na imagem seguinte.....	49
Figura 6.1.1 - 1 Objectos de estudo	51
Figura 6.1.1 - 2 Ilustração dos possíveis erros ocorridos	53
Figura 6.1.2 - 1 Ilustração do erro ocorrido	54
Figura 6.1.2 - 2 Ilustração dos erros ocorridos.....	56
Figura 6.2.1 - 1 Ilustração dos erros ocorridos.....	59
Figura 6.2.2 - 1 Ilustração do erro ocorrido	61

Figura 6.3 - 1 Etapa de comparação do SIFT	62
Figura 6.5 - 1 Da esquerda para à direita: Escala ajustável e escala fixa	72
Figura 6.5 - 2 Imagens superiores: Ilustração de momentos em que a <i>frame rate</i> é inconstante; Imagens inferiores: Ilustração das várias posições assumidas pelo mesmo objecto ao longo de uma sequência	73

Índice de Gráficos

Gráfico 4.2 - 1 Variação da luminosidade com imagens naturais	33
Gráfico 4.2 - 2 Variação da luminosidade com imagens infravermelhas sem luz infravermelha artificial	33
Gráfico 4.2 - 3 Variação da luminosidade com imagens infravermelhas com luz infravermelha artificial	34
Gráfico 4.2 - 4 Variação da luminosidade com imagens termográficas.....	34
Gráfico 5.1 - 1 Avaliação dos algoritmos de segmentação: escala ajustável	43
Gráfico 5.1 - 2 Avaliação dos algoritmos de segmentação: escala fixa	43
Gráfico 6.1.1 - 1 Comparação entre objectos diferentes na mesma imagem	50
Gráfico 6.1.1 - 2 Comparação entre objectos em imagens diferentes	52
Gráfico 6.1.2 - 1 Comparação entre objectos diferentes na mesma imagem	53
Gráfico 6.1.2 - 2 Comparação entre objectos em imagens diferentes	55
Gráfico 6.2.1 - 1 Comparação entre objectos diferentes na mesma imagem	57
Gráfico 6.2.1 - 2 Comparação entre objectos em imagens diferentes	58
Gráfico 6.2.2 - 1 Comparação entre objectos diferentes na mesma imagem	59
Gráfico 6.2.2 - 2 Comparação entre objectos em imagens diferentes	60
Gráfico 6.3.1 - 1 Comparação entre objectos diferentes na mesma imagem	63
Gráfico 6.3.1 - 2 Comparação entre objectos em imagens diferentes	64
Gráfico 6.3.2 - 1 Comparação entre objectos diferentes na mesma imagem	65
Gráfico 6.3.2 - 2 Comparação entre objectos em imagens diferentes	66
Gráfico 6.4.1 - 1 Comparação entre objectos diferentes na mesma imagem	68
Gráfico 6.4.1 - 2 Comparação entre objectos em imagens diferentes	69
Gráfico 6.4.2 - 1 Comparação entre objectos na mesma imagem.....	70
Gráfico 6.4.2 - 2 Comparação entre objectos em imagens diferentes	71

Índice de Tabelas

Tabela 4.2 - 1 Tabela de eventos	32
Tabela 4.2 - 2 Dados extraídos dos gráficos	35
Tabela 6.1.1 - 1 Histograma de cor (escala variável): Dados obtidos para a situação a)	51
Tabela 6.1.1 - 2 Histograma de cor (escala variável): Dados obtidos para a situação b) e c)	52
Tabela 6.1.2 - 1 Histograma de cor (escala fixa): Dados obtidos para a situação a)	54
Tabela 6.1.2 - 2 Histograma de cor (escala fixa): Dados obtidos para a situação b) e c)	55
Tabela 6.2.1 - 1 Histograma de Gradientes (escala ajustável): Dados obtidos para a situação a)	57
Tabela 6.2.1 - 2 Histograma de Gradientes (escala ajustável): Dados obtidos para a situação b) e c)	58
Tabela 6.2.2 - 1 Histograma de Gradientes (escala fixa): Dados obtidos para a situação a)	60
Tabela 6.2.2 - 2 Histograma de Gradientes (escala fixa): Dados obtidos para a situação b) e c)	61
Tabela 6.3.1 - 1 SIFT (escala ajustável): Dados obtidos para a situação a)	63
Tabela 6.3.1 - 2 SIFT (escala ajustável): Dados obtidos para a situação b) e c)	64
Tabela 6.3.2 - 1 SIFT (escala fixa): Dados obtidos para a situação a)	66
Tabela 6.3.2 - 2 SIFT (escala fixa): Dados obtidos para a situação b) e c)	67
Tabela 6.4.1 - 1 SURF (escala ajustável): Dados obtidos para a situação a)	68
Tabela 6.4.1 - 2 SURF (escala ajustável): Dados obtidos para a situação b) e c)	69
Tabela 6.4.2 - 1 SURF (escala fixa): Dados obtidos para a situação a)	71
Tabela 6.4.2 - 2 SURF (escala fixa): Dados obtidos para a situação b) e c)	72

Acrónimos

FIR	<i>Far-Infrared</i>
HOG	<i>Histogram of Oriented Gradients</i>
MIR	<i>Middle-Infrared</i>
NIR	<i>Near-Infrared</i>
ROI	<i>Region of Interest</i>
SIFT	<i>Scale-Invariant Features Transform</i>
SURF	<i>Speeded Up Robust Features</i>
SVM	<i>Support Vector Machine</i>
Wi-Fi	<i>Wireless-Fidelity</i>

1. Introdução

Termografia é o nome atribuído à técnica de detectar variações de temperaturas de um objecto através de imagens, utilizando a radiação térmica emitida pelo objecto. Esta técnica desempenha um papel de grande importância em vários sectores, na medida em que a capacidade de identificar situações de perigo permite ajudar em termos de prevenção e protecção. Exemplos da aplicabilidade da termografia em vários ramos de actividade são: na indústria automobilística e aeronáutica na manutenção preditiva eléctrica e mecânica; no controle de reactores e torres de refrigeração na indústria química; na engenharia civil para a avaliação do isolamento térmico de edifícios, identificação de zonas de infiltração e fugas, identificação de sistemas de aquecimento em pavimentos e identificação de colónias de insectos em elementos de madeira; na área militar e policial para o combate a crimes em ambientes de total escuridão ou quando os criminosos se encontram dissimulados no ambiente local; permite auxiliar os bombeiros na localização de vítimas em locais de fumo intenso e escuridão; na área da segurança rodoviária no auxílio da visão nocturna dos automobilistas. Pelo facto de ser uma técnica não invasiva, a termografia é, também, aplicada na área da medicina para a detecção de síndromas gripais em grandes aeroportos e espaços públicos ou como um método imagiológico para o diagnóstico de inúmeras doenças, muitas vezes não detectadas por outros métodos e de forma ainda mais precoce. Uma aplicação recente deste método está relacionada com a detecção do cancro da mama.

1.1 CONTEXTO

Na área de segurança civil, a automatização da monitorização dos sistemas de videovigilância é uma necessidade crescente para as empresas ou particulares que necessitem de um sistema de videovigilância contra eventuais intrusos. Este destaque

deve-se ao facto de a monitorização destes sistemas requerer operadores humanos qualificados, o que implica elevados custos a quem adquira este tipo de serviço. Além disso, a produtividade humana é afectada por factores como o sono, a fadiga ou o trabalho monótono. Tais factores podem originar possíveis erros na execução da tarefa de monitorização. De modo a reduzir os erros humanos e os custos representativos para as empresas, tem sido dada ênfase ao desenvolvimento de algoritmos de detecção e seguimento de objectos, em particular, de pessoas.

A instalação de sistemas de videovigilância, no interior ou exterior, em locais como aeroportos, centros comerciais, escritórios, edifícios estatais, bases militares ou casas privadas tem o intuito de auxiliar na tarefa de monitorização do local contra eventuais intrusos. Com estes sistemas é possível realizar a detecção e o seguimento das pessoas que se encontram no ambiente local, tornando a monitorização mais eficiente. No entanto, algumas condições ambientais adversas como o nível de luminosidade existente no local (luzes muito fortes ou escuridão total), a presença de chuva, de nevoeiro ou de fumo dificultam a tarefa de monitorização das pessoas. Deste modo, tornou-se necessário realizar estudos e apresentar soluções que aumentem a eficácia dos sistemas de videovigilância quando sujeitos a condições ambientais adversas, ou seja, em ambientes não controlados.

Um dos possíveis cenários de aplicação das imagens termográficas na área de vigilância é um projecto que se encontra em fase de desenvolvimento no INESC Porto. Este projecto consiste em desenvolver uma plataforma móvel, na qual irá estar instalada a câmara termográfica e que terá como função auxiliar os operacionais na área de segurança. Devido ao facto de um dos requisitos da plataforma móvel ser a autonomia, a utilização de uma câmara termográfica face a uma câmara dia e noite, em modo nocturno, torna-se uma grande vantagem na medida em que a câmara termográfica não necessita de uma fonte luminosa dedicada para situações de escuridão, levando a um menor consumo das baterias da plataforma móvel. Esta plataforma móvel permitirá reduzir custos na contratação de operacionais da área de vigilância, na medida em que será possível realizar rondas, fazer o reconhecimento de objectos e seguir intrusos.

As imagens captadas pelos sistemas de videovigilância provêm de dois tipos de câmaras: câmaras naturais ou de câmaras intituladas de câmara dia e noite. Estas últimas, em

comparação com a câmara natural, têm a particularidade de permitir a monitorização durante o dia e durante a noite. Tal é possível porque a câmara dia e noite opera não só na gama de luz visível, como as câmaras naturais, mas também na gama do infravermelho mais próximo da luz visível (*Near Infrared Rays*) do espectro electromagnético, ilustrado na Figura 1.2 - 1.

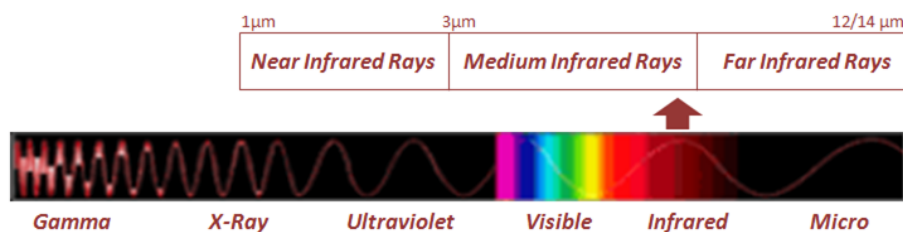


Figura 1.2 - 1 Espectro electromagnético

Para efeitos de clareza, ao longo do documento o termo imagem natural será utilizado para referir às imagens captadas pela câmara dia e noite em modo diurno e as imagens infravermelhas às imagens captadas em modo nocturno.

A câmara dia e noite possui um filtro de corte infravermelho que é colocado durante o dia para não distorcer as cores percebidas pelo olho humano e removido durante a noite para a captação de luz infravermelha proveniente de fontes de luz natural ou fontes artificiais de luz infravermelha, como se pode constatar na Figura 1.2 - 2.



Figura 1.2 - 2 Imagem à esquerda: Filtro de corte infravermelho numa câmara; Ao centro: Posição do filtro durante o dia e posição do filtro durante a noite; À direita: Exemplo de uma fonte de luz artificial

Na Figura 1.2 - 3 encontram-se exemplos de imagens provenientes da câmara dia e noite em modo diurno (imagem natural) e em modo nocturno (imagem infravermelha). Na imagem infravermelha verifica-se uma zona de maior intensidade que é provocada pela luz infravermelha incidente nessa região proveniente de uma fonte de luz infravermelha artificial.



Figura 1.2 - 3 Esquerda: Imagem Natural; Direita: Imagem Infravermelha

As imagens convencionais (imagem natural e imagem infravermelha) são afectadas pela condição de luminosidade existente no local, na medida em que em situações de ausência de luz, as imagens naturais e as imagens infravermelhas sem uma fonte de luz infravermelha artificial tornam-se inutilizáveis. No caso das imagens infravermelhas com uma fonte de luz artificial, no momento da variação luminosa ocorrem duas situações: a) quando a luz se encontra acesa e é desligada, há um tempo de ajuste da câmara até apresentar a imagem; b) quando a luz se encontra desligada e é acesa há um tempo de clarão registado até a imagem ser apresentada, situação ilustrada na sub-secção 4.2 do capítulo 4. Desta forma, a aplicação das imagens convencionais nos algoritmos de detecção e seguimento das pessoas torna-se, por vezes, impraticável.

Como alternativa para contornar as desvantagens apresentadas pelas câmaras dia e noite, são apresentadas soluções utilizando câmaras termográficas nos sistemas de videovigilância, dado estas apresentarem-se menos influenciáveis perante situações de fumos, chuva e nevoeiro e nada susceptíveis à variação luminosa do local.

Quer a câmara termográfica, quer a câmara dia e noite são câmaras infravermelhas. No entanto, estas operam em regiões diferentes da gama do infravermelho. A câmara termográfica utiliza a região do infravermelho médio e longo enquanto a câmara dia e noite opera na região do infravermelho curto. Nesta dissertação o termo infravermelho será utilizado para referir as imagens captadas pela câmara dia e noite em modo nocturno e o termo termográfico para as imagens captadas pela câmara termográfica.

1.2 OBJECTIVOS

Pretende-se com este trabalho realizar um estudo sobre a detecção e seguimento de objectos em imagens termográficas, dando ênfase à análise experimental de modelos de descrição.

O trabalho encontra-se dividido em 4 etapas. Na primeira etapa é feito um estudo das propriedades das câmaras termográficas.

A segunda etapa consiste na caracterização de sequências típicas num cenário de vigilância. Por exemplo, numa garagem teremos como elementos carros, pessoas a movimentarem-se, fontes de iluminação, postes e portas; num corredor, uma ou mais pessoas a movimentarem-se, portas e os revestimentos desse local.

Na terceira etapa é realizada uma análise de diferentes técnicas para a detecção e seguimento de pessoas e é desenvolvido um código que permita identificar os elementos considerados como sendo de interesse, nomeadamente, as pessoas presentes num cenário.

Na última etapa analisa-se e comparam-se diferentes modelos de descrição de objectos de forma a determinar qual deles será o melhor a diferenciar os vários elementos entre si.

1.3 CONCEITOS BÁSICOS

De forma a tornar clara a sua utilização ao longo da dissertação, apresenta-se a seguir uma lista e ilustrações (Figura 1.4 - 1) das definições de alguns conceitos básicos na área da detecção e seguimento

<i>Background</i>	Termo utilizado para referir o fundo (tipicamente estático) de um cenário.
<i>Bounding Box</i>	Termo utilizado para referir uma representação aproximada do objecto que indica a sua posição na imagem. Normalmente, a representação é feita recorrendo ao uso de um rectângulo.
<i>Foreground</i>	Termo utilizado para referir o que não pertence ao <i>background</i> , contendo informação de interesse.

Ocultação Do termo em Inglês *Occlusion*, utilizado para indicar que um objecto é oculto por um outro ou pelo fundo. Este acontecimento pode ser parcial ou total.

Segmentação Do termo em Inglês *Segmentation*, não há uma definição concreta, mas refere-se ao processo de separação da informação relevante (*foreground*) da não relevante (*background*).

Seguimento Do termo em Inglês *Tracking*, refere-se à identificação consistente de um objecto ao longo de uma sequência.

Threshold Termo utilizado para referir o valor que serve de decisão.

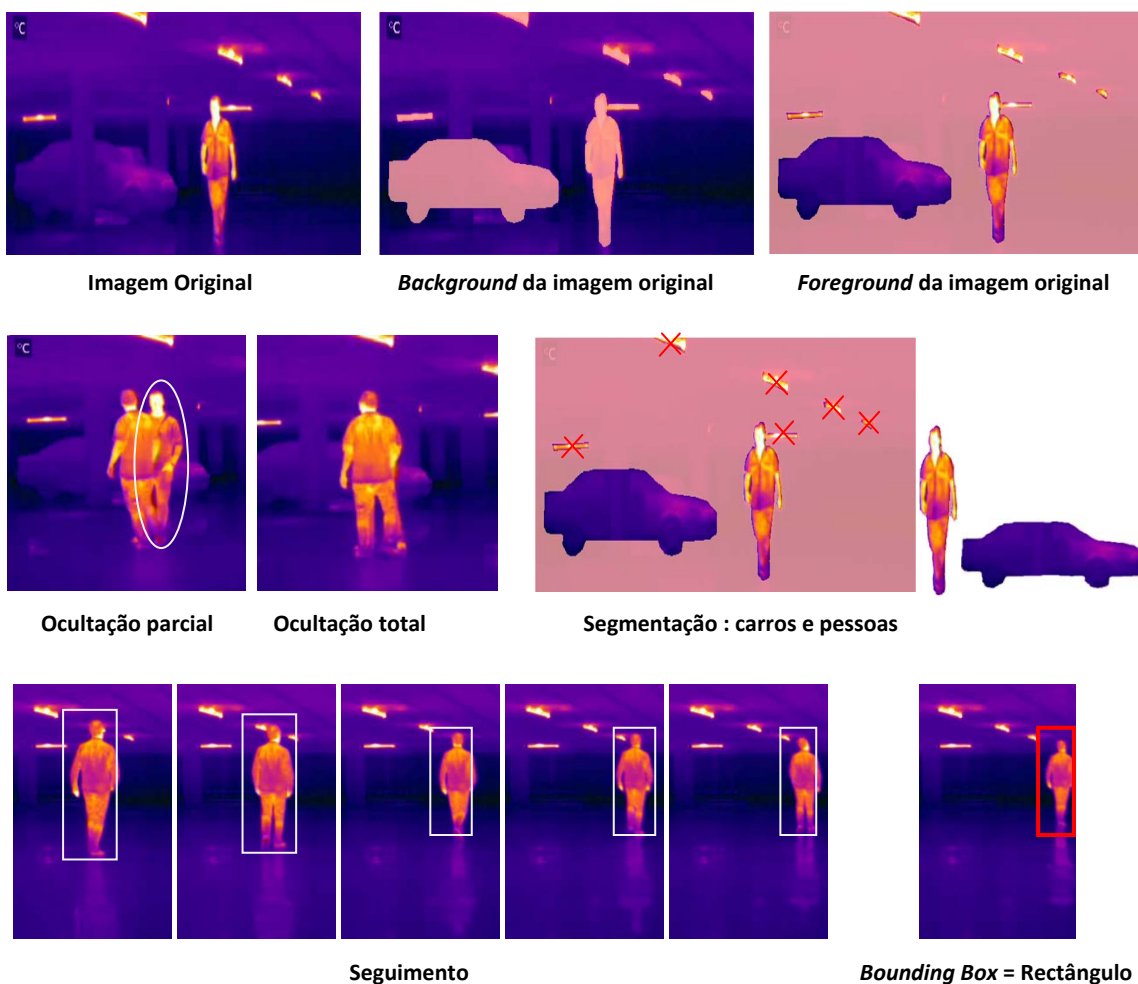


Figura 1.4 - 1 Ilustração dos conceitos básicos associados à detecção e seguimento

1.4 ESTRUTURA DO RELATÓRIO

Este relatório encontra-se estruturado em 7 capítulos. No primeiro capítulo fornece-se ao leitor uma breve introdução ao tema, o contexto desta tese, a motivação que levou à sua realização e os objectivos estabelecidos para a sua prossecução.

No segundo capítulo é realizado um estudo dos elementos da termografia como as propriedades das câmaras termográficas e das imagens termográficas.

O capítulo 3 descreve algumas técnicas utilizadas para a detecção e seguimento de pessoas em áreas como a segurança rodoviária e vigilância com uma câmara fixa ou móvel.

No capítulo seguinte, o leitor encontrará uma caracterização de sequências típicas num cenário de vigilância. Por exemplo, numa garagem teremos como elementos carros, pessoas a movimentarem-se, fontes de iluminação, postes e portas; num cenário interior teremos um corredor, uma ou mais pessoas a movimentarem-se, portas e os revestimentos desse local. É feita ainda uma demonstração dos benefícios da utilização da câmara termográfica em situações de variações luminosas face às imagens naturais e infravermelhas.

No capítulo 5 são apresentados resultados do *software* desenvolvido para a segmentação de objectos. Comparam-se os resultados obtidos com os resultados provenientes de dois algoritmos de segmentação distintos e efectua-se uma avaliação dos algoritmos de segmentação anteriores.

No capítulo 6 apresenta-se a análise experimental de modelos de descrição e no último capítulo apresentam-se as principais conclusões obtidas e perspectiva-se alguns desenvolvimentos futuros.

2. Termografia

Neste capítulo é apresentado ao leitor os principais aspectos dos elementos associados à termografia, nomeadamente as câmaras termográficas e as imagens resultantes da mesma (imagens termográficas).

2.1 CÂMARAS TERMOGRÁFICAS

As câmaras termográficas, tal como as câmaras dia e noite, fazem uso da gama de infravermelhos. Como se pode observar na Figura 2.1 - 1, esta gama encontra-se situada entre a luz visível e as microondas no espectro electromagnético. O seu intervalo de comprimento de onda vai desde $1\mu\text{m}$ até $14\mu\text{m}$, estando subdividido em três partes: infravermelhos curtos (*near-infrared* - NIR) que vai desde $1\mu\text{m}$ até $3\mu\text{m}$; infravermelhos médios (*middle-infrared* - MIR) que vai dos $3\mu\text{m}$ até os $5\mu\text{m}$ e infravermelhos longos (*far-infrared* - FIR) que vai desde os $8\mu\text{m}$ até $14\mu\text{m}$. A zona entre os $5\mu\text{m}$ a $8\mu\text{m}$ é incomum para fins de geração de imagens térmicas devido à alta absorção espectral da atmosfera nesta faixa.

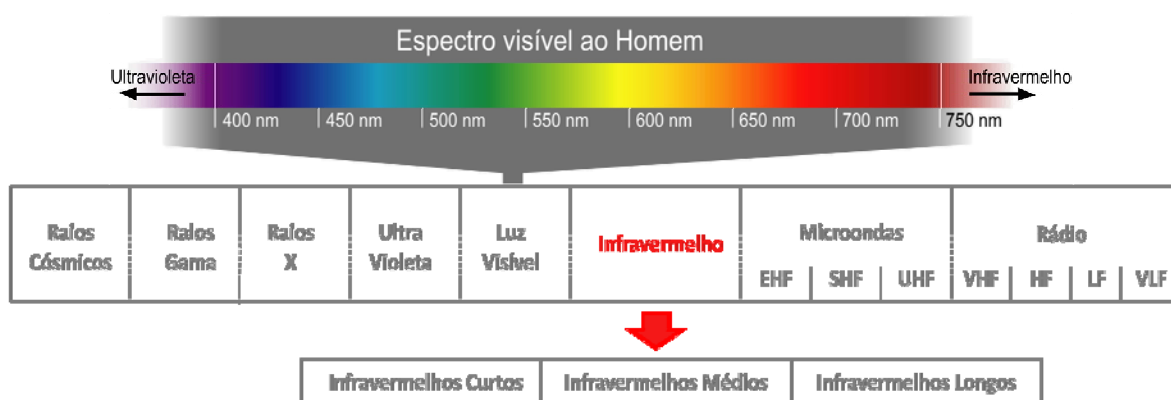


Figura 2.1 - 1 Espectro Electromagnético

Os raios infravermelhos apesar de não serem detectados pela visão humana são sentidos sob a forma de calor. Como tal, podem ser utilizados como uma forma de medir o calor irradiado por um objecto. Este pode ser classificado como sendo um corpo negro ou um corpo real. O primeiro é um objecto capaz de absorver toda a radiação que incide sobre ele em qualquer comprimento de onda enquanto o segundo é um objecto capaz de emitir uma determinada parte da energia. O parâmetro que determina a capacidade de emissão de energia é a emissividade (ϵ). Qualquer objecto (orgânico ou inorgânico) que possua uma temperatura acima do zero absoluto (0° Kelvin, $-273,15^\circ$ C ou -459° F) emite uma certa quantidade de radiação infravermelha em função da sua temperatura. Essa radiação emitida pode incidir sobre a superfície de outro objecto podendo ser reflectida, absorvida ou transmitida que será posteriormente captada pela câmara termográfica, (Figura 2.1 - 2).

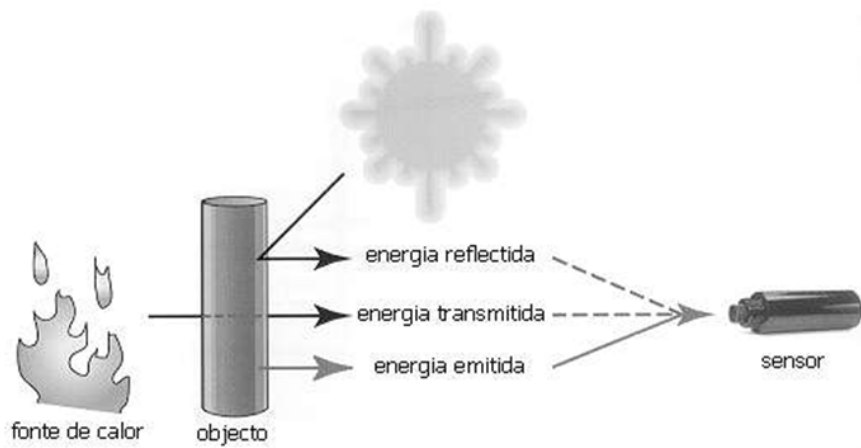


Figura 2.1 - 2 Reflexão, absorção e transmissão de energia

O princípio de funcionamento de uma câmara termográfica baseia-se na lei de Stefan-Boltzmann. A lei enuncia que a energia radiante total emitida por um corpo negro por unidade de superfície é proporcional à quarta potência da temperatura absoluta. Sendo expressa pela seguinte fórmula matemática:

$$W = \sigma\epsilon T^4 \quad (1)$$

Onde:

W, Potência radiante [W/m^2];

σ , constante de Stefan-Boltzmann [$5.7 \times 10^{-8} W/K^4m^2$];

ϵ , emissividade;

T, temperatura absoluta [K].

As imagens resultantes deste tipo de câmaras são denominadas por imagens termográficas e são apresentadas, normalmente, numa escala cinza. No entanto, devido ao facto de o olho humano ser mais sensível à detecção de variações de sombras de cores do que sombras de cinza, estas podem ser convertidas para uma escala de pseudo-cores em que, tipicamente, os tons roxo e azul representam as zonas mais frias e o vermelho e amarelo as regiões mais quentes sendo o branco utilizado para representar a temperatura máxima dessa escala. A Figura 2.1 - 3 apresenta um exemplo da utilização desta gama de cores numa imagem termográfica.

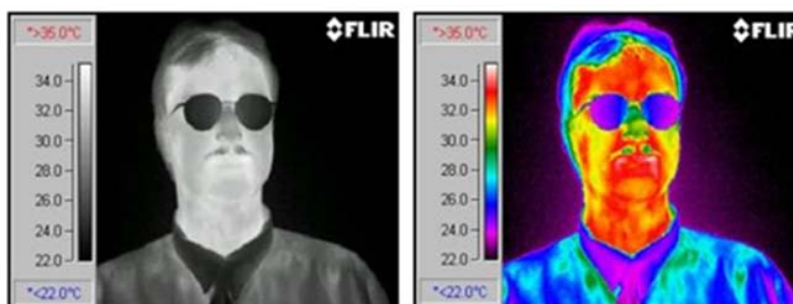


Figura 2.1 - 3 Câmara termográfica: Imagem resultante em escala cinza e pseudo-cores

As câmaras termográficas não requerem, ao contrário das câmaras dia e noite, a presença de uma fonte de luz infravermelha (natural ou artificial). No entanto, pelo facto de o vidro comum bloquear a radiação térmica é necessário recorrer a materiais especiais (*Germanium*) para a produção dos elementos ópticos das câmaras termográficas, tornando-as, deste modo, mais dispendiosas face às câmaras dia e noite.

2.2 CAPTURA DE IMAGENS TERMOGRÁFICAS

A câmara utilizada no trabalho descrito nesta dissertação para a captação das imagens termográficas foi uma FLIR A300. Mais informação sobre as características da câmara adquirida podem ser encontrada em [PAG01]. O *software* que acompanha a câmara permite, entre outras funcionalidades, escolher a paleta de cores atribuída à imagem e optar entre uma escala de temperatura ajustável ou uma escala com valores fixos.

Na Figura 2.2 - 1 encontram-se exemplos das três opções de paletes de cores: “Gray”, “Iron”, “Rainbow”, podendo estas ser invertidas. Esta opção torna-se importante, na medida em que a escolha da paleta de cores representa a escolha das cores que irão estar associadas às altas e baixas temperaturas.



Figura 2.2 - 1 Paleta de cores: Imagem da esquerda em Gray; Imagem ao centro em Iron; Imagem da direita em Rainbow

Após uma análise das três opções, escolheu-se para as imagens termográficas a paleta de cor denominada por “Iron” pelo facto de ser a que melhor permite distinguir as variações de temperatura apresentadas pelo objecto, na medida em que atribui as cores que o ser humano consegue melhor associar às temperaturas como o azul e roxo às temperaturas mais baixas, cores como amarelo, laranja e vermelho às temperaturas mais altas e à temperatura máxima da escala a cor branca.

A segunda opção, a da definição da gama de temperaturas a ser utilizada é de extrema importância, na medida em que numa imagem com uma escala de temperaturas ajustável é possível ver detalhes como a maçaneta de uma porta, como se pode observar na Figura 2.2 - 2.

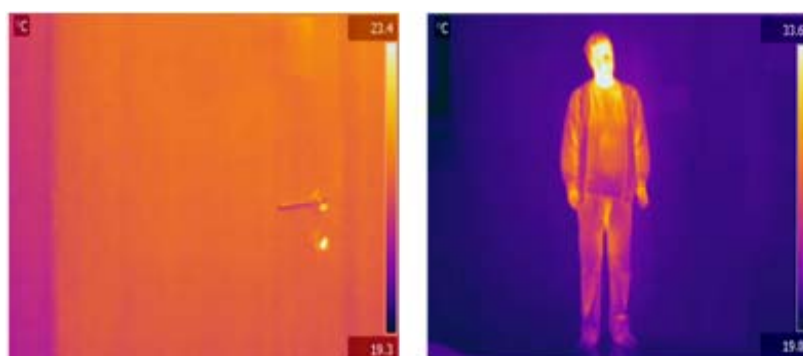


Figura 2.2 - 2 Escala ajustável: Detalhe observado

Contudo, na mesma figura podemos observar que se um objecto com uma temperatura superior aparecer na imagem, a escala ajusta-se para o elemento com a temperatura mais

elevada, não sendo possível observar os detalhes da porta. Ou seja, com a utilização de uma escala ajustável obtêm-se imagens mais descritivas das pessoas, mas é necessário ter em atenção para o caso de aparecerem objectos cuja temperatura é superior à de um humano, a escala de temperaturas é ajustada ao objecto que possui a maior temperatura e as pessoas aparecerão em tons mais escuros e menos detalhados.

Com uma escala fixa que vai dos 20° C até os 40° C, como se encontra ilustrado na Figura 2.2 - 3, é possível contornar a situação, uma vez que se aparecer um objecto cuja temperatura seja superior à da escala, não afectará a descrição da pessoa.



Figura 2.2 - 3 Imagem termográfica com uma escala fixa

3. Áreas de aplicação e técnicas de processamento de imagens termográficas

Neste capítulo encontram-se descritas as soluções para o tratamento das imagens termográficas em áreas como a segurança rodoviária e vigilância com o uso de uma câmara fixa e com uma câmara instalada numa plataforma móvel com o intuito de auxiliar os humanos no terreno.

O tratamento das imagens é um processo que envolve etapas como a detecção e o seguimento de objecto. As técnicas associadas às etapas têm vindo a ser aprofundadas nas duas últimas décadas devido à crescente necessidade em automatizar os sistemas de videovigilância. O leitor poderá encontrar conceitos básicos de detecção e seguimento de pessoas em [FOR02], [GAV98], [MOE00], [SHA00], [YIL06] e uma visão genérica de um sistema de videovigilância em [VEN09].

Face às desvantagens apresentadas aquando da utilização das imagens convencionais (imagem natural e imagem infravermelha) em ambientes com condições adversas, em [KRO08], [LEY08] e [KUM06] é apresentada uma possível solução que consiste na utilização conjunta de imagens naturais e imagens termográficas, de forma a fazer uso da informação extraída de ambas as imagens com o intuito de tornar o algoritmo de detecção e seguimento em cenários de videovigilância mais robusto. Informação relativa ao processo de combinação de dados provenientes dos dois tipos de imagens pode ser encontrada em [KUM10], no qual os autores propõem uma solução genérica.

3.1 SEGURANÇA RODOVIÁRIA

Dentro do tema de segurança rodoviária, as câmaras termográficas podem ser instaladas nos carros, (Figura 3.1 - 1), para auxiliar o condutor na sua condução nocturna, alertando-o para eventuais objectos e pessoas que possam encontrar-se em locais de total escuridão, ou seja, fora do alcance de vista do mesmo. Desta forma, garante-se que o condutor tenha tempo para reagir, evitando, deste modo, acidentes graves. Um exemplo da utilização de câmaras termográficas na área de protecção civil, nomeadamente para a detecção dos peões durante a condução nocturna é encontrado em [OMA10].

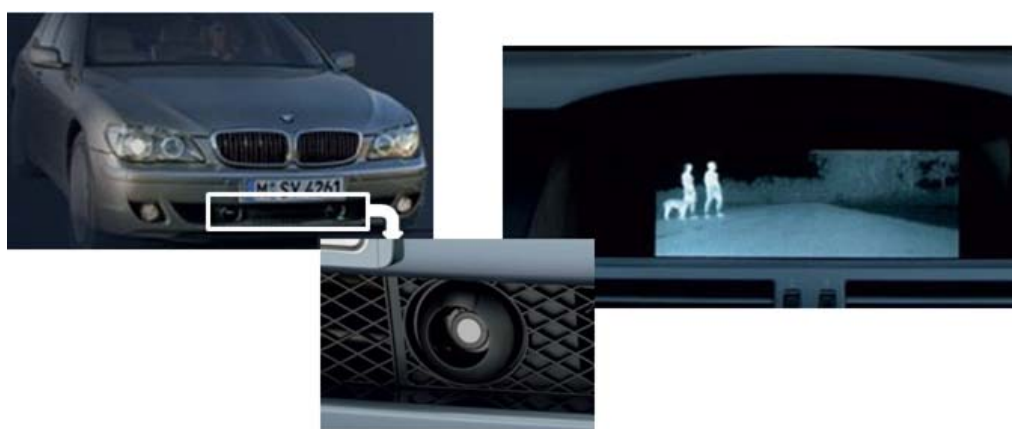


Figura 3.1 - 1 Instalação da câmara termográfica num carro

Um dos pontos focados pelos autores neste artigo é a distorção causada pelo tipo de vestuário utilizados pelos peões em ambientes cuja temperatura ambiente se situa abaixo dos 8° C. O algoritmo proposto pelos autores começa por determinar a região de interesse que definem como sendo o processo de estudo da imagem para destacar determinadas zonas que poderão vir a conter informações sobre o alvo e à qual será realizado um estudo mais pormenorizado, à posteriori.

Durante a noite, a temperatura corporal dos peões captada pela câmara termográfica ronda os 33° C, o decréscimo de 4° C deve-se ao facto de a temperatura ambiental ser relativamente baixa, factor que influencia a captação da radiação térmica de um objecto. Contudo, os peões quando comparados com o meio ambiente, continuam a ser os elementos mais brilhantes nas imagens termográficas. Mas durante épocas de maior frio, as pessoas recorrem a tipos de vestuário que as protejam contra o frio, ou seja, roupas de um tipo de material que isole o calor, fazendo com que o tronco do peão, na imagem capturada, não seja tão brilhante quando comparada com as restantes partes do corpo que se

encontram menos protegidas. Tal facto poderá levar à não identificação do peão como sendo uma zona de interesse e, conseqüentemente, à classificação de como não sendo um peão.

Para compensar esta ligeira perda devido ao tipo de vestuário utilizado pelo peão, os autores realizam uma operação morfológica sobre a imagem na escala cinza de forma a aumentar a intensidade do tom na parte do tronco. As operações morfológicas consistem em sucessivas operações elementares de dilatação e erosão de uma imagem. A aplicação destas operações permitirá atenuar artefactos escuros ou ruídos presentes numa imagem, deixando intactos os restantes detalhes brilhantes. De uma forma mais detalhada, sob a imagem termográfica capturada, os autores criam um gráfico de intensidade em 3D. Ao elevarem a intensidade do tronco, verificou-se, como se pode observar na Figura 3.1 - 2, que no caso de dois peões estarem relativamente próximos um do outro, após este processo, não seria possível fazer a sua separação. Para contornar este problema, os autores resolveram elevar a intensidade segundo uma figura rectangular. Desta forma, a distância entre os peões era assegurada, assim como, os demais detalhes importantes. A única alteração sofrida na imagem original seria a do tom da intensidade do tronco. A forma rectangular possui duas dimensões, uma maior para o caso do peão se encontrar muito perto do carro e uma menor para o caso do peão se encontrar relativamente afastado do carro.

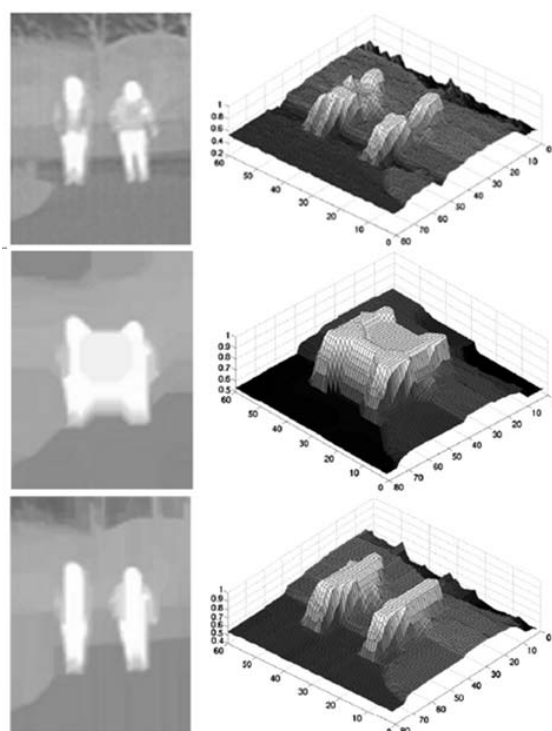


Figura 3.1 - 2 Tratamento de artefactos escuros, [OMA10]

Concluída a etapa de compensação do efeito provocado pela roupa, como se pode observar na imagem anterior, (Figura 3.1 - 2), o peão encontrar-se-á mais destacado face ao *background* e, conseqüentemente, na etapa de segmentação toda a região pertencente ao peão aparecerá unida. A determinação de zonas de interesse é baseada na utilização de um valor fixo elevado de *threshold* para detectar pontos com maior nível de intensidade, que irão corresponder às zonas de temperaturas mais elevadas. As regiões resultantes após a aplicação do valor de *threshold* máximo são classificadas pelos autores como sendo sementes. Na etapa seguinte, os autores aplicam um novo valor de *threshold*. Este valor é menor face ao valor atribuído ao primeiro *threshold*, o que fará com que a semente em cada iteração se una aos pixels à sua volta que cumpram o critério do valor atribuído ao segundo *threshold*. O critério de paragem do crescimento da semente terá como base a análise de dois factores: o modo de crescimento e a área do contorno da figura. Ou seja, o modo de crescimento da figura irá ter em conta se a semente se expandiu mais em termos de altura do que largura, uma vez que os autores assumem que a maioria da população possui mais altura do que largura. À medida que a semente vai aumentando, a sua área é comparada com valores de referência que se encontram numa base de dados. Se o valor da área da semente atingir o intervalo de valores caracterizado como sendo possíveis pessoas, o crescimento da semente cessa.

A etapa de classificação das zonas anteriormente extraídas como sendo regiões de interesse, é considerada como sendo uma etapa de extrema sensibilidade na medida em que caso resultem falsos positivos, a credibilidade do sistema será posta em causa para futuras decisões e no pior dos casos, se resultarem falsos negativos, poderá originar graves acidentes rodoviários. Para esta etapa, é proposta a utilização de histogramas de gradientes orientados (*Histogram of Oriented Gradients* - HOG), em conjunto com um conceito relacionado com a aprendizagem e reconhecimento de padrões utilizado para a classificação de objectos denominado por SVM (*Support Vector Machine*).

HOG é um método utilizado para a descrição de um objecto com base nos seus gradientes. Este divide a imagem em células equivalentes e para cada célula extrai as características com base nos contornos encontrados nessa mesma célula, como ilustra a Figura 3.1 - 3. Para a etapa da classificação, foi criada uma base de dados com 800 imagens, 400 das quais representam peões e as restantes 400 representam outros objectos.

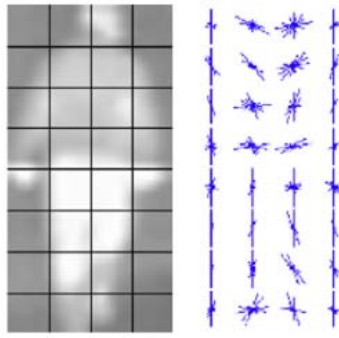


Figura 3.1 - 3 Esquerda: Imagem dividida em células de 4x8; Direita: Resultado do descritor HOG, [OMA10]

As regiões de interesse extraídas foram todas convertidas para o tamanho de 20 x 40. O histograma de gradiente orientado da imagem resulta da combinação dos resultados provenientes de uma convolução entre o gradiente horizontal e o gradiente vertical. De seguida, a imagem é repartida em células ou sub-regiões onde será aplicado um histograma de magnitude a cada uma delas.

Na etapa da classificação, o SVM irá comparar o resultado com os vários modelos que se encontram na base de dados, previamente criada, e determinará se a região corresponde ou não a um peão.

Para o seguimento, os autores aplicam um método matemático baseado em filtros de Kalman, [WEL06]. Com este método é possível estimar a posição de um objecto de uma forma linear num sistema dinâmico que poderá ter medições incompletas ou cobertas de ruído.

De modo a evitar o seguimento de falsos positivos, o seguimento de um objecto só é iniciado após a detecção e confirmação do mesmo em três imagens consecutivas. No caso de o objecto sair da imagem, o seu seguimento é removido após um tempo pré-definido. Desta forma, o sistema é tolerante durante um período de tempo a falhas de detecção originadas por oclusões parciais ou totais.

Em suma, foi apresentado uma proposta para a detecção de peões na área da segurança rodoviária utilizando câmaras termográficas. Contudo, factores como a detecção de pessoas sentadas, crianças, ciclistas, animais ou grupo de pessoas não foram alvo de estudo por parte dos autores.

3.2 VIGILÂNCIA COM RECURSO A UMA CÂMARA FIXA

As imagens termográficas encontram-se, também, aplicadas a cenários de vigilância nocturna ou cenários que possuam condições ambientais adversas. Proposta para realizar a etapa de detecção de pessoas utilizando uma câmara termográfica fixa é encontrada em [FER11].

O algoritmo proposto para a segmentação de pessoas consiste em três etapas: numa primeira é feita a detecção de possíveis formas humanas. Essas formas são melhoradas numa segunda etapa e numa última é obtida a confirmação se essa forma é realmente uma pessoa ou não.

Na primeira etapa, o algoritmo começa por uniformizar todas as imagens capturadas, convertendo-as para a mesma escala de valores em tons de cinza.

A etapa seguinte consiste em eliminar pontos ou zonas incandescentes, isto é, pontos correspondentes a lâmpadas ou outro tipo de fonte de calor. O critério de eliminação destes pontos baseia-se no facto de estes possuírem uma temperatura acima da suportada pelo corpo humano, indicando, deste modo, que não se trata de uma pessoa. Para tal, foi criada uma zona de delimitação ao redor dessa área e esta é preenchida num tom de cinza que o algoritmo irá assumir como sendo um factor não relevante. Finda a etapa de eliminação, é aplicado um *threshold* à imagem de modo a isolar regiões pertencentes a uma figura humana. De seguida, o algoritmo realiza operações morfológicas para eliminar zonas com pixéis isolados e unir áreas que eventualmente foram separadas aquando da aplicação do *threshold*. A validação das zonas será feita caso a área do objecto atinja o valor mínimo que foi atribuído como sendo uma área que possa conter uma pessoa, como se pode observar na Figura 3.2 - 1.



Figura 3.2 - 1 Da esquerda para a direita: Imagem termográfica capturada; Imagem normalizada; Eliminação de zonas incandescentes; Imagem binária, [FER10]

A segunda etapa corresponde ao refinamento ou otimização dos contornos, ou seja, consiste em analisar as formas que previamente foram seleccionadas e inferir se esta corresponde a um ou mais do que um humano. Para tal, a tarefa foi subdividida em duas etapas: delimitação vertical e delimitação horizontal.

Na primeira sub-etapa é feita a delimitação vertical utilizando uma projecção vertical da região seleccionada. Ou seja, percorre-se a imagem por colunas e vai-se adicionando o valor de cinza correspondente a cada pixel dessa coluna, obtendo um histograma que irá indicar as zonas da imagem que possuem uma maior concentração de calor. Através da análise da projecção vertical é possível separar os vários elementos que se encontram num grupo em sub-imagens com apenas um indivíduo, como se pode observar na Figura 3.2 - 2.

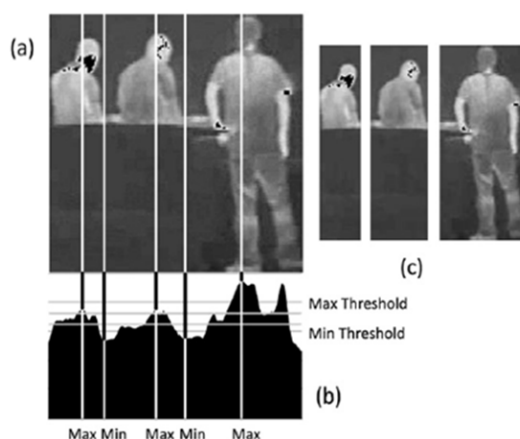


Figura 3.2 - 2 a) Região de interesse seleccionada b) Histograma aplicado c) Resultado da inferência, [FER10]

A sub-etapa seguinte é realizada sobre cada uma das sub-regiões obtidas anteriormente. Para tal, aplica-se um *threshold* e ajusta-se o limite superior e inferior da sub-região de acordo com a informação extraída com a aplicação do *threshold*. Desta sub-etapa pode concluir-se se a pessoa se encontra numa situação de ocultação, como se pode constatar em dois casos na Figura 3.2 - 3.

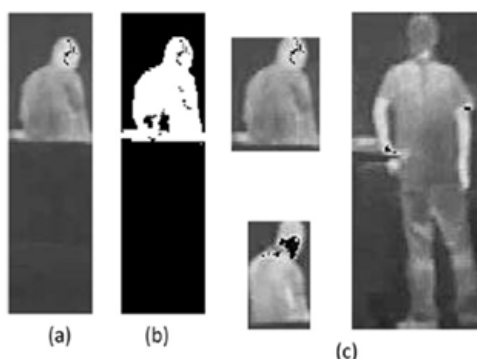


Figura 3.2 - 3 a) Sub-regiões obtidas na sub-etapa de delimitação vertical; b) Imagem resultante após aplicação do *threshold*; c) Resultados obtidos, [FER10]

Dos vários testes realizados, os autores concluem que o algoritmo consegue de uma forma eficiente detectar pessoas que estão muito próximas umas das outras e/ou em situações de ocultação parcial. Da mesma forma, o algoritmo demonstra-se eficaz ao não apresentar falsos positivos na presença de objectos que possuam temperatura o suficiente para serem classificados como humanos. Contudo, situações de ocultação total ou quando duas pessoas estão quase sobrepostas, não foram tido em conta neste estudo.

3.3 VIGILÂNCIA COM RECURSO A UMA CÂMARA MÓVEL

A instalação de câmaras termográficas em plataformas móveis permite auxiliar os operacionais no terreno, na medida que será possível fazer rondas para a vigilância do local e detectar intrusos. Podendo ser um acréscimo à segurança do ser humano que actue em ambientes perigosos como incêndios. Soluções para a detecção e seguimento de pessoas são apresentadas em [FER10] e [TRE06].

Em [FER10] é descrita uma proposta para a detecção de pessoas através de imagens capturadas por uma câmara termográfica que se encontram instalada numa plataforma móvel.

Este robô, denominado por mSecurityTM, é uma plataforma móvel especialmente concebida para tarefas de videovigilância. Nele encontram-se instaladas duas câmaras: uma que captura imagens naturais e outra que captura imagens termográficas. Possui, também, um módulo que permite ao utilizador controlar determinados parâmetros do robô, tais como: o seu estado actual, níveis de bateria, qualidade do sinal Wi-Fi (*Wireless-Fidelity*), entre outros. Possui ainda sensores de ultra-som para evitar colisões com objectos. O caminho a ser patrulhado é inicialmente explorado pelo robô, que os vai marcando como pontos conhecidos. A opção de traçar o caminho inicial de patrulha do robô está também disponível. No caso da detecção de um intruso, o robô envia um sinal de alarme ao gestor do sistema, pára todas as restantes funções e fica no modo de gravação de imagens.

O algoritmo proposto começa por uniformizar todas as imagens capturadas, convertendo-as para a mesma escala de valores em tons de cinza. De seguida é feita a eliminação de pontos incandescentes e as restantes zonas deverão cumprir um valor de área mínimo para que possam ser considerados como regiões de interesse, da mesma forma que foi apresentado, pelo mesmo autor, no artigo anterior [FER10].

Para a detecção de pessoas, foi criado um sistema inteligente que adapta a forma como o sistema realiza a detecção de pessoas mediante a mobilidade da plataforma. Se o robô estiver parado o método utilizado será o denominado por subtração de imagens. O método de subtração de imagens consiste em utilizar as imagens capturadas nos instantes t e $t-1$ e realizar a subtração entre as duas para obter o trajecto feito pelo indivíduo. A imagem resultante indicará a posição actual da pessoa e a anterior. De modo a eliminar as posições anteriores, aplica-se um determinado *threshold* à imagem, (Figura 3.3 - 1).



Figura 3.3 - 1 Da esquerda para a direita: Imagem no instante $t-1$; Imagem no instante t ; Posição actual e posição anterior; aplicação do *threshold* para eliminar a posição fantasma, [FER10A]

No caso de o robô estar em movimento, o método aplicado é intitulado de *optical flow* de Lucas-Kanade, [KAN81]. Este método é utilizado como sendo um intermediário para obter uma representação aproximada do movimento dos objectos, uma correspondência entre as características das imagens, correlações ou propriedades da intensidade das estruturas, permitindo deste modo separar o *foreground* do *background*.

Nas Figura 3.3 - 2 e Figura 3.3 - 3 encontram-se os resultados obtidos com o uso do método de subtração de imagens e com o método de *optical flow* de Lucas-Kanade, respectivamente.



Figura 3.3 - 2 Resultados obtidos com um método de subtração de imagens

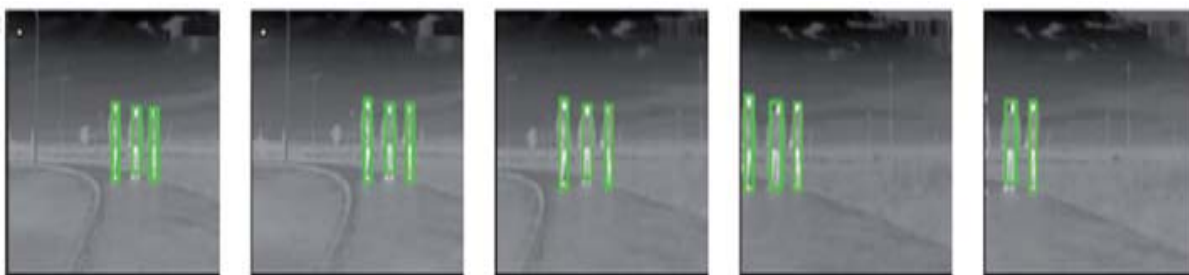


Figura 3.3 - 3 Resultados obtidos com um método de *optical flow* de Lucas-Kanade

Outra solução para a detecção de pessoas em tempo real num cenário de vigilância utilizando imagens termográficas pode ser encontrada em [TRE06]. Tal como no artigo anterior, a câmara encontra-se instalada num robô que terá como função identificar pessoas enquanto patrulha um edifício. O robô terá três tipos de comportamento: estático, a patrulhar pelo corredor ou a seguir uma pessoa. Neste tipo de cenário o robô deverá ser capaz de identificar pessoas mesmo a longas distâncias e não assumir que as pessoas estão sempre numa posição de frente para o robô.

Para a etapa de detecção das pessoas, os autores criaram dois modelos, denominando-os por *Elliptic contour model* e *Feature model*. O primeiro modelo de medida apresentado é utilizado para estimar a posição da pessoa, utilizando duas metades de elipses: uma mais pequena que servirá para indicar a posição da cabeça da pessoa e uma segunda maior para estimar a posição do tronco. A implementação desta medida permitirá detectar pessoas mesmo quando estas não se encontrem viradas de frente para o robô, como se pode observar na Figura 3.3 - 4.



Figura 3.3 - 4 *Elliptic contour model* e divisão do contorno elíptico, [TRE06]

Na segunda medida - *Feature model* - os autores baseiam-se nas características dos valores de cinza, utilizando um algoritmo de aprendizagem proposto por Viola & Jones, [VIO01]. Este modelo é utilizado para melhorar o desempenho do *Elliptic contour model*. Devido à sua

rapidez de cálculo, o *Feature model* é aplicado à imagem para localizar possíveis regiões que possuam características dos valores de cinza pertencentes a uma pessoa, posteriormente é aplicado o *Elliptic contour model* para confirmar se é ou não uma pessoa.

Ambos os modelos estão integrados no algoritmo de seguimento. Para esta etapa, os autores recorrem a uma das variantes dos *Bayesian Filters*, intitulado por *Particle Filter*. Os *Bayesian Filters* recorrem a estatísticas para estimar o comportamento de um sistema dinâmico, que poderá ser uma pessoa ou um objecto, através da observação do ruído desse sistema. Para a estimação do estado futuro do sistema, toda a informação deve ser somente retirada da posição actual e não de posições passadas e da observação feita do sistema. O *particle filter* estima a posição seguinte do sistema com base na filtragem da distribuição das partículas. Uma explicação mais pormenorizada com exemplos ilustrativos do *Bayesian Filters* e as suas variantes pode ser encontrada em **[FOX03]**.

4. Caracterização de cenários de teste e impacto da variação luminosa

Neste capítulo serão apresentados os vários cenários em que foram realizadas as capturas das imagens termográficas. De forma a obter imagens com características e elementos distintos, efectuaram-se capturas numa garagem e em dois corredores com características diferentes. Apresenta-se também um estudo do impacto da variação luminosa que permite demonstrar os benefícios do uso das imagens termográficas face às três variantes de imagens que podem ser utilizadas em sistemas de vigilância: imagem natural, imagem infravermelha com e sem luz infravermelha artificial.

4.1 DESCRIÇÃO E ANÁLISE DE CENÁRIOS

Devido ao facto de não existirem sequências típicas para as imagens termográficas, foi necessário proceder à aquisição de conteúdos que retratem possíveis casos de videovigilância. A título de exemplo, num cenário de videovigilância numa garagem teremos como elementos carros, pessoas a movimentarem-se, fontes de iluminação, postes e portas; num cenário interior teremos um corredor, uma a várias pessoas a movimentarem-se, portas e os revestimentos desse local.

Para a captura das sequências foi necessário retratar acontecimentos que ilustrem casos do quotidiano, que se descrevem a seguir e se encontram ilustrados na Figura 4.1 - 1.

- a) Numa primeira experiência, a cena incluiu uma pessoa a fazer um trajecto e a inverter a marcha fazendo depois o mesmo percurso no sentido inverso. Nesta cena foi também incluído um terceiro elemento (carro).
- b) Esta experiência foi realizada com três pessoas que simularam casos de ocultação parcial por se encontrar a caminhar relativamente próximas umas das outras e um terceiro elemento, o carro. O trajecto a ser realizado pelo grupo de pessoas será o mesmo que na experiência anterior.
- c) A terceira experiência consistiu em duas pessoas que inicialmente caminhavam em sentidos opostos e parando quando se cruzassem. Este momento de pausa representará a situação de ocultação parcial a total de um objecto. Posteriormente, as pessoas regressaram à sua posição inicial.
- d) Na última experiência, à semelhança da cena anterior, duas pessoas caminhavam em sentidos opostos e pararam quando se cruzaram. Após o momento de ocultação, as pessoas seguiram em frente.

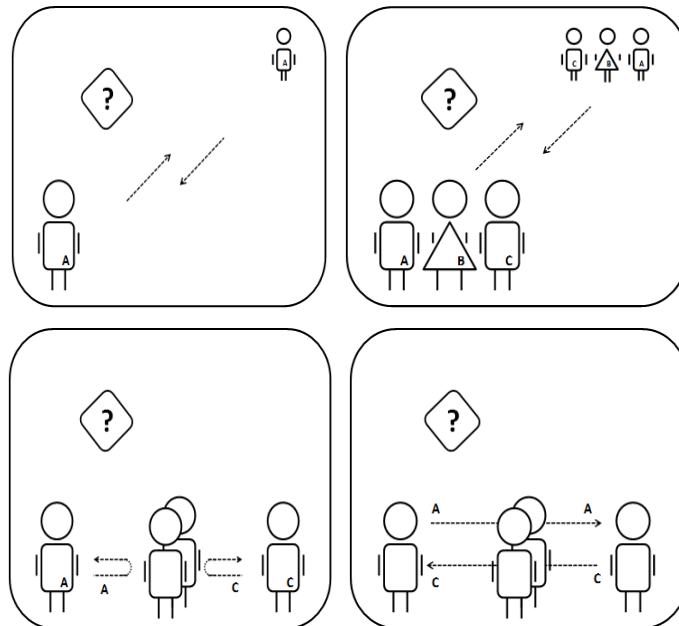


Figura 4.1 - 1 Recriação de acontecimentos

Devido ao facto de na captura das imagens com uma escala ajustável poder ocorrer o aparecimento de um objecto que possua uma temperatura muito superior à de um corpo humano, o que fará com que não seja possível separar a pessoa dos detalhes pertencentes

ao *background*, para cada cenário foram gravadas diversas sequências utilizando duas escalas de temperatura: escala ajustável e com uma escala fixa. O valor mínimo e máximo escolhido para a representação da escala fixa foi 20° C e 40°C, respectivamente. Com este intervalo, a descrição de objectos que possuam uma temperatura ligeiramente inferior à de um humano e que sejam relevantes para a segmentação é assegurada. Assim como a descrição da pessoa dado a temperatura de um humano rondar os 37° C e não ultrapassar os 40° C, em casos normais.

Os cenários escolhidos para a gravação das sequências incluem três ambientes interiores: um corredor revestido só por paredes e acesso a outras salas; um outro corredor que dá acesso a outras salas e que difere do primeiro cenário por possuir zonas com vidros; e uma garagem, como se pode observar na Figura 4.1 - 2.

Imagens infravermelhas



Imagens termográficas com a escala de temperatura ajustável



Imagens termográficas com a escala de temperatura fixa: 20° C - 40° C



Figura 4.1 - 2 Esquerda: corredor comum; Ao centro: corredor com zonas revestidas de vidro; Esquerda: garagem

Após a análise das imagens captadas verificou-se a existência de diversos tipos de reflexões. Nos cenários correspondentes aos corredores, esse efeito foi sentido nas zonas que possuem vidros, uma vez que o vidro comum bloqueia as radiações térmicas, como se pode observar na figura anterior. Ainda no ambiente relacionado com os corredores, verificou-se o efeito da reflexão da temperatura emitida pelo objecto no chão. No caso de o cenário ser a garagem, apesar de ocorrer o efeito de reflexão, esta não é tão nítida como nos cenários dos corredores. Tal deve-se ao facto de o material utilizado no pavimento da garagem diferir do material utilizado nos corredores, sendo o chão da garagem de cimento e o dos corredores de um material brilhante. Como tal, dado as sequências captadas nos corredores não serem as melhores para efeitos de estudo devido à reflexão, as sequências que irão ser utilizadas para fins de estudo, nomeadamente, para a segmentação e aplicação dos modelos de descrição serão as imagens captadas no cenário da garagem com a escala ajustável e a escala fixa entre os 20° C e os 40° C.

4.2 ANÁLISE DO EFEITO DE VARIAÇÕES ABRUPTAS DE ILUMINAÇÃO

Nesta fase pretendeu-se demonstrar o impacto da variação da luminosidade presente no local nas imagens convencionais (imagem natural e imagem infravermelha) utilizadas nos sistemas de vigilância, assim como as vantagens apresentadas pelas imagens termográficas neste tipo de situações, na medida em que não são influenciáveis por este factor. Para este estudo, comparam-se as imagens termográficas com imagens naturais e imagens infravermelhas com e sem o uso de uma fonte luminosa infravermelha.

A Figura 4.2 - 1 apresenta diferentes imagens captadas em diferentes momentos e que ilustram a transição entre uma situação de luz acesa e luz apagada. Como se pode verificar, nas imagens naturais e infravermelha, quando a luz se encontra acesa e posteriormente é apagada, há um momento de ajuste antes de a cena ficar totalmente escura. No caso da imagem infravermelha com luz artificial, a câmara é capaz de recuperar a imagem após um momento de ajuste em que a cena fica preta. No momento em que se volta a acender a luz, constata-se que há um clarão nas imagens naturais e nas imagens infravermelhas até à estabilização da imagem. No entanto, com as imagens termográficas, verifica-se que estas não sofrem de qualquer influência com a variação luminosa do local.

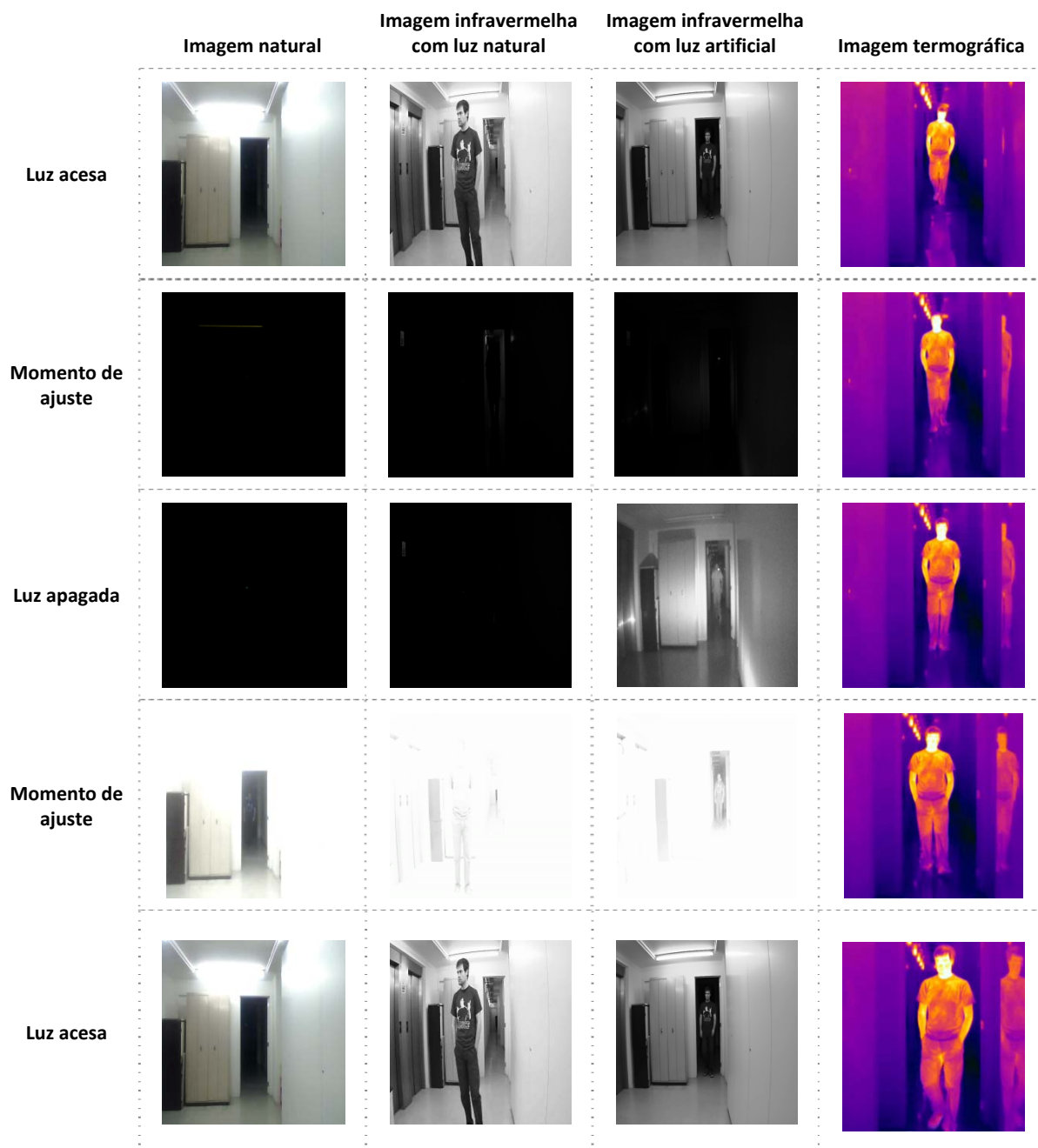


Figura 4.2 - 1 Comparação entre as imagens da câmara dia e noite e termográfica em situações de variação da luminosidade

Para uma análise mais qualitativa dos eventos, apresenta-se de seguida a comparação entre as quatro variantes de imagens em quatro eventos que se encontram resumidos na Tabela 4.2 - 1. O ambiente em que as imagens foram captadas é constituído por dois corredores que foram denominados por zona 1 (Z1) e zona 2 (Z2). Para complementar a compreensão do cenário, a Figura 4.2 - 2 ilustra a posição das câmaras e a restante planta do local.

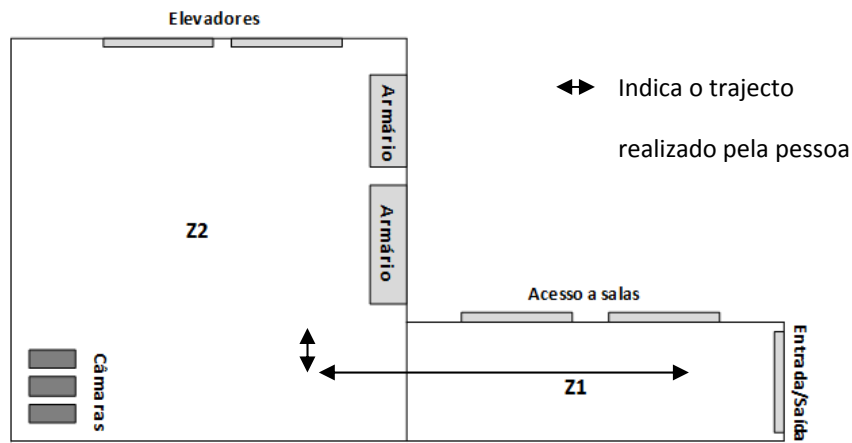


Figura 4.2 - 2 Planta do local da gravação

Inicialmente todas as luzes da zona 1 e da zona 2 se encontram ligadas. O evento 1 é caracterizado pelo momento em que só as luzes da zona 2 são desligadas. No evento 2, o estado da luz inicial das duas zonas é ligada e, posteriormente, a luz da zona 1 é apagada. No evento 3, a luz da zona 1 mantém-se apagada do início até ao fim e a luz da zona 2 que se encontrava acesa é apagada. Para o evento 4, todas as luzes da zona 1 e da zona 2 são apagadas e de seguida ligam-se as luzes.

Tabela 4.2 - 1 Tabela de eventos

	Estado inicial	Estado Final	Estado inicial	Estado Final
	Zona 1	Zona 1	Zona2	Zona2
Evento 1	Luz Acesa	Luz Acesa	Luz Acesa	Luz Apagada
Evento 2	Luz Acesa	Luz Apagada	Luz Acesa	Luz Acesa
Evento 3	Luz Apagada	Luz Apagada	Luz Acesa	Luz Apagada
Evento 4	Luz Apagada	Luz Acesa	Luz Apagada	Luz Acesa

■ Sinaliza uma alteração de estado

De seguida serão apresentados quatro gráficos referentes aos quatro tipos de imagens: imagem natural, imagem infravermelha com e sem luz infravermelha artificial e imagem termográfica. Nos gráficos encontra-se indicado o erro associado à segmentação dos objectos devido à ocorrência dos eventos anteriormente apresentados. O algoritmo utilizado para realizar a segmentação, assim como o algoritmo utilizado para a avaliação do algoritmo de segmentação são referenciados no capítulo seguinte.

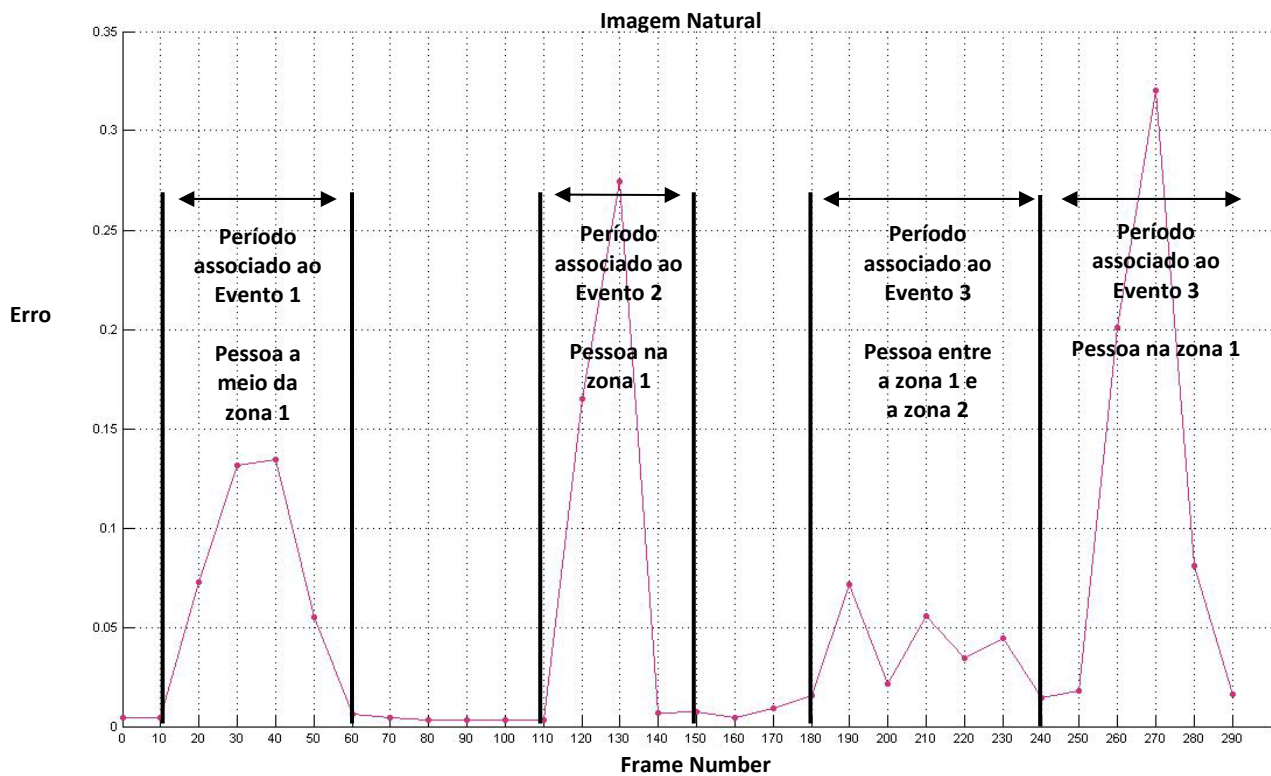


Gráfico 4.2 - 1 Variação da luminosidade com imagens naturais

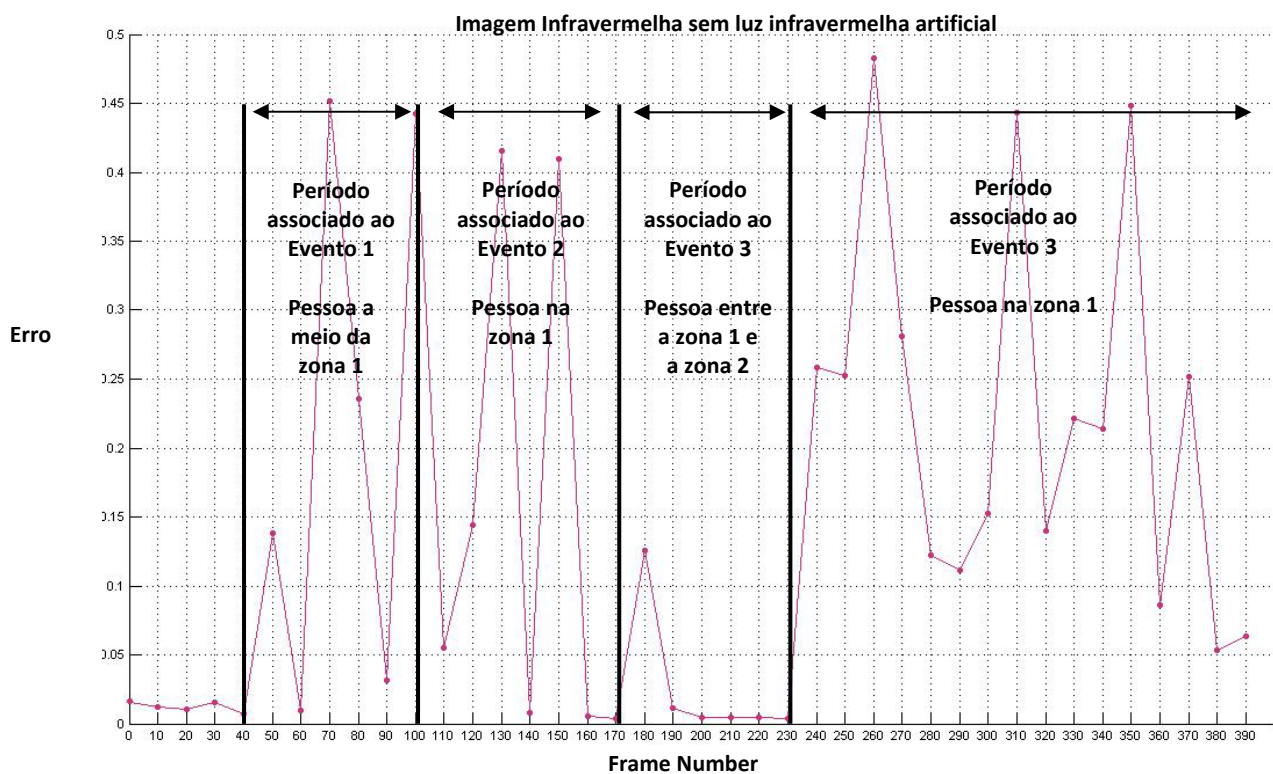


Gráfico 4.2 - 2 Variação da luminosidade com imagens infravermelhas sem luz infravermelha artificial

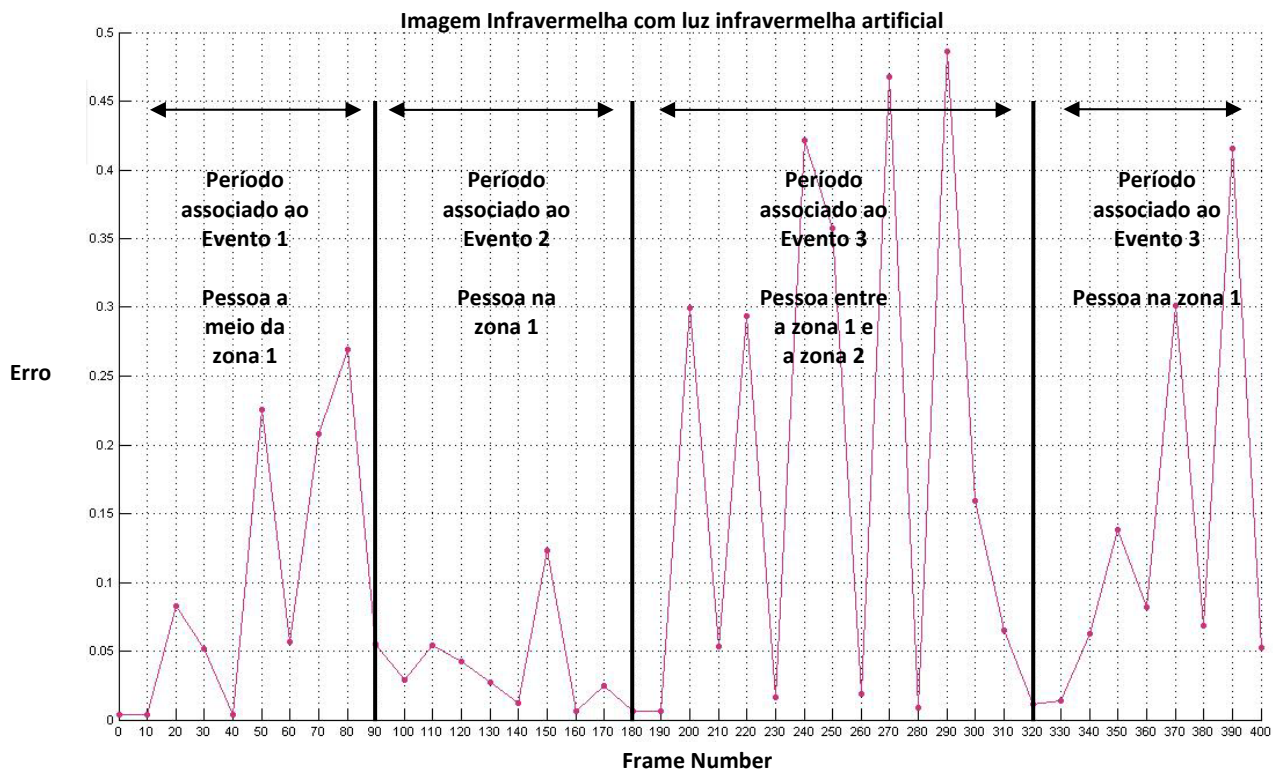


Gráfico 4.2 - 3 Variação da luminosidade com imagens infravermelhas com luz infravermelha artificial

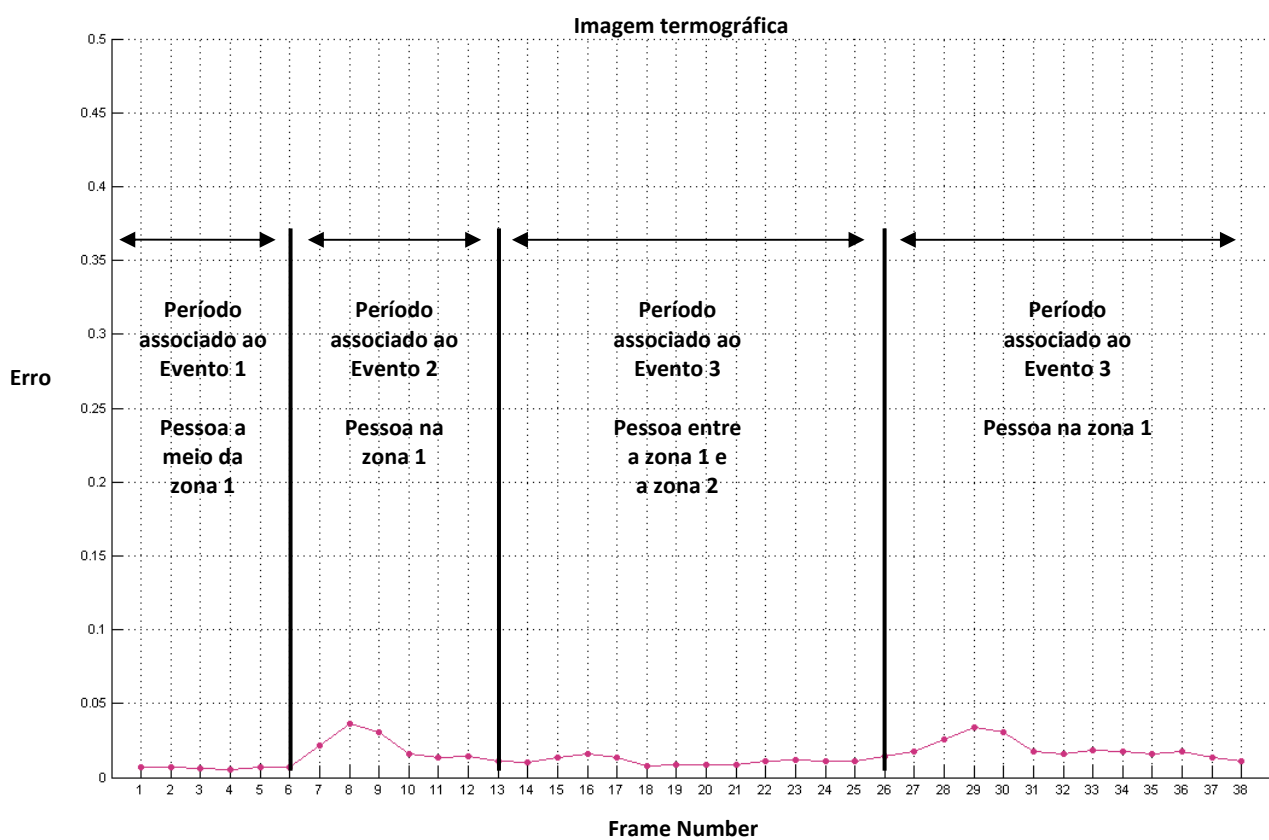


Gráfico 4.2 - 4 Variação da luminosidade com imagens termográficas

Tabela 4.2 - 2 Dados extraídos dos gráficos

Tipo de imagem	Período	Descrição do evento	Frame Início	Frame Final	Nº Total de Frames	Escala de erro
Natural	Período 1	Evento 1: Apaga luz da zona2	10	60	50	[0; 0.35]
	Período 2	Evento 2: Acende luz da zona 2	110	150	40	
	Período 3	Evento 3: Apaga tudo	180	240	60	
	Período 4	Evento 4: Acende tudo	250	290	40	
Infravermelha sem projector	Período 1	Evento 1: Apaga luz da zona2	40	100	60	[0; 0.5]
	Período 2	Evento 2: Acende luz da zona 2	100	170	70	
	Período 3	Evento 3: Apaga tudo	170	230	60	
	Período 4	Evento 4: Acende tudo	230	310	80	
Infravermelha com projector	Período 1	Evento 1: Apaga luz da zona2	0	90	90	[0; 0.5]
	Período 2	Evento 2: Acende luz da zona 2	90	180	90	
	Período 3	Evento 3: Apaga tudo	180	320	140	
	Período 4	Evento 4: Acende tudo	320	400	80	
Termográfica	Período 1	Evento 1: Apaga luz da zona2	1	6	5	[0; 0.5]
	Período 2	Evento 2: Acende luz da zona 2	6	13	7	
	Período 3	Evento 3: Apaga tudo	13	26	13	
	Período 4	Evento 4: Acende tudo	27	38	11	

Na Tabela 4.2 - 2, são apresentados os dados mais relevantes extraídos dos quatro gráficos. Da análise dos dados e dos gráficos podemos constatar que as imagens naturais assim como as imagens infravermelhas com e sem luz infravermelha artificial são afectadas pela variação da luminosidade no local. Dos dados tabelados verifica-se que quando há uma variação da condição luminosa no sentido de a luz estar acesa e posteriormente ser totalmente apagada, o momento de estabilização é maior com a imagem infravermelha com o projector do que nas outras duas (natural e infravermelha sem o projector), tal deve-se ao facto de esta necessitar de se ajustar até apresentar a imagem final. O mesmo acontece quando a luz se encontra apagada e é acesa, a estabilização após o clarão que é captado, é maior com a imagem infravermelha com o projector do que na imagem natural e infravermelha sem o projector.

Apesar de à primeira vista no gráfico que representa as imagens termográficas os picos encontrados nos gráficos poderem estar associados aos eventos, tal não é verdade porque os erros relatados devem-se ao facto de as imagens captadas com a câmara termográfica

possuírem um *frame rate* não constante. A Figura 4.2 - 3 apresenta o exemplo de um erro que ocorreu entre a imagem 6 à imagem 10 e que poderá estar associado ao evento, mas que na realidade é uma falha no algoritmo de segmentação, na medida em que o algoritmo de segmentação compara os histogramas de cor do objecto na imagem actual e a seguinte e que devido ao salto que ocorre, é esperado que haja um erro associado na comparação dos histogramas para inferir se é ou não o mesmo objecto.

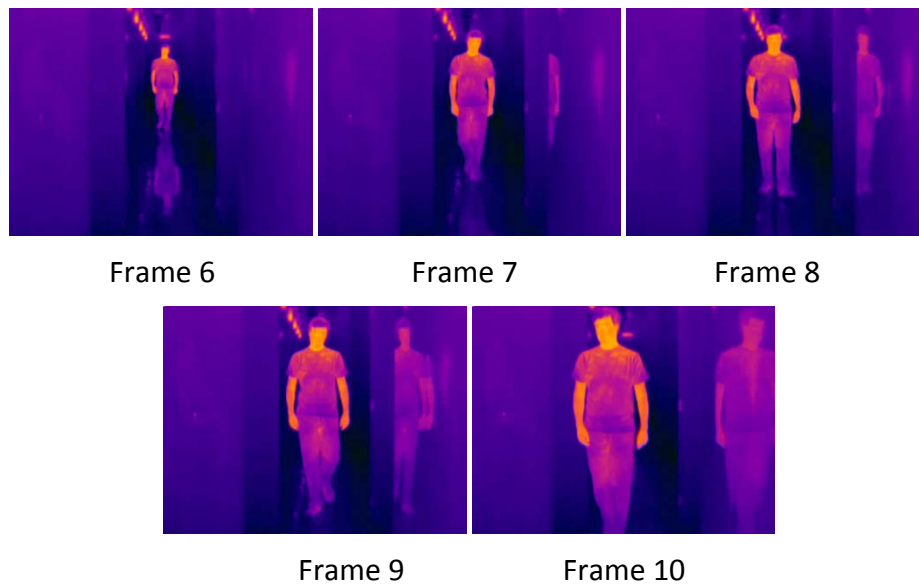


Figura 4.2 - 3 Ilustração do porquê da ocorrência do erro

5. Segmentação e detecção de objectos

Neste capítulo apresenta-se os resultados provenientes do algoritmo desenvolvido para a etapa da segmentação baseado num valor de *threshold*. Dos resultados obtidos, é realizada uma comparação com os resultados provenientes de dois algoritmos de segmentação que são baseados no movimento do objecto, de forma a avaliar qual dos algoritmos permite obter uma melhor segmentação.

5.1 SEGMENTAÇÃO BASEADA EM THRESHOLD

Na Figura 5.1 - 1 apresenta-se o fluxograma que representa as várias etapas do algoritmo desenvolvido para a etapa da segmentação de pessoas baseado num valor de *threshold*. As funções utilizadas para o processamento de imagem provêm da biblioteca OpenCv [INT99].

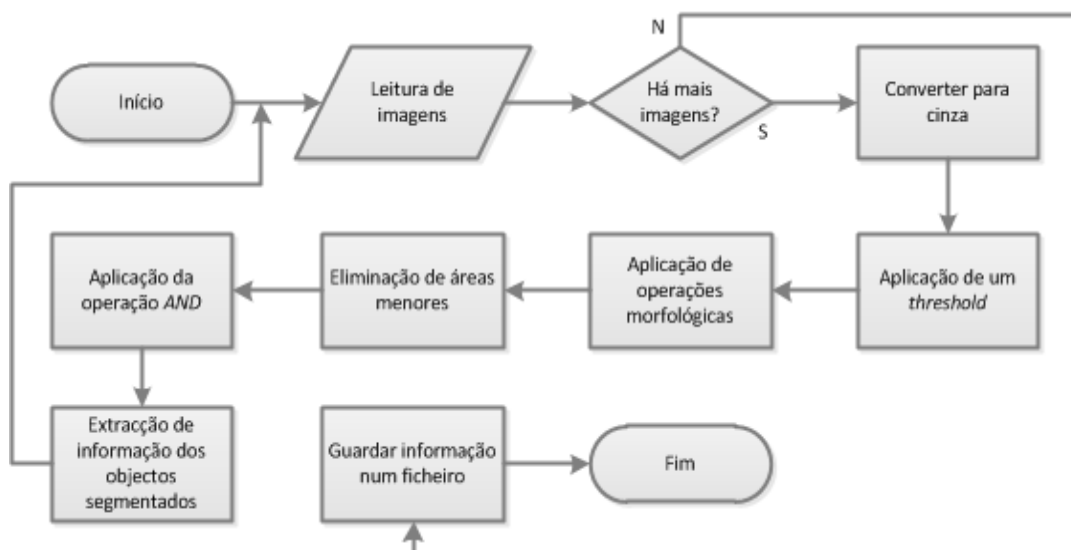


Figura 5.1 - 1 Fluxograma do código de segmentação

O algoritmo começa por fazer uma leitura das imagens de uma sequência, convertendo-as para uma escala cinza. De seguida é aplicado um valor de *threshold*. Os valores de *threshold* utilizados resultam de uma análise empírica na qual se constatou que com o uso dos valores apresentados resultavam em melhores resultados para a etapa de segmentação. No caso de ter sido utilizado a escala ajustável, o valor foi de 60. Para as imagens gravadas com a escala fixa entre os 20° C a 40° C, o valor de *threshold* aplicado foi de 40. Esta diferença no valor aplicado ao *threshold* deve-se ao facto de que com uma escala ajustável, a diferença entre o valor mínimo e o valor máximo atribuído para a escala de temperatura ser menor do que a diferença apresentada com uma escala fixa. Tal facto leva a que com uma escala ajustável haja um maior destaque da pessoa face aos restantes elementos, o que permitirá aplicar um valor de *threshold* maior. Desta forma, é possível seleccionar os elementos de interesse sem que a imagem possua muito ruído térmico proveniente de objectos não relevantes. Com uma escala fixa, apesar de garantirmos que na presença de um elemento com uma temperatura superior à apresentada por um ser humano, a descrição da pessoa é assegurada, o que não acontece com uma escala ajustável, pois a pessoa não se encontra tão destacada face aos outros elementos da imagem. Como tal, o valor aplicado ao *threshold* terá que ser menor. E, conseqüentemente, a imagem resultante desta etapa irá possuir mais ruído face à imagem resultante captada com uma escala ajustável, como se pode observar na Figura 5.1 - 2.

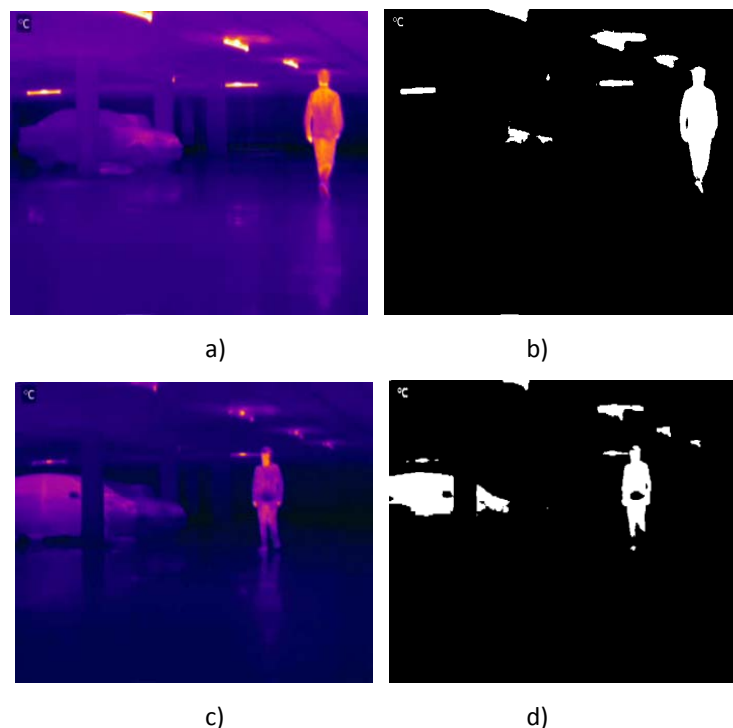


Figura 5.1 - 2 a) Imagem termográfica com escala ajustável; b) Aplicação do *threshold* = 60; c) Imagem termográfica com escala fixa; d) Aplicação do *threshold* = 40

À imagem resultante da aplicação do *threshold* são aplicadas operações morfológicas para eliminar ruído e unir ou separar pixéis que na passagem para uma imagem binária tenham sido separados ou unidos indevidamente. As operações morfológicas consistem em sucessivas operações elementares de dilatação e erosão de uma imagem aplicadas a um rectângulo de 3x3, por omissão. A aplicação destas operações permitirá atenuar artefactos escuros ou ruídos presentes numa imagem, deixando intactos os restantes detalhes brilhantes. Com esta etapa foi possível eliminar o símbolo de indicação de temperatura (°C) que estava presente e que não era informação relevante. Na Figura 5.1 - 3, podemos constatar que para a imagem com uma escala ajustável utilizada para o exemplo, não se verifica alterações significativas. Mas com o exemplo da imagem com uma escala fixa, é possível verificar que com a aplicação das operações morfológicas, foi possível proceder a uma melhor separação entre a lâmpada e a pessoa.

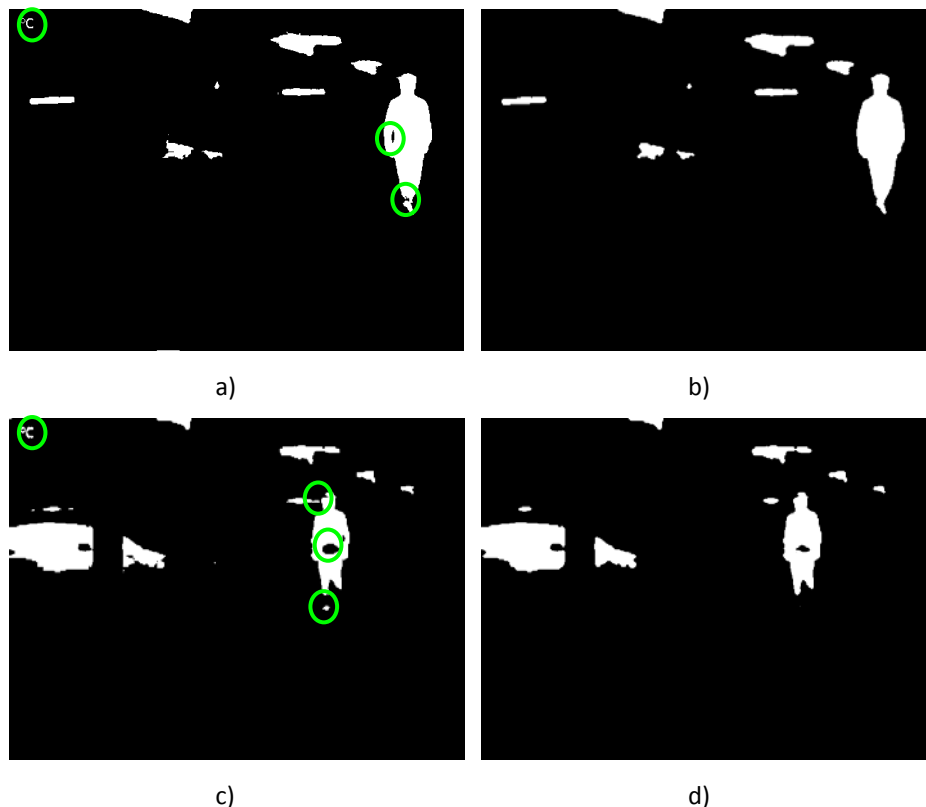


Figura 5.1 - 3 a) *Threshold* da imagem com escala ajustável; b) Operações morfológicas na imagem com escala ajustável; c) *Threshold* da imagem com escala fixa; d) Operações morfológicas na imagem com escala fixa;

As imagens resultantes das duas etapas anteriores ainda apresentam um certo ruído térmico proveniente de objectos não relevantes como as lâmpadas, que foram contabilizadas por apresentarem uma temperatura semelhante ou superior à de um humano ou a presença de partes de um carro na imagem captada com a uma escala fixa pelo facto da sua temperatura

à superfície ser superior ao valor mínimo estabelecido (20° C). Devido a estes factos, foi necessário implementar uma etapa que permitisse proceder à remoção do ruído térmico. Para tal, utilizou-se uma função da biblioteca OpenCv, denominada por `floodfill`. Esta função analisa toda a imagem e permite extrair informação dos objectos, nomeadamente, a área, o seu contorno e a localização do objecto na imagem. A localização é expressa em formato de uma *bounding box*, ou seja, é constituída um ponto inicial - $P(x,y)$ - e por duas variáveis: altura e largura. Através da informação da área e da localização foi possível introduzir um critério de selecção dos objectos. O critério consiste em analisar a área e a localização, no caso de o objecto possuir mais largura do que altura e o valor da área for menor que o dado como referência, então o objecto será removido da imagem. A remoção é feita recorrendo ao uso da função `floodfill`, através da qual os pixéis na imagem associados a um objecto são alterados para a cor preta. Na Figura 5.1 - 4 são apresentados os resultados obtidos nesta etapa.

Imagem termográfica com uma escala ajustável



Imagem termográfica com uma escala fixa

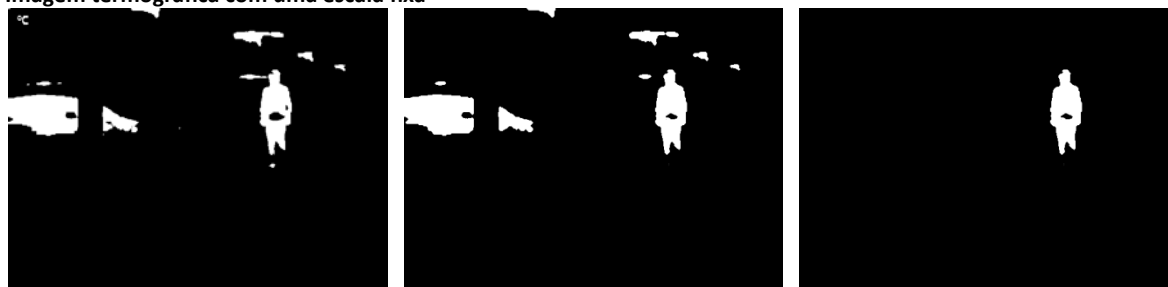


Figura 5.1 - 4 Da coluna esquerda para a direita: Resultado da aplicação do *threshold*; Resultado da aplicação das operações morfológicas; Resultado obtido com os critérios de selecção

À imagem resultante é extraída a informação do objecto e gravada num ficheiro. Esta servirá como máscara na operação lógica `AND` com a imagem original, como se pode observar na Figura 5.1 - 5. O resultado desta etapa é gravado e utilizado em conjunto com o ficheiro na etapa de comparação dos descritores, que será abordado no capítulo seguinte, terminando o algoritmo.



Figura 5.1 - 5 Da esquerda para a direita: Imagem original; Máscara; Resultado obtido com a operação AND

5.2 AVALIAÇÃO DOS ALGORITMOS DE SEGMENTAÇÃO

Terminada a etapa da segmentação, procedeu-se à comparação dos resultados obtidos com o algoritmo baseado no *threshold*, com os resultados provenientes de dois algoritmos de segmentação já implementados. Estes dois algoritmos de segmentação diferem do algoritmo de segmentação desenvolvido por se basearem no movimento e não num valor de *threshold*. O objectivo desta análise consistiu em verificar se a utilização de um valor de *threshold* para a etapa de segmentação traz ou não benefícios face aos métodos convencionais para realizar a segmentação que são baseados no movimento. Informação acerca destes algoritmos pode ser encontrada em [TEI07] e em [INT99], que ao longo do documento irá ser referido como ISegmentation e OpenCv, respectivamente. A diferença entre os dois algoritmos de segmentação com base no movimento reside no facto de o algoritmo ISegmentation possuir mais funções que permitirá eliminar mais ruído que sejam contemplados nas imagens. O algoritmo baseado em *threshold* será referido como Threshold ao longo do documento.

A *framework* de avaliação que permitirá avaliar os três algoritmos de segmentação com base em imagens de referência, encontra-se descrita em [CAR09]. Esta recorre ao uso das métricas do *partition-distance* para avaliar os algoritmos de segmentação e seguimento. No processo principal da métrica encontra-se o gráfico de intersecção entre duas segmentações, que pode ser definido como um gráfico sub-dividido com um nó para cada região da segmentação. Dois nós são ligados por uma borda não-direccionada e pesada se e só se as duas regiões se intersectarem mutuamente. O gráfico de intersecção associado às imagens segmentadas será utilizado como uma referência de índices de similaridade entre as partições. O *partition-distance* aborda o cálculo do máximo pesado resultante da correspondência do gráfico de intersecção. A soma pesada das bordas sem correspondência

no processo de *matching* providencia a distância entre duas segmentações, este valor encontra-se normalizado entre [0, 1].

Para a geração das imagens de referência encontravam-se disponíveis duas ferramentas: uma desenvolvida pela universidade de Berkeley e uma desenvolvida por um colaborador da INESC Porto e denomina-se por *Bounding Box*. A ferramenta desenvolvida pela universidade de Berkeley permite obter informação acerca do contorno do objecto a ser segmentado, mais informação sobre esta ferramenta pode ser encontrada em [PAG02]. A segunda ferramenta não se encontra documentada, sendo o seu resultado um ficheiro que contém a informação da *Bounding Box* gerada, que poderá ser representada como uma máscara na imagem. Para a avaliação dos resultados dos algoritmos de segmentação optou-se por utilizar as máscaras resultantes da ferramenta de Berkeley por possuírem uma informação mais exacta do objecto, apesar desta exigir um esforço adicional pelo facto de o contorno do objecto ter que ser gerado manualmente. Exemplos dos resultados gerados pelas ferramentas encontram-se ilustrados na Figura 5.1 - 6.



Figura 5.1 - 6 Da esquerda para a direita: Imagem original; Imagem gerada com a ferramenta de Bounding Box; Imagem gerada com a ferramenta de Berkeley

O Gráfico 5.1 - 1 e Gráfico 5.1 - 2 apresentam a avaliação feita com o ISegmentation, OpenCv, Threshold e o resultado da segmentação obtido com a combinação dos resultados do OpenCv com o do Threshold. Os valores apresentados no eixo das ordenadas representam o erro que resulta da distância do resultado obtido dos algoritmos face à imagem de referência de segmentação gerada.

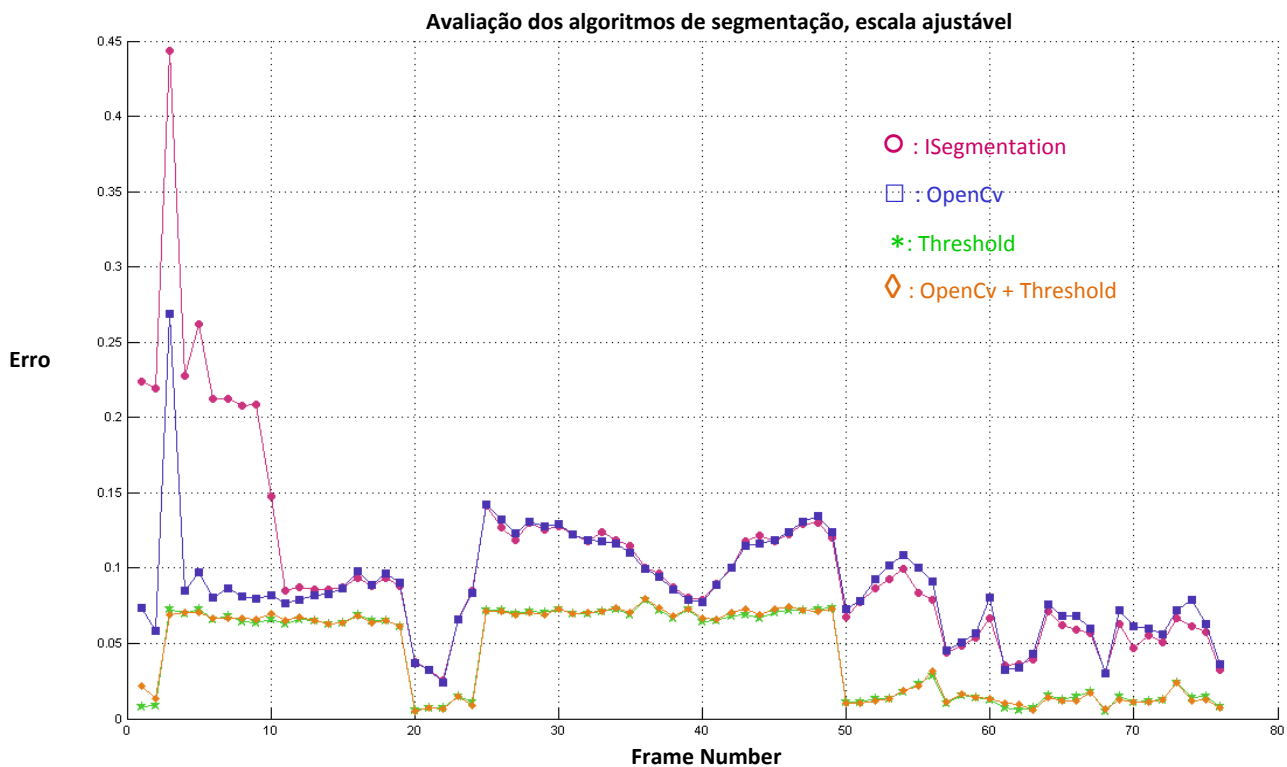


Gráfico 5.1 - 1 Avaliação dos algoritmos de segmentação: escala ajustável

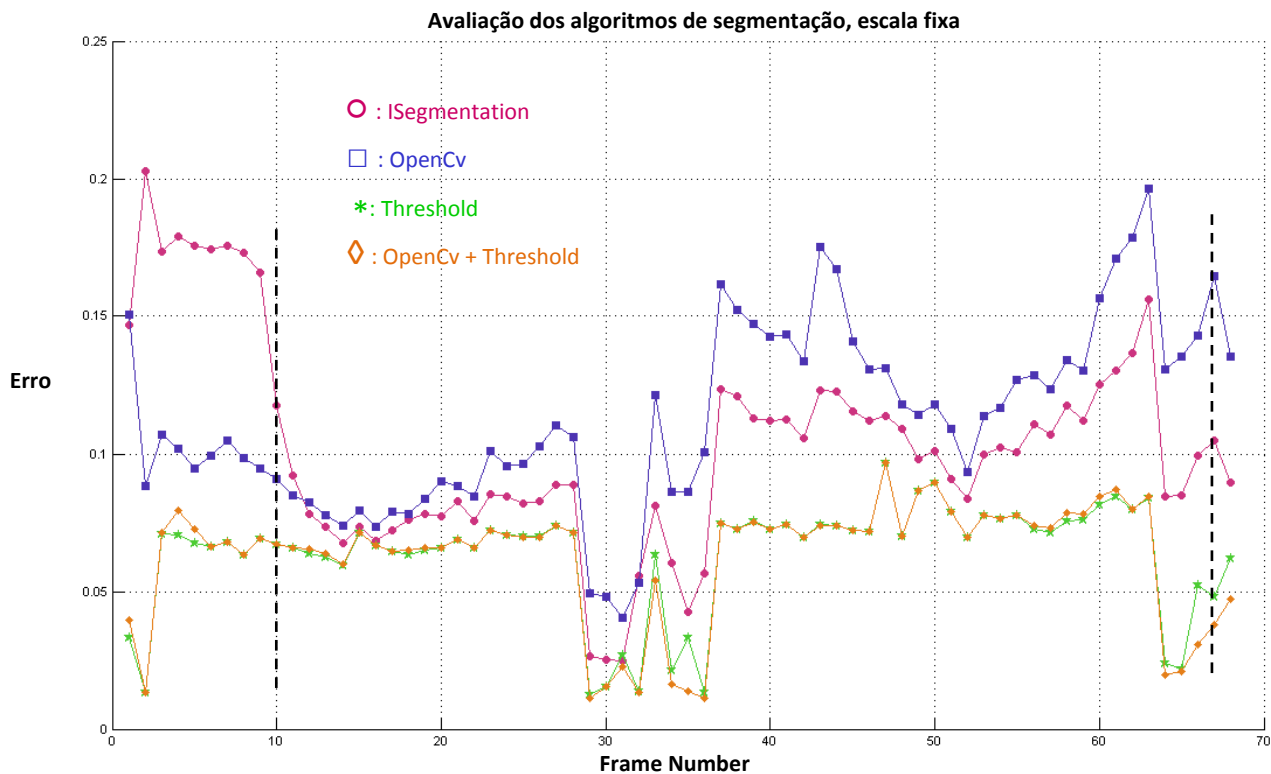


Gráfico 5.1 - 2 Avaliação dos algoritmos de segmentação: escala fixa

Como se pode constatar, com a aplicação dos dois algoritmos que são baseados no movimento, o erro associado à segmentação é maior do que no algoritmo baseado no

threshold. Analisando o gráfico do Threshold, as imagens em que o valor do erro é maior devem-se ao facto de o algoritmo não ter segmentado o objecto carro que é contemplado na imagem de segmentação de referência. O mesmo erro também ocorre com o ISegmentation e o OpenCv. Tal deve-se ao facto de o carro ser um objecto estático, o que faz com que algoritmos baseados em movimento para efectuar a segmentação de um objecto falhem. Da combinação do OpenCv, que obteve em geral melhores resultados que o ISegmentation, com o Threshold pode-se verificar que há uma ligeira melhoria nos valores dos erros apresentados para a escala fixa. No entanto, para a imagem com a escala ajustável não se verificou grandes melhorias nos valores de erros apresentados.

A Figura 5.1 - 7 e a Figura 5.1 - 8 ilustram os resultados obtidos com a aplicação dos três algoritmos de segmentação e a combinação dos resultados provenientes do OpenCv e do Threshold. Analisando a Figura 5.1 - 7, podemos verificar que na imagem de referência encontram-se dois objectos, um carro e uma pessoa.

Com a aplicação do algoritmo ISegmentation para a mesma imagem, verificou-se que: o carro não é segmentado, facto que seria de esperar, uma vez que o carro é um objecto estático nesta sequência e o algoritmo baseia-se no movimento dos objectos para os segmentar; foi possível segmentar a pessoa, apesar de o algoritmo contemplar o efeito de reflexão que é observado no chão da garagem; é também apresentada a posição inicial da pessoa devido ao facto de o algoritmo estar projectado para reter informação acerca de posições anteriores para casos de movimentos bruscos, o que leva a que apesar do algoritmo de ISegmentation possuir mais funções que permita eliminar mais ruído apresente resultados piores que o algoritmo OpenCv.

Os resultados obtidos com a aplicação do algoritmo OpenCv apresentam mais ruído face aos outros algoritmos de segmentação, como se pode observar melhor na Figura 5.1 - 8.

Com o algoritmo baseado no *threshold*, pode-se verificar na Figura 5.1 - 7 que a pessoa é segmentada sem ser contemplada a sua reflexão. Contudo, devido ao facto de o objecto carro possuir uma distribuição de cor que se situa abaixo do valor de *threshold* aplicado, este não é segmentado, facto que leva a que haja um erro associado na avaliação do algoritmo de segmentação.

Combinando o algoritmo baseado no movimento que melhores resultados apresentou (OpenCv) com o algoritmo Threshold, que foi feita recorrente à operação lógica AND, verificou-se que na imagem com uma escala ajustável (Figura 5.1 - 7) não apresenta melhorias face ao resultado obtido com o Threshold. Mas, com as imagens captadas com uma escala fixa, como se pode observar na Figura 5.1 - 8, o resultado da combinação dos dois algoritmos apresentou grandes melhorias. Na medida em que com o algoritmo baseado no *threshold* a pessoa do lado esquerdo encontra-se associada a uma região adicional que não faz parte da pessoa enquanto no resultado da combinação essa região não é contemplada.

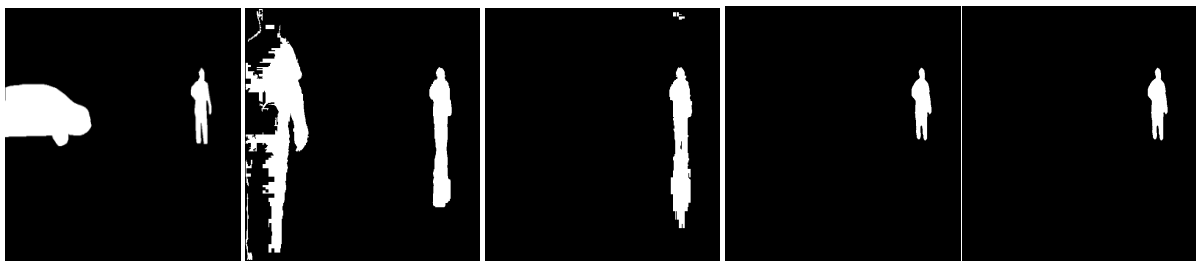


Figura 5.1 - 7 Imagem10 - Da esquerda para a direita: Imagem de referência de segmentação; Resultado do ISegmentation; Resultado do OpenCv; Resultado do *threshold*; Resultado da combinação do OpenCv com o *threshold*

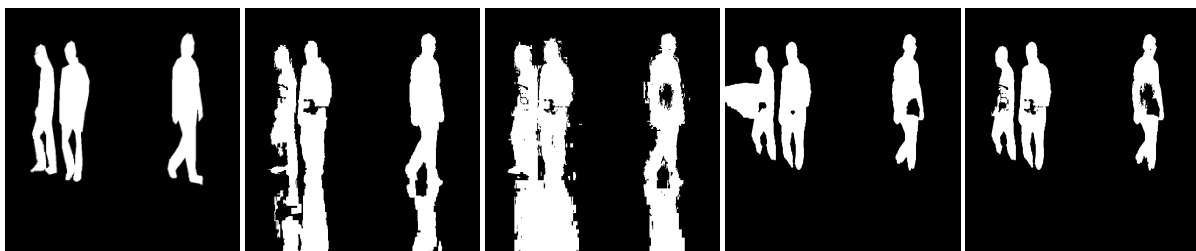


Figura 5.1 - 8 Imagem67 - Da esquerda para a direita: Imagem de referência de segmentação; Resultado do ISegmentation; Resultado do OpenCv; Resultado do *threshold*; Resultado da combinação do OpenCv com o *threshold*

6. Descrição e Seguimento

Neste capítulo descreve-se a comparação realizada entre os quatro modelos de descrição: Histograma de Cor [INT99], Histograma de Gradientes (*Histogram of Gradients – HOG*) [DAL05], SIFT (*Scale-Invariant Feature Transform*) [LOW04] e SURF (*Speeded Up Robust Features*) [BAY06]. A análise que será feita para as imagens captadas com uma escala ajustável e com uma escala fixa permitirá verificar qual dos modelos de descrição possui o melhor desempenho na distinção entre objectos que se apresentam numa sequência. Para tal, foram considerados três casos de estudo: a) distinguir entre objectos diferentes na mesma imagem; b) analisar a capacidade de inferir se é a mesma pessoa ao longo de uma sequência; c) distinguir entre objectos diferentes em imagens consecutivas.

6.1 DESCRITOR: HISTOGRAMA DE COR

Nesta fase foi necessário recorrer ao uso de uma função da biblioteca do OpenCv para efectuar o cálculo dos histogramas de cor das imagens previamente segmentadas, em que as localizações das regiões de interesse na imagem se encontram definidas no ficheiro gerado na etapa da segmentação. Após a extracção da informação de cor do objecto para a construção do histograma, realizou-se a comparação de dois histogramas e calculou-se a similaridade entre os histogramas recorrendo à expressão matemática do Qui-Quadrado [ERD04], que se encontra expressa na equação (2). Os valores resultantes encontram-se normalizados.

$$0 \leq X^2 (H_t, H_{ref}) = \frac{\sum_{j=1}^B \frac{[r_1 H_t(j) - r_2 H_{ref}(j)]^2}{H_t(j) + H_{ref}(j)}}{N_{H_t} + N_{H_{ref}}} \leq 1 \quad (2)$$

Onde:

$$r_1 = \sqrt{\frac{N_{H_{ref}}}{N_{H_t}}}, r_2 = \frac{1}{r_1}, N_{H_t} = \sum_{j=1}^B H_t(j), N_{H_{ref}} = \sum_{j=1}^B H_{ref}(j)$$

H_t - Refere-se ao histograma da imagem no instante t

H_{ref} - Refere-se ao histograma de referência, $\{t-1, \dots, t-1\}$

r_1 e r_2 - Parcelas para a normalização dos dados

A escolha dos dois histogramas a serem comparados dependerá do caso de estudo. Ou seja, no caso em que se comparam objectos diferentes que estejam na mesma imagem, o primeiro histograma referir-se-á ao objecto A e o segundo histograma ao objecto B. O cálculo só é realizado se ambos os objectos estiverem presentes na mesma imagem. A Figura 6.1 - 1 pretende ilustrar a comparação do objecto A com o objecto B e com o objecto C.

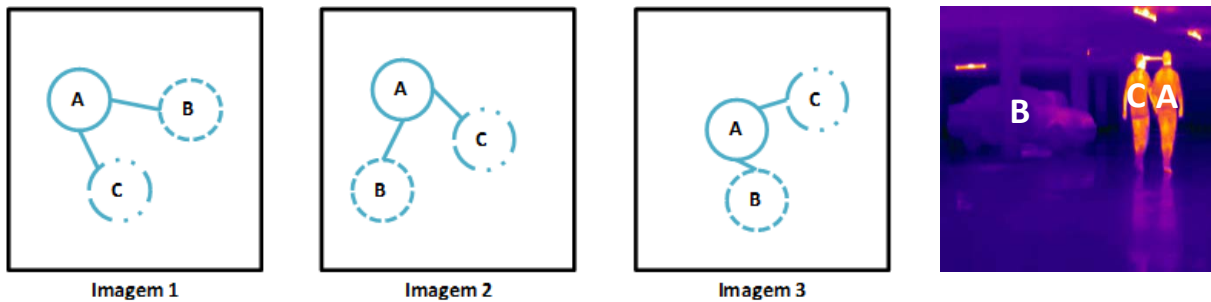


Figura 6.1 - 1 Comparação entre objectos diferentes na mesma imagem

Para o segundo caso de estudo em que se compara a mesma pessoa ao longo de uma sequência, o primeiro histograma será da Pessoa A na imagem 1 e o segundo será da mesma pessoa (Pessoa A) na imagem seguinte, que poderá ser a imediatamente a seguir (imagem 2) ou a imagem em que a Pessoa A voltará a entrar na cena (por exemplo: imagem 15). A Figura 6.1 - 2 pretende ilustrar a comparação do objecto A ao longo de uma sequência.

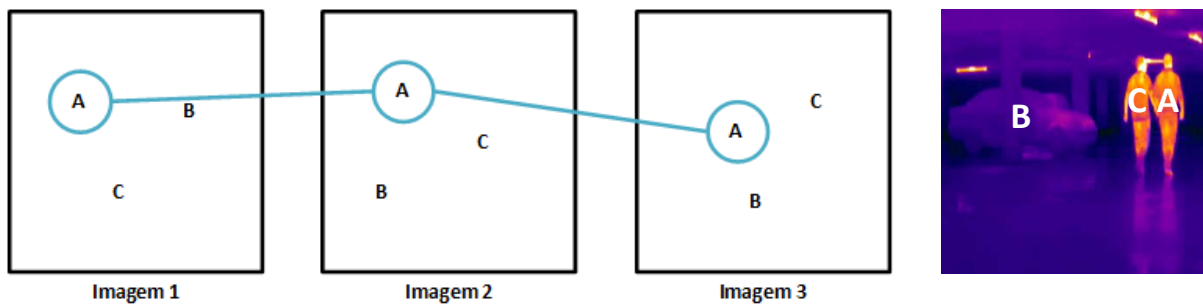


Figura 6.1 - 2 Comparação entre o mesmo objecto, mas em imagens diferentes

No caso da comparação de objectos diferentes em frames consecutivas, o primeiro histograma será o do objecto A e o segundo de um objecto B. O cálculo é realizado se e só se o objecto B estiver presente na imagem imediatamente a seguir à imagem em que se encontra o objecto A. A Figura 6.1 - 3 pretende ilustrar a comparação entre objectos diferentes em imagens consecutivas.

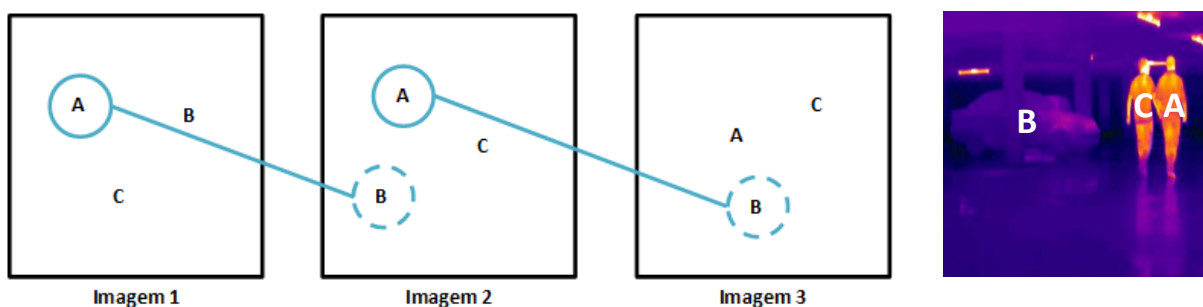


Figura 6.1 - 3 Comparação de o objecto A com o objecto B na imagem seguinte

A mesma metodologia de estudo de comparação entre os objectos foi aplicada para os descritores HOG, SIFT e SURF.

6.1.1 ESCALA AJUSTÁVEL

Para uma melhor análise dos gráficos que serão apresentados, o eixo das ordenadas indica os valores das distâncias obtidas entre os histogramas. Ou seja, quanto menor for o valor obtido da distância, maior será a similaridade entre os histogramas. Da mesma forma, quanto maior for o valor da distância, maior será a diferença entre os histogramas.

No eixo das abcissas é representado apenas as imagens em que as condições são satisfeitas para cada situação de estudo. Isto é, para o primeiro gráfico apresentado, que retratará a situação de utilizar os modelos para distinguir entre objectos diferentes que se encontram

na mesma imagem, o número apresentado no eixo das abcissas representará apenas as situações em que ambos os objectos a serem comparados existam na mesma imagem.

No segundo gráfico serão retratadas duas situações: analisar a mesma pessoa ao longo de uma sequência e a situação de ser ou não possível distinguir objectos em imagens consecutivas. Para este caso, o eixo das abcissas representará para a primeira situação as imagens em que o objecto alvo de estudo aparece, na medida em que o valor da distância representado para a imagem 1 corresponde à comparação do objecto na imagem 1 e na imagem 2. Para a segunda situação, os valores apresentados no eixo correspondem à condição de um segundo objecto de estudo se encontrar na imagem imediatamente a seguir à imagem em que o primeiro objecto de estudo se encontra.

Para todos os descritores, o objecto em estudo será uma pessoa, denominada por Pessoa C e que será comparada com três objectos: um carro e duas pessoas (Pessoa D e Pessoa E). No Gráfico 6.1.1 - 1 é apresentada a comparação da Pessoa C com um carro, Pessoa C com Pessoa D e pessoa C com Pessoa E na mesma imagem. Os dados mais relevantes do gráfico como a média de todos os valores, a variância, o valor mínimo e máximo registados encontram-se na Tabela 6.1.1 - 1.

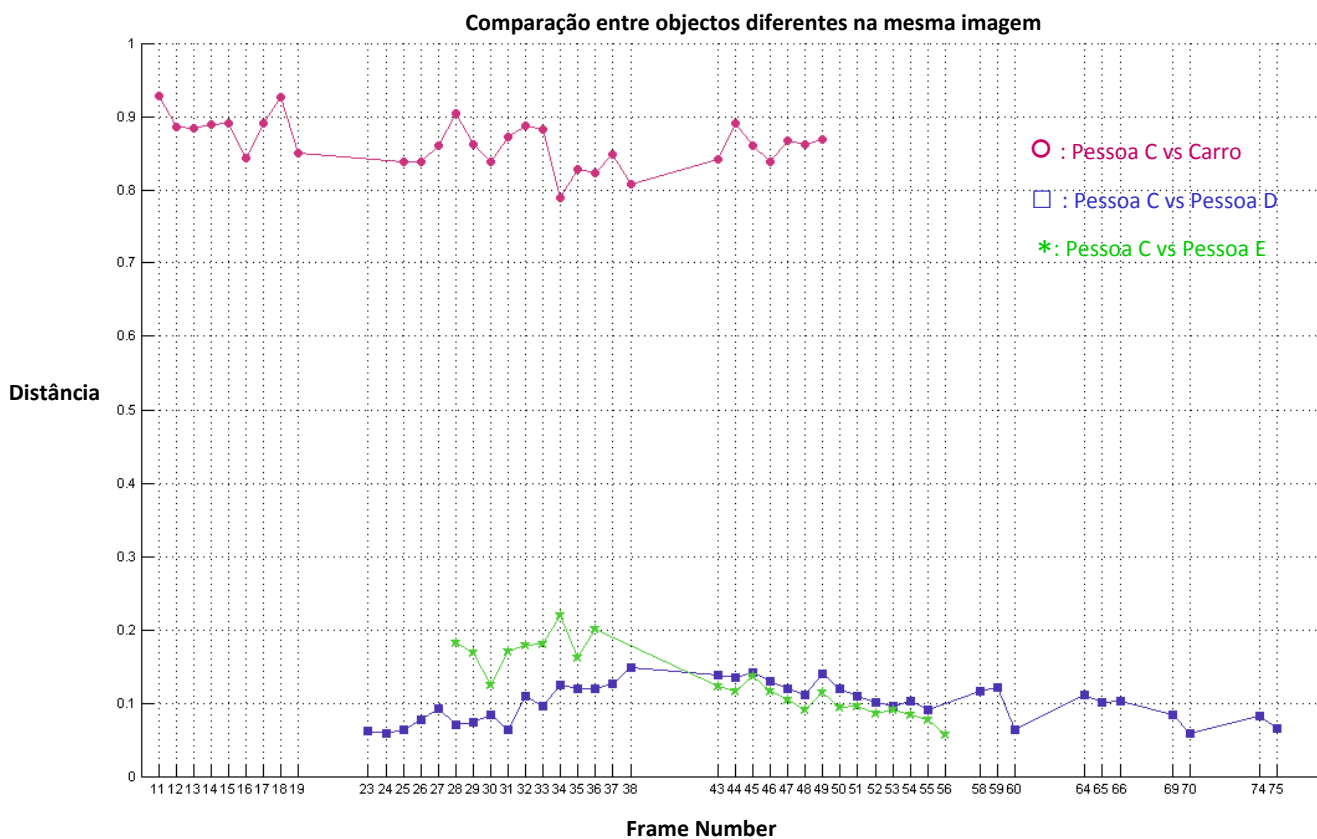


Gráfico 6.1.1 - 1 Comparação entre objectos diferentes na mesma imagem

Tabela 6.1.1 - 1 Histograma de cor (escala variável): Dados obtidos para a situação a)

Escala	Modelo	Descrição		Média	Variância	Mínimo	Máximo
Ajustável	Histograma de Cor	Mesma imagem	Pessoa C vs Carro	0.863	0.001	0.789	0.929
			Pessoa C vs Pessoa D	0.100	0.001	0.058	0.148
			Pessoa C vs Pessoa E	0.129	0.002	0.057	0.219

Analisando a tabela e o gráfico, pode-se constatar que, como seria de esperar, há uma menor similaridade entre uma pessoa e um carro do que entre pessoas. No caso da comparação entre pessoas verifica-se quer graficamente, quer no valor da média que se encontra na tabela anterior que há uma maior similaridade entre a Pessoa C e Pessoa D do que entre a Pessoa C e a pessoa E. Tal acontece devido ao facto de a Pessoa E face à Pessoa C possuir mais altura e a distribuição da temperatura corporal ser ligeiramente diferente, como se pode observar na Figura 6.1.1 - 1.

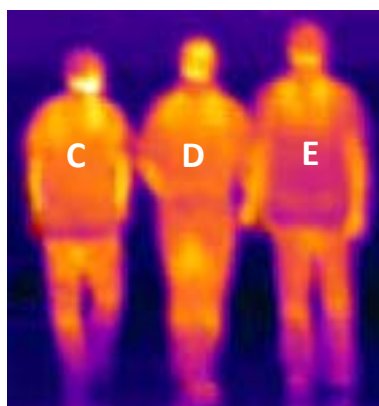


Figura 6.1.1 - 1 Objectos de estudo

Devido ao facto de o processamento de imagens ser feito com base na imagem em que o objecto é detectado e identificado com a imagem seguinte, ou seja, nos instantes $t-1$ e t , tornou-se necessário analisar a similaridade entre os histogramas de objectos diferentes em imagens consecutivas. Ou seja, tendo o histograma de cor da Pessoa C e ao analisar a imagem seguinte que possui mais três objectos (duas pessoas e um carro), se é possível quando comparado com outro objecto inferir que se trata de um objecto diferente ou se não consegue fazer essa distinção, comprometendo a etapa de seguimento.

No Gráfico 6.1.1 - 2 só irão ser contemplados os casos de comparação da mesma pessoa ao longo de uma sequência e a comparação entre pessoas distintas em imagens consecutivas. A comparação entre uma pessoa e um carro não foi analisada pelo facto de já se ter verificado que é possível distinguir entre uma pessoa e um carro. Os valores mais relevantes encontram-se na Tabela 6.1.1 - 2.

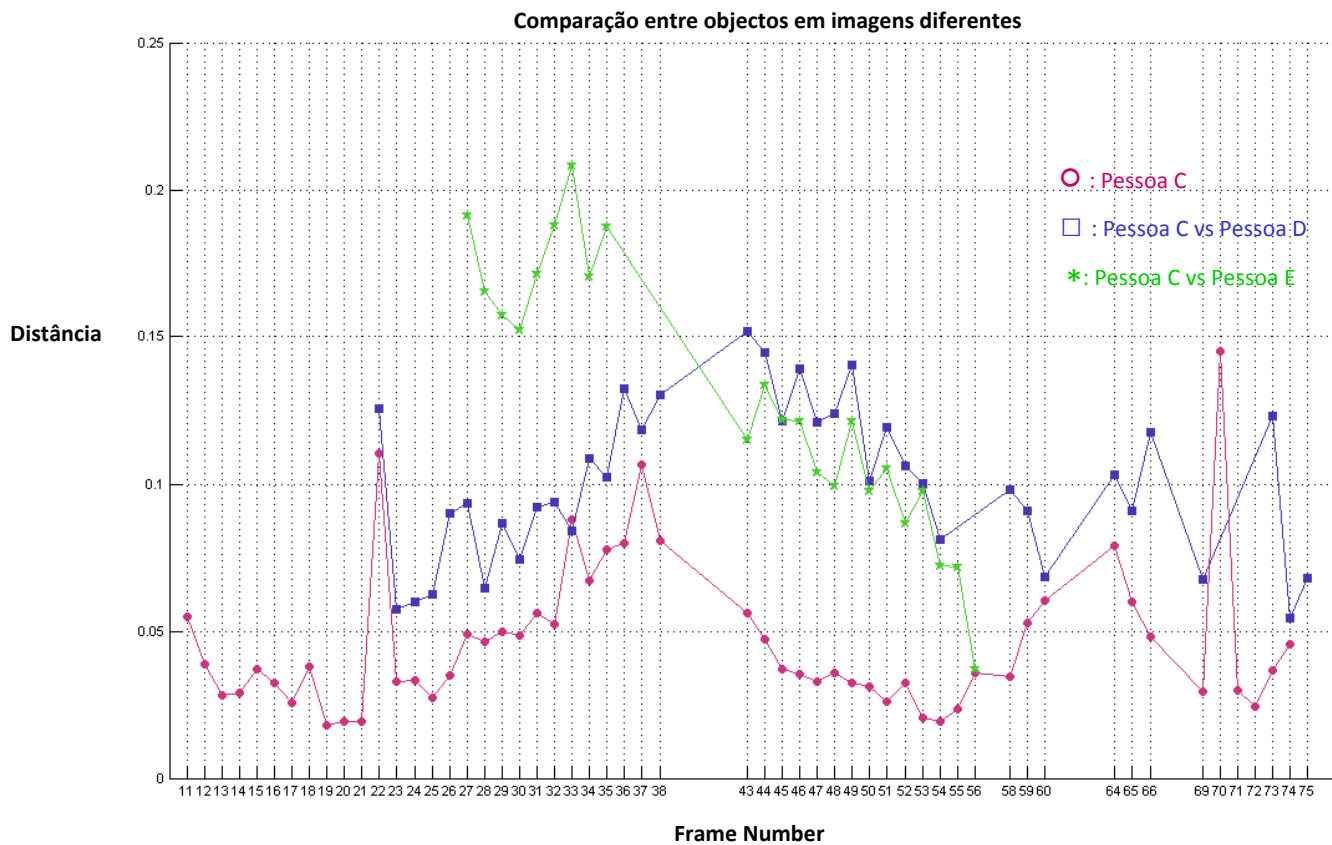


Gráfico 6.1.1 - 2 Comparação entre objectos em imagens diferentes

Tabela 6.1.1 - 2 Histograma de cor (escala variável): Dados obtidos para a situação b) e c)

Escala	Modelo	Descrição	Média	Variância	Mínimo	Máximo
Ajustável	Histograma de Cor	Mesma Pessoa (C)	0.046	0.001	0.018	0.144
		Pessoas Diferentes (C vs D)	0.100	0.001	0.055	0.152
		Pessoas Diferentes (C vs E)	0.129	0.002	0.037	0.208

Da análise dos resultados obtidos, pode-se verificar que se a decisão, de ser a mesma pessoa ou não, fosse baseada num conjunto de imagens, ou seja, na média, seria possível inferir com uma margem de segurança de que se tratava da mesma pessoa, na medida em que a média dos valores para a mesma pessoa é de 0.05, metade do valor da média da

comparação da pessoa C com outra pessoa [0.1; 0.13]. Mas devido ao facto de a análise ser feita com base na imagem nos instantes t e $t-1$, como se pode constatar no gráfico há períodos em que a decisão poderá ser comprometida, na medida em que o valor da distância para a mesma pessoa é superior à distância calculada entre a Pessoa C e uma outra pessoa. Tal deve-se ao facto da sequência captada não possuir uma *frame rate* constante, facto que origina saltos abruptos de uma imagem para a outra como se pode observa na Figura 6.1.1 - 2, que ilustra os casos mais extremos apresentados no gráfico - imagens 22, 23, 70, 71.

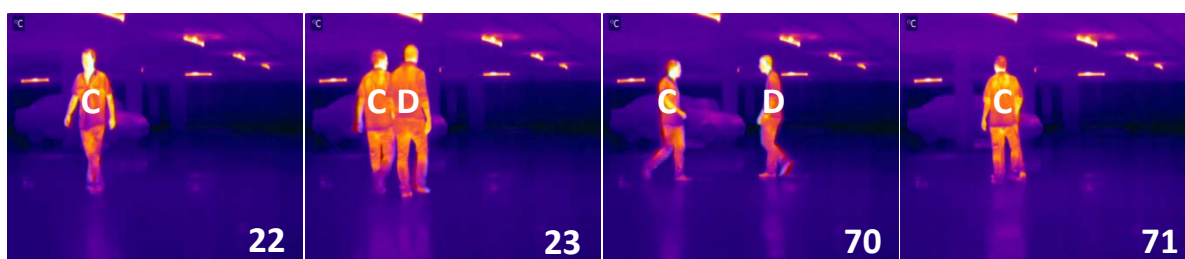


Figura 6.1.1 - 2 Ilustração dos possíveis erros ocorridos

6.1.2 ESCALA FIXA

Foi feito um estudo semelhante ao anterior com as imagens captadas com a escala fixa. No Gráfico 6.1.2 - 1 é apresentada a comparação entre objectos diferentes na mesma imagem. Os dados mais relevantes do gráfico encontram-se na Tabela 6.1.2 - 1.

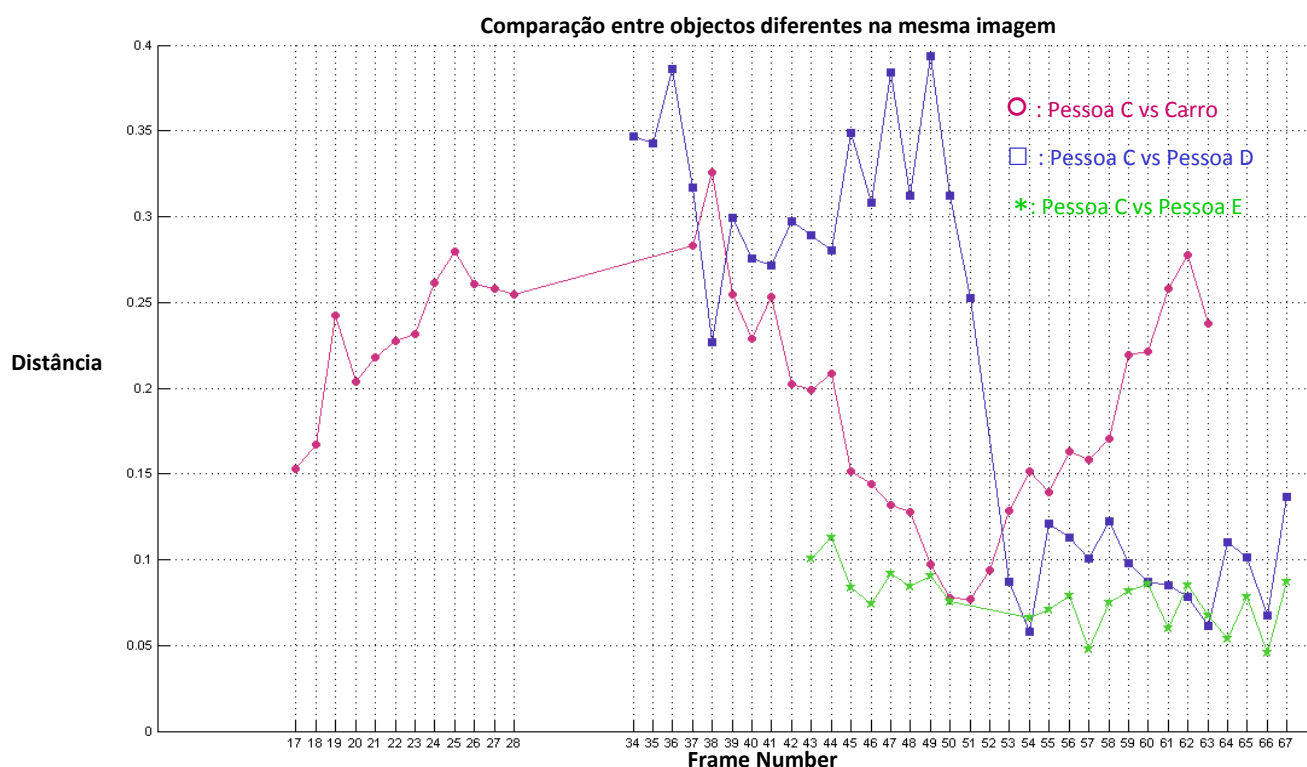


Gráfico 6.1.2 - 1 Comparação entre objectos diferentes na mesma imagem

Tabela 6.1.2 - 1 Histograma de cor (escala fixa): Dados obtidos para a situação a)

Escala	Modelo	Descrição		Média	Variância	Mínimo	Máximo
2040	Histograma de Cor	Mesma Imagem	Pessoa C vs Carro	0.198	0.004	0.077	0.326
			Pessoa C vs Pessoa D	0.214	0.013	0.058	0.393
			Pessoa C vs Pessoa E	0.077	0.001	0.046	0.113

Da análise dos resultados obtidos e em comparação com os resultados anteriores, pode-se verificar que com o uso de uma escala fixa a distinção entre uma pessoa e um carro não foi tão evidente tendo sido a similaridade entre duas pessoas, neste caso entre a Pessoa C e a Pessoa D maior do que a comparação entre um carro e uma pessoa. Tal facto deve-se à distribuição de cores da Pessoa C estar mais próxima da do carro do que da Pessoa D, que possui tons mais brilhantes, como se pode observar na Figura 6.1.2 - 1.



Figura 6.1.2 - 1 Ilustração do erro ocorrido

O Gráfico 6.1.2 - 2 contempla todos os casos, ou seja a comparação da mesma pessoa ao longo de uma sequência, a comparação entre pessoas distintas em imagens consecutivas e a comparação da pessoa e um carro devido ao facto de no gráfico não ter sido possível distinguir tão bem como na escala ajustável entre estes dois objectos. Os valores mais relevantes encontram-se na Tabela 6.1.2 - 2.

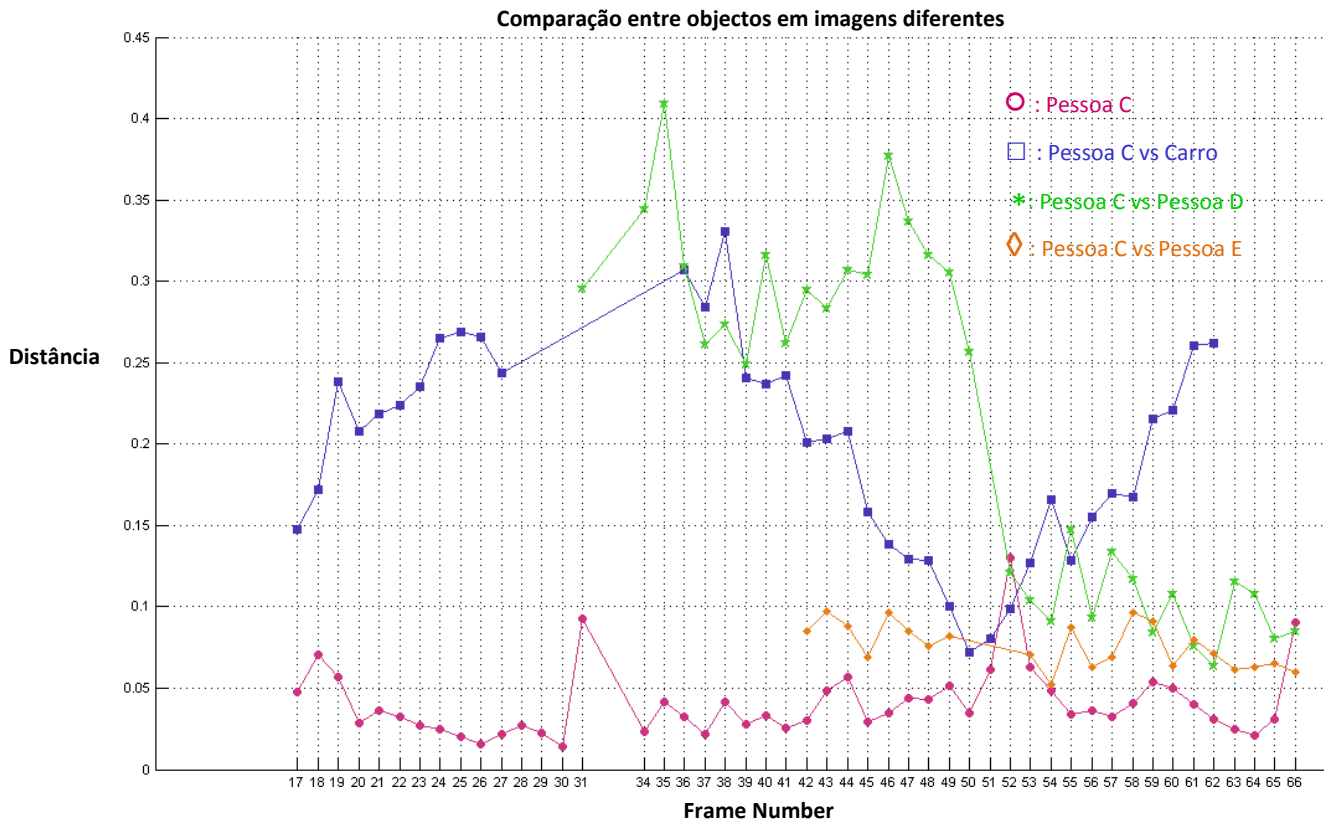


Gráfico 6.1.2 - 2 Comparação entre objectos em imagens diferentes

Tabela 6.1.2 - 2 Histograma de cor (escala fixa): Dados obtidos para a situação b) e c)

Escala	Modelo	Descrição	Média	Variância	Mínimo	Máximo
2040	Histograma de Cor	Imagens Consecutivas				
		Mesma Pessoa (C)	0.041	0.001	0.014	0.1302
		Pessoa C vs Carro	0.198	0.004	0.071	0.330
		Pessoas Diferentes (C vs D)	0.213	0.012	0.063	0.409
		Pessoas Diferentes (C vs E)	0.076	0.0001	0.052	0.097

Analisando o gráfico e os valores que se encontram na Tabela 6.1.2 - 2, podemos verificar que apesar de a distância entre a comparação de uma pessoa e um carro não ser tão discrepante como no gráfico das imagens com uma escala ajustável, é possível, de uma forma geral, distinguir entre um carro e uma pessoa. Constata-se, também, que nesta sequência, a similaridade entre duas pessoas é menor, levando a que seja possível utilizar um valor limite de decisão. Para o caso de se estar a determinar se é a mesma pessoa ao longo da sequência, pode-se verificar que de um modo geral foi possível inferir que seria a mesma pessoa, com excepção na imagem 52 em comparação com a imagem 53. Tal deveu-

se ao facto de a pessoa C estar de costas voltadas para a câmara e na imagem seguinte se encontrar de frente. Uma situação semelhante ocorre nas imagens 66 e 67, pelo facto de a pessoa estar numa posição de frente e na imagem seguinte numa posição lateral, como se pode observar na Figura 6.1.2 - 2.

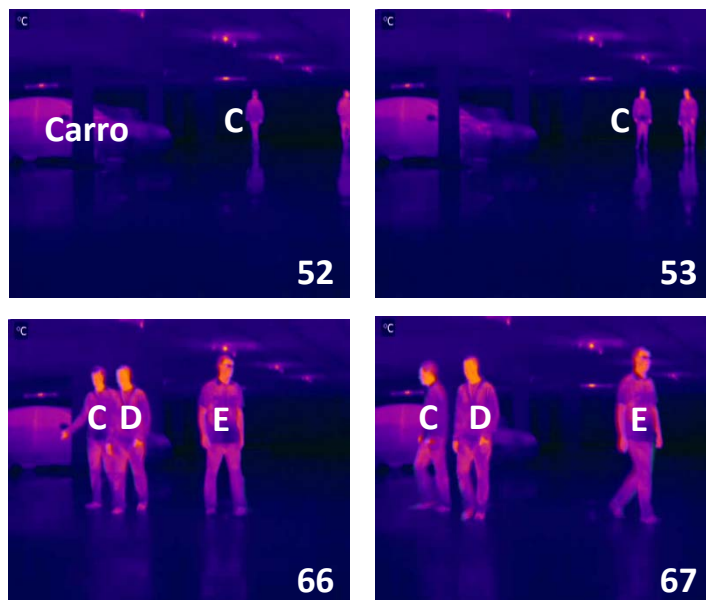


Figura 6.1.2 - 2 Ilustração dos erros ocorridos

6.2 DESCRITOR: HISTOGRAMA DE GRADIENTES (HOG)

Nesta fase será realizada um estudo semelhante ao realizado com o descritor anterior (Histograma de Cor). O descritor HOG já se encontra implementado, tendo sido apenas necessário adaptá-lo para o caso em estudo. O uso do histograma de gradientes difere do histograma de cor pela forma como é realizada a análise da imagem e na informação em que se baseia para a realizar. Informações mais detalhadas acerca da forma como é feita esta análise pode ser encontrada em [DAL05]. Os valores resultantes da comparação dos histogramas de gradientes encontram-se normalizados, tendo sido utilizado o mesmo método matemático para a comparação.

6.2.1 ESCALA AJUSTÁVEL

No Gráfico 6.2.1 - 1 é apresentado a comparação da Pessoa C com um carro, Pessoa C com Pessoa D e pessoa C com Pessoa E na mesma imagem. Os dados mais relevantes do gráfico como a média de todos os valores, a variância, o valor mínimo e máximo registados encontram-se na Tabela 6.2.1 - 1.

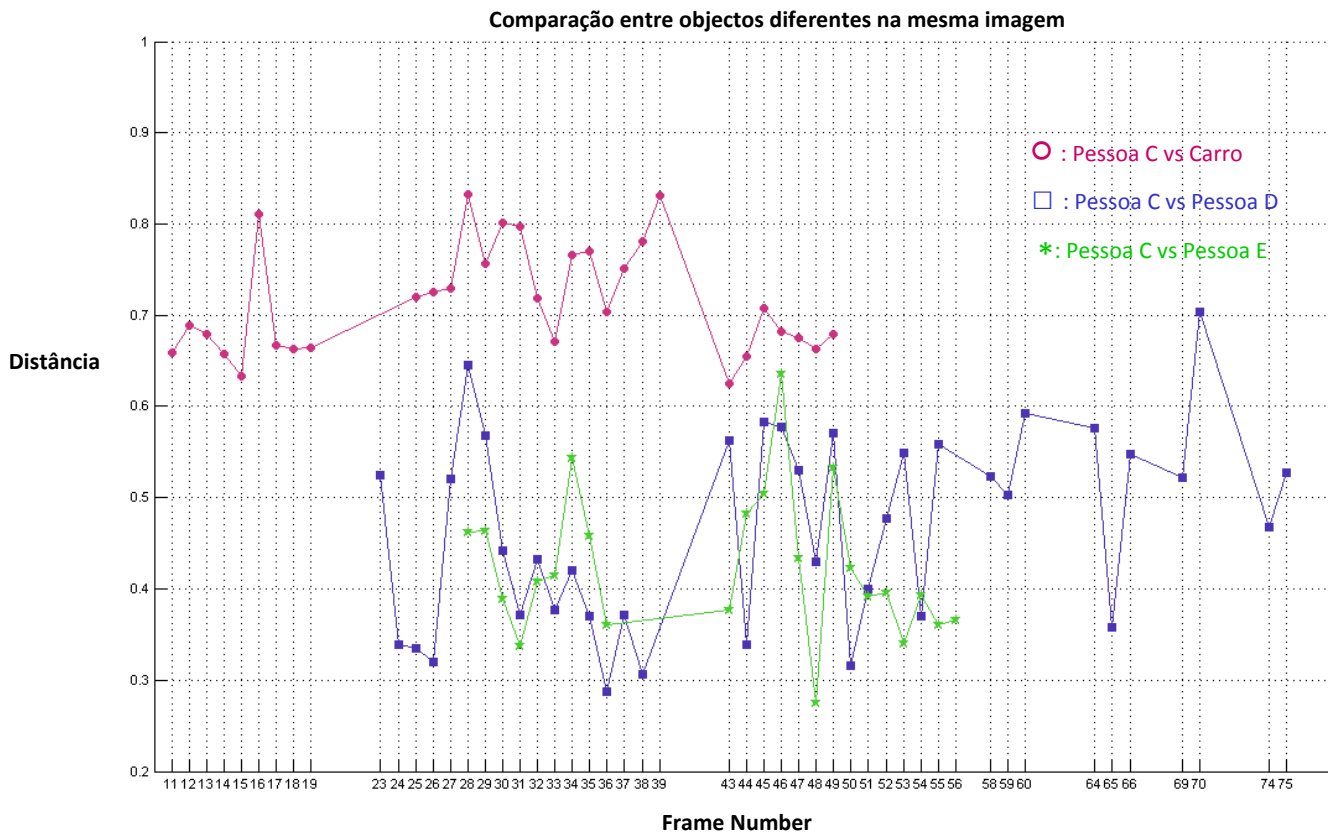


Gráfico 6.2.1 - 1 Comparação entre objectos diferentes na mesma imagem

Tabela 6.2.1 - 1 Histograma de Gradientes (escala ajustável): Dados obtidos para a situação a)

Escala	Modelo	Descrição	Média	Variância	Mínimo	Máximo
Ajustável	Histograma de Gradiente	Mesma Imagem				
		Pessoa C vs Carro	0.715	0.003	0.625	0.833
		Pessoa C vs Pessoa D	0.467	0.011	0.287	0.703
		Pessoa C vs Pessoa E	0.424	0.006	0.276	0.635

Observando os resultados obtidos com a aplicação do descritor HOG, verifica-se que a distância entre os gradientes de uma pessoa e os gradientes de um carro diferem de quase o dobro do valor da média da distância obtida entre a comparação de duas pessoas. Em relação à distância entre duas pessoas, os valores apresentados são muito próximos, o que indica que a similaridade entre duas pessoas é maior, como seria de esperar.

No Gráfico 6.2.1 - 2 só irão ser contemplados os casos de comparação da mesma pessoa ao longo de uma sequência e a comparação entre pessoas distintas em imagens consecutivas. A comparação entre uma pessoa e um carro não foi analisada pelo facto de já se ter verificado

que é possível distinguir entre uma pessoa e um carro. Os valores mais relevantes encontram-se na Tabela 6.2.1 - 2.

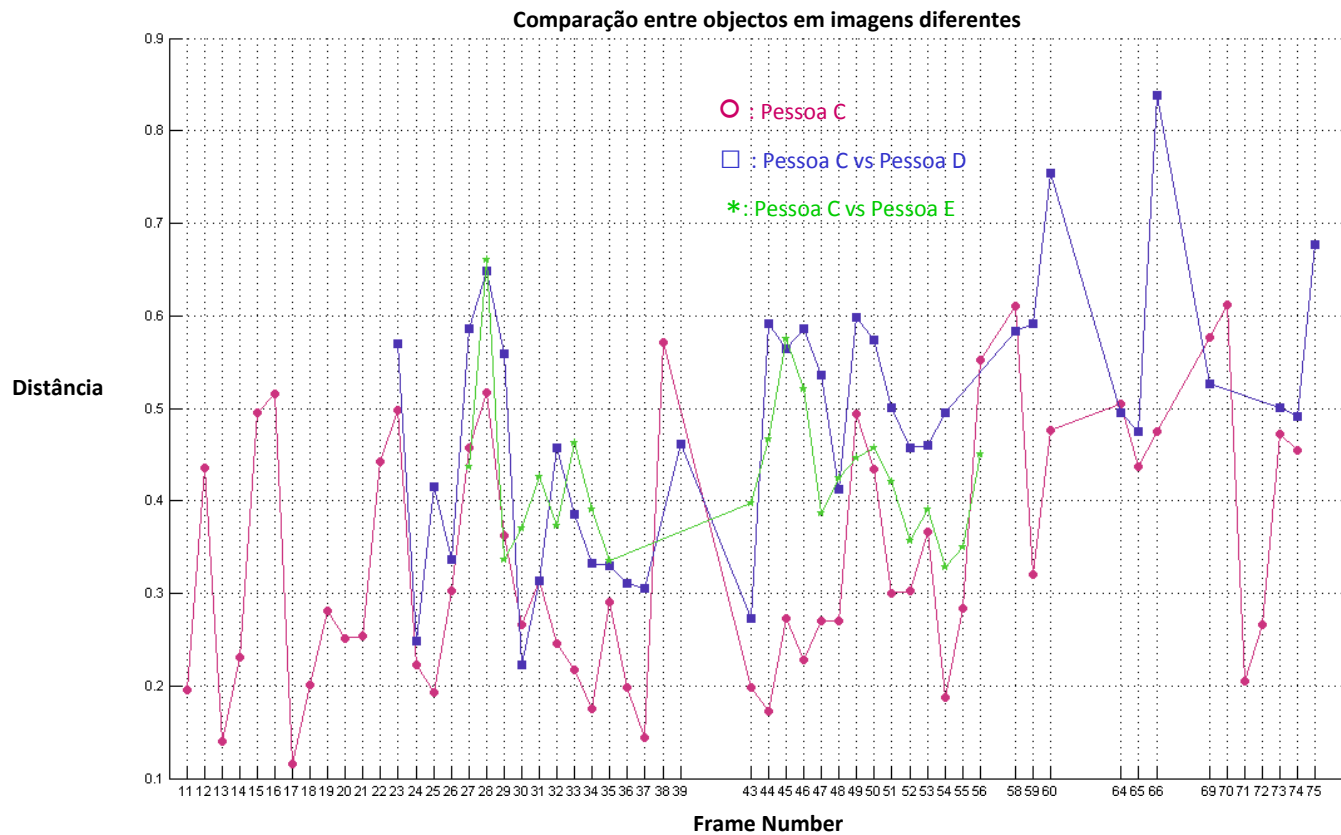


Gráfico 6.2.1 - 2 Comparação entre objectos em imagens diferentes

Tabela 6.2.1 - 2 Histograma de Gradientes (escala ajustável): Dados obtidos para a situação b) e c)

Escala	Modelo	Descrição	Média	Variância	Mínimo	Máximo
Ajustável	Histograma de Gradiente	Mesma Pessoa (C)	0.338	0.019	0.115	0.612
		Pessoas Diferentes (C vs D)	0.485	0.019	0.2223	0.838
		Pessoas Diferentes (C vs E)	0.424	0.006	0.327	0.661

Analisando o gráfico que representa a comparação da mesma pessoa ao longo de uma sequência e a comparação dessa mesma pessoa com outras, verifica-se que com o histograma de gradientes não é possível distinguir entre duas pessoas. A Figura 6.2.1 - 1 representa algumas das situações (imagem 15, 16, 17, 70 e 71) em que o descritor não consegue inferir correctamente que é a mesma pessoa. Pode-se constatar pelas imagens que o histograma de gradientes é afectado pela posição da pessoa, ou seja, se a pessoa se encontra de frente, de lado ou costas.

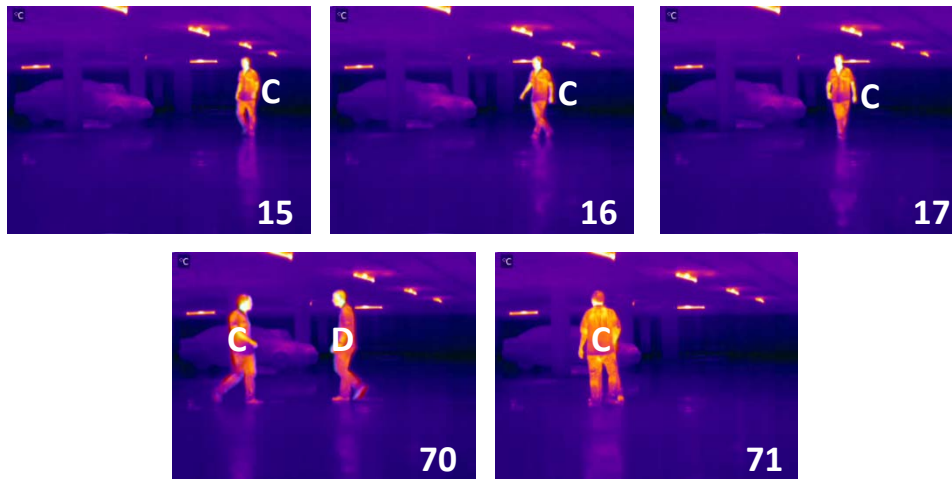


Figura 6.2.1 - 1 Ilustração dos erros ocorridos

6.2.2 ESCALA FIXA

Foi feito um estudo semelhante ao anterior com as imagens captadas com a escala fixa. No Gráfico 6.2.2 - 1 é apresentado a comparação da Pessoa C com um carro, Pessoa C com Pessoa D e pessoa C com Pessoa E na mesma imagem. Os dados mais relevantes do gráfico encontram-se na Tabela 6.2.2 - 1.

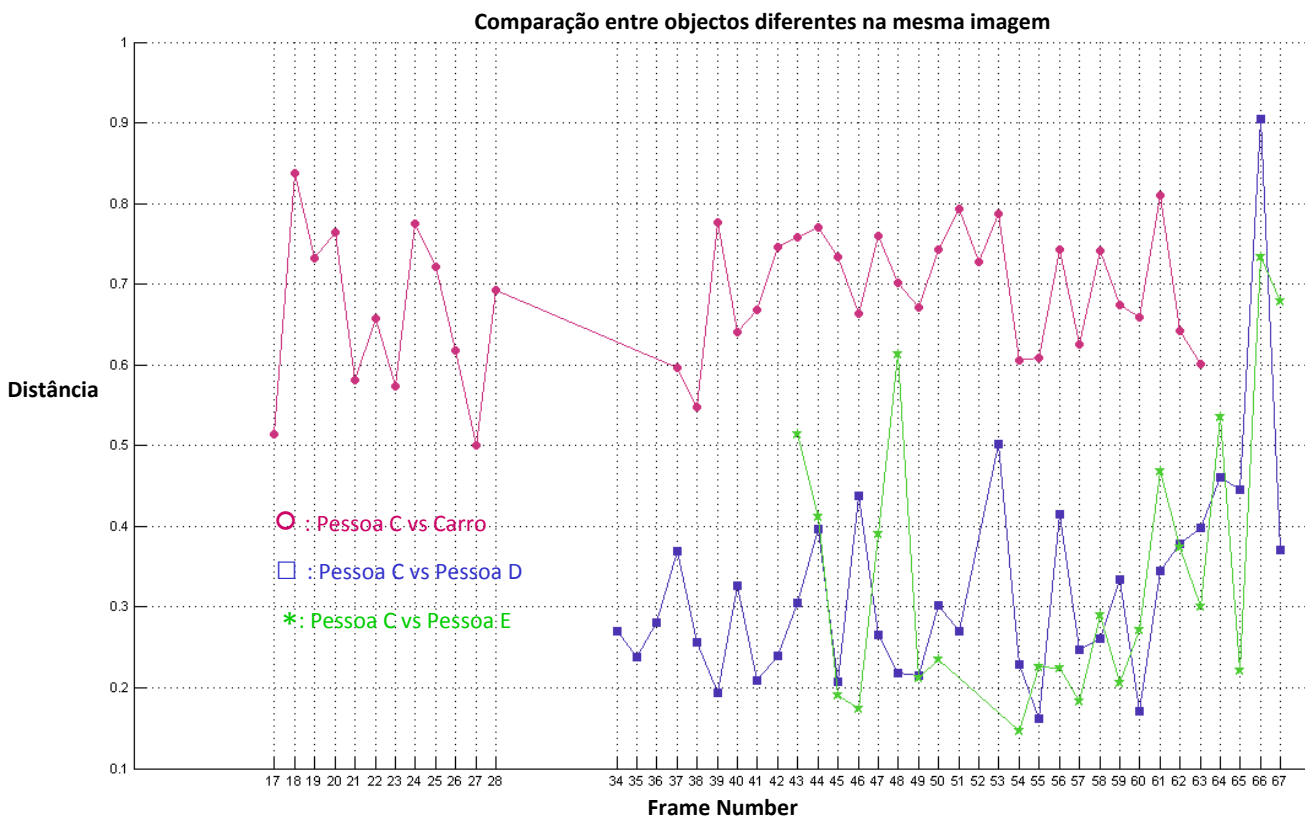


Gráfico 6.2.2 - 1 Comparação entre objectos diferentes na mesma imagem

Tabela 6.2.2 - 1 Histograma de Gradientes (escala fixa): Dados obtidos para a situação a)

Escala	Modelo	Descrição	Média	Variância	Mínimo	Máximo
2040	Histograma de Gradiente	Pessoa C vs Carro	0.686	0.007	0.501	0.837
		Pessoa C vs Pessoa D	0.322	0.019	0.161	0.904
		Pessoa C vs Pessoa E	0.346	0.030	0.145	0.734

Da análise do gráfico, podemos constatar que utilizando uma escala fixa com um descritor HOG, ao contrário do que sucedeu com o uso do histograma de cor para uma escala de temperatura fixa, é possível distinguir entre uma pessoa e um carro. No entanto, para a comparação entre duas pessoas, o resultado é semelhante.

No Gráfico 6.2.2 - 2 encontram-se representados os casos de comparação da mesma pessoa ao longo de uma sequência e a comparação entre pessoas distintas em imagens consecutivas. A comparação entre uma pessoa e um carro não foi analisada pelo facto de já se ter verificado que é possível distinguir entre uma pessoa e um carro. Os valores mais relevantes encontram-se na Tabela 6.2.2 - 2.

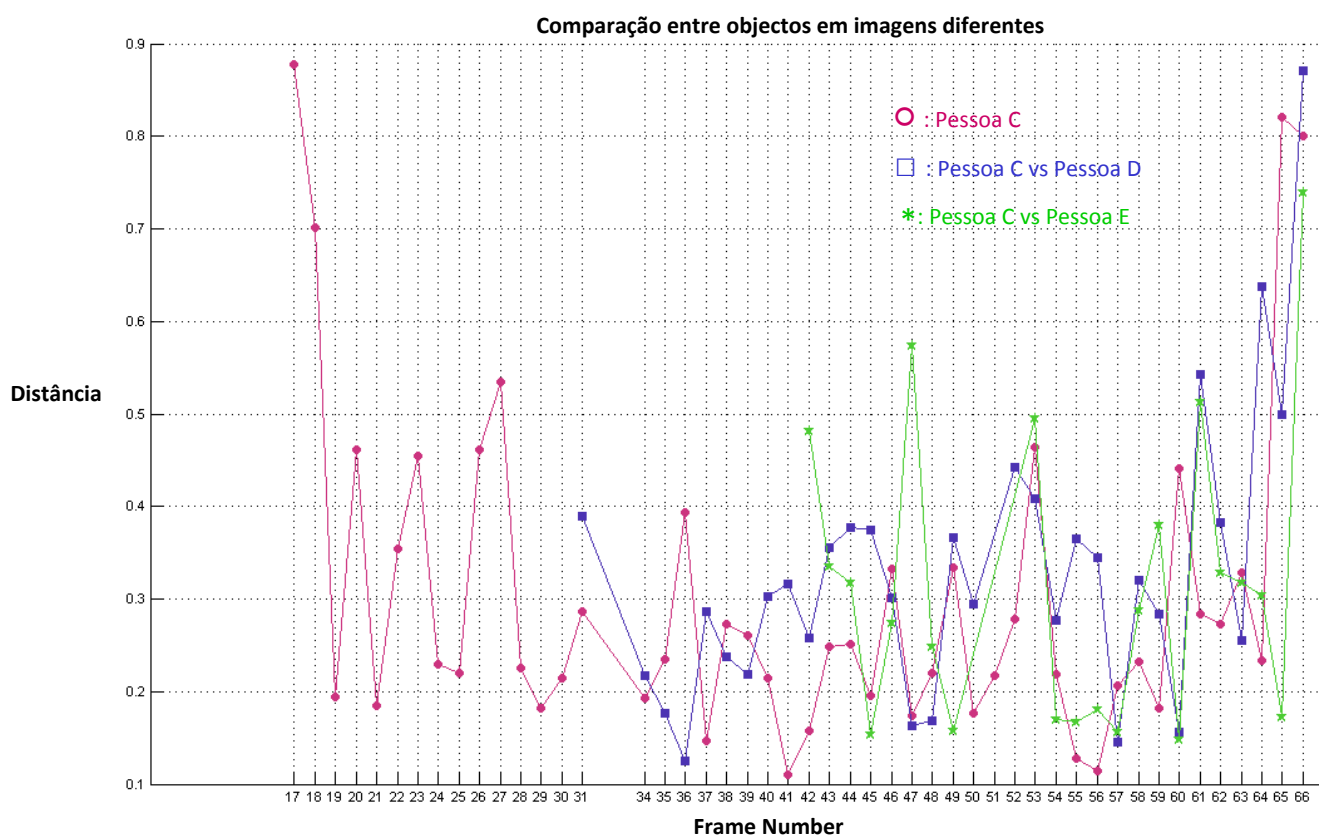


Gráfico 6.2.2 - 2 Comparação entre objectos em imagens diferentes

Tabela 6.2.2 - 2 Histograma de Gradientes (escala fixa): Dados obtidos para a situação b) e c)

Escala	Modelo	Descrição		Média	Variância	Mínimo	Máximo
2040	Histograma de Gradiente	Imagens Consecutivas	Mesma Pessoa (C)	0.306	0.032	0.110	0.878
			Pessoas Diferentes (C vs D)	0.329	0.022	0.125	0.870
			Pessoas Diferentes (C vs E)	0.313	0.025	0.148	0.738

Com uma mudança na escala de temperatura utilizada, pode-se constatar que para a comparação entre a mesma pessoa ao longo de uma sequência e para a comparação da pessoa com uma outra, os resultados obtidos são mais semelhantes entre si do que os resultados obtidos com uma escala de temperatura ajustável, em que para a mesma pessoa o valor da média foi de 0,34 e para a comparação com outras pessoas o valor situa-se nos [0,43; 0,49].

À semelhança da escala ajustável, os erros ocorridos devem-se ao facto da variação da posição apresentada pela pessoa ao longo das imagens, como se pode observar na Figura 6.2.2 - 1.



Figura 6.2.2 - 1 Ilustração do erro ocorrido

6.3 DESCRITOR: SIFT

Scale-Invariant Features Transform é um modelo de descrição que detecta um conjunto de pontos de interesse numa imagem, [LOW04]. São considerados bons candidatos a pontos de interesse, pontos que sejam invariantes a deformações locais numa imagem. O método de comparação é realizado da seguinte forma: supondo que no objecto 1 foram detectados 20 pontos e no objecto 2 foram detectados 25, (Figura 6.3 - 1), na etapa de comparação é fixado um ponto do objecto 1 e realiza-se uma comparação com todos os pontos do objecto2. Dessas comparações resultam distâncias euclidianas, no qual são guardadas as

duas distâncias menores em que a distância final seleccionada terá que ser 60% menor que o valor da segunda distância seleccionada para que haja um correspondência. O processo repete-se para todos os pontos do objecto 1. Como resultados poderão haver pontos diferentes do objecto 1 que correspondam ao mesmo ponto no objecto 2, pontos do objecto 1 que só tenham uma correspondência no objecto 2, assim como, poderão existir pontos que não terão nenhuma correspondência. Para os pontos que não possuam nenhuma correspondência é-lhes atribuído um valor de penalização, como se pode observar pela equação (3). A normalização do valor resultante da comparação encontra-se normalizada numa etapa posterior.

$$P = \sum_{i=1}^M d_i + P_{M\acute{a}x} * (\max(ND_1, ND_2) - M) \quad (3)$$

Onde:

P - Penalização final

d_i - Distância entre os descritores

$P_{M\acute{a}x}$ - Penalização máxima, valor muito superior a qualquer distância possível

M - Número de correspondências

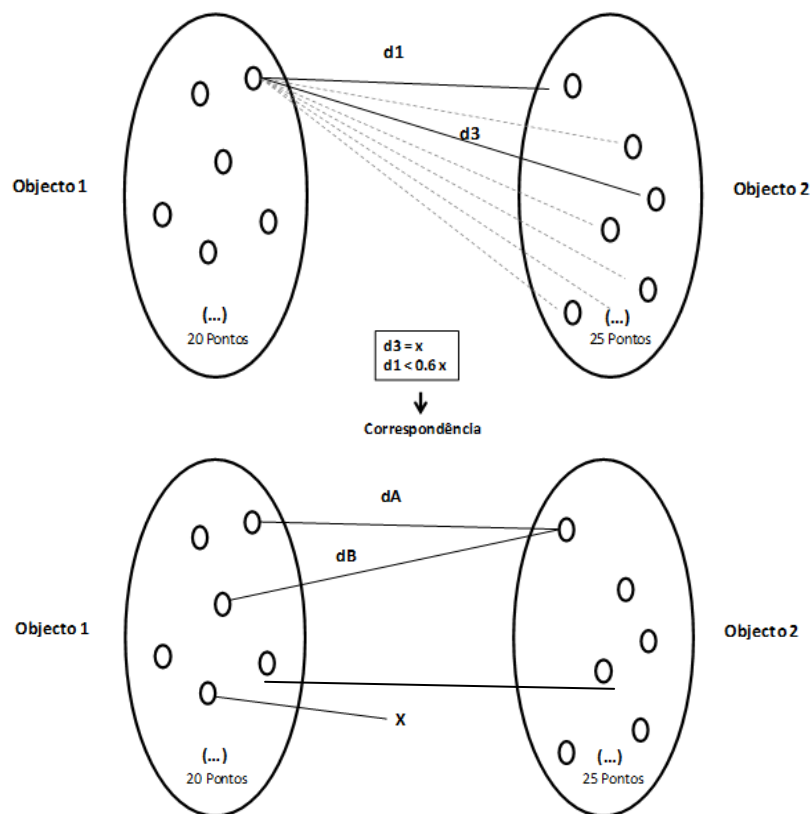


Figura 6.3 - 1 Etapa de comparação do SIFT

6.3.1 ESCALA AJUSTÁVEL

À semelhança do estudo dos dois modelos de descrição anteriores, no Gráfico 6.3.1 - 1 é apresentada a comparação da Pessoa C com um carro, Pessoa C com Pessoa D e pessoa C com Pessoa E na mesma imagem. Os dados mais relevantes do gráfico encontram-se na Tabela 6.3.1 - 1.

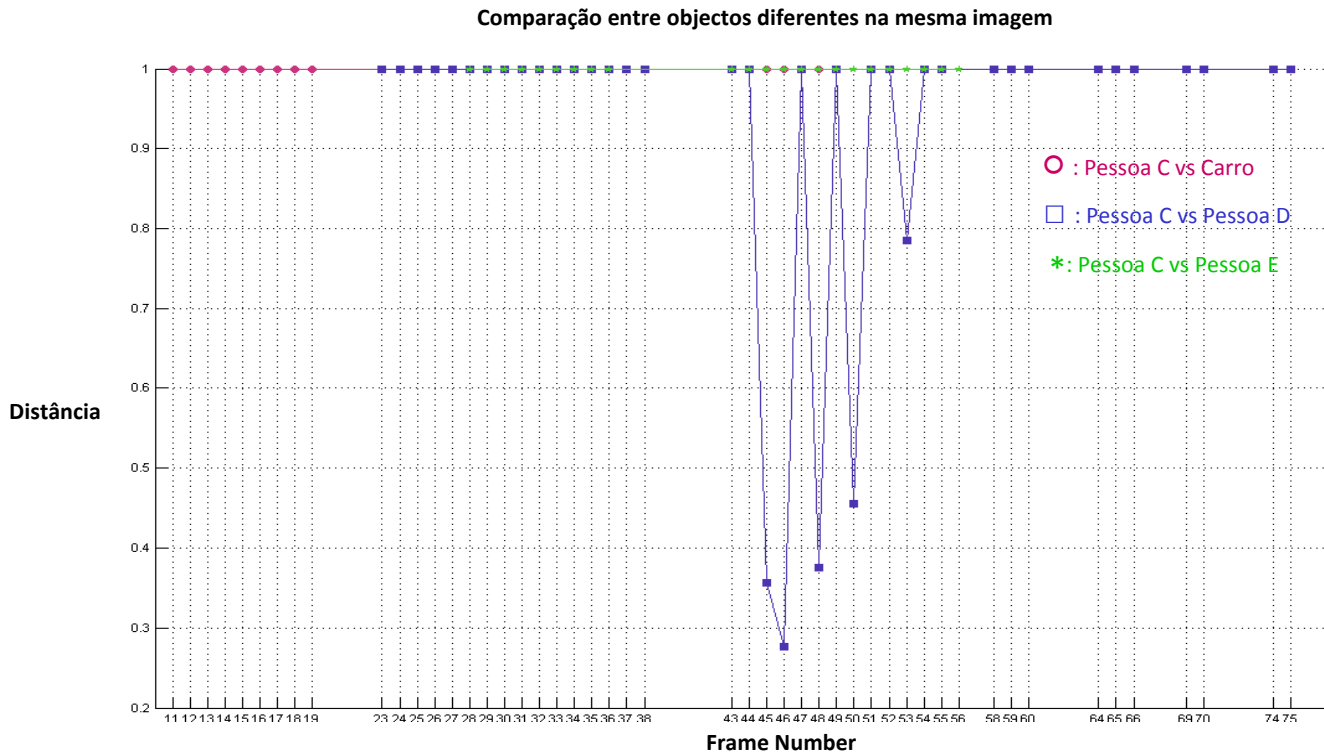


Gráfico 6.3.1 - 1 Comparação entre objectos diferentes na mesma imagem

Tabela 6.3.1 - 1 SIFT (escala ajustável): Dados obtidos para a situação a)

Escala	Modelo	Descrição	Média	Variância	Mínimo	Máximo
Ajustável	SIFT	Mesma Imagem				
		Pessoa C vs Carro	1	0	1	1
		Pessoa C vs Pessoa D	0.930	0.038	0.277	1
		Pessoa C vs Pessoa E	1	0	1	1

No gráfico verifica-se que os resultados obtidos para a comparação entre a Pessoa C e Pessoa D e para a comparação entre a Pessoa C e um carro encontram-se todos com o valor máximo, o que indica que poderá ter ocorrido um dos três possíveis casos: a) foram detectados pontos em ambas imagens mas não houve nenhuma correspondência; b) não foram detectados nenhuns pontos em ambas as imagens; c) foi apenas detectados pontos numa das imagens. Da depuração dos resultados obtidos verificou-se que o que aconteceu

deveu-se ao facto de apesar de terem sido detectados pontos em ambas as imagens, não ter havido nenhuma correspondência entre os mesmos. Verifica-se, também, que há uma maior similaridade entre a Pessoa C e a Pessoa D do que entre a Pessoa C e a Pessoa E, pelo facto de haver correspondência entre alguns dos pontos de interesse detectados aquando da sua comparação. De uma forma global, os resultados obtidos indicam que seria possível utilizar o modelo SIFT para a discriminação entre objectos.

O Gráfico 6.3.1 - 2 representará apenas os casos de comparação da mesma pessoa ao longo de uma sequência e entre pessoas distintas em imagens consecutivas. Os valores mais relevantes encontram-se na Tabela 6.3.1 - 2.

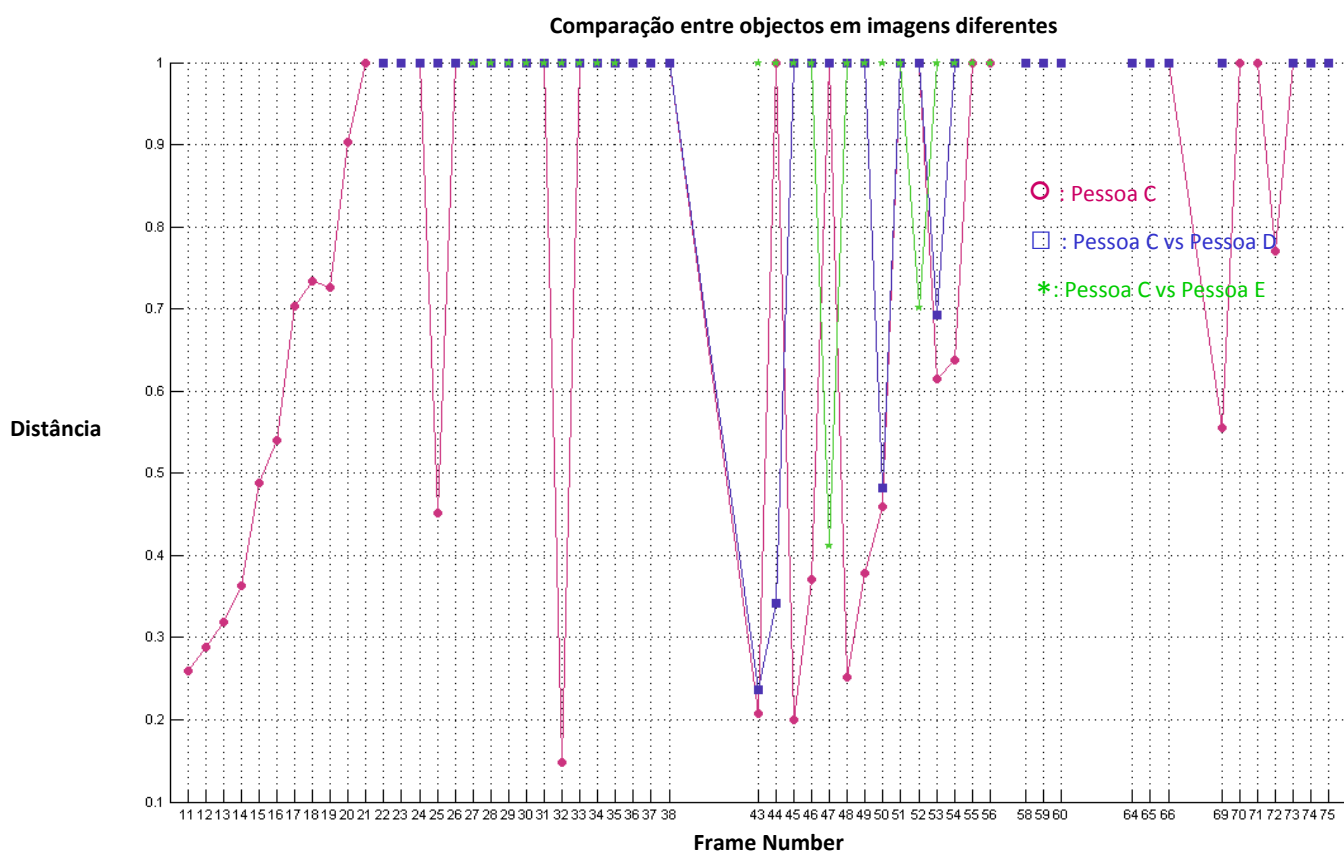


Gráfico 6.3.1 - 2 Comparação entre objectos em imagens diferentes

Tabela 6.3.1 - 2 SIFT (escala ajustável): Dados obtidos para a situação b) e c)

Escala	Modelo	Descrição	Média	Variância	Mínimo	Máximo
Ajustável	SIFT	Mesma Pessoa (C)	0.785	0.087	0.148	1
		Pessoas Diferentes (C vs D)	0.942	0.032	0.237	1
		Pessoas Diferentes (C vs E)	0.961	0.018	0.411	1

Analisando os valores obtidos, verifica-se que de uma forma geral é possível distinguir entre duas pessoas distintas, com excepção de alguns momentos em que ocorre uma correspondência parcial entre os pontos detectados nas imagens. Constatou-se que para a comparação da mesma pessoa ao longo de uma sequência apesar de existirem alguns momentos em que não há nenhuma correspondência entre os pontos detectados, de um modo global observa-se que poderá ser possível determinar que se trata da mesma pessoa, na medida em que se poderia utilizar para a etapa de seguimento, como no caso anterior, um algoritmo que permita restringir a área de pesquisa. Essa condição para o caso da comparação da mesma pessoa poderia resultar em detecção de pontos de interesse mais semelhantes, por estar a focar a pesquisa numa zona bastante semelhante à da imagem anterior, o que levaria a que os momentos em que se ilustram uma não correspondência passassem a possuir uma correspondência parcial ou total. Para o caso da comparação entre duas pessoas, a restrição da área de pesquisa, levaria a que não fossem detectados pontos de interesse noutras regiões e que levassem a uma falsa correspondência entre imagens.

6.3.2 ESCALA FIXA

No Gráfico 6.3.2 - 1 é apresentado a comparação entre objectos diferentes na mesma imagem. Os dados mais relevantes do gráfico encontram-se na Tabela 6.3.2 - 1.

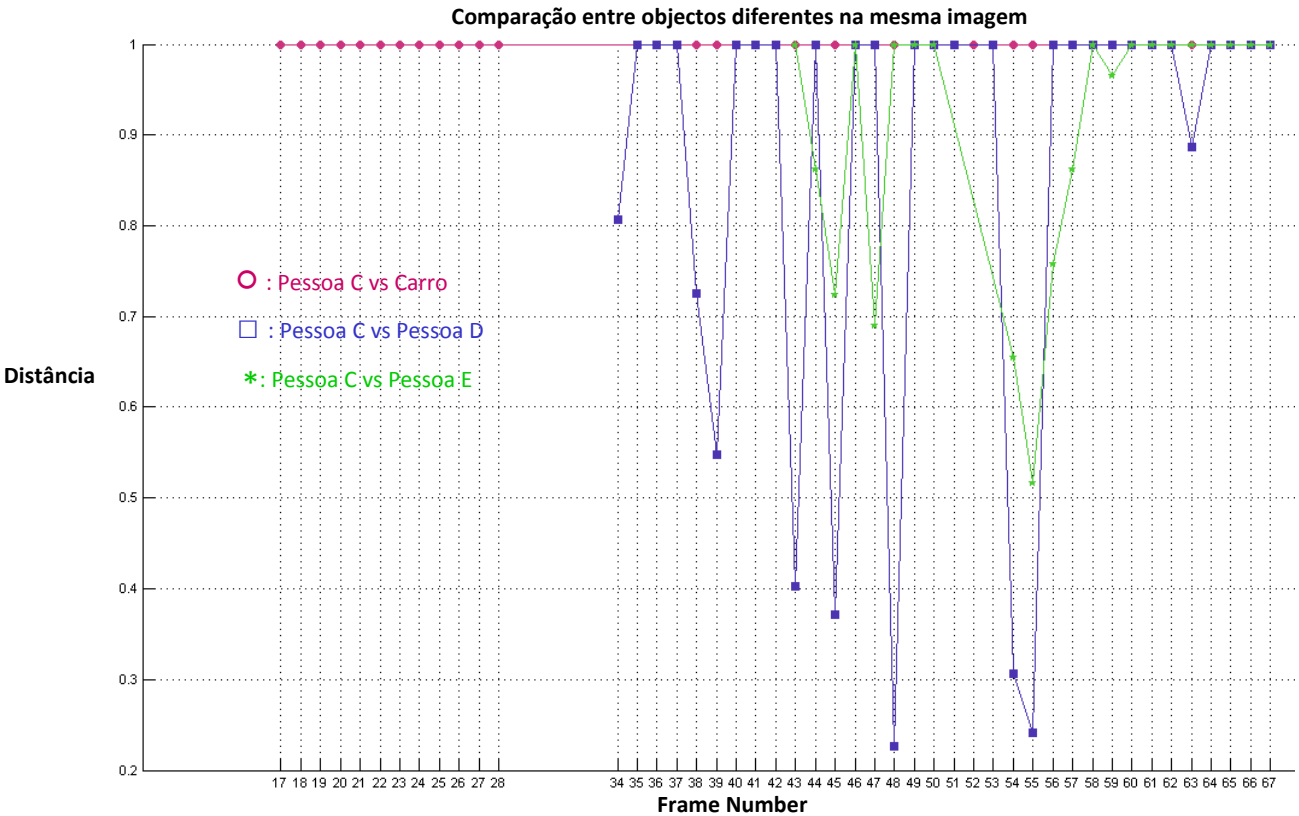


Gráfico 6.3.2 - 1 Comparação entre objectos diferentes na mesma imagem

Tabela 6.3.2 - 1 SIFT (escala fixa): Dados obtidos para a situação a)

Escala	Modelo	Descrição	Média	Variância	Mínimo	Máximo
2040	SIFT	Pessoa C vs Carro	1	0	1	1
		Pessoa C vs Pessoa D	0.864	0.066	0.225	1
		Pessoa C vs Pessoa E	0.911	0.021	0.517	1

Da mesma forma que as imagens captadas com uma escala de temperatura ajustável, com uma escala fixa é possível distinguir entre um carro e uma pessoa. O mesmo não acontece quando se é comparada duas pessoas diferentes, em que os resultados apresentados com o uso de uma escala fixa, são piores face aos obtidos com o uso de uma escala ajustável que permitiu distinguir entre duas pessoas distintas. Tal acontece devido ao facto da distribuição de cores ser mais homogênea com o uso de uma escala fixa, levando a que haja uma maior correspondência entre os pontos de interesse detectados em ambas as pessoas.

O gráfico a seguir só representará a comparação da mesma pessoa ao longo de uma sequência e a comparação entre pessoas distintas em imagens consecutivas. Os valores mais relevantes encontram-se na Tabela 6.3.2 - 2.

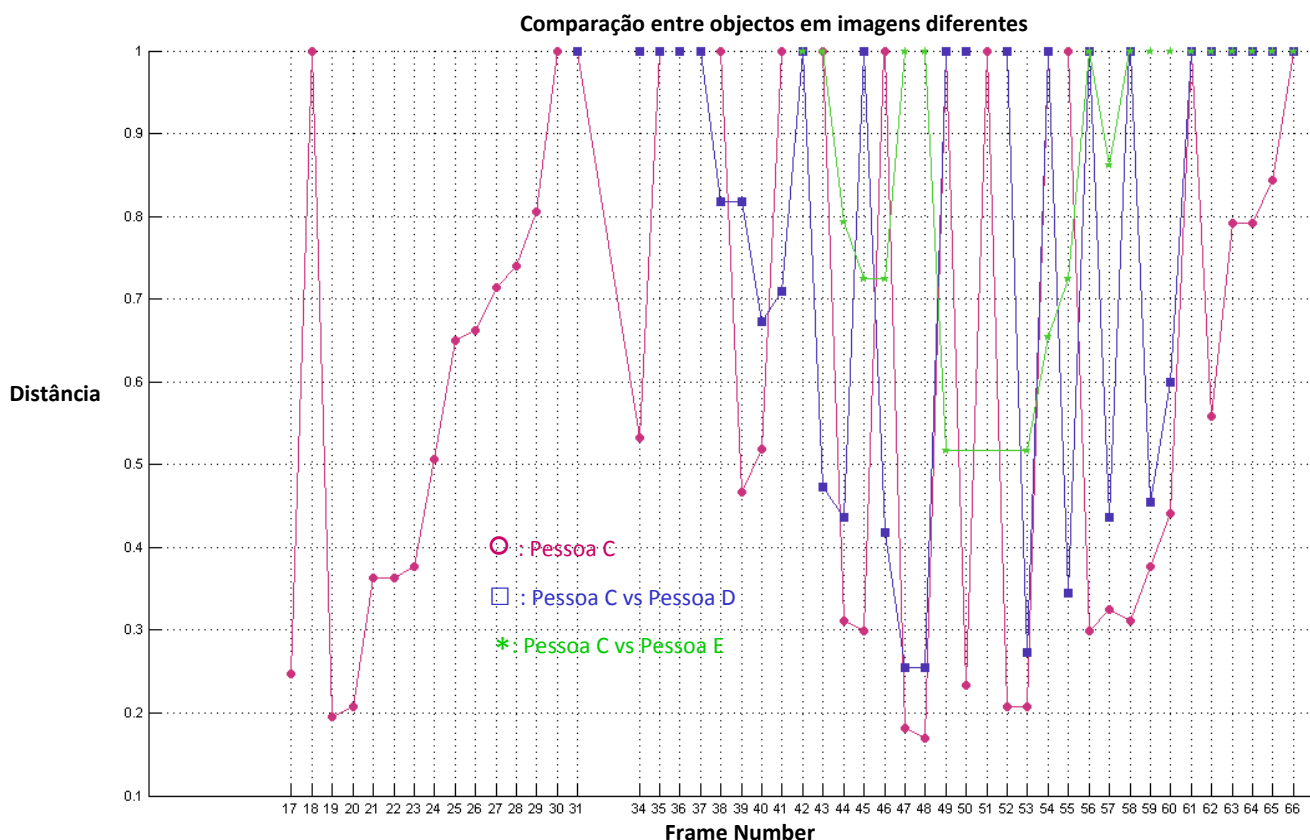


Gráfico 6.3.2 - 2 Comparação entre objectos em imagens diferentes

Tabela 6.3.2 - 2 SIFT (escala fixa): Dados obtidos para a situação b) e c)

Escala	Modelo	Descrição		Média	Variância	Mínimo	Máximo
2040	SIFT	Imagens Consecutivas	Mesma Pessoa (C)	0.640	0.102	0.168	1
			Pessoas Diferentes (C vs D)	0.787	0.079	0.255	1
			Pessoas Diferentes (C vs E)	0.987	0.028	0.517	1

Analisando os resultados obtidos, pode-se verificar que em comparação com os resultados obtidos utilizando uma escala ajustável, há uma maior correspondência entre os pontos detectados com uma escala fixa quer para a comparação entre duas pessoas distintas, quer para a comparação da mesma pessoa. Tal acontece devido ao facto da distribuição de cores ser mais homogénea com o uso de uma escala fixa, tornando, deste modo, mais difícil a tarefa de determinar se se trata da mesma pessoa ou não.

6.4 DESCRITOR: SURF

Speeded-Up Robust Features é um modelo de descrição que detecta um conjunto de pontos de interesse numa imagem, [BAY06]. Da mesma forma que o modelo SIFT, bons candidatos a pontos de interesse são os pontos que são invariantes às deformações locais aplicadas numa imagem. A comparação dos pontos é realizada do mesmo modo que no SIFT. A diferença entre o modelo SURF e o modelo SIFT reside no facto de o SURF ser um processo mais acelerado pelo facto de detectar menos pontos.

À semelhança do SIFT e do HOG, este modelo já se encontrava implementado, tendo sido apenas necessário acrescentar uma parte de código que permitisse receber as imagens e ler a localização dos objectos do ficheiro criado anteriormente.

6.4.1 ESCALA AJUSTÁVEL

No Gráfico 6.4.1 - 1 é apresentado a comparação da Pessoa C com um carro, Pessoa C com Pessoa D e pessoa C com Pessoa E na mesma imagem. Os dados mais relevantes do gráfico como a média de todos os valores, a variância, o valor mínimo e máximo registados encontram-se na Tabela 6.4.1 - 1.

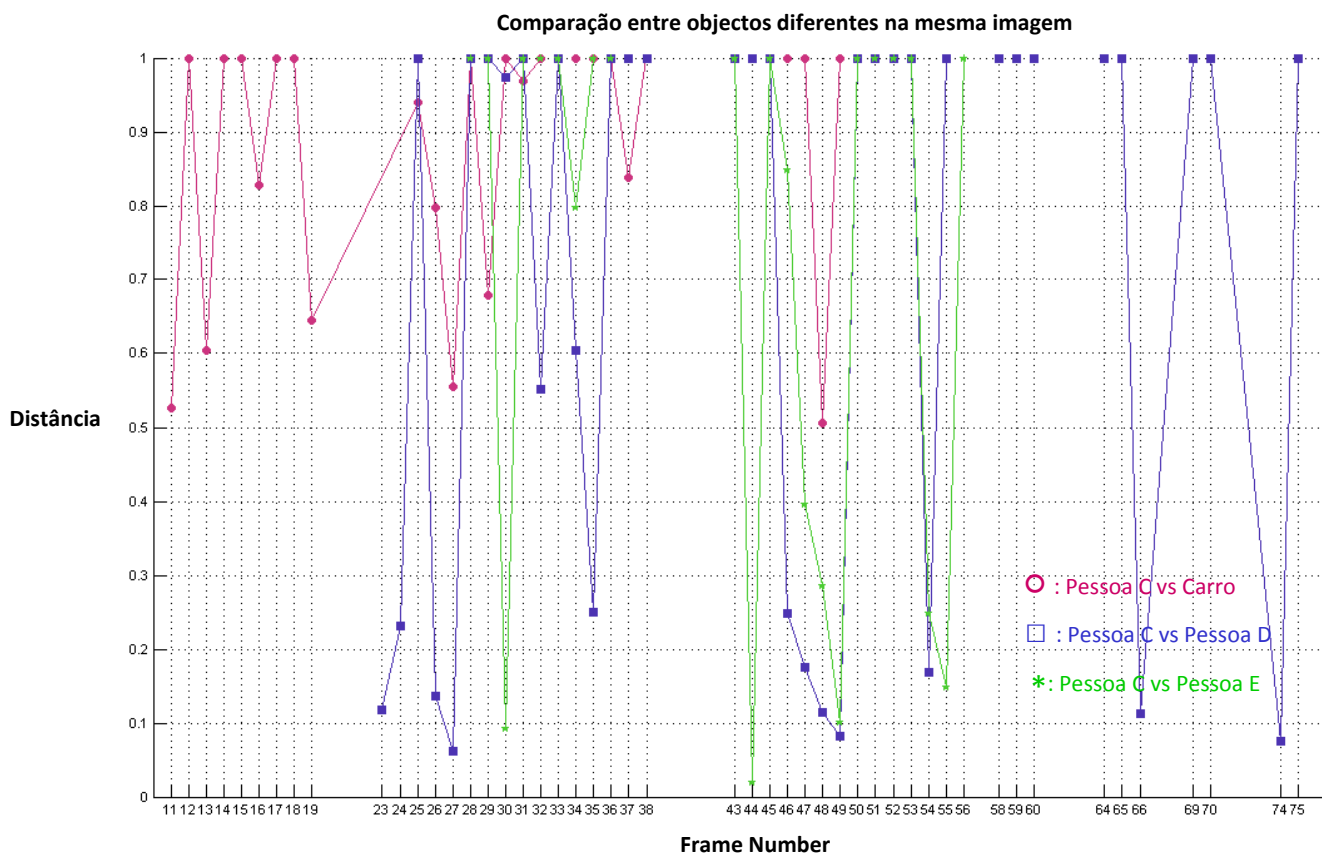


Gráfico 6.4.1 - 1 Comparação entre objectos diferentes na mesma imagem

Tabela 6.4.1 - 1 SURF (escala ajustável): Dados obtidos para a situação a)

Escala	Modelo	Descrição	Média	Variância	Mínimo	Máximo
Ajustável	SURF	Pessoa C vs Carro	0.896	0.029	0.505	1
		Pessoa C vs Pessoa D	0.716	0.157	0.062	1
		Pessoa C vs Pessoa E	0.649	0.273	-1	1

Analisando os valores obtidos e a representação gráfica, pode-se observar que com o modelo SURF, a distinção entre uma pessoa e um carro não é tão evidente, havendo uma correspondência parcial entre os pontos detectados num carro e numa pessoa. De uma forma geral, verifica-se que não é possível distinguir entre duas pessoas distintas, na medida em que há, em certos momentos, correspondência quase total dos pontos detectados nas duas imagens que indica que há uma grande similaridade. Da mesma forma que no modelo SIFT, o valor máximo indica três possíveis casos: a) foram detectados pontos em ambas imagens mas não houve nenhuma correspondência; b) não foram detectados nenhuns pontos em ambas as imagens; c) foi apenas detectados pontos numa das imagens. Tendo-se

verificado que houve apenas uma situação em que não houve extracção de pontos em ambas as imagens, que foi aquando da comparação da Pessoa C na imagem 35 com a Pessoa E na imagem 36.

O Gráfico 6.4.1 - 2 contempla os casos de comparação da mesma pessoa ao longo de uma sequência e a comparação entre pessoas distintas em imagens consecutivas. Os valores mais relevantes encontram-se na Tabela 6.4.1 - 2.

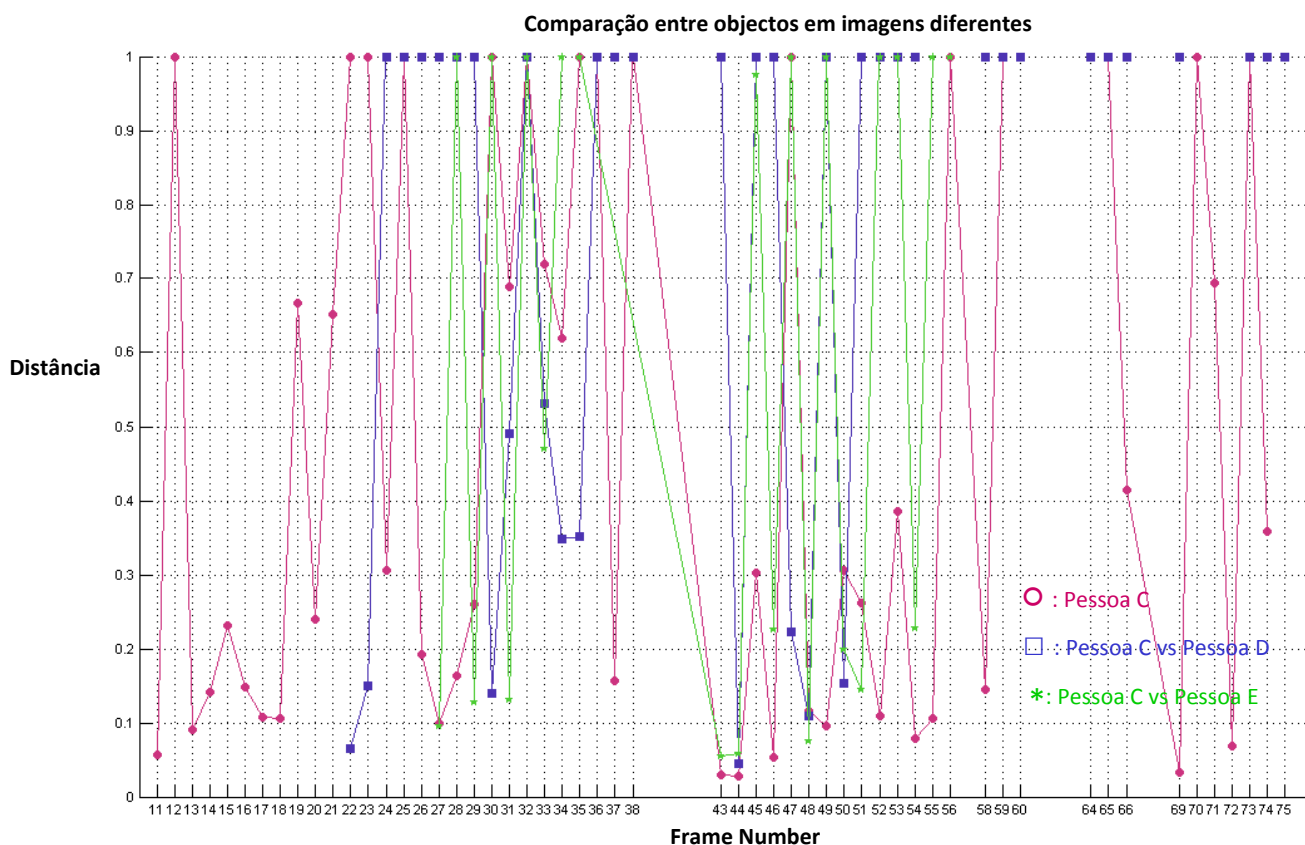


Gráfico 6.4.1 - 2 Comparação entre objectos em imagens diferentes

Tabela 6.4.1 - 2 SURF (escala ajustável): Dados obtidos para a situação b) e c)

Escala	Modelo	Descrição	Média	Variância	Mínimo	Máximo
Ajustável	SURF	Mesma Pessoa (C)	0.486	0.155	0.028	1
		Pessoas Diferentes (C vs D)	0.734	0.208	-1	1
		Pessoas Diferentes (C vs E)	0.599	0.187	0.055	1

Da análise do gráfico, pode-se verificar que com o modelo SURF a determinação de ser ou não a mesma pessoa torna-se uma tarefa mais difícil na medida em que ocorre uma maior sobreposição do intervalo dos valores em que houve uma correspondência dos pontos detectados pertencentes à comparação da mesma pessoa ao longo de uma sequência com os valores representativos da comparação de duas pessoas distintas. À semelhança do gráfico anterior, neste estudo, aquando da comparação entre duas pessoas distintas (Pessoa C e Pessoa D) ocorreu na imagem 34 e imagem 35 um caso em que não houve extracção de pontos nas duas imagens.

6.4.2 ESCALA FIXA

No Gráfico 6.4.2 - 1 é apresentado a comparação da Pessoa C com um carro, Pessoa C com Pessoa D e pessoa C com Pessoa E na mesma imagem. Os dados mais relevantes do gráfico como a média de todos os valores, a variância, o valor mínimo e máximo registados encontram-se na Tabela 6.4.2 - 1.

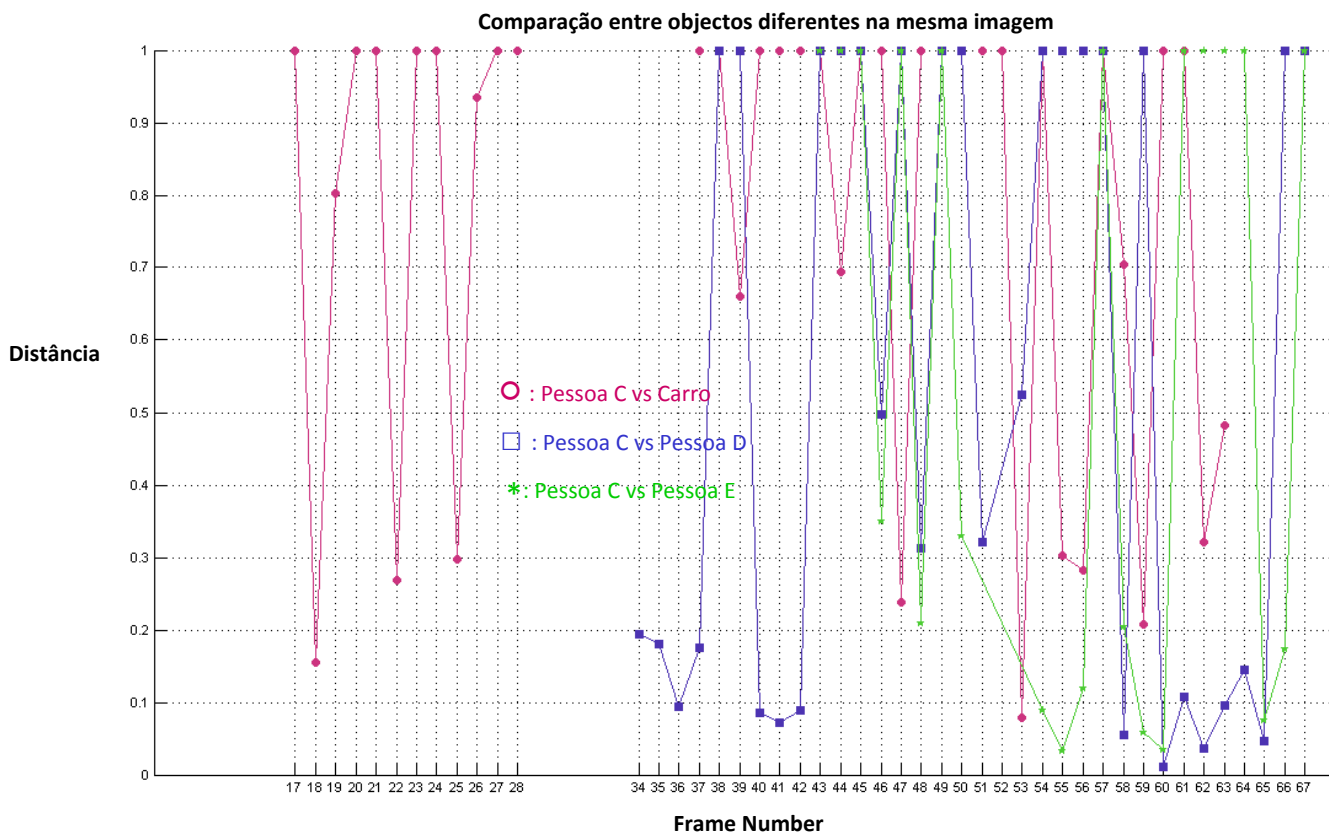


Gráfico 6.4.2 - 1 Comparação entre objectos na mesma imagem

Tabela 6.4.2 - 1 SURF (escala fixa): Dados obtidos para a situação a)

Escala	Modelo	Descrição	Média	Variância	Mínimo	Máximo
2040	SURF	Pessoa C vs Carro	0.780	0.105	0.078	1
		Pessoa C vs Pessoa D	0.547	0.188	0.011	1
		Pessoa C vs Pessoa E	0.576	0.194	0.032	1

Analisando o gráfico correspondente à comparação de objectos diferentes na mesma imagem com o modelo de descrição SURF com imagens captadas com uma escala fixa, podemos observar que não é possível distinguir um carro de uma pessoa e entre duas pessoas distintas apesar de existirem momentos em que o resultado da comparação entre os objectos indicar que não há uma correspondência entre os pontos detectados nas imagens a serem analisadas.

No Gráfico 6.4.2 - 2 só irão ser contemplados os casos de comparação da mesma pessoa ao longo de uma sequência e a comparação entre pessoas distintas em imagens consecutivas. Os valores mais relevantes encontram-se na Tabela 6.4.2 - 2.

Comparação entre objectos em imagens diferentes

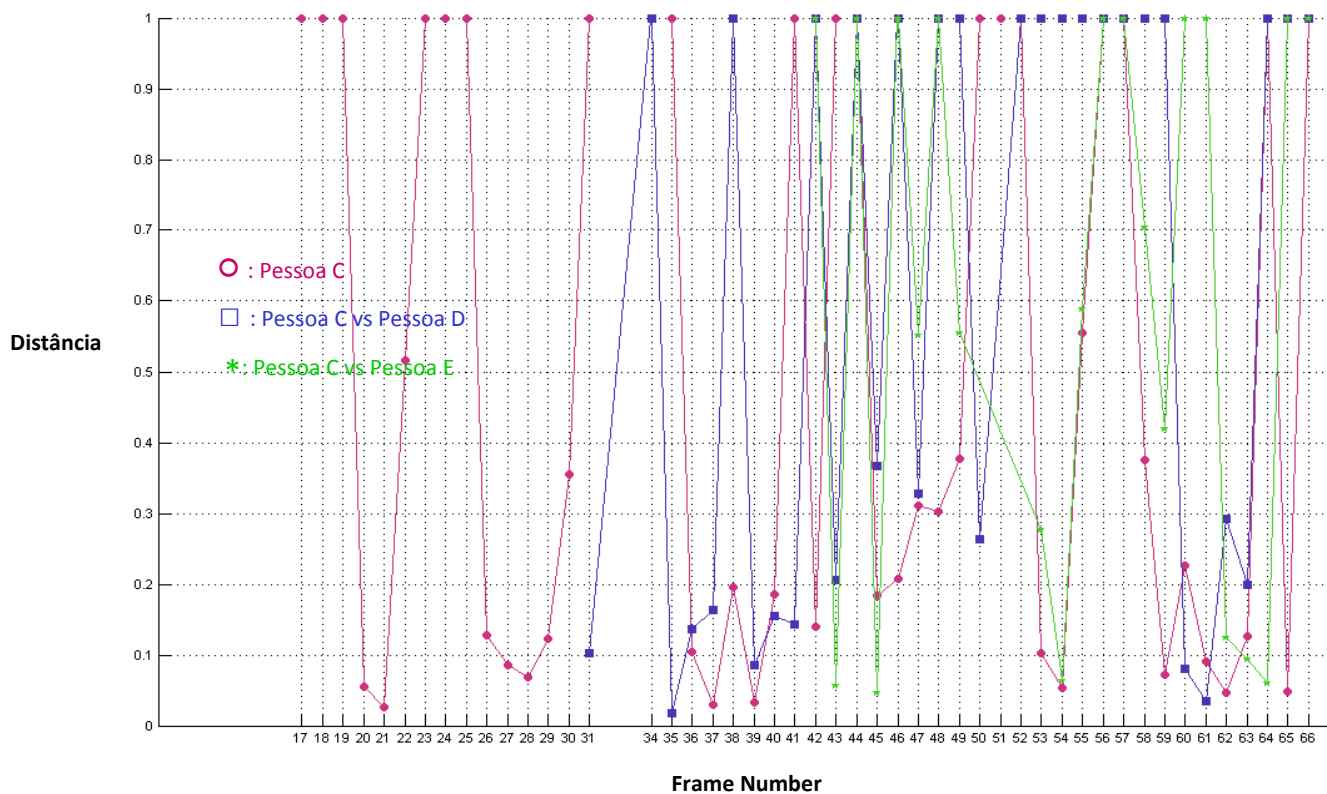


Gráfico 6.4.2 - 2 Comparação entre objectos em imagens diferentes

Tabela 6.4.2 - 2 SURF (escala fixa): Dados obtidos para a situação b) e c)

Escala	Modelo	Descrição		Média	Variância	Mínimo	Máximo
2040	SURF	Imagens Consecutivas	Mesma Pessoa (C)	0.502	0.178	0.026	1
			Pessoas Diferentes (C vs D)	0.623	0.180	0.017	1
			Pessoas Diferentes (C vs E)	0.616	0.162	0.047	1

Da análise dos resultados obtidos, pode-se constatar que com o uso de uma escala fixa não é possível inferir que se trata da mesma pessoa ao longo de uma sequência na medida em que ocorre uma sobreposição dos valores dos intervalos referentes à comparação entre duas pessoas distintas e a comparação da mesma pessoa.

6.5 CONCLUSÕES DECORRENTES DA ANÁLISE EFECTUADA

Terminado o estudo dos quatro modelos de descrição (Histograma de Cor, HOG, SIFT e SURF) com as imagens captadas com uma escala ajustável e com uma escala fixa, pôde-se constatar que com a aplicação do modelo de descrição baseado no histograma de cor é possível distinguir entre um carro e uma pessoa, sendo essa distinção maior quando é utilizada a escala ajustável em que a pessoa aparece mais destacada do que com a utilização da escala fixa, visto que a distribuição de cores apresentada pelo carro é bastante semelhante à de uma pessoa, como se pode observar na Figura 6.5 - 1.



Figura 6.5 - 1 Da esquerda para à direita: Escala ajustável e escala fixa

Verificou-se, também, que este modelo é afectado pela posição da pessoa na imagem, ou seja, se a pessoa se encontra de lado, numa posição de frente ou de costas. O facto de a *frame rate* não ser constante implica que ocorra situações de variações abruptas na distribuição da cor do objecto o que leva a que este modelo de descrição não obtenha

melhores resultados em determinar se se trata da mesma pessoa ou não. Exemplos destes acontecimentos encontram-se na Figura 6.5 - 2.

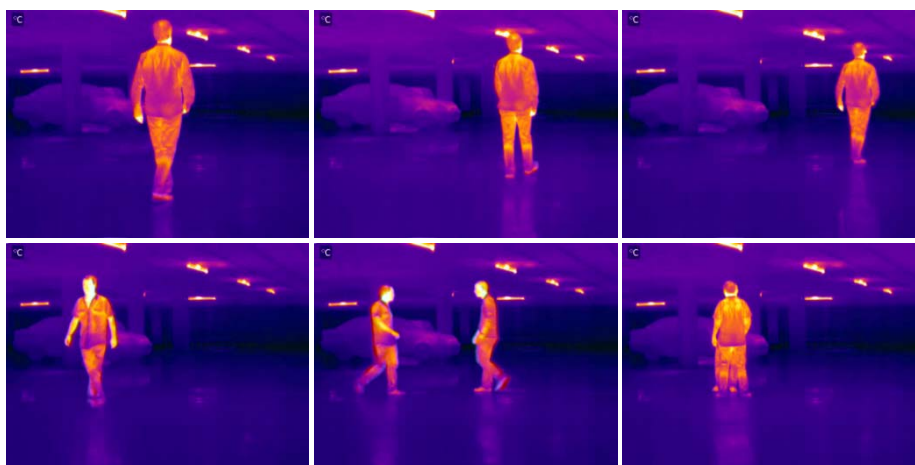


Figura 6.5 - 2 Imagens superiores: Ilustração de momentos em que a *frame rate* é inconsistente; Imagens inferiores: Ilustração das várias posições assumidas pelo mesmo objecto ao longo de uma sequência

Com o uso do modelo de descrição baseado nos histogramas de gradientes verificou-se que quer para as imagens captadas com uma escala ajustável quer para as imagens captadas com uma escala fixa, foi possível distinguir entre uma pessoa e um carro. No entanto, para a distinção entre duas pessoas diferentes, constatou-se que tal não era possível. À semelhança do modelo baseado no histograma de cor, o HOG também é afectado pela posição da pessoa na imagem, mas não é tão afectado pela *frame rate* inconsistente que provoca variações abruptas nos objectos.

Analisando os modelos de descrição SIFT e SURF, constatou-se que apesar de os resultados não serem afectados pela posição da pessoa (de frente, de lado ou de costas) ou pela *frame rate* inconsistente, com o primeiro modelo é possível distinguir entre os vários objectos que se encontram na imagem. O mesmo não acontece com o modelo SURF pelo facto de serem detectados menos pontos, o que leva a que possua menos informação de comparação face ao modelo SIFT.

7. Conclusão e Desenvolvimentos Futuros

Neste capítulo faz-se o balanço do trabalho realizado face aos objectivos inicialmente propostos e sugerem-se eventuais desenvolvimentos futuros para o estudo das imagens termográficas aplicadas à área da vigilância.

7.1 CONCLUSÃO

Com este trabalho realizou-se um estudo sobre a detecção e seguimento de objectos utilizando imagens termográficas, tendo sido dado maior ênfase na análise experimental de modelos de descrição (Histograma de cor, HOG, SIFT e SURF).

Na primeira etapa em que foi feito um estudo das propriedades das câmaras termográficas, assim como das propriedades das imagens termográficas concluiu-se que com o uso de uma escala ajustável seria possível descrever melhor os elementos mais relevantes como uma pessoa do que com uma escala fixa. Contudo, foram captadas imagens com uma escala fixa para verificar se haveria mais alguma vantagem para além da de poder continuar a descrever uma pessoa da mesma forma no caso de aparecer um objecto que possuísse uma temperatura superior à do corpo humano. O que não aconteceria se fosse utilizada uma escala ajustável, uma vez que a pessoa apareceria menos destacada face aos elementos do *background* e o destaque seria dado ao objecto que possuísse a temperatura maior. Outra conclusão que se pode tirar prende-se com o facto de com a câmara termográfica adquirida, a *frame rate* não é constante, o que origina movimentos abruptos dos objectos entre imagens consecutivas.

A segunda etapa consistiu na caracterização de sequências típicas num cenário de vigilância devido ao facto de não existirem sequências típicas que poderiam ter sido utilizadas para a prossecução do estudo. Após a análise das imagens captadas com as duas escalas de temperatura nos dois corredores com características diferentes e na garagem, constatou-se que devido ao material utilizado no pavimento dos corredores, ocorria o efeito de reflexão da pessoa no chão, assim como nas zonas que eram revestidas por vidro num dos corredores. A reflexão captada no chão da garagem, face à dos corredores não era tão nítida, o que se deveu ao material utilizado para revestir o pavimento (cimento). Para efeitos de estudo optou-se pelas sequências gravadas na garagem com ambas as escalas (ajustável e fixa) por não sofrerem tanta influência do efeito de reflexo presente.

Na terceira etapa foi realizada uma análise de diferentes técnicas para a detecção e seguimento de pessoas, tendo sido desenvolvido um algoritmo que permitiu segmentar os elementos considerados como sendo de interesse, nomeadamente, as pessoas presentes num cenário. Na mesma etapa foi realizada uma comparação com outros dois algoritmos baseados em movimento e concluiu-se que para as imagens termográficas os algoritmos baseados em movimento não eram tão eficazes como o algoritmo baseado num valor de *threshold*.

Na última etapa analisaram-se e compararam-se os diferentes modelos de descrição de objectos de forma a determinar qual deles seria o melhor a diferenciar os vários elementos entre si. Da comparação dos quatro modelos de descrição, pode-se concluir que o melhor modelo para descrever objectos é o modelo SIFT. Deste estudo, conclui-se também que só utilizando o modelo de descrição não é o suficiente para distinguir entre os vários objectos, sendo necessário utilizar na etapa de seguimento um algoritmo capaz de restringir a área de pesquisa para a zonas semelhantes nas duas imagens para que possam ser detectados os pontos de interesse e devidamente comparados.

Relativamente ao uso de uma escala fixa, verificou-se que não traz benefícios, na medida que com certos modelos, o resultado apresentado para distinguir entre objectos diferentes era pior face aos apresentados com o uso de uma escala ajustável. Deste modo, conclui-se que todos os objectivos inicialmente propostos foram cumpridos.

7.2 DESENVOLVIMENTOS FUTUROS

Perspectivam-se como desenvolvimentos futuros a implementação de um valor de *threshold* adaptativo no algoritmo de segmentação, dado ter-se constatado que a utilização de uma escala fixa não trazer benefícios. Com esta implementação seria possível no uso de uma escala ajustável segmentar os elementos de interesse (pessoas) mesmo que apareça na imagem um elemento que possua uma temperatura superior à de um ser humano, visto a câmara termográfica reajustar os valores com base na temperatura maior captada e que com um valor de *threshold* fixo, não era possível segmentar a pessoa, por esta não se encontrar destacada face ao *background*. Para tal, obter-se-ia como *output* da câmara a imagem e o valor mínimo e máximo atribuído para a escala de temperatura ajustável em cada instante.

Devido ao facto de a *frame rate* não ser constante, verificou-se nas imagens captadas que entre imagens consecutivas os movimentos realizados pelas pessoas nem sempre eram suaves, havendo movimentos abruptos que originam a que sejam considerados nos resultados erros associados à segmentação. Para anular esse efeito seria necessário adquirir uma câmara termográfica que permitisse gravar com uma *frame rate* superior e voltar a realizar os mesmos testes para concluir se as falhas do algoritmo de segmentação se mantêm e se os resultados com a aplicação dos modelos de descrição melhoram. Deste modo, seria possível concluir se se justificaria a aquisição de uma câmara termográfica mais dispendiosa.

Como foi possível constatar, aplicando os quatro modelos de descrição o caso mais difícil em que todos os modelos tiveram dificuldade foi o de distinguir entre duas pessoas distintas em imagens consecutivas. Tal deve-se ao facto de as imagens termográficas não possuírem as melhores características que possam ser utilizadas como um factor de decisão, tais como associação de cores diferentes a pessoas distintas, como nas se verifica nas imagens naturais em que uma pessoa utiliza uma camisola com um tom branco e uma outra com um tom verde. Como tal, será interessante num trabalho futuro estudar a utilização conjunta de informação proveniente de imagens termográficas e imagens naturais. Por exemplo, utilizar a informação das imagens termográficas para realizar a segmentação dos objectos dado este tipo de imagens não serem tão susceptíveis a condições ambientais adversas e nada

influenciáveis pela condição luminosa do local, como foi verificado. E utilizar a informação extraída das imagens naturais para a etapa de descrição dos objectos por possuírem características diferentes das imagens termográficas, como por exemplo: seria mais fácil distinguir entre uma pessoa com uma camisola vermelha de uma pessoa com uma camisola azul do que entre duas pessoas que na zona da camisola possuam tons entre o amarelo, laranja e vermelho, que se devem à distribuição da temperatura corporal registada pela câmara termográfica e, conseqüentemente, apresentada nas imagens termográficas.

Apesar de se ter verificado que o modelo SIFT foi o que apresentou, de uma forma geral, o melhor desempenho, o seu custo computacional torna-se uma grande desvantagem para a utilização deste modelo de descrição. Como alternativa, uma vez que o SURF também faz uso de um conjunto de pontos de interesse para a representação de objectos e possui um custo computacional inferior que é devido ao facto de detectar menos pontos ou por vezes nenhum ponto o que leva a ter resultados piores face ao SIFT, seria interessante complementar a forma de como os pontos de interesse são detectados no modelo SURF. Ou seja, implementar uma grelha que force a detecção de mais pontos para além dos que já são detectados. Com esta implementação contornava-se os dois problemas do SURF. O primeiro problema estaria associado à não detecção de pontos nas imagens que induzia na comparação da mesma pessoa a inferir que não haveria correspondência e, conseqüentemente, que não seria a mesma pessoa. O segundo problema é a detecção de poucos pontos, o que implica pouca informação de comparação, com a aplicação de uma grelha, a informação extraída para a comparação seria maior, o que poderá levar a menos erros no processo de correspondência.

Referências Documentais

- [BAS06] BASHIR, Faisal; PORIKLI, Faith - *Performance Evaluation of Object Detection and Tracking Systems. IEEE International Workshop on Performance Evaluation of Tracking and Surveillance (PETS)*, June 2006 (PETS 2006)
- [BAY06] BAY, Herbert; TUYTELAARS, Tinne; VAN GOOL, Luc - *SURF: Speeded Up Robust Features. Springer-Verlag Berlin Heidelberg 2006*
- [CAR05] CARDOSO, Jaime; CORTE-REAL, Luís – *Toward a Generic Evaluation of Image Segmentation. IEEE Transactions on Image Processing, vol. 14, issue 11, pp. 1773-1782*
- [CAR09] CARDOSO, Jaime; CARVALHO, Pedro; TEIXEIRA, Luís F.; CORTE-REAL, Luís - *Partition-distance methods for assessing spatial segmentations of images and videos. Computer Vision and Image Understanding 113 (2009) pp. 811-823*
- [CAR10] CARVALHO, Pedro; CARDOSO, Jaime; CORTE-REAL, Luís - *Hybrid Framework for evaluating video object tracking algorithms. Electronics Letters - ELECTRON LETT, vol. 46, no. 6, pp 411-412, 2010*
- [DAL05] DALAL, Navneet; TRIGGS, Bill - *Histograms of Gradients for Human Detection. Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on, Vol. 1 (2005), pp. 886-893*
- [ERD04] ERDEM, Cigdem Eroglu; SANKUR, Bulent; TEKALP, A. Murat - *Performance Measures for Video Object Segmentation and Tracking. IEEE Transactions on Image Processing, Vol. 13, No. 7, pp. 937 - 951 July 2004.*
- [FER10] FERNÁNDEZ-CABALLERO, Antonio; CASTILLO, José Carlos; MARTÍNEZ-CANTOS, Javier; MARTINÉZ-TOMÁS, Rafael - *Optical flow or image subtraction in human detection from infrared camera on mobile robot. Journal Robotics and Autonomous Systems. Volume 58, Issue 12, December 2010.*

- [FER11] FERNÁNDEZ-CABALLERO, Antonio; CASTILLO, José Carlos; SERRANO-CUERDA, Juan; MALDONADO-BASCÓN, Saturnino - *Real-time huma segmentation in infrared vídeos. Expert Systems with Applications. Volume 38, Issue 3, March 2011, pp. 2577-2584.*
- [FOR02] FORSYTH, D.A.; PONCE, J. - *Computer Vision: A Modern Approach. Prentice Hall, August 2002*
- [FOX03] FOX, Dieter; HIGHTOWER, Jeffrey; LIAO, Lin; SCHULZ, Dirk; BORRIELLO, Gaetano - *Bayesian Filtering for Location Estimation. IEEE Pervasive Computing, Vol. 2, No. 3. (July 2003), pp. 24-33*
- [GAV98] GAVRILA, D. M. - *The visual Analysis of Human movement: A Survey. Computer Vision and Image Understanding: Vol. 73, No. 1, January, pp. 82–98, 1999.*
- [INT99] INTEL - *Open Source Computer Vision Library, Reference Manual. 1999-2001 Intel Corporation*
- [KAN81] KANADE, Takeo; LUCAS, Bruce D. - *An Iterative Image Registration Technique with an Application to Stereo Vision. Proceedings DARPA Image Understanding Workshop, April 1981, pp. 121-130*
- [KRO08] KROTOSKY, Stephen J.; TRIVEDI, Mohan Manubhai - *Person surveillance using visual and infrared imagery. IEEE transactions on circuits and systems for video technology, vol. 18, no. 8, August, 2008.*
- [KUM06] KUMAR, Praveen; MITTAL, Ankush, KUMAR, Padam - *Fusion of Thermal Infrared and Visible Spectrum Video for Robust Surveillance. Indian Conference on Computer Vision, Graphics & Image Processing - ICVGIP , pp. 528-539, 2006*
- [KUM10] KUMAR, Praveen; MITTAL, Ankush, KUMAR, Padam - *Addressing uncertainty in multi-modal fusion for improved object detection in dynamic environment. Information Fusion. Volume 11, Issue 4, October 2010, pp. 311-324*

- [LEY08] LEYKIN, Alex; HAMMOUD, Riad - *Pedestrian tracking by fusion of thermal-visible surveillance. Machine vision and applications (2010) 21: pp.587-595, 10 October, 2008.*
- [LOW04] LOWE, David G. – *Distinctive Image Features from Scale-Invariant Keypoints. International Journal of Computer Vision In International Journal of Computer Vision, Vol. 60, No. 2. (1 November 2004), pp. 91-110*
- [MOE00] MOESLUND, Thomas B.; GRANUM, Erik - *A Survey of Computer Vision-Based Human Motion Capture. Computer Vision and Image Understanding 81, pp.231–268*
- [OMA10] O’MALLEY, Ronan; JONES, Edward; GLAVIN, Martin - *Detection of pedestrians in far-infrared automotive night vision using region-growing and clothing distortion compensation. Infrared Physics & Technology. Volume 53, Issue 6, November 2010, pp. 439-449.*
- [PAG01] Página de Internet – data de último acesso: 8 de Junho de 2011,
http://www.flir.com/uploadedFiles/Thermography_APAC/Products/Product_Literture/100930%20A300%20datasheet_en.pdf
- [PAG02] Página de Internet – data de último acesso: 12 de Junho de 2011,
<http://www.eecs.berkeley.edu/Research/Projects/CS/vision/grouping/segbench/>
- [SHA00] SHAPIRO, Linda; STOCKAMN, George - *Computer Vision, March 2000*
- [TEI07] TEIXEIRA, Luís F.; CARDOSO, Jaime; CORTE-REAL, Luís - *Object Segmentation Using Background Modelling and Cascaded Change Detection. Journal of Multimedia, Vol2, No.5, September 2007.*
- [TRE06] TREPTOW, André; CIELNIAK, Grzegorz, DUCKETT, Tom - *Real-time people tracking for mobile robots using thermal vision. Robotics and Autonomous Systems. Volume 54, Issue 9, 30 September 2006, pp. 729-739*

- [VEN09] VENETIANER, Péter L.; DENG, Hongli - *Performance evaluation of an intelligent video surveillance system, A case study. Computer vision and image understanding. Volume 114, Issue 11, 30 November 2010*
- [VIO01] VIOLA, Paul; JONES, Michael J. - *Robust real-time object detection. IEEE Workshop on Statical and Theories of Computer Vision, 2001.*
- [WEL06] WELCH, Greg; BISHOP, Gary - *An Introduction to the Kalman Filter. Department of computer Science, University Of North Carolina at Chapel Hill. July 24, 2006.*
- [YIL06] YILMAZ, A.; JAVED, O.; SHAH, M. - *Object tracking: A survey. ACM Comput. Surv. 38, 4, Article 13 (Dec. 2006), 45 pages.*