

**UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ**  
**DIRETORIA DE PESQUISA E PÓS-GRADUAÇÃO**  
**PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO**

**FELIPE BUENO**

**MÉTODO PARA MAPEAMENTO TRIDIMENSIONAL TÉRMICO**  
**APLICADO À SISTEMAS DE AUXÍLIO AO CONDUTOR**

**DISSERTAÇÃO**

**PONTA GROSSA**

**2018**

**FELIPE BUENO**

**MÉTODO PARA MAPEAMENTO TRIDIMENSIONAL TÉRMICO  
APLICADO À SISTEMAS DE AUXÍLIO AO CONDUTOR**

Dissertação apresentada como requisito parcial à obtenção do título de Mestre em Ciência da Computação do Programa de Pós-Graduação em Ciência da Computação da Universidade Tecnológica Federal do Paraná - Câmpus Ponta Grossa. Área de Concentração: Sistemas e Métodos de Computação

Orientador: Prof. Dr. Erikson Freitas de Moraes

Coorientador: Prof. Dr. Max Mauro Dias Santos

**PONTA GROSSA**

**2018**

Ficha catalográfica elaborada pelo Departamento de Biblioteca  
da Universidade Tecnológica Federal do Paraná, Campus Ponta Grossa  
n.02/19

B928 Bueno, Felipe

Método para mapeamento tridimensional térmico aplicado à sistemas de auxílio ao  
condutor. / Felipe Bueno. 2018.

97 f.; il. 30 cm

Orientador: Prof. Dr. Erikson Freitas de Moraes

Coorientador: Prof. Dr. Max Mauro Dias Santos

Dissertação (Mestrado em Ciência da Computação) - Programa de Pós-  
Graduação em Ciência da Computação. Universidade Tecnológica Federal do Paraná,  
Ponta Grossa, 2018.

1. Motoristas. 2. Visão estereoscópica. 3. Visão por computador. 4. Termologia. 5.  
Algoritmos genéticos. I. Moraes, Erikson Freitas de. II. Santos, Max Mauro Dias. III.  
Universidade Tecnológica Federal do Paraná. IV. Título.

CDD 004



**Ministério da Educação**  
**Universidade Tecnológica Federal do Paraná**  
**Câmpus Ponta Grossa**  
Diretoria de Pesquisa e Pós-Graduação  
Programa de Pós-Graduação em Ciência da Computação



## **FOLHA DE APROVAÇÃO**

Título de Dissertação Nº 9/2018

### **MÉTODO PARA MAPEAMENTO TRIDIMENSIONAL TÉRMICO APLICADO À SISTEMAS DE AUXÍLIO AO CONDUTOR**

Por

**Felipe Bueno**

Esta dissertação foi apresentada às **14:30 horas de 18 de dezembro de 2018**, na sala da **DIREC**, como requisito parcial para a obtenção do título de MESTRE EM CIÊNCIA DA COMPUTAÇÃO, Programa de Pós-Graduação em Ciência da Computação. O candidato foi arguido pela Banca Examinadora, composta pelos professores abaixo assinados. Após deliberação, a Banca Examinadora considerou o trabalho APROVADO.

---

**Prof. Dr. Luciano José Senger (UEPG)**

---

**Prof. Dr. Ionildo José Sanches  
(UTFPR)**

---

**Prof. Dr. Erikson Freitas de Moraes  
(UTFPR)**

*Orientador e presidente da banca*



Visto da Coordenadora:

---

**Profª. Drª. Sheila Moraes de Almeida**  
Coordenadora do PPGCC  
UTFPR - Câmpus Ponta Grossa

- A Folha de Aprovação assinada encontra-se na Coordenação do Programa -

Dedico este trabalho à Deus, que me  
guia, sustenta e molda através das  
alegrias e lutas desta vida.

## AGRADECIMENTOS

Certamente estes parágrafos não irão atender a todas as pessoas que fizeram parte dessa importante fase de minha vida. Portanto, desde já peço desculpas àquelas que não estão presentes entre essas palavras, mas elas podem estar certas que fazem parte do meu pensamento e de minha gratidão.

Agradeço a Deus, meu sustento, a Força que me move.

Aos meus irmãos do Grupo de Oração Jovem Restauração, que estiveram ao meu lado nos momentos até então mais difíceis da minha vida.

Agradeço ao meu orientador Prof. Dr. Erikson, pela compreensão e sabedoria com que me guiou nesta trajetória.

Agradeço ao meu coorientador Prof. Dr. Max, pelo incentivo para que eu iniciasse o mestrado e apoio em todo o projeto de pesquisa e pós-graduação.

Aos meus amigos, colegas de mestrado e laboratório, em especial à Thissiany, Mayara e Lucas, bem como Matheus e Othon, Fábila e Lin, por todo o apoio e auxílio.

Aos meus amigos e colegas de laboratório do Grupo de Sistemas Automotivos, em especial ao Rodrigo, Neme e David, por me auxiliarem e motivarem nas etapas iniciais do mestrado, bem como proverem ferramentas que me ajudaram nesta tarefa.

Gostaria de deixar registrado também, o meu reconhecimento à minha família, pois acredito que sem o apoio deles seria muito difícil vencer esse desafio.

Ao Prof. Dr. Thiago do Departamento Acadêmico de Mecânica, da UTFPR-PG, por nos disponibilizar a câmera infravermelha e permitir que realizássemos os experimentos necessários.

Enfim, a todos os que por algum motivo contribuíram para a realização desta pesquisa.

Mais vale a paciência que o heroísmo,  
mais vale quem domina o coração do que  
aquele que conquista uma cidade.  
(Provérbios 16:32)

## RESUMO

BUENO, Felipe. **Método para mapeamento tridimensional térmico aplicado à sistemas de auxílio ao condutor**. 2018. 97 f. Dissertação (Mestrado em Ciência da Computação) - Universidade Tecnológica Federal do Paraná, Ponta Grossa, 2018.

A maioria dos acidentes de trânsito são causados por falha humana. Dentre estes acidentes, muitos envolvem pedestres. Uma forma de minimizar esse cenário consiste em utilizar sistemas de auxílio ao motorista. Tais sistemas poderiam atuar de forma a impedir o acidente, caso recebam informações confiáveis e em tempo hábil acerca do ambiente onde o veículo se encontra. A utilização de informações do espectro infravermelho possibilita uma visão privilegiada do ambiente, pois não está sujeita à ruídos relacionados à iluminação, fumaça e neblina. Nesse sentido, no presente trabalho propõe-se avaliar a estimativa de distâncias de pedestres, a partir de um método capaz de gerar um modelo tridimensional térmico, o qual pode servir como entrada para um sistema de auxílio ao condutor. Este modelo consiste numa série de etapas, sendo estas, aquisição das imagens térmicas, calibração, correspondência estéreo, e otimização da correspondência. O resultado do método estabelece um mapa de profundidade capaz de definir distâncias até os objetos observados. Foi utilizado um algoritmo genético para se otimizar os parâmetros da etapa de correspondência do método. Foram avaliadas as distâncias de pedestres em cenários variados, e utilizou-se a métrica da raiz do erro médio quadrático para se comparar os valores obtidos com os reais. Os resultados obtidos foram satisfatórios para distâncias de até 10m, onde obteve-se um erro de 12,59%. A aplicação do presente método em sistemas de auxílio ao condutor poderia prevenir colisões com pedestres em vias coletoras (40km/h), e minimizar o impacto em vias locais (30km/h) e arteriais (60km/h).

**Palavras-chave:** Visão estéreo. Termografia. Sistemas de auxílio ao condutor. Visão computacional. Algoritmo genético.

## ABSTRACT

BUENO, Felipe. **Method for Three-dimensional thermal mapping applied to driver assistance systems**. 2018. 97 p. Dissertation (Master Degree in Computer Science) - Federal University of Technology - Paraná, Ponta Grossa, 2018.

Most traffic accidents are caused by human failure. Among these accidents, many involve pedestrians. One way to minimize this scenario is to use driver assistance systems. Such systems could act to prevent the accident if they receive reliable and timely information about the environment in which the vehicle is located. The use of information from the infrared spectrum provides a privileged view of the environment, as it is not subject to noise related to lighting, smoke and fog. In this sense, the present work intends to evaluate the estimation of pedestrian distances, based on a method capable of generating a three-dimensional thermal model, which can serve as input for a driver assistance system. This model consists of a series of steps, such as acquisition of thermal images, calibration, stereo matching, and matching optimization. The result of the method establishes a depth map capable of defining distances up to the observed objects. A genetic algorithm was used to optimize the parameters of the matching step of the method. The distances of pedestrians were evaluated in varied scenarios, and the root mean squared error metric was used to compare the values obtained with the real ones. The results obtained were satisfactory for distances up to 10m, where an error of 12.59% was obtained. Applying this method to driver assistance systems could prevent collisions with pedestrians in collector roads (40km/h) and minimize impact on local (30km/h) and arterial (60km/h) roads.

**Keywords:** Stereo vision. Thermography. Driver assistance systems. Computer vision. Genetic algorithm.

## LISTA DE FIGURAS

Figura 1 - Exemplo de atuação da função PAEB. ....	18
Figura 2 - Níveis de automação veicular. ....	19
Figura 3 - Funcionamento do LiDAR. ....	21
Figura 4 - Funcionamento do RADAR. ....	22
Figura 5 - Faixas do espectro eletromagnético. ....	26
Figura 6 - Exemplos de corpo negro, onde ocorrem múltiplas reflexões internas, e consequentemente a absorção de toda a radiação recebida. ....	28
Figura 7 - Demonstração da Lei de Wien. ....	30
Figura 8 - Exemplo de termograma. a) Imagem original, e b) Termograma. ....	31
Figura 9 - Modelo de câmera <i>pinhole</i> . ....	35
Figura 10 - Geometria epipolar. ....	36
Figura 11 - Retificação estéreo. ....	38
Figura 12 - Correspondência baseada em correlação. Avaliam-se janelas de pixels na imagem da direita, buscando-se a correspondência com a janela destacada na imagem da esquerda. ....	39
Figura 13 - Correspondência baseada em características. Avaliam-se as características relacionadas às bordas da figura na imagem da direita, buscando-se a correspondência com as características extraídas da imagem da esquerda. ....	39
Figura 14 - Geração do mapa de disparidade a partir de imagens retificadas. a) Imagem captada pela câmera da esquerda. b) Imagem captada pela câmera da direita. c) Mapa de disparidade gerado pelo algoritmo SGBM ( <i>semi global block matching</i> ). ....	40
Figura 15 - Funcionamento do algoritmo genético. ....	46
Figura 16 - Fluxograma proposto. A classificação acima de cada processo indica a seção na qual será abordado. ....	51
Figura 17 - Partes do par estéreo nº1 a) Estrutura de fixação, e b) Câmeras convencionais. ....	53
Figura 18 - Par estéreo nº2 utilizado. ....	53
Figura 19 - Câmera térmica FLIR T440. ....	54
Figura 20 - Suporte para deslocamento controlado da câmera térmica. ....	55
Figura 21 - Alvo de calibração convencional, com as quinas destacadas em colorido. ....	57
Figura 22 - Alvo de calibração térmico utilizando pastilhas de Peltier, em a) Imagem convencional e b) Termograma. ....	58
Figura 23 - Etapas para estabelecimento da profundidade a) Imagem original da câmera na posição esquerda, b) Após retificação, c) Mapa de disparidade original e d) Mapa de disparidade após aplicação do filtro WLS. ....	62

Figura 24 - Pessoa em espectros distintos - pouca iluminação. a) Imagem original, e b) Termograma.....	67
Figura 25 - Pessoa em espectros distintos - quase nenhuma iluminação. a) Imagem original, e b) Termograma. ....	68
Figura 26 - Pessoa em espectros distintos - ofuscamento. a) Imagem original, e b) Termograma.....	68
Figura 27 - Diferentes posições para calibração, sendo estas a) centro, b) abaixo, c) acima, d) direita, e) esquerda, f) inclinado direita, g) inclinado esquerda, h) inclinado cima e i) inclinado baixo. ....	70
Figura 28 - Diferentes posições para teste, em 3 colunas (convencional, térmico <i>grayscale</i> e térmico <i>rainbow</i> respectivamente): frontal, direita, costas, esquerda e inclinado. O retângulo foi definido manualmente e delimita a região avaliada. ....	71

## LISTA DE GRÁFICOS

Gráfico 1 - RMSD x Profundidade para diferentes gerações - dataset 2018_06_07__14_00__320x240__thermal.....	79
Gráfico 2 - RMSD x Profundidade para diferentes gerações - dataset 2018_09_04__19_00__320x240__thermal.....	80
Gráfico 3 - RMSD x Profundidade para diferentes gerações - dataset capts_1_320x240__thermal.....	81
Gráfico 4 - RMSD x Profundidade para diferentes gerações - dataset capts_2_320x240__thermal_rainbow.....	82
Gráfico 5 - RMSD x Profundidade para diferentes gerações - dataset capts_6_320x240__thermal_rainbow_dist.....	83
Gráfico 6 - RMSD x Profundidade para diferentes gerações - dataset 2018_08_21__12_30__480x640.....	84
Gráfico 7 - RMSD x Profundidade para diferentes gerações - dataset 2018_08_21__12_30__720x960.....	85
Gráfico 8 - RMSD x Profundidade para diferentes gerações - dataset 2018_09_04__15_00__480x640.....	86
Gráfico 9 - RMSD x Profundidade para diferentes gerações - dataset capts_3_640x480.....	87
Gráfico 10 - RMSD x Profundidade para diferentes gerações - dataset capts_8_640x480__dist.....	88

## LISTA DE QUADROS

Quadro 1 - Principais aspectos de trabalhos relacionados à detecção de pedestres e termografia. ....	34
Quadro 2 - Principais aspectos de trabalhos relacionados à visão estéreo. ....	43
Quadro 3 - Principais aspectos de trabalhos que utilizam algoritmos genéticos aplicados à visão estéreo. ....	49
Quadro 4 - Dados de saída da calibração. ....	60
Quadro 5 - Parâmetros do algoritmo genético.....	65
Quadro 6 - Parâmetros de correspondência estéreo que são genes do AG.....	66

## LISTA DE TABELAS

Tabela 1 - Subdivisões do infravermelho. ....	26
Tabela 2 - Parâmetros da câmera térmica. ....	55
Tabela 3 - Análise da profundidade calculada em imagens térmicas para distância de 2m. ....	72
Tabela 4 - Análise da profundidade calculada em imagens convencionais para distância de 2m. ....	73
Tabela 5 - Análise da profundidade calculada em imagens térmicas para distância de 3m. ....	74
Tabela 6 - Análise da profundidade calculada em imagens convencionais para distância de 3m. ....	74
Tabela 7 - Análise da profundidade calculada em imagens térmicas para distância de 4m. ....	75
Tabela 8 - Análise da profundidade calculada em imagens convencionais para distância de 4m. ....	75
Tabela 9 - Análise da profundidade calculada em imagens térmicas para distância de 5m. ....	76
Tabela 10 - Análise da profundidade calculada em imagens convencionais para distância de 5m. ....	76
Tabela 11 - Análise da profundidade calculada em imagens térmicas para distância de 10m. ....	77
Tabela 12 - Análise da profundidade calculada em imagens convencionais para distância de 10m. ....	77
Tabela 13 - Análise da profundidade calculada em imagens térmicas para distância de 15m. ....	78
Tabela 14 - Análise da profundidade calculada em imagens convencionais para distância de 15m. ....	78
Tabela 15 - Análise das distâncias de frenagens e resultados deste trabalho. ....	89

## LISTA DE ABREVIATURAS E SIGLAS

ADAS	<i>Advanced Driver Assistance Systems</i>
AG	Algoritmo Genético
bbox	<i>Bounding box</i>
ELAS	<i>Efficient Large-Scale Stereo Matching</i>
Euro NCAP	<i>European New Car Assessment Program</i>
FIR	<i>Far infrared</i>
FPS	<i>Frames per Second</i>
GOLD	<i>Generic Obstacle and Lane Detection system</i>
GT	<i>Ground-truth</i>
HOG	<i>Histogram Oriented Gradient</i>
ICP	<i>Iterative Closest Point</i>
IPM	<i>Inverse Perspective Mapping</i>
LiDAR	<i>Light Detection and Ranging</i>
LKA	<i>Lane Keeping Assist</i>
LWIR	<i>Long-wavelength infrared</i>
MIM	<i>Mutual Information Maximization</i>
MTT	Modelo Tridimensional Térmico
MWIR	<i>Mid-wavelength infrared</i>
NASA	<i>National Aeronautics and Space Administration</i>
NASA-TLX	<i>NASA Task Load Index</i>
NHTSA	<i>National Highway Traffic Safety Administration</i>
NIR	<i>Near infrared</i>
OMS	Organização Mundial de Saúde
PASAT	<i>Paced Auditory Serial Addition Test</i>
RADAR	<i>Radio Detection and Ranging</i>
RMSD	<i>Root Mean Squared Deviation</i>
SAE	<i>Society of Automotive Engineers</i>
SGBM	<i>Semi-Global Block Matching</i>
SI	Sistema Internacional de Unidades
SWIR	<i>Short-wavelength infrared</i>
ToF	<i>Time of Flight</i>
WLS	<i>Weightened Last Squares</i>

## SUMÁRIO

<b>1 INTRODUÇÃO</b> .....	<b>15</b>
1.1 DEFINIÇÃO DO PROBLEMA .....	17
1.2 OBJETIVOS.....	23
1.2.1 Objetivo Geral.....	23
1.2.2 Objetivos Específicos.....	23
1.3 ORGANIZAÇÃO DO TRABALHO.....	24
<b>2 REVISÃO DA LITERATURA</b> .....	<b>25</b>
2.1 TERMOGRAFIA INFRAVERMELHA .....	25
2.1.1 Trabalhos Relacionados .....	31
2.2 VISÃO ESTÉREO.....	34
2.2.1 Trabalhos Relacionados .....	41
2.3 ALGORITMO GENÉTICO.....	45
2.3.1 Trabalhos Relacionados .....	48
<b>3 MÉTODO PARA MAPEAMENTO TRIDIMENSIONAL TÉRMICO</b> .....	<b>51</b>
3.1 AQUISIÇÃO DE IMAGENS.....	52
3.2 CALIBRAÇÃO.....	56
3.3 CORRESPONDÊNCIA ESTÉREO .....	61
3.4 AVALIAÇÃO E OTIMIZAÇÃO .....	62
<b>4 EXPERIMENTOS E RESULTADOS</b> .....	<b>67</b>
4.1 ANÁLISE QUALITATIVA.....	67
4.2 ANÁLISE QUANTITATIVA.....	69
<b>5 CONCLUSÃO</b> .....	<b>90</b>
5.1 TRABALHOS FUTUROS .....	91
<b>REFERÊNCIAS</b> .....	<b>92</b>

## 1 INTRODUÇÃO

A visão é um dos principais sentidos utilizados pelos animais para captar informação do mundo ao seu redor. Ela apresenta diferentes níveis de desenvolvimento dependendo da espécie, e em conjunto com os demais sentidos, garante a sobrevivência dos indivíduos, permitindo a identificação de ameaças bem como de alimento. Águias tem uma visão aguçada, o que permite a identificação de presas a grandes distâncias e o estabelecimento prévio da rota de voo para a abordagem. Moscas têm olhos adaptados que possibilitam uma visão ampla, aumentando sua percepção ao redor de si. Na espécie humana, a visão tem papel fundamental, pois facilita a percepção de objetos e indivíduos, bem como sua localização, o que permite, por exemplo, o deslocamento por meio de determinado espaço.

A noção de posicionamento de objetos é possível devido à disparidade da posição dos olhos, ou seja, a partir da observação da mesma cena de pontos de vista ligeiramente diferentes. A aquisição de informações tridimensionais, geradas com base na disparidade, constitui o plano de estudo da visão estéreo ou estereoscopia (HARTLEY; ZISSERMAN, 2004a, p. 239; RUSS, 2002, p. 64-65).

A análise de informações obtidas pela visão se mostra insuficiente em determinadas condições onde, por exemplo, o campo de visão está de algum modo limitado/ofuscado. Desta forma alguns animais desenvolveram habilidades que lhes permitem sentir radiação térmica infravermelha, como é o caso das cobras, que podem utilizar esta característica para a identificação de presas e predadores. Esse sensor infravermelho natural possibilita a estimativa da posição da fonte de calor mesmo em situações de baixa iluminação e até mesmo oclusão da presa (GORIS, 2011).

Em um sistema de visão computacional, uma representação desta informação térmica captada é o mapa de calor, no qual fica visível a energia térmica irradiada na faixa do infravermelho no espectro eletromagnético com base na variação na coloração dos pixels da imagem. A detecção desta energia e sua conversão para uma imagem faz parte da termografia infravermelha (MEOLA, 2012, cap. 1).

Várias tecnologias são inspiradas na natureza como é o caso do sonar inspirado dos morcegos, o velcro inspirado em filamentos de plantas, o trem bala que

possui aerodinâmica inspirada em pássaros pescadores e submarinos inspirados nos peixes, dentre outros.

Sabe-se que a visão no espectro infravermelho representa informação privilegiada, pois é capaz de atravessar obstáculos como fumaça e neblina. Assim como uma cobra consegue perseguir sua presa no escuro e percebê-la atrás de um arbusto, um sistema de visão artificial que capte informações no espectro infravermelho seria capaz de permitir a um sistema de auxílio à condução, detectar um pedestre oculto devido à falta de luminosidade, antes mesmo que estivesse visível ao motorista devido à incidência da luz dos faróis. Nesse caso, várias decisões poderiam ser tomadas com antecedência, como emitir um alerta sonoro ou exibir o ponto de risco em uma tela auxiliar.

Além disso, um sistema de auxílio ao motorista equipado com tecnologias inteligentes de visão térmica, seria capaz de diminuir acidentes de trânsito em situações específicas onde o motorista tem a visão prejudicada. Em momentos de forte neblina, o condutor precisa dirigir com atenção redobrada e mesmo assim o risco de acidentes com engavetamento é grande. A existência de focos de incêndio próximo à via de trânsito e a fumaça decorrente obstrui a visão do motorista, aumentando também o risco de ocorrer um acidente em função da visão desfavorecida. O mesmo ocorre em situações de grande volume de chuva, onde é aconselhado parar o veículo em local seguro devido à falta de visão.

Em todos esses casos, um sistema de visão auxiliar capaz de captar informações no espectro infravermelho conseguiria ver além dos obstáculos citados anteriormente e antecipar informações que o condutor ainda não consegue perceber. Esse sistema de visão possui parâmetros que visam adequar sua resposta a determinado ambiente. A escolha correta da parametrização pode aprimorar os resultados obtidos, relacionados à visão do ambiente no espectro infravermelho. Um método bastante empregado em otimizações é a aplicação de meta-heurísticas, que possibilitam a partir da utilização de um determinado modelo, a busca de configurações de parâmetros melhores. Um desses modelos é o algoritmo genético, que se baseia no modelo evolutivo de Darwin, e consegue através da evolução das populações, que representam o conjunto de parâmetros a ser otimizado, chegar a melhores resultados.

Considerando o sistema calibrado, com seus parâmetros otimizados, e contando com redundância de informação oriundas de, pelo menos duas câmeras

térmicas, caracteriza-se um sistema estéreo de onde é possível extrair um mapeamento tridimensional térmico que permite observar as informações de temperatura emanadas pelo ambiente associadas a coordenadas 3D do mundo real. Isso significa saber em qual ponto da cena observada está qual fonte de calor, inclusive com a informação de profundidade ou a distância em relação ao ponto de observação.

Deste modo, analogamente a uma cobra que intui a qual distância está sua presa, um sistema de visão auxiliar contando com um mapeamento tridimensional térmico, seria capaz de saber de forma antecipada, a distância e o posicionamento dos obstáculos quentes a sua frente, tais como pedestres, animais na margem da pista, outros veículos emitindo calor por seus sistemas de exaustão, entre outros.

Na seção seguinte será tratado especificamente o problema deste trabalho, bem como a justificativa do tema e objetivos a serem atingidos.

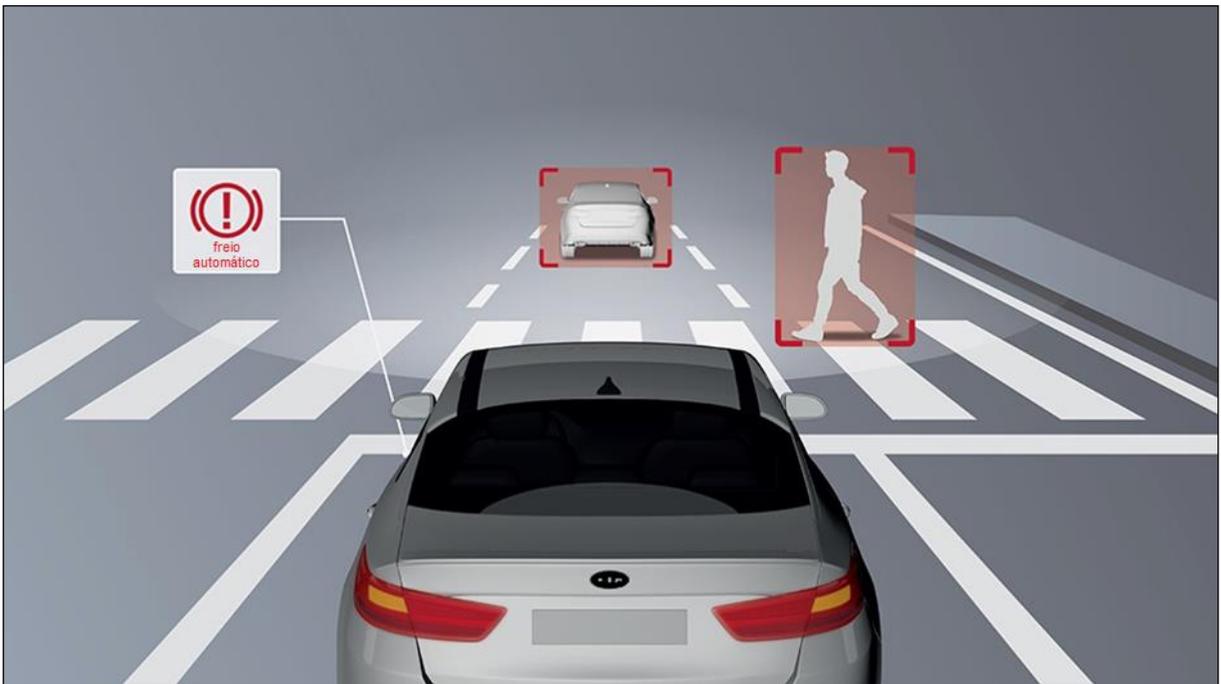
## 1.1 DEFINIÇÃO DO PROBLEMA

No Brasil, cerca de 90% dos acidentes de trânsito são causados por falha do condutor (ONSV, 2015). Nos Estados Unidos, essa estatística chega a aproximadamente 94% (NHTSA, 2015b). Os motivos variam, podendo ser classificados de forma relacionada à imprudência (desrespeito à legislação, como a realização de ultrapassagens em lugares proibidos) e imperícia (falta de experiência do condutor, o que leva o mesmo a não reagir de forma adequada perante situações de risco no trânsito). Segundo a OMS (Organização Mundial de Saúde), 1,25 milhões de pessoas morrem nas estradas todos os anos, sendo que 31% são ocupantes de veículo, 23% motociclistas (duas ou três rodas), 22% pedestres, 4% ciclistas e 21% outros (WHO, 2015). Nos Estados Unidos, cerca de 25% dos acidentes com veículos resultam em ferimentos ou mortes de pedestres (NHTSA, 2015a).

Uma forma de minimizar essa estatística consiste na utilização de tecnologia para automatizar funções de direção ou auxiliar o motorista nesta tarefa. Neste contexto atuam os sistemas de assistência avançada ao motorista (ADAS - *Advanced Driver Assistance Systems*), os quais são capazes de auxiliar o motorista na tomada de decisão, gerar alertas durante situações de risco, e até mesmo assumir o controle do veículo, como é o caso da função de estabilidade de faixa (LKA - *Lane Keeping*

*Assist*), a qual auxilia o motorista a permanecer na faixa onde o mesmo transita (BISHOP, 2005, p. 109; GERONIMO et al., 2010). Outra função ADAS importante é a de acionamento automático de freios de emergência para prevenção de colisões com pedestres (PAEB - *Pedestrian Automatic Emergency Breaking*), que avalia se o pedestre está à frente do veículo e se há possibilidade de colisão baseada na velocidade e direção do automóvel, acionando os freios de emergência se necessário (SAFERCAR, 2018), conforme ilustrado na Figura 1.

**Figura 1 - Exemplo de atuação da função PAEB.**



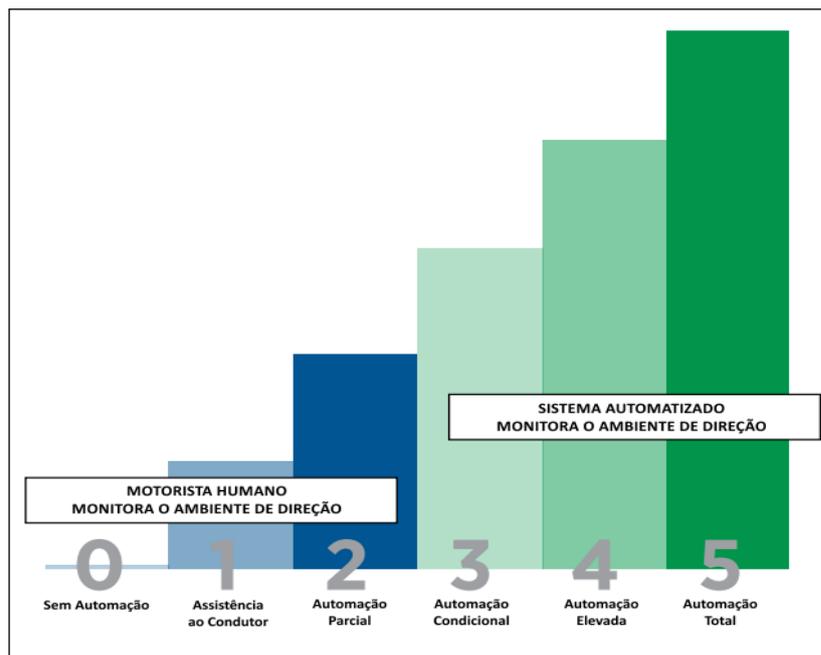
Fonte: Adaptado de KIA (2017) .

Em função deste cenário, políticas de incentivo à utilização de automação veicular vêm sendo empregadas, com o intuito de fomentar seu uso, como por exemplo a aplicação de tecnologia de visão traseira, sendo uma exigência da agência norte americana NHTSA (*National Highway Traffic Safety Administration*), para todos os veículos novos abaixo de 4536 kg (aprox. 10000 lb.), fabricados a partir de maio de 2018 (NAYLOR, 2014). De forma semelhante, a Euro NCAP (*European New Car Assessment Program*) desenvolveu um ranking e protocolos de avaliação para estabelecer o nível de segurança dos veículos, o qual pode variar de uma a cinco estrelas. Esse ranking leva em consideração quatro eixos, os quais são 1) proteção à ocupantes adultos do veículo, 2) proteção à crianças ocupantes do veículo, 3) proteção à pedestres (a qual foi expandida para incluir ciclistas, tornando-se proteção

à usuários vulneráveis da estrada) e 4) atuação de sistemas de auxílio ao condutor (EURO NCAP, 2018).

O crescimento da utilização de tecnologia veicular tornou necessário uma padronização da nomenclatura, de forma a especificar os níveis de tecnologia que se aplicam, desde veículos sem automação nenhuma até os veículos autônomos, totalmente guiados por computador. Desta forma, em 2014 a SAE (*Society of Automotive Engineers*) criou uma norma regulamentadora, a J3016, “Taxonomia e Definições para Termos Relacionados à Sistemas de Direção Automatizada para Veículos Motorizados Rodoviários”<sup>1</sup> (SAE INTERNATIONAL, 2018), a qual estabeleceu padrões mais tarde adotados pelo congresso dos Estados Unidos (SAE INTERNATIONAL, 2017). A ideia central dessa norma é apresentada na Figura 2, onde se subdivide os níveis de automação, e o escopo de atuação humano/máquina em cada nível.

**Figura 2 - Níveis de automação veicular.**



Fonte: Adaptado de SAE International (2017).

Nos níveis iniciais, do zero ao dois, o motorista é responsável pelo monitoramento do ambiente, e conforme a classificação aumenta, o sistema veicular ganha responsabilidades acessórias. O nível zero equivale a um veículo sem nenhum sistema de automação. O nível um representa a atuação de um sistema de auxílio ao

<sup>1</sup> *Taxonomy and Definitions for Terms Related to On-Road Motor Vehicle Automated Driving Systems.*

condutor, como o *Lane Keep Assist*, que aplica um torque no veículo caso este esteja saindo da faixa de condução. O nível dois representa a atuação de mais de um sistema ADAS simultaneamente, por exemplo o LKA e a desaceleração automática do veículo diante de uma situação de risco. Dos níveis três ao cinco, o sistema é responsável pelo monitoramento do ambiente, e conforme maior o nível, menos é necessária a intervenção humana. No nível três, o sistema é capaz de navegação autônoma em certas circunstâncias, por exemplo em rodovias, porém caso seja identificada uma condição que este não abrange, o condutor precisa assumir a direção. No nível quatro, o sistema também pode navegar autonomamente em certas circunstâncias, e consegue se adaptar caso seja necessária a intervenção do condutor, por exemplo executando uma parada em local seguro. Por fim, no nível cinco, o sistema é capaz de direção autônoma em todos os ambientes que um motorista humano poderia dirigir, executando todas as funções inerentes à direção (SAE INTERNATIONAL, 2018).

No contexto de acidentes de trânsito, um desafio consiste na detecção de objetos que possam atuar como obstáculos para o condutor na via, como no caso de pedestres, veículos, animais e outros objetos. Além da detecção, é interessante identificar sua posição no espaço, o que possibilita ao sistema ADAS atuar de forma preventiva, evitando assim a colisão.

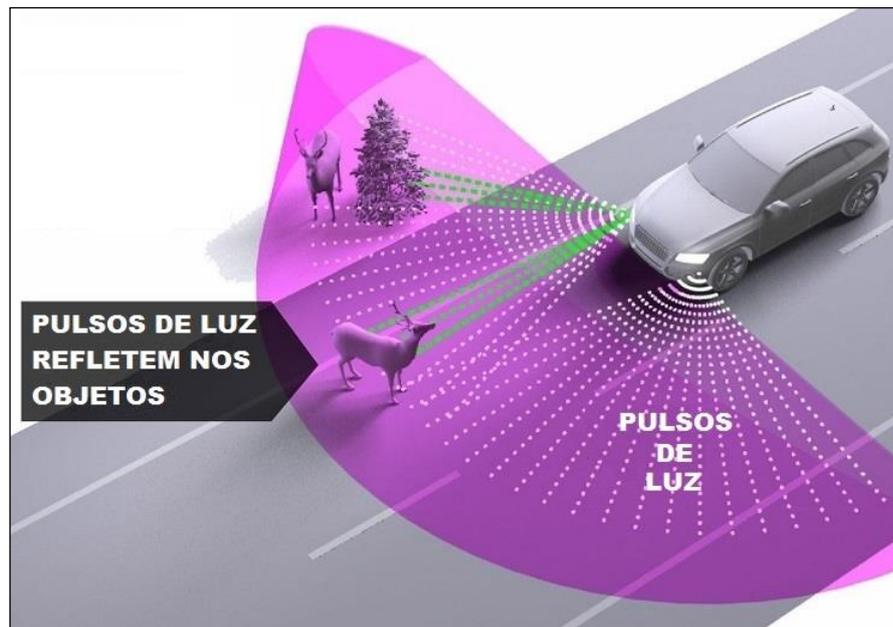
Esta identificação pode ser realizada de diferentes formas, dentre elas:

- Fusão de dados provenientes de radar e câmeras convencionais (ALESSANDRETTI; BROGGI; CERRI, 2007; WANG et al., 2011);
- Fusão de dados provenientes de escaneadores laser e câmeras convencionais (BROGGI et al., 2009);
- Análise do modelo de movimento (JAZAYERI et al., 2011);
- Análise de imagens obtidas de diferentes perspectivas (imagens estéreo) (BERTOZZI; BROGGI, 1998; HUH et al., 2008; TOULMINET et al., 2006);
- Fusão de dados entre imagens estéreo e informação térmica (câmera infravermelha) (KROTOSKY; TRIVEDI, 2007).

Dentre as diversas tecnologias que se propõem estabelecer a distância entre os objetos da via e o veículo, convém destacar três tipos de sensores que podem ser empregados no meio automotivo além das câmeras convencionais, sendo eles o LiDAR (*Light Detection and Ranging*), RADAR (*Radio Detection and Ranging*) e sensor térmico. O LiDAR funciona a partir da emissão, reflexão e análise de feixes de

luz a partir do sensor equipado no veículo. A Figura 3 indica o funcionamento deste sensor, onde inicialmente são emitidos pulsos de luz em uma determinada frequência. Esses pulsos trafegam pelo ambiente até atingirem um objeto em seu caminho, sendo refletidos novamente para o sensor, conforme destacado em verde na figura. Conhecendo-se a velocidade do deslocamento do feixe de luz no meio (velocidade da luz), e analisando o tempo entre a emissão e recepção do sinal, é possível estimar a distância dos objetos na cena.

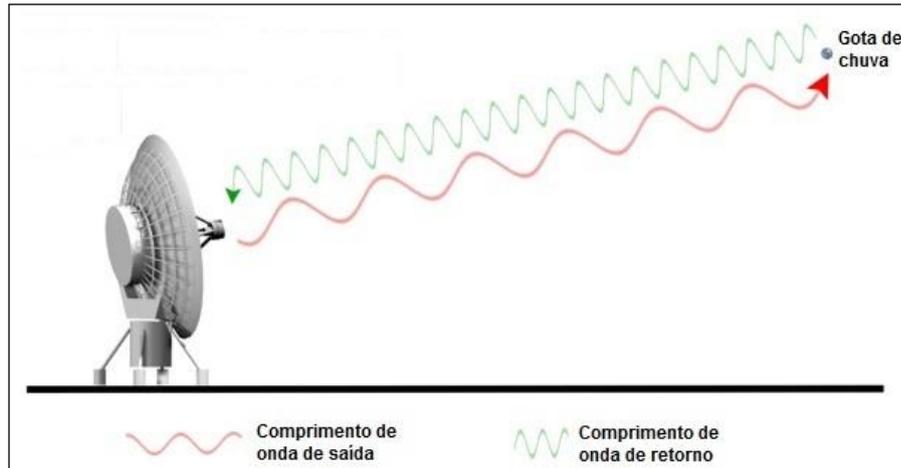
**Figura 3 - Funcionamento do LiDAR.**



**Fonte: Adaptado de Automotive News (2015).**

O RADAR, por sua vez, tem um funcionamento muito semelhante ao LiDAR. A principal diferença refere-se que este atua em uma seção diferente do espectro eletromagnético (ondas de rádio ao invés de luz). Além disso, o RADAR possui um alcance maior, porém menor precisão se comparado com o LiDAR (cobre maiores distâncias, porém objetos menores que seu comprimento de onda podem não ser detectados) (KIDONO et al., 2011). A Figura 4 indica o funcionamento deste sensor.

**Figura 4 - Funcionamento do RADAR.**



**Fonte: Adaptado de EOL - Earth Observing Laboratory (2017).**

Por fim o sensor térmico diferencia-se dos demais pois se utiliza da captação da radiação térmica natural dos corpos que possuem temperatura acima do zero absoluto. Este sensor capta a quantidade de energia radiante emitida na faixa infravermelha do espectro eletromagnético, e baseado em leis da Física possibilita a identificação da temperatura dos corpos que emitiram a radiação, possibilitando a formação de imagens térmicas. A interferência que este sensor estaria sujeito está relacionada a fontes de calor próximas da cena observada, diferentemente dos demais sensores aqui descritos.

No sentido de auxiliar o desenvolvimento e melhoramento de sistemas ADAS, este trabalho propõe a geração de um modelo tridimensional térmico (MTT), com o foco de estabelecer a distância de pedestres na cena, o qual representa uma ferramenta importante para sistemas de auxílio ao motorista, principalmente por apresentar as diferentes temperaturas da cena mapeadas a uma estimativa confiável de distância e posicionamento 3D.

O trabalho realizado configura uma solução prática, visto que se baseia em sensores térmicos capazes de produzir o termograma, ou seja uma representação em formato de imagem do mapa de calor da cena. O resultado é uma visão privilegiada capaz de identificar corpos quentes mapeados em coordenadas 3D, além de não apresentar sensibilidade a interferências e ruídos tais como fumaça, neblina e variações na iluminação, os quais são fatores que comumente caracterizam condições de difícil análise em sistemas de visão computacional, como detalhado no trabalho de Andrade et al. (2018) onde os autores realizaram experimentos também em condições adversas da rodovia, como presença de sombra, ofuscamento e contraste na imagem.

## 1.2 OBJETIVOS

Nesta seção serão apresentados o objetivo geral e os objetivos específicos deste trabalho. A identificação do pedestre na imagem, ou seja, o reconhecimento do mesmo na cena, não faz parte do escopo deste trabalho e não será abordada.

### 1.2.1 Objetivo Geral

Avaliar a estimativa de distância de pedestres utilizando a raiz do erro médio quadrático, em comparação com as medidas de distância reais da cena observada, a partir da geração de um mapa de profundidade baseado em imagens contendo informação de temperatura, organizadas em um par estéreo calibrado, considerando otimizações nos parâmetros de correspondência decorrentes da aplicação de um algoritmo genético.

### 1.2.2 Objetivos Específicos

Como objetivos específicos, pode-se citar:

- Estabelecer a geometria epipolar da cena térmica;
- Estabelecer os mapas de disparidade e profundidade da cena avaliada;
- Avaliar a profundidade calculada em comparação com a medida real, utilizando a raiz do erro médio quadrático, para a região de interesse relativa à posição do pedestre na imagem;
- Investigar o número de gerações do algoritmo genético, que otimiza os parâmetros de correspondência, e seu impacto na qualidade da estimativa de profundidade;
- Analisar a variação de pixels válidos nas regiões de interesse delimitadas, em função das distâncias avaliadas.

Na seção seguinte será detalhado a distribuição do conteúdo ao longo dos capítulos deste trabalho.

### 1.3 ORGANIZAÇÃO DO TRABALHO

O trabalho está organizado da seguinte forma: o Capítulo 2 apresenta uma revisão dos conceitos principais envolvidos acerca de termografia, visão estéreo e algoritmo genético, bem como evidencia aplicações destes temas encontradas na literatura. No Capítulo 3 é descrito o método utilizado, sua implementação e as fases que o compõem. No Capítulo 4 são apresentados os resultados obtidos, bem como o detalhamento dos experimentos realizados. Por fim, no Capítulo 5 são apresentadas as conclusões acerca do trabalho realizado.

## 2 REVISÃO DA LITERATURA

Esta seção tem como objetivo revisar os principais conceitos utilizados neste trabalho, bem como apresentar aplicações a partir da análise de trabalhos correlatos. Na seção 2.1 é tratado a respeito da termografia, abordando inicialmente o conceito de temperatura, e finalizando com emissividade. Na seção 2.1.1 são descritos trabalhos que aplicam a termografia em cenários variados. Na seção 2.2 são revisados os conceitos acerca da visão estéreo, como o modelo de câmera *pinhole* e geometria epipolar. Na seção 2.2.1, por sua vez, são descritas aplicações diversas da visão estéreo. Na seção 2.3 é apresentada a teoria acerca do algoritmo genético (AG), seus operadores e funcionamento. Na seção 2.3.1, por sua vez, são elencados trabalhos que utilizam o AG como passo de otimização de processos relacionados à sistemas de visão estéreo.

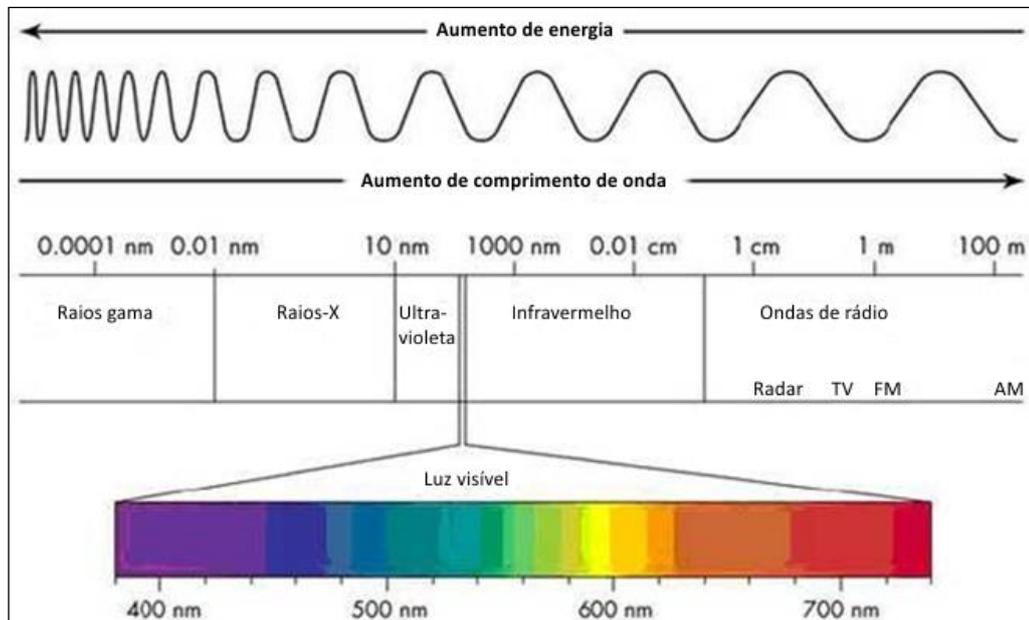
### 2.1 TERMOGRAFIA INFRAVERMELHA

A termografia, ou o registro da temperatura de um determinado corpo, já vem sendo empregada há muito tempo. Por volta de 400 a.C. existem registros referentes à Hipócrates, o “pai da medicina”, acerca da utilização da diferença de temperatura de certas regiões do corpo como um indicador de enfermidade da parte analisada (ŽIVČÁK et al., 2013, p. 1). Com a evolução da história, percebe-se que uma das principais aplicações da termografia é no ramo da medicina, como um fator importante no diagnóstico do paciente, sendo utilizada no estudo da oncologia, desordens vasculares e respiratórias, análise da dor, entre outras (DIAKIDES et al., 2008, p. 21).

A medição da radiação infravermelha, por sua vez, tem sua origem atribuída a Herschel que, por volta de 1800, realizou um experimento onde observou os efeitos do calor incidentes na porção localizada logo após a faixa vermelha da parte visível do espectro eletromagnético, a partir da decomposição da luz do sol incidente em um prisma. Além de provar a existência de faixas do espectro além das visíveis, Herschel observou que as temperaturas registradas na porção além da faixa vermelha eram mais altas que as registradas nas seções visíveis (MINKINA; DUDZIK, 2009, p. 15). Essa seção invisível, até então desconhecida, corresponde ao que se conhece hoje

como espectro infravermelho. A Figura 5 exibe a posição desse espectro em relação ao visível.

**Figura 5 - Faixas do espectro eletromagnético.**



Fonte: Adaptado de Taib, Jadin e Kabir (2012, p. 210).

Esse espectro infravermelho, por sua vez, também é dividido em regiões baseado na faixa de comprimento de onda no qual atua. Essa divisão normalmente é feita em cinco seções, conforme indicado na Tabela 1. Dentre estas seções, a do infravermelho longo que geralmente é utilizada em sistemas de visão térmica, pois é possível obter uma imagem térmica completa sem a necessidade de iluminação ou fonte de calor externa (D'AMICO et al., 2009, p. 22).

**Tabela 1 - Subdivisões do infravermelho.**

Tipo	Comprimento de onda ( $\mu\text{m}$ )
Próximo ( <i>Near infrared</i> - NIR)	0,75 - 1,4
Curto ( <i>Short-wavelength infrared</i> - SWIR)	1,4 - 3
Médio ( <i>Mid-wavelength infrared</i> - MWIR)	3 - 8
Longo ( <i>Long-wavelength infrared</i> - LWIR)	8 - 15
Distante ( <i>Far infrared</i> - FIR)	15 - 1000

Fonte: Adaptado de D'Amico et al. (2009).

Pode-se definir a termografia infravermelha como uma metodologia não invasiva que permite o registro de forma gráfica da energia térmica irradiada de objetos na faixa infravermelha do espectro eletromagnético. Para seu melhor

entendimento, é necessária antes a definição de conceitos físicos acerca da transferência de calor pela radiação, os quais foram sendo estabelecidos e aprimorados desde o experimento de Herschel. Esse estudo pode ser realizado a partir da abordagem dos seguintes conceitos (ŽIVČÁK et al., 2013, cap. 3):

- Temperatura;
- Calor (Energia Térmica);
- Transferência de Calor;
- Primeira Lei da Termodinâmica;
- Radiação Infravermelha;
- Corpo Negro;
- Emissividade;
- Lei de Planck;
- Lei de Deslocamento de Wien;
- Termograma.

Temperatura ( $T$ ) é a medida estatística da energia cinética molecular/atômica devido ao ganho ou perda de calor. No Sistema Internacional de Unidades (SI) é medida em Kelvin (K), onde o ponto de solidificação da água de 273 K, e o ponto de vaporização 373 K, ao nível do mar (ŽIVČÁK et al., 2013, cap. 3). A detecção da energia térmica na termografia é traduzida em unidades de temperatura.

Calor ( $Q$ ) é a medida da troca de energia térmica transferida entre sistemas quando há diferença de temperatura, a qual ocorre até que seja atingido o equilíbrio térmico (HALLIDAY; RESNICK; WALKER, 2012, p. 191). É medido em Joules (J).

A transferência de calor ocorre entre regiões de maior temperatura para regiões de menor temperatura, de acordo com a segunda lei da termodinâmica. Pode ocorrer de três formas: condução (no interior do corpo, molécula a molécula, ou entre corpos em contato), convecção (movimentação de massa devido à diferença de densidade, característica de fluidos) e radiação (propagação de energia por meio de ondas eletromagnéticas, capaz de ocorrer no vácuo) (YOUNG; FREEDMAN; FORD, 2008, p. 199, 202, 203).

A primeira lei da termodinâmica é baseada no princípio de conservação de energia, o qual afirma que a quantidade total de energia em um sistema isolado permanece constante. A primeira lei afirma que a energia interna de um sistema ( $U$ )

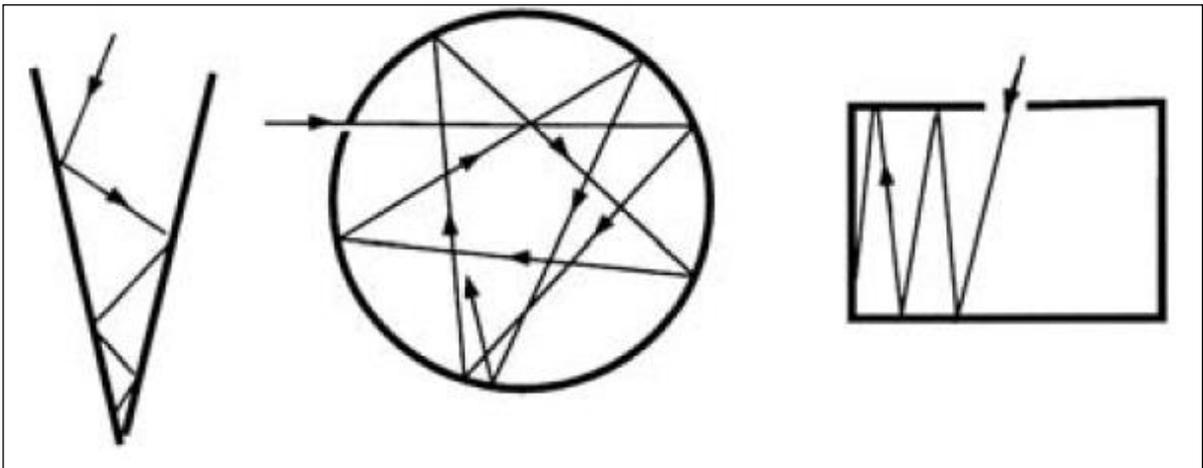
varia com base na diferença de calor (recebido ou cedido) e realização de trabalho ( $W$ ), como mostra a equação (1),

$$\Delta U = \Delta Q - \Delta W \quad (1)$$

onde  $\Delta U$  representa a mudança da energia interna do sistema,  $\Delta Q$  a variação de calor e  $\Delta W$  a variação do trabalho realizado ou consumido (ŽIVČÁK et al., 2013, cap. 3).

A radiação infravermelha representa a transferência de calor através da radiação. Todo objeto com temperatura superior ao zero absoluto (0 Kelvin) emite este tipo de radiação. Além disso, quando a radiação atua sobre um corpo, parte dela é absorvida, parte passa através do corpo e parte é refletida (MINKINA; DUDZIK, 2009, cap. 2). Um corpo no qual toda a radiação é absorvida é chamado de corpo negro. Isso acontece devido à ocorrência de múltiplas reflexões internas no corpo. A Figura 6 exibe modelos geométricos onde poderia ocorrer tal comportamento, visto que o corpo negro é um corpo idealizado.

**Figura 6 - Exemplos de corpo negro, onde ocorrem múltiplas reflexões internas, e consequentemente a absorção de toda a radiação recebida.**



Fonte: Minkina e Dudzik (2009).

O modelo de corpo negro, além de absorver toda a radiação incidente sobre ele, possui o maior índice de emissividade, a qual representa a capacidade da superfície do material de possibilitar a emissão de radiação. A emissividade é calculada com base na razão entre a quantidade de energia emitida pelo corpo analisado ( $Z'$ ) e seu respectivo corpo negro ( $Z$ ), dentro das mesmas condições de temperatura e comprimento de onda (MINKINA; DUDZIK, 2009, cap. 2), sendo representada pela letra  $\varepsilon$ , conforme descrito na equação (2).

$$\varepsilon = \frac{Z'}{Z} \quad (2)$$

A emissividade é uma medida adimensional importante para a análise termográfica de qualquer corpo, pois serve de parâmetro para se calcular a temperatura apresentada em função da quantidade de energia radiante emitida. Para melhor entender essa questão, serão definidos os conceitos de leis de radiação, que estabelecem proporcionalidades entre temperatura, comprimento de onda e fluxo radiante. Antes disso, porém, é interessante a definição de três conceitos, que são, em linhas gerais (TANG; LI, 2014, p. 7):

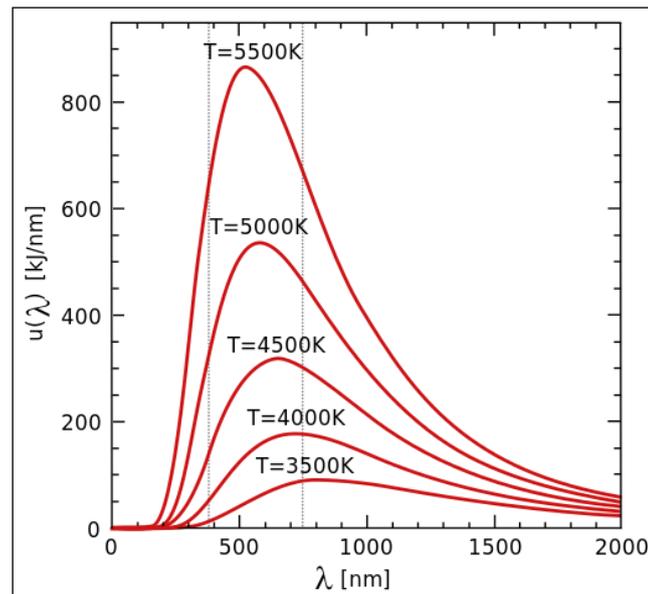
- Fluxo radiante: quantidade de energia radiante emitida por unidade de tempo;
- Radiância: quantidade de fluxo radiante em relação à unidade dimensional - área;
- Radiância espectral: radiância de uma superfície, podendo ser representada de diversas formas, dentre elas por comprimento de onda ou frequência.

As leis supracitadas são a Lei de Planck e Lei de Deslocamento de Wien. A Lei de Planck define a radiância espectral para um corpo negro. A partir dela, a radiância espectral é calculada apenas em função da temperatura de um corpo, desconsiderando o material que o constitui ou sua forma. Esta lei se aplica à chamada radiação de corpo negro, que é a radiação emitida por um corpo negro em equilíbrio térmico. Essa lei é importante porque estabelece uma relação direta entre a radiância espectral, e dessa forma a quantidade de energia radiante, e a temperatura de um corpo, em um cenário ideal tendo por base o corpo negro. A equação (3) representa a lei de Planck, onde  $I_\lambda$  é a radiância espectral em termos de comprimento de onda ( $\lambda$ ),  $T$  a temperatura,  $h$  a constante de Planck ( $6,62606876 \times 10^{-34}$  J·s),  $c$  a velocidade da luz ( $2,99792458 \times 10^8$  m/s) e  $k$  a constante de Boltzmann ( $1,3806503 \times 10^{-23}$  J/K) (LIENHARD IV; LIENHARD V, 2008, p. 31-32).

$$I_\lambda(T) = \frac{2\pi h c^2}{\lambda^5} \frac{1}{e^{\frac{hc}{\lambda kT}} - 1} \quad (3)$$

A Lei de Deslocamento de Wien estabelece que a emissão máxima da radiação espectral para um determinado comprimento de onda é inversamente proporcional à temperatura apresentada. A Figura 7 demonstra esse conceito, onde percebe-se a alteração do pico de emissão de energia em função da mudança de temperatura, para diferentes comprimentos de onda.

**Figura 7 - Demonstração da Lei de Wien.**



Fonte: Wiki Commons (2008).

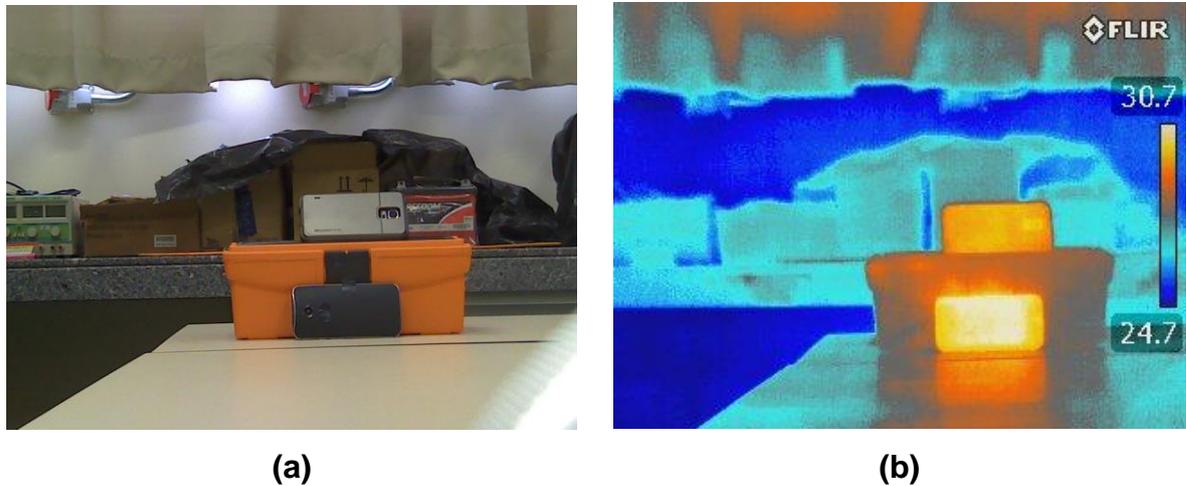
A partir dessa lei, é possível estabelecer a temperatura apenas baseado no conhecimento do comprimento de onda. Ela é representada na equação (4), onde  $\lambda_{max}$  é o comprimento de onda de máxima emissão,  $b$  é a constante de deslocamento de Wien ( $2,8977729 \times 10^{-3} \text{ m}\cdot\text{K}$ ), e  $T$  a temperatura.

$$\lambda_{max} = \frac{b}{T} \quad (4)$$

Uma ferramenta importante na análise computacional da temperatura de um ambiente é o termograma, que é uma representação da informação térmica em forma de imagem, onde se utiliza uma paleta de cores para identificar a temperatura das regiões na cena. A Figura 8 representa um exemplo de termograma que utiliza uma paleta onde os tons azuis representam uma temperatura menor na cena (conforme a escala de temperatura definida, presente na imagem térmica) e os tons de laranja representam uma temperatura maior. Neste exemplo, quanto mais próximo de  $24,7^\circ\text{C}$ , mais azul escuro o pixel, até atingir a cor preta, e quanto mais próximo de  $30,7^\circ\text{C}$ ,

mais laranja claro, até atingir a cor branca. Temperaturas abaixo de 24,7 °C e acima de 30,7 °C assumem a coloração mínima e máxima, respectivamente.

**Figura 8 - Exemplo de termograma. a) Imagem original, e b) Termograma.**



**Fonte: Autoria própria.**

A seção seguinte irá apresentar trabalhos que fazem uso da termografia em cenários variados.

### 2.1.1 Trabalhos Relacionados

A termografia se trata de uma tecnologia que abrange várias áreas distintas com diversas aplicações diferentes entre si. Dessa forma, esta seção lista alguns destes trabalhos afim de elucidar a aplicabilidade do tema e sua vastidão, sempre tendo em vista os objetivos finais do objeto de estudo deste trabalho.

Na linha biomédica, de Souza, Sanches e Gamba (2012) apresentaram um método para geração de um modelo tridimensional térmico do crânio, utilizando imagens infravermelhas, ressonância magnética e/ou tomografia computadorizada. Esta fusão de dados possibilita a visualização conjunta da temperatura da pele (superfície) e informação interna (ressonância magnética) a partir do mesmo modelo 3D, o qual pode ser utilizado no auxílio do diagnóstico e monitoramento dos pacientes.

Uma abordagem com foco na indústria pode ser encontrada no trabalho de Anzengruber e Riener (2012), onde os autores apresentam uma metodologia capaz de avaliar o nível de stress de um condutor de veículos baseado na análise de imagens térmicas faciais. Os testes foram feitos a partir da condução do veículo em

um ambiente simulado, com situações variadas de distração, como a aplicação simultânea do teste acústico/cognitivo PASAT (*Paced Auditory Serial Addition Test*), que mede a velocidade de processamento de informações, e tarefas numéricas visuais/motoras baseado em um *heads-up display*<sup>2</sup>, que exibe uma sequência de números, onde, caso apareça um determinado número, o usuário deve acionar um botão. Para avaliar a carga de stress, os participantes utilizaram o índice NASA (*National Aeronautics and Space Administration*) *Task Load Index* (NASA-TLX), o qual classifica a carga de trabalho percebida após a execução de cada tarefa. Foram consideradas para avaliação os valores de temperatura absoluta mais altos de cada participante. As diferenças térmicas registradas e as respectivas tarefas executadas (carga de trabalho) possibilitaram a criação de um classificador de stress probabilístico. Os resultados evidenciaram uma relação entre o nível de stress e a temperatura facial, porém não foram conclusivos a ponto de estabelecer essa dependência.

A seguir serão listados trabalhos relacionados à detecção de pedestres e, em seguida, será apresentada uma discussão envolvendo os mesmos.

Goubet, Katz e Porikli (2006) realizaram uma análise da eficácia da detecção de pedestres, a partir de experimentos com uma câmera infravermelha baseada em detectores ferroelétricos, comparando o desempenho da detecção em diferentes estações (verão e inverno) e condições de iluminação (dia e noite). Evidenciou-se que a análise puramente termográfica se mostra insuficiente em condições onde a temperatura dos pedestres e do ambiente ao seu redor está próxima (dias quentes), porém sendo mais eficiente em condições de baixa iluminação. Por outro lado, a fusão de dados entre a imagem infravermelha e convencional possibilitou melhor segmentação dentre os cenários estudados, devido ao contraste apresentado entre o pedestre e o plano de fundo da imagem após a fusão.

Krotosky e Trivedi (2007), por sua vez, propõem a fusão entre imagens convencionais geradas a partir de um par estéreo e infravermelhos, para detecção de pedestres. Foi demonstrado que a incorporação de imagens estéreo no treinamento de detectores de pedestres aumentou seu desempenho em relação à detectores treinados apenas com imagens convencionais, ou apenas imagens infravermelhas.

Dickens, Van Wyk e Green (2011) fazem uso da termografia (câmera térmica) aliada a sensores de profundidade (câmera *Time of Flight* - ToF e Xbox Kinect), para

---

<sup>2</sup> Monitor de alertas, que fornece informações ao condutor sem que este precise desviar o olhar do para-brisa do veículo.

identificação de pedestres e prevenção de colisões de veículos em minas. Esse processo se divide em identificação da assinatura térmica do pedestre, estimativa de distâncias e cálculo do tempo de colisão. Demonstrou-se que a segmentação da região de interesse baseada em limiar (*threshold*) de temperatura, a partir da análise da imagem da câmera térmica, foi satisfatória em comparação a outros métodos de segmentação apresentados (método Otsu e segmentação baseada em intensidade e borda). Também foram evidenciados os problemas da câmera ToF relacionados ao cálculo de distância, devido à um borrão de movimento e névoa característica do local, o que demonstra sensibilidade do método a variações do ambiente.

Por fim, Kim e Kwon (2015) apresentam um método para detecção e rastreamento (*tracking*) de pedestres em imagens térmicas. A detecção é realizada após a identificação da temperatura característica dos indivíduos, a qual apresenta contraste em relação ao plano de fundo. O *tracking*, que utiliza como detector de bordas o Histograma do Gradiente Orientado (HOG - *Histogram Oriented Gradient*), analisa o deslocamento dos pedestres com base na comparação das bordas destacadas entre os *frames* do vídeo. O método demonstrado tem como ponto negativo o fato de não tratar sobreposições de pedestres na mesma imagem, como a situação onde duas pessoas caminham lado a lado. Neste caso não foi possível distinguir a presença de dois pedestres no mesmo *frame*.

O Quadro 1 elenca os principais aspectos dos trabalhos referentes à detecção de pedestres.

**Quadro 1 - Principais aspectos de trabalhos relacionados à detecção de pedestres e termografia.**

<b>Autores</b>	<b>Cenário de captação</b>	<b>Técnica</b>	<b>Pontos Positivos</b>	<b>Pontos Negativos</b>
(GOUBET; KATZ; PORIKLI, 2006)	Externo. Verão e inverno. Dia e noite.	Fusão entre imagens infravermelhas e convencionais (unimodais) para segmentação dos pedestres.	A fusão aumentou o desempenho da segmentação de pedestres no verão.	Não trata sobreposições de pedestres na imagem.
(KROTOSKY; TRIVEDI, 2007)	Externo. Dia.	Sobreposição entre imagens infravermelhas e convencionais estereoscópicas para segmentação dos pedestres.	Alto percentual de desempenho na detecção (98%).	<i>Setup</i> complexo (duas câmeras térmicas e duas convencionais) pode limitar a aplicação do método.
(DICKENS; WYK; GREEN, 2011)	Subterrâneo (minas).	Fusão entre imagens infravermelhas e informações de profundidade (câmera <i>Time of Flight</i> ou Microsoft Kinect).	O Kinect demonstrou-se superior para captação da profundidade no ambiente subterrâneo	Câmera ToF sensível a névoa característica do ambiente.
(KIM; KWON, 2015)	Externo. Dia e noite.	Rastreamento de pedestres em sequência de imagens infravermelhas, a partir da eliminação do <i>background</i> e cálculo do HOG.	<i>Setup</i> simples (apenas uma câmera infravermelha) e efetivo para rastreamento.	A eliminação do <i>background</i> está sujeita a erro caso a cena tenha movimento.

**Fonte: Autoria própria.**

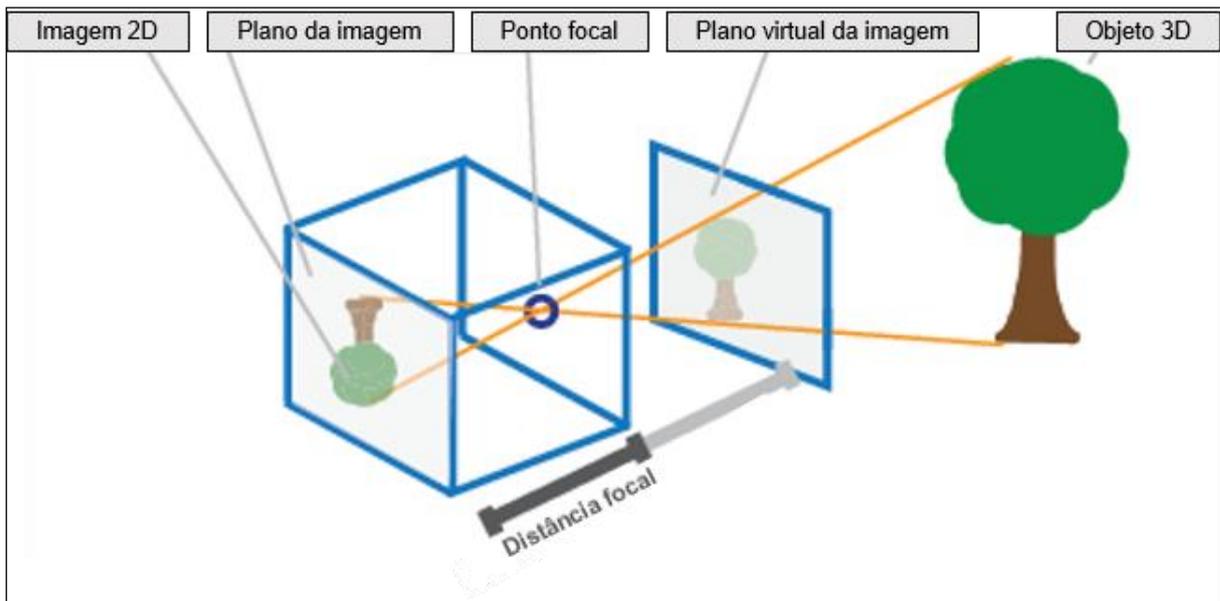
A partir da análise conjunta de trabalhos acerca do mesmo tema, percebe-se a vantagem obtida a partir da incorporação de termogramas à identificação de pedestres. Nesse sentido, permanece um desafio a utilização de fusão de dados provenientes de diversos sensores, de forma a manter a robustez em condições distintas de luminosidade e calor, sem encarecer a solução final para que a mesma seja aplicável no contexto automobilístico.

## 2.2 VISÃO ESTÉREO

A visão estéreo, também chamada de estereoscopia, permite a aquisição de informações tridimensionais e distância do ponto observador até determinado objeto,

a partir de sua visualização de dois ou mais pontos de vista (TRUCCO; VERRI, 1998a, p. 140). Essas informações são adquiridas a partir do estabelecimento de relações geométricas entre a cena observada e a imagem referente à sua projeção. A geometria da visão estéreo é chamada de geometria epipolar, e se baseia no modelo de câmera *pinhole*, ilustrado na Figura 9.

**Figura 9 - Modelo de câmera *pinhole*.**



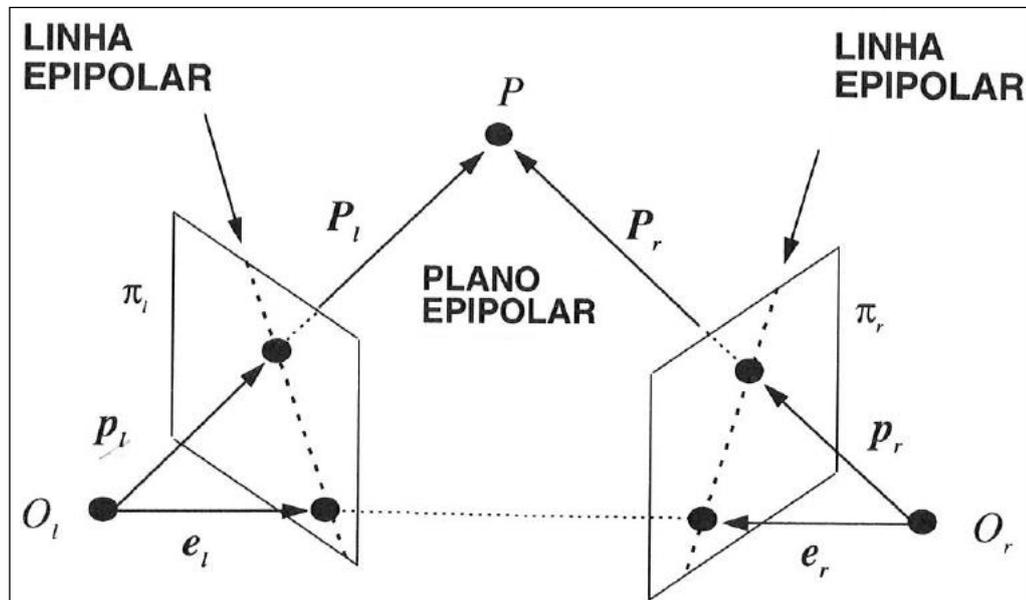
Fonte: Adaptado de Mathworks (2017).

Esse modelo define que em uma câmara escura com um orifício com diâmetro tendendo a zero (ponto focal), a convergência dos raios de luz captados através deste orifício projeta no fundo da câmara a imagem observada (FORSYTH; PONCE, 2012, p. 4). Esse modelo de projeção é importante para se entender como uma imagem nas câmeras atuais é formada, em função do que é captado através da lente da câmera. É interessante perceber que a imagem projetada no fundo também pode ser obtida antes de ser invertida pela projeção através do orifício, como pode ser observado na Figura 9, no plano da imagem e plano virtual da imagem. A distância focal, por sua vez, é a distância entre o ponto de convergência da luz (ponto focal) e a projeção da imagem no plano da imagem. Em câmeras convencionais essa distância é determinada a partir do modelo da lente utilizada.

Com base no entendimento da projeção de uma imagem a partir de uma câmera, pode-se adotar o mesmo modelo, porém para projeções a partir de  $n$  câmeras, e então traçar relações geométricas que possibilitem a descoberta de informações

acerca da posição de um objeto no espaço e suas dimensões. A Figura 10 apresenta um modelo básico da geometria epipolar a partir de dois pontos de vista distintos.

Figura 10 - Geometria epipolar.



Fonte: Adaptado de Trucco e Verri (1998a).

Baseado na Figura 10, pode-se definir terminologias que serão importantes para a compreensão deste trabalho. Os símbolos  $O_l$  e  $O_r$  representam o ponto focal, para as câmeras da esquerda e direita. Os símbolos  $\pi_l$  e  $\pi_r$  representam os planos da imagem. Os caracteres  $e_l$  e  $e_r$  são chamados de epipolos, pois indicam os pontos onde ocorre a interseção entre as projeções dos pontos focais. As linhas representadas por  $P_l$  e  $P_r$  indicam o mesmo ponto nas coordenadas 3D, e as linhas indicadas por  $p_l$  e  $p_r$  indicam as projeções desses pontos nos planos da imagem. As linhas epipolares indicadas na Figura 10 indicam a localização do plano epipolar. Ambos fazem parte da definição de restrição epipolar, a qual diz que para um determinado ponto no plano da imagem de uma das câmeras, o ponto correspondente na outra deve estar ao longo da linha epipolar (TRUCCO; VERRI, 1998a, p. 152). A identificação do exato ponto correspondente ao longo da linha epipolar caracteriza o chamado problema de correspondência.

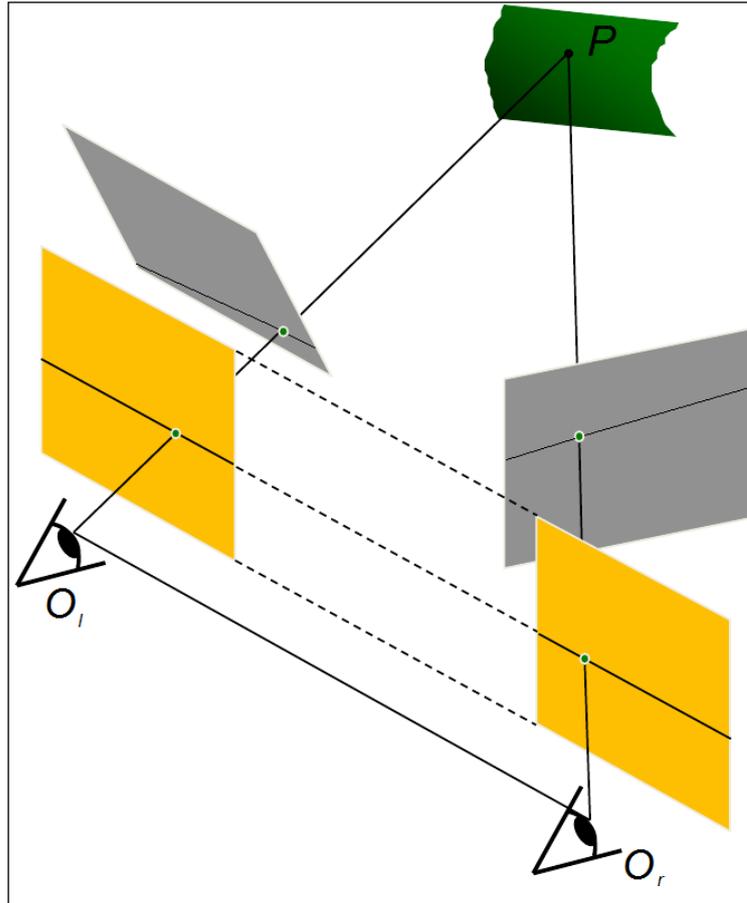
A partir da análise do modelo básico de geometria epipolar para duas câmeras, representado na Figura 10, pode-se iniciar o entendimento de como são traçadas as relações geométricas que correlacionam as imagens captadas em ambos os planos de imagem. O primeiro passo consiste em considerar características das câmeras relacionadas à sua posição no espaço e estrutura física, pois estas afetam diretamente

os cálculos relacionados ao estabelecimento de informações tridimensionais da cena. Essas características são chamadas de parâmetros extrínsecos e intrínsecos, respectivamente, e o processo de identificação desses parâmetros é chamado de calibração de câmera.

A calibração de câmera refere-se ao processo de extração (identificação) dos parâmetros intrínsecos e extrínsecos da câmera. Os parâmetros intrínsecos são distância focal (em pixels), coordenadas do ponto principal (centro do plano da imagem), e tamanho do pixel (vertical e horizontal). Os parâmetros extrínsecos são os valores de rotação e translação que permitem retificar as imagens para um mesmo plano (TRUCCO; VERRI, 1998a, p. 144), e estão relacionados à posição da câmera nas coordenadas do mundo 3D. Esses parâmetros podem ser representados a partir de uma matriz, chamada de matriz fundamental.

A retificação estéreo consiste em alinhar as imagens em um mesmo plano de modo que os pares das linhas epipolares conjugadas tornem-se colineares e paralelos a um dos eixos (TRUCCO; VERRI, 1998a, p. 157). Com isso, é possível promover um alinhamento 2D, de modo a restringir a busca de pontos de correspondência entre as imagens 1 e 2 ao longo da linha epipolar. A Figura 11 mostra um exemplo de retificação estéreo, onde os símbolos  $O_l$  e  $O_r$  representam o ponto focal, para as câmeras da esquerda e direita. O ponto P representa o ponto tridimensional captado pelas imagens. Percebe-se que as linhas epipolares (linhas pretas que cruzam horizontalmente as imagens destacadas em cinza na Figura 11) tornam-se paralelas após a retificação (imagens amarelas na Figura 11).

Figura 11 - Retificação estéreo.



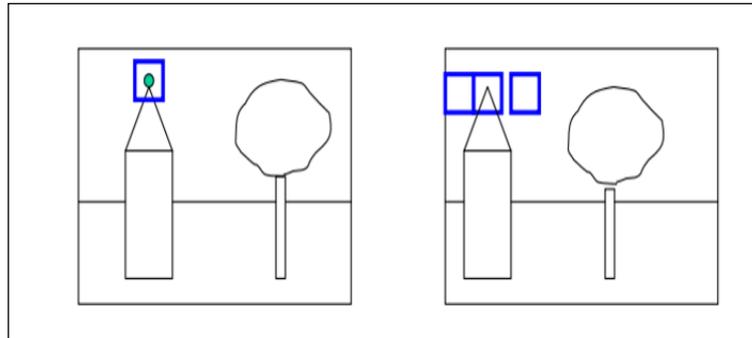
Fonte: Adaptado de Denning (2018).

Matriz fundamental é a matriz capaz de correlacionar um ponto na imagem A do par estéreo com a linha epipolar da imagem B (transformação de espaço) (HARTLEY; ZISSERMAN, 2004a, p. 242), dessa forma otimizando a análise da imagem para resolver o problema de correspondência, que significa identificar qual ponto exato da linha epipolar da imagem B corresponde ao ponto da imagem A. Existem diferentes abordagens para se tratar do problema de correspondência, porém podem ser classificadas de duas formas gerais, sendo elas métodos baseados em correlação e métodos baseados em características (TRUCCO; VERRI, 1998a, p. 145-150).

Os métodos baseados em correlação utilizam sub-regiões da imagem, também chamadas de janelas, para correspondência. A métrica de similaridade é dada pela sub-região que apresenta melhor resultado, dessa forma definindo o ponto correspondente entre as imagens como o centro da janela utilizada (TRUCCO; VERRI, 1998a, p. 146). A Figura 12 apresenta um exemplo de correspondência por correlação,

onde procura-se encontrar o conjunto de pixels na imagem da direita, correspondentes à janela do ponto verde localizado na imagem da esquerda.

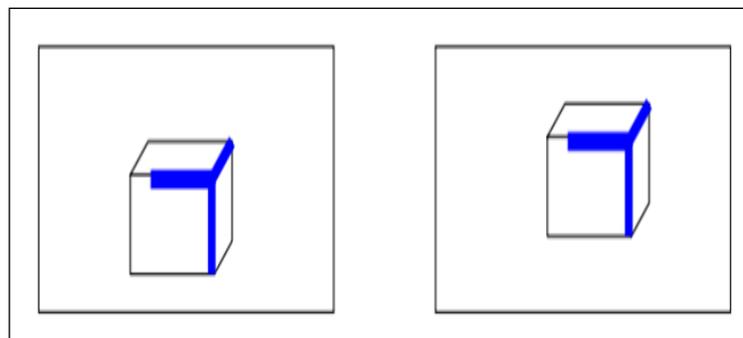
**Figura 12 - Correspondência baseada em correlação. Avaliam-se janelas de pixels na imagem da direita, buscando-se a correspondência com a janela destacada na imagem da esquerda.**



Fonte: Rao (2017).

Os métodos baseados em características, por sua vez, utilizam representações numéricas das características de interesse de uma imagem, como comprimento e ângulo de linhas detectadas na imagem, e estabelecem uma medida de similaridade entre essas representações. Dessa forma é necessária a aplicação de um algoritmo descritor de características (*feature descriptor*), que extraia as informações necessárias para então ser aplicada a correspondência baseada em características (TRUCCO; VERRI, 1998a, p. 148). A Figura 13 exibe um exemplo desse tipo de correspondência, onde procura-se comparar informações referentes às bordas dos cubos em ambas as imagens para se determinar a correspondência.

**Figura 13 - Correspondência baseada em características. Avaliam-se as características relacionadas às bordas da figura na imagem da direita, buscando-se a correspondência com as características extraídas da imagem da esquerda.**

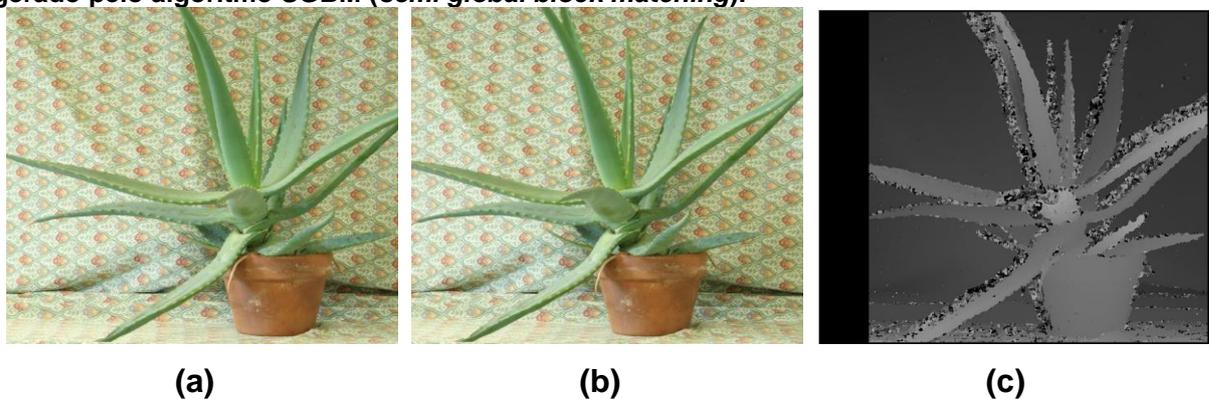


Fonte: Rao (2017).

A partir da correta correspondência dos pontos entre as imagens, e utilizando-se de um processo de triangulação, é possível determinar o mapa de disparidade da

cena. Este mapa pode ser caracterizado como uma imagem que denota a diferença da localização de um mesmo ponto tridimensional projetado a partir da perspectiva de pontos de vista distintos (SHAPIRO; STOCKMAN, 2001, p. 431,432). Se um ponto A em uma imagem 1 está localizado na coordenada (100, 20), e B (135, 20) é seu correspondente na imagem 2, a disparidade será de 35 pixels. Dessa forma, no mapa de disparidade, pode ser aplicada uma normalização nos valores de disparidade para que estes assumam um valor no intervalo de 0 a 255 permitindo sua visualização em tons de cinza. A Figura 14 exibe um exemplo de mapa de disparidade.

**Figura 14 - Geração do mapa de disparidade a partir de imagens retificadas. a) Imagem captada pela câmera da esquerda. b) Imagem captada pela câmera da direita. c) Mapa de disparidade gerado pelo algoritmo SGBM (*semi global block matching*).**



Fonte: Autoria própria.

O mapa de profundidade, por sua vez, é uma imagem que representa as distâncias dos objetos observados em relação ao ponto observador, que pode ser considerado como sendo o ponto médio entre as câmeras. Nesse caso a profundidade é o valor da distância de um ponto da cena ao ponto observador e o mapa é uma matriz de profundidades referentes à cena. O mapa de profundidade é inversamente proporcional ao mapa de disparidade, ou seja, quanto maior a disparidade menor a profundidade. Em imagens retificadas, a profundidade pode ser calculada como demonstrado na equação (5), onde  $Z$  representa a profundidade,  $b$  a distância entre os centros de projeção das câmeras (distância *baseline*),  $f$  a distância focal e  $d$  o valor da disparidade para determinado pixel.

$$Z = bf \frac{1}{d} \quad (5)$$

Isso fica mais claro ao analisar-se a Figura 11, na qual a variável  $b$  da equação anterior representa a distância entre os dois pontos focais  $O_l$  e  $O_r$ , e a variável  $f$

representa as distâncias entre cada ponto focal e a imagem semitransparente (retificada). Essa fórmula não se aplica em imagens não retificadas pois considera que a disparidade foi previamente calculada em imagens alinhadas (retificadas), pois tal processo aumenta a performance do cálculo dos pontos correspondentes na imagem.

Com base no mapa de profundidade é possível então obter um modelo tridimensional que represente a cena captada. A seção seguinte apresenta trabalhos que utilizam conceitos relacionados à visão estéreo em aplicações variadas.

### 2.2.1 Trabalhos Relacionados

Esta seção destaca abordagens relacionadas à estereoscopia com o intuito de ressaltar a importância e aplicabilidade do tema em questão. Ao fim da seção será apresentada uma discussão envolvendo os trabalhos elencados.

Bertozzi e Broggi (1998) propõem o sistema GOLD (*Generic Obstacle and Lane Detection system*), que é uma arquitetura de hardware e software veicular baseada em visão estéreo, voltado para sistemas de auxílio ao condutor. Os autores aplicam em seu sistema o algoritmo IPM (*Inverse Perspective Mapping*), o qual é capaz de remover o efeito de perspectiva de uma imagem, tendo como saída a imagem vista a partir de uma visão aérea. Uma das imagens é utilizada na detecção de faixas a partir da aplicação de operadores morfológicos, enquanto que ambas são utilizadas para o estabelecimento do espaço livre à frente do condutor. O sistema foi avaliado em um veículo de teste por cerca de 3000 km, e apresentou robustez na detecção mesmo em situações adversas de iluminação.

Abraham, Jasiobedzki e Umasuthan (2001) propõem a utilização da visão estéreo em operações relacionadas à proximidade de satélites, utilizando o algoritmo ICP (*Iterative Closest Point*), e realizando uma construção 3D do satélite a partir das imagens capturadas e comparação com um modelo pré-estabelecido do satélite. O processo é dividido em duas fases, uma *offline* e uma *online*. A etapa *offline* é responsável pela calibração das câmeras e determinação do modelo tridimensional do satélite, além do cálculo prévio das distâncias das bordas do modelo para cada unidade de uma *octree*<sup>3</sup>. O modelo *octree* foi utilizado para subdividir o espaço tridimensional em seções pré-determinadas, as quais servem de referência para

---

<sup>3</sup> Estrutura de dados do tipo árvore, onde cada nó que não é folha possui oito nós filhos.

armazenamento das distâncias do modelo 3D. O módulo *online* é responsável pelo *matching* entre as imagens adquiridas e o modelo pré-determinado, e geração de modelo 3D utilizando o algoritmo ICP. Foram realizados testes em um simulador que gerava as condições características do espaço, simulando a movimentação do satélite bem como condições de iluminação. Os autores concluem que o sistema apresentou desempenho satisfatório (40% a 100% de correspondência na geração do modelo), mesmo em condições de baixa iluminação e oclusão.

Cheng, Park e Trivedi (2005) em seu trabalho propõem uma fusão de dados entre pares de câmeras calibradas térmicas e convencionais com o intuito de identificar partes do corpo (cabeça, direção do olhar e mãos) em um cenário de condução de veículo. Foram realizados testes de identificação das partes do corpo baseado em funções comuns realizadas pelo motorista durante a direção do veículo, como trocar de marcha e manobrar, onde o sistema se mostrou confiável na estimativa da posição das partes do corpo (30 mm de erro médio em distância espacial, e cerca de 10 pixels na imagem).

Krotosky e Trivedi (2006) propõem um método capaz de identificar pedestres em cenários onde os mesmos se encontram em diferentes profundidades. Sua abordagem utiliza fusão de dados entre imagens de câmeras convencionais e térmicas calibradas, e aplica maximização de informação mútua (MIM) e janelas móveis na resolução da correspondência entre as imagens. Os resultados de disparidade gerados são comparados com os provenientes de algoritmos baseados apenas em imagens convencionais (coloridas). O algoritmo se mostrou robusto em situações de oclusão, configurando uma abordagem útil em cenários de análise de pedestres.

Prakash, Lee e Caelli (2006) propõem um método de geração tridimensional de um mapa de temperatura, realizando a correspondência entre imagens provenientes de câmeras térmicas calibradas. Essa correspondência é feita a partir da identificação de linhas isotérmicas, ou seja, linhas que apresentam a mesma temperatura, as quais servem de critério para estabelecimento da geometria epipolar nas imagens térmicas. Os testes foram realizados a partir da análise de um cilindro aquecido, onde foi possível distinguir as linhas isotérmicas. Para fins de comparação foi inserido um ruído gaussiano na geração do modelo tridimensional, e comparado com a projeção gerada a partir das dimensões do cilindro. Os autores concluem que

seu método é eficaz para a reconstrução 3D térmica, indicando a temperatura da superfície do cilindro.

Toulminet et al. (2006) propõem uma estratégia focada na detecção e cálculo da distância do veículo à frente, a partir da extração de características relevantes de um modelo tridimensional. É realizado um processo de correspondência de regiões da imagem de uma das câmeras com um modelo de traseira de veículo, desta forma identificando-o na cena. Os testes foram realizados em um veículo instrumentado, onde demonstrou-se a eficácia na detecção da distância dos veículos (com uma margem de erro variável), mesmo em situações adversas como má iluminação.

Huh et al. (2008) propõem um sistema baseado na visão estéreo para detecção de obstáculos em rodovias. O sistema utiliza conceitos de correspondência de características, bem como a restrição epipolar para estabelecimento dos pontos correspondentes entre as imagens da direita e esquerda. Após o estabelecimento da geometria epipolar, são aplicados uma série de técnicas, como o filtro da média e agrupamento de distâncias, de forma a caracterizar os obstáculos na via, para então se estabelecer a distância até eles. O sistema se mostrou capaz de detectar obstáculos à frente, bem como veículos ultrapassando e mudando de faixa. Os experimentos na detecção de obstáculos apresentaram 5% de erro com variação de 1m, em distâncias abaixo de 45m.

O Quadro 2 elenca os principais aspectos dos trabalhos referentes à visão estéreo.

**Quadro 2 - Principais aspectos de trabalhos relacionados à visão estéreo.**

<b>Autores</b>	<b>Cenário</b>	<b>Técnica</b>	<b>Pontos Positivos</b>	<b>Pontos Negativos</b>
(BERTOZZI; BROGGI, 1998)	Externo (rodovia). Dia.	Fusão entre imagens de diferentes pontos de vista (para geometria epipolar). Aplicação do IPM e operadores morfológicos para detecção de obstáculos e faixas.	Tempo de execução aceitável para uso em veículos. Resultados robustos mesmo com variações de luminosidade.	Resultados não confiáveis para mais que 45m. Não foram abordados cenários em condições adversas.
(ABRAHAM; JASIOBEDZKI; UMASUTHAN, 2001)	Vácuo, próximo à órbita da Terra, simulado.	Construção 3D de modelo de satélite utilizando ICP e comparação com modelo pré-estabelecido, para operações relativas à aproximação de satélites.	40% a 100% de correspondência na geração do modelo, mesmo em condições ruins de iluminação, e oclusão.	Sistema fixo para um modelo de satélite. Necessário treinamento <i>off-line</i> dedicado ao modelo.

<b>Autores</b>	<b>Cenário</b>	<b>Técnica</b>	<b>Pontos Positivos</b>	<b>Pontos Negativos</b>
(CHENG; PARK; TRIVEDI, 2005)	Interior de protótipo de veículo.	Fusão entre imagens de pares de câmeras térmicas e comuns para identificação de partes do corpo do condutor (cabeça, direção do olhar e mãos).	30mm de erro médio em distância espacial, cerca de 10 pixels na imagem.	Setup complexo encarece a solução (quatro câmeras térmicas e quatro comuns). Testes apenas em protótipos.
(KROTOSKY; TRIVEDI, 2006)	Exterior. Dia. Via urbana.	Fusão entre imagens de câmeras convencionais e térmicas, aplicando a técnica MIM e janelas móveis na solução da correspondência entre as imagens, aplicados à detecção de pedestres.	Resultados robustos em situações de oclusão, e diferentes profundidades.	Pode não detectar se a pose/tamanho não corresponde com <i>template</i> . Limiar de temperatura. Pode falhar em cenários muito quentes.
(PRAKASH; LEE; CAELLI, 2006)	Ambiente controlado (cilindro aquecido).	Geração de mapa de temperatura 3D, realizando a correspondência entre imagens de câmeras térmicas calibradas.	Setup simples (duas câmeras térmicas fixas, alinhadas e calibradas).	Resultados não são claros. Escopo para avaliação limitado (apenas um cilindro).
(TOULMINET et al., 2006)	Rodovia. Dia.	Detecção e cálculo da distância do veículo à frente, a partir da extração de <i>features</i> de um modelo 3D. Correspondência de regiões com um modelo de traseira de veículo.	Detecção de distância eficaz, mesmo em situações adversas de iluminação.	Detecção de apenas um veículo por vez.
(HUH et al., 2008)	Rodovia. Dia.	Aplicação de <i>feature matching</i> e restrição epipolar para cálculo da geometria epipolar, e uma sequência de técnicas para detectar obstáculos na via.	5% de erro com variação de 1m, em distâncias abaixo de 45m.	Baixa eficácia acima de 65m.

Fonte: Autoria própria.

Com base na análise dos trabalhos supracitados, percebe-se que as informações tridimensionais obtidas a partir da visão estéreo encontram aplicabilidade em cenários distintos, onde algoritmos utilizam os conceitos de geometria epipolar, em conjunto a outras técnicas para diversos fins. Já existem soluções capazes de promover a geração do modelo tridimensional em tempo aceitável, caracterizando sua viabilidade no contexto automobilístico. Encara-se como desafio a utilização destas

técnicas em cenários não ideais, como detecção de múltiplos objetos em cenários de iluminação crítica (como ofuscamento por exemplo).

### 2.3 ALGORITMO GENÉTICO

Os algoritmos genéticos (AG) referem-se a uma classe de algoritmos empregados na computação chamados de meta-heurísticas, que se propõem a utilizar uma abordagem não determinística para a resolução de problemas, muitas vezes não oferecendo a solução ótima para o caso em questão, mas baseado em um percentual de assertividade. Essas meta-heurísticas são assim chamadas pois empregam conceitos de pseudoaleatoriedade em seu funcionamento. Eles pertencem também a uma categoria chamada de algoritmos bio-inspirados, pois têm sua inspiração baseada no funcionamento de processos naturais da biologia. Foram introduzidos por Holland (1992), que em seu trabalho apresenta a fundamentação teórica e aplicações do AG.

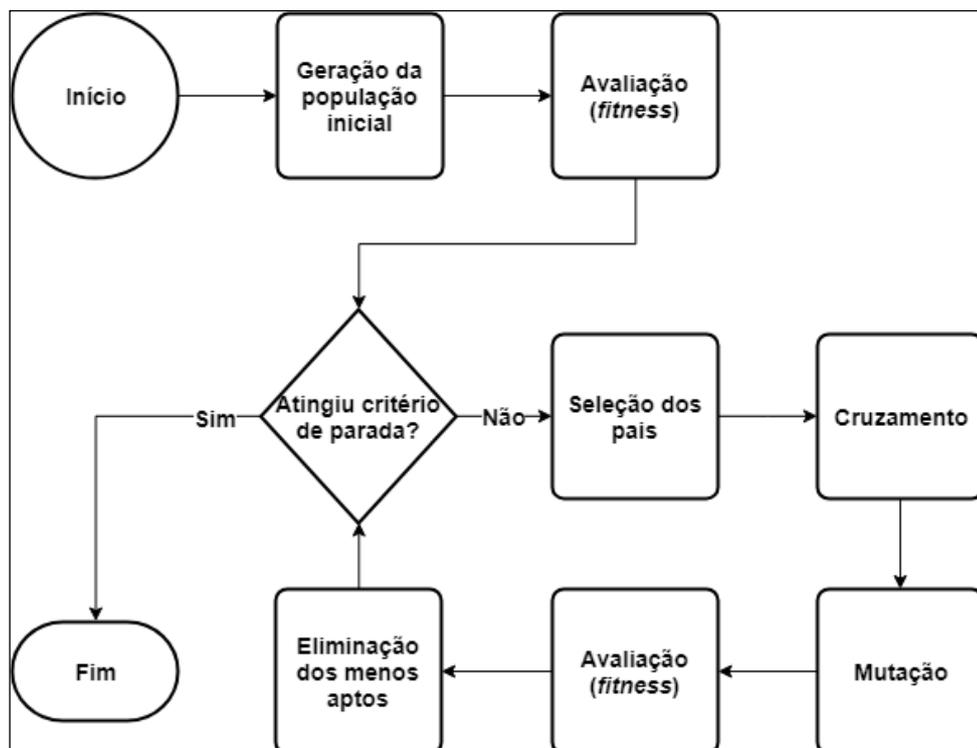
Seu funcionamento tem como base o processo evolutivo das espécies chamado de seleção natural, definido por Darwin, onde há uma disputa entre indivíduos e os mais aptos tendem a sobreviver e reproduzir-se, dessa forma passando seus genes adiante e assim conduzindo a um aprimoramento da espécie. Essa evolução genética apresentada é então aplicada como um problema de otimização (minimização ou maximização de um valor), onde pretende-se melhorar uma medida de avaliação (a qual define o quão apto é um indivíduo), a partir da seleção natural (RUSSELL; NORVIG; DAVIS, 2010). Os seguintes elementos fazem parte da estrutura do AG:

- Gene: Cada gene em um AG representa um dos parâmetros a serem otimizados. Sua representação pode ser binária (mais comum), em valores inteiros ou ponto flutuante. O conjunto de valores dos genes, bem como os novos valores que venham a surgir, precisam estar dentro de uma faixa de valores aceitável para o problema em questão;
- Cromossomo ou indivíduo: O conjunto de genes é chamado de cromossomo ou indivíduo. Ele representa o conjunto de parâmetros que definem uma solução para o problema abordado;

- População: A população, por sua vez, representa o conjunto de indivíduos. É o espaço de possíveis soluções do problema, o qual vai sofrendo combinações e mutações e convergindo para soluções que obtém maior desempenho;
- Geração: A geração compreende uma iteração do AG, ou seja, a execução de todas as fases deste, as quais serão abordadas na sequência. Um dos critérios de parada do AG normalmente é o número de gerações executadas pelo mesmo. Teoricamente quanto maior o número, maior a chance de se otimizar o resultado;
- Convergência: Representa o momento em que as populações começam a ter seus genes repetidos, estagnando em um único conjunto de valores. A convergência geralmente é analisada e definida como um dos critérios de parada dos algoritmos.

A Figura 15 exemplifica o fluxo de funcionamento do algoritmo genético, a ser detalhado na sequência.

**Figura 15 - Funcionamento do algoritmo genético.**



Fonte: Autoria própria.

O AG se inicia com a definição do tamanho da população e dos valores dos genes desta, os quais podem ser definidos aleatoriamente, porém é interessante que

cubram a maior parte do espaço de busca (LINDEN, 2008, p. 64). Na sequência, atribui-se um valor a cada indivíduo da população, que representa seu nível de aptidão no contexto, ou seja, é aplicada uma medida de desempenho, a partir de uma função (função *fitness*), que serve para ranquear os indivíduos da população.

A próxima etapa consiste no processo de seleção dos pais, onde dentro da população atual são escolhidos indivíduos que irão dar origem a um descendente. Esta seleção normalmente emprega o método da roleta ou torneio. No método da roleta, divide-se a probabilidade de seleção de um indivíduo baseado em sua medida de avaliação, ou seja, dentre toda a população, escolhe-se aleatoriamente, porém com maior chance de se selecionar aqueles indivíduos que possuem valor *fitness* maior. No método de torneio, seleciona-se aleatoriamente dois ou mais indivíduos, e o melhor dentre estes torna-se o “pai”. Os indivíduos não selecionados podem ou ser deletados ou mantidos.

Após a seleção dos pais, ocorre a operação de cruzamento (*crossover*) que é baseada no cruzamento genético da biologia, onde combina-se os genes dos pais de um novo indivíduo a ser gerado. Existem diversas estratégias de *crossover*, sendo a mais simples delas a baseada em um ponto de corte, a qual une determinados genes de um dos pais, ao restante de genes do outro progenitor.

Na sequência do algoritmo, é avaliado a possibilidade de ocorrência de mutação em cada gene do novo indivíduo. Essa avaliação pode ser feita a partir do sorteio de um valor entre 0 e 1, e caso o valor sorteado seja menor que a probabilidade definida de mutação (parâmetro do AG), é definido um novo valor para o gene de forma aleatória.

Ao final de cada iteração, pode ser aplicado um critério de eliminação de indivíduos, visando manter a população com o tamanho original. Isso pode ser feito com base em um novo *ranking* dos indivíduos, levando em conta seu valor da função *fitness*. Após a execução das etapas citadas, verifica-se os critérios de parada do algoritmo, que podem ser número de gerações, tempo de processamento, ou similaridade dos indivíduos. Sendo atingido esse critério de parada, seleciona-se o indivíduo que possui o melhor valor *fitness*, obtendo como resultado o valor contido em seus genes.

A seção seguinte irá apresentar trabalhos que fazem uso do algoritmo genético em um contexto semelhante ao deste trabalho.

### 2.3.1 Trabalhos Relacionados

Esta seção destaca abordagens relacionadas à utilização do algoritmo genético no contexto da visão estéreo, com o intuito de ressaltar sua aplicabilidade e eficácia. Ao fim da seção será apresentada uma discussão envolvendo os trabalhos elencados.

Chai e Ma (1998) em seu trabalho estabelecem um *framework* capaz de realizar a correspondência estéreo em imagens não calibradas. Eles utilizam um algoritmo genético para estabelecer as correspondências e recuperar a geometria epipolar da cena, fazendo uso de um novo operador genético, chamado de “adaptação”, que integra esses dois processos. Esse operador favorece a sobrevivência de genes que se enquadram na geometria epipolar da cena. Em seus resultados, é ressaltado a rapidez com que a correspondência é realizada, para até 100 pontos de característica, porém para cenários mais complexos sugere-se a utilização de clusterização do processamento. Além disso, foram demonstrados poucos testes, em condições controladas de iluminação.

Silva, Simoni e Bharadwaj (2000) propõem uma abordagem hierárquica para encontrar pontos correspondentes entre imagens candidatas a formar um par estéreo. O processo possui duas etapas, a de processamento de baixo nível, que encontra um conjunto inicial de pontos correspondentes, baseado em uma análise híbrida (correlação da vizinhança dos pontos e características que definem esta região). Além disso, os autores definem um novo critério de correspondência, sendo esta necessária de ocorrer para um conjunto de pontos simultaneamente, definindo um polígono de correspondência. Neste trabalho o algoritmo genético é utilizado para acelerar o processo de identificação da melhor correspondência dos polígonos, visto que é necessária a análise de diversas combinações em um grande espaço de busca. Os resultados apresentados indicaram uma aceleração no processo de correspondência, contudo não foram detalhados de forma quantitativa as melhorias apontadas.

Também na linha de otimização da correspondência, Issa, Ruichek e Postaire (2002) propõem a utilização do algoritmo genético para aumentar o desempenho na identificação do ponto correspondente ao longo do espaço de soluções. Os cromossomos gerados são considerados válidos caso se enquadrem em duas restrições locais, sendo estas a geométrica, que é baseada na posição da borda detectada, sendo que  $x_l > x_r$ , ou seja, a posição em  $x$  de um ponto identificado deve

ser maior na imagem da esquerda do que seu possível ponto correspondente na imagem da direita, e de inclinação, baseado na correspondência da medição do gradiente dos pontos selecionados. Além disso, são definidas estratégias específicas para se avaliar os cromossomos, baseada na comparação de características dos pixels de pontos correspondentes. Os resultados apresentados indicaram bons índices de correspondência, bem como baixo tempo para convergência da solução do AG, devido à implementação de um esquema de codificação do cromossomo baseado em números inteiros, divergindo da abordagem clássica binária. Apenas um caso de teste foi abordado, não havendo uma análise quantitativa.

Ruichek *et al.* (2004) apresentam uma estratégia baseada em algoritmo genético para a detecção de obstáculos em tempo real. A correspondência estéreo é abordada como um processo de otimização, onde procura-se encontrar a melhor solução que atenda uma série de critérios. O algoritmo genético é então aplicado de forma a analisar o espaço de busca e encontrar a melhor correspondência possível, baseado em um esquema de codificação dos genes utilizando valores inteiros ao invés de binários (tradicional). Para se adequar ao requisito de processamento em tempo real, a busca dos pontos correspondentes é dividida em diferentes níveis baseado na magnitude do gradiente apresentado, começando pelas bordas mais significativas. Os testes realizados a partir de vídeos de veículos em movimento demonstraram a robustez da solução proposta, e adequação ao requisito de tempo de processamento almejado, contudo foi demonstrado apenas um tratamento parcial de oclusões, durante a aplicação do algoritmo genético multinível.

O Quadro 3 elenca os principais aspectos dos trabalhos referentes à utilização do algoritmo genético no contexto de correspondência estéreo.

**Quadro 3 - Principais aspectos de trabalhos que utilizam algoritmos genéticos aplicados à visão estéreo.**

<b>Autores</b>	<b>Cenário</b>	<b>Técnica</b>	<b>Pontos Positivos</b>	<b>Pontos Negativos</b>
(CHAI; MA, 1998)	Dados simulados; cena estática <i>indoor</i> .	<i>Framework</i> genético de correspondência estéreo para imagens não calibradas, baseado em novo operador genético.	Tempo de execução e eficiência para até 100 pontos de características.	Poucos testes exibidos, apenas em condições controladas de iluminação.
(SILVA; SIMONI; BHARADWAJ, 2000)	Cenas <i>indoor</i> . Iluminação controlada.	Identificação da correspondência estéreo baseado na correlação genética de polígonos de pontos	Aceleração no processo de correspondência.	Faltou detalhar quantitativamente as melhorias apontadas.

Autores	Cenário	Técnica	Pontos Positivos	Pontos Negativos
		correspondentes, a partir de abordagem hierárquica e híbrida.		
(ISSA; RUICHEK; POSTAIRE, 2002)	Movimentação de pedestre em uma cena urbana.	Restrições para definição de pontos correspondentes utilizando os conceitos de evolução do algoritmo genético para estabelecimento da correspondência cálculo da profundidade.	Tempo de convergência do AG baixo. Correspondência considerada adequada pelo autor.	É apenas ilustrado um caso de teste. Não há uma análise quantitativa da assertividade.
(RUICHEK et al., 2004)	Tráfego urbano (imagens captadas a partir de um veículo em movimento).	Detecção de obstáculos em tempo real. Utiliza AG para otimizar a análise do espaço de busca na correspondência estéreo.	Comparação quantitativa entre AG binário, com valores inteiros e multinível, apontando o ganho do último no tempo de processamento.	Tratamento parcial de oclusões, quando aplicado o algoritmo genético multinível.

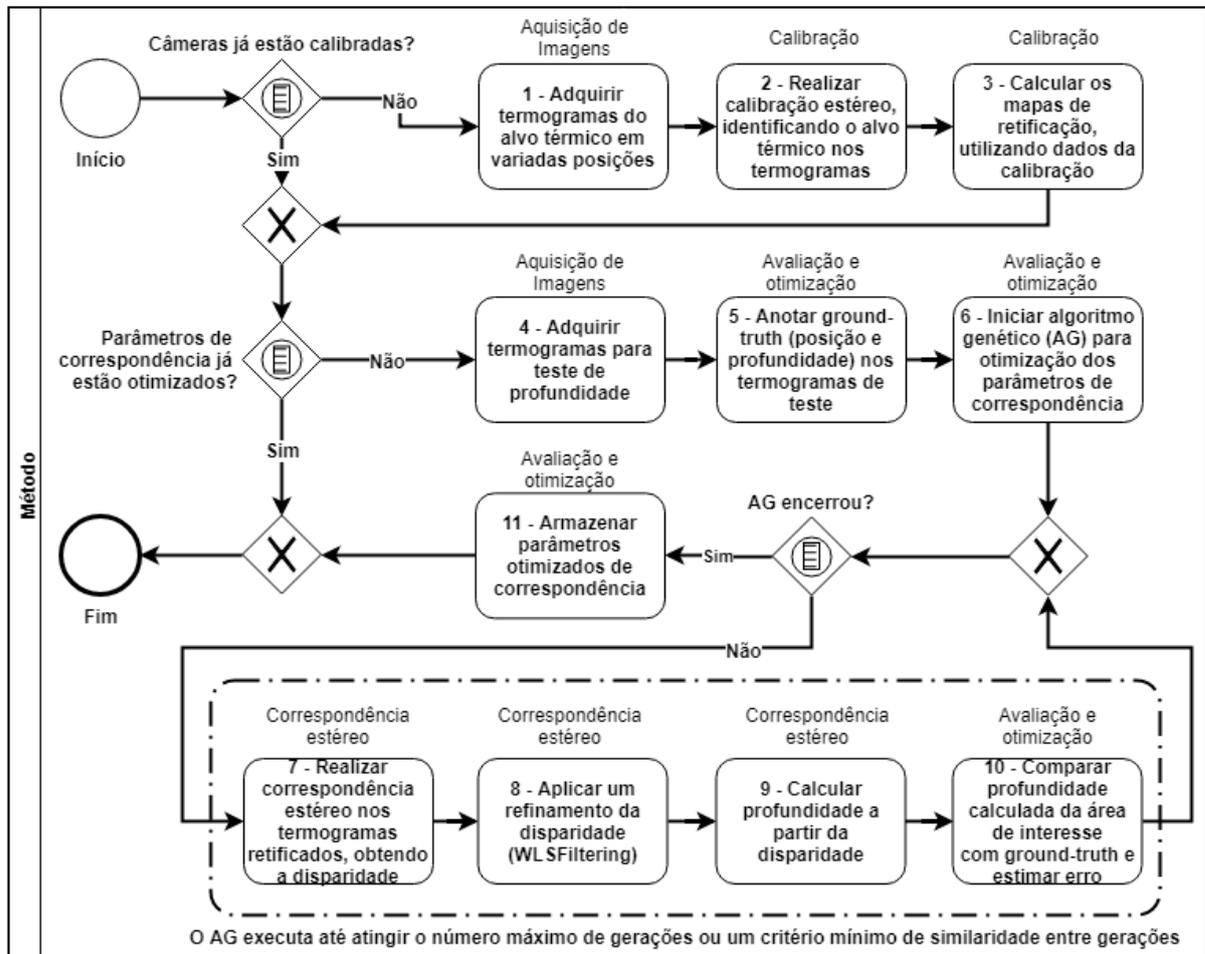
Fonte: Autoria própria.

Baseado na análise destes trabalhos relacionados, percebe-se a que técnicas de algoritmos evolutivos como o algoritmo genético encontram um grande campo de aplicação no contexto de visão estéreo. Nesse sentido, é válida sua utilização no contexto deste trabalho, como um meio de otimizar a parametrização dos algoritmos de correspondência aqui empregados. O próximo capítulo irá abordar o método desenvolvido, bem como a utilização do AG realizada neste trabalho.

### 3 MÉTODO PARA MAPEAMENTO TRIDIMENSIONAL TÉRMICO

A análise dos termogramas representa a obtenção de informações privilegiadas acerca de objetos que contenham uma assinatura térmica distinguível em uma cena. A junção desta informação a conceitos de visão estéreo permite o estabelecimento do chamado mapeamento tridimensional térmico. Neste capítulo serão detalhadas todas as etapas inerentes ao processo de obtenção de um mapa de profundidade a partir de imagens de origem térmica. Para a obtenção deste mapa, é necessário completar uma série de etapas, descritas na Figura 16.

**Figura 16 - Fluxograma proposto. A classificação acima de cada processo indica a seção na qual será abordado.**



Fonte: Autoria própria.

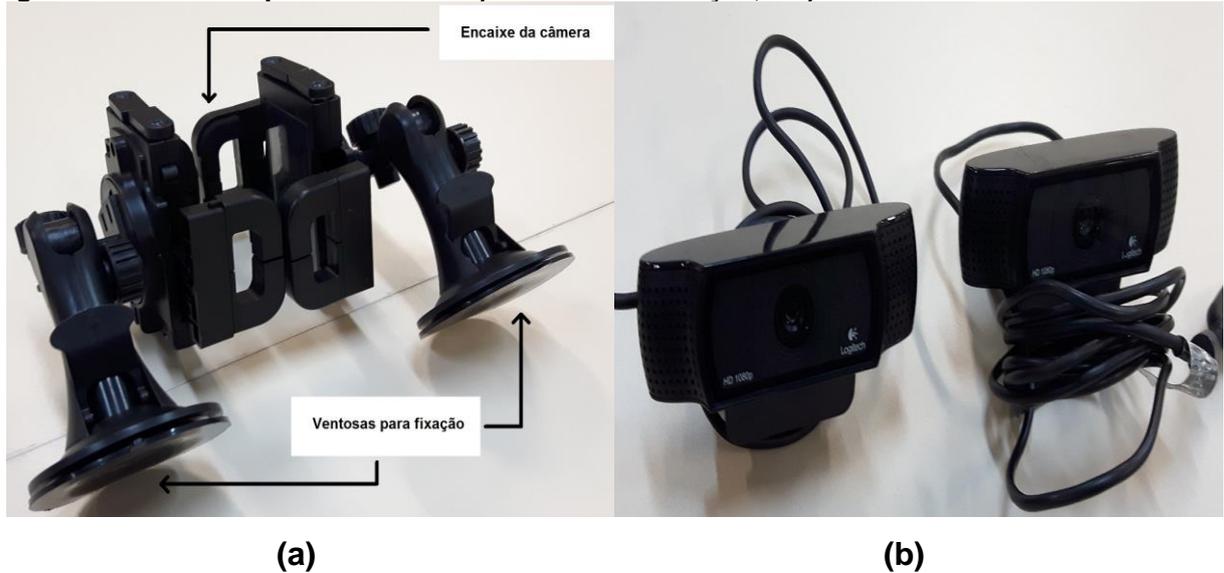
Este capítulo tem como objetivo apresentar o fluxo desenvolvido neste trabalho, indicando os detalhes referentes à cada etapa nas seções subsequentes. Na seção 3.1 é tratado a respeito do processo de aquisição de imagens termográficas, e da estrutura necessária para realizar os experimentos de captura, as quais

compreendem os processos 1 e 4 do fluxograma. Na seção 3.2 aborda-se o tema de calibração e a composição do alvo térmico, que compreendem os processos 2 e 3. Na seção 3.3 é abordado o processo de correspondência estéreo e seu aprimoramento, englobando os processos 7, 8 e 9. Por fim, seção 3.4 é tratado acerca da avaliação do método, bem como da otimização dos parâmetros realizada a partir da utilização de um algoritmo genético, incluindo os processos 5, 6, 10 e 11 do fluxograma.

### 3.1 AQUISIÇÃO DE IMAGENS

Durante o desenvolvimento deste trabalho, foram adquiridas imagens convencionais e termogramas, de modo a demonstrar a execução do método em ambos os casos. No caso das imagens convencionais, utilizou-se duas câmeras Logitech C920, montadas em um suporte configurando assim um par estéreo, conforme exibido na Figura 17, no item a). Esse suporte possui duas ventosas, de modo a permitir a fixação na parte interna do vidro frontal de um veículo. Foram realizados testes com foco fixo, e resoluções de 320x240, 640x480, 960x720 pixels. Em função da posição das câmeras (vertical ao invés de horizontal), foi necessária uma fase de pré-processamento, onde se rotacionam as imagens de modo que estas representem a mesma cena alinhada horizontalmente, de pontos de vista distintos. Essa etapa resultou na alteração da área útil das imagens, mudando suas resoluções para 240x320, 480x640 e 720x960, respectivamente. Esse posicionamento incomum das câmeras no suporte foi adotado em função da minimização da distância *baseline* entre os pontos focais da câmera, que para este suporte foi de 0,07m.

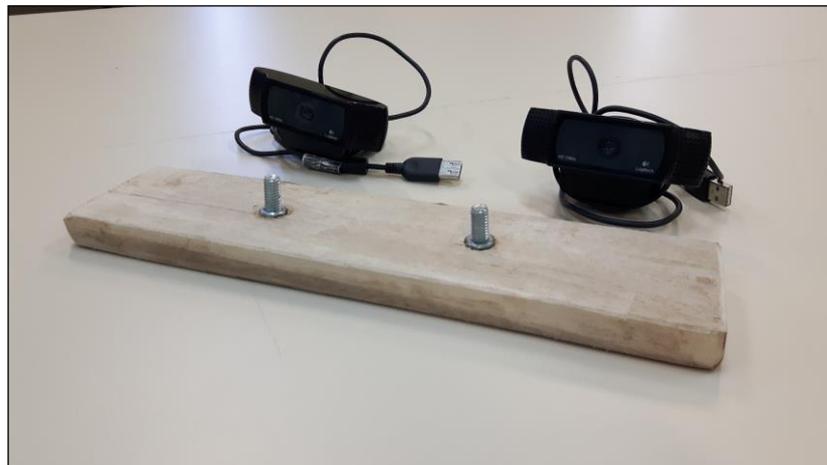
Figura 17 - Partes do par estéreo nº1 a) Estrutura de fixação, e b) Câmeras convencionais.



Fonte: Aatoria própria.

Também foi utilizado um outro suporte, onde as câmeras são posicionadas lado a lado, não sendo necessária a etapa de rotação das imagens capturadas. Este par estéreo nº2 possui distância *baseline* de 0,10m, e é exibido na Figura 18.

Figura 18 - Par estéreo nº2 utilizado.



Fonte: Aatoria própria.

Para a obtenção dos termogramas, utilizou-se a câmera infravermelha FLIR T440, na resolução de 320x240 pixels, a qual possui sensibilidade térmica de 0,045°C à 30°C, e atua na faixa do espectro de 7,5 a 13µm (FLIR, 2018). Os termogramas obtidos não contêm a informação pixel a pixel referente à temperatura, pois essa informação só pode ser extraída a partir dos softwares da FLIR ou utilizando sua SDK, os quais possibilitam a leitura do formato JPEG radiométrico que contém tal informação. A Figura 19 apresenta a câmera térmica utilizada.

**Figura 19 - Câmera térmica FLIR T440.**



Fonte: Ivy Tools (2018).

Para este trabalho optou-se pela análise da temperatura na cena apenas baseado nas variações das cores na imagem, pois este método se assemelha à análise que é feita em imagens convencionais, durante os processos de correspondência estéreo, baseados nos valores de intensidade dos pixels. Foi desenvolvido um software em C++ que integra com a câmera, utilizando a biblioteca OpenCV (BRADSKI, 2000) na versão 3.2.0, exportando o termograma na paleta de cores indicada para uma imagem comum. Nesta imagem as variações de cores indicam as diferentes temperaturas da cena, conforme a escala de temperatura adotada. A temperatura aproximada poderia ser obtida a partir de uma transformação linear, que correlaciona a escala da paleta *grayscale* da câmera térmica e os 256 níveis de cinza da imagem, conforme demonstrado na equação (6), onde  $T$  é a temperatura que se pretende encontrar,  $T_{max}$  e  $T_{min}$  as temperaturas máxima e mínima da paleta do termograma,  $i_{max}$  e  $i_{min}$  os valores máximos e mínimos dos níveis de cinza da imagem (255 e 0), e  $i$  o valor de intensidade (nível de cinza) do pixel analisado.

$$T = \frac{T_{max}}{i_{max} - i_{min}} (i - T_{min}) \quad (6)$$

Para padronização das capturas, utilizou-se a seguinte configuração de parâmetros da câmera térmica, descritos na Tabela 2.

**Tabela 2 - Parâmetros da câmera térmica.**

Parâmetro	Valor
Emissividade	0,98
Distância do objeto	2m a 5m, 10m e 15m
Temperatura atmosférica	20°C
Temperatura mínima da escala da paleta de cores	10°C
Temperatura máxima da escala da paleta de cores	60°C

**Fonte: Autoria própria.**

Como dispõe-se apenas de uma câmera térmica, foi desenvolvido um suporte para movimentação controlada do dispositivo, de forma a obter as imagens do par estéreo. A Figura 20 exhibe o suporte construído, que permite o deslocamento em um dos eixos. A câmera é parafusada na parte superior do suporte, que utiliza o mesmo parafuso padrão para fixação em tripé. A parte superior do suporte pode ser movimentada na direção horizontal. A cada 1cm, é possível travar o deslocamento com a haste de metal ilustrada na figura. Dessa forma, durante os experimentos obtinha-se as imagens térmicas a partir de dois pontos controlados, havendo a movimentação da câmera para que se pudesse obter as imagens do par estéreo térmico. Nos experimentos utilizou-se deslocamento de 1cm, ou seja, a movimentação mínima que o suporte permitia.

**Figura 20 - Suporte para deslocamento controlado da câmera térmica.**

**Fonte: Autoria própria.**

A utilização desse suporte acaba inserindo ruído no processo de calibração e geração da correspondência estéreo, devido principalmente a dois pontos: inicialmente porque há a possibilidade de que a posição da câmera não seja

exatamente a mesma, entre as capturas, para as posições da esquerda ou direita, ou seja, a distância *baseline* pode variar devido ao posicionamento manual do suporte. Mesmo sendo pequena essa variação (da ordem de milímetros), há um impacto na estimativa da correspondência e consequentemente da profundidade calculada, principalmente para objetos mais distantes. Além disso, durante os experimentos procurou-se manter a cena das imagens o mais estável possível entre as capturas da imagem da direita e esquerda, porém até mesmo pequenas modificações, seja na posição da pessoa de referência, ou no ambiente, ou na temperatura apresentada, podem levar a distorções na correspondência estéreo, pois haverá diferenças entre a imagem da esquerda e direita, que a princípio deveriam retratar exatamente a mesma cena, porém de pontos de vista distintos.

Dispondo desse aparato, foram realizados diversos experimentos de captura, que serão detalhados no Capítulo 4, os quais geraram *datasets* convencionais e térmicos a serem utilizados posteriormente.

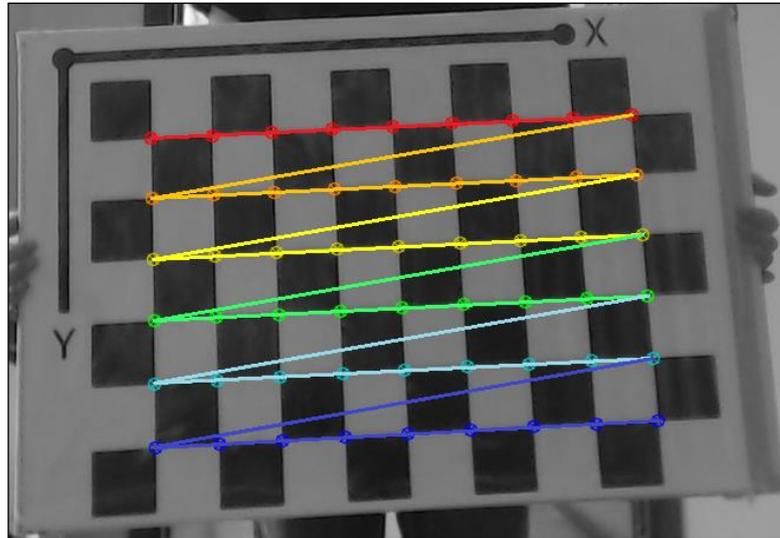
## 3.2 CALIBRAÇÃO

Após a definição da estrutura necessária para a realização dos experimentos, e, portanto, a aquisição das imagens, o primeiro passo necessário no fluxo deste método consiste na calibração das câmeras utilizadas. Conforme descrito na seção 2.2, existem diversos métodos na literatura que tratam do problema de calibração. No nosso caso, utilizou-se o método de Zhang (2000), amplamente utilizado e que já está incorporado nos módulos de visão computacional da biblioteca OpenCV. Além disso, neste trabalho não se aborda a calibração radiométrica, e do sensor térmico, abordando-se apenas a calibração geométrica, que configura uma etapa fundamental no fluxo estéreo utilizado.

O processo de calibração utilizado consiste na análise de imagens de um alvo com um padrão de calibração de geometria conhecida e que possua características visuais que possam ser localizadas facilmente. Para esse processo normalmente utiliza-se um objeto planar com dimensões conhecidas tal como um tabuleiro de xadrez (TRUCCO; VERRI, 1998b, p. 124), como de fato foi feito durante os experimentos. Para a calibração convencional, utilizou-se um padrão de tabuleiro de xadrez, ilustrado na Figura 21, que apresenta 9x6 quinas de quadrados de dimensão

(9 quinas internas na horizontal por 6 quinas internas na vertical), totalizando 54 quinas, com 7,45cm de comprimento de cada quadrado. Esse formato foi escolhido devido a possibilitar a identificação do padrão de tabuleiro mesmo em imagens onde o alvo se encontra distante das câmeras.

**Figura 21 - Alvo de calibração convencional, com as quinas destacadas em colorido.**



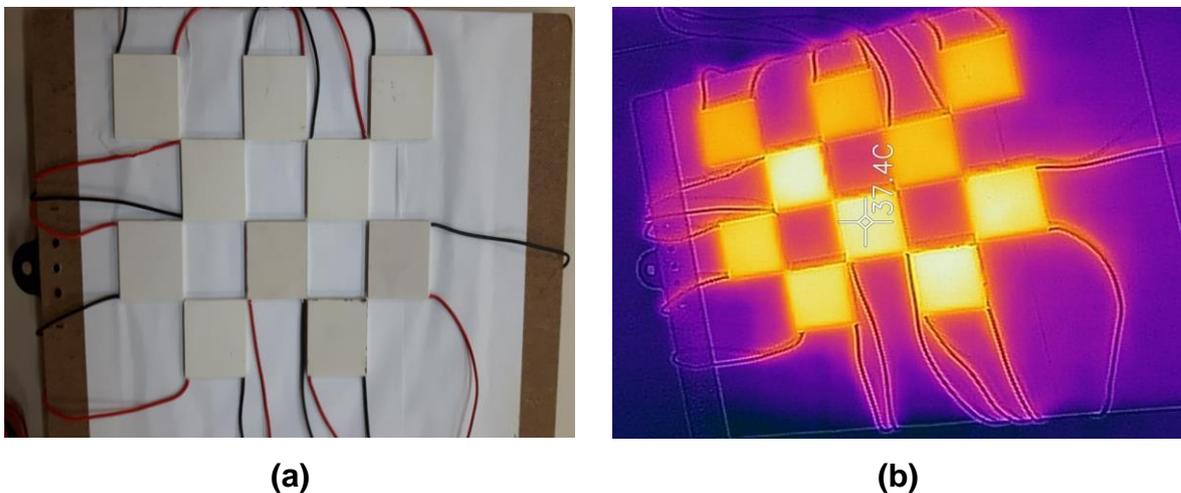
Fonte: Autoria própria.

Para a calibração geométrica dos termogramas, foi necessário utilizar um alvo de calibração apropriado onde seja possível destacar as diferenças de temperatura. Saponaro et al. (2015), utiliza um alvo impresso de papel semelhante ao da Figura 21, que é colocado sobre uma peça de cerâmica revestida e lisa, de dimensões suficientes para comportar o alvo. Após a sobreposição, o alvo e a cerâmica são aquecidos por uma lâmpada direcionada de 250W. Deste modo as regiões de maior calor na cerâmica equivalem aos quadrados pretos da Figura 21, possibilitando a calibração das câmeras térmicas. A cerâmica foi escolhida devido a ser acessível, e por apresentar baixo *momentum* térmico<sup>4</sup>, o que faz com que demore para esquentar e esfriar. Baseado nessa característica, nos experimentos apresentados, o alvo se mostrou utilizável por cerca de 10 minutos, com aquecimentos realizados durante faixas alteradas de tempo.

<sup>4</sup> Conceito físico, que indica a velocidade de aquecimento/resfriamento de determinado objeto.

Neste trabalho, utilizou-se um alvo desenvolvido por um aluno do mesmo laboratório de pesquisa, baseado em placas de Peltier<sup>5</sup> em uma superfície plana, caracterizando o formato de alvo de calibração, conforme detalhado na Figura 22. O efeito Peltier, que dá nome às pastilhas, tem como característica principal o aquecimento de um dos lados da pastilha e esfriamento do outro lado, mediante inserção de corrente elétrica. Baseado no mesmo princípio, também seria possível gerar energia elétrica mediante aquecimento de um dos lados da pastilha e resfriamento do lado oposto. Esse alvo térmico é conectado em uma fonte de alimentação, e após sua ativação as pastilhas já iniciam o aquecimento. O alvo possui 4x3 quinas de quadrados de dimensão, totalizando 12 quinas a serem detectadas, com 4cm de comprimento de cada quadrado. Esse tamanho foi definido em função da limitação da quantidade de placas acessíveis. Devido ao menor número de quinas a serem detectadas em comparação ao alvo convencional, é necessário um maior número de imagens em posições distintas a serem analisadas para que os resultados obtidos sejam satisfatórios, pois normalmente utiliza-se alvos com a dimensão 9x6 (54 quinas).

**Figura 22 - Alvo de calibração térmico utilizando pastilhas de Peltier, em a) Imagem convencional e b) Termograma.**



**Fonte: Autoria própria.**

Observou-se a partir dos experimentos que há uma pequena variação da temperatura das células no decorrer do tempo, havendo um aumento contínuo da temperatura. Para averiguar este comportamento, uma das células do alvo térmico

<sup>5</sup> Pastilha termoelétrica capaz de gerar o efeito Peltier, caracterizado pela produção de uma alteração de temperatura mediante inserção de tensão elétrica em um circuito contendo semicondutores de materiais diferentes.

teve sua temperatura registrada a cada 30 segundos, por um período de 25 minutos. Destes, foram considerados na análise as informações do minuto 5 até o 25, pois o período inicial de aquecimento não é considerado nos testes de calibração. A média de temperatura calculada foi de aprox. 39,568°C, com desvio padrão de aprox. 1,103°C. Além disso, há uma diferença de temperatura de cerca de 3°C entre conjuntos de pastilhas presentes no alvo, devido à distribuição do circuito elétrico que cada conjunto contempla. Essas diferenças também podem levar ao não reconhecimento do alvo, dependendo de como está estruturado o processo de calibração (e conseqüentemente análise da imagem onde está presente o alvo). No caso deste trabalho, para as imagens térmicas onde a paleta de cores era a escala de cinza, essa pequena variação não teve grande influência no processo de calibração, pois, a escala de temperatura fixa adotada (10 a 60°C) minimiza a diferença de tom de cinza que estas temperaturas apresentam na imagem.

Todas as imagens recebidas pelo algoritmo são tratadas como imagens monocromáticas (tons de cinza). Além disso, para que se pudesse utilizar a mesma rotina de calibração tanto para o termograma quanto para as imagens convencionais, aplicou-se nas imagens de origem térmica o efeito chamado negativo, onde a escala de cores da imagem é invertida (temperatura maior agora fica representada pela cor preta, e menor pela cor branca). Dessa forma, as células de Peltier, que baseado na configuração da paleta de cores térmica, inicialmente eram brancas, são exibidas na cor preta, aproximando a imagem da figura de um alvo convencional. No capítulo 4 serão detalhadas quais as paletas de cores utilizadas para aquisição das imagens térmicas.

O processo de calibração é executado baseado na leitura de sequências de pares de imagens (referente à visão da esquerda e direita) que exibem um alvo de calibração convencional ou térmico, dependendo da natureza do teste executado. Esses pares de imagens são então transformados em escala de cinza, e se identifica a posição de cada interseção dos quadrados do alvo. Essas intercessões recebem uma numeração crescente, baseada na posição na imagem, lida da esquerda para direita e de cima para baixo.

As coordenadas (em pixel) das quinas identificadas são armazenadas em listas de vetores de coordenadas, onde para cada imagem há uma sequência de coordenadas das quinas, tanto para a imagem da esquerda como da direita. Além

disso, são geradas as coordenadas do mundo real, considerando o tamanho real (em alguma unidade de distância) do lado do quadrado do alvo de calibração. Essas informações são utilizadas durante a calibração do OpenCV, que aplica um processo de minimização de erro, gerando uma matriz que realiza uma transformação geométrica capaz de se adequar às diversas transformações espaciais geradas em função da aquisição de imagens do alvo em diferentes profundidades e ângulos.

A saída desta etapa consiste em matrizes e vetores que serão utilizados nas próximas fases. Esses dados são armazenados em arquivo de modo a permitir a execução separadamente da calibração e correspondência estéreo. Vale lembrar que esta etapa do método, devido à natureza dos cálculos realizados, consiste no maior gargalo de tempo de execução, pois quanto mais imagens utilizadas, ou quanto maior a resolução da imagem ou do alvo, mais difícil computacionalmente é gerar matrizes que generalizem as transformações geométricas envolvidas. O Quadro 4 exibe os dados armazenados em arquivo, e uma breve descrição de seus significados no contexto.

**Quadro 4 - Dados de saída da calibração.**

<b>Parâmetro</b>	<b>Significado</b>
M1 e M2	Matrizes dos parâmetros intrínsecos para cada câmera (esquerda e direita).
D1 e D2	Vetores de distorção para cada câmera.
R	Matriz de rotação entre os sistemas de coordenadas de cada câmera.
T	Vetor de translação entre os sistemas de coordenadas de cada câmera.
R1 e R2	Matrizes de transformação de retificação, para ambas as câmeras.
P1 e P2	Matrizes de projeção para o novo (retificado) sistema de coordenadas.
Q	Matriz de mapeamento da disparidade para a profundidade.
mapl1, mapl2, mapr1 e mapr2	Mapas de remapeamento, indicando as novas coordenadas em x e y, para as imagens da câmera da esquerda e direita.
validRoi_0 e validRoi_1	Região de interesse válida para as imagens da esquerda e direita, após retificação.

**Fonte: Autoria própria.**

Após a execução da calibração, e cálculo dos mapas de retificação, é informada uma medida de erro de reprojeção, capaz de avaliar a qualidade da

calibração realizada. Essa medida se baseia na distância entre um ponto no mundo real 3D, e o ponto que é estimado pela calibração (HARTLEY; ZISSERMAN, 2004b). Experimentalmente verificou-se que mesmo um erro de reprojeção baixo, ou seja, próximo à 1,00 ou inferior, não significa que a calibração será bem-sucedida, alcançando as transformações desejadas nas próximas etapas, pois é possível se obter um bom percentual mesmo com poucas imagens de calibração. Nessa etapa uma sugestão para se alcançar melhores resultados é a análise do erro de reprojeção calculado, em conjunto com a avaliação visual da retificação de algumas imagens de calibração, para se constatar se sua transformação geométrica está de acordo com o esperado.

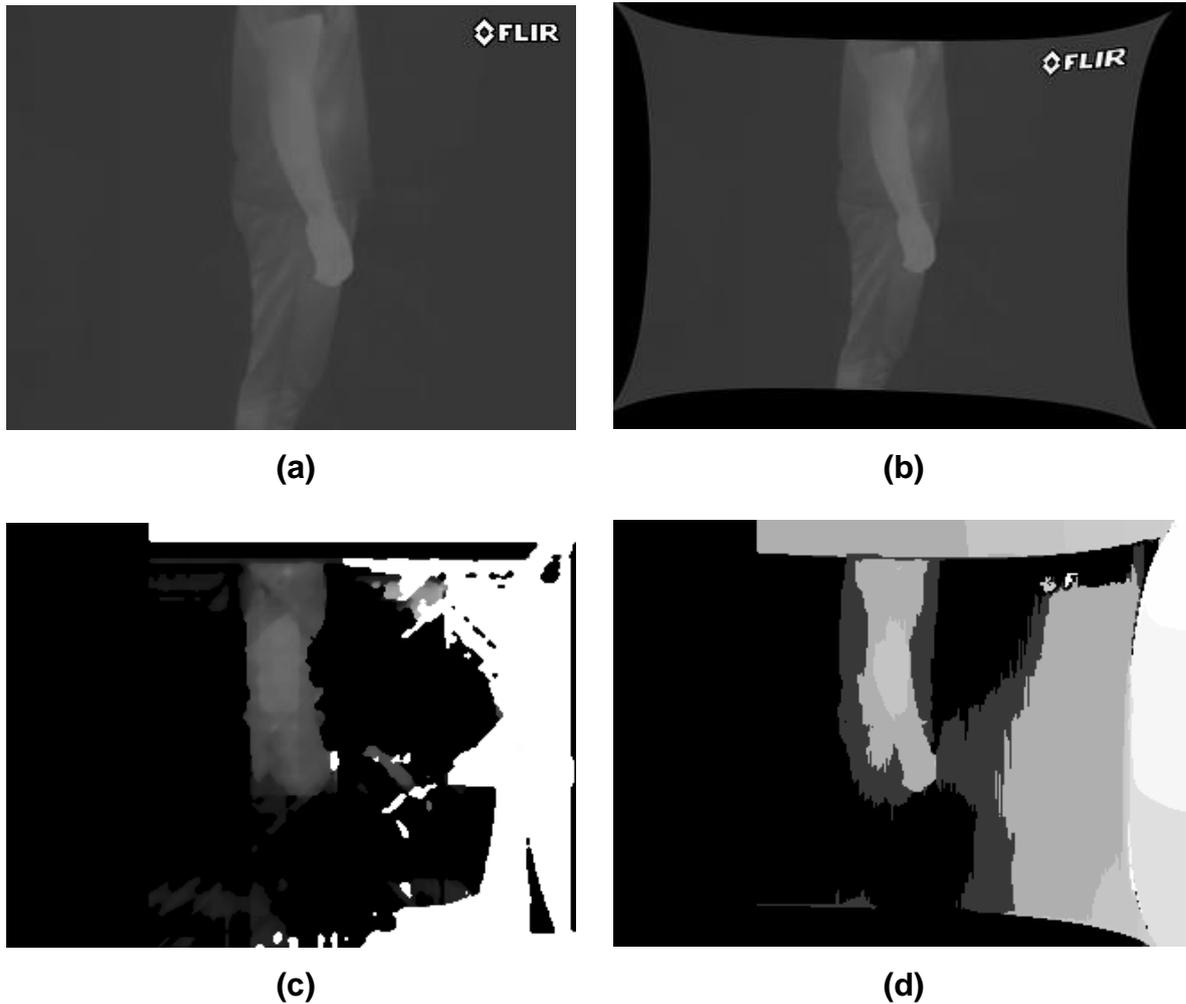
De posse das informações de saída da calibração, a próxima etapa consiste na correspondência estéreo, a ser abordada na próxima seção.

### 3.3 CORRESPONDÊNCIA ESTÉREO

O processo de correspondência estéreo tem início com o carregamento das informações provenientes do processo de calibração, ou seja, as matrizes de calibração. Com essas matrizes é possível realizar o processo de retificação, que consiste em alinhar as imagens em um dos eixos. Realizado este processo, a correspondência é feita utilizando uma técnica chamada *semi global block matching* (SGBM), disponível nas bibliotecas de visão computacional do OpenCV. Como saída desta etapa, obtém-se a imagem de disparidade, que consiste em uma imagem em tons de cinza que representa as distâncias entre os pontos onde ocorreu a correspondência, de forma que quanto maior a distância em pixels dos pontos que correspondem ao mesmo objeto observado, mais claro o pixel é representado na escala de tons de cinza.

Essa imagem de disparidade é então submetida a um filtro de refinamento, onde se aplica a técnica WLS (*weighted last squares*). A partir dela obtém-se uma imagem de disparidade refinada, com atenuação de ruído e, portanto, maior confiabilidade. A Figura 23 exemplifica as etapas realizadas para se obter o mapa de disparidade, a partir de imagens térmicas.

Figura 23 - Etapas para estabelecimento da profundidade a) Imagem original da câmera na posição esquerda, b) Após retificação, c) Mapa de disparidade original e d) Mapa de disparidade após aplicação do filtro WLS.



Fonte: Autoria própria.

Na sequência, a imagem refinada é reprojeta para o 3D, utilizando-se da matriz  $Q$  (matriz de projeção da disparidade para profundidade). Com essa projeção, é possível extrair então a profundidade (em metros, por exemplo) da região de interesse de uma imagem. Na próxima seção será tratado da definição dessa região de interesse, bem como da avaliação da profundidade calculada.

### 3.4 AVALIAÇÃO E OTIMIZAÇÃO

Baseado no desenvolvimento realizado até então, percebeu-se que a parametrização dos métodos de calibração e correspondência tem grande influência

na exatidão da profundidade calculada. Existem diversos pontos que precisam ser corretamente parametrizados para que o resultado seja definido como aceitável em diferentes escopos, ou seja para que se obtenha um cálculo de profundidade com uma medida de erro minimamente aceitável. Além disso, para fins de determinação da acurácia dos resultados obtidos, foi necessário o desenvolvimento de uma métrica de modo a estabelecer os resultados, a serem discutidos no Capítulo 4. Esta seção irá detalhar os processos de definição dos parâmetros da correspondência estéreo, bem como o método utilizado para sua avaliação.

Tomando como objeto de estudo o resultado do método, que é o mapa de profundidade, percebemos que sua avaliação poderia contemplar diversos eixos, como precisão das distâncias calculadas e tempo de processamento. Como o contexto de aplicação deste trabalho tem como foco a identificação de pedestres, para a avaliação foi desenvolvido um módulo onde o usuário pode fazer a marcação (*bounding box*) da posição onde o pedestre se encontra na cena. Conforme descrito no fluxograma na Figura 16, a geração do GT (*ground-truth*) é uma etapa fundamental na execução do algoritmo, identificando assim a pessoa de referência e a profundidade real na qual se encontra, em relação ao ponto médio da distância *baseline* onde as câmeras estão posicionadas. Essa distância real foi medida com uma trena e marcada no solo, de modo que os voluntários que auxiliaram nos testes se posicionavam em frente a cada marcação, para que pudesse ser feita a captura em diferentes profundidades. O procedimento executado em cada captura será detalhado no Capítulo 4.

Após a geração do *ground-truth* para todas as imagens dos testes, foi possível avaliar o método, sendo que esta avaliação consiste na média dos valores de profundidade dos pixels, posicionados dentro do *bbox* (*bounding box*) definido (desconsiderando as bordas).

Para a extração dessa região, inicialmente desenha-se o *bbox* antes da retificação das imagens, para que o retângulo também sofra as distorções geradas a partir do processo de calibração. Após a retificação, binariza-se a imagem, onde todo ponto em uma faixa de vermelho pré-definida torna-se 1 e os demais 0. Na sequência aplica-se uma técnica de inundação de regiões (*floodFill*), para dividir a região de interesse das demais. Em seguida classifica-se as regiões (*labeling*) de modo que a região que não é o *background* da cena (baseado no valor do pixel na coordenada

$[0,0]$ ) é considerada como a do *bbox*. Tomou-se o cuidado durante o processo de geração do GT para que os retângulos não se sobrepusessem, da mesma forma que se avaliou a presença do tom de vermelho do *bbox* nas imagens para não gerar erros no processo de segmentação. Convencionou-se gerar o *bbox* apenas nas imagens da câmera da esquerda do par estéreo.

É interessante ressaltar que nem todos os pixels do mapa de profundidade apresentam valores válidos, sendo que a média só considera os valores ditos válidos. Para cada *bounding box* da imagem avaliada, o valor médio de profundidade obtido era comparado com o valor real anotado no *ground-truth*, obtendo-se um valor de desvio. Baseado nisso, para cada conjunto de dados de teste, calculava-se a raiz do erro médio quadrático (*root mean squared deviation* - RMSD), que é uma medida amplamente usada em contextos onde se mede o desvio de um valor obtido com o esperado. O RMSD é calculado conforme consta na equação (7), onde  $p_c$  é a profundidade calculada pelo algoritmo,  $p_r$  a profundidade real, definida no *ground-truth*,  $b$  é o índice referente aos *bounding boxes* avaliados nas imagens, e  $B$  é o número total de *bounding box* a serem avaliados, considerando todos presentes nas imagens do *dataset*.

$$RMSD = \sqrt{\frac{\sum_{b=1}^B (p_c - p_r)^2}{B}} \quad (7)$$

Com a obtenção de uma medida clara de desempenho, verificou-se a necessidade de otimização dos parâmetros de correspondência, presentes no Quadro 6, de modo a obter resultados consistentes com autores correlatos. Dentre as possíveis escolhas, optou-se pela utilização do algoritmo genético (AG), pois a abstração dos genes convém com a estrutura dos parâmetros que precisa ser otimizada. Baseado nisso, foi desenvolvida uma estrutura que estabelecia limites mínimos e máximos para os valores possíveis dos genes, onde cada parâmetro a ser otimizado é representado por um gene. A medida de comparação do algoritmo genético (*fitness*) se baseou no RMSD calculado. Para cada indivíduo de determinada geração do AG, calcula-se o RMSD do *dataset* utilizado considerando o conjunto de genes (parâmetros) do indivíduo, e este valor servia de base para a manutenção ou não da existência do indivíduo e possível mutação dos valores.

O AG é executado até que se encerre um número pré-definido de gerações, ou baseado em um critério de similaridade entre o RMSD do melhor indivíduo de uma geração, e os demais indivíduos. Caso a diferença entre os RMSD citados seja menor que um determinado parâmetro, é encerrada a execução. Após sua finalização, são exibidos os valores dos genes (parâmetros) que obtiveram o menor RMSD no conjunto avaliado. Dentre todos os parâmetros do fluxo deste método, optou-se pelos relacionados à correspondência estéreo, pois o processo de cálculo da medida de avaliação do AG (*fitness*) necessita que cada indivíduo seja avaliado dentro da população, portanto é necessário executar todos os testes novamente. Como mencionado na seção 3.2, o processo de calibração é identificado como o gargalo de processamento, pois é o mais pesado computacionalmente, portanto essa etapa iria tornar inviável a otimização de seus parâmetros, da forma como este método foi projetado. O Quadro 5 indica os parâmetros do algoritmo genético, lembrando que o número de gerações variou nos experimentos, pois houve casos onde o critério de similaridade do algoritmo genético fez com que não executasse todas as gerações.

**Quadro 5 - Parâmetros do algoritmo genético.**

Parâmetro	Valor	Significado
int generations	20	Número de gerações do AG. Critério de parada.
int populationSize	1000	Tamanho da população, quantidade de indivíduos de cada geração que são avaliados.
int discardSize	250	Taxa de mortalidade dos indivíduos. Quantidade de indivíduos que é descartada de geração em geração.
float mutationProb	0,01	Probabilidade de ocorrer mutação em cada gene, à cada recombinação.
vector<MyParam*> genValidatorVec	-	Vetor dos parâmetros de correspondência estéreo, que representa os genes de cada indivíduo. Será detalhado no Quadro 6.
float minimumSimilarity	0,00001	Similaridade mínima, a ser avaliada ao fim de cada geração. Critério de parada.

Fonte: Autoria própria.

O Quadro 6, por sua vez, exhibe os genes do algoritmo genético, ou seja, os parâmetros de correspondência estéreo a serem otimizados. Durante as recombinações e mutações, é aplicado um validador para que os novos parâmetros criados sejam válidos. Por exemplo, o parâmetro *numberOfDisparities* precisa ser múltiplo de 16. Dessa forma, caso seja gerado um valor não válido, este parâmetro é incrementado em 1 até que adquira um valor válido. As descrições dos parâmetros se baseiam na documentação do OpenCV.

**Quadro 6 - Parâmetros de correspondência estéreo que são genes do AG.**

Parâmetro	Significado	Valor mínimo	Valor máximo
int P1	Parâmetro relacionado à suavidade da disparidade.	1	5000
int P2	Parâmetro relacionado à suavidade da disparidade.	1	5000
int blockSize	Tamanho do bloco avaliado pelo algoritmo.	3	20
int disp12MaxDiff	Máxima diferença em pixels permitida na avaliação da disparidade entre as imagens da esquerda/direita.	0	5
int minDisparity	Valor mínimo de disparidade possível.	0	600
int numberOfDisparities	Disparidade máxima menos a disparidade mínima.	16	304
int preFilterCap	Valor de truncamento para os pixels da imagem pré-filtrada.	1	500
int speckleRange	Parâmetro de uma função de pós processamento, referente à filtragem de ruído disperso.	1	2
int speckleWindowSize	Parâmetro de uma função de pós processamento, referente à filtragem de ruído.	0	200
int uniquenessRatio	Percentual relativo à seleção da função de custo calculada, para definir se a correspondência é correta.	5	15
int remapInterpolationFlag	Tipos de interpolação utilizadas na função <i>rema</i> .	1	4
float wlsLambda	Parâmetro utilizado no algoritmo WLS. Define a regularização durante a filtragem. Valores maiores forçam as bordas do mapa de disparidade a aderirem à imagem original.	1000,0	16000,0
float wlsSigma	Parâmetro utilizado no algoritmo WLS. Define a sensibilidade à ruídos referentes às bordas da imagem filtrada.	0,8	2,0

Fonte: Autoria própria.

Ao fim da execução do algoritmo genético, o indivíduo selecionado com melhor medida *fitness* apresenta o conjunto de genes com os melhores valores até então, referente aos parâmetros apresentados. No capítulo seguinte serão abordados os procedimentos adotados acerca da avaliação deste trabalho, a partir de perspectivas qualitativa e quantitativa.

## 4 EXPERIMENTOS E RESULTADOS

Neste capítulo será tratado acerca da eficácia do método proposto, a partir de dois aspectos, qualitativo e quantitativo. O aspecto qualitativo visa identificar a utilidade do método em situações de iluminação adversas, como ofuscamento e escuro, e discutir resultados obtidos durante etapas do processo a partir da análise das imagens resultantes. O aspecto quantitativo foca na análise dos resultados a partir do ponto de vista das métricas estabelecidas para medição das distâncias.

### 4.1 ANÁLISE QUALITATIVA

A Figura 24 exibe lado a lado a mesma imagem de uma pessoa adquirida a partir de espectros diferentes, em uma condição de pouca iluminação. Essas imagens registram o mesmo instante, visto que foram adquiridas a partir da câmera térmica da FLIR, que com um disparo registra simultaneamente a imagem de seu sensor térmico e a imagem convencional.

**Figura 24 - Pessoa em espectros distintos - pouca iluminação. a) Imagem original, e b) Termograma.**

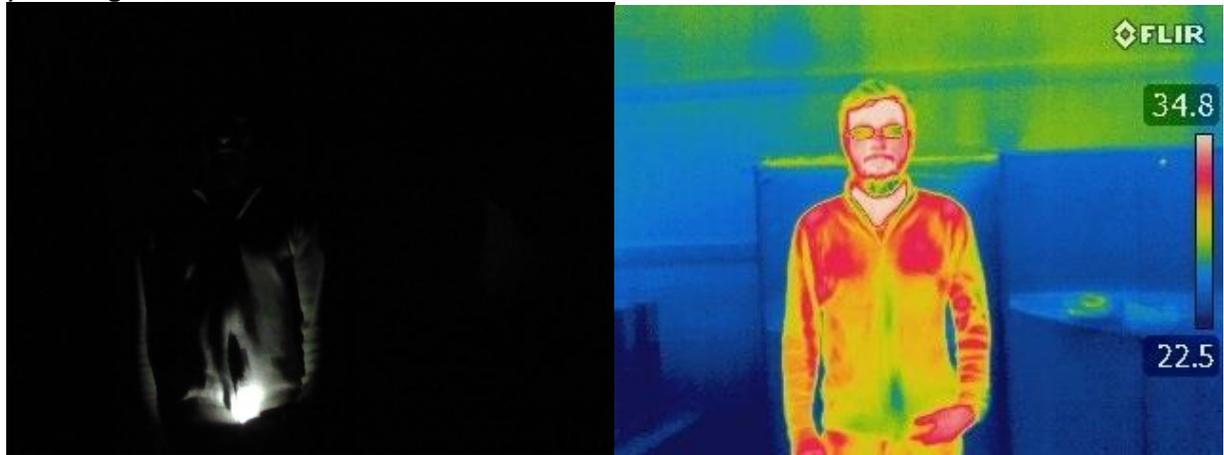


Fonte: Autoria própria.

A não identificação de um pedestre em uma condição de baixa iluminação pode ocorrer por exemplo quando este se aproxima da estrada antes de ser iluminado pelo farol do veículo, em um cenário à noite. Neste exemplo, a visão térmica ajudaria a prover antecipadamente a informação acerca da pessoa se aproximando. A Figura

25 ilustra um cenário onde não há praticamente nenhuma fonte luminosa, e como a imagem térmica é aplicável neste cenário.

**Figura 25 - Pessoa em espectros distintos - quase nenhuma iluminação. a) Imagem original, e b) Termograma.**



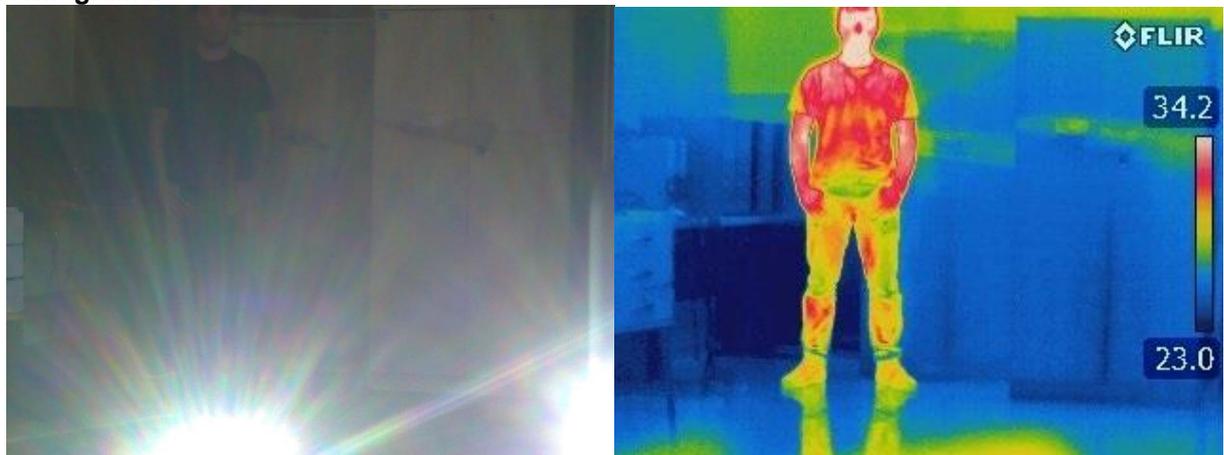
(a)

(b)

Fonte: Autoria própria.

Por sua vez, a Figura 26 exibe lado a lado a mesma imagem de uma pessoa adquirida a partir de espectros diferentes, em uma condição de ofuscamento.

**Figura 26 - Pessoa em espectros distintos - ofuscamento. a) Imagem original, e b) Termograma.**



(a)

(b)

Fonte: Autoria própria.

Observa-se que a identificação de pontos correspondentes é prejudicada no espectro convencional devido à presença do ruído luminoso na cena. Esse tipo de ruído é comum no contexto de direção, pois representa o ofuscamento causado por um veículo que trafega em sentido oposto com a luz alta. Mais uma vez percebe-se que a imagem térmica, e, portanto, a informação estéreo obtida a partir desta,

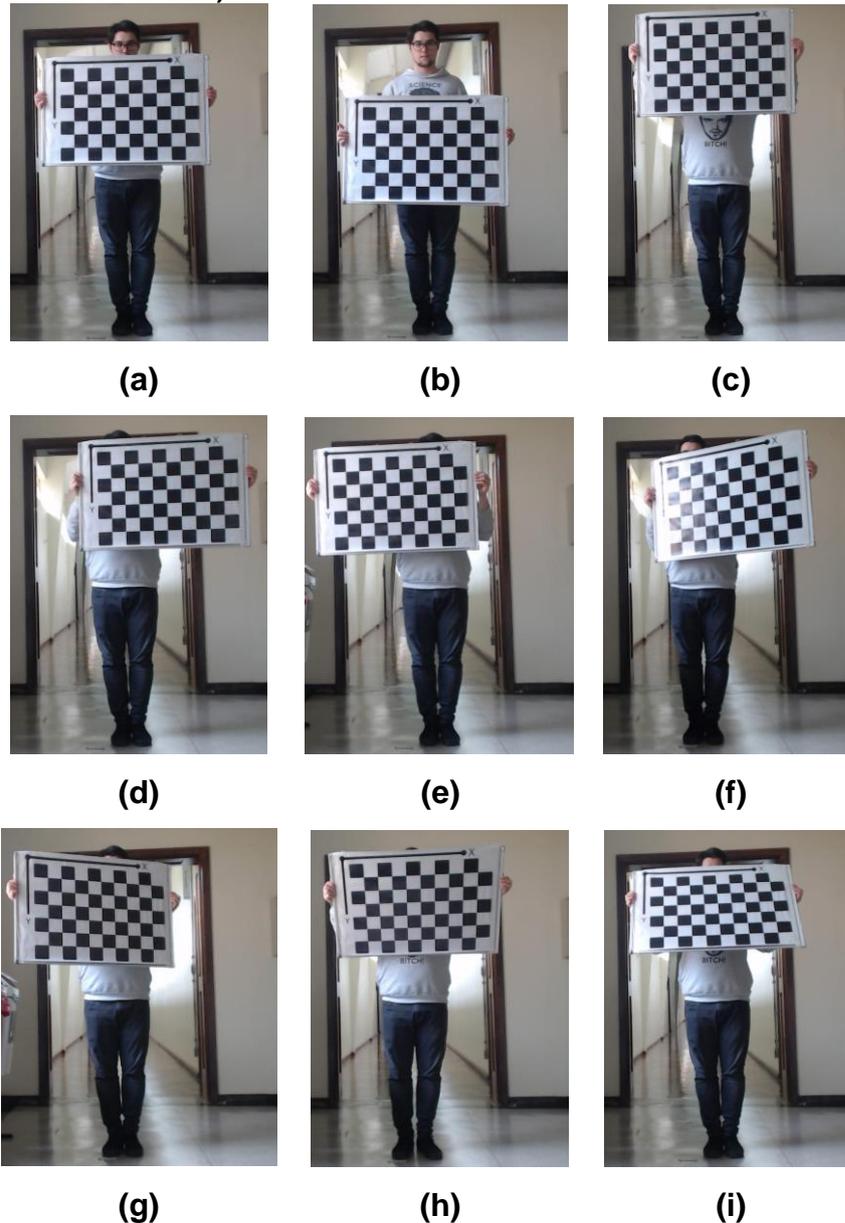
configura um meio importante de obtenção de dados acerca do monitoramento dos arredores do veículo.

## 4.2 ANÁLISE QUANTITATIVA

Os experimentos foram divididos em baterias de captura de imagens térmicas e convencionais, seguidas da geração do *ground-truth*, otimização dos parâmetros e avaliação final a partir das métricas descritas na seção 3.4. Na sequência serão descritos os protocolos para obtenção das imagens de calibração e teste, e resultados inerentes a cada conjunto de dados obtido.

Em cada teste, após montagem dos equipamentos (alvo, bancada) e posicionamento da câmera, eram colocadas marcações no solo para identificar a distância em relação às câmeras. As seguintes distâncias foram utilizadas: 2, 3, 4, 5, 10 e 15m. Nem todos os testes foram realizados em todas as distâncias, devido às limitações dos espaços de captura. Foi definido como máxima distância 15m pois as limitações da estrutura do alvo e resolução das imagens térmicas dificultavam a correta avaliação em maiores distâncias. Após a marcação das distâncias no solo, realizava-se então o processo de captura das imagens do alvo de calibração. Um voluntário se posicionava com o alvo em cada uma das marcações, e era orientado a mover o alvo de modo que se conseguisse captar imagens em posições distintas para aquela profundidade. A orientação era que se posicionasse o mesmo em nove posições, sendo estas acima (alvo cobre os olhos), centro (alvo na altura do tórax), abaixo (alvo abaixo do tórax), esquerda e direita (alvo ao lado do tórax), inclinado direita e inclinado esquerda (alvo na altura do tórax, porém com um dos lados mais à frente que o outro), e alvo inclinado acima e inclinado (centro do tórax, porém a parte superior ou inferior do alvo era posicionada mais à frente). A Figura 27 exhibe exemplos de posições adotadas.

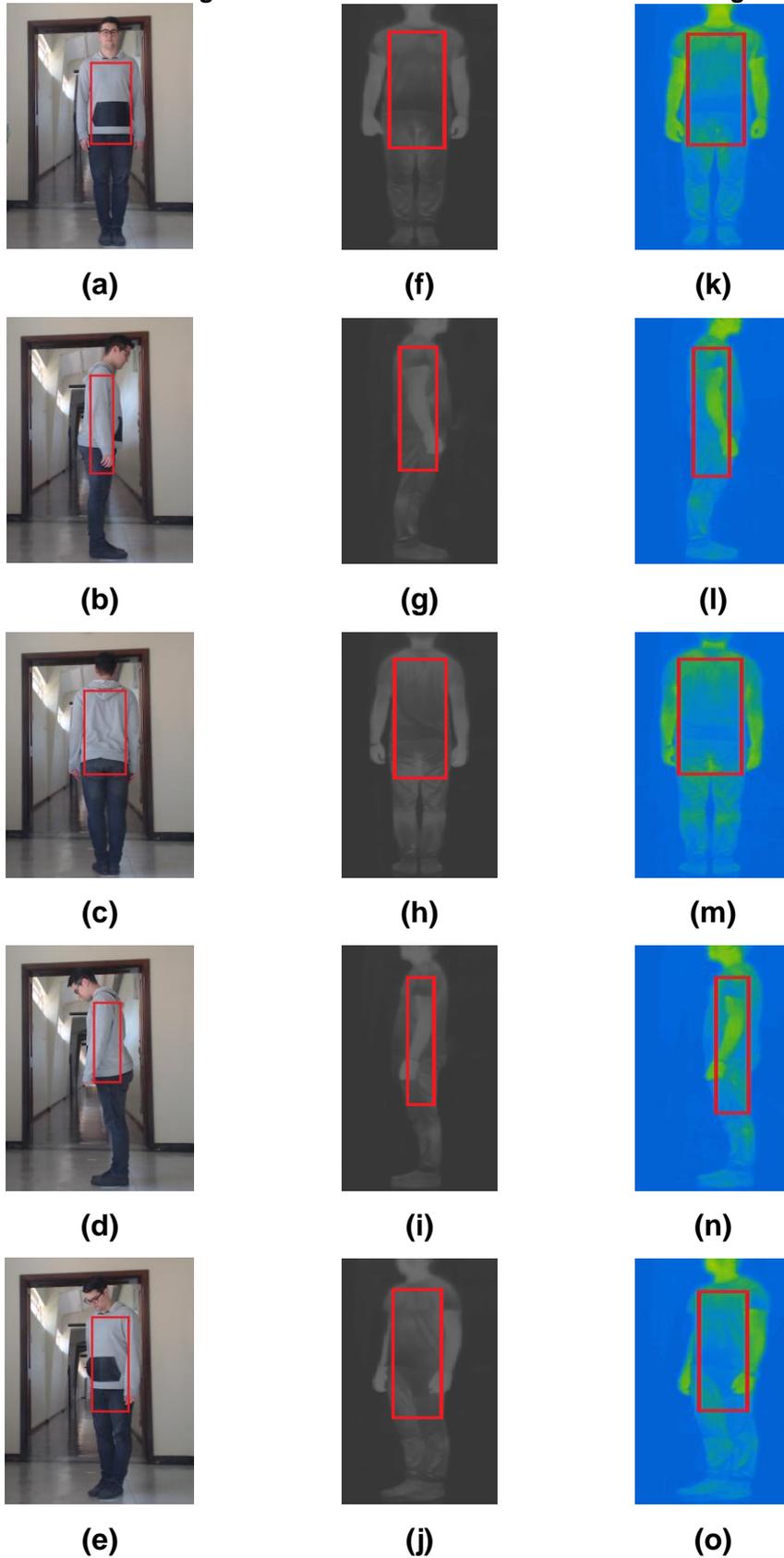
**Figura 27 - Diferentes posições para calibração, sendo estas a) centro, b) abaixo, c) acima, d) direita, e) esquerda, f) inclinado direita, g) inclinado esquerda, h) inclinado cima e i) inclinado baixo.**



**Fonte: Autoria própria.**

Após a calibração, o voluntário se posicionava novamente em todas as marcações, porém sem o alvo de calibração. Em cada marcação, o voluntário alternava entre cinco posições, sendo elas frontal, costas, lado esquerdo, lado direito e inclinado. Esse posicionamento será posteriormente demarcado com o *ground-truth*, e essas imagens serão utilizadas no processo de avaliação, conforme descrito na seção 3.4. A Figura 28 exhibe os posicionamentos adotados, com um retângulo em vermelho exemplificando a região a ser avaliada posteriormente, a qual é definida manualmente através de um software desenvolvido para tal finalidade.

Figura 28 - Diferentes posições para teste, em 3 colunas (convencional, térmico *grayscale* e térmico *rainbow* respectivamente): frontal, direita, costas, esquerda e inclinado. O retângulo foi definido manualmente e delimita a região avaliada.



Fonte: Autoria própria.

Na sequência serão apresentados os resultados obtidos em cada conjunto de dados utilizado, com o intuito de se avaliar as diferenças de acurácia baseada na variação da resolução, espectro (infravermelho ou convencional), e distância real. Para cada profundidade mensurada, serão exibidos:

- *Dataset*: conjunto de dados de um experimento, envolvendo imagens de calibração e teste;
- Total de pixels: quantidade total de pixels que compreende a região interna da caixa de seleção vermelha destacada na Figura 28;
- Total de pixels válidos: quantos pixels, dentro da região de interesse do item anterior, que foram utilizados para se calcular a profundidade (ou seja, que possuem estimativa válida de profundidade);
- RMSD: desvio médio quadrático em relação à medida a ser alcançada. As medidas de erro aqui apresentadas representam o melhor resultado obtido após a aplicação do AG para otimização dos parâmetros, aplicados à cada uma das medições de distância;
- RMSD %: percentual do erro medido em relação à distância real.

Todos os conjuntos térmicos aqui apresentados utilizaram uma escala de temperatura fixa, de 10 a 60°C. Foram realizados experimentos com escala variável, porém a diferença de temperatura apresentada pelas células do alvo térmico é ressaltada nesse tipo de escala, inviabilizando a calibração pelo método utilizado.

A Tabela 3 apresenta os resultados obtidos para distância de 2m, em imagens térmicas.

**Tabela 3 - Análise da profundidade calculada em imagens térmicas para distância de 2m.**

<i>Dataset</i>	Total de pixels	Total de pixels válidos	RMSD (m)	RMSD (%)
2018_06_07__14_00__320x240__thermal	40301	40301	0,272705	13,635250
2018_09_04__19_00__320x240__thermal	34861	34861	0,379768	18,988400
capt_1_320x240__thermal	37995	37995	0,006132	0,306600
capt_2_320x240__thermal_rainbow	37596	29728	0,554436	27,721800
capt_6_320x240__thermal_rainbow_dist	40079	40079	0,457729	22,886450

Fonte: Autoria própria.

O melhor resultado obtido para essa distância ocorreu no *dataset* “capt\_1\_320x240\_\_thermal”, que apresentou cerca de 6mm de erro em relação à

medida desejada. Este conjunto de dados foi adquirido utilizando-se a paleta de escala de cinza, e ajustando-se o foco da câmera a cada captura de distância (assim como os demais conjuntos iniciados por “caps”). Por outro lado, o conjunto “caps\_2\_320x240\_\_thermal\_rainbow” apresentou o pior resultado, obtendo cerca de 28% de erro. É notável que essa alta taxa de erro esta correlacionada com a paleta de cores utilizada nessa captura, chamada de paleta arco-íris. É interessante também destacar que este *dataset* com maior taxa de erro apresenta uma diferença entre o total de pixels e o total válido, indicando dificuldades que o algoritmo encontrou para realizar a correta estimativa de profundidade na região de interesse. A Tabela 4 apresenta os resultados obtidos para estimativa de profundidade de 2m, em imagens convencionais.

**Tabela 4 - Análise da profundidade calculada em imagens convencionais para distância de 2m.**

<i>Dataset</i>	Total de pixels	Total de pixels válidos	RMSD (m)	RMSD (%)
2018_08_21__12_30__480x640	59634	59634	0,065389	3,269450
2018_08_21__12_30__720x960	126459	126404	0,074762	3,738100
2018_09_04__15_00__480x640	29209	29209	0,034796	1,739800
caps_3_640x480	40276	40276	0,042649	2,132450
caps_8_640x480__dist	41939	41939	0,032388	1,619400

**Fonte: Autoria própria.**

O menor erro apresentado foi de aproximadamente 1,62%, e o maior de aprox. 3,74%. Percebe-se que nas imagens convencionais a margem de erro apresentada entre os datasets é pequena em comparação com as variações apresentadas nos datasets térmicos. As prováveis divergências se dão devido a fatores característicos de cada cenário de aquisição de imagens, como a luminosidade, que pode afetar a correta correspondência das características levando a distorções no valor calculado. A Tabela 5 apresenta os resultados obtidos para estimativa de profundidade de 3m, em imagens térmicas.

**Tabela 5 - Análise da profundidade calculada em imagens térmicas para distância de 3m.**

<i>Dataset</i>	Total de pixels	Total de pixels válidos	RMSD (m)	RMSD (%)
2018_06_07__14_00__320x240__therma 	32665	32665	0,627396	20,913200
2018_09_04__19_00__320x240__therma 	23612	23612	0,442881	14,762700
caps_1_320x240__thermal	21349	21349	0,012209	0,406967
caps_2_320x240__thermal_rainbow	23783	23783	0,298234	9,941133
caps_6_320x240__thermal_rainbow_dist	20649	19164	0,081696	2,723200

**Fonte: Aatoria própria.**

O menor erro apresentado foi de aprox. 0,41%, e o maior de aprox. 20,91%, essa diferença se destaca em relação às medições anteriores, para 2m de distância. O alto percentual de erro do *dataset* pode também ser explicado devido à uma convergência muito rápida nos testes realizados com o algoritmo genético. A Tabela 6 apresenta os resultados obtidos para estimativa de profundidade de 3m, em imagens convencionais.

**Tabela 6 - Análise da profundidade calculada em imagens convencionais para distância de 3m.**

<i>Dataset</i>	Total de pixels	Total de pixels válidos	RMSD (m)	RMSD (%)
2018_08_21__12_30__480x640	27573	27573	0,111857	3,728567
2018_08_21__12_30__720x960	58485	58479	0,039130	1,304333
2018_09_04__15_00__480x640	12519	12519	0,035835	1,194500
caps_3_640x480	20850	20850	0,169547	5,651567
caps_8_640x480__dist	22494	22494	0,156690	5,223000

**Fonte: Aatoria própria.**

Também neste caso houve uma pequena variação nos percentuais de erro obtidos, com uma maior incidência de erro nos últimos *dataset* apresentados. A Tabela 7 apresenta os resultados obtidos para estimativa de profundidade de 4m, em imagens térmicas.

**Tabela 7 - Análise da profundidade calculada em imagens térmicas para distância de 4m.**

<i>Dataset</i>	Total de pixels	Total de pixels válidos	RMSD (m)	RMSD (%)
2018_06_07__14_00__320x240__thermal	21079	21079	0,132571	3,314275
2018_09_04__19_00__320x240__thermal	14033	14033	0,694065	17,351625
capt_1_320x240__thermal	15522	15346	0,008983	0,224575
capt_2_320x240__thermal_rainbow	12893	12666	0,034363	0,859075
capt_6_320x240__thermal_rainbow_dist	13862	13862	1,123076	28,076900

**Fonte: Autoria própria.**

O menor erro apresentado foi de aprox. 0,22%, e o maior de aprox. 28,08%. Também nesse caso o conjunto “capt\_1\_320x240\_\_thermal” obteve os melhores resultados, e um dos conjuntos que utiliza a paleta arco-íris obteve o pior resultado. A Tabela 8 apresenta os resultados obtidos para estimativa de profundidade de 4m, em imagens convencionais.

**Tabela 8 - Análise da profundidade calculada em imagens convencionais para distância de 4m.**

<i>Dataset</i>	Total de pixels	Total de pixels válidos	RMSD (m)	RMSD (%)
2018_08_21__12_30__480x640	15769	15769	0,291804	7,295100
2018_08_21__12_30__720x960	32713	32713	0,073407	1,835175
2018_09_04__15_00__480x640	7958	7958	0,054726	1,368150
capt_3_640x480	9594	9594	0,053444	1,336100
capt_8_640x480__dist	12483	12483	0,049346	1,233650

**Fonte: Autoria própria.**

Em geral houve pequena variação de erro, com exceção do primeiro conjunto de dados. A Tabela 9 apresenta os resultados obtidos para estimativa de profundidade de 5m, em imagens térmicas.

**Tabela 9 - Análise da profundidade calculada em imagens térmicas para distância de 5m.**

<i>Dataset</i>	Total de pixels	Total de pixels válidos	RMSD (m)	RMSD (%)
2018_06_07__14_00__320x240__thermal	14104	14104	0,770141	15,402820
2018_09_04__19_00__320x240__thermal	9046	9046	0,910130	18,202600
caps_1_320x240__thermal	19187	6447	0,794325	15,886496
caps_2_320x240__thermal_rainbow	9852	9667	0,772019	15,440380
caps_6_320x240__thermal_rainbow_dist	9093	8981	0,629927	12,598540

**Fonte: Autoria própria.**

Para essa distância houve uma alta taxa de erro em todos os *datasets*, sendo o maior erro apresentado pelo “caps\_1\_320x240\_\_thermal” (15,89%). É interessante ressaltar a quantidade de pixels válidos apresentados, pois estes representam cerca de um terço do total para esse conjunto.

Pode-se deduzir que algum ruído ocorreu neste momento de captura, conduzindo ao baixo aproveitamento de pixels na região de interesse, refletido em um percentual de erro equivalente aos dos demais *datasets*, para a distância de 5m. A Tabela 10 apresenta os resultados obtidos para estimativa de profundidade de 5m, em imagens convencionais.

**Tabela 10 - Análise da profundidade calculada em imagens convencionais para distância de 5m.**

<i>Dataset</i>	Total de pixels	Total de pixels válidos	RMSD (m)	RMSD (%)
2018_08_21__12_30__480x640	10447	10447	0,373607	7,472140
2018_08_21__12_30__720x960	19692	19690	0,063803	1,276060
2018_09_04__15_00__480x640	4692	4692	0,171239	3,424780
caps_3_640x480	7264	7264	0,313654	6,273080
caps_8_640x480__dist	7727	7727	0,419728	8,394560

**Fonte: Autoria própria.**

Para esta medição, a menor taxa de erro apresentada foi de aprox. 1,28%, pelo *dataset* “2018\_08\_21\_\_12\_30\_\_720x960”. À medida que as distâncias avaliadas aumentam, este conjunto que trabalha com imagens com resolução maior apresenta um desempenho superior aos demais. A Tabela 11 apresenta os resultados obtidos para estimativa de profundidade de 10m, em imagens térmicas.

**Tabela 11 - Análise da profundidade calculada em imagens térmicas para distância de 10m.**

<i>Dataset</i>	Total de pixels	Total de pixels válidos	RMSD (m)	RMSD (%)
2018_06_07__14_00__320x240__thermal	-	-	-	-
2018_09_04__19_00__320x240__thermal	9046	9046	1,810552	18,105520
caps_1_320x240__thermal	2453	676	0,706727	7,067270
caps_2_320x240__thermal_rainbow	2601	417	1,560218	15,602180
caps_6_320x240__thermal_rainbow_dist	2184	2184	2,851712	28,517120

Fonte: Autoria própria.

Mais uma vez o conjunto *caps\_1* apresenta o melhor resultado (aprox. 7,07%), condizente com o que vinha apresentando. O pior resultado fica por conta de um dos conjuntos da escala arco-íris (28,51%). A Tabela 12 apresenta os resultados obtidos para estimativa de profundidade de 10m, em imagens convencionais.

**Tabela 12 - Análise da profundidade calculada em imagens convencionais para distância de 10m.**

<i>Dataset</i>	Total de pixels	Total de pixels válidos	RMSD (m)	RMSD (%)
2018_08_21__12_30__480x640	2657	2657	1,006544	10,06544
2018_08_21__12_30__720x960	5676	5676	0,205531	2,055310
2018_09_04__15_00__480x640	1265	1265	0,203645	2,036450
caps_3_640x480	1764	1764	0,487370	4,873700
caps_8_640x480__dist	2257	2256	0,305750	3,057500

Fonte: Autoria própria.

O conjunto “2018\_09\_04\_\_15\_00\_\_480x640” apresenta o melhor resultado (aprox. 2,04%), e o conjunto “2018\_08\_21\_\_12\_30\_\_480x640” o pior, de 10,07%. Um ponto a se destacar é que em geral para a medição de 10m o erro encontrado foi menor do que os resultantes da medição de 5m. A Tabela 13 apresenta os resultados obtidos para estimativa de profundidade de 15m, em imagens térmicas.

**Tabela 13 - Análise da profundidade calculada em imagens térmicas para distância de 15m.**

<i>Dataset</i>	Total de pixels	Total de pixels válidos	RMSD (m)	RMSD (%)
2018_06_07__14_00__320x240__thermal	-	-	-	-
2018_09_04__19_00__320x240__thermal	983	660	2,633346	17,555640
caps_1_320x240__thermal	1461	771	3,315202	22,101347
caps_2_320x240__thermal_rainbow	1117	581	7,243855	48,292367
caps_6_320x240__thermal_rainbow_dist	1007	613	4,816380	32,109200

**Fonte: Autoria própria.**

As medições de 15m apresentaram altas taxas de erro em relação às demais, variando entre 17 a 48%. Percebe-se uma queda acentuada no número de pixels, e uma diferença crescente entre os pixels considerados válidos e o total apresentado. Esse comportamento tende a aumentar os percentuais de erro em todos os *datasets*, contudo esse aumento é mais intenso nas imagens de origem térmica, devido à resolução das mesmas, e também à natureza da emissão térmica, que tende a equalizar regiões com a mesma temperatura, diminuindo possíveis pontos úteis na realização da correspondência. A Tabela 14 apresenta os resultados obtidos para estimativa de profundidade de 15m, em imagens convencionais.

**Tabela 14 - Análise da profundidade calculada em imagens convencionais para distância de 15m.**

<i>Dataset</i>	Total de pixels	Total de pixels válidos	RMSD (m)	RMSD (%)
2018_08_21__12_30__480x640	987	987	0,765677	5,104513
2018_08_21__12_30__720x960	2138	2138	0,406965	2,713100
2018_09_04__15_00__480x640	438	438	0,742898	4,952653
caps_3_640x480	453	363	0,938305	6,255367
caps_8_640x480__dist	703	703	0,837780	5,585200

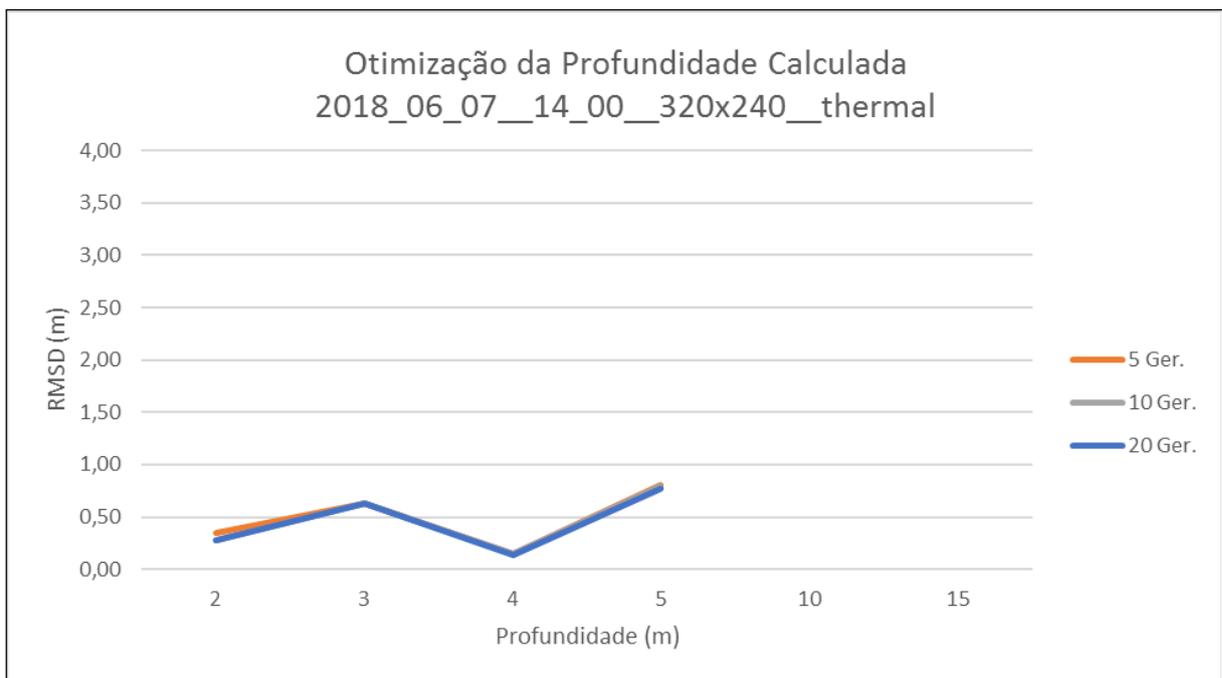
**Fonte: Autoria própria.**

Conforme afirmado anteriormente, mais uma vez o conjunto que apresenta as imagens com maior resolução obteve o percentual de erro menor (2,71%).

Na sequência serão apresentados os resultados referente à utilização do algoritmo genético no contexto de otimização dos parâmetros da correspondência estéreo, e seu impacto na medição da profundidade.

O Gráfico 1 apresenta a análise da otimização dos parâmetros de correspondência, baseada na profundidade calculada, utilizando o *dataset* térmico “2018\_06\_07\_\_14\_00\_\_320x240\_\_thermal”, que apresenta termogramas utilizando uma paleta em escala de cinza, e resolução de 320x240 pixels, e 0,01m de distância *baseline* entre a posição da câmera ao se adquirir as imagens. Pode-se perceber que a variação de gerações do algoritmo genético não teve impacto significativo no RMSD calculado, pois mesmo com 5 gerações já se obteve um resultado próximo ao obtido com 20 gerações. Esse *dataset* foi avaliado até apenas 5m em função da limitação física do espaço de testes. O menor RMSD percentual calculado ocorreu para a distância de 4 metros, atingindo 3,315% de erro, ou seja, 0,132571m, com 20 gerações de otimização. O maior RMSD percentual foi de 20,91%, para a distância de 3m, representando 0,627396m. A distância de 2m obteve o erro de 13,63% (0,272705m), e a distância de 5m obteve o erro de 15,40% (0,770141m).

**Gráfico 1 - RMSD x Profundidade para diferentes gerações - dataset 2018\_06\_07\_\_14\_00\_\_320x240\_\_thermal.**

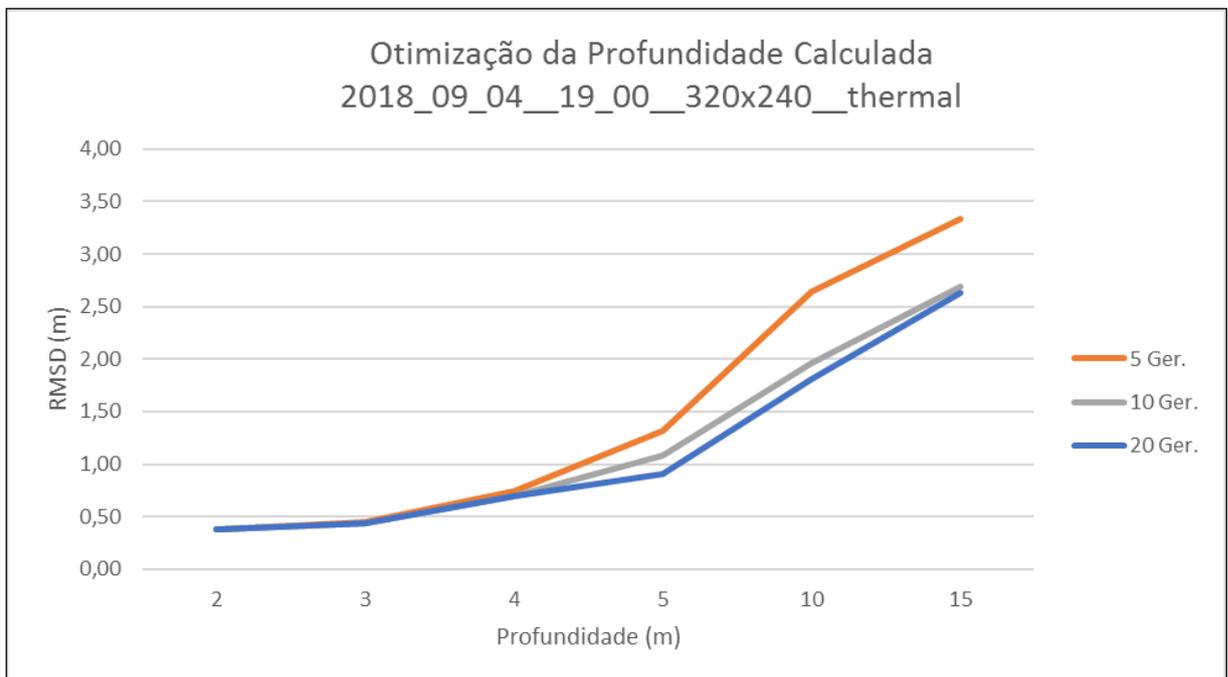


**Fonte: Autoria própria.**

O Gráfico 2 apresenta a análise do *dataset* térmico “2018\_09\_04\_\_19\_00\_\_320x240\_\_thermal”, que também apresenta termogramas utilizando uma paleta em escala de cinza, e resolução de 320x240 pixels, também com distância *baseline* de 0,01m.

O menor RMSD percentual calculado foi de 14,76%, obtendo 0,442881m de erro para a distância de 3m, em 20 gerações. O maior RMSD percentual foi de 18,99%, para a distância de 2m, representando 0,379768m. Para a profundidade de 15m, o RMSD calculado foi de 2,633346m (17,56%). Atribui-se aos altos índices de erro a baixa resolução da imagem, e a própria natureza do termograma em escala de cinza, que tende a apresentar menos pontos de correspondência comparando-se com uma imagem em tons de cinza adquirida a partir de uma câmera convencional. Esses pontos destacados ganham maior relevância ao analisar distâncias maiores do que 5m. Neste *dataset* é visível uma melhora nos resultados no decorrer das gerações do AG.

**Gráfico 2 - RMSD x Profundidade para diferentes gerações - dataset 2018\_09\_04\_19\_00\_320x240\_thermal.**

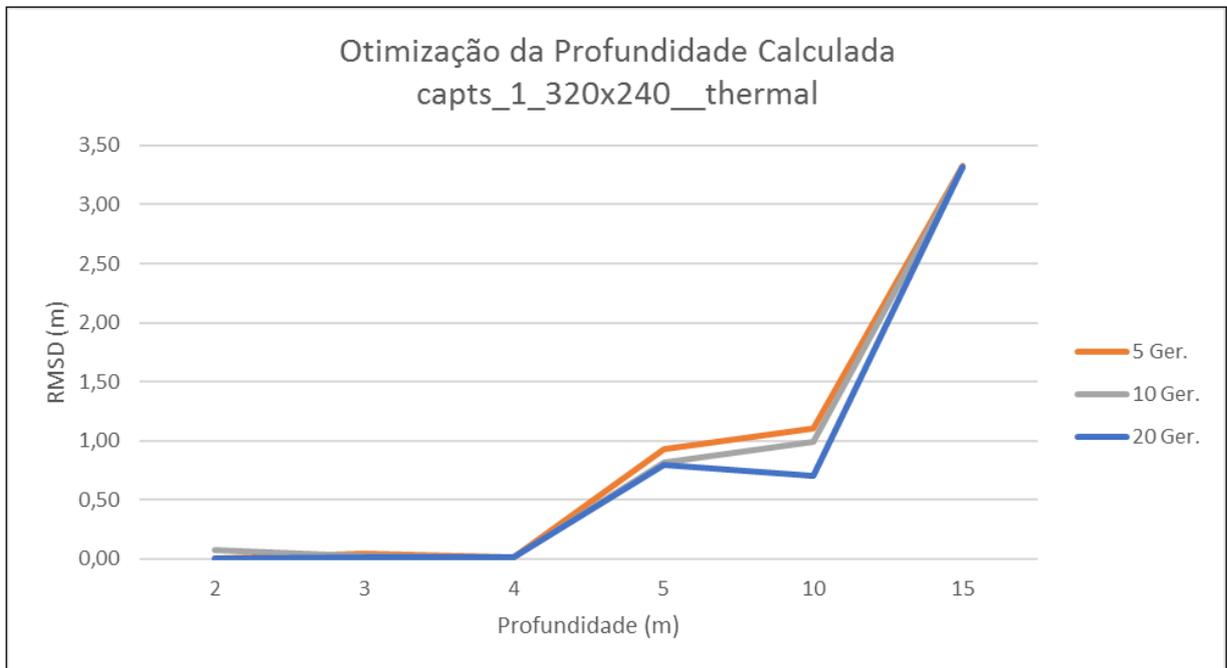


**Fonte: Autoria própria.**

O Gráfico 3 apresenta a análise do *dataset* térmico "caps\_1\_320x240\_thermal", que também apresenta termogramas utilizando uma paleta em escala de cinza, e resolução de 320x240 pixels, e distância *baseline* de 0,01m. Durante sua geração, a cada alteração de posição do voluntário, era feito um ajuste no foco da câmera. O menor RMSD percentual foi de 0,23%, representando 0,008983m de erro para a profundidade de 4m. O maior RMSD percentual foi de 22,10% para 15m (3,315202m de desvio). Neste conjunto de dados, há uma variação brusca

da taxa de erro entre as medições de 10m e 15m, que chegam a 22,10% no último caso. A faixa de erro deste *dataset* variou entre 0,23% a 22,10%.

**Gráfico 3 - RMSD x Profundidade para diferentes gerações - dataset capts\_1\_320x240\_\_thermal.**

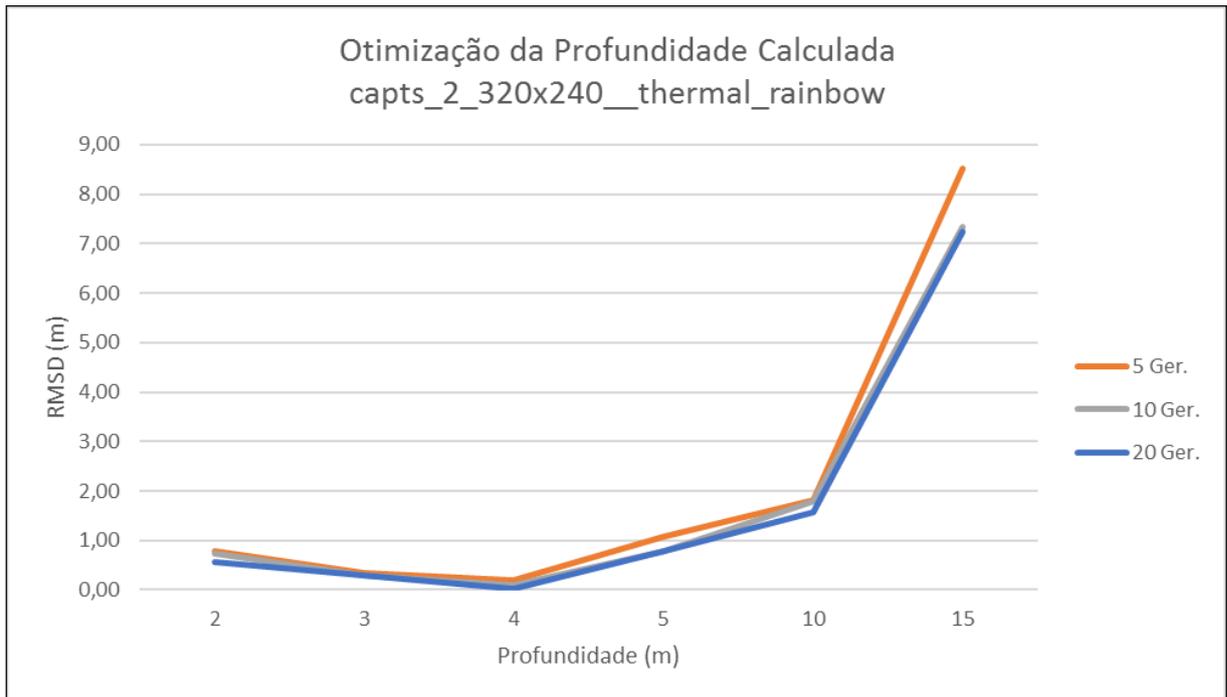


**Fonte: Autoria própria.**

O Gráfico 4 apresenta a análise do *dataset* térmico capts\_2\_320x240\_\_thermal\_rainbow, composto por imagens térmicas que usam uma paleta arco-íris, e resolução de 320x240 pixels, e distância *baseline* de 0,01m. Essa paleta atribui tons de vermelho a regiões na imagem que retratam uma temperatura mais elevada (baseada na escala adotada), e tons azuis para regiões mais frias. Durante sua geração, a cada alteração de posição do voluntário, era feito um ajuste no foco da câmera.

O menor RMSD percentual foi de 0,86%, ou seja, 0,034363m para a distância de 4m. Neste conjunto de dados, há uma variação brusca da taxa de erro entre as medições de 10m e 15m, que chegam a 48,29% no último caso, ou seja 7,243855m de erro para a medição de 15m. A distância de 2m obteve um erro de 0,554436m, ou seja, 27,72%.

**Gráfico 4 - RMSD x Profundidade para diferentes gerações - dataset capts\_2\_320x240\_\_thermal\_rainbow.**

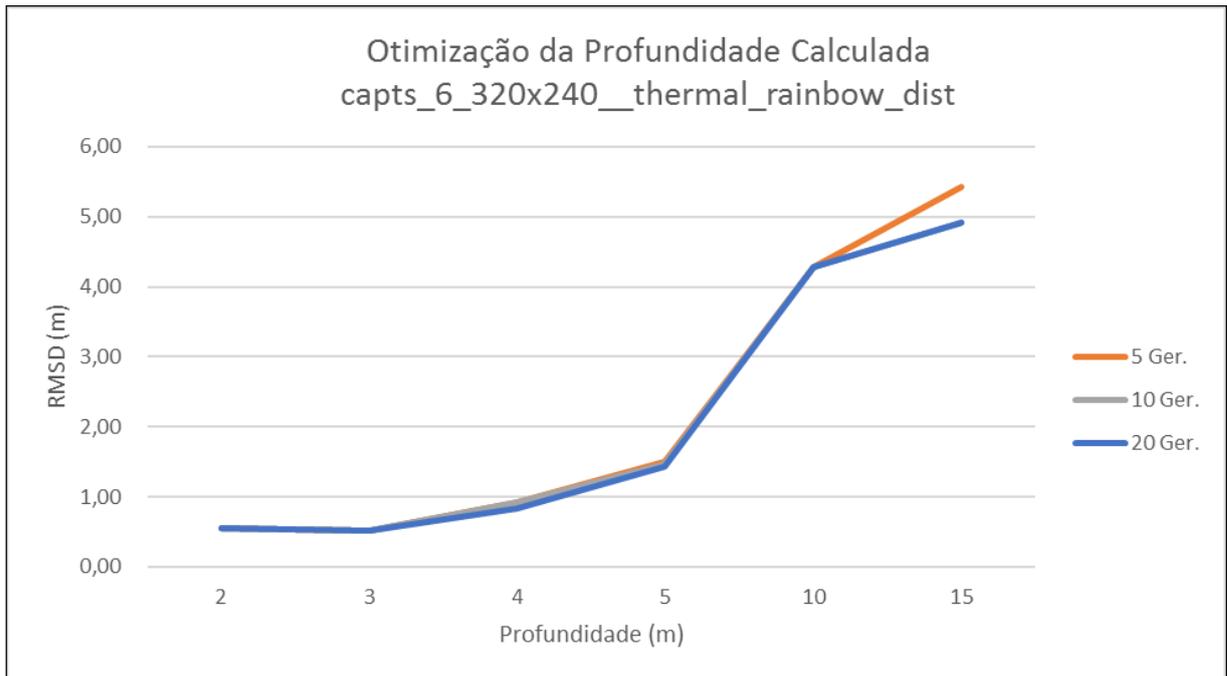


Fonte: Autoria própria.

O Gráfico 5 apresenta a análise do *dataset* térmico *capt\_s\_2\_320x240\_\_thermal\_rainbow\_dist*, composto também por imagens térmicas que usam uma paleta arco-íris, e resolução de 320x240 pixels, obtidas com distância *baseline* de 0,01m. Durante sua geração, a cada alteração de posição do voluntário, era feito um ajuste no foco da câmera. O menor erro percentual encontrado no *dataset* ocorreu na medição de 3m, sendo este 17,06% - 0,511954m, e o maior na de 10m, 42,80% - 4,280444m.

Percebe-se também nesse *dataset* um aumento acentuado dos índices de erro, conforme aumenta-se a distância das medições. Esse comportamento não é tão expressivo nos dados provenientes de paletas em tons de cinza. Baseado nisso, pode-se concluir que imagens que utilizam essa paleta arco-íris necessitam de um pré-processamento específico que se utilize de suas características (tons de matizes diferentes para regiões quentes e frias) para então transformá-las em imagens monocromáticas, a serem utilizadas na geração da correspondência estéreo.

**Gráfico 5 - RMSD x Profundidade para diferentes gerações - dataset capts\_6\_320x240\_\_thermal\_rainbow\_dist.**



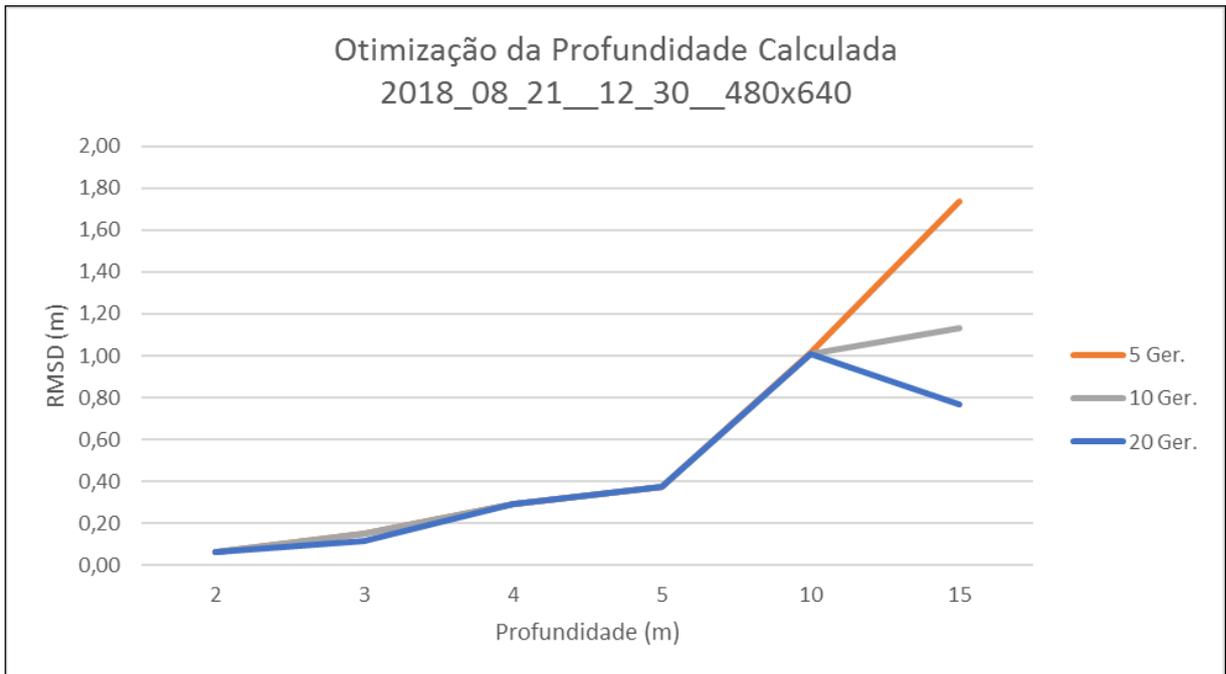
Fonte: Autoria própria.

Na sequência serão apresentados os gráficos referentes à cinco *datasets* convencionais, construídos durante a execução deste trabalho.

O Gráfico 6 apresenta a análise da otimização dos parâmetros aplicada no *dataset* “2018\_08\_21\_\_12\_30\_\_480x640”, que contém imagens convencionais na resolução de 480x640 pixels. Essa resolução invertida do padrão (640x480) foi adotada em função da estrutura do suporte estéreo inicial utilizado. A distância *baseline* para aquisição das imagens foi de 0,07m.

O menor RMSD percentual calculado foi de 3,27%, para a profundidade real de 2 metros, representando 0,065389m, a partir da otimização realizada pelo AG em 20 gerações. A maior taxa de erro foi de 10,07%, ou seja 0,765677m em 10m. A medida de 2m obteve 3,27% de erro (0,065389m), e a de 15m obteve 5,10% (0,765677m). Em geral as medições de erro desse *dataset* variam entre 3 a 10%, apresentando uma faixa de erros inferior à das imagens do espectro térmico, fato que pode ser explicado devido à maior resolução das imagens e maior variabilidade de informações para correta realização da correspondência estéreo, em comparação ao espectro térmico.

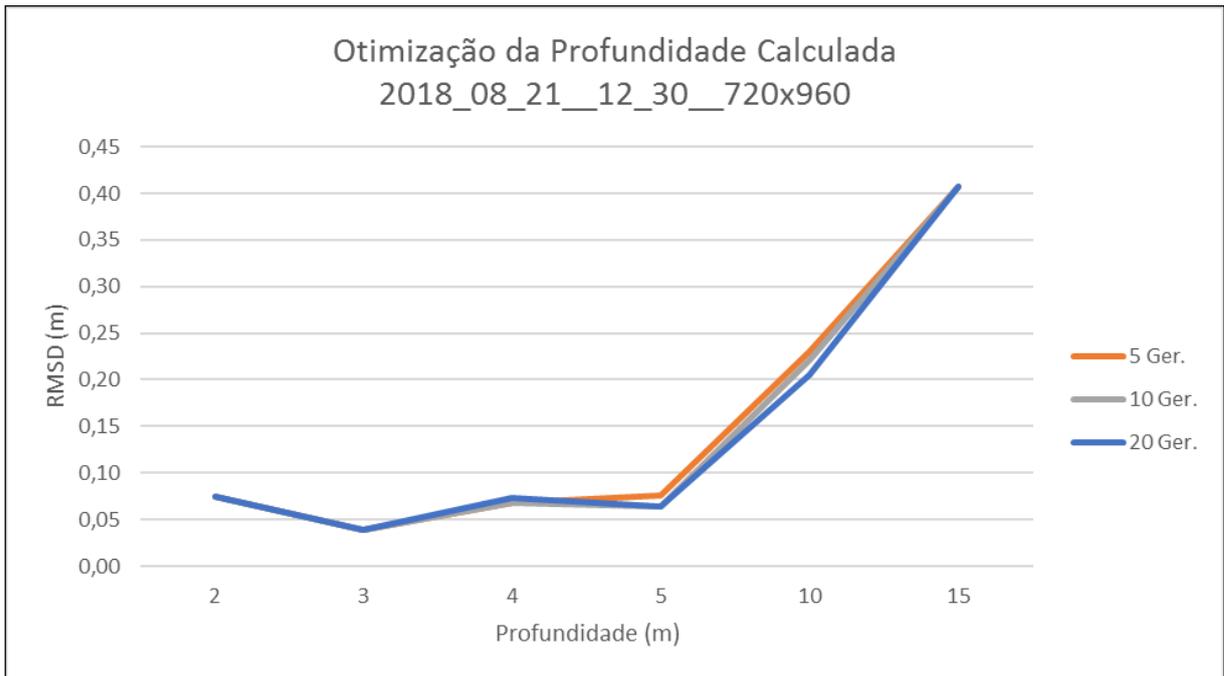
**Gráfico 6 - RMSD x Profundidade para diferentes gerações - dataset 2018\_08\_21\_\_12\_30\_\_480x640.**



Fonte: Autoria própria.

O Gráfico 7 apresenta a análise aplicada no *dataset* “2018\_08\_21\_\_12\_30\_\_720x960”, que contém imagens convencionais na resolução de 720x960 pixels, e 0,07m de distância *baseline*. Essa resolução em modo retrato foi adotada em função da estrutura do suporte estéreo inicial utilizado. O menor RMSD percentual calculado no *dataset* foi de 1,28% para a distância de 5m, representando 0,063803m de erro. O maior percentual ocorreu nos 2m, atingindo 3,74%, o que representou 0,075m de erro. Para profundidade de 15m, o RMSD obtido foi de 0,406965m, ou 2,71% de erro. A faixa de erro desse *dataset* variou entre 1,28 e 3,74%, um resultado expressivo se comparado aos obtidos pelos conjuntos de dados do espectro térmico. A partir dessa comparação destaca-se a importância da resolução das imagens para que se mantenham bons índices de erro à medida que se calcula a profundidade para distâncias maiores.

**Gráfico 7 - RMSD x Profundidade para diferentes gerações - dataset 2018\_08\_21\_\_12\_30\_\_720x960.**

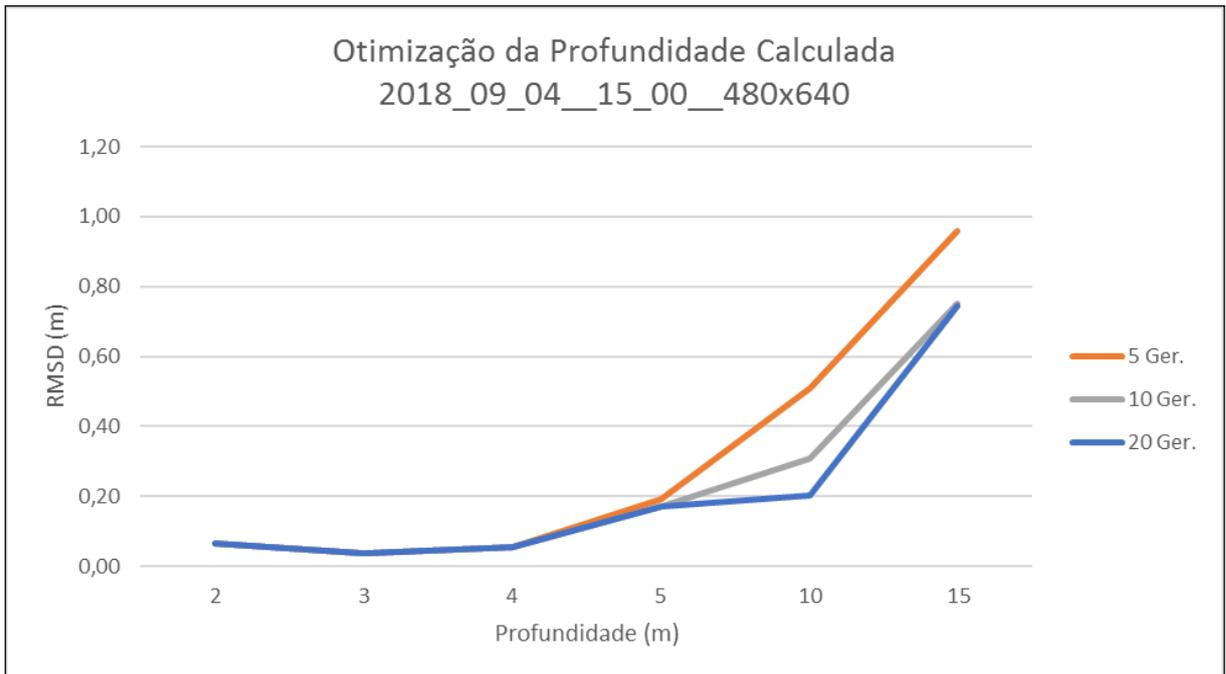


Fonte: Autoria própria.

O Gráfico 8 apresenta a análise aplicada no *dataset* “2018\_09\_04\_\_15\_00\_\_480x640”, que contém imagens convencionais na resolução de 480x640 pixels, e 0,07m de distância *baseline*. O menor RMSD percentual calculado no *dataset* foi de 1,19% para a distância de 3m, representando 0,035835m de erro. O maior percentual ocorreu nos 15m, atingindo 4,95%, o que representou 0,742898m de erro. Para profundidade de 2m, o erro obtido foi de 0,065980m, ou 3,30% de erro.

Mantém-se também neste *dataset* a reduzida variação de erro entre as distâncias avaliadas. Também se destaca uma variabilidade de resultados, ao comparar-se as gerações do algoritmo genético.

**Gráfico 8 - RMSD x Profundidade para diferentes gerações - dataset 2018\_09\_04\_15\_00\_480x640.**

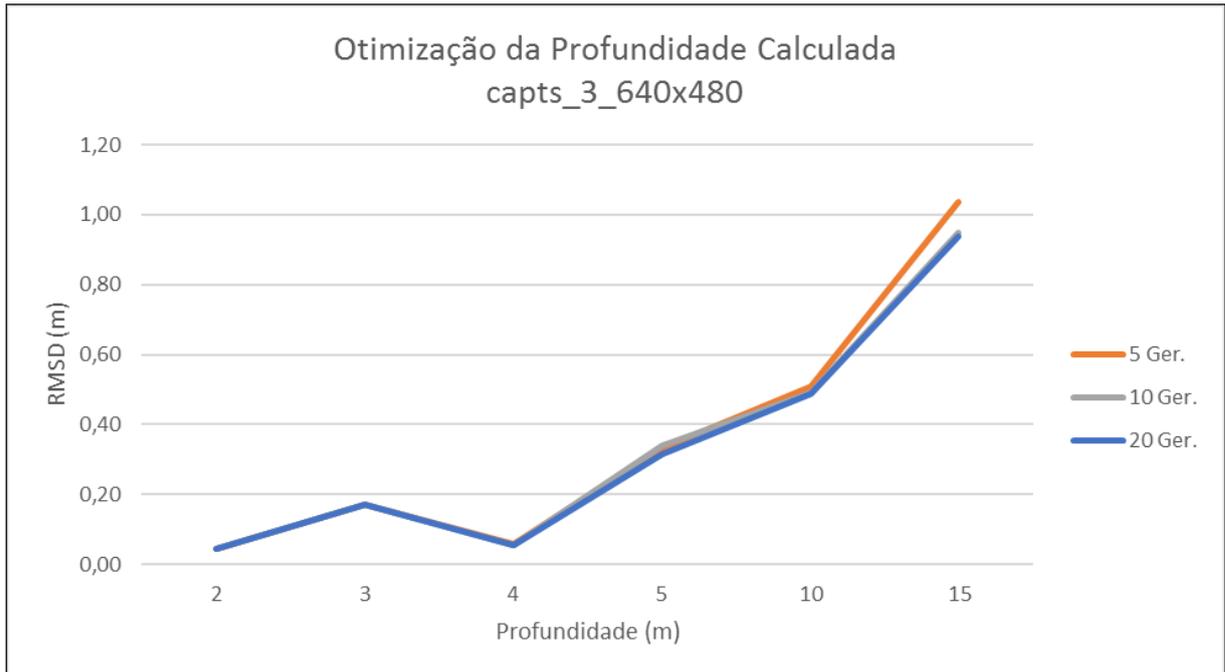


**Fonte: Autoria própria.**

O Gráfico 9 apresenta a análise aplicada no *dataset* “*capt3\_3\_640x480*”, que contém imagens convencionais na resolução de 640x480 pixels, e 0,10m de distância *baseline*.

O menor RMSD percentual calculado no *dataset* foi de 1,34% para a distância de 4m, representando 0,053444m de erro. O maior percentual ocorreu nos 5m, atingindo 6,27%, o que representou 0,313654m de erro. Para profundidade de 2m, o erro obtido foi de 0,042649m, ou 2,13%, e para 15m 0,938305m, ou 6,26% de erro.

**Gráfico 9 - RMSD x Profundidade para diferentes gerações - dataset capts\_3\_640x480.**

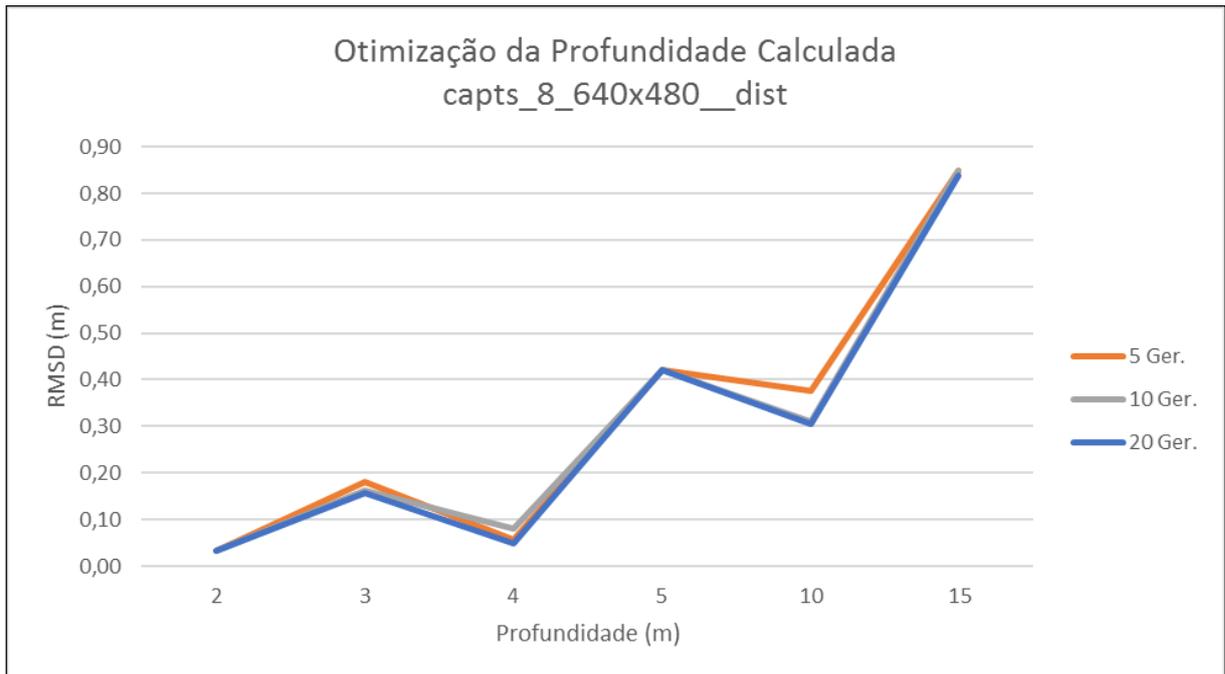


**Fonte: Autoria própria.**

O Gráfico 10 apresenta a análise aplicada no *dataset* “capt\_8\_640x480\_\_dist”, que contém imagens convencionais na resolução de 640x480 pixels, e 0,10m de distância *baseline*.

O menor RMSD percentual calculado no *dataset* foi de 1,23% para a distância de 4m, representando 0,049346m de erro. O maior percentual ocorreu nos 5m, atingindo 8,39%, o que representou 0,419728m de erro. Para profundidade de 2m, o erro obtido foi de 0,032388m, ou 1,62%, e para 15m 0,83778m, ou 5,59% de erro.

**Gráfico 10 - RMSD x Profundidade para diferentes gerações - dataset capts\_8\_640x480\_\_dist.**



**Fonte: Autoria própria.**

Esta seção demonstrou de forma objetiva as avaliações realizadas no método de mapeamento tridimensional, aplicado tanto no espectro térmico quanto no espectro convencional. Os resultados foram apresentados a partir de dois formatos, um deles levando em consideração a quantidade de pixels válidos nas amostras, detalhado a partir das tabelas apresentadas nesta seção, e o outro indicando a utilização do algoritmo genético, e o comportamento das funções de erro em cada *dataset* avaliado, observado a partir de uma análise gráfica. Em média, constatou-se um tempo de processamento de 30ms para a geração da correspondência estéreo, em imagens 640x480 pixels, utilizando um ambiente Linux (Ubuntu 16.04 LTS 64-bits), em um notebook com processador Intel®Core™ i7-3537U CPU @ 2.00GHz × 4, memória de 8 GB, compilador utilizado foi o G++ (GNU C++ Compiler) 5.4.0 em sua configuração padrão. A resolução das imagens é proporcional ao tempo de processamento gasto nessa etapa, o qual também é afetado pela estrutura de hardware utilizada.

A partir da análise dos resultados obtidos, pode-se concluir que a geração de um mapa de profundidade térmico nos termos aqui descritos é satisfatória para distâncias até 10m, pois para essa distância obteve-se uma média de percentual de erro de 12,59%, considerando os dois *datasets* térmicos que apresentavam essa medição de 10m, com paleta em escala de cinza.

Wang, Wu e Lin (2005) em seu trabalho realizam uma análise da identificação da distância percorrida por veículos de passageiros que utilizam freio ABS. Em seus resultados, são apresentadas distâncias percorridas após a frenagem, para velocidades de 30, 50 e 70km/h, em 2 cenários distintos de pista seca. Aplicando uma média nos resultados obtidos, tem-se que para a velocidade de 30km/h o deslocamento do veículo é de 4,59m, para 50km/h o deslocamento é de 11,59m, e para 70km/h, 21,65m. Utilizando uma interpolação quadrática nos resultados, obtém-se que para 40km/h o deslocamento seria de 7,125m, e para 60km/h, 15,625m. A Tabela 15 exibe uma comparação entre as medições realizadas por Wang, Wu e Lin, e os resultados obtidos neste trabalho, onde exibe-se os melhores resultados para as distâncias avaliadas que mais se aproximam aos deslocamentos de frenagem, supondo uma situação onde o pedestre seria detectado nestas distâncias (5, 10 e 15m), e o intervalo de confiança das respectivas medições baseado no RMSD obtido em metros.

**Tabela 15 - Análise das distâncias de frenagens e resultados deste trabalho.**

Velocidade (km/h)	Deslocamento (m)	Distância avaliada (m)	RMSD (m)	Intervalo (m)	RMSD (%)
30	4,59	5	0,852	4,148 - 5,852	17,04
40	7,125 <sup>6</sup>	10	1,259	8,741 - 11,259	12,59
50	11,59	10	1,259	8,741 - 11,259	12,59
60	15,625 <sup>6</sup>	15	2,974	12,0255 - 17,9745	19,83

Fonte: Autoria própria.

Baseado nesses dados, a implementação de um modelo tridimensional térmico em sistemas de auxílio ao condutor poderia prevenir colisões em vias coletoras (40km/h), pois caso o pedestre estivesse em trajetória de colisão, a 10m do veículo, mesmo considerando o intervalo de erro da medição do sistema (o qual indicaria a posição do indivíduo entre 8,741 a 11,259m), a estimativa seria suficiente para acionar a tempo o sistema de auxílio ao condutor, executando a frenagem e impedindo a colisão (após deslocar 7,125m). Em vias locais (30km/h) e arteriais (60km/h), supondo um cenário equivalente, porém com as respectivas distâncias, poderia ocorrer ou a prevenção da colisão, ou a minimização do impacto, contribuindo assim para a preservação da vida e integridade do pedestre.

<sup>6</sup> Deslocamento estimado através de interpolação quadrática.

## 5 CONCLUSÃO

Este trabalho investiga um método capaz de gerar um mapa de profundidade térmico de uma cena, passível de utilização por sistemas de auxílio ao condutor. Para geração deste modelo, é necessário o estabelecimento da geometria epipolar e correspondências entre termogramas captados de pontos de vista diferentes. Tal tarefa implica na calibração das câmeras, de forma a obter os parâmetros necessários para estabelecimento da geometria epipolar. Feito isto, e resolvido o problema de correspondência, é possível estabelecer um mapa de disparidade, e consequentemente, um mapa de profundidade.

Além da obtenção da profundidade da cena, este trabalho fornece uma estrutura projetada para a otimização dos parâmetros relacionados à correspondência estéreo, baseada na utilização de um algoritmo genético, com o intuito de preparar o melhor conjunto de parâmetros que se enquadrem a um pré-determinado escopo de atuação. Isso significa que dependendo das condições das imagens adquiridas, a variação dos parâmetros pode favorecer a correspondência para cenários específicos. A partir dos experimentos verificou-se que a parametrização do AG acarretou uma rápida convergência, pois os resultados obtidos para diferentes gerações ficaram muito próximos.

A partir da análise dos resultados obtidos, pode-se concluir que a geração de um mapa de profundidade térmico nos termos aqui descritos é satisfatória para distâncias até 10m, pois para essa distância obteve-se uma média de percentual de erro de 12,59%, considerando os dois *datasets* térmicos que apresentavam essa medição de 10m, com paleta em escala de cinza. Baseado nisso constatou-se que a implementação de um modelo tridimensional térmico em sistemas de auxílio ao condutor poderia prevenir colisões com pedestres em vias coletoras (40km/h), e minimizar o impacto em vias locais (30km/h) e arteriais (60km/h).

A partir dos testes realizados constatou-se qualitativamente a eficiência na utilização de imagens de origem térmica no contexto de identificação de pedestres, visto que em situações reais as câmeras sofrem diariamente influência de condições de iluminação.

Como pontos de melhoria indica-se a utilização de um alvo térmico com mais pontos de calibração, pois a baixa resolução do alvo utilizado faz com que sejam necessárias mais imagens para obtenção de resultados satisfatórios. Além disso,

havendo a possibilidade de montagem do par estéreo com duas câmeras, poderia ser eliminado o suporte de deslocamento, otimizando assim os testes realizados e reduzindo a inserção e erro no processo de correspondência.

Identificou-se como a maior vantagem a característica do sensor térmico de não depender de iluminação para atuar corretamente, conforme demonstrado na análise qualitativa. Além dos pontos de melhoria já apontados, a fusão de sensores entre a imagem térmica e convencional tende a aproveitar essa característica térmica e utilizar as vantagens do espectro convencional (maior quantidade de pontos de correspondência), configurando uma solução interessante para ser aplicada no contexto de sistemas de auxílio ao condutor.

## 5.1 TRABALHOS FUTUROS

Sugere-se como trabalhos futuros a aplicação de outros formatos de parametrização do AG, e avaliação de seu impacto na estimativa de profundidade. Além disso, sugere-se o estudo da identificação do pedestre ou objeto de interesse na cena térmica, visto que este trabalho tem como escopo a análise da profundidade no espectro infravermelho. Essa identificação consiste em, a partir da imagem térmica, conseguir-se determinar se existe ou não um pedestre na cena, e sua localização (em coordenadas de pixels), o que substituiria por exemplo o processo de geração do GT. Sugere-se também a aplicação de outras meta-heurísticas, ou redes neurais profundas, a fim de se obter um aprimoramento na otimização dos parâmetros, em comparação com o AG. Também se propõe a utilização de *datasets* públicos na avaliação do método, além da abrangência de mais parâmetros ligados ao fluxo estéreo, como os relacionados à etapa de calibração, e também parametrização de outros algoritmos de correspondência, além do SGBM.

## REFERÊNCIAS

- ABRAHAM, M.; JASIOBEDZKI, P.; UMASUTHAN, M. **Robust 3D Vision for Autonomous Space Robotic Operations**. . In: PROCEEDINGS OF THE 6TH INTERNATIONAL SYMPOSIUM ON ARTIFICIAL INTELLIGENCE AND ROBOTICS \& AUTOMATION IN SPACE: I-SAIRAS. Montreal, Canada: 2001
- ALESSANDRETTI, G.; BROGGI, A.; CERRI, P. Vehicle and Guard Rail Detection Using Radar and Vision Data Fusion. **IEEE Transactions on Intelligent Transportation Systems**, v. 8, n. 1, p. 95-105, mar. 2007.
- ANDRADE, D. C. et al. A Novel Strategy for Road Lane Detection and Tracking Based on a Vehicle's Forward Monocular Camera. **IEEE Transactions on Intelligent Transportation Systems**, p. 1-11, 2018.
- ANZENGRUBER, B.; RIENER, A. **"FaceLight": Potentials and Drawbacks of Thermal Imaging to Infer Driver Stress**. Proceedings of the 4th International Conference on Automotive User Interfaces and Interactive Vehicular Applications. **Anais...**: AutomotiveUI '12. New York, NY, USA: ACM, 2012
- AUTOMOTIVE NEWS. **Delphi, Quanergy team up on low-cost lidar**. Disponível em: <<https://www.autonews.com/article/20151026/OEM06/310269985/delphi-quanergy-team-up-on-low-cost-lidar>>. Acesso em: 2 jan. 2019.
- BERTOZZI, M.; BROGGI, A. GOLD: a parallel real-time stereo vision system for generic obstacle and lane detection. **IEEE Transactions on Image Processing**, v. 7, n. 1, p. 62-81, jan. 1998.
- BISHOP, R. Lane Keeping Assist Systems. In: **Intelligent vehicle technology and trends**. Boston: Artech House, 2005.
- BRADSKI, G. The OpenCV Library. **Dr. Dobb's Journal of Software Tools**, 2000.
- BROGGI, A. et al. A New Approach to Urban Pedestrian Detection for Automatic Braking. **IEEE Transactions on Intelligent Transportation Systems**, v. 10, n. 4, p. 594-605, dez. 2009.
- CHAI, J.; MA, S. **An evolutionary framework for stereo correspondence**. Proceedings. Fourteenth International Conference on Pattern Recognition (Cat. No.98EX170). **Anais...** In: PROCEEDINGS. FOURTEENTH INTERNATIONAL CONFERENCE ON PATTERN RECOGNITION (CAT. NO.98EX170). ago. 1998
- CHENG, S. Y.; PARK, S.; TRIVEDI, M. M. **Multiperspective Thermal IR and Video Arrays for 3D Body Tracking and Driver Activity Analysis**. 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05) - Workshops. **Anais...** In: 2005 IEEE COMPUTER SOCIETY CONFERENCE ON COMPUTER VISION AND PATTERN RECOGNITION (CVPR'05) - WORKSHOPS. jun. 2005
- D'AMICO, A. et al. Volatile Compounds Detection by IR Acousto-Optic Detectors. In: BYRNES, J. (Ed.). . **Unexploded ordnance detection and mitigation**. NATO

science for peace and security series Series B : Physics and biophysics. Dordrecht: Springer, 2009.

DENNING, J. **Stereo Matching**. Disponível em: <[https://cse.taylor.edu/~jdenning/classes/cos351/slides/14\\_stereomatching.html](https://cse.taylor.edu/~jdenning/classes/cos351/slides/14_stereomatching.html)>. Acesso em: 2 jan. 2019.

DIAKIDES, N. A. et al. Advances in Medical Infrared Imaging. In: DIAKIDES, N. A.; BRONZINO, J. D. (Eds.). . **Medical Infrared imaging**. Boca Raton: CRC Press/Taylor & Francis, 2008.

DICKENS, J. S.; WYK, M. A. VAN; GREEN, J. J. **Pedestrian detection for underground mine vehicles using thermal images**. AFRICON, 2011. **Anais...** In: AFRICON, 2011. set. 2011

EOL - EARTH OBSERVING LABORATORY. **How Do Radars Work?** Disponível em: <<https://www.eol.ucar.edu/content/how-do-radars-work>>. Acesso em: 2 jan. 2019.

EURO NCAP. **The Ratings Explained**. Disponível em: <<https://www.euroncap.com:443/en/vehicle-safety/the-ratings-explained/>>. Acesso em: 19 out. 2018.

FLIR. **FLIR T440 Datasheet**. Disponível em: <[http://www.flirmedia.com/MMC/THG/Brochures/IND\\_003/IND\\_003\\_EN.pdf](http://www.flirmedia.com/MMC/THG/Brochures/IND_003/IND_003_EN.pdf)>. Acesso em: 12 nov. 2018.

FORSYTH, D. A.; PONCE, J. Geometric Camera Models. In: **Computer Vision A Modern Approach**. 2. ed. Boston: Pearson Prentice Hall, 2012.

GERONIMO, D. et al. Survey of Pedestrian Detection for Advanced Driver Assistance Systems. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, v. 32, n. 7, p. 1239-1258, jul. 2010.

GORIS, R. C. Infrared Organs of Snakes: An Integral Part of Vision. **Journal of Herpetology**, v. 45, n. 1, p. 2-14, 1 mar. 2011.

GOUBET, E.; KATZ, J.; PORIKLI, F. **Pedestrian tracking using thermal infrared imaging**. (B. F. Andresen, G. F. Fulop, P. R. Norton, Eds.)2006

HALLIDAY, D.; RESNICK, R.; WALKER, J. Temperatura, Calor e a Primeira Lei da Termodinâmica. In: **Fundamentos de física. volume 2: gravitação, ondas e termodinâmica**. 9. ed. Rio de Janeiro: LTC, 2012. v. 2.

HARTLEY, R.; ZISSERMAN, A. Camera Models. In: **Multiple view geometry in computer vision**. Second ed. Cambridge, UK: Cambridge University Press, 2004a.

HARTLEY, R.; ZISSERMAN, A. Estimation - 2D Projective Transformations. In: **Multiple view geometry in computer vision**. Second ed. Cambridge, UK: Cambridge University Press, 2004b.

HOLLAND, J. H. **Adaptation in natural and artificial systems: an introductory analysis with applications to biology, control, and artificial intelligence**. 1st MIT Press ed ed. Cambridge, Mass: MIT Press, 1992.

HUH, K. et al. A stereo vision-based obstacle detection system in vehicles. **Optics and Lasers in Engineering**, v. 46, n. 2, p. 168-178, fev. 2008.

ISSA, H.; RUICHEK, Y.; POSTAIRE, J.-. **Stereo correspondence using a genetic scheme with a new solution encoding**. IEEE International Conference on Systems, Man and Cybernetics. **Anais...** In: IEEE INTERNATIONAL CONFERENCE ON SYSTEMS, MAN AND CYBERNETICS. out. 2002

IVY TOOLS. **FLIR T440 Infrared Thermal Imaging Camera**. Disponível em: <<http://www.IvyTools.com/product-p/flir-t440.htm>>. Acesso em: 9 nov. 2018.

JAZAYERI, A. et al. Vehicle Detection and Tracking in Car Video Based on Motion Model. **IEEE Transactions on Intelligent Transportation Systems**, v. 12, n. 2, p. 583-595, jun. 2011.

KIA. **Driving Assist Technologies**. Disponível em: <<http://pr.kia.com/en/wow/drive-wise/drive-wise-technologies/driving-assist.do>>. Acesso em: 2 jan. 2019.

KIDONO, K. et al. **Pedestrian recognition using high-definition LIDAR**. 2011 IEEE Intelligent Vehicles Symposium (IV). **Anais...** In: 2011 IEEE INTELLIGENT VEHICLES SYMPOSIUM (IV). jun. 2011

KIM, D. E.; KWON, D. S. **Pedestrian detection and tracking in thermal images using shape features**. 2015 12th International Conference on Ubiquitous Robots and Ambient Intelligence (URAI). **Anais...** In: 2015 12TH INTERNATIONAL CONFERENCE ON UBIQUITOUS ROBOTS AND AMBIENT INTELLIGENCE (URAI). out. 2015

KROTOSKY, S. J.; TRIVEDI, M. M. On Color-, Infrared-, and Multimodal-Stereo Approaches to Pedestrian Detection. **IEEE Transactions on Intelligent Transportation Systems**, v. 8, n. 4, p. 619-629, dez. 2007.

KROTOSKY, S.; TRIVEDI, M. **Multimodal Stereo Image Registration for Pedestrian Detection**. 2006 IEEE Intelligent Transportation Systems Conference. **Anais...** In: 2006 IEEE INTELLIGENT TRANSPORTATION SYSTEMS CONFERENCE. 2006

LIENHARD IV, J. H.; LIENHARD V, J. H. The General Problem of Heat Exchange. In: **A heat transfer textbook**. 3rd ed ed. Cambridge, Mass: Phlogiston Press, 2008.

LINDEN, R. O GA Mais Básico. In: **Algoritmos genéticos**. 2a. ed. [s.l.] Brasport, 2008.

MATHWORKS. **What is Camera Calibration?** Disponível em: <<https://www.mathworks.com/help/vision/ug/camera-calibration.html>>. Acesso em: 23 ago. 2017.

MEOLA, C. Origin and Theory of Infrared Thermography. In: MEOLA, C. (Ed.). . **Infrared Thermography Recent Advances and Future Trends**. [s.l.] Bentham Science Publishers, 2012.

MINKINA, W.; DUDZIK, S. Measurements in Infrared Thermography. In: **Infrared Thermography: Errors and Uncertainties**. Chichester, UK: John Wiley & Sons, Ltd, 2009.

NAYLOR, N. **NHTSA Announces Final Rule Requiring Rear Visibility Technology**. Text. Disponível em: <<https://www.transportation.gov/briefing-room/nhtsa-announces-final-rule-requiring-rear-visibility-technology>>. Acesso em: 18 out. 2018.

NHTSA. **TRAFFIC SAFETY FACTS 2015**. [s.l.] NHTSA - National Highway Traffic Safety Administration, 2015a. Disponível em: <<https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/812384>>.

NHTSA. **Critical Reasons for Crashes Investigated in the National Motor Vehicle Crash Causation Survey**. Washington: NHTSA - National Highway Traffic Safety Administration, fev. 2015b. Disponível em: <<https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/812115>>.

ONSV. **OBSERVATÓRIO | 90% dos acidentes são causados por falhas humanas, alerta OBSERVATÓRIO**. Disponível em: <<https://www.onsv.org.br/90-dos-acidentes-sao-causados-por-falhas-humanas-alerta-observatorio/>>. Acesso em: 1 maio. 2017.

PRAKASH, S.; LEE, P. Y.; CAELLI, T. **3D Mapping of Surface Temperature Using Thermal Stereo**. Robotics and Vision 2006 9th International Conference on Control, Automation. **Anais...** In: ROBOTICS AND VISION 2006 9TH INTERNATIONAL CONFERENCE ON CONTROL, AUTOMATION. dez. 2006

RAO, R. **Stereo and 3D Vision**. Disponível em: <<https://courses.cs.washington.edu/courses/cse455/09wi/Lects/lect16.pdf>>. Acesso em: 29 out. 2018.

RUICHEK, Y. et al. **Towards real-time obstacle detection using a hierarchical decomposition methodology for stereo matching with a genetic algorithm**. 16th IEEE International Conference on Tools with Artificial Intelligence. **Anais...** In: 16TH IEEE INTERNATIONAL CONFERENCE ON TOOLS WITH ARTIFICIAL INTELLIGENCE. nov. 2004

RUSS, J. C. Acquiring Images. In: **The image processing handbook**. Boca Raton, FL: CRC Press, 2002.

RUSSELL, S. J.; NORVIG, P.; DAVIS, E. Beyond Classical Search. In: **Artificial intelligence: a modern approach**. Prentice Hall series in artificial intelligence. 3rd ed ed. Upper Saddle River: Prentice Hall, 2010.

SAE INTERNATIONAL. **Seeking a common language for vehicle automation**. Disponível em: <<http://articles.sae.org/15462/>>. Acesso em: 17 out. 2018.

SAE INTERNATIONAL. **Taxonomy and Definitions for Terms Related to On-Road Motor Vehicle Automated Driving Systems**. [s.l.] SAE Technical Paper, jun. 2018.

SAFERCAR. **Pedestrian Automatic Emergency Braking**. Disponível em: <<https://www.safercar.gov/Vehicle-Shoppers/Safety-Technology/paeb>>. Acesso em: 2 jan. 2019.

SAPONARO, P. et al. **Improving calibration of thermal stereo cameras using heated calibration board**. 2015 IEEE International Conference on Image Processing (ICIP). **Anais...** In: 2015 IEEE INTERNATIONAL CONFERENCE ON IMAGE PROCESSING (ICIP). set. 2015

SHAPIRO, L.; STOCKMAN, G. C. Perceiving 3D from 2D Images. In: **Computer Vision**. [s.l.] Prentice Hall, 2001.

SILVA, J. D. S. DA; SIMONI, P. O.; BHARADWAJ, K. K. **Multiple correspondences in stereo vision under a genetic algorithm approach**. Proceedings 13th Brazilian Symposium on Computer Graphics and Image Processing (Cat. No.PR00878). **Anais...** In: PROCEEDINGS 13TH BRAZILIAN SYMPOSIUM ON COMPUTER GRAPHICS AND IMAGE PROCESSING (CAT. NO.PR00878). out. 2000

SOUZA, M. A. DE; SANCHES, I. J.; GAMBA, H. R. **A New Method for Generating 3D Thermography Models**. ACTAPRESS, 2012 Disponível em: <<http://www.actapress.com/PaperInfo.aspx?paperId=453249>>. Acesso em: 1 maio. 2017

TAIB, S.; JADIN, M. S.; KABIR, S. Thermal Imaging for Enhancing Inspection Reliability: Detection and Characterization. In: **Infrared Thermography**. [s.l.] INTECH Open Access Publisher, 2012.

TANG, H.; LI, Z.-L. Basic Concepts. In: **Quantitative remote sensing in thermal infrared: theory and applications**. Springer remote sensing/photogrammetry. Heidelberg: Springer, 2014.

TOULMINET, G. et al. Vehicle detection by means of stereo vision-based obstacles features extraction and monocular pattern analysis. **IEEE Transactions on Image Processing**, v. 15, n. 8, p. 2364-2375, ago. 2006.

TRUCCO, E.; VERRI, A. Stereopsis. In: **Introductory techniques for 3-D computer vision**. Englewood Cliffs: Prentice Hall, 1998a. v. 201.

TRUCCO, E.; VERRI, A. Camera Calibration. In: **Introductory techniques for 3-D computer vision**. Englewood Cliffs: Prentice Hall, 1998b. v. 201.

WANG, T. et al. Integrating Millimeter Wave Radar with a Monocular Vision Sensor for On-Road Obstacle Detection Applications. **Sensors**, v. 11, n. 9, p. 8992-9008, 21 set. 2011.

WANG, Y.; WU, J.; LIN, C. Skidmark Patterns and Identification of ABS-Equipped Passenger car. **Journal of the Eastern Asia Society for Transportation Studies**, v. 6, p. 3401-3412, 2005.

WHO. **Global status report on road safety 2015**. Disponível em:  
<[http://www.who.int/violence\\_injury\\_prevention/road\\_safety\\_status/2015/en/](http://www.who.int/violence_injury_prevention/road_safety_status/2015/en/)>.  
Acesso em: 23 jul. 2017.

WIKI COMMONS. **Wien's law of radiation**. Disponível em:  
<[https://commons.wikimedia.org/wiki/File:Wiens\\_law\\_vis\\_limits.svg](https://commons.wikimedia.org/wiki/File:Wiens_law_vis_limits.svg)>. Acesso em: 29  
out. 2018.

YOUNG, H. D.; FREEDMAN, R. A.; FORD, A. L. Temperatura e calor. In: **Sears & Zemansky Física II: Termodinâmica e Ondas**. 12. ed. São Paulo: Pearson Addison Wesley, 2008.

ZHANG, Z. A flexible new technique for camera calibration. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, v. 22, n. 11, p. 1330-1334, nov. 2000.

ŽIVČÁK, J. et al. Physical Nature of Thermography. In: **Methodology, models and algorithms in thermographic diagnostics**. Berlin; New York: Springer Science & Business Media, 2013. p. 31--54.