

PAULO CEZAR DE QUEIROZ HERMIDA

**Reconhecimento Automático de Armas de
Fogo no Interior de Veículos.**

Manaus

2017

PAULO CEZAR DE QUEIROZ HERMIDA

**Reconhecimento Automático de Armas de Fogo no
Interior de Veículos.**

Dissertação apresentada ao Programa de Pós-Graduação em Engenharia Elétrica da Universidade Federal do Amazonas, como requisito parcial para a obtenção do título de Mestre em Engenharia Elétrica.

Área de concentração: Controle e Automação de Sistemas.

Orientador: Prof. Dr. José Luiz de Souza Pio

Manaus

2017

Ficha Catalográfica

Ficha catalográfica elaborada automaticamente de acordo com os dados fornecidos pelo(a) autor(a).

H554r Hermida, Paulo Cezar de Queiroz
Reconhecimento Automático de Armas de Fogo no Interior de Veículos. / Paulo Cezar de Queiroz Hermida. 2017
89 f.: il. color; 31 cm.

Orientador: Prof. Dr. José Luiz de Souza Pio.
Dissertação (Mestrado em Engenharia Elétrica) - Universidade Federal do Amazonas.

1. Descritores Locais.. 2. Identificação de assalto a veículos.. 3. Reconhecimento de armas de fogo.. 4. Segurança em veículos.. I. Pio., Prof. Dr. José Luiz de Souza II. Universidade Federal do Amazonas III. Título

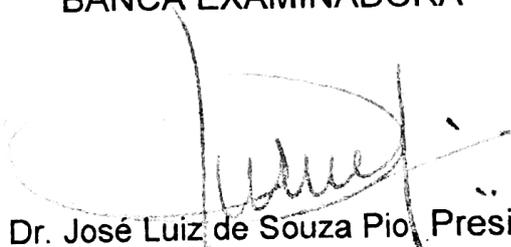
PAULO CEZAR DE QUEIROZ HERMIDA

RECONHECIMENTO AUTOMÁTICO DE ARMAS DE FOGO NO INTERIOR DE VEÍCULOS.

Dissertação apresentada ao Programa de Pós-Graduação em Engenharia Elétrica da Universidade Federal do Amazonas, como requisito parcial para obtenção do título de Mestre em Engenharia Elétrica na área de concentração Controle e Automação de Sistemas.

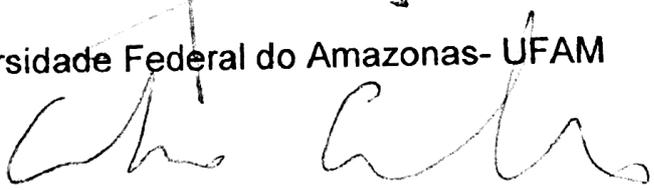
Aprovado em 09 de junho de 2017.

BANCA EXAMINADORA



Prof. Dr. José Luiz de Souza Pio, Presidente

Universidade Federal do Amazonas- UFAM



Prof. Dr. Celso Barbosa Carvalho, Membro

Universidade Federal do Amazonas- UFAM



Profª. Dra. Eulanda Miranda dos Santos, Membro

Universidade Federal do Amazonas- UFAM

*Dedico esse trabalho à minha mãe Maria Olga, à minha esposa Suelly e a meus filhos
Juliana e Paulo.*

AGRADECIMENTOS

Ao meu orientador professor doutor José Luiz de Souza Pio, pela formação, orientação e incentivo;

Aos professores do curso de pós-graduação em engenharia elétrica pelo aprendizado adquirido;

À Universidade Federal do Amazonas e em especial ao Programa Pós Graduação em Engenharia Elétrica – PPGEE;

A todos meus familiares e amigos pelo apoio;

Resumo

O aumento da violência urbana no Brasil tem destacado o crescente número de assaltos a mão armada no interior de veículos. Os dados registrados impressionam, somando o alarmante número de 57 veículos assaltados por hora no país. Manaus apresenta-se como uma das cidades brasileiras com maior número de assaltos a veículos. O Sindicato das Empresas de Transporte de Passageiros do Estado do Amazonas (Sinetram) já registra, apenas nos primeiros quatro meses de 2017 o alarmante número de 1.120 assaltos a ônibus em Manaus. Por outro lado, o Sindicato dos Taxistas do Amazonas (Sintax-AM) aponta que, pelo menos dez taxistas são assaltados por dia na cidade. O objetivo deste trabalho é o desenvolvimento de um método que faça o reconhecimento automático de assaltos a mão armada no interior de veículos automotivos. Contribuindo dessa forma para o combate a violência urbana e viabilizando a atuação mais rápida e efetiva dos agentes de segurança pública. A abordagem que será adotada consiste na criação de um conjunto de descritores locais, gerados a partir de uma sequência de imagens de armas de fogo (revólveres e pistolas). Esses descritores fornecem uma base de informações capaz de identificar a presença de armas de fogo, nas imagens capturadas do interior de veículos de passeio. Essa abordagem dispensa a localização da arma no espaço da imagem e a reconhece a partir de um conjunto de características otimizadas. Os resultados obtidos mostram que o método desenvolvido reconhece a arma de fogo em diferentes situações de movimento, com índices de acerto acima de 80% em todas as métricas utilizadas. O método é integrável aos sistemas de segurança veicular modernos e suficientemente robusto para o monitoramento contínuo do interior de carros de passeio.

Palavras-chaves: Descritores Locais; Reconhecimento de armas fogo; Roubos em Veículos.

Abstract

The increase in urban violence in Brazil has highlighted the growing number of armed robberies inside vehicles. The data recorded impressed, adding to the alarming number of 57 vehicles robbed per hour in the country. Manaus presents itself as one of the Brazilian cities with the highest number of vehicle assaults. The Sinetram Passenger Transport Companies Union (Sinetram) already registers in the first four months of 2017 the alarming number of 1,120 bus robberies in Manaus. On the other hand, the Union of Taxistas do Amazonas (Sintax-AM) points out that at least ten taxi drivers are assaulted per day in the city. The objective of this work is the development of a method that automatically recognizes armed robberies inside automotive vehicles. Contributing in this way to the fight against urban violence and enabling the quicker and more effective action of public security agents. The approach that will be adopted is the creation of a set of local descriptors, generated from a sequence of images of firearms (revolvers and pistols). These descriptors provide an information base capable of identifying the presence of firearms in the images captured from the interior of walking vehicles. This approach dispenses with the location of the weapon in the image space and recognizes it from a set of optimized features. The obtained results show that the developed method recognizes the firearm in different situations of movement, with hit rates above 80 % in all metrics used. The method is integrable with modern vehicle safety systems and robust enough for continuous monitoring of the interior of passenger cars.

Keywords: Local Descriptors; Recognition of firearms; Robbery in Vehicles.

Lista de figuras

Figura 1 – Espaço de escalas com duas oitavas, na segunda oitava a escala foi reduzida de um fator k	22
Figura 2 – Processo de criação da DoG.	24
Figura 3 – Visão da DoG com a definição de um ponto extremo através da intensidade do seu brilho em relação a seus vizinhos.	25
Figura 4 – Máximos e mínimos utilizados para encontrar pontos extremos nas bordas.	27
Figura 5 – Extração do descritor de um ponto chave em função da amplitude e da orientação do gradiente de seus vizinhos, a esquerda o mapa de gradientes e a direita o descritor do ponto chave.	29
Figura 6 – Histograma com as orientações de um ponto chave, os pontos mais altos definem a orientação dominante.	31
Figura 7 – Resultado encontrado com a variação no número de orientações e na dimensão da matriz de histogramas.	33
Figura 8 – Abaixo o resultado da fusão de imagens, a esquerda a imagem normal no centro a imagem do sensor infravermelho e a direita a imagem fundida.	35
Figura 9 – O resultado da fusão de imagens a esquerda a imagem normal e a direita a com infravermelho.	35
Figura 10 – Equipamento MIRAE utilizado para a detecção de armas a curta distância.	37
Figura 11 – A esquerda a imagem natural e a direita a imagem em escala de cinza gerada pela recepção dos dados do MIRAE.	37
Figura 12 – <i>Setup</i> experimental montado para a realização dos testes, considerando a distância entre a arma e a antena.	38
Figura 13 – Nos gráficos abaixo a faixa de distância em que a arma foi detectada, a esquerda sem a correção e a direita com a correção no posicionamento.	38
Figura 14 – Resultado da técnica de multi polarização, (a) imagem original, (b) Imagem com antenas de polaridades complementares e em (c) antenas com polaridades cruzadas.	39
Figura 15 – O sistema utiliza Modelo Oculto de Markov para identificar a partir das características do áudio o tipo de arma.	40
Figura 16 – Durante o procedimento experimental foram utilizadas quatro armas para validação.	41

Figura 17 – Na Figura o processo de segmentação da imagem em (a) o momento anterior ao evento em (b) o evento e em (c) a segmentação da região de interesse.	42
Figura 18 – Templates utilizados no processo de busca de pessoas com braços levantados na imagem capturada.	42
Figura 19 – Resultado do <i>matching</i> entre a imagem capturada e segmentada e o <i>template</i>	43
Figura 20 – Fluxograma do método proposto por Tiwari e Verma (2015a).	44
Figura 21 – Segmentação da cor preta sendo aplicada na imagem.	45
Figura 22 – Resultados obtidos com a) Diferentes fundos; b) Variação na iluminação; c) Armas diferentes; d) Oclusão; e) Variação na escala e rotação; f) Múltiplas armas.	45
Figura 23 – Resumo dos trabalhos relacionados.	46
Figura 24 – Método utilizado para validar o uso de descritores locais como ferramenta principal para a identificação de armas de fogo no interior de veículos.	48
Figura 25 – Fluxograma com o processo de detecção de pontos chave.	49
Figura 26 – Exemplo de um espaço de escala com uma oitava formada por uma imagem original e quatro imagens geradas a partir da aplicação do filtro gaussiano para $\sigma = 5$, a sequência vai da esquerda para a direita e de cima para baixo.	50
Figura 27 – Cálculo da DoG utilizando a oitava apresentada anteriormente.	51
Figura 28 – Detecção de pontos chave em seis imagens de armas da base de dados IMFDB.	52
Figura 29 – Na imagem foram encontrados 151 pontos chave, logo abaixo uma lista parcial destes pontos com suas características.	53
Figura 30 – Na Figura 29 foram encontrados 151 pontos chave, abaixo uma listagem parcial dos 128 descritores locais extraídos de cada ponto chave.	54
Figura 31 – Listagem parcial do arquivo da base de descritores locais.	55
Figura 32 – Processo de otimização da base de descritores utilizando a diferença Euclidiana e a qualidade do ponto chave.	56
Figura 33 – Processo de monitoramento do ambiente interno do veículo em busca de armas de fogo, com o processo de <i>matching</i> entre os pontos da base de descritores e os pontos da imagem capturada do interior do veículo.	57

Figura 34 – Procedimento que verifica a semelhança entre dois descritores, o valor da distância Euclidiana foi calculado de forma empírica.	58
Figura 35 – Comportamento do método em função da variação do parâmetro <i>limiarDistancia</i>	59
Figura 36 – Matriz de confusão	61
Figura 37 – Alguns exemplos de imagens de armas gerados a partir da base de dados IMFDB (2017) utilizados nos experimentos.	62
Figura 38 – Abaixo o fluxo do experimento I, que visa obter a resposta do método quando houver uma mudança no ângulo da arma.	63
Figura 39 – Resultados obtidos para uma rotação de 0^0	64
Figura 40 – Resultados obtidos para uma rotação de 90^0	65
Figura 41 – Resultados obtidos para uma rotação de 180^0	66
Figura 42 – Resultados obtidos para uma rotação de 270^0	67
Figura 43 – Gráfico com a variação nos resultados obtidos pelo método em função das rotações na imagem.	68
Figura 44 – Gráfico com o desvio padrão dos resultados obtidos pelo método em função das rotações na imagem.	68
Figura 45 – Abaixo o fluxo do experimento II, que visa obter a resposta do método quando ocorre ruído na imagem.	69
Figura 46 – Resultados obtidos para ruído com $\sigma = 0$	70
Figura 47 – Resultados obtidos para ruído com $\sigma = 80$	71
Figura 48 – Resultados obtidos para ruído com $\sigma = 90$	72
Figura 49 – Resultados obtidos para ruído com $\sigma = 120$	73
Figura 50 – Gráfico com a variação nos resultados obtidos pelo método em função da variação no nível de ruído na imagem.	74
Figura 51 – Gráfico com o desvio padrão dos resultados obtidos pelo método em função variação no nível de ruído na imagem.	74
Figura 52 – Fluxograma do experimento III, aonde será feita comparação entre os resultados obtidos por Tiwari e Verma (2015a) e o do método desenvolvido.	75
Figura 53 – Exemplos de imagens utilizadas por Tiwari e Verma (2015a).	76
Figura 54 – Resultado final do método, com a classificação de 12 <i>frames</i> que contém a arma e que possuem variação no fundo da imagem.	77

Figura 55 – Resultado final do método, com a classificação de 9 <i>frames</i> que contém a arma e que possuem variação na iluminação da imagem.	77
Figura 56 – Resultado final do método, com a classificação de 11 <i>frames</i> que contém a arma e que possuem variação na classe da arma	78
Figura 57 – Resultado final do método, com a classificação de 17 <i>frames</i> que contém a arma e que possuem oclusão	79
Figura 58 – Resultado final do método, com a classificação de 10 <i>frames</i> que contém a arma e que possuem variação na escala e rotação	79
Figura 59 – Resultado final do método, com a classificação de 6 <i>frames</i> que contém múltiplas armas.	80
Figura 60 – Resultado final com os resultados alcançados pelo método desenvolvido nos 6 cenários.	81
Figura 61 – Fluxograma do experimento IV aonde será feita a análise dos resultados do método desenvolvido classificando um vídeo da internet.	82
Figura 62 – Resultados obtidos da classificação do vídeo pelo método desenvolvido.	83

Lista de tabelas

Tabela 1 – Resumo da relação entre os autovalores de matriz Hessiana com o tipo de ponto extremo. Fonte Moroneze (2013).	27
Tabela 2 – Resultados alcançados nos itens da Figura 22.	46
Tabela 3 – Tabela com as métricas usadas para avaliar a performance dos experimentos.	61

—

Lista de abreviaturas e siglas

IMFDB	<i>Internet Movie Firearms Database;</i>
DoG	<i>Differences of Gaussian;</i>
LoG	<i>Laplacian of Gaussian;</i>
ORB	<i>Oriented FAST,rotated BRIEF;</i>
OpenCV	<i>Open Source Computer Vision Library;</i>
SIFT	<i>Scale Invariant Feature Transform;</i>
SURF	<i>Robust Independent Elementary Features;</i>
ROC	<i>Receiver Operating Characteristic.</i>

Sumário

1	INTRODUÇÃO	17
1.1	Contextualização	17
1.2	Justificativa	18
1.3	Objetivos	19
1.3.1	Geral	19
1.3.2	Específicos	19
1.4	Organização do trabalho	20
2	REVISÃO BIBLIOGRÁFICA	21
2.1	Fundamentação teórica	21
2.1.1	O algoritmo SIFT	21
2.1.2	O Espaço de escalas	21
2.1.3	Detecção de pontos extremos.	23
2.1.4	Convertendo pontos extremos em pontos chaves	25
2.1.4.1	Eliminação de pontos extremos com baixo contraste	26
2.1.4.2	Eliminação de pontos extremos localizados em bordas	26
2.1.5	Extração dos descritores locais	29
2.1.6	Extraindo as características que formam os descritores de um ponto chave	30
2.1.7	Parametrização do descritor local	32
2.1.8	Correspondência entre pontos chaves	33
2.2	Trabalhos relacionados	34
2.2.1	Detecção de armas de fogo utilizando infravermelho	34
2.2.2	Detecção de armas de fogo utilizando ondas milimétricas	36
2.2.3	Detecção de armas de fogo utilizando áudio	39
2.2.4	Detecção de armas de fogo utilizando <i>template matching</i>	41
2.2.5	Detecção de armas de fogo utilizando segmentação e ex- tração de características	43
3	RECONHECIMENTO DE ARMA DE FOGO	47

3.1	Fase de criação e otimização da base de descritores locais	47
3.2	Fase de monitoramento	57
3.2.1	Procedimento para definição do limiar da distância Euclidiana	58
4	APLICAÇÕES EXPERIMENTAIS	61
4.1	Descrição dos experimentos e métricas	61
4.2	Base da dados de imagens utilizada nos experimentos	62
4.3	Procedimento experimental I - Análise da resposta do método com relação mudanças no ângulo da arma . . .	63
4.3.1	Resultados do Experimento I	63
4.3.1.1	Resultados da Fase 1 - Rotação de 0^0	63
4.3.1.2	Resultados da Fase 2 - Rotação de 90^0	64
4.3.1.3	Resultados da Fase 3 - Rotação de 180^0	65
4.3.1.4	Resultados da Fase 4 - Rotação de 270^0	66
4.3.2	Discussão dos resultados do Experimento I	67
4.4	Procedimento experimental II - Análise da resposta do método com relação a ocorrência de ruídos na imagem	69
4.4.1	Resultados do Experimento II	70
4.4.1.1	Resultados da Fase 1 - $\sigma = 0$	70
4.4.1.2	Resultados da Fase 2 - $\sigma = 80$	70
4.4.1.3	Resultados da Fase 3 - $\sigma = 90$	71
4.4.1.4	Resultados da Fase 4 - $\sigma = 120$	72
4.4.2	Discussão dos resultados do Experimento II	73
4.5	Procedimento experimental III - Comparação dos resultados entre o experimento realizado por Tiwari e Verma (2015a) e o método desenvolvido	75
4.5.1	Resultados do Experimento III	76
4.5.2	Discussão dos resultados do Experimento III	80
4.6	Procedimento experimental IV - Análise da resposta do método desenvolvido aplicado a um vídeo real . . .	81

4.6.1	Resultados do Experimento IV	82
4.6.2	Discussão dos resultados do Experimento IV	83
5	CONCLUSÕES E TRABALHOS FUTUROS . . .	84
5.1	Conclusões	84
5.2	Limitações do trabalho	85
5.2.1	Ruído na imagem	85
5.2.2	Movimentos bruscos	86
5.3	Trabalhos futuros	86
5.4	Considerações finais	86
	Referências ¹	87

¹ De acordo com a Associação Brasileira de Normas Técnicas. NBR 6023.

1 INTRODUÇÃO

1.1 Contextualização

A violência urbana no Brasil tem contribuído para o aumento considerável de assaltos e roubos a mão armada no interior de veículos. Os jornais trazem diariamente notícias de assaltos a coletivos, táxis e veículos particulares. Segundo a Associação de seguradoras de veículos do Brasil Privados (2016), no último semestre de 2016 havia 7.584.171 veículos segurados em todo o Brasil e nesse período foram contabilizados 136.348 eventos de roubos. Em Manaus é alarmante o número de ocorrências de assaltos a mão armada em ônibus e táxis. O Sindicato das Empresas de Transporte de Passageiros do Estado do Amazonas (Sinetram) registrou, apenas nos primeiros quatro meses de 2017, o alarmante número de 1.120 assaltos a ônibus. Por outro lado, o Sindicato dos Taxistas do Amazonas (Sintax-AM) aponta que, pelo menos dez taxistas são assaltados por dia em Manaus. E no ano de 2016, 14 motoristas de táxi foram mortos em assaltos a mão armada.

O tema deste trabalho está relacionado com as ocorrências de assaltos a mão armada no interior de carros de passeio. O problema a ser tratado pode ser sintetizado por meio da seguinte pergunta: *Como reconhecer automaticamente a presença ou o porte de uma arma de fogo, a partir de dados visuais da cena capturada do interior de veículos?*

Este problema está relacionado com a área de Visão Computacional no tema específico de reconhecimento de objetos. A Visão computacional é a área científica que estuda e desenvolve soluções para a extração de informações do mundo real a partir de uma ou mais imagens, Trucco e Verri (1998). Trata explicitamente da obtenção e manipulação dos dados de uma ou mais imagens e do seu uso para diferentes propósitos. No caso particular do reconhecimento de objetos, busca-se por soluções que localizem os objetos alvo no espaço da imagem e em seguida o reconheçam, identificando-os a partir de diferentes tipos de características, como: forma, cor, casamento de modelos entre outras. Parte-se do princípio de que os objetos são reconhecidos quando são encontradas correspondências entre a imagem do objeto e as informações das características do objeto que estão armazenadas na memória.

Algumas soluções para a identificação da ocorrência de assaltos tem sido propostas, como em Sreedevi e Nair (2011), Manjunath et al. (2013), Bavya e Mohanamurali (2014), Liu, Zhang e Li (2013), Manjunath et al. (2013), Das e Sengupta (2016), Shruthi et

al. (2015), todos estes trabalhos apresentam soluções alternativas ao problema. Usando diferentes técnicas, tipos de informação ou empregando diferentes tipos de sensores. Os descritores locais, que podem ser definidos como um conjunto de características extraídas no entorno de um ponto específico na imagem, foram utilizados por Lowe (1999), Lowe (2004), Suga et al. (2008), Höferlin e Heidemann (2010), Coutinho e Maroquim (2010), Silva et al. (2012), Karami, Prasad e Shehata (2010), Huang, Mu e Zeng (2016), Li et al. (2016b), Lee, Rhee e Lee (2016), Fang et al. (2016), Ma et al. (2017), e tem se apresentado como uma abordagem mais apropriada para o reconhecimento de objetos.

O método desenvolvido dividi-se em duas etapas: A primeira etapa consiste na criação de um conjunto de descritores locais, que são obtidos a partir de imagens de armas de fogo (revólveres e pistolas). Esses descritores fornecem uma base de informações otimizada com as características mais relevantes do conjunto de imagens. A segunda etapa é o reconhecimento do porte ou a presença da arma de fogo no interior de um veículo. Isso é feito através da extração *online* de descritores locais que são gerados a partir de sequências de vídeo capturadas do interior do veículo. Este novo conjunto de descritores é comparado com os que estão armazenados na base. Dentro de critérios pré definidos a semelhança entre eles é estabelecida, indicando assim a presença da arma. Essa abordagem dispensa a localização da arma no espaço da imagem pois a reconhece a partir de um conjunto de características otimizadas.

A principal contribuição deste trabalho é o desenvolvimento de um método de reconhecimento de armas de fogo totalmente integrável aos modernos sistemas de segurança. Fornecendo para a sociedade uma alternativa segura e eficaz contra a violência urbana. Sob o ponto de vista científico, essa abordagem avança no dispensando a localização do objeto procurado no espaço de busca na imagem. A estruturação de uma base de características otimizadas no espaço e no tempo de reconhecimento, também é uma contribuição deste trabalho.

1.2 Justificativa

Diferente das outras abordagens utilizadas para a identificação de armas de fogo como o uso de sensores Infravermelho(Xue e Blum (2003), Upadhyay e Rana (2014), Xu e Wu (2015), Hussein e Hu (2016), Chen et al. (2005)) ou de Ondas Milimétricas utilizando equipamentos que enviam e recebem essas ondas (Kim et al. (2011), Kumar, Upadhyay

e Singh (2015), Li et al. (2016a), Kapilevich et al. (2013), Moulder et al. (2016)), ou utilizando características do sinal de Áudio gerado pelo disparo da arma de fogo (Kiktova et al. (2015), Ahmed, Uppal e Muhammad (2013), Vozáriková et al. (2011)), ou usando a técnica de *Template Matching* (Hung e Pan (2015), Dever, Lobo e Shah (2002)) o método desenvolvido neste trabalho, não utiliza nenhum tipo de sensor ou hardware adicional além da câmera. Sendo que o método utiliza como fonte de informação o conjunto de características obtidas de sequencias de vídeo geradas a partir do interior do veículo. Nota-se ainda que as abordagens que utilizam descritores locais, partem de uma base de referência com o foco no objeto que se procura identificar, sem a preocupação de estabelecer critérios de otimização das características extraídas. A base de descritores gerada pelo método desenvolvido, passa por um processo de otimização, que visa aumentar a performance durante a fase de *matching*.

1.3 Objetivos

Como forma de melhor organizar os objetivos do trabalho ele foram subdivididos em geral e específicos.

1.3.1 Geral

O objetivo geral deste trabalho é desenvolver um identificador de arma de fogo, revólveres e pistolas presentes no interior de veículos de passeio.

1.3.2 Específicos

1. Criar e otimizar uma base de descritores locais de imagens de armas de fogo;
2. Desenvolver um extrator de características da imagem capturada do interior de um veículo;
3. Criar um identificador capaz de reconhecer a semelhança entre as imagens de armas de fogo da base de descritores e as imagens capturadas do interior do veículo.

1.4 Organização do trabalho

O presente trabalho encontra-se organizado da seguinte forma. No Capítulo 2 será apresentada a revisão bibliográfica e a fundamentação teórica, com o detalhamento do processo de detecção de pontos chave e extração de descritores locais. Também pertence a este capítulo o processo de correspondência entre pontos chave da base de treinamento, e os da imagem capturada no interior do veículo na qual se busca a arma. Também estão neste capítulo a revisão dos trabalhos relacionados que tem como objetivo a identificação de armas, mas que utilizam técnicas diferentes. No Capítulo 3 está a descrição do método desenvolvido detalhando as suas duas fases, que são: Criação e otimização da base de descritores locais e Monitoramento. No Capítulo 4 estão descritos os experimentos realizados e seus resultados. No Capítulo 5 está a conclusão e as considerações finais assim como os trabalhos futuros.

2 REVISÃO BIBLIOGRÁFICA

Neste capítulo serão apresentados os conceitos fundamentais que suportam o desenvolvimento deste trabalho. A hipótese de que a análise da imagem no interior do veículo, utilizando descritores locais irá solucionar o problema, será aqui detalhada. Também serão apresentados trabalhos que tem como objetivo identificar armas, porém utilizam técnicas diferentes para alcançar este objetivo.

2.1 Fundamentação teórica

Serão apresentados os conceitos de espaço de escalas, pontos chave sua localização e orientação, extração de descritores locais entre outros. O processo de correspondência entre pontos, utilizando o cálculo de distâncias Euclidianas, também será detalhado.

O processo de detecção de pontos chave e de extração de descritores locais usados neste trabalho, utiliza como base o algoritmo SIFT (*Scale Invariant Feature Transform*) criado por Lowe (1999), que é descrito a seguir.

2.1.1 O algoritmo SIFT

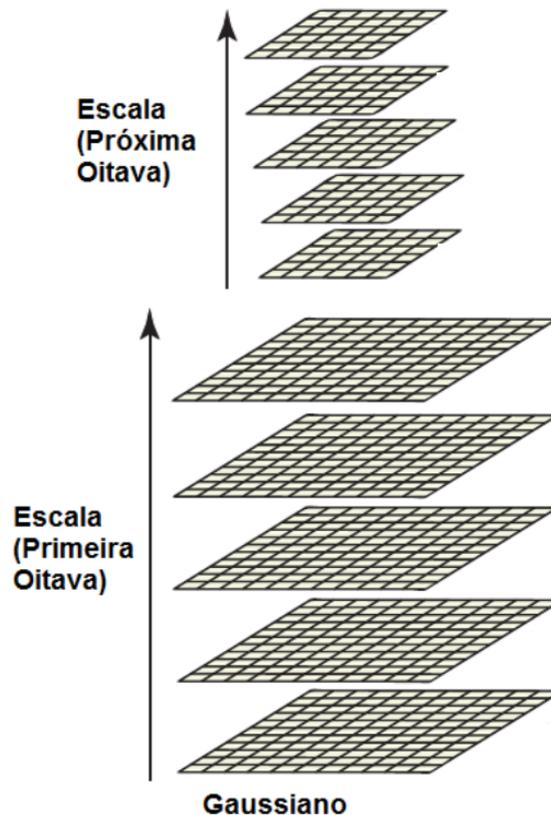
O algoritmo SIFT é um algoritmo de visão computacional criado por Lowe (1999) e é formado por duas fases bem definidas: Detecção de pontos chave e Extração de descritores locais. A fase de detecção envolve o processo de criação do espaço de escalas, geração da diferença de gaussianas e o processo de filtragem para a eliminação de pontos com baixo contraste ou que estejam localizados próximos a bordas. A fase de extração consiste no uso dos pontos chave encontrados na fase de detecção, para extrair as características de amplitude e direção do gradiente nos vizinhos destes pontos. Estes dados serão utilizados para gerar os descritores locais. Todos estes procedimentos serão detalhados a seguir.

2.1.2 O Espaço de escalas

O espaço de escalas é formado por imagens sobrepostas, estas camadas formam grupos que são chamados de oitavas. Todas as imagens pertencentes a mesma oitava possuem a mesma escala, e o que as difere é a aplicação sucessiva de um filtro gaussiano,

que está definido na Equação 2. Na Figura 1 um exemplo de espaço de escala com duas oitavas.

Figura 1 – Espaço de escalas com duas oitavas, na segunda oitava a escala foi reduzida de um fator k .



Fonte: Lowe (2004).

Segundo Teixeira (2001) um sistema de visão genérico não tem informações sobre as escalas presentes na imagem de entrada, portanto deve-se realizar o processamento da imagem em baixo nível em todas as escalas possíveis, deixando para depois a decisão de quais escalas são mais adequadas. Em Silva (2012) afirma-se que o uso do espaço de escalas tende a eliminar as estruturas menores e a encontrar estruturas maiores, que mantém sua visibilidade em diferentes escalas. O espaço de escala gerado através da convolução das imagens com o núcleo Gaussiano é conhecido como Espaço de Escala Linear ou Gaussiano. Segundo Velho, Teixeira e Gomes (2000) suas principais propriedades são:

1. Invariância a translação: $L(x, y, \sigma)$ foi gerada a partir de uma convolução que por definição é invariante a translação;

2. Linearidade: O mapeamento de L na imagem original i pertencente ao espaço de escala é uma transformação linear, pois: $L_{f+\lambda g}(x, y, \sigma) = L_f(x, y, \sigma) + \lambda L_g(x, y, \sigma)$;
3. Causalidade: O número de pontos críticos não aumenta a medida que σ cresce, isso é válido apenas para o caso unidimensional.

A criação do Espaço de Escalas é uma das fases do processo de detecção de pontos chave. O seu principal objetivo é encontrar estruturas invariantes a mudanças de escalas.

Uma imagem original $I(x, y)$ passar a ser definida no espaço de escalas como $L(x, y, \sigma)$, portanto:

$$L(x, y, \sigma) = G(x, y, \sigma) * I(x, y). \quad (1)$$

Dado que,

$$G(x, y, \sigma) = \frac{1}{2\pi\sigma^2} e^{-(x^2+y^2/2\sigma^2)}, \quad (2)$$

temos que $L(x, y, \sigma)$ é o resultado da convolução de $I(x, y)$ com um filtro Gaussiano $G(x, y, \sigma)$ de variância σ arbitrária, conforme as equações (1) e (2).

O processo de criação das oitavas começa quando é aplicado na primeira imagem $I(x, y)$ da primeira oitava, um filtro gaussiano $G(x, y, \sigma)$. Sendo que a próxima imagem $L(x, y, \sigma)$ será o resultado dessa convolução, e esse processo se repete até a última imagem dessa oitava. Na próxima oitava a escala é alterada de um fator k e o processo recomeça, sendo que esse fator passa a fazer parte do filtro gaussiano $G(x, y, k\sigma)$. E esse processo continua até a última imagem da última oitava ser suavizada.

O número de oitavas e a quantidade de imagens presentes em cada uma delas, depende do tamanho da imagem original, Lowe (2004) sugere que quatro oitavas com 5 imagens em cada uma é uma boa configuração.

2.1.3 Detecção de pontos extremos.

Após a criação do espaço de escalas, as imagens que formam as oitavas serão utilizadas para gerar um novo conjunto de imagens, chamado de DoG (Diferença de Gaussianas). Esse novo conjunto é gerado pela subtração entre as imagens da mesma oitava no espaço de gaussianas. As imagens resultantes das subtrações serão usadas para identificar pontos extremos, pois durante seu processo de formação as bordas e os cantos se destacam. Segundo Lindeberg (1994) o correto seria utilizar o Laplaciano das Gaussianas

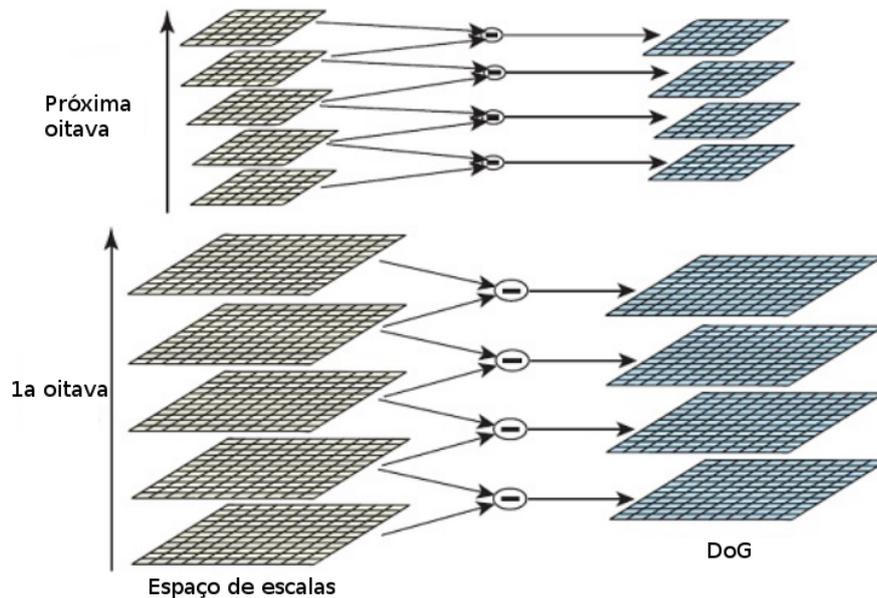
para encontrar esses pontos extremos, pois é um descritor invariante para rotação e translação de coordenadas e seus zeros são invariantes a mudanças de brilho, além disso é um operador linear. No entanto Lowe (2004) mostrou que esse procedimento pode ser aproximado pela DoG. A DoG é calculada conforme a Equação 3, e o resultado está na Figura 2.

$$D(x, y) = L_{i+1}(x, y, \sigma) - L_i(x, y, \sigma), \quad (3)$$

onde:

- D : Imagem resultante de diferença;
- L_{i+1} : É a imagem no espaço de gaussianas;
- L_i : É a imagem anterior no espaço de gaussianas;
- x, y : Coordenada dos *pixels*;
- σ : Parâmetro do filtro gaussiano ao qual a imagem foi submetida.

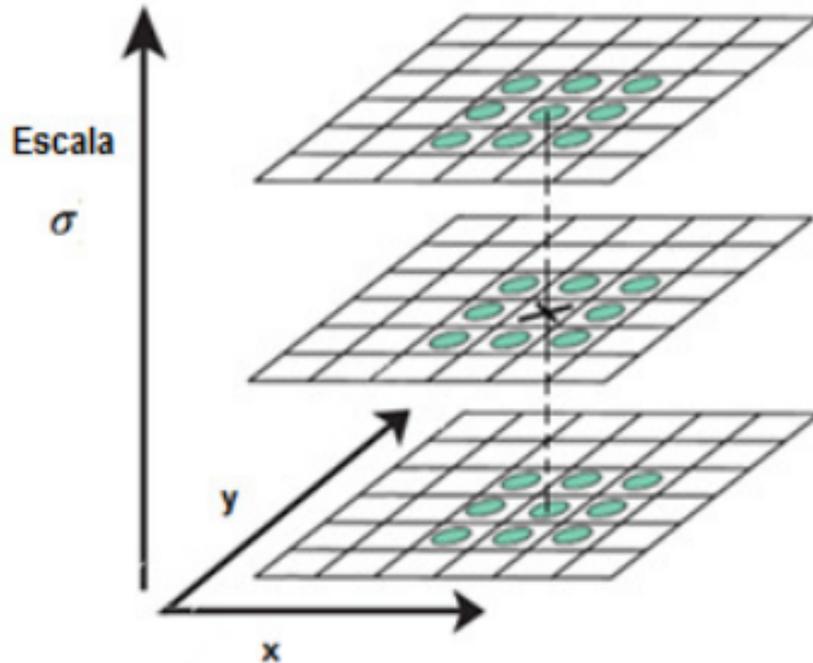
Figura 2 – Processo de criação da DoG.



Fonte: Lowe (2004).

Neste ponto o Espaço de escala já está finalizado assim como o cálculo das DoG's. O próximo passo é detectar os pontos extremos em cada DoG. O critério para a escolha destes pontos, é determinar se eles tem maior ou menor intensidade de brilho em relação a seus vizinhos na sua imagem, na imagem anterior e na posterior. Um exemplo desse processo pode ser observado na Figura 3.

Figura 3 – Visão da DoG com a definição de um ponto extremo através da intensidade do seu brilho em relação a seus vizinhos.



Fonte: Lowe (2004).

Na sua própria imagem o ponto extremo tem 8 pontos vizinhos e na imagem de cima são 9 pontos e na de baixo também 9 pontos. Então para um determinado ponto são analisados $8+9+9 = 26$ pontos vizinhos, o ponto X é marcado como candidato a ser um ponto extremo se for mínimo ou máximo entre esses 26 vizinhos. Aplicados os critérios de máximos e mínimos serão encontrados muitos pontos extremos, porém nem todos vão se transformar em pontos chaves. Isto ocorre por que muitos pontos estão localizados próximos a bordas, ou apresentam baixo contraste e por isso devem ser eliminados.

2.1.4 Convertendo pontos extremos em pontos chaves

Até aqui foram identificados os pontos extremos através da análise da máximos e mínimos no espaço de imagens DoG, no entanto nem todos os pontos vão ser considerados pontos chave. É preciso eliminar aqueles que possuem baixo contraste ou que estejam localizados próximos a bordas. Desta forma os que não forem eliminados se tornarão pontos chave, entorno dos quais serão extraídos os descritores locais, que é o nosso objetivo. A seguir será detalhado esse processo de eliminação de pontos extremos inadequados.

2.1.4.1 Eliminação de pontos extremos com baixo contraste

O primeiro passo desse processo é a localização exata do ponto extremo, isso ocorre pois os pontos extremos detectados podem estar entre dois *pixels*. A solução para esse problema foi dada por Brown e Lowe (2002) que propôs um método que tem como objetivo, encontrar a localização interpolada do ponto extremo. Isso é feito utilizando a série de Taylor da função do espaço de escala $D(x, y, \sigma)$ até o termo quadrático, sendo assim:

$$D(x, y, \sigma) = D + \frac{\delta D^T}{\delta X} X + \frac{1}{2} X^T \frac{\delta^2 D}{\delta X^2} X. \quad (4)$$

Onde D e suas derivadas são obtidas no ponto de amostragem e $X = (x, y, \sigma)^T$ é o deslocamento a partir desse ponto. A localização do ponto extremo \hat{x} é determinada tomando a derivada da Equação 4 em relação a X e igualando a zero, resultando em:

$$\hat{x} = -\frac{\delta^2 D^{-1} \delta D}{\delta X^2 \delta X}. \quad (5)$$

Para obter a estimativa interpolada da localização do extremo o deslocamento \hat{x} é adicionado ao valor da função na localização do ponto de amostragem.

$D(\hat{x})$ é utilizado para descartar extremos instáveis com baixo contraste e é calculado abaixo:

$$D(\hat{x}) = D + \frac{1}{2} \frac{\delta D^T}{\delta X} \hat{x}. \quad (6)$$

Segundo Lowe (2004) os extremos com valor de $|D(\hat{x})| < 0.03$ devem ser descartados, considerando que os valores dos *pixels* estão normalizados entre 0 e 1.

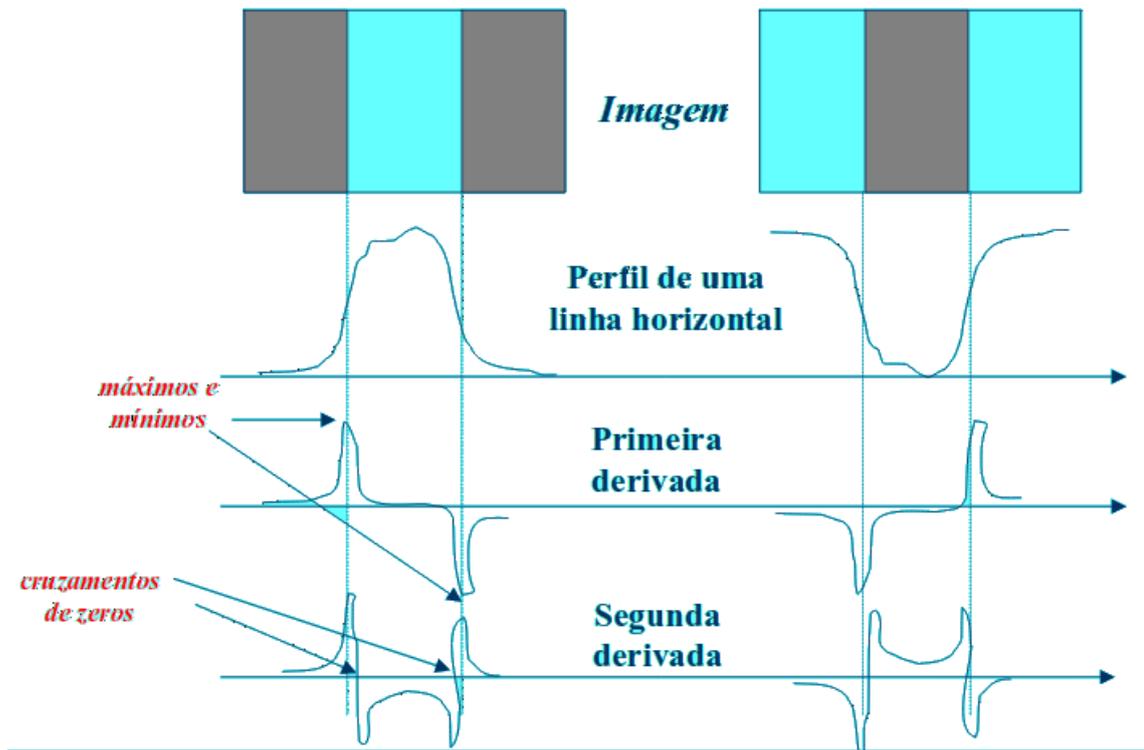
2.1.4.2 Eliminação de pontos extremos localizados em bordas

Até aqui os pontos extremos instáveis com relação ao contraste foram eliminados utilizando o critério de Lowe (2004), porém além deles ainda precisam ser eliminados os pontos extremos que estão localizados próximos a bordas.

A Equação (3) que foi utilizada para a geração da DoG tem uma resposta forte ao longo das bordas. Segundo Thome (2010) em termos de intensidade os pontos extremos tem um valor alto na borda, porém na sua vizinhança perpendicular os valores são baixos. A Figura 4 mostra as propriedades das bordas em relação aos máximos e mínimos. Para

encontrar uma relação entre esses dois grupos serão utilizadas as propriedades da matriz Hessiana, Bishop (2006) e Silva (2012).

Figura 4 – Máximos e mínimos utilizados para encontrar pontos extremos nas bordas.



Fonte: Thome (2010).

A matriz Hessiana guarda uma relação entre seus autovalores e a situação de um ponto extremo que está sendo analisado, essa dedução pode ser vista em Moroneze (2013) e a Tabela 1 mostra um resumo dessa relação.

Tabela 1 – Resumo da relação entre os autovalores de matriz Hessiana com o tipo de ponto extremo. Fonte Moroneze (2013).

Sinais dos autovalores	O ponto extremo é..
Todos positivos	Ponto de mínimo
Ao menos um positivo e um negativo	Ponto de sela
Todos negativos	Ponto de Máximo

Para que os pontos extremos localizados nas bordas sejam eliminados é calculada uma matriz Hessiana 2x2 gerada no local e na escala em que se encontra o ponto, essa matriz está posta na Equação 7.

$$H = \begin{bmatrix} D_{xx} & D_{xy} \\ D_{xy} & D_{yy} \end{bmatrix}, \quad (7)$$

$$D_{xx} = D(x + 1, y, \sigma) - 2D(x, y, \sigma) + D(x - 1, y, \sigma), \quad (8)$$

$$D_{yy} = D(x, y + 1\sigma) - 2D(x, y, \sigma) + D(x, y - 1, \sigma), \quad (9)$$

$$D_{xy} = \left[\frac{D(x-1,y+1\sigma) - D(x+1,y+1,\sigma) + D(x+1,y-1,\sigma) - D(x-1,y-1,\sigma)}{4} \right], \quad (10)$$

onde:

D_{xy} é a derivada de $D(x, y, \sigma)$ na localização e escala em relação a x e y;

D_{xx} é a derivada segunda em relação a x; e

D_{yy} é a derivada segunda em relação a y.

A diferença dos vizinhos é usada para o cálculo das derivadas como visto nas equações (8),(9) e (10). Os autovalores de H são proporcionais a curvatura de D e são calculados a partir do traço e do determinante de H:

$$Tr(H) = D_{xx} + D_{yy} = \alpha + \beta, \quad (11)$$

$$Det(H) = D_{xx}D_{yy} - (D_{xy})^2 = \alpha\beta, \quad (12)$$

onde

- α : autovalor de maior magnitude;
- β : autovalor de menor magnitude;
- $Tr(H)$: Traço da matriz Hessiana H;
- $Det(H)$: Determinante da matriz Hessiana H.

Considerando r a razão entre os autovalores:

$$r = \frac{\alpha}{\beta}. \quad (13)$$

Então:

$$\frac{Tr(H)^2}{Det(H)} = \frac{(\alpha + \beta)^2}{\alpha\beta} = \frac{(r + 1)^2}{r}. \quad (14)$$

Na Equação (14) a relação fica dependente apenas da razão entre os autovalores $\frac{(r+1)^2}{r}$, sendo mínima se os autovalores forem iguais e aumenta em função de r . Então é definido um limiar acima do qual o ponto está em uma borda e deve ser descartado. Este limiar está posto na Equação 15.

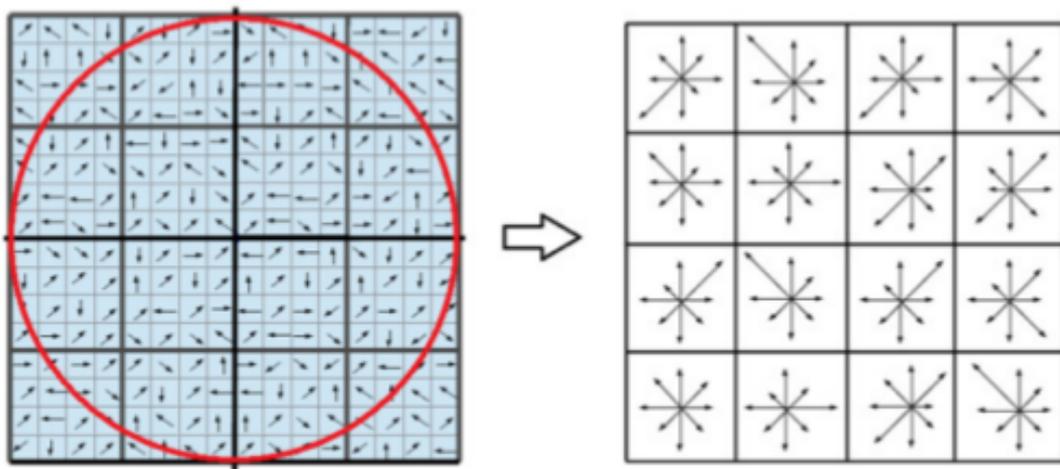
$$\frac{Tr(H)^2}{Det(H)} < \frac{(r+1)^2}{r}. \quad (15)$$

Em seu trabalho Silva (2012) sugere $r = 10$, o que elimina pontos extremos que tem razão entre as curvaturas principais maior que 10.

2.1.5 Extração dos descritores locais

Nessa fase já foram descartados os pontos extremos que não possuem as características adequadas e tem início o processo de extração dos descritores locais. O descritor local de um ponto chave é calculado em função da magnitude e orientação do gradiente de seus vizinhos.

Figura 5 – Extração do descritor de um ponto chave em função da amplitude e da orientação do gradiente de seus vizinhos, a esquerda o mapa de gradientes e a direita o descritor do ponto chave.



Fonte: Meggiolaro (2014).

Na Figura 5 o processo de extração do descritor pode ser visto com mais detalhes, o círculo em vermelho representa uma janela Gaussiana cujo nível (σ) é função da escala na oitava onde se encontra o ponto chave, esse procedimento se repete para todas as imagens das oitavas. A janela Gaussiana é utilizada para prevenir mudanças rápidas no descritor em

função de pequenas variações na posição da janela, a ideia é focar nos gradientes que estão próximos ao ponto chave. No lado direito da Figura 5 está o resultado dos histogramas das regiões 8x8 que estão no lado esquerdo da imagem, cada histograma possui 8 direções. O tamanho de cada vetor corresponde a magnitude do histograma de entrada na direção correspondente. O descritor é construído então com um vetor contendo os valores das resultantes do histograma de entrada. A Figura 5 mostra a resultante do histograma de entrada como uma matriz 4x4, este formato foi recomendado por Lowe (2004) baseado em seus experimentos com melhores resultados. Assim para cada ponto chave o vetor de característica tem a dimensão de $4 \times 4 \times 8 = 128$ elementos. O vetor característico na última fase do processo é normalizado para reduzir os efeitos da variação de luminosidade. Quando ocorre uma mudança no contraste da imagem o valor de cada pixel é multiplicado por uma constante que também será usada para multiplicar os gradientes, dessa forma a mudança no contraste será cancelada.

Uma mudança linear no brilho não terá efeito sobre o descritor pois ele é calculado pela diferença de *pixels*, porém quando a mudança é não linear e valores elevados de gradientes são criados é preciso reduzir a influência desses grandes valores, e isso é feito utilizando um limiar no vetor de características unitárias para que cada um não seja maior que 0,2, normalizando-os ao comprimento da unidade. Isto significa que combinar as magnitudes para grandes gradientes já não é tão importante, e que a distribuição de orientações tem maior ênfase. O valor de 0,2 foi determinado experimentalmente usando imagens contendo diferentes iluminações para os mesmos objetos 3D.

2.1.6 Extraíndo as características que formam os descritores de um ponto chave

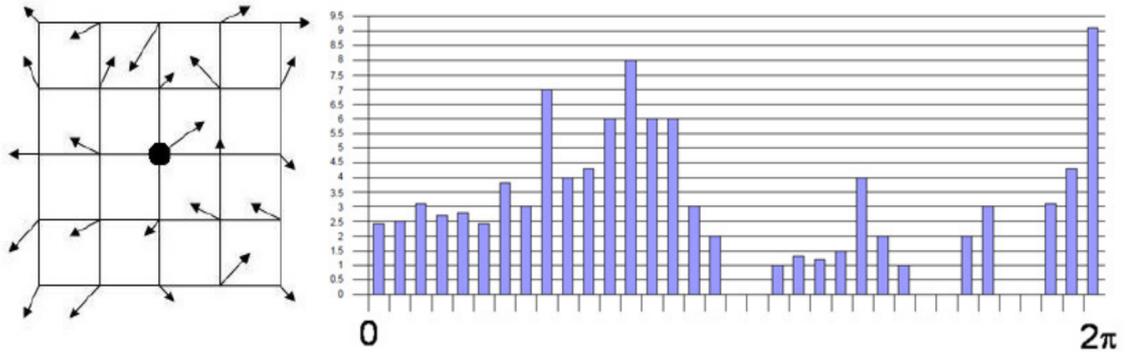
Agora serão atribuídos uma orientação e uma magnitude para cada ponto chave restante, isso é feito utilizando a magnitude e a orientação do gradiente de cada ponto para cada oitava no espaço de escalas $L(x, y, \sigma)$, como visto na Equação (1):

$$m(x, y) = \sqrt{\frac{(L(x+1, y) - L(x-1, y))^2}{(L(x, y+1) - L(x, y-1))^2}}, \quad (16)$$

$$\theta(x, y) = \tan^{-1} \frac{(L(x, y+1) - L(x, y-1))}{(L(x+1, y) - L(x-1, y))}, \quad (17)$$

onde $m(x, y)$ é a magnitude do gradiente no ponto e $\theta(x, y)$ a sua orientação, utilizando para o cálculo a diferença de *pixels*.

Figura 6 – Histograma com as orientações de um ponto chave, os pontos mais altos definem a orientação dominante.



Fonte: Meggiolaro (2014).

Com as equações (16) e (17) é criado um histograma com as orientações dos pontos em uma região vizinha ao ponto chave. Segundo Meggiolaro (2014) o histograma representa todas as orientações possíveis variando de $(0$ a $2\pi)$ e pode ser visto na Figura 6.

Cada ponto é adicionado ao histograma usando dois pesos o primeiro é dado pela valor de sua magnitude (16) e o segundo é dado por uma janela Gaussiana circular com σ' igual a 1,5 vezes maior que a escala da oitava em que se encontra o ponto chave, essa janela é definida como:

$$G(x, y, \sigma) = \frac{1}{2\pi\sigma'^2} e^{-(\Delta x^2 + \Delta y^2 / 2\sigma'^2)}. \quad (18)$$

Aonde Δx e Δy são as distâncias entre cada ponto e o ponto chave.

O valor total dos pesos para cada ponto na vizinhança (x, y) é dado por:

$$h'_\theta = h_\theta + \alpha m(x, y) \cdot g(\Delta x, \Delta y, \sigma), \quad (19)$$

com:

$$\alpha = d/i, d < i,$$

$$\alpha = 0, d > i.$$

Sendo h'_θ a atualização de h_θ e d é a distância absoluta em graus entre a orientação do ponto e o Θ discretizado, e i é o intervalo em graus entre os Θ discretizados.

Valores altos no histograma definem as direções dominantes dos gradientes locais, são considerados picos que correspondem a valores de até 80% do valor máximo. Dessa

forma o mesmo ponto chave pode ter múltiplas orientações, isso irá torna-lo mais estável durante o processo de identificação. Os três valores mais próximo do pico são interpolados utilizando uma parábola para obter uma valor mais exato da sua posição.

Agora cada ponto chave tem quatro informações:

- Id;
- Posição x;
- Posição y;
- Magnitude do gradiente;
- Orientação do gradiente.

2.1.7 Parametrização do descritor local

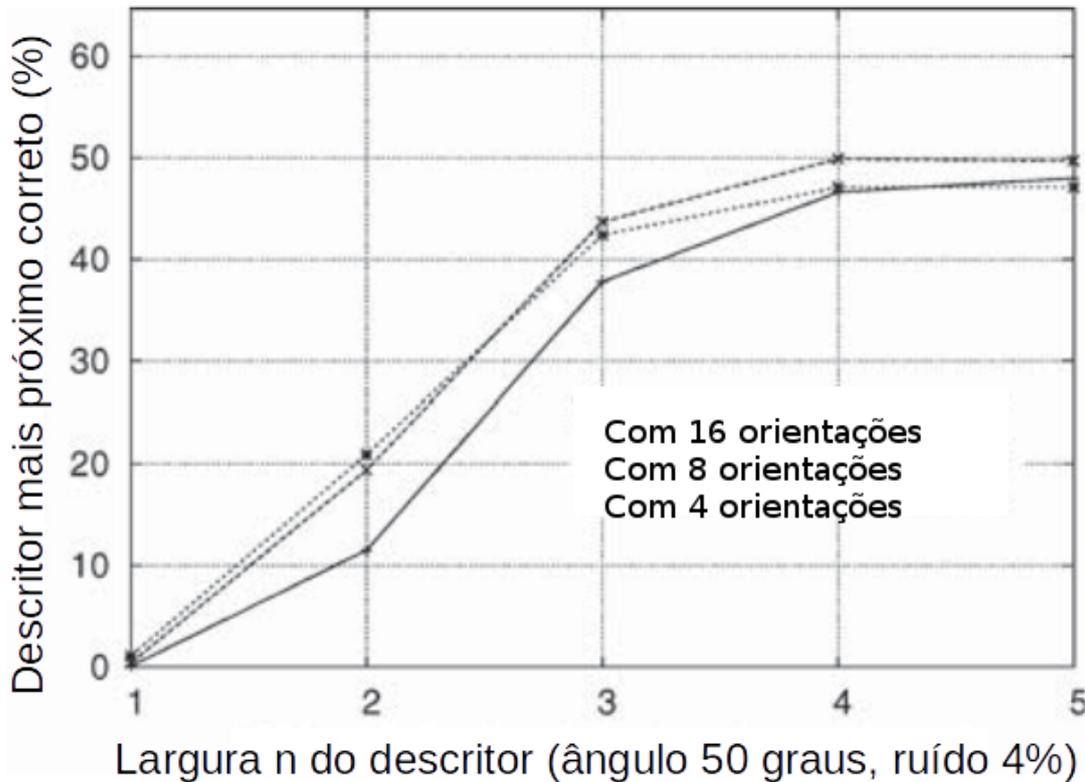
Existem dois parâmetros que podem ser usados para variar a complexidade do descritor local:

- r : O número de orientações no histograma;
- n : A dimensão do *array* de orientação dos histogramas.

No exemplo dado na Figura 5 temos que $r = 8$ e $n = 2$, de uma forma geral o descritor resultante terá o tamanho de rn^2 .

A Figura 7 mostra os resultados experimentais considerando uma amostra de 40000 pontos chaves em uma condição crítica para a realização da correspondência (rotacionados de 50° em relação a imagem original e com a adição de 4% de ruído) no eixo x temos a dimensão n da matriz de histogramas e são montadas três curvas para quantidades diferentes de orientações (4,8 e 16) os resultados da variação estão em percentual no eixo y . A medida que o tamanho do descritor aumenta ele irá discriminar melhor em um número maior de imagens no entanto ficará mais sensível a distorções de forma e a ocorrência de oclusão. No gráfico da Figura 7 o ponto ótimo ocorre com uma matriz 4x4 de histogramas com 8 orientações, a partir daí o aumento da complexidade do descritor já não é tão eficiente. Segundo Lowe (2004) este padrão se repete para outros graus de variação.

Figura 7 – Resultado encontrado com a variação no número de orientações e na dimensão da matriz de histogramas.



Fonte: Lowe (2004).

2.1.8 Correspondência entre pontos chaves

A correspondência entre pontos chaves também conhecida como *matching* é o processo de encontrar pontos semelhantes entre duas imagens.

Dadas duas imagens I_1 e I_2 com seus descritores definidos como $d1_i$ e $d2_j$, respectivamente, onde i e j são os índices dos descritores de cada imagem. Temos que,

$$d1_i = (\gamma_1, \gamma_2, \gamma_3, \dots, \gamma_{128}), \quad (20)$$

$$d2_j = (\xi_1, \xi_2, \xi_3, \dots, \xi_{128}), \quad (21)$$

onde γ e ξ são as magnitudes de cada elemento do descritor. O ponto que mais se assemelha do descritor $d1_i$ para um dado i é definido pelo descritor $d2_j$ que possui a menor distância Euclidiana em relação a $d1_i$, a Equação (22) define a função que deve ser minimizada para um determinado valor de j :

$$|d1_i - d2_j| = \sqrt{((\gamma_1 - \xi_1)^2 + (\gamma_2 - \xi_2)^2 + \dots + (\gamma_{128} - \xi_{128})^2)}. \quad (22)$$

Esse processo se repete para todo i , para que se encontre todos os pares de descritores semelhantes. Durante esse procedimento pode ocorrer a geração de falsas correspondências isso ocorre por causa de pontos chaves instáveis. Para contornar esse problema compara-se a menor distância com a segunda menor distância, definindo-se um limiar para essa relação e selecionando apenas os pontos que atendem a essa condição.

2.2 Trabalhos relacionados

Na pesquisa realizada foram encontrados cinquenta e quatro trabalhos que tem como objetivo o reconhecimento de armas de fogo, esses trabalhos foram agrupados de acordo com a técnica utilizada, e os mais significativos serão detalhados a seguir:

Abaixo as principais técnicas utilizadas no reconhecimento de armas de fogo:

- Infravermelho;
- Ondas milimétricas;
- Áudio;
- *Template Matching*;
- Segmentação e extração de características.

2.2.1 Detecção de armas de fogo utilizando infravermelho

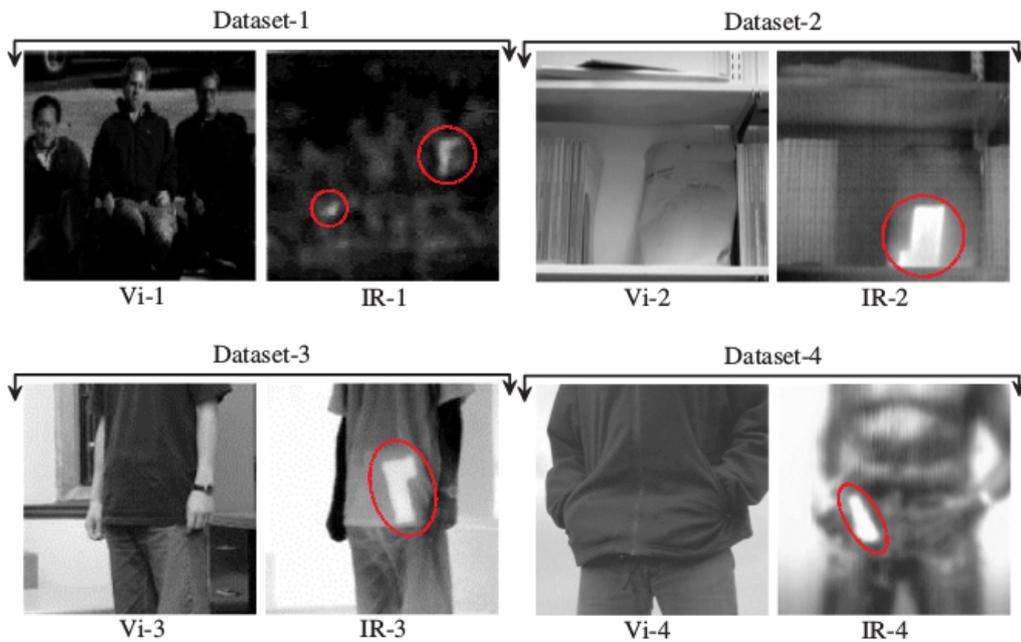
Radiação infravermelha são ondas de comprimento que variam de 1 milímetro até 700 nanômetros e não são visíveis para o olho humano. É uma radiação não ionizante sem efeitos danosos e sem riscos de causar males como o câncer. No espectro de luz está localizada depois da luz vermelha, daí surgiu seu nome. Apesar de não poder ser vista a radiação infravermelha pode ser notada no corpo em forma de calor: terminações nervosas, chamadas termorreceptores, conseguem captar essa radiação. O trabalho de Xue e Blum (2003) utiliza um sensor infravermelho como forma de identificar armas que estão sob a roupa, a abordagem é fundir a imagem gerada pelo sensor de infravermelho com a imagem RGB gerada por uma câmera, o resultado desse processo está na Figura 8.

Figura 8 – Abaixo o resultado da fusão de imagens, a esquerda a imagem normal no centro a imagem do sensor infravermelho e a direita a imagem fundida.



Fonte: Xue e Blum (2003).

Figura 9 – O resultado da fusão de imagens a esquerda a imagem normal e a direita a com infravermelho.



Fonte: Bhatnagar e Wu (2011).

Em Bhatnagar e Wu (2011) é utilizada uma abordagem semelhante, e o resultado pode ser visto na Figura 9. Seguindo essa mesma abordagem temos ainda os trabalhos de Upadhyay e Rana (2014), Xu e Wu (2015), Hussein e Hu (2016), Chen et al. (2005). O foco destes trabalhos é auxiliar os operadores de sistemas de segurança na localização de armas, o processo não é automático e depende da atenção do operador. Nesse aspecto a nossa proposta seria uma melhoria para esse sistema pois a imagem fundida poderia ser analisada pelo algoritmo de identificação da armas e o alerta seria emitido sem a necessidade da intervenção de um operador.

2.2.2 Detecção de armas de fogo utilizando ondas milimétricas

Ondas milimétricas estão situadas no espectro eletromagnético entre as ondas de rádio e infravermelha e possuem um propriedade única de passar por entre materiais leves como roupas. Essa propriedade é utilizada nos *scanners* instalados em aeroportos para a detecção de armas e outros objetos metálicos, também são utilizadas em sistemas de alarmes contra intrusos e no auxílio de motoristas quando estacionam. O foco dos trabalhos a seguir é a análise da imagem gerada de uma pessoa a partir de sua exposição a estas ondas.

Em Kim et al. (2011) foi utilizado o *Millimeter-wave Imaging Radiometer Equipment* (MIRAE) que está na Figura 10 para gerar ondas milimétricas na frequência de 94 GHz e que tem acoplado um módulo de recepção com 30 canais. Este módulo converte diretamente o sinal recebido, em impulsos elétricos que são posteriormente convertidos em uma imagem na escala de cinza. Essa imagem está na Figura 11.

Figura 10 – Equipamento MIRAE utilizado para a detecção de armas a curta distância.



Fonte: Kim et al. (2011).

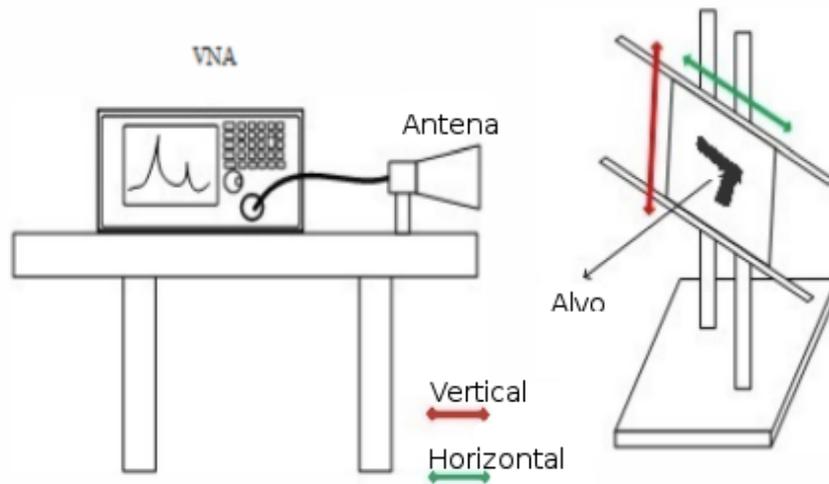
Figura 11 – A esquerda a imagem natural e a direita a imagem em escala de cinza gerada pela recepção dos dados do MIRAE.



Fonte: Kim et al. (2011).

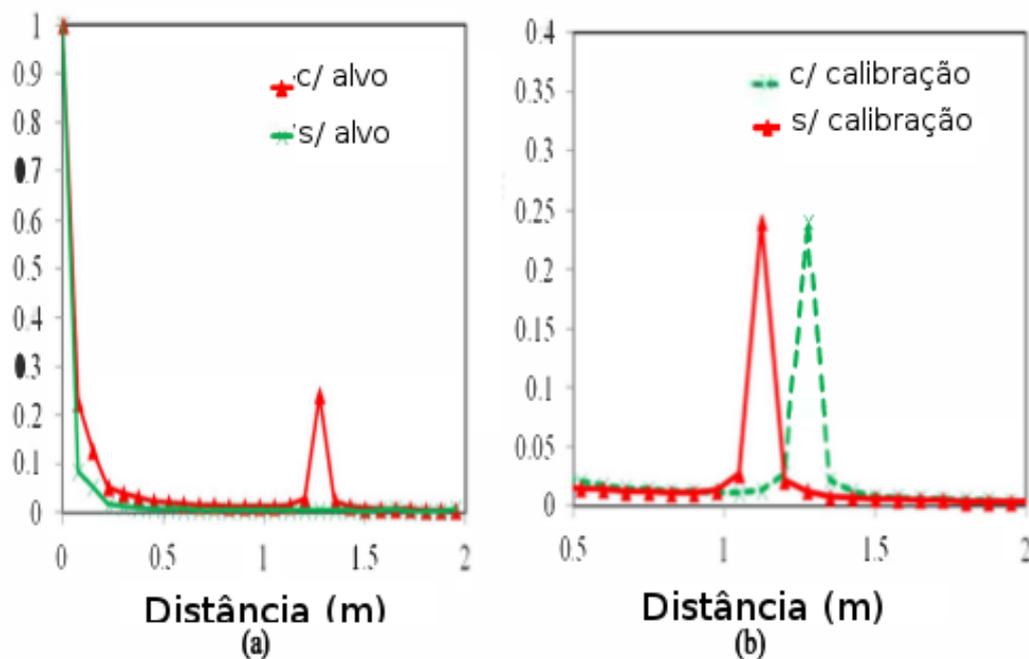
O trabalho de Kumar, Upadhyay e Singh (2015) também utiliza ondas milimétricas para a detecção de armas, na Figura 12 o *setup* que foi montado para este processo. Um avanço em relação aos outros trabalhos é que essa abordagem consegue medir a distância em que se encontra a arma e a sua forma.

Figura 12 – *Setup* experimental montado para a realização dos testes, considerando a distância entre a arma e a antena.



Fonte: Kumar, Upadhyay e Singh (2015).

Figura 13 – Nos gráficos abaixo a faixa de distância em que a arma foi detectada, a esquerda sem a correção e a direita com a correção no posicionamento.

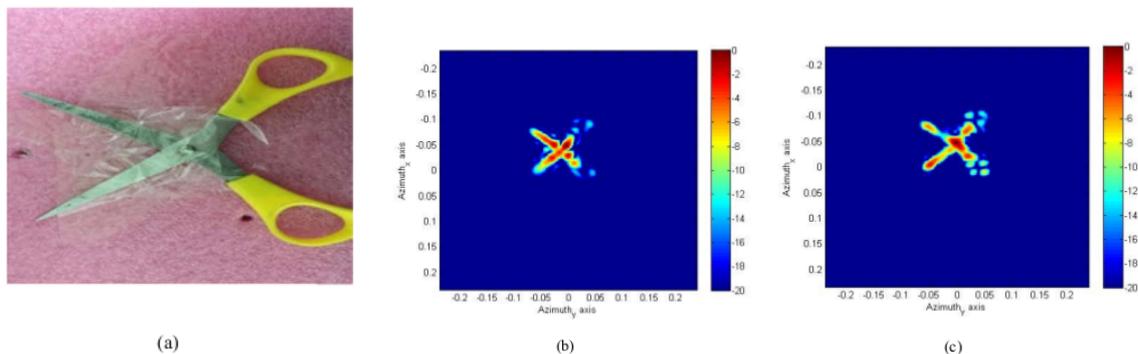


Fonte: Kumar, Upadhyay e Singh (2015).

Na Figura 13 está a forma de como se mede a distância da arma em função do pico detectado, que no caso corresponde a reflexão do metal. Para aumentar a precisão foi introduzido um fator de correção. Na imagem (a) a curva verde representa o sinal sem a presença da arma, não ocorre o pico nessa situação, a curva vermelha representa o sinal com a presença da arma, o pico foi detectado a uma distância de pouco mais de 1 metro. Na imagem (b) é aplicado o fator de correção para chegar aos exatos 1.12 metros de distância.

No trabalho de Li et al. (2016a) foi utilizada a técnica de multi polarização das antenas com o objetivo de obter mais informações sobre o objeto detectado.

Figura 14 – Resultado da técnica de multi polarização, (a) imagem original, (b) Imagem com antenas de polaridades complementares e em (c) antenas com polaridades cruzadas.



Fonte: Li et al. (2016a).

Na Figura 14 os resultados obtidos com essa técnica. Na conclusão do trabalho o autor reconhece a baixa eficiência dessa abordagem e sugere mudanças no tipo de antena.

Ainda seguindo no uso de ondas milimétricas para a identificação de armas temos os trabalhos de Kapilevich et al. (2013) e Moulder et al. (2016).

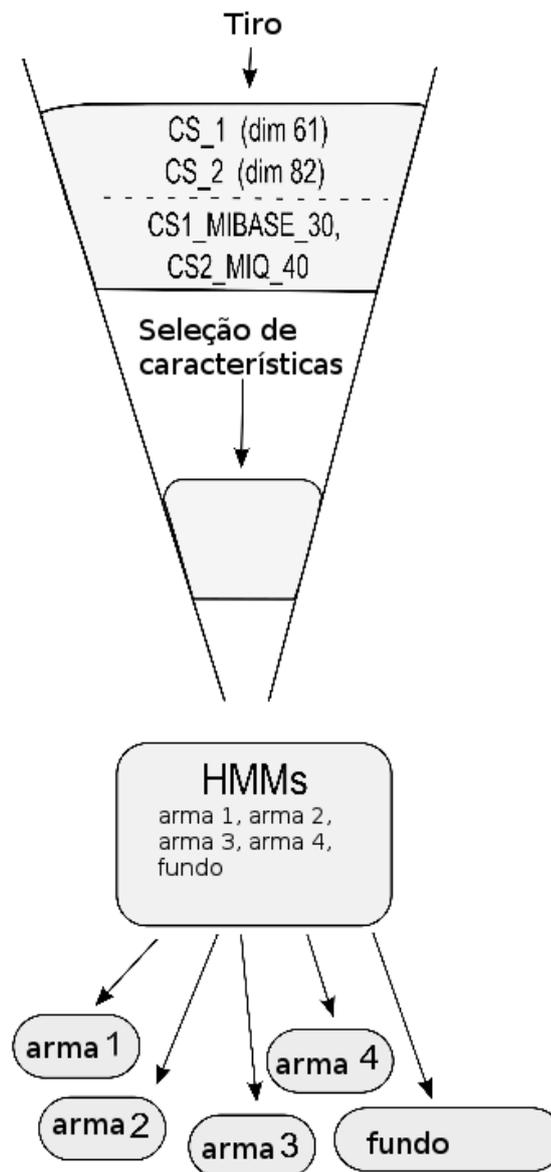
Em todos os trabalhos vistos não foi feita a fusão das imagens e nem o reconhecimento automático da arma o que seria um ponto de melhoria, outro aspecto é a distância de detecção que é bem limitada além dos custos elevados envolvidos para a aquisição de equipamento e sua manutenção.

2.2.3 Detecção de armas de fogo utilizando áudio

No trabalho de Kiktova et al. (2015) a identificação da arma de fogo é feita utilizando o áudio gerado com o seu disparo. Na Figura 15 o fluxo desse sistema, CS.1 e CS.2 são

dois conjuntos de características com dimensões de 61 e 82 elementos respectivamente, e CS1_MIBASE e CS2_MIQ são as mesmas características porém com dimensões reduzidas para 30 e 40 elementos respectivamente e formam a base de dados. Em seguida é utilizado o modelo oculto de Markov para estimar a que arma pertence o áudio.

Figura 15 – O sistema utiliza Modelo Oculto de Markov para identificar a partir das características do áudio o tipo de arma.



Fonte: Kiktova et al. (2015).

Figura 16 – Durante o procedimento experimental foram utilizadas quatro armas para validação.



Fonte: Kiktova et al. (2015).

Na Figura 16 os quatro modelos de armas utilizados nos experimentos, da esquerda para a direita, Pi vz.82, 1911, Ruger LCP e uma pistola sinalizadora. Os resultados alcançados foram de 80.61% de acertos no reconhecimento da arma.

A identificação da ocorrência de um disparo de arma de fogo é uma ação importante e a identificação do tipo de arma seria o ideal. No entanto existem fatores externos que podem minar estas duas abordagens, o ruído inerente a espaços abertos, a distância que ocorreu o disparo e até a forma como isso foi feito podem fazer variar bastante os resultados encontrados. Além disso essa ação de reconhecimento é posterior ao fato, ou seja, a arma já foi usada. O que elimina a possibilidade de uma ação preventiva. Seguindo essa abordagem temos os trabalhos de Ahmed, Uppal e Muhammad (2013), Vozáriková et al. (2011).

2.2.4 Detecção de armas de fogo utilizando *template matching*

No trabalho de Hung e Pan (2015) a abordagem utilizada foi analisar o comportamento das pessoas, e a partir daí identificar a presença de uma arma de fogo no ambiente. A hipótese adotada foi que se uma pessoa está com seus dois braços levantados, então ela está sob ameaça de uma arma.

Figura 17 – Na Figura o processo de segmentação da imagem em (a) o momento anterior ao evento em (b) o evento e em (c) a segmentação da região de interesse.



Fonte: Hung e Pan (2015)

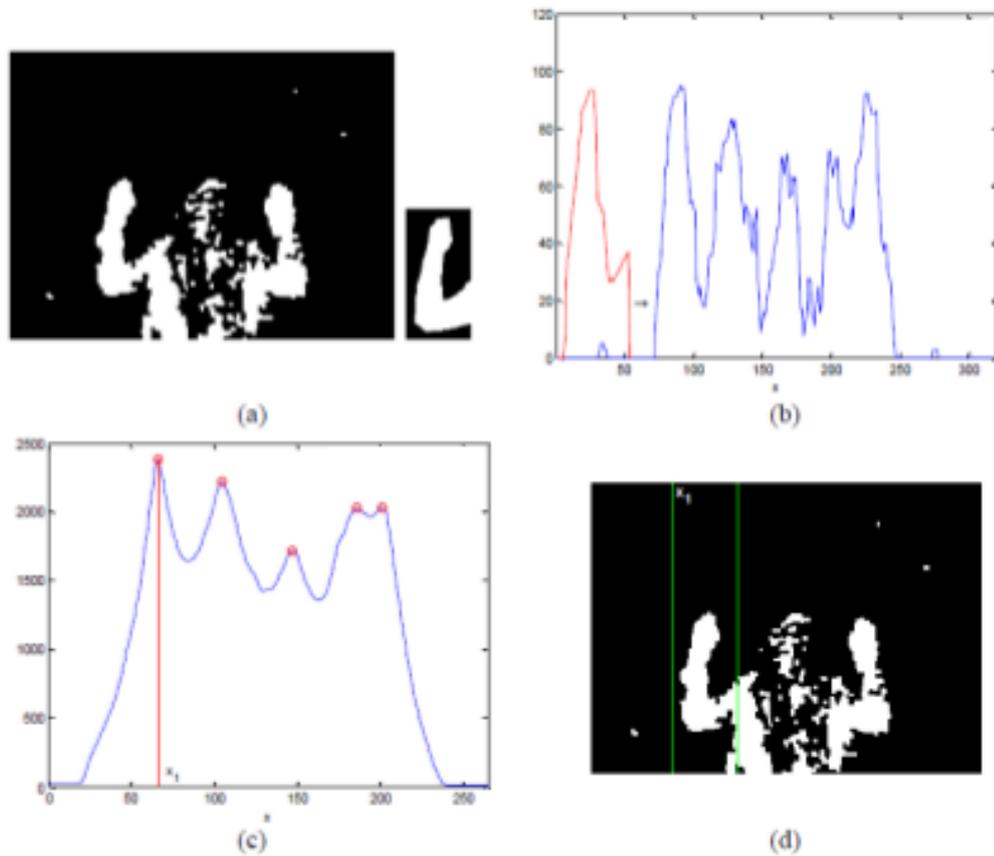
Figura 18 – Templates utilizados no processo de busca de pessoas com braços levantados na imagem capturada.



Fonte: Hung e Pan (2015).

Na Figura 17 o processo de retirada do fundo e segmentação da pessoa com as mãos levantadas pode ser visto. Nesse ponto é aplicado a técnica de *Template Matching*, utilizando padrões de braços levantados conforme visto na Figura 18.

Figura 19 – Resultado do *matching* entre a imagem capturada e segmentada e o *template*.



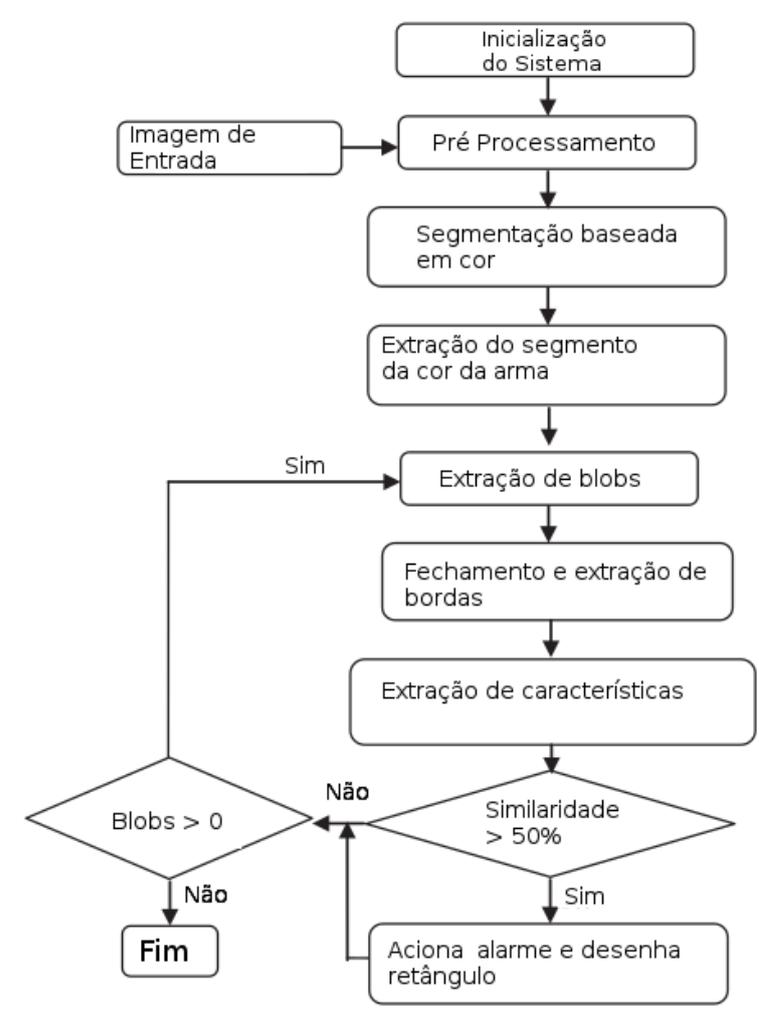
Fonte: Hung e Pan (2015).

O resultado deste processo está na Figura 19. Aonde está posto o casamento entre o *template* do braço direito erguido e a imagem da pessoa com ambos os braços erguidos. Esse trabalho obteve uma taxa de precisão de 89%. Esta abordagem é interessante porém depende da reação das pessoas quando expostas ao evento de um assalto o que pode comprometer o resultado. Essa mesma abordagem já foi utilizada em Dever, Lobo e Shah (2002) sendo nesse caso foi utilizado o processo de extração de silhueta e segmentação do esqueleto para identificar os braços levantados.

2.2.5 Detecção de armas de fogo utilizando segmentação e extração de características

O trabalho de Tiwari e Verma (2015a) utiliza uma abordagem com segmentação de cor e detector de pontos de interesse Harris usando o extrator de característica FREAK para detectar armas de fogo em uma imagem, na Figura 20 o fluxo desse sistema.

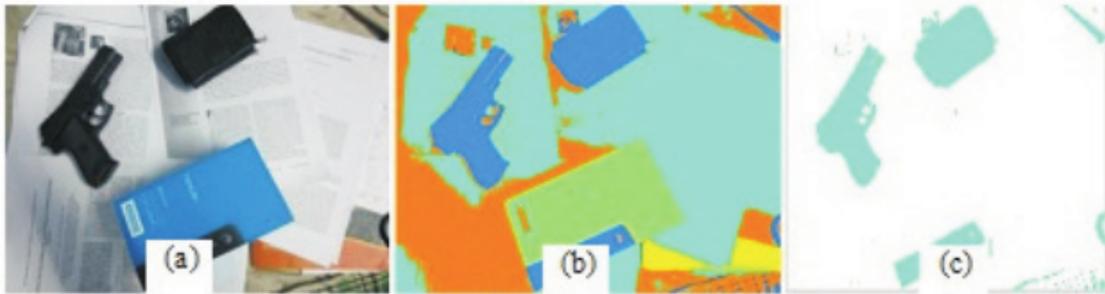
Figura 20 – Fluxograma do método proposto por Tiwari e Verma (2015a).



Fonte: Tiwari e Verma (2015a).

Na inicialização do sistema são carregados os descritores das armas previamente extraídos e armazenados e em seguida é inserida a imagem de entrada que sofre um pré-processamento e é segmentada pela cor preta. Os objetos resultantes da segmentação tem suas bordas fechadas e extraídas para então passar pelo processo de detecção de pontos de interesse usando o método Harris. Em seguida ocorre a extração de suas características utilizando o método FREAK e após esse ponto ocorre o *matching* com os dados lidos no início do processo e se ocorrer uma correspondência maior que 50% assume-se que a arma foi encontrada e é criado um retângulo verde sobre a área na imagem.

Figura 21 – Segmentação da cor preta sendo aplicada na imagem.



Fonte: Tiwari e Verma (2015a).

Na Figura 21 o processo de segmentação por cor o método assume que as armas de fogo tem a cor preta.

Figura 22 – Resultados obtidos com a) Diferentes fundos; b) Variação na iluminação; c) Armas diferentes; d) Oclusão; e) Variação na escala e rotação; f) Múltiplas armas.



Fonte: Tiwari e Verma (2015a).

Na Figura 22 os resultados obtidos em diferentes situações, na Tabela 2 os valores alcançados para cada item.

Tabela 2 – Resultados alcançados nos itens da Figura 22.

Item	Imagens Positivas	Imagens corretamente classificadas	Taxa de acerto (%)
a	12	11	91.66
b	9	7	77.77
c	11	9	81.81
d	17	14	82.35
e	10	8	80.00
f	6	5	83.33

Esse método proposto por Tiwari e Verma (2015a) tem uma acuracidade de 84.26%, e tem alguns pontos de melhoria como o processo de segmentação por cor que limita a detecção de armas da cor preta, outro resultado na Tabela 2 que pode ser melhorado é a taxa de acerto quando ocorre uma variação na luminosidade. Seguindo essa mesma abordagem temos os trabalhos de Tiwari e Verma (2015b), Zhang e Miao (2014), Ondroušek, Kolomazník e Vytečka (2014). Na Figura 23 um resumo dos trabalhos relacionados mostrados anteriormente.

Figura 23 – Resumo dos trabalhos relacionados.

Autores	Técnica	Observações
(Xue e Blum (2003), Upadhyay e Rana (2014), Xu e Wu (2015), Hussein e Hu (2016), Chen et al. (2005))	Infravermelho	- Operação manual. - Uso de sensores. - Fusão de imagens.
(Kim et al. (2011), Kumar, Upadhyay e Singh (2015), Li et al. (2016a), Kapilevich et al. (2013), Moulder et al. (2016))	Ondas milimétricas	- Uso de equipamentos dedicados. - Operação manual. - Curta distância.
(Kiktova et al. (2015), Ahmed, Uppal e Muhammad (2013), Vozáriková et al. (2011))	Áudio	- Depende de fatores externos. - Ruído. - Distância. - Ação de reconhecimento é posterior ao fato.
(Hung e Pan (2015), Dever, Lobo e Shah (2002))	Template Matching	- Depende da reação das pessoas. - Gestos semelhantes podem confundir.
Tiwari e Verma (2015b), Zhang e Miao (2014), Ondroušek, Kolomazník e Vytečka (2014)	Segmentação e extração de características	- Segmentação por cor, detecção de armas da cor preta.

Fonte: O Autor.

3 RECONHECIMENTO DE ARMA DE FOGO

Neste capítulo será apresentado o método desenvolvido para realizar o reconhecimento de armas de fogo no interior de veículos. Trata-se de uma abordagem desenvolvida em duas fases:

1^a Fase: É a criação e otimização da base de descritores locais. Essa fase começa com a geração de um grupo de imagens de armas de fogo que servirão de modelo, delas serão extraídos seus descritores locais que irão formar a base original de descritores locais. No segundo momento dessa fase a base original de descritores locais será otimizada, com o objetivo de aumentar a performance do método na fase seguinte.

2^a Fase: É o processo de monitoramento do ambiente interno do veículo em busca da presença de armas de fogo. Nesse momento os descritores locais extraídos das imagens do interior do veículo serão comparados com os descritores da base otimizada gerada na fase anterior. O objetivo dessa comparação é encontrar similaridades entre os descritores locais que indiquem a presença de armas no interior do veículo.

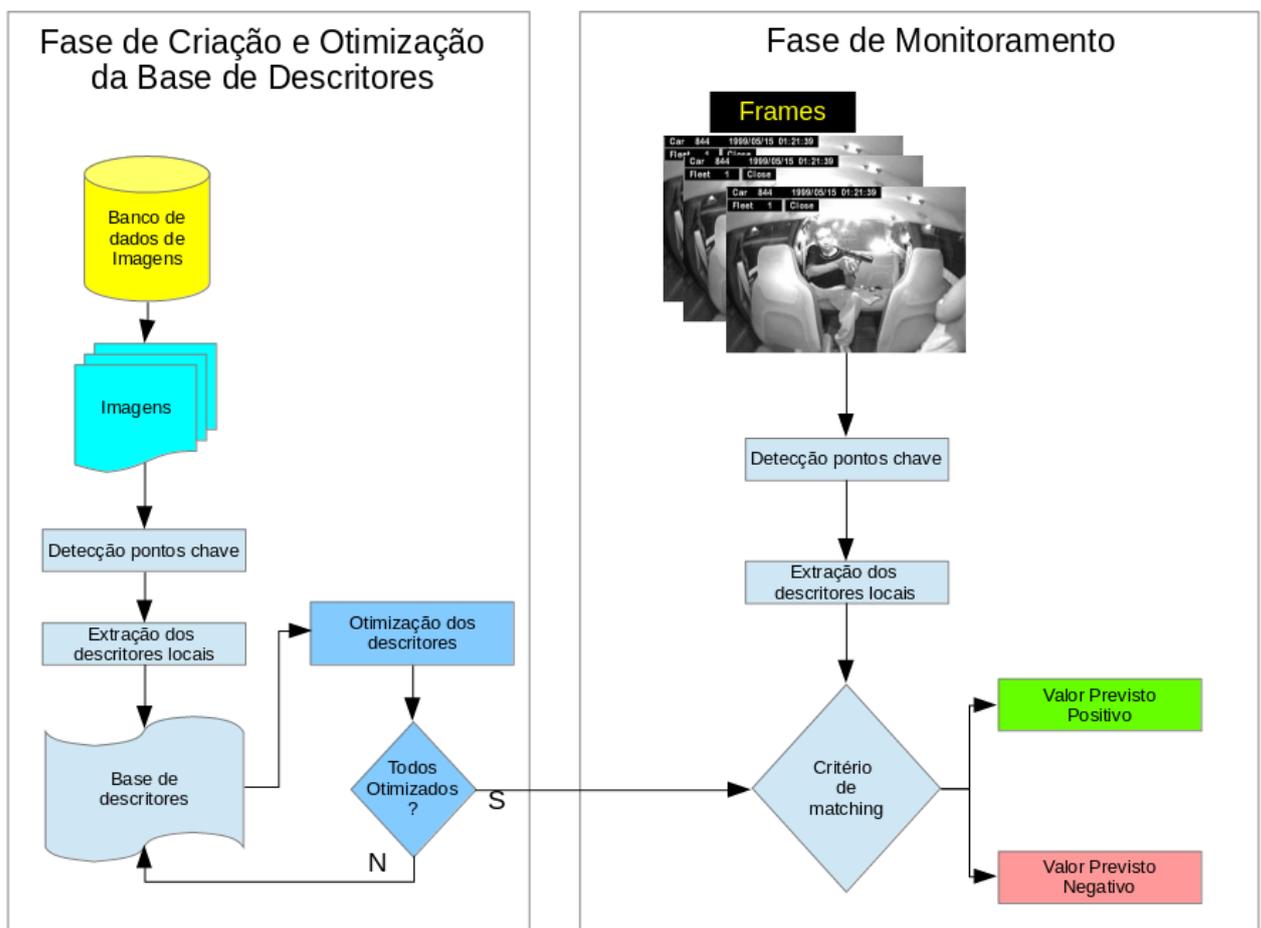
A Figura 24 ilustra esquematicamente a estrutura do método, que será detalhada nas seções seguintes.

3.1 Fase de criação e otimização da base de descritores locais

Essa fase começa com a captura de um grupo de imagens de armas de fogo. Cada imagem passa primeiramente pelo processo de detecção de pontos chave, que está apresentado no fluxograma da Figura 25.

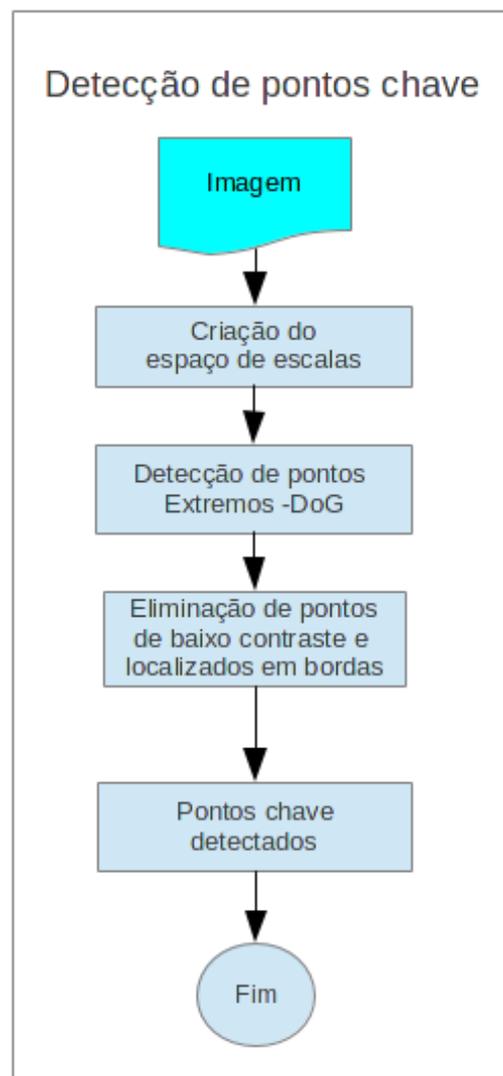
O primeiro passo é a criação do espaço de escalas, conforme visto anteriormente no Capítulo 2, na Figura 26 está uma imagem do banco de dados de imagem (IMFDB, 2017) com o detalhe da criação de uma oitava pertencente a um espaço de escala. O fluxo começa na imagem original que é submetida a um filtro gaussiano com $\sigma = 5$, em seguida a imagem resultante que se posiciona acima da original também é submetida a um filtro gaussiano, como visto na Equação (2) com $\sigma = 5$ e o processo se repete até a última imagem da oitava.

Figura 24 – Método utilizado para validar o uso de descritores locais como ferramenta principal para a identificação de armas de fogo no interior de veículos.



Fonte: O Autor

Figura 25 – Fluxograma com o processo de detecção de pontos chave.



Fonte: O Autor.

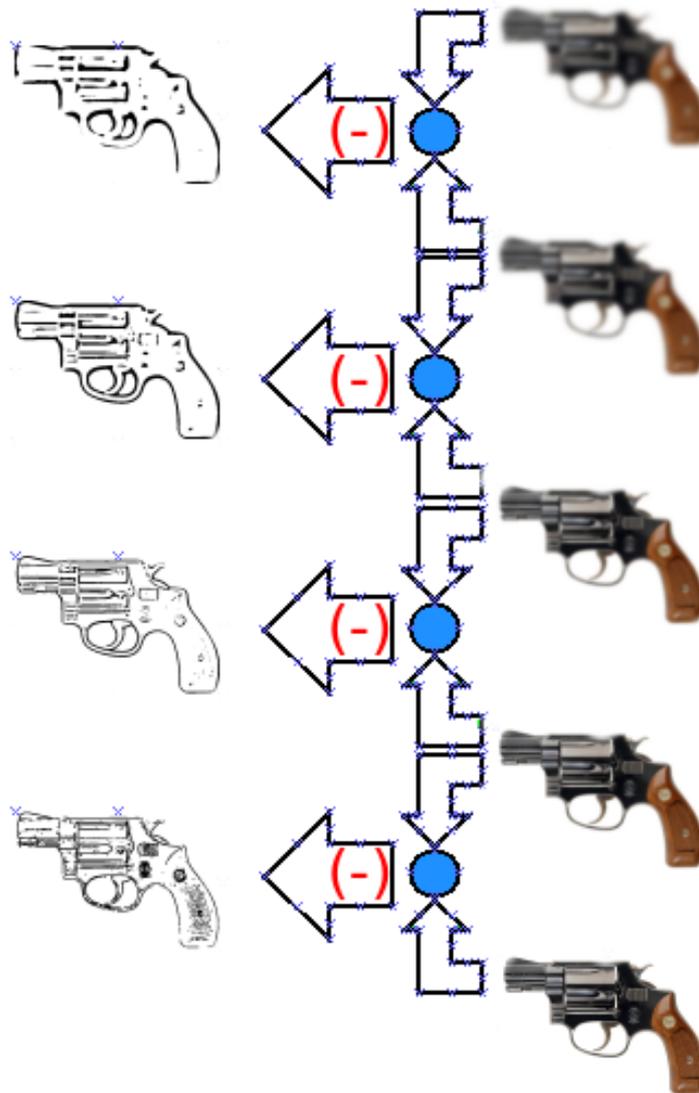
Figura 26 – Exemplo de um espaço de escala com uma oitava formada por uma imagem original e quatro imagens geradas a partir da aplicação do filtro gaussiano para $\sigma = 5$, a sequência vai da esquerda para a direita e de cima para baixo.



Fonte: O autor.

Após a criação do espaço de escalas tem início o processo de detecção de pontos extremos (máximos e mínimos). Para isso é utilizada a DoG (Diferença de Gaussianas), o processo de formação da DoG está na Figura 27. No lado direito está uma oitava pertencente ao espaço de escala, com um fluxo que parte da imagem original que está em baixo e segue com a aplicação do filtro gaussiano com $\sigma = 5$ nas imagens acima. No lado esquerdo da Figura 27 está a DoG com as quatro imagens que foram geradas a partir da subtração das imagens do espaço de escalas.

Figura 27 – Cálculo da DoG utilizando a oitava apresentada anteriormente.



Fonte: O Autor.

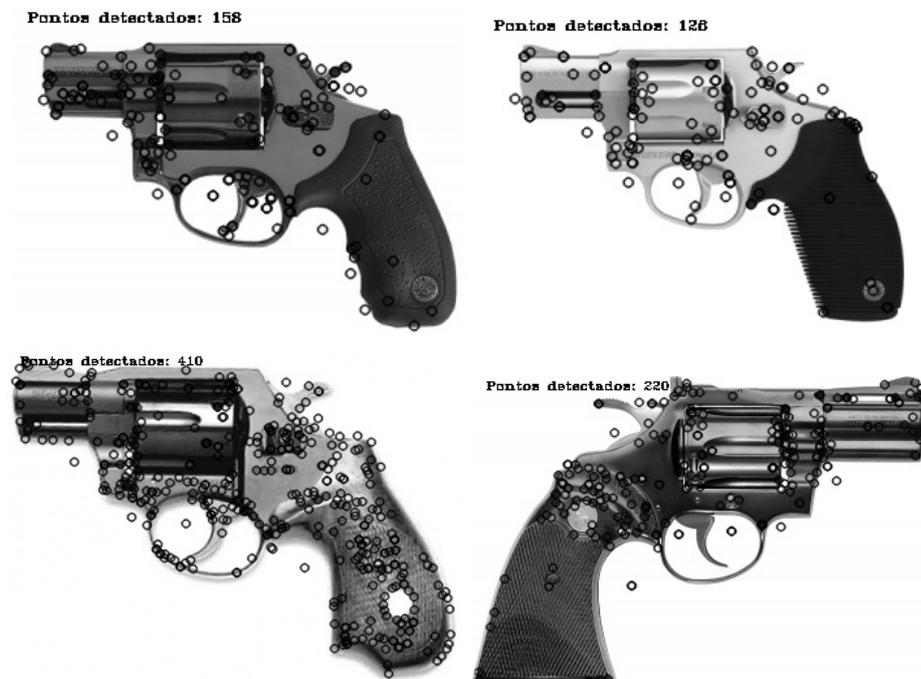
As imagens resultantes do processo de DoG tem como característica evidenciar os pontos de máximo e mínimo na imagem, como isso foram detectados vários pontos

extremos. Porém nem todos vão ser considerados, é preciso eliminar os pontos que tem baixo contraste ou que estejam localizados em bordas. A Equação 6 é utilizada para descartar extremos instáveis com baixo contraste e é calculada de forma a eliminar pontos que tenham valor de $|D(\hat{x})| < 0.03$.

Agora falta apenas eliminar os pontos extremos que se localizam sobre bordas, isso é feito utilizando as propriedades da matriz Hessiana. Para isso foi utilizada a Equação 15, sendo que em seu trabalho Silva (2012) sugere o valor de $r=10$ para que se eliminem pontos extremos que tem razão entre as curvaturas principais maior que 10. Isso é suficiente para o descarte dos pontos que estão sobre bordas. A partir desse ponto restaram apenas os pontos que estão em condições aceitáveis para a extração dos descritores locais, que passam a se chamar pontos chave.

Na Figura 28 está o resultado do processo mostrado na Figura 25 na detecção de pontos chave em algumas imagens da base de dados IMFDB (2017).

Figura 28 – Detecção de pontos chave em seis imagens de armas da base de dados IMFDB.



Fonte: O Autor.

Após a eliminação dos pontos inadequados restam os pontos chave definitivos e para cada um deles são extraídas as informações abaixo:

1. Id do ponto: É um número inteiro que vai de 0 até a quantidade total de pontos chave detectados.

2. Coordenada x: Valor da coordenada x do ponto chave na imagem;
3. Coordenada y: Valor da coordenada y do ponto chave na imagem;
4. Magnitude ou Peso: Que é a magnitude do gradiente no ponto chave; definida pela Equação (16);
5. Angulo: É a orientação do gradiente no ponto chave dada pela Equação (17).

Figura 29 – Na imagem foram encontrados 151 pontos chave, logo abaixo uma lista parcial destes pontos com suas características.



Ponto	Pos X	Pos Y	Peso	Angulo
0	9.07076	9.60436	0.050872	212.078
1	187.859	13.6339	0.0197721	259.339
2	11.7001	13.9173	0.032714	269.755
3	243.49	13.8767	0.030118	132.059
4	281.028	14.6891	0.0464308	30.9603
5	189.958	16.762	0.0174225	187.322
6	80.4129	17.0241	0.0211527	270.661
7	84.2063	17.0174	0.0189015	270.133
8	153.496	17.4947	0.0219189	269.919
9	234.011	36.101	0.0278442	230.043
10	79.4718	37.1764	0.0668116	167.439
.
.
.
140	276.68	37.8287	0.0593027	323.143
141	261.124	86.8795	0.0636363	83.7625
142	243.892	90.1623	0.0686942	128.463
143	243.892	90.1623	0.0686942	90.7361
144	261.017	26.6018	0.0640261	356.45
145	261.017	26.6018	0.0640261	291.307
146	286.969	42.4189	0.0605286	296.089
147	278.854	112.488	0.0586834	151.91
148	278.854	112.488	0.0586834	110.558
149	299.003	81.6379	0.0709793	307.69
150	229.386	49.6914	0.0711299	118.322

Fonte: O Autor.

Na Figura 29 está um exemplo do resultado alcançado com o fluxo da Figura 25, foram detectados 151 pontos chave na imagem e logo abaixo a listagem parcial destes pontos com suas respectivas informações.

Seguindo o fluxo da Figura 25 está finalizado o processo de detecção de pontos chave. E voltando para a Figura 24 será iniciado o processo de extração dos descritores locais. Estes descritores são calculados em regiões ao redor de pontos chave, como estamos

seguindo o modelo SIFT vamos utilizar o padrão de descritores mostrado na Figura 5. Cada ponto chave está associado a 128 regiões vizinhas e em cada região são aplicadas as Equações (16) e (17) para se encontrar um vetor cujo valor é a magnitude do gradiente, e a direção é a orientação do gradiente. Estes dois valores são associados em um histograma que está na Figura 6, cada ponto é associado ao histograma utilizando dois pesos o primeiro é dado pela sua magnitude e o segundo é dado por uma janela Gaussiana circular que esta representada na Figura 6 por um circulo vermelho. O valor final desse cálculo está definido na Equação (19)

O resultado desse processo é que cada ponto chave mostrado na Figura 29 está associado um vetor com 128 posições. Na Figura 30 está o resultado com a listagem parcial desses vetores extraídos da imagem da Figura 29, que são os próprios descritores locais.

Figura 30 – Na Figura 29 foram encontrados 151 pontos chave, abaixo uma listagem parcial dos 128 descritores locais extraídos de cada ponto chave.

```
Ponto Pos X Pos Y Peso Angulo
0 9.07076 9.60436 0.050872 212.078
[0, 0, 1, 1, 0, 0, 0, 44, 134, 5, 1, 2, 1, 0, 0, 41, 133, 107, 8,
0, 0, 0, 0, 15, 145, 96, 8, 0, 0, 0, 0, 1, 48, 13, 8, 35, 47,
0, 0, 19, 53, 145, 18, 15, 7, 0, 0, 14, 145, 33, 1, 0, 0, 0, 0,
45, 145, 8, 0, 0, 0, 0, 5, 53, 43, 14, 33, 49, 0, 0, 0, 11,
145, 47, 4, 1, 0, 0, 0, 43, 123, 10, 0, 0, 0, 0, 2, 28, 1, 0, 0,
0, 0, 0, 1, 2, 26, 0, 2, 3, 0, 0, 9, 83, 145, 1, 0, 0, 0, 0, 0,
133, 85, 2, 0, 0, 0, 0, 19, 1, 2, 0, 0, 0, 0, 0, 0, 0]

1 187.859 13.6339 0.0197721 259.339
[14, 30, 9, 2, 0, 0, 0, 25, 12, 6, 13, 3, 0, 0, 1, 141, 11, 1,
5, 5, 0, 0, 8, 141, 4, 2, 11, 15, 0, 0, 11, 29, 49, 8, 5, 0, 0, 0,
0, 51, 20, 19, 33, 0, 0, 0, 1, 141, 15, 2, 9, 11, 0, 0, 8, 141,
2, 1, 14, 17, 0, 0, 36, 6, 92, 98, 4, 0, 1, 1, 0, 94, 69, 81, 14,
0, 0, 0, 3, 141, 22, 4, 9, 9, 0, 0, 16, 141, 1, 1, 8, 11, 0, 0,
71, 0, 22, 51, 4, 5, 64, 53, 0, 141, 84, 41, 1, 2, 19, 14, 25, 141,
29, 2, 7, 2, 1, 0, 24, 141, 1, 1, 3, 1, 0, 0, 72]

.
.
.
147 278.854 112.488 0.0586834 151.91
[1, 2, 20, 5, 5, 3, 0, 0, 70, 12, 43, 12, 2, 1, 0, 32, 39, 0, 0, 0,
0, 0, 0, 31, 0, 0, 0, 0, 0, 0, 0, 12, 1, 4, 3, 121, 59, 0, 1, 136,
38, 7, 3, 34, 7, 0, 66, 113, 33, 12, 0, 0, 0, 0, 47, 1, 6, 7, 0, 0,
0, 0, 0, 5, 2, 0, 0, 136, 134, 0, 0, 112, 136, 7, 0, 50, 72, 0, 2,
32, 136, 41, 0, 0, 0, 0, 1, 11, 65, 16, 0, 0, 0, 0, 0, 0, 0, 0, 8,
83, 108, 0, 0, 3, 35, 2, 9, 131, 107, 0, 0, 17, 136, 13, 5, 7, 4,
0, 0, 62, 136, 1, 0, 0, 0, 0]

148 278.854 112.488 0.0586834 110.558
[0, 0, 0, 55, 40, 0, 0, 0, 3, 14, 7, 72, 50, 0, 3, 24, 13, 50, 19, 3, 1,
0, 21, 77, 0, 1, 1, 0, 0, 0, 4, 9, 5, 0, 0, 144, 144, 1, 0, 3, 96, 3, 0,
87, 48, 0, 11, 144, 31, 4, 0, 1, 0, 0, 79, 144, 0, 0, 0, 0, 0, 16, 24,
18, 1, 4, 126, 144, 1, 0, 2, 144, 39, 0, 8, 30, 1, 0, 47, 115, 53, 0, 0,
0, 0, 3, 34, 4, 5, 0, 0, 0, 0, 1, 2, 14, 5, 11, 82, 60, 5, 0, 1, 144, 27,
4, 4, 3, 0, 0, 10, 127, 22, 0, 0, 0, 0, 0, 16, 13, 3, 0, 0, 0, 0, 0, 3]

149 299.003 81.6379 0.0709793 307.69
[0, 0, 0, 10, 145, 42, 0, 0, 42, 20, 2, 7, 145, 77, 4, 17, 145, 28, 0, 0,
1, 2, 1, 78, 18, 0, 0, 0, 0, 0, 19, 0, 0, 0, 4, 103, 103, 0, 0, 68, 5,
0, 9, 145, 145, 2, 13, 145, 30, 0, 0, 12, 6, 1, 75, 49, 4, 0, 0, 0, 0, 0,
22, 0, 0, 0, 85, 48, 16, 0, 0, 57, 5, 1, 145, 145, 24, 0, 8, 145, 15, 0, 5,
13, 1, 0, 22, 21, 2, 0, 0, 0, 0, 0, 0, 0, 0, 21, 4, 0, 0, 0, 6, 2, 0,
44, 23, 5, 0, 0, 9, 2, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0]

150 229.386 49.6914 0.0711299 118.322
[0, 0, 0, 2, 5, 0, 0, 0, 8, 1, 0, 16, 110, 21, 2, 12, 67, 3, 0, 3, 25, 10,
13, 154, 5, 0, 0, 0, 0, 1, 14, 0, 0, 0, 23, 13, 0, 0, 0, 49, 18, 5, 168,
168, 10, 1, 8, 168, 79, 4, 31, 43, 4, 2, 80, 66, 7, 0, 0, 0, 0, 14, 0,
0, 0, 3, 0, 0, 0, 45, 4, 1, 58, 49, 13, 18, 35, 168, 42, 1, 6, 10, 6, 12,
114, 83, 12, 0, 0, 0, 0, 13, 0, 0, 0, 0, 0, 0, 0, 0, 11, 1, 0, 0, 0, 0,
3, 8, 168, 107, 0, 0, 0, 2, 16, 79, 65, 0, 0, 0, 0, 0, 0]
```

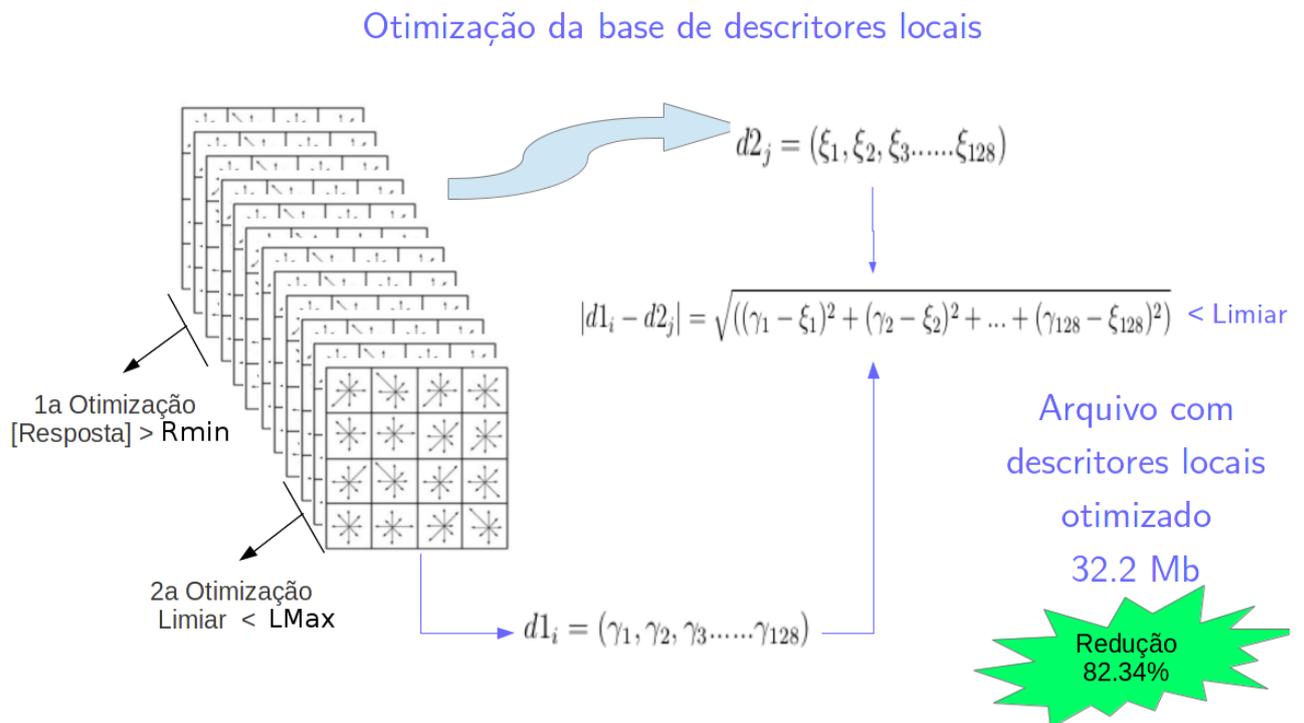
Fonte: O Autor.

O primeiro procedimento é a utilização do parâmetro [Resposta], que é gerado durante a fase de extração das características e informa a qualidade do ponto em relação a instabilidade de contraste e proximidade de bordas. Quanto maior esse valor melhor será esse ponto.

O segundo procedimento é a eliminação de descritores que são semelhantes, isto ocorre por que algumas armas tem partes que são parecidas e irão gerar também descritores com valores parecidos.

Com esses dois procedimentos ocorreu uma redução de 82.34% no tamanho da base de descritores sem perda na qualidade da informação, o novo arquivo está com o tamanho de 32.2 Mb.

Figura 32 – Processo de otimização da base de descritores utilizando a diferença Euclidiana e a qualidade do ponto chave.



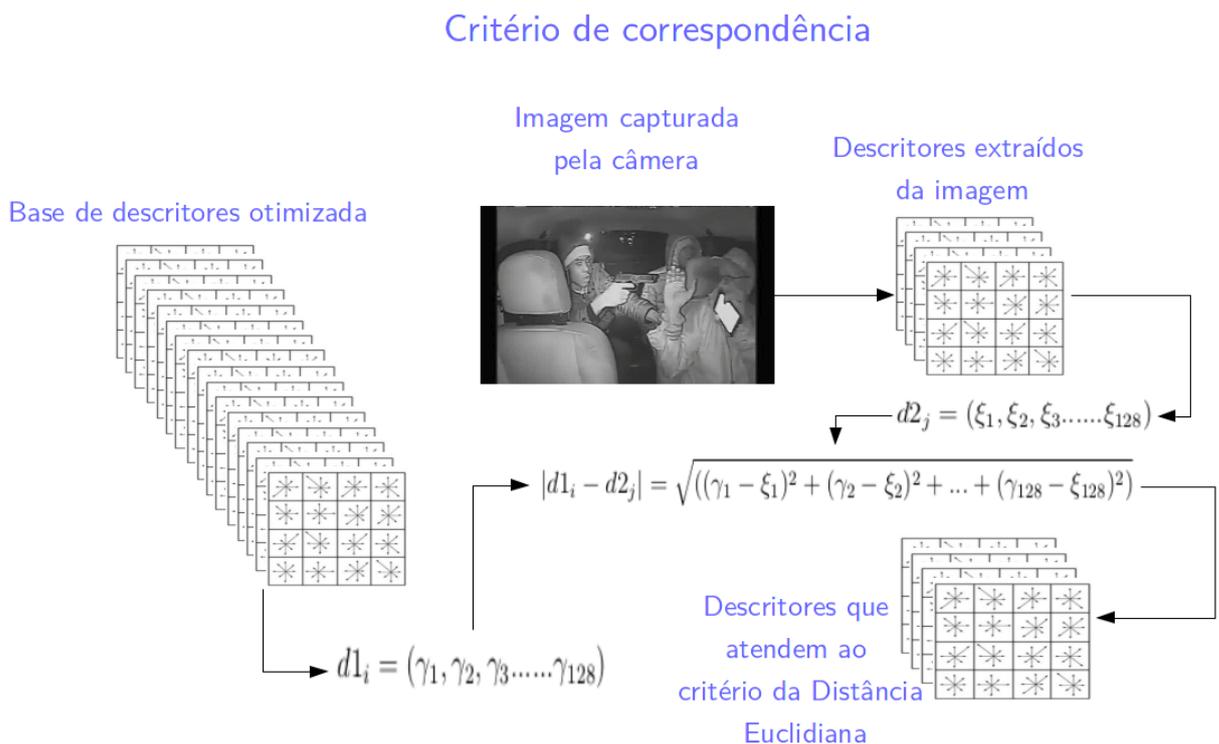
Fonte: O Autor.

Na Figura 32 está o fluxo do processo de otimização da base de descritores. Com esse processo o tamanho da base e consequentemente o tempo de acesso foi reduzido. Com a otimização finalizada também termina a fase de criação e otimização da base de descritores locais.

3.2 Fase de monitoramento

Nessa fase do método é feita a captura das sequências de imagens do interior do veículo e a verificação e reconhecimento da existência de armas de fogo. A Figura 33 mostra esse processo, aonde cada *frame* é comparado com o conjunto de descritores na base de dados. Esse processo de correspondência (*matching*) é realizado utilizando a Equação (22), que calcula a distância Euclidiana entre os pontos do *frame* atual como os pontos da base de descritores.

Figura 33 – Processo de monitoramento do ambiente interno do veículo em busca de armas de fogo, com o processo de *matching* entre os pontos da base de descritores e os pontos da imagem capturada do interior do veículo.



Fonte: O Autor.

Na Figura 34 o procedimento que faz a correspondência e define a semelhança entre dois descritores um pertencente a imagem corrente capturada do interior do veículo e outro pertencente a base de descritores. A função recebe como parâmetros dois *arrays* de vetores, *descAtual* que contém todos os descritores da imagem capturada do interior do veículo e *descRef* que contém todos os descritores da base de descritores. São criados dois

Figura 34 – Procedimento que verifica a semelhança entre dois descritores, o valor da distância Euclidiana foi calculado de forma empírica.

```

/**
 * Verifica semelhanças entre a base e a imagem capturada
 * descAtual: Array de vetores contendo todos os descritores encontrados na imagem capturada.
 * descRef: Array de vetores contendo todos os descritores da base de descritores.
 * desc1 e desc2: Vetores temporários que recebem um descritor dos arrays para o cálculo da distância.
 * nn: Distância euclidiana entre desc1 e desc2.
 * limiarDistancia: É o valor máximo da distância que considera dos descritores semelhantes.
 * qtdMinP: É a quantidade mínima de pontos que o método assume a presença da arma na imagem capturada.
 */
int achaPar1(Mat& descAtual, Mat& descRef) {
    int cont, cont_ng;
    cont = 0;
    cont_ng = 0;
    for (int i = 0; i < descAtual.rows; i++) {
        qtdMinP = 0;

        for (int k = 0; k < descRef.rows; k++) {
            Mat desc1 = descAtual.row(i);
            Mat desc2 = descRef.row(k);
            int nn = eudclidDistancia(desc1, desc2);
            if (nn < limiarDistancia) {
                qtdMinP = qtdMinP + 1;
                cont = cont + 1;
            }
        }
    }
    return (1);
}

```

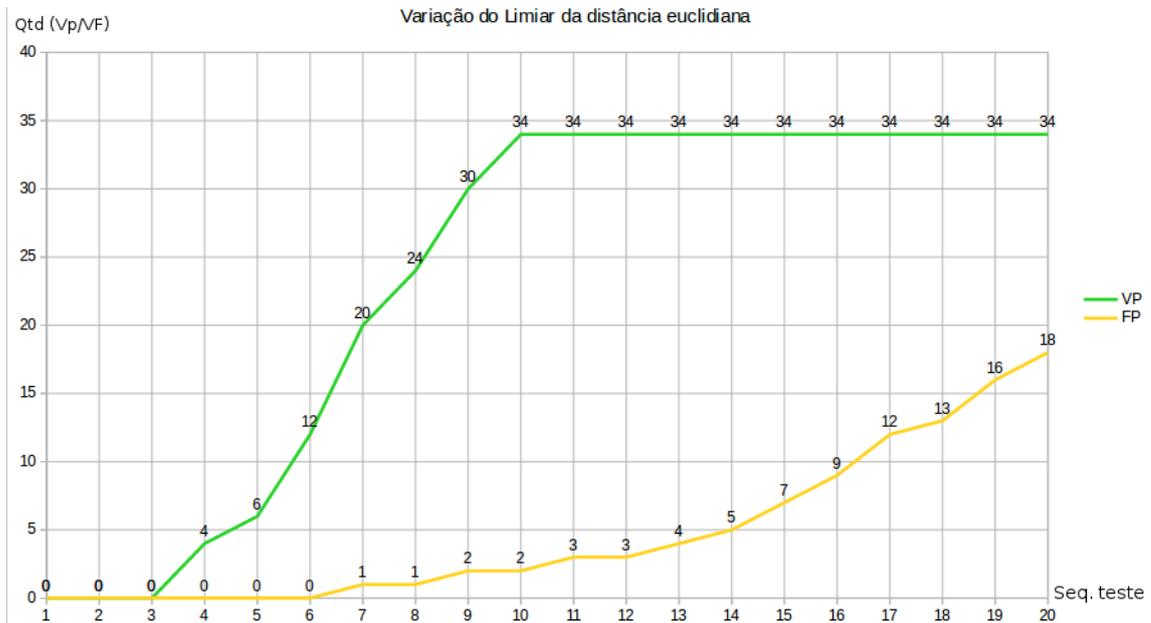
Fonte: O Autor.

laços que fazem a varredura dos dois *arrays*, no laço mais interno é feito o cálculo de cada distância Euclidiana e o resultado é comparado a variável global *limiarDistancia*, que define a máxima distância que indica a semelhança entre os descritores. Seu valor é obtido de forma empírica como será visto a seguir.

3.2.1 Procedimento para definição do limiar da distância Euclidiana

Como visto no procedimento da Figura 34 o valor de *limiarDistancia* foi definido de forma experimental. O experimento consiste em variar o seu valor e acompanhar o comportamento do método em relação a estas variações. Na Figura 35 está o resultado do experimento realizado para definir esse valor.

Figura 35 – Comportamento do método em função da variação do parâmetro *limiarDistancia*.



Fonte: O Autor.

No gráfico da Figura 35 o valor do *limiarDistancia* foi incrementado de 1, o objetivo é avaliar a resposta do método para cada valor de *limiarDistancia*. Na primeira fase foram testadas 36 imagens verdadeiras positivas, ou seja que possuíam arma de fogo. Para valores abaixo de 3 o método não conseguiu identificar a arma nessas imagens, no valor 4 das 36 imagens foram encontradas 4, no valor 5 foram encontradas 6 e assim sucessivamente sendo que para o valor 10 foram encontradas 34 imagens das 36. A partir daí até o valor 20 o método não conseguiu encontrar mais imagens verdadeiras positivas. Refazendo esse experimento com 36 imagens verdadeiras negativas, na valor 10 ocorreram 2 imagens falsos positiva das 36 imagens, após esse valor a quantidade de falsos positivo continua aumentando. Observando o gráfico resultante o valor ótimo para *limiarDistancia* é 10. E este valor foi adotado para a realização dos experimentos.

Outro parâmetro observado no procedimento da Figura 34 é a *qtdMinP* que representa a quantidade mínima de pontos que possuem a sua distância Euclidiana menor que o valor de *limiarDistancia*. Essa quantidade de pontos irá indicar ao método que a imagem é verdadeira positiva, ou seja a arma de fogo está presente na imagem. Essa quantidade é obtida pela observação da média de pontos chave (*mp*), detectados nas imagens que formam a base de descritores. Considerando que apenas 1 ponto de correspondência já é indicativo da presença da arma na imagem, temos que:

$$1 \leq qtdMinP \leq 0.05mp, \quad (23)$$

Conforme mostra a Equação 23 esse valor foi definido como sendo maior ou igual a 1 e menor ou igual a 5% do valor da médias de pontos chave detectados nas imagens que formam a base de descritores.

4 APLICAÇÕES EXPERIMENTAIS

4.1 Descrição dos experimentos e métricas

Neste capítulo são detalhados os experimentos realizados. fase foram realizados os experimentos. Foi utilizado o sistema operacional Linux (Ubuntu 16.04) e o ambiente de programação NetBeans 8.2 a linguagem de programação utilizada foi C++ e a biblioteca padrão foi a OpenCV 2.4, uma câmera Logitech 720p, um computador com processador Intel Core i7 2.4 GHz, 8Gb de memória e placa de vídeo GeForce 940MX de 2Gb.

Para avaliar os resultados encontrados nos experimentos realizados foram definidos algumas métricas. Na Figura 36 está a matriz de confusão que é a base para as definições das métricas que estão na Tabela 3.

Figura 36 – Matriz de confusão .

		Valor Verdadeiro (confirmado por análise)	
		positivos	negativos
Valor Previsto (predito pelo teste)	positivos	VP Verdadeiro Positivo	FP Falso Positivo
	negativos	FN Falso Negativo	VN Verdadeiro Negativo

Fonte: O Autor.

Tabela 3 – Tabela com as métricas usadas para avaliar a performance dos experimentos.

Indicador	Descrição
Acuracidade	$ACC = (VP + VN) / (P + N)$
Sensibilidade	$SENS = VP / (VP + FN)$
Especificidade	$SPEC = VN / (VN + FP)$
Eficiência	$EFF = (SENS + SPEC) / 2$

Fonte: Souza (2010).

Observando a Figura 36 considera-se que o valor previsto positivo ocorre quando o método indica a presença da arma em um *frame*, caso a arma realmente esteja presente então o método acertou e se contabiliza um verdadeiro positivo (VP=acertos), caso a arma não esteja presente então se contabiliza um falso positivo (FP=erros) e o método errou. Considera-se que o valor previsto negativo ocorre quando o método indica a não presença da arma em um *frame*, caso a arma realmente não esteja presente então o método acertou e se contabiliza um verdadeiro negativo (VN=acertos), caso a arma esteja presente então se contabiliza um falso negativo (FN=erros) e o método errou.

4.2 Base da dados de imagens utilizada nos experimentos

A base de dados de imagens que será utilizada em todos os experimentos foi criada a partir de imagens de revólveres e pistolas coletadas de IMFDB (2017). Essa base passou pelo processo de otimização descrito anteriormente. Na Figura 37 alguns exemplos de imagens da base de dados.

Figura 37 – Alguns exemplos de imagens de armas gerados a partir da base de dados IMFDB (2017) utilizados nos experimentos.

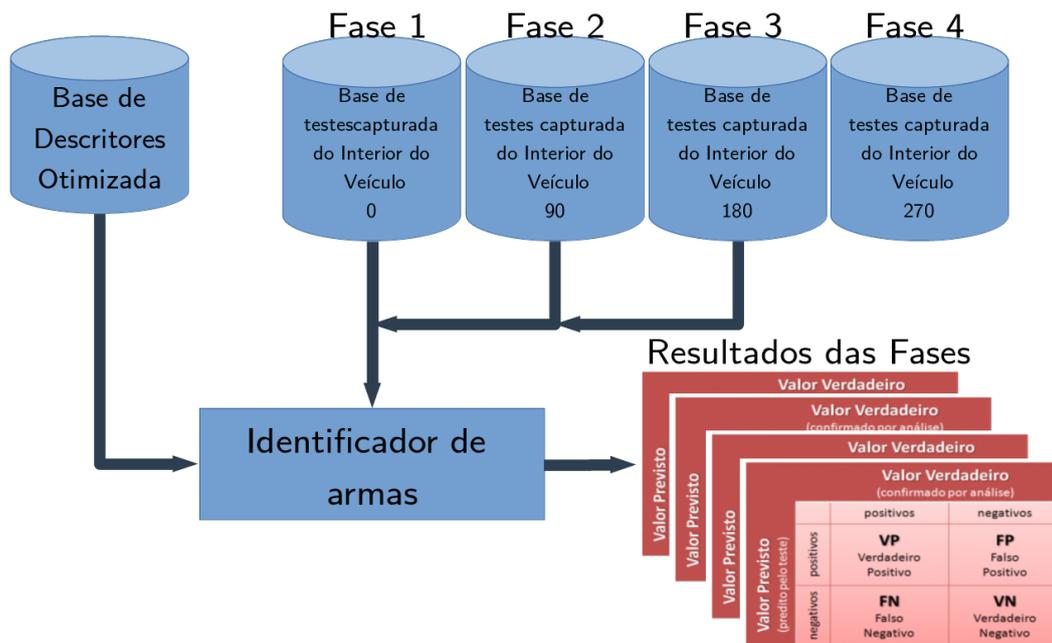


Fonte: O Autor.

4.3 Procedimento experimental I - Análise da resposta do método com relação mudanças no ângulo da arma

Durante o processo de captura da imagem é bastante razoável considerar que a arma não irá permanecer estática ao longo do tempo. Desta forma o objetivo deste experimento é avaliar a resposta do método quanto aos indicadores da Tabela 3 quando ocorre uma rotação da arma na imagem. O experimento I se dividiu em 4 fases, na primeira fase as imagens da base de teste sofreram uma rotação de 0^0 e em seguida foram classificadas pelo método. Na segunda fase o processo se repetiu sendo que o ângulo da rotação foi 90^0 , na terceira fase o ângulo foi 180^0 e na quarta fase o ângulo foi 270^0 . Na Figura 38 o fluxograma deste experimento.

Figura 38 – Abaixo o fluxo do experimento I, que visa obter a resposta do método quando houver uma mudança no ângulo da arma.



Fonte: O Autor.

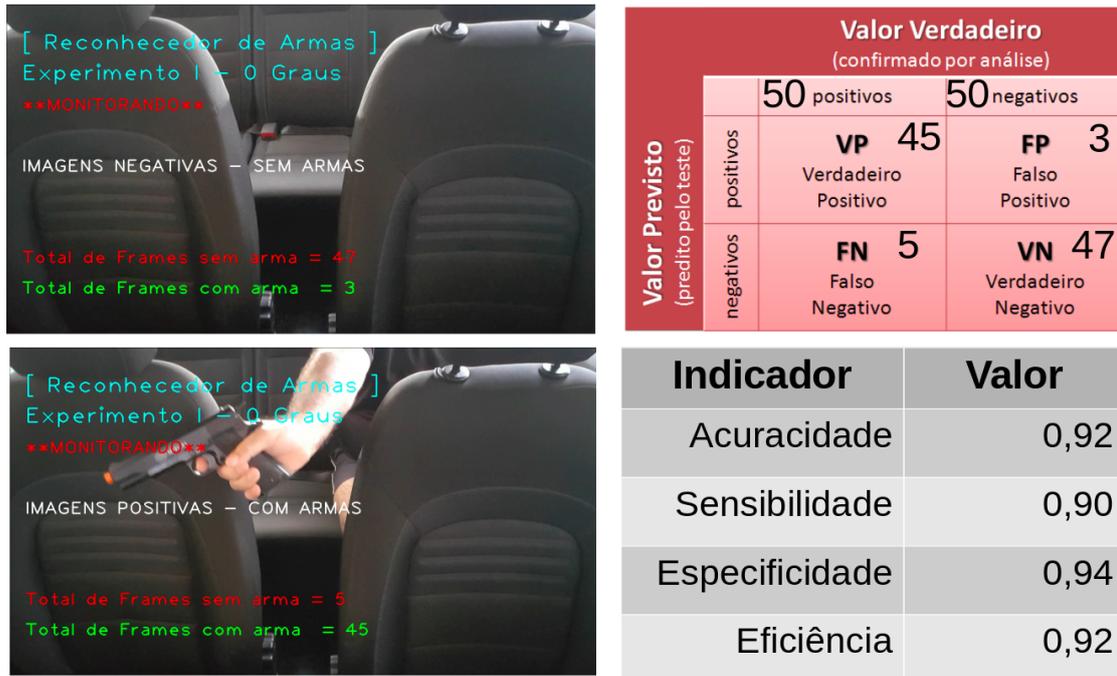
4.3.1 Resultados do Experimento I

4.3.1.1 Resultados da Fase 1 - Rotação de 0^0

Na Figura 39 estão os resultados obtidos pelo método com relação aos dois grupos de *frames* que foram rotacionado em 0^0 . Nos primeiros 50 *frames* havia a presença da

arma, no entanto o método fez o reconhecimento em 45. No segundo bloco de *frames* nenhum deles possuía a imagem da arma, o método confirmou isso em 47 no entanto apontou a presença da arma em 3. No canto direito inferior da Figura 39 estão os valores das métricas obtidas neste experimento de acordo com o que foi definido na Tabela 3.

Figura 39 – Resultados obtidos para uma rotação de 0^0 .

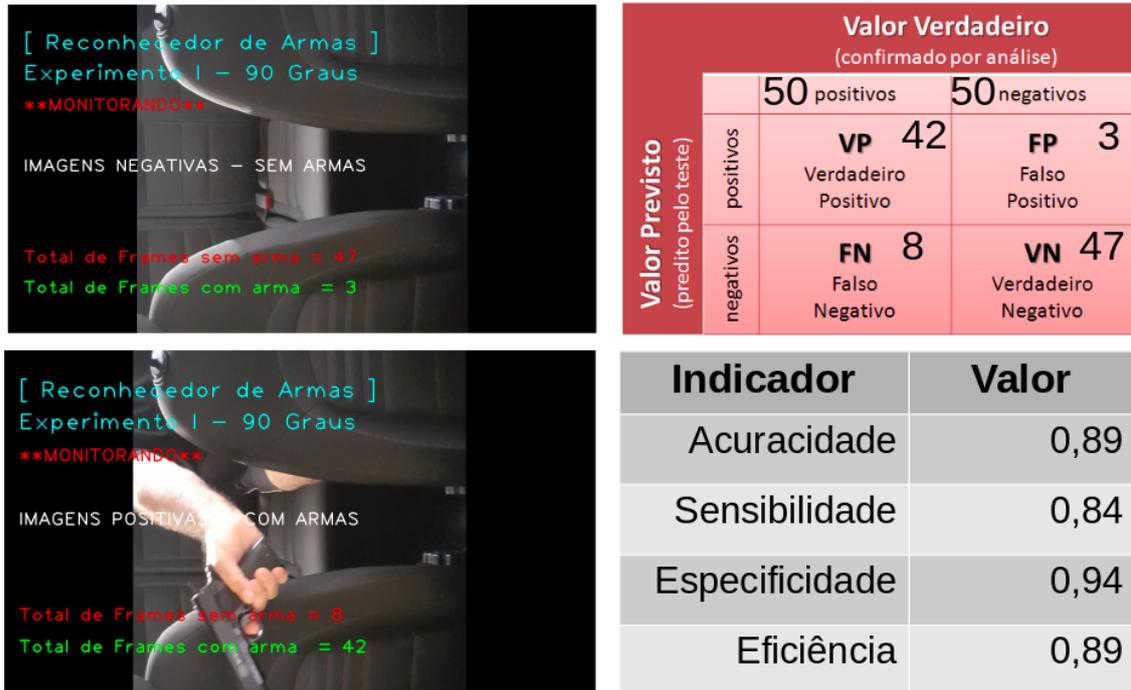


Fonte: O Autor.

4.3.1.2 Resultados da Fase 2 - Rotação de 90^0

Na Figura 40 estão os resultados obtidos pelo método com relação aos dois grupos de *frames* que foram rotacionado em 90^0 . Nos primeiros 50 *frames* havia a presença da arma, no entanto o método fez o reconhecimento em 42. No segundo bloco de *frames* nenhum deles possuía a imagem da arma, o método confirmou isso em 47 no entanto apontou a presença da arma em 3. No canto direito inferior da Figura 40 estão os valores das métricas obtidas neste experimento de acordo com o que foi definido na Tabela 3.

Figura 40 – Resultados obtidos para uma rotação de 90°.

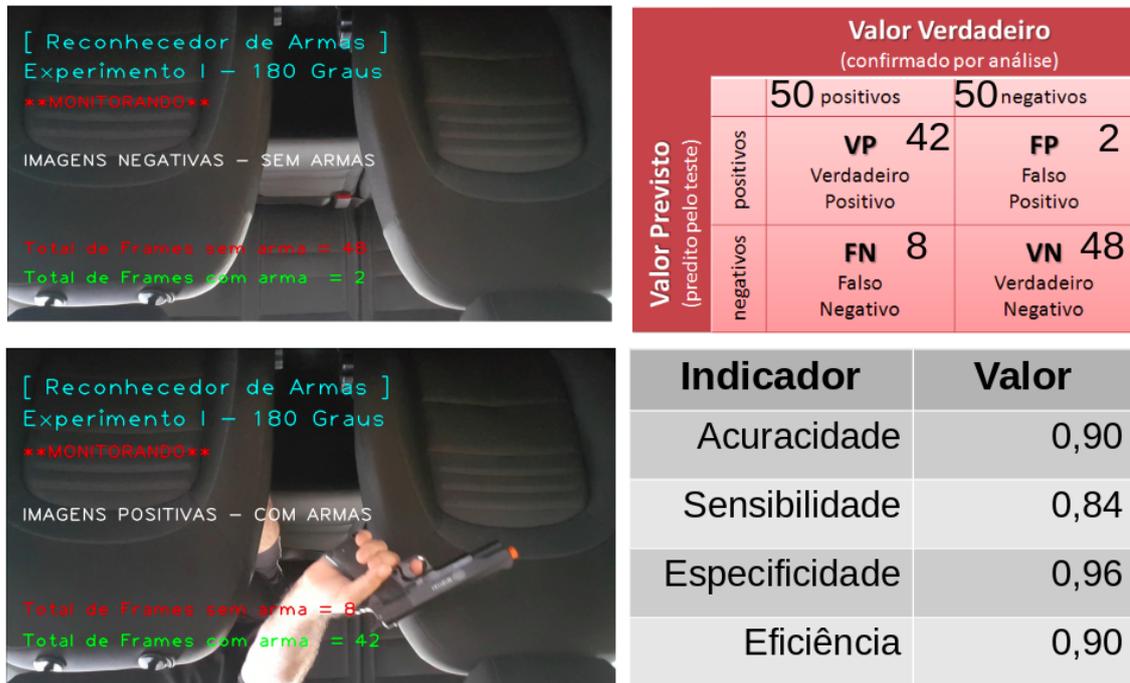


Fonte: O Autor.

4.3.1.3 Resultados da Fase 3 - Rotação de 180°

Na Figura 41 estão os resultados obtidos pelo método com relação aos dois grupos de *frames* que foram rotacionado em 180°. Nos primeiros 50 *frames* havia a presença da arma, no entanto o método fez o reconhecimento em 42. No segundo bloco de *frames* nenhum deles possuía a imagem da arma, o método confirmou isso em 48 no entanto apontou a presença da arma em 2. No canto direito inferior da Figura 41 estão os valores das métricas obtidas neste experimento de acordo com o que foi definido na Tabela 3.

Figura 41 – Resultados obtidos para uma rotação de 180°.

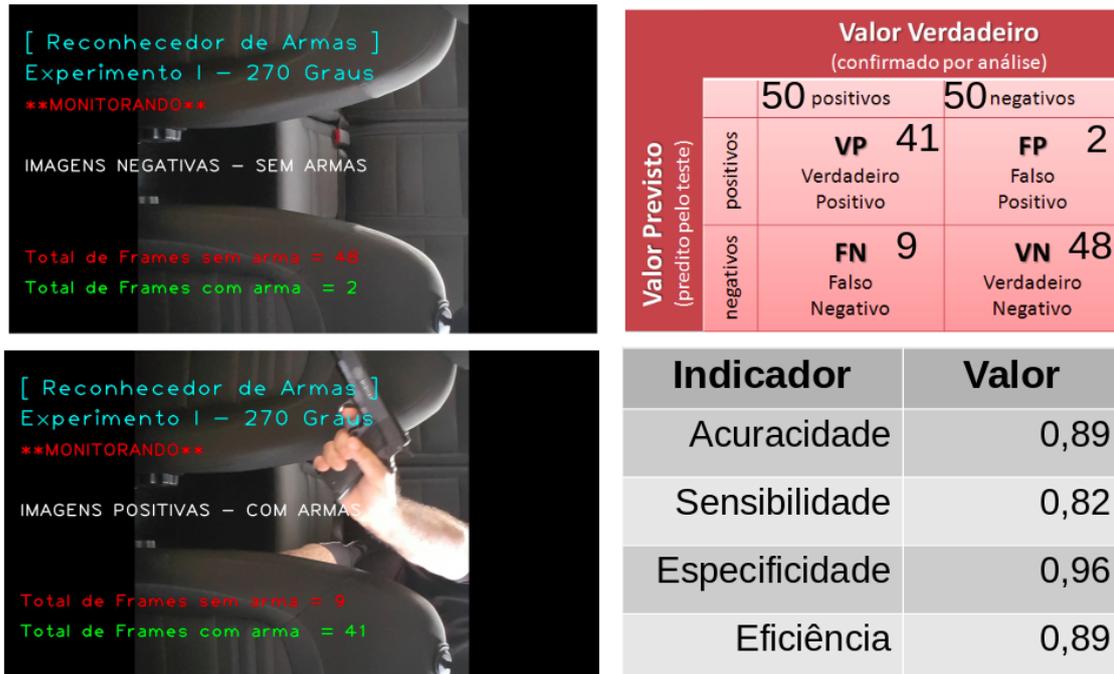


Fonte: O Autor.

4.3.1.4 Resultados da Fase 4 - Rotação de 270°

Na Figura 42 estão os resultados obtidos pelo método com relação aos dois grupos de *frames* que foram rotacionado em 270°. Nos primeiros 50 *frames* havia a presença da arma, no entanto o método fez o reconhecimento em 41. No segundo bloco de *frames* nenhum deles possuía a imagem da arma, o método confirmou isso em 48 no entanto apontou a presença da arma em 2. No canto direito inferior da Figura 42 estão os valores das métricas obtidas neste experimento de acordo com o que foi definido na Tabela 3.

Figura 42 – Resultados obtidos para uma rotação de 270⁰.



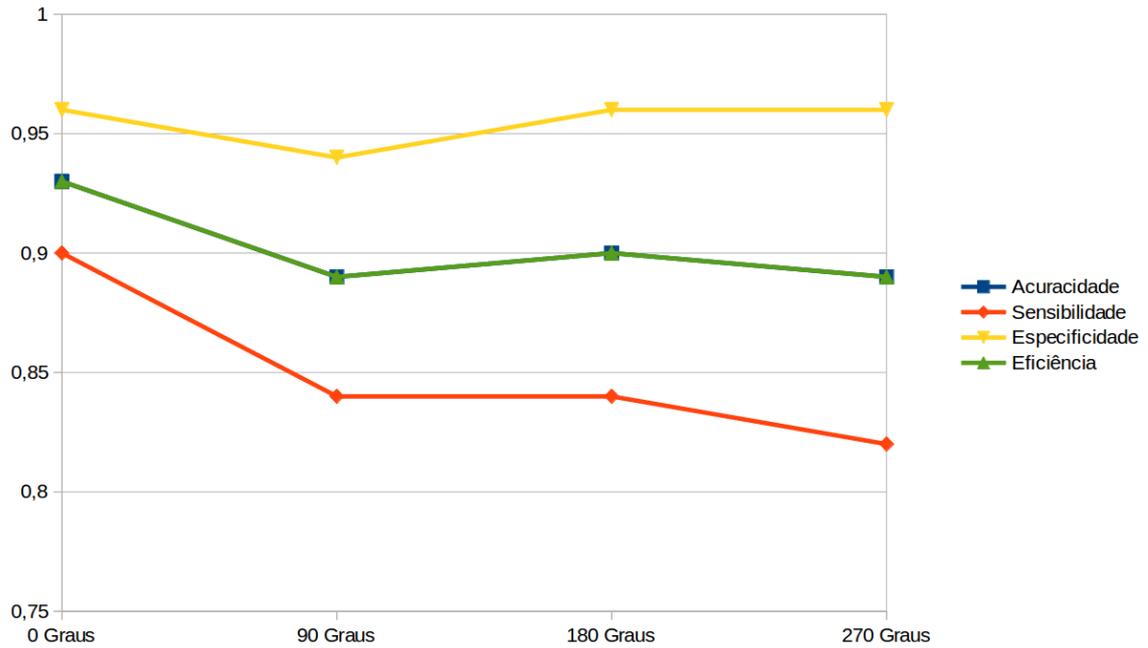
Fonte: O Autor

4.3.2 Discussão dos resultados do Experimento I

A Figura 43 mostra a variação nas métricas de acordo com a variação nos ângulos da imagem. Os valores mostram que ocorreram perdas em todas as métricas ou seja as rotações afetaram a resposta do método. Na Figura 44 está o desvio padrão calculado. No caso mais crítico que ocorreu em 270⁰, a variação ficou em torno de 0,06 que é uma variação três vezes maior que a ocorreu com o ângulo 0⁰. Houveram perdas significativas no valores das métricas, porém o método em todos os casos obteve valores acima de 80% para todos os ângulos testados.

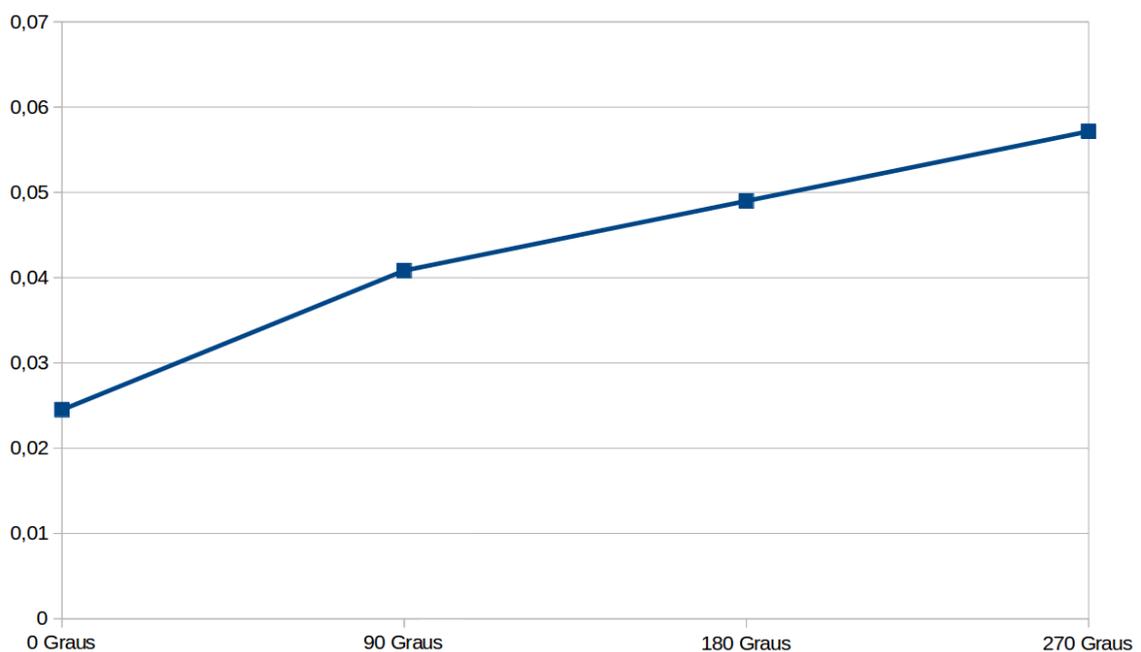
Dessa forma a conclusão que se chega é que o método teve um aumento na variabilidade das métricas quando ocorrem rotações na imagem. Em termos absolutos essa variação é baixa e não compromete a viabilidade prática do método.

Figura 43 – Gráfico com a variação nos resultados obtidos pelo método em função das rotações na imagem.



Fonte: O Autor.

Figura 44 – Gráfico com o desvio padrão dos resultados obtidos pelo método em função das rotações na imagem.

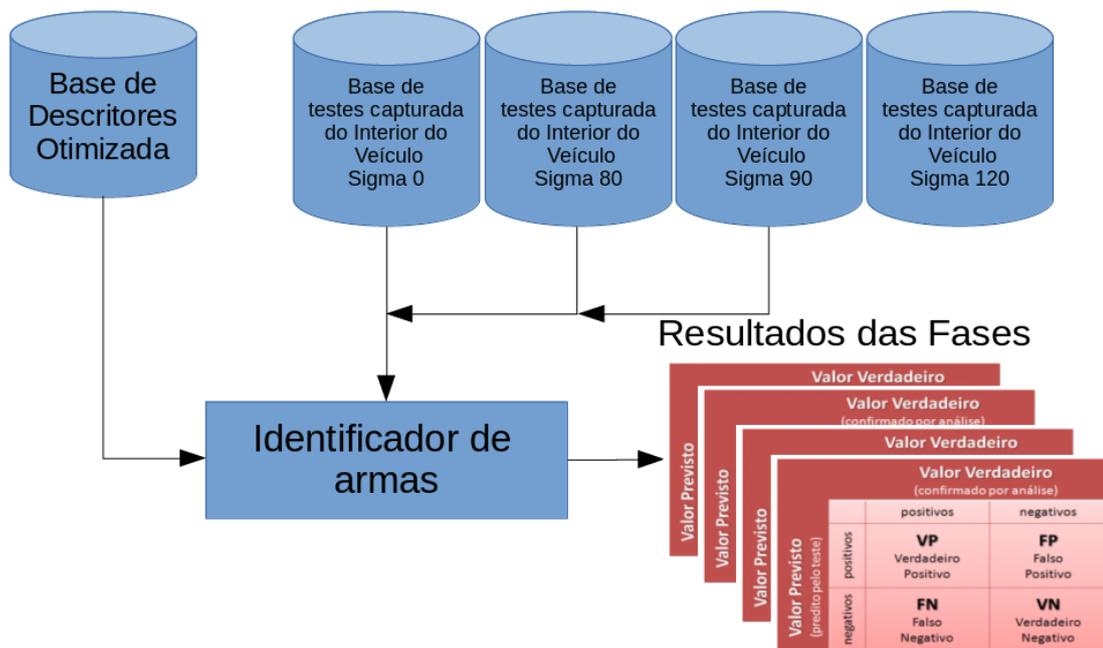


Fonte: O Autor.

4.4 Procedimento experimental II - Análise da resposta do método com relação a ocorrência de ruídos na imagem

O ambiente interno de um veículo principalmente em função da temperatura pode afetar a qualidade da imagem capturada pela câmera. O efeito mais comum é a geração de ruído na imagem e isso pode afetar a resposta do método. Desta forma o objetivo deste experimento é avaliar a resposta do método quanto aos indicadores da Tabela 3 quando ocorre ruído na imagem. O experimento II se dividiu em 4 fases, na primeira fase as imagens da base de teste passaram a ter um ruído Gaussiano com $\sigma = 0$ e em seguida foram classificadas pelo método. Na segunda fase o processo se repetiu sendo que agora $\sigma = 80$ na terceira fase $\sigma = 90$ e na quarta fase $\sigma = 120$. Na Figura 45 o fluxograma deste experimento.

Figura 45 – Abaixo o fluxo do experimento II, que visa obter a resposta do método quando ocorre ruído na imagem.



Fonte: O Autor.

4.4.1 Resultados do Experimento II

4.4.1.1 Resultados da Fase 1 - $\sigma = 0$

Na Figura 47 estão os resultados obtidos pelo método com relação aos dois grupos de *frames* que foram submetidos a ruído Gaussiano com $\sigma = 0$. Nos primeiros 50 *frames* havia a presença da arma, no entanto o método fez o reconhecimento em 45. No segundo bloco de *frames* nenhum deles possuía a imagem da arma, o método confirmou isso em 47 no entanto apontou a presença da arma em 3. No canto direito inferior da Figura 46 estão os valores das métricas obtidas neste experimento de acordo com o que foi definido na Tabela 3.

Figura 46 – Resultados obtidos para ruído com $\sigma = 0$.



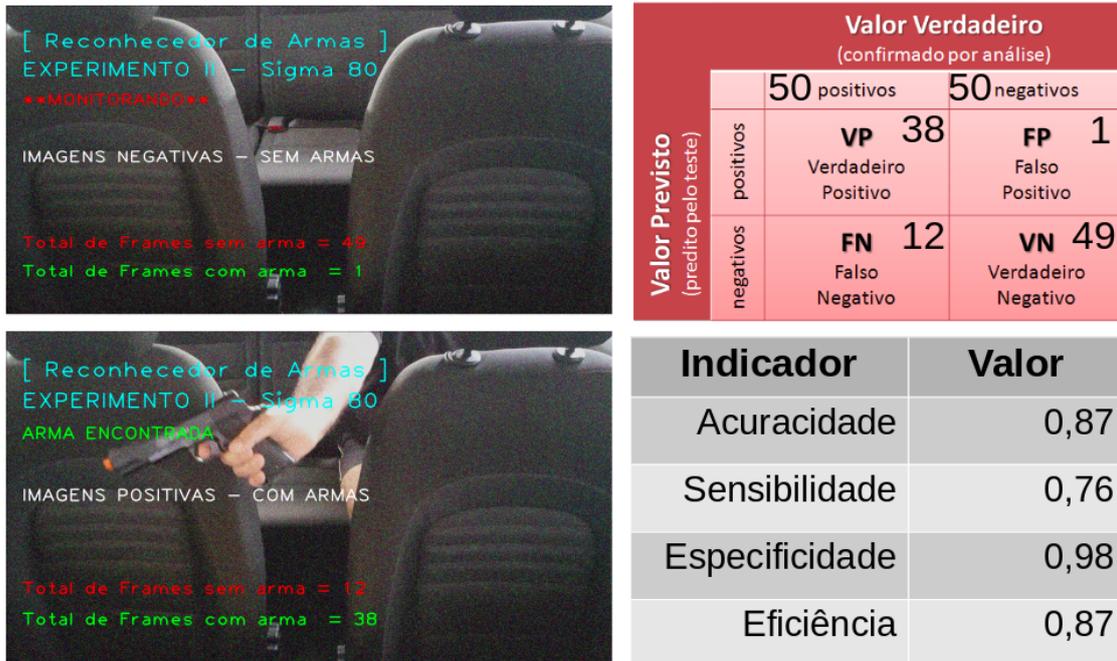
Fonte: O Autor.

4.4.1.2 Resultados da Fase 2 - $\sigma = 80$

Na Figura 47 estão os resultados obtidos pelo método com relação aos dois grupos de *frames* que foram submetidos a ruído Gaussiano com $\sigma = 80$. Nos primeiros 50 *frames* havia a presença da arma, no entanto o método fez o reconhecimento em 38. No segundo bloco de *frames* nenhum deles possuía a imagem da arma, o método confirmou isso em 49

no entanto apontou a presença da arma em 1. No canto direito inferior da Figura 47 estão os valores das métricas obtidas neste experimento de acordo com o que foi definido na Tabela 3.

Figura 47 – Resultados obtidos para ruído com $\sigma = 80$.

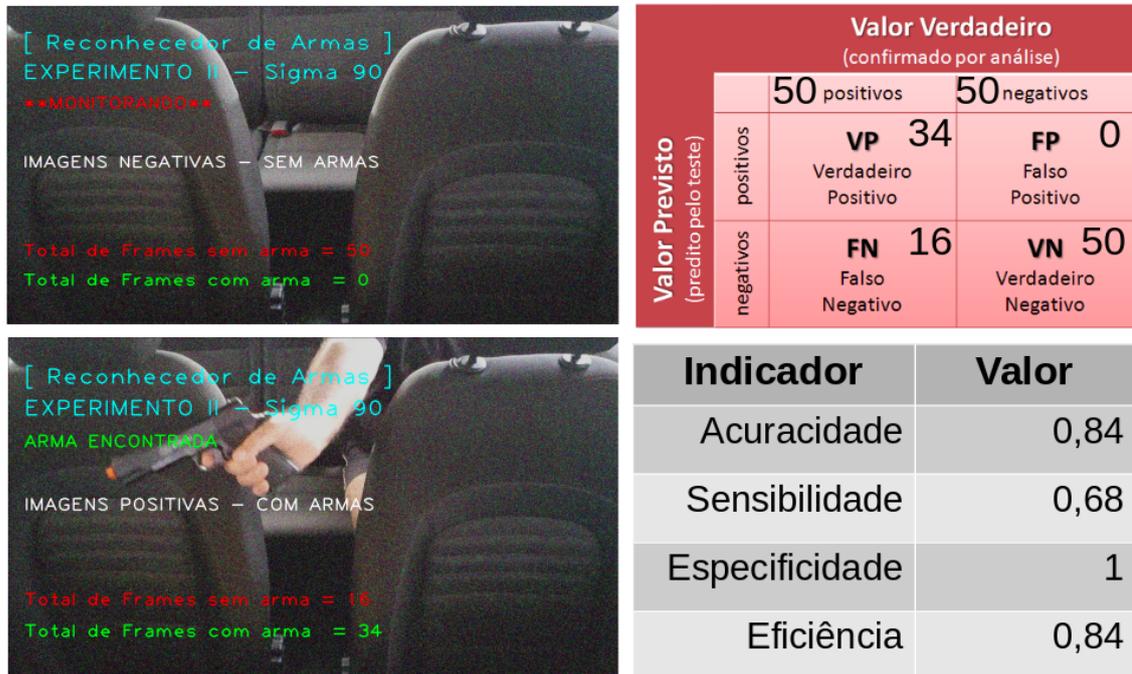


Fonte: O Autor.

4.4.1.3 Resultados da Fase 3 - $\sigma = 90$

Na Figura 48 estão os resultados obtidos pelo método com relação aos dois grupos de *frames* que foram submetidos a ruído Gaussiano com $\sigma = 90$. Nos primeiros 50 *frames* havia a presença da arma, no entanto o método fez o reconhecimento em 34. No segundo bloco de *frames* nenhum deles possuía a imagem da arma, o método confirmou isso em 50. No canto direito inferior da Figura 48 estão os valores das métricas obtidas neste experimento de acordo com o que foi definido na Tabela 3.

Figura 48 – Resultados obtidos para ruído com $\sigma = 90$.

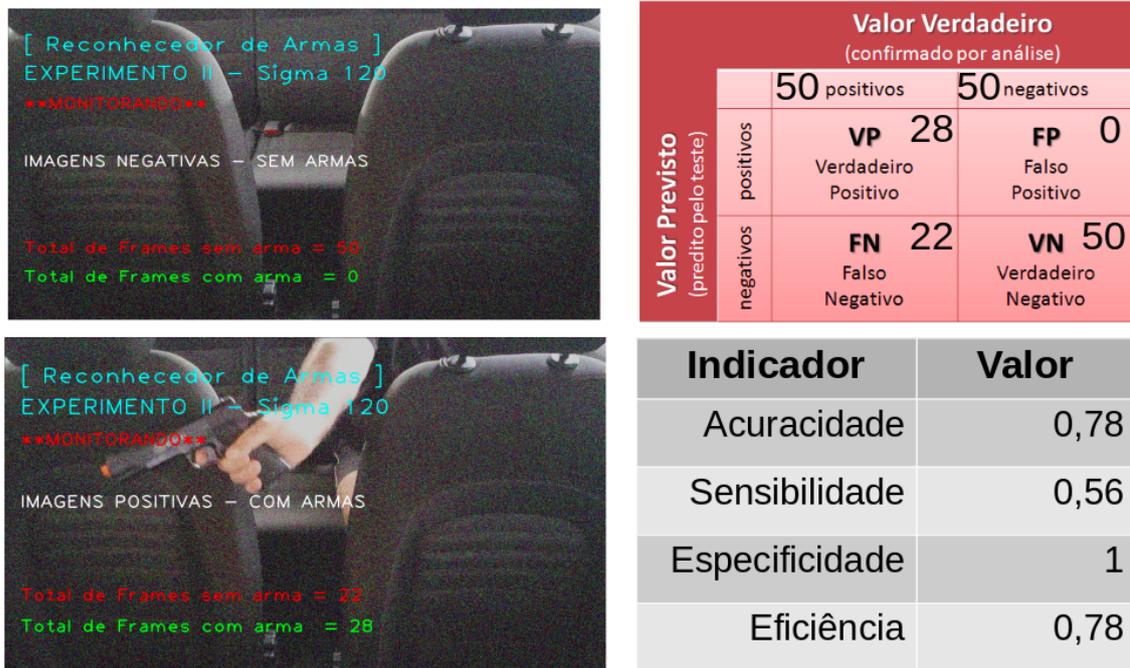


Fonte: O Autor.

4.4.1.4 Resultados da Fase 4 - $\sigma = 120$

Na Figura 49 estão os resultados obtidos pelo método com relação aos dois grupos de *frames* que foram submetidos a ruído Gaussiano com $\sigma = 120$. Nos primeiros 50 *frames* havia a presença da arma, no entanto o método fez o reconhecimento em 28. No segundo bloco de *frames* nenhum deles possuía a imagem da arma, o método confirmou isso em 50. No canto direito inferior da Figura 49 estão os valores das métricas obtidas neste experimento de acordo com o que foi definido na Tabela 3.

Figura 49 – Resultados obtidos para ruído com $\sigma = 120$.

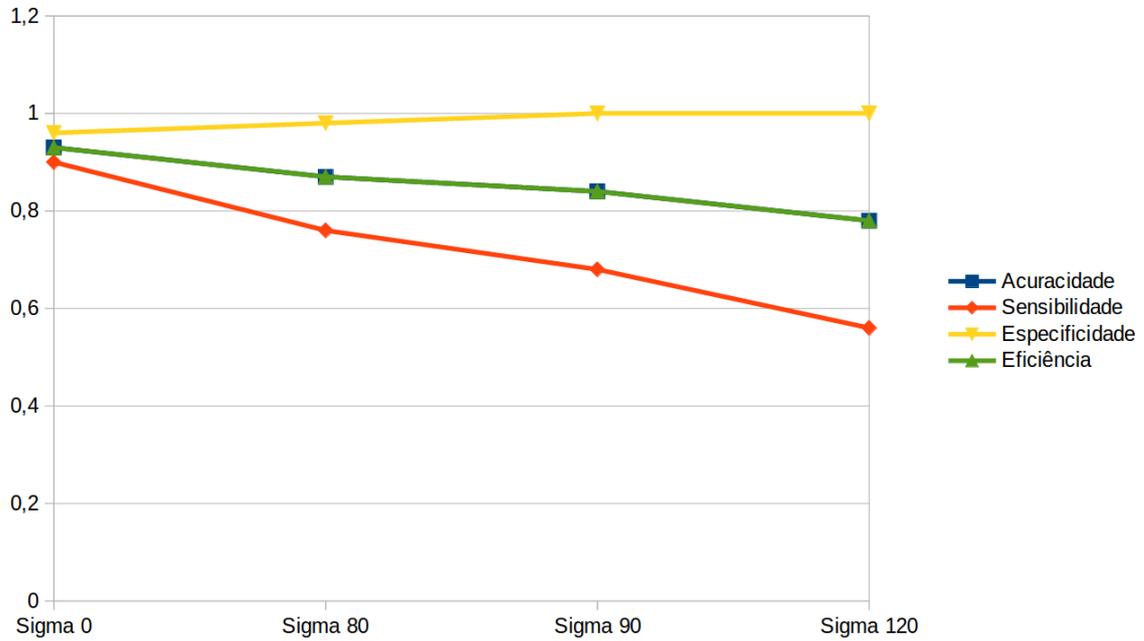


Fonte: O Autor

4.4.2 Discussão dos resultados do Experimento II

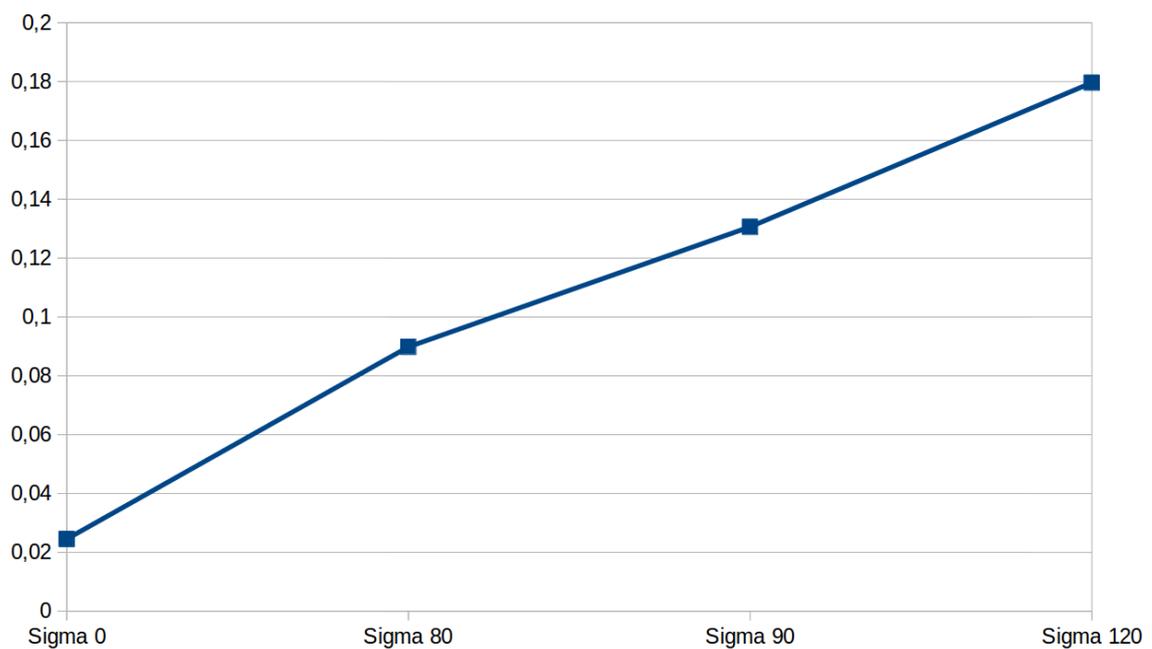
A Figura 50 mostra a variação nas métricas de acordo com a variação na intensidade do ruído na imagem. Os valores mostram que ocorreram perdas em todas as métricas ou seja o ruído afetou a resposta do método. Na Figura 51 está o desvio padrão calculado. No caso mais crítico que ocorreu em $\sigma = 120$, a variação ficou em torno de 0,018 que é uma variação quase dez vezes maior que quando $\sigma = 0$. Houveram perdas significativas no valores das métricas, em três delas os valores ficaram abaixo de 80%. Dessa forma a conclusão que se chega é que o método teve um aumento na variabilidade das métricas e uma redução significativa nos seus valores. Pelo gráfico da Figura 50 para que o método opere com os valores das métricas acima de 80% o valor de $\sigma < 40$ deve ser obedecido, pois valores acima deste comprometem a viabilidade prática do método.

Figura 50 – Gráfico com a variação nos resultados obtidos pelo método em função da variação no nível de ruído na imagem.



Fonte: O Autor.

Figura 51 – Gráfico com o desvio padrão dos resultados obtidos pelo método em função variação no nível de ruído na imagem.

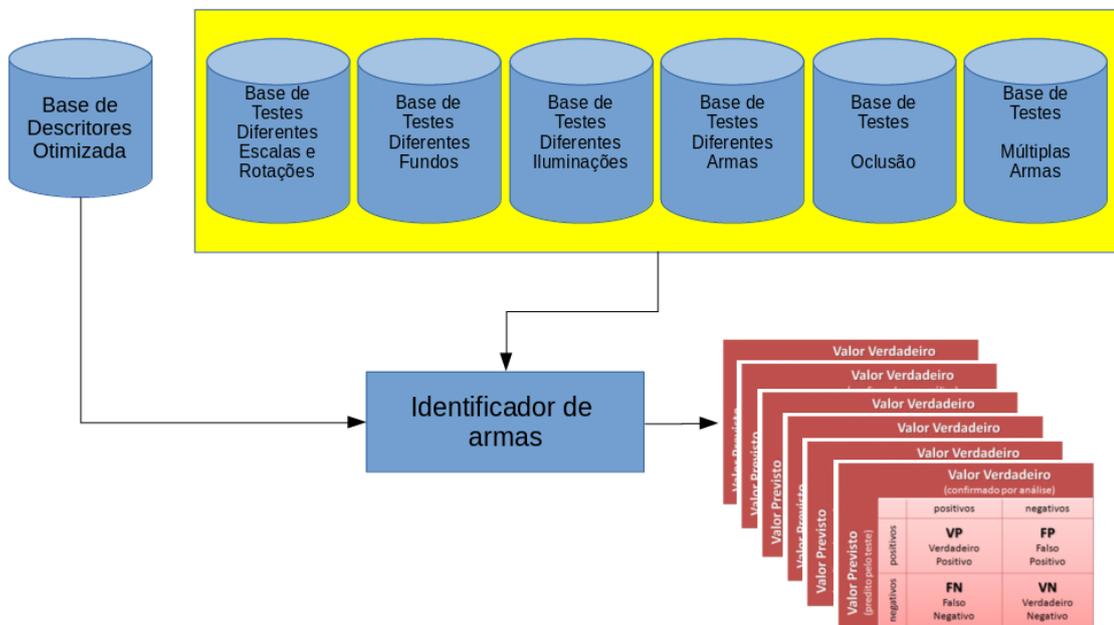


Fonte: O Autor.

4.5 Procedimento experimental III - Comparação dos resultados entre o experimento realizado por Tiwari e Verma (2015a) e o método desenvolvido

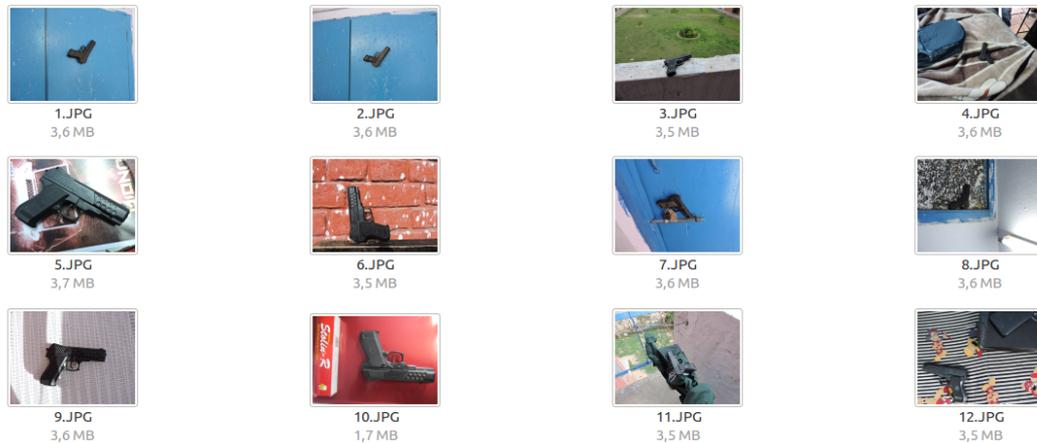
O objetivo deste experimento é comparar os resultados obtidos com o método desenvolvido em relação aos resultados do trabalho de Tiwari e Verma (2015a), utilizando a mesma base de imagens de testes utilizada em seu experimento e as mesmas métricas. O fluxograma deste experimento está na Figura 52.

Figura 52 – Fluxograma do experimento III, aonde será feita comparação entre os resultados obtidos por Tiwari e Verma (2015a) e o do método desenvolvido.



Fonte: O Autor.

Figura 53 – Exemplos de imagens utilizadas por Tiwari e Verma (2015a).



Fonte: Tiwari e Verma (2015a).

Na Figura 53 alguns exemplos da base de imagens gentilmente cedidas por Tiwari e Verma (2015a) para a realização deste experimento.

4.5.1 Resultados do Experimento III

O método fez a classificação das imagens presentes na base de imagens do trabalho de Tiwari e Verma (2015a), de forma que foram avaliadas pelo método um total de 12 imagens que possuíam a imagem da arma, mas com uma variação na imagem de fundo de cada uma delas, 9 imagens que possuíam a imagem da arma mas com uma variação na iluminação de cada uma delas, 11 imagens que possuíam a imagem da arma mas com a inclusão de objetos da mesma cor da arma, 17 imagens que possuíam a imagem da arma mas com a ocorrência de oclusão em cada uma delas, 10 imagens que possuíam a imagem da arma mas com uma variação na escala da arma em cada uma delas e 6 imagens que possuíam a imagem de mais de uma arma.

Abaixo os resultados obtidos.

- a) Diferentes fundos - Figura 54;
- b) Variação na iluminação -Figura 55;
- c) Armas diferentes -Figura 56;
- d) Oclusão - Figura 57;
- e) Variação na escala e rotação - Figura 58;
- f) Múltiplas armas - Figura 59.

Figura 54 – Resultado final do método, com a classificação de 12 *frames* que contém a arma e que possuem variação no fundo da imagem.



Fonte: O Autor.

Na Figura 54 o resultado do método quando ocorre uma mudança no imagem de fundo. Em todos os casos a arma foi identificada sem problemas alcançando um resultado de 100% de acerto.

Figura 55 – Resultado final do método, com a classificação de 9 *frames* que contém a arma e que possuem variação na iluminação da imagem.



Fonte: O Autor.

Na Figura 55 o resultado do método quando ocorre uma mudança na iluminação. Em todos os casos a arma foi identificada sem problemas alcançando um resultado de 100% de acerto.

Figura 56 – Resultado final do método, com a classificação de 11 *frames* que contém a arma e que possuem variação na classe da arma



Fonte: O Autor.

Na Figura 56 o resultado do método quando ocorre a inclusão de outros objetos da mesma cor da arma. Em todos os casos a arma foi identificada sem problemas alcançando um resultado de 100% de acerto e não houve nenhuma identificação incorreta.

Figura 57 – Resultado final do método, com a classificação de 17 *frames* que contém a arma e que possuem oclusão



Fonte: O Autor.

Na Figura 57 o resultado do método quando ocorre oclusão de parte da arma. Em todos os casos a arma foi identificada sem problemas alcançando um resultado de 100% de acerto.

Figura 58 – Resultado final do método, com a classificação de 10 *frames* que contém a arma e que possuem variação na escala e rotação



Fonte: O Autor.

Na Figura 58 o resultado do método quando ocorre uma mudança na escala da arma. Em todos os casos a arma foi identificada sem problemas alcançando um resultado de 100% de acerto.

Figura 59 – Resultado final do método, com a classificação de 6 *frames* que contém múltiplas armas.



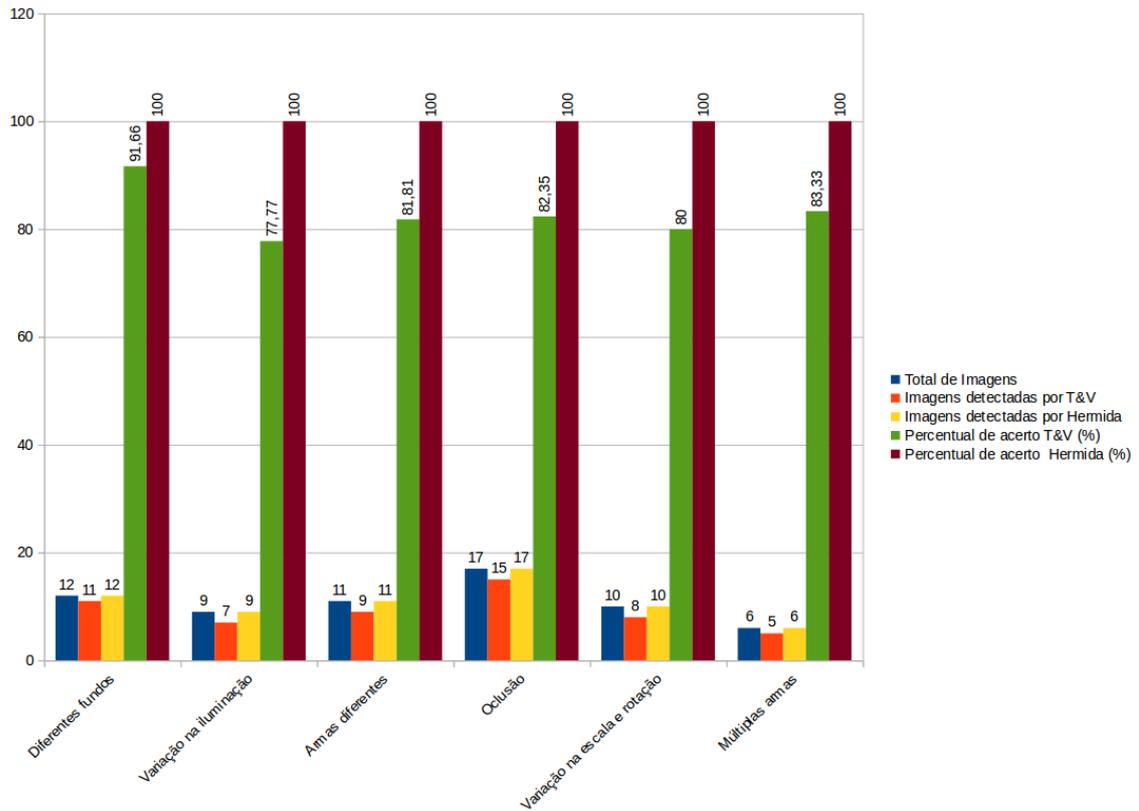
Fonte: O Autor.

Na Figura 59 o resultado do método quando ocorre a presença de mais de uma arma na imagem. Em todos os casos a arma foi identificada sem problemas alcançando um resultado de 100% de acerto.

4.5.2 Discussão dos resultados do Experimento III

Os resultados do Experimento III estão no gráfico da Figura 60. Em todos os casos a abordagem de descritores locais utilizando o método SIFT obteve resultados superiores.

Figura 60 – Resultado final com os resultados alcançados pelo método desenvolvido nos 6 cenários.

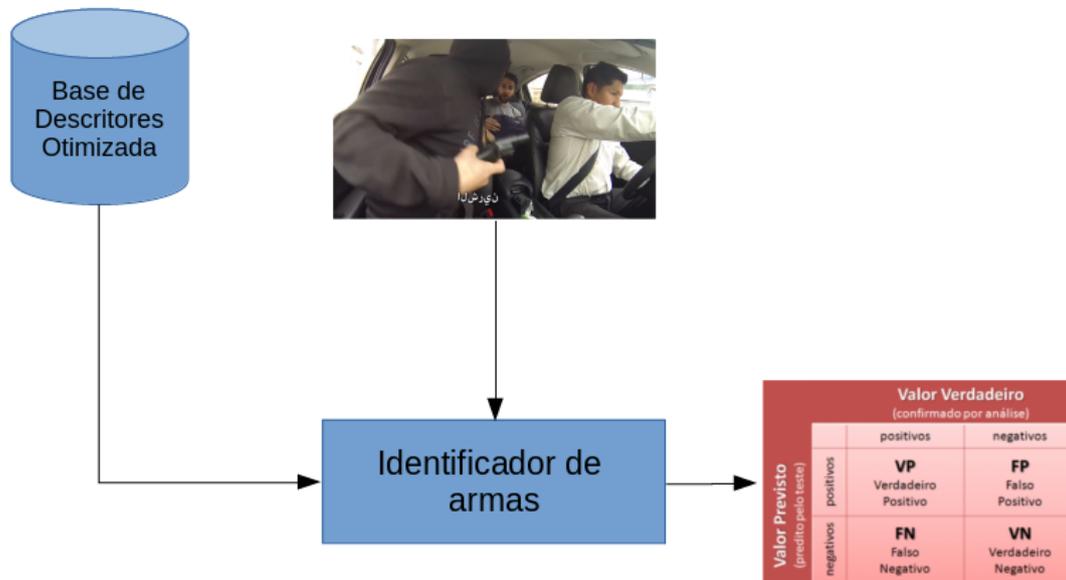


Fonte: O Autor

4.6 Procedimento experimental IV - Análise da resposta do método desenvolvido aplicado a um vídeo real

O objetivo deste experimento é avaliar os resultados do método desenvolvido no processo de classificação de um vídeo da internet que possui armas em alguns *frames*. O fluxograma deste experimento está na Figura 61.

Figura 61 – Fluxograma do experimento IV aonde será feita a análise dos resultados do método desenvolvido classificando um vídeo da internet.

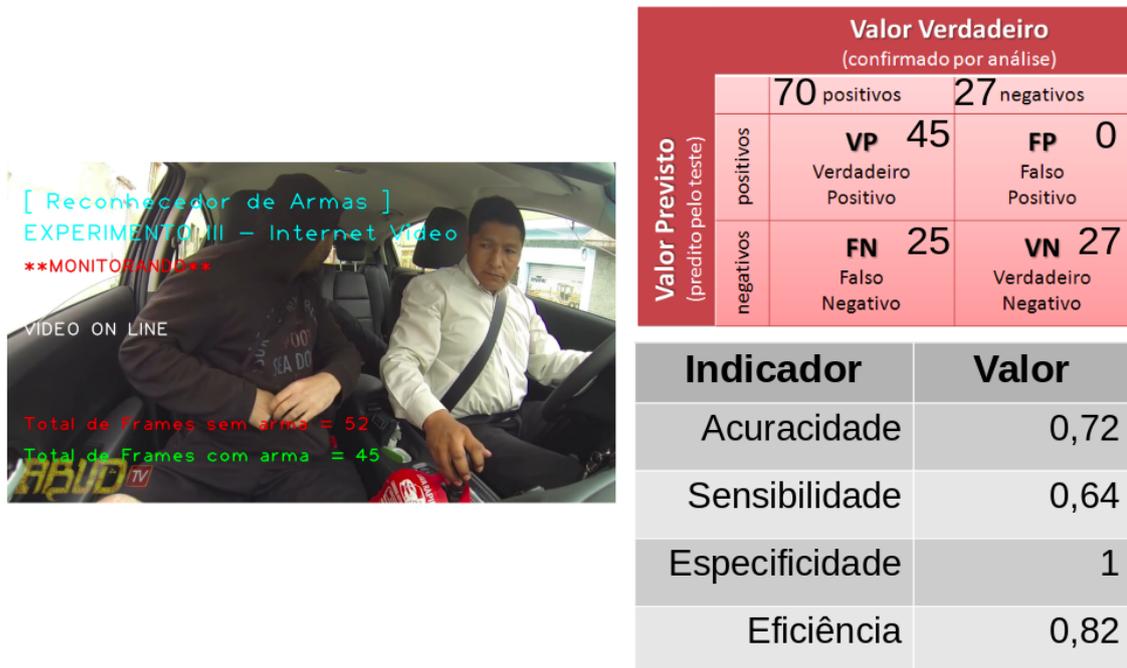


Fonte: O Autor.

4.6.1 Resultados do Experimento IV

Na Figura 62 estão os resultados obtidos pelo método com relação ao vídeo. O vídeo possui 70 *frames* aonde a arma está visível e 27 *frames* aonde a arma não está visível. O método reconheceu a presença da arma em 45 *frames* e não reconheceu a arma em nenhum dos 27 *frames* em que ela não estava presente.

Figura 62 – Resultados obtidos da classificação do vídeo pelo método desenvolvido.



Fonte: O Autor.

4.6.2 Discussão dos resultados do Experimento IV

O experimento IV mostrou a viabilidade prática do método desenvolvido, as condições do vídeo estavam boas quanto ao nível de ruído, qualidade da imagem e iluminação. Neste experimento o método mostrou uma eficiência de mais de 80% em condições não controladas. Embora tenha ocorrido um grande número de falsos negativo, ou seja o sistema não encontra a arma, no geral a detecção ocorreu de forma satisfatória.

5 CONCLUSÕES E TRABALHOS FUTUROS

5.1 Conclusões

Neste trabalho foi desenvolvido um método para reconhecer automaticamente a presença e o porte de armas de fogo, revólveres e pistolas, comumente utilizados nos assaltos a mão armada no interior de carros de passeio. A aplicabilidade do método é no combate a violência urbana e a viabilização de uma abordagem facilmente integrável aos sistemas de segurança e monitoramento interno de veículos, visando a atuação mais rápida e efetiva dos agentes de segurança pública ou privada.

A abordagem desenvolvida baseou-se na adaptação do algoritmo SIFT (*Scale Invariant Feature Detector*), Lowe (1999) para a obtenção de características das armas de fogo por meio da formação de uma base de descritores. Os descritores locais são utilizados para comparar imagens reais de armas de fogo, capturadas do interior de um veículo com as imagens armazenadas na base de descritores.

O algoritmo SIFT faz uso de uma estrutura conhecida como espaço de escala, que consiste na realização de borrarmentos Gaussianos sucessivos e operações de redução de escala, para se obter uma espécie de pirâmide de gaussianas da imagem original. Com essa pirâmide, o SIFT busca aproximar o valor do Laplaciano da imagem em cada estágio da pirâmide. Os pontos extremos em cada nível são encontrados por meio da operação de diferença de imagens nas gaussianas de cada nível. De posse dos pontos de máximo e mínimo locais, o algoritmo calcula então orientações e magnitudes (ou seja, em coordenadas polares) para um conjunto de pontos próximo a um desses extremos. Em seguida, monta-se um histograma dos pontos nessa janela gaussiana cujos valores possíveis são intervalos de tamanho uniforme no comprimento do círculo trigonométrico e usando como pesos para a inserção no histograma as magnitudes desses pontos. A partir da obtenção dessa orientação e magnitude representativa, o algoritmo monta então vetores que objetivam descrever a imagem através das orientações mais comuns ao redor deles e das magnitudes dessas orientações.

O método foi desenvolvido em duas fases. A primeira fase envolveu a criação de uma base de descritores locais, gerada a partir da aplicação do algoritmo SIFT em um grupo de imagens de armas de fogo. Esta base posteriormente foi otimizada para melhorar a performance do método. A segunda fase envolve o processo de monitoramento do ambiente

interno do veículo em busca da presença de armas de fogo. Nesse momento os descritores locais extraídos das imagens do interior do veículo são comparados com os descritores da base otimizada gerada na fase anterior, como o objetivo de encontrar similaridades que indiquem a presença de armas.

O processo de correspondência ou *matching* consiste em encontrar pontos semelhantes entre as imagens que geraram a base de descritores e a imagem que foi extraída do interior do veículo. Isso é feito analisando a proximidade entre os descritores locais destes pontos, essa proximidade é medida através do cálculo da distância Euclidiana.

O processo de otimização da base de descritores alcançou uma redução de 82.34% no tamanho da base, que passou de 185.4 Mb para 32.2 Mb. Durante a fase experimental o método desenvolvido utilizando descritores locais obteve resultados na média acima de 80%. Destaca-se ainda os resultados obtidos utilizando a base de imagens gentilmente cedida por Tiwari e Verma (2015a), foram avaliados os mesmos cenários do seu trabalho: a) Diferentes fundos; b) Variação na iluminação; c) Armas diferentes; d) Oclusão; e) Variação na escala e rotação; f) Múltiplas armas, e em todos eles os resultados do método desenvolvido foram superiores.

5.2 Limitações do trabalho

5.2.1 Ruído na imagem

Durante o experimento II foi detectado que o método teve uma redução nos valores absolutos e um aumento da variabilidade nas suas métricas. Isso ocorreu quando foi acrescentado ruído na imagem, sendo que para valores de $\sigma = 120$ a maioria das métricas ficou abaixo de 80% e a sensibilidade chegou a 56%. A principal causa no aumento do ruído na imagem é o aumento da temperatura no interior do veículo. Pelo gráfico da Figura 50 para que o método opere com os valores das métricas acima de 80% o valor de $\sigma < 40$ deve ser obedecido, pois valores acima deste comprometem a viabilidade prática do método.

5.2.2 Movimentos bruscos

A velocidade da movimentação dentro do ambiente interno do veículo pode gerar falhas no processo de identificação em função da captura de imagens borradas e sem nitidez. Isso pode ser resolvido desenvolvendo um método para realizar o descarte dessas imagens, evitando que passem pelo processo de identificação ou investindo em uma câmera com mais recursos para tratar esse problema.

5.3 Trabalhos futuros

Como trabalhos futuros temos a possibilidade de aprimorar o processo de otimização da base de descritores locais, para com isso melhorar a performance geral do método. A detecção de pontos chave é um processo que também pode ser trabalhado no sentido de gerar uma menor variabilidade nas suas características, utilizando para isso pontos adequadamente escolhidos. A integração do método com sistemas de segurança embarcados em veículos, também é um ponto que deve ser visto.

5.4 Considerações finais

Este trabalho é uma pequena contribuição para o desenvolvimento de soluções para o problema da violência urbana. A integração do método de detecção de armas de fogo aos modernos sistemas de segurança dos veículos atuais, pode tornar-se uma funcionalidade que já venha como um opcional de fábrica, garantindo mais segurança aos motoristas e usuários do transporte público.

Por outro lado, esse trabalho de pesquisa incentiva o desenvolvimento de novas tecnologias e de produtos inovadores, promovendo a criação de empresas de base tecnológica e favorecendo o crescimento da economia local. Por fim, entende-se que este trabalho é apenas um começo, uma etapa inicial a partir da qual novas pesquisas sobre temas correlacionados serão derivadas e desenvolvidas com o auxílio do método e das técnicas aqui desenvolvidas.

Referências¹

- AHMED, T.; UPPAL, M.; MUHAMMAD, A. Improving efficiency and reliability of gunshot detection systems. In: IEEE. *Acoustics, Speech and Signal Processing (ICASSP), 2013 IEEE International Conference on*. [S.l.], 2013. p. 513–517. Citado 2 vezes nas páginas 19 e 41.
- BAVYA, R.; MOHANAMURALI, R. Next generation auto theft prevention and tracking system for land vehicles. In: *International Conference on Information Communication and Embedded Systems (ICICES2014)*. [S.l.: s.n.], 2014. p. 1–5. Citado 2 vezes nas páginas 17 e 18.
- BHATNAGAR, G.; WU, Q. J. Human visual system based framework for concealed weapon detection. In: IEEE. *Computer and Robot Vision (CRV), 2011 Canadian Conference on*. [S.l.], 2011. p. 250–256. Citado 2 vezes nas páginas 35 e 36.
- BISHOP, C. M. Pattern recognition. *Machine Learning*, v. 128, 2006. Citado na página 27.
- BROWN, M.; LOWE, D. G. *Invariant Features from Interest Point Groups*. 2002. Citado na página 26.
- CHEN, H.-M. et al. Imaging for concealed weapon detection: a tutorial overview of development in imaging sensors and processing. *IEEE signal processing Magazine*, IEEE, v. 22, n. 2, p. 52–61, 2005. Citado 2 vezes nas páginas 18 e 36.
- COUTINHO; MAROQUIM. *SIFT - Scale Invariant Feature Transform*. 2010. Citado na página 18.
- DAS, P. D.; SENGUPTA, S. Proposing the systems to provide protection of vehicles against theft and accident. In: *2016 IEEE International Conference on Recent Trends in Electronics, Information Communication Technology (RTEICT)*. [S.l.: s.n.], 2016. p. 1681–1685. Citado 2 vezes nas páginas 17 e 18.
- DEVER, J.; LOBO, N. da V.; SHAH, M. Automatic visual recognition of armed robbery. In: IEEE. *Pattern Recognition, 2002. Proceedings. 16th International Conference on*. [S.l.], 2002. v. 1, p. 451–455. Citado 2 vezes nas páginas 19 e 43.
- FANG, X. et al. Automatic gui test by using sift matching. *China Communications*, IEEE, v. 13, n. 9, p. 227–236, 2016. Citado na página 18.
- HÖFERLIN, B.; HEIDEMANN, G. *Selection of an Optimal Set of Discriminative and Robust Local Features with Application to Traffic Sign Recognition*. [S.l.]: Václav Skala-UNION Agency, 2010. Citado na página 18.
- HUANG, M.; MU, Z.; ZENG, H. Efficient image classification via sparse coding spatial pyramid matching representation of sift-wcs-ltp feature. *IET Image Processing*, IET, v. 10, n. 1, p. 61–67, 2016. Citado na página 18.
- HUNG, M.-H.; PAN, J.-S. A real-time action detection system for surveillance videos using template matching. *Journal of Information Hiding and Multimedia Signal Processing*, v. 6, n. 6, p. 1088–1099, 2015. Citado 4 vezes nas páginas 19, 41, 42 e 43.

¹ De acordo com a Associação Brasileira de Normas Técnicas. NBR 6023.

- HUSSEIN, N. J.; HU, F. An alternative method to discover concealed weapon detection using critical fusion image of color image and infrared image. In: IEEE. *Computer Communication and the Internet (ICCCI), 2016 IEEE International Conference on*. [S.l.], 2016. p. 378–383. Citado 2 vezes nas páginas 18 e 36.
- IMFDB. *the Internet Movie Firearms Database*. 2017. Disponível em: <<http://www.imfdb.org/wiki/>>. Citado 5 vezes nas páginas 10, 47, 52, 55 e 62.
- KAPILEVICH, B. et al. Portable passive millimeter-wave sensor for detecting concealed weapons and explosives hidden on a human body. *IEEE Sensors Journal*, IEEE, v. 13, n. 11, p. 4224–4228, 2013. Citado 3 vezes nas páginas 18, 19 e 39.
- KARAMI, E.; PRASAD, S.; SHEHATA, M. *Image Matching Using SIFT, SURF, BRIEF and ORB: Performance Comparison for Distorted Images*. 2010. Citado na página 18.
- KIKTOVA, E. et al. Gun type recognition from gunshot audio recordings. In: IEEE. *Biometrics and Forensics (IWBF), 2015 International Workshop on*. [S.l.], 2015. p. 1–6. Citado 4 vezes nas páginas 19, 39, 40 e 41.
- KIM, W.-G. et al. System design of focal plane array based millimeter-wave imaging radiometer for concealed weapon detection. In: IEEE. *Geoscience and Remote Sensing Symposium (IGARSS), 2011 IEEE International*. [S.l.], 2011. p. 2258–2261. Citado 4 vezes nas páginas 18, 19, 36 e 37.
- KUMAR, B.; UPADHYAY, R.; SINGH, D. Critical analysis of signal processing techniques for concealed weapon identification with mmw (60 ghz) imaging radar system. In: IEEE. *Industrial and Information Systems (ICIIS), 2015 IEEE 10th International Conference on*. [S.l.], 2015. p. 463–468. Citado 3 vezes nas páginas 18, 19 e 38.
- LEE, C.; RHEE, C. E.; LEE, H.-J. Complexity reduction by modified scale-space construction in sift generation optimized for a mobile gpu. *IEEE Transactions on Circuits and Systems for Video Technology*, IEEE, 2016. Citado na página 18.
- LI, X. et al. Multi-polarized millimeter-wave imaging for concealed weapon detection. In: IEEE. *Microwave and Millimeter Wave Technology (ICMMT), 2016 IEEE International Conference on*. [S.l.], 2016. v. 2, p. 892–894. Citado 3 vezes nas páginas 18, 19 e 39.
- LI, Y. et al. Sift keypoint removal and injection via convex relaxation. *IEEE Transactions on Information Forensics and Security*, IEEE, v. 11, n. 8, p. 1722–1735, 2016. Citado na página 18.
- LINDBERG, T. Scale-space theory: A basic tool for analysing structures at different scales. In: CITESEER. *Journal of applied statistics*. [S.l.], 1994. Citado na página 23.
- LIU, Z.; ZHANG, A.; LI, S. Vehicle anti-theft tracking system based on internet of things. In: *Proceedings of 2013 IEEE International Conference on Vehicular Electronics and Safety*. [S.l.: s.n.], 2013. p. 48–52. Citado 2 vezes nas páginas 17 e 18.
- LOWE, D. G. *Object recognition from local scale-invariant features*. 1999. 1150–1157 p. Citado 3 vezes nas páginas 18, 21 e 84.
- LOWE, D. G. *Distinctive image features from scale-invariant keypoints*. [S.l.]: Springer, 2004. 91–110 p. Citado 9 vezes nas páginas 18, 22, 23, 24, 25, 26, 30, 32 e 33.

- MA, W. et al. Remote sensing image registration with modified sift and enhanced feature matching. *IEEE Geoscience and Remote Sensing Letters*, IEEE, v. 14, n. 1, p. 3–7, 2017. Citado na página 18.
- MANJUNATH, T. K. et al. Locking and unlocking of theft vehicles using can (theft control system). In: *2013 International Conference on Green High Performance Computing (ICGHPC)*. [S.l.: s.n.], 2013. p. 1–7. Citado 2 vezes nas páginas 17 e 18.
- MEGGIOLARO, M. A. *Medição de deformações elastoplásticas em regiões de concentração de tensões utilizando métodos sem malha e visão computacional*. Tese (Doutorado) — PUC-Rio, 2014. Citado 2 vezes nas páginas 29 e 31.
- MORONEZE, A. O. Uma aplicação da fórmula de taylor, com análise de autovalores da matriz hessiana. 2013. Citado 2 vezes nas páginas 12 e 27.
- MOULDER, W. F. et al. Development of a high-throughput microwave imaging system for concealed weapons detection. In: IEEE. *Phased Array Systems and Technology (PAST), 2016 IEEE International Symposium on*. [S.l.], 2016. p. 1–6. Citado 3 vezes nas páginas 18, 19 e 39.
- ONDROUŠEK, V.; KOLOMAZNÍK, J.; VYTEČKA, M. Recognition of the objects on the conveyor belt using graph matching algorithms. In: IEEE. *Mechatronics-Mechatronika (ME), 2014 16th International Conference on*. [S.l.], 2014. p. 715–720. Citado na página 46.
- PRIVADOS, S.-S. de S. Disponível em: <http://www.susep.gov.br>. Acesso em: mar, 2016. Citado na página 17.
- SHRUTHI, K. et al. Design of an anti-theft vehicle tracking system with a smartphone application. In: *2015 International Conference on Information Processing (ICIP)*. [S.l.: s.n.], 2015. p. 755–760. Citado 2 vezes nas páginas 17 e 18.
- SILVA, F. A. da et al. *Uma Metodologia para Detectar e Reconhecer Placas de Sinaliza de Transito*. 2012. Citado na página 18.
- SILVA, R. L. d. S. *Indexação de Faces em Estruturas de Dados Métricas*. [S.l.]: UFIJ, 2012. Citado 4 vezes nas páginas 22, 27, 29 e 52.
- SOUZA, C. Análise de poder discriminativo através de curvas roc. *Disponível em: http://crsouza.blogspot.com/2009/07/analise-de-poder-discriminativo-atraves.html* Acesso em: Maio de, 2010. Citado na página 61.
- SREEDEVI, A. P.; NAIR, B. S. S. Image processing based real time vehicle theft detection and prevention system. In: *2011 International Conference on Process Automation, Control and Computing*. [S.l.: s.n.], 2011. p. 1–6. Citado 2 vezes nas páginas 17 e 18.
- SUGA, A. et al. Object recognition and segmentation using sift and graph cuts. In: IEEE. *Pattern Recognition, 2008. ICPR 2008. 19th International Conference on*. [S.l.], 2008. p. 1–4. Citado na página 18.
- TEIXEIRA, R. *Introdução aos espaços de escala: EDPs em processamento de imagens*. [S.l.]: IMPA, 2001. Citado na página 22.

THOME, G. A. Processamento de imagenssegmentação. *UFRJ*, 2010. Citado 2 vezes nas páginas 26 e 27.

TIWARI, R. K.; VERMA, G. K. A computer vision based framework for visual gun detection using harris interest point detector. *Procedia Computer Science*, Elsevier, v. 54, p. 703–712, 2015. Citado 10 vezes nas páginas 9, 10, 15, 43, 44, 45, 46, 75, 76 e 85.

TIWARI, R. K.; VERMA, G. K. A computer vision based framework for visual gun detection using surf. In: IEEE. *Electrical, Electronics, Signals, Communication and Optimization (EESCO), 2015 International Conference on*. [S.l.], 2015. p. 1–5. Citado na página 46.

TRUCCO, E.; VERRI, A. *Introductory techniques for 3-D computer vision*. [S.l.]: Prentice Hall Englewood Cliffs, 1998. v. 201. Citado na página 17.

UPADHYAY, E. M.; RANA, N. Exposure fusion for concealed weapon detection. In: IEEE. *Devices, Circuits and Systems (ICDCS), 2014 2nd International Conference on*. [S.l.], 2014. p. 1–6. Citado 2 vezes nas páginas 18 e 36.

VELHO, L.; TEIXEIRA, R.; GOMES, J. *Introdução aos Espaços de Escala*. [S.l.]: Escola de Computação, 2000. Citado na página 22.

VOZÁRIKOVÁ, E. et al. Surveillance system based on the acoustic events detection. *Journal of Electrical and Electronics Engineering*, University of Oradea, v. 4, n. 1, p. 255, 2011. Citado 2 vezes nas páginas 19 e 41.

XU, T.; WU, Q. J. Multisensor concealed weapon detection using the image fusion approach. *IET*, 2015. Citado 2 vezes nas páginas 18 e 36.

XUE, Z.; BLUM, R. S. Concealed weapon detection using color image fusion. In: *Proceedings of the 6th International Conference on Information Fusion*. [S.l.: s.n.], 2003. v. 1, p. 622–627. Citado 3 vezes nas páginas 18, 34 e 35.

ZHANG, Y.; MIAO, Z. Object recognition based on orb and self-adaptive kernel clustering algorithm. In: IEEE. *Signal Processing (ICSP), 2014 12th International Conference on*. [S.l.], 2014. p. 1397–1402. Citado na página 46.