

Universidade Federal do Pampa

Vinícius Bittencourt da Silva

# **Remoção de sombras em vídeos aplicado à vigilância**

Alegrete

2013



Vinícius Bittencourt da Silva

## **Remoção de sombras em vídeos aplicado à vigilância**

Trabalho de Conclusão de Curso apresentado ao Curso de Graduação em Ciência da Computação da Universidade Federal do Pampa como requisito parcial para a obtenção do título de Bacharel em Ciência da Computação.

Orientador: Prof. Me. Alessandro Bof de Oliveira

Alegrete

2013



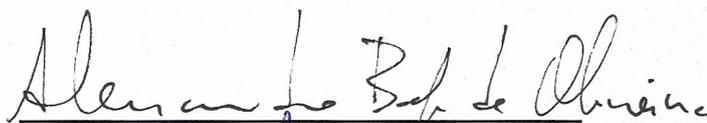
Vinícius Bittencourt da Silva

## Remoção de sombras em vídeos aplicado à vigilância

Trabalho de Conclusão de Curso apresentado ao Curso de Graduação em Ciência da Computação da Universidade Federal do Pampa como requisito parcial para a obtenção do título de Bacharel em Ciência da Computação.

Trabalho de Conclusão de Curso defendido e aprovado em 11 de Outubro de 2013

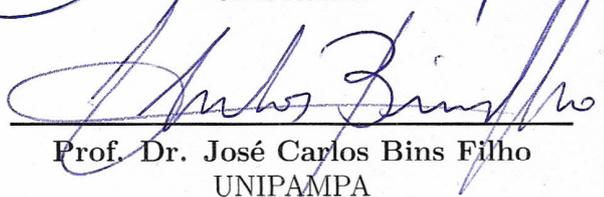
Banca examinadora:



Prof. Me. Alessandro Bof de Oliveira  
Orientador



Prof. Dr. Juliano Fontoura Kazienko  
UNIPAMPA



Prof. Dr. José Carlos Bins Filho  
UNIPAMPA



*Dedico este trabalho à memória de meu avô Horaci Oliveira da Silva.*



# Agradecimentos

Agradeço primeiramente à minha família pelo apoio e carinho, principalmente à minha mãe e ao meu pai que estiveram do meu lado para que eu conseguisse continuar em frente durante todos esses anos e alcançar o objetivo de minha graduação.

Também agradeço ao meu orientador, professor Alessandro, pela compreensão, ensinamentos, e apoio para o desenvolvimento desse trabalho.

Não há como fazer um agradecimento individual à todos que fizeram parte de minha jornada aqui em Alegrete, mas fiz muitas amizades, algumas se desfizeram, outras fortaleceram-se, e essas são as quais eu sei que irei carregar por toda a vida.

Um agradecimento especial ao meu companheiro canino, Scooby, que está ao meu lado à mais de uma década.

Enfim, à todos que fizeram parte disso, direta ou indiretamente, muito obrigado!



*“ Stay hungry, stay foolish.”*  
*(Steve Jobs)*



# Resumo

Em diversas aplicações na área de visão computacional o tratamento de regiões de sombra, ocasionadas pela posição de uma fonte de luz em relação a um objeto de interesse a ser segmentado é necessário. Tal tarefa constitui uma importante etapa de pré-processamento de imagens para uma posterior análise do objeto. O presente trabalho visa estudar o método proposto por Jacques, Jung e Musse (2005) para a segmentação de sombras em sequências de vídeos e propor uma melhoria à detecção dessas regiões com o uso de métodos probabilísticos juntamente com o método correlação de regiões. Inicialmente, gera-se um modelo de fundo a partir do qual, com o uso da subtração de *background* detectar-se-á todos os objetos de interesse que surgirem no cenário. Nesses objetos obtidos aplica-se o algoritmo proposto para a detecção e remoção de sombras com o intuito de obter-se uma detecção desse objeto. Os resultados obtidos com o uso de informação de probabilidade juntamente com a correlação cruzada normalizada mostraram-se visualmente melhores do que o método sem o uso de tal refinamento.

**Palavras-chave:** Segmentação de Sombras. Processamento de Vídeo. Método Probabilístico. Correlação de Regiões.



# Abstract

In many applications in computer vision area, treating shadow regions caused by the source light position relative to an object of interest to be segmented is necessary. This task is an important step of preprocessing of images for further analysis of the object. The present work aims to study the method proposed by [Jacques, Jung e Musse \(2005\)](#) for segmentation of shadows in video sequences and propose an improvement to the detection of these regions using probabilistic methods together with the method for correlation of regions. Initially, it generates a model of background from which, using the background subtraction will detect all objects of interest that arise in the scenario. An algorithm for the detection and removal of shadows is applied to such objects in order to obtain a detection of this object. The results obtained with use of information of probability along with the normalized cross correlation visually proved to be better than the method without the use of such refinement.

**Key-words:** Shadows Segmentation. Video Processing. Probabilistic Methods. Correlation of Regions.



# Lista de ilustrações

Figura 1 – Exemplo de imagem em tons de cinza. . . . .	31
Figura 2 – Cubo do <i>Red Green Blue</i> (RGB). . . . .	32
Figura 3 – Representação geométrica em hexágono do modelo de cores <i>Hue Saturation Value</i> (HSV). . . . .	32
Figura 4 – Espaço de cores <i>Luminance Chroma</i> (YUV). . . . .	33
Figura 5 – Elemento estruturante em forma de diamante. . . . .	34
Figura 6 – Exemplo de obtenção de objetos de <i>foreground</i> . . . . .	37
Figura 7 – Fluxograma do algoritmo de remoção de sombras. . . . .	38
Figura 8 – Fluxograma do algoritmo de detecção de Sun. . . . .	39
Figura 9 – Diagrama de cromaticidade. . . . .	40
Figura 10 – Formação de sombras. . . . .	43
Figura 11 – Detecção de sombras. . . . .	44
Figura 12 – Representação de alguns <i>frames</i> de <i>Walk1.mpg</i> . . . . .	47
Figura 13 – <i>Background</i> com fundo fixo. . . . .	49
Figura 14 – Objeto de interesse no cenário. . . . .	49
Figura 15 – Resultado da subtração de <i>background</i> com fundo fixo. . . . .	50
Figura 16 – Imagem de <i>foreground</i> binarizada. . . . .	51
Figura 17 – Imagem após a remoção da sombra. . . . .	52
Figura 18 – Imagem após o processamento morfológico. . . . .	53
Figura 19 – Refinamento morfológico aprimorado. . . . .	54
Figura 20 – Resultado da geração de <i>background</i> por mediana. . . . .	55
Figura 21 – Resultado da subtração de <i>background</i> binarizada obtida por mediana. . . . .	55
Figura 22 – Remoção de sombra obtida por mediana. . . . .	56
Figura 23 – Objeto final obtido por mediana. . . . .	57
Figura 24 – Objetos detectados em subtração de imagem RGB. . . . .	58
Figura 25 – Regiões de sombra identificadas por <i>Normalized Cross Correlation</i> (NCC). . . . .	58
Figura 26 – Reclassificação de pixels. . . . .	61
Figura 27 – Sequência de identificação de regiões de sombra . . . . .	62
Figura 28 – Fundo fixo em escala de cinza . . . . .	69
Figura 29 – Fundo fixo RGB . . . . .	69
Figura 30 – Sequência em escala de cinza 1 - <i>Frame</i> 76 . . . . .	70
Figura 31 – Sequência em escala de cinza 2 - <i>Frame</i> 316 . . . . .	70
Figura 32 – Sequência em escala de cinza 3 - <i>Frame</i> 137 . . . . .	70
Figura 33 – Sequência escala de cinza 4 - <i>Frame</i> 46 . . . . .	71
Figura 34 – Quadro do vídeo e segmentação de objetos de <i>foreground</i> . . . . .	71

Figura 35 –NCC x Refinamento Probabilístico . . . . .	72
Figura 36 –Objeto final: Operadores morfológicos aplicados à NCC e Refinamento probabilístico . . . . .	72
Figura 37 –Quadro do vídeo e segmentação de objetos de <i>foreground</i> . . . . .	72
Figura 38 –NCC x Refinamento Probabilístico . . . . .	73
Figura 39 –Objeto final: Operadores morfológicos aplicados à NCC e Refinamento probabilístico . . . . .	73
Figura 40 –Quadro do vídeo e segmentação de objetos de <i>foreground</i> . . . . .	73
Figura 41 –NCC x Refinamento Probabilístico . . . . .	74
Figura 42 –Objeto final: Operadores morfológicos aplicados à NCC e Refinamento probabilístico . . . . .	74
Figura 43 –Quadro do vídeo e segmentação de objetos de <i>foreground</i> . . . . .	74
Figura 44 –NCC x Refinamento Probabilístico . . . . .	75
Figura 45 –Objeto final: Operadores morfológicos aplicados à NCC e Refinamento probabilístico . . . . .	75

# Lista de tabelas

Tabela 1 – Informação técnica do vídeo <i>Walk1.mpg</i> . . . . .	47
Tabela 2 – Valores de pixels . . . . .	60



# Lista de siglas

**AVI** Audio Video Interleave

**HSI** *Hue Saturation Intensity*

**HSV** *Hue Saturation Value*

**ITS** *Intelligent Transport System*

**MPG** Moving Picture Group

**NCC** *Normalized Cross Correlation*

**RGB** *Red Green Blue*

**YUV** *Luminance Chroma*



# Sumário

<b>1</b>	<b>Introdução</b>	<b>23</b>
<b>2</b>	<b>Objetivos</b>	<b>27</b>
<b>3</b>	<b>Conceitos Gerais e Revisão da Literatura</b>	<b>29</b>
3.1	Objetos de Interesse (Objetos de <i>Foreground</i> )	29
3.2	Modelo de Fundo	29
3.3	Subtração de <i>Background</i>	30
3.4	Sistemas de Cores	30
3.5	Morfologia Matemática e Elementos Estruturantes	33
3.6	Correlação Cruzada Normalizada (NCC)	35
3.7	Trabalhos Relacionados	36
<b>4</b>	<b>Apresentação da Pesquisa e Análise dos Resultados</b>	<b>45</b>
4.1	Abordagem com <i>Background</i> Gerado Por Fundo Fixo	48
4.2	Abordagem com <i>Background</i> Gerado por Mediana	54
4.3	Abordagem com uso de Informações de RGB	57
4.4	Refinamento por Meio de Informação de Probabilidade	59
<b>5</b>	<b>Considerações Finais</b>	<b>63</b>
	<b>Referências</b>	<b>65</b>
	<b>Apêndices</b>	<b>67</b>
	<b>APÊNDICE A RESULTADOS</b>	<b>69</b>
	<b>Índice</b>	<b>77</b>



# 1 Introdução

A visão é tida como um dos principais sentidos do seres humanos. A captura de imagens por meio de câmeras, visa simular essa percepção sensorial. A obtenção dessas imagens abriu caminho para muitas aplicações em visão computacional.

Atualmente as câmeras possuem uma ótima qualidade tanto para captura de imagens, assim como para a geração de vídeos, além de, capacidade de armazenamento interno cada vez maiores e com preços mais acessíveis.

Quando utilizadas para observar constantemente determinada região, as câmeras, são conhecidas como câmeras de vigilância (muitas vezes também chamadas de câmeras de monitoramento). Tais ferramentas têm a capacidade de manter um registro de toda a movimentação de determinado local.

Se algo indevido ou inesperado ocorrer em uma região sob o monitoramento de uma câmera de vigilância, tem-se um registro do fato, além de que, talvez o mesmo possa ser evitado (uma vez que haja um acompanhamento a partir de uma sala de monitoramento). Esse tipo de situação, segundo [Bovik \(2009\)](#)

“está estimulando um grande interesse em desenvolver sistemas baseados em vídeos capazes de observar e interpretar cenas específicas, com objetivo de detectar anomalias ou eventos que podem afetar a segurança, economia, ou outros aspectos vitais da atividade humana” (BOVIK, Alan Conrad. *The Essential Guide To Video Processing*. Austin, Texas, 2009, p. 141).

Para ser viável a obtenção de uma capacidade de interpretação, há a necessidade de que as máquinas sejam construídas para ver, ou seja, têm-se que gerar para tais instrumentos a capacidade de possuir uma visão semelhante ao sentido dos seres humanos.

[Shapiro e Stockman \(2001\)](#) afirmam que máquinas podem ser construídas para ver, e que Alan Turing (tido como um dos pais da computação) acreditava que um computador poderia alcançar a inteligência e habilidade para interpretar cenas.

Afirma [Bovik \(2009\)](#), que

“esforços nessa área tem feito com que os sistemas para a detecção de anomalias tenham evoluído bastante na questão do hardware, além de possuir softwares cada vez de melhor qualidade, assim como câmeras melhores, e redes de sensores visuais.” (BOVIK, Alan Conrad. *The Essential Guide To Video Processing*. Austin, Texas, 2009, p. 141).

O primeiro passo rumo a um sistema de visão computacional que possua a habilidade de interpretação de cenas é prover a capacidade de identificação precisa de objetos

contidos na mesma.

Porém, uma vez realizada a detecção dos objetos de interesse contidos nas sequências de imagens obtidas pelas câmeras, esses, podem trazer agregados à si, a sua informação de sombra. No âmbito da visão computacional, o tratamento das informações de sombras no pré-processamento das imagens obtidas é algo fundamental uma vez que deseja-se fazer uma análise futura mais precisa dos objetos.

Informações de sombra, agregadas ao objeto de interesse, o deformam de forma a dificultar a identificação do formato do mesmo, assim como podem ocasionar no agrupamento de dois ou mais objetos.

Segundo [Matsushita \(2002\)](#),

“Sistemas de vigilância por vídeo envolvendo detecção e rastreamento de objetos requerem robustez contra mudanças de iluminação causados pela variação de, por exemplo, condições climáticas. Obstáculos importunos incluem não só a alteração das condições de iluminação, mas também as grandes sombras projetadas pelas estruturas vizinhas, como grandes edifícios e árvores altas.” ([MATSUSHITA, Yasuyuki, Shadow elimination and interpolation for computer vision and graphics, Tóquio, 2002 p. 24](#)).

A variação da iluminação do ambiente, como citado anteriormente, é um contra-tempo para a identificação de objetos de forma precisa, uma vez que o algoritmo deve ser autoadaptativo às diversas condições de iluminação para tornar-se robusto.

Inicialmente é realizada a detecção ou a segmentação dos objetos em determinada cena. [Ponce et al. \(2011\)](#), diz que o termo segmentação é utilizado para diversas atividades, porém, que para todas a motivação é a mesma: obter uma representação compacta do que é de interesse no âmbito de uma imagem.

Após a obtenção de uma segmentação de objetos, é viável aplicar um dos métodos descritos na literatura para realizar a identificação e remoção de áreas nas quais haja presença de sombra obtendo uma identificação robusta do objeto.

O uso de técnicas de remoção de sombras é necessário em aplicações de diversas áreas, como por exemplo, sistemas de monitoramento de tráfego e sistemas para a detecção de pessoas em cenários de vídeo. Essas sombras são removidas para que não haja uma detecção errônea dos objetos de interesse, como agrupamentos e deformidades no seu contorno, o que pode ser crucial em algumas aplicações. ([WANG et al., 2007](#)) ([SUN; LI, 2010](#)) ([YUAN; YANG; XU, 2010](#)) ([SANIN; SANDERSON; LOVELL, 2010](#)) ([JACQUES; JUNG; MUSSE, 2005](#)).

Uma vez sabendo da importância da remoção de sombras em sequências de vídeos, o presente trabalho propõe a implementação do método descrito por [Jacques, Jung e Musse \(2005\)](#), e a geração de um refinamento com o uso de probabilidade à fim de obter melhorias em relação ao método já existente, aprimorando assim a sua aplicabilidade.

O restante do trabalho está organizado na seguinte sequência:

- Capítulo 2: Descreve de forma mais precisa cada um dos objetivos trabalhados.
- Capítulo 3: Mostra os conceitos gerais juntamente com uma revisão da literatura sobre alguns dos métodos existentes focados à remoção de sombras em sequências de vídeos.
- Capítulo 4: Demonstra a pesquisa juntamente com a análise dos resultados obtidos.
- Capítulo 5: Apresenta uma conclusão do trabalho.



## 2 Objetivos

Com o intuito de obter um algoritmo de remoção de sombras em sequências de vídeo que apresente melhorias em relação ao método da literatura proposto por [Jacques, Jung e Musse \(2005\)](#) o presente trabalho visa alcançar os seguintes objetivos:

- Implementar um algoritmo de detecção de objetos em vídeo e remoção de regiões de sombra contidas nesses objetos.
- Testar a aplicabilidade do algoritmo desenvolvido, verificando visualmente sua eficiência na segmentação de regiões de sombras. Tais testes dar-se-ão em vídeos obtidos por câmeras de vigilância.
- Identificar objetos de interesse através da diferença entre modelo de fundo e os demais quadros do vídeo.
- Realizar a identificação das regiões de sombras, utilizando a técnica de correlação de regiões, nos objetos de interesse obtidos através da segmentação.
- Aplicar testes com o uso informações probabilísticas a fim de obter um melhor refinamento dos objetos segmentados.
- Efetuar comparações entre o método da literatura de [Jacques, Jung e Musse \(2005\)](#) em relação ao refinamento proposto, objetivando demonstrar a eficácia do método desenvolvido.



## 3 Conceitos Gerais e Revisão da Literatura

A detecção e remoção de regiões de sombras compõe um elemento muito importante na área de visão computacional uma vez que através dela é possível obter-se uma detecção mais precisa de objetos de interesse em sequências de vídeos. O presente capítulo é organizado da seguinte forma:

- Seção 3.1: Busca descrever o que são os objetos de interesse.
- Seção 3.2: Define o que é modelo de fundo e destaca algumas das técnicas para a obtenção do mesmo.
- Seção 3.3: Demonstra a utilização da subtração de *background*.
- Seção 3.4: Apresenta os principais sistemas de cores utilizados nos trabalhos relacionados (e nesse trabalho).
- Seção 3.5: Mostra a definição de morfologia matemática e descreve o que são elementos estruturantes.
- Seção 3.6: Visa elucidar o conceito de correlação cruzada normalizada, técnica utilizada posteriormente para remoção de sombras em vídeos.
- Seção 3.7: Abrange o conteúdo da literatura especializada a fim de apresentar alguns dos principais trabalhos relacionados à remoção de sombras.

### 3.1 Objetos de Interesse (Objetos de *Foreground*)

Define-se como objetos de interesse, também conhecidos como objetos de *foreground*, elementos que não fazem parte de um cenário específico, ou seja, que estão presentes na sequência de vídeo durante apenas um intervalo de tempo, e, provavelmente, são pessoas, carros, animais, etc.

Esses objetos podem possuir agregados à si, a informação de sombra. Tal informação geralmente é indesejada, pois causa distorções na detecção, e identificação do mesmo, assim sendo busca-se a identificação dessas regiões e aplicação de remoção de sombras à esses objetos.

### 3.2 Modelo de Fundo

Muitos trabalhos da literatura especializada em processamento de imagens, voltados à segmentação de objetos descrevem a criação de um modelo de fundo para que

seja possível a obtenção dos objetos de interesse, como em Jacques, Jung e Musse (2005), Yuan, Yang e Xu (2010) e Sanin, Sanderson e Lovell (2010).

Utiliza-se várias técnicas para a geração do modelo de fundo, que nada mais é que o cenário do vídeo, sem elementos de *foreground* presentes.

Dentre as principais técnicas, pode-se citar o uso de modelo de fundo fixo Yuan, Yang e Xu (2010) e o modelo de mediana Jacques, Jung e Musse (2005), entretanto, há muitas outras abordagens para a geração do modelo de fundo.

A modelagem de fundo fixo, é uma modelagem mais simples e se comparada à técnica que faz uso do cálculo da mediana de *pixels*.

A técnica de fundo fixo gera resultados mais imprecisos e maior ruído. Entretanto, a técnica de modelagem de fundo a partir da mediana apesar de ser uma modelagem mais precisa, possui um maior custo computacional agregado.

### 3.3 Subtração de *Background*

Uma vez definido um modelo de fundo, é necessário realizar a segmentação dos objetos de *foreground* para realizar o tratamento das regiões de sombra. A subtração de *background* é um dos métodos mais difundidos para a obtenção desses objetos.

Basicamente, trata-se de uma diferença entre um quadro do vídeo pelo modelo de fundo, geralmente com o uso de um limiar para criar uma separação entre o que é pertencente ao fundo, e o que é objeto de interesse.

### 3.4 Sistemas de Cores

A forma pela qual as câmeras de vídeo obtêm as imagens de um ambiente podem variar. Entre algumas variações estão a resolução (tamanho) do quadro capturado, as cores, e também a forma que essas imagens são armazenadas.

Existem diversos modelos de cores definidos para imagens digitais, entretanto pode-se citar entre os principais modelos, o modelo em escala de cinza, o modelo RGB, o modelo HSV, e o modelo YUV. É fundamental um conhecimento sobre os diversos modelos de cores, pois eles interferem diretamente no modo em que trabalha-se com a imagem.

Em Davies (2004) é levantado o questionamento sobre o tipo de imagem a ser usado em processamento de imagens. É comparado o uso do sistema de cores no padrão RGB com o modelo em escala de cinza e chega-se a constatação de que necessita-se verificar o âmbito ao qual o processamento é direcionado.

Imagens coloridas trazem a necessidade de um maior armazenamento pois são maiores, por conterem a informação de três faixas de cores, e conseqüentemente um maior processamento agregado. Por outro lado, imagens coloridas podem ser muito úteis em algumas aplicações, como por exemplo na medicina em imagens obtidas durante uma cirurgia.

Para o presente trabalho, verifica-se que o uso de imagens em escala de cinza inicialmente são suficientes para alcançar os primeiros objetivos propostos pois o método de Jacques, Jung e Musse (2005), trabalha sob esse sistema de cores.

Pode-se observar em Russ (2006), a definição de resolução em tons de cinza como a impressão ou exibição de uma imagem, a partir de um número de 8 bits, cujo os valores vão de 0 à 255. Apesar de existirem outras escalas maiores o uso de imagens que utilizam tal escala é aplicável ao método de Jacques, Jung e Musse (2005) a partir do qual posteriormente serão gerados os refinamentos.



Figura 1 – Exemplo de imagem em tons de cinza.

Fonte: <http://www.tecgraf.puc-rio.br/mgattass/fcg/trb13/DavidsonFellipe/t1.html>

Os refinamentos propostos por esse trabalho, utilizam o espaço de cores RGB (vermelho, verde e azul, respectivamente), que é um dos mais difundidos, e é geralmente representado por três círculos aninhados, onde cada um representa uma das bandas de cores. Outra possível representação para esse modelo é um cubo cuja as coordenadas de seus eixos representam os valores dos tons de cores.

Em Souto (2003), é demonstrado que as cores em tom de cinza dentro do RGB são representadas pela diagonal entre os vértices que representam as cores branco e preto, como pode ser visto na Figura 2 de Envi (2000).

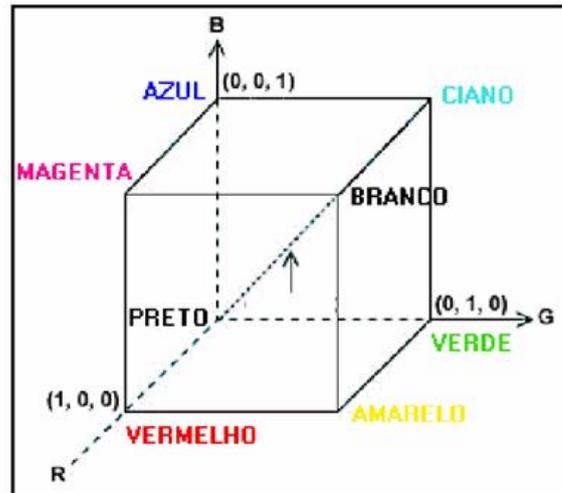


Figura 2 – Cubo do RGB.

Fonte: Envi (2000)

Outro modelo de cores bastante utilizado é o HSV que pode ser representado pela Figura 3, onde os eixos  $H$ ,  $S$  e  $V$  representam, a tonalidade (indo de valores de  $0^\circ$  até  $360^\circ$ ), a saturação (com valores de  $0\%$  à  $100\%$ ) o brilho (com valores de  $0\%$  à  $100\%$ ), respectivamente.

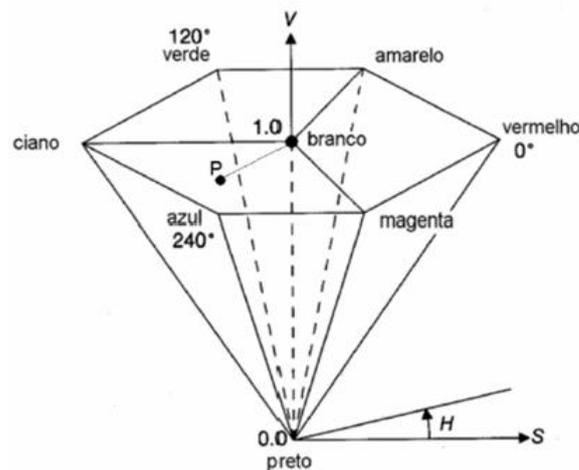


Figura 3 – Representação geométrica em hexágono do modelo de cores HSV.

Fonte: Envi (2000)

Assim como os sistemas de cores descritos anteriormente, o sistema de cores YUV é bastante difundido na literatura. Tal sistema é composto pelo componente  $Y$  (cujo os valores vão de 0 à 255) que é a representação do brilho da cor (luminância), e o componentes  $U$  e  $V$  (que possuem valores que vão de  $-128$  à  $127$  ou  $0$  à  $255$ , dependendo

se a codificação permite valores negativos, ou não) e que representam a cromaticidade.



Figura 4 – Espaço de cores YUV.

Fonte: <http://softpixel.com/~cwright/programming/colorspace/yuv/>

A Figura 4, baseada em [Whright \(2013\)](#), exibe de forma visual o sistema de cores YUV, onde 4(a) define o cubo que contém todas as cores representadas pelo sistema YUV, e 4(b) e 4(c), são representações do cubo particionado, a fim de demonstrar as cores contidas em seu interior.

### 3.5 Morfologia Matemática e Elementos Estruturantes

A morfologia matemática tem base na geometria dos elementos de uma imagem. É uma forma de análise da forma de um objeto, extraindo as informações relativas à geometria da imagem.

O foco da morfologia no entanto é aplicar operações sobre conjuntos de pixels em uma imagem utilizando a teoria dos conjuntos.

Através do uso de elementos estruturantes, que podem ser definidos como conjuntos matemáticos bem definidos, é possível gerar transformações na forma de uma imagem para obter um melhor realce, realizar segmentações, entre diversas outras operações.

Um elemento estruturante nada mais é que um conjunto de coordenadas de pixels e pode possuir diversos formatos, diamante, caixa, cruz, entre outros.

A forma de um elemento estruturante define a quantidade de pixels que serão trabalhados nas operações sobre esses conjuntos, por exemplo o elemento em forma de cruz de tamanho 1 possui as coordenadas de pixels denotadas pelo o conjunto definido por  $C = \{(0, 0), (-1, 0), (1, 0), (0, -1), (0, 1)\}$ , assim sendo, o valor do elemento estruturante define o deslocamento dele sobre os eixos da imagem.

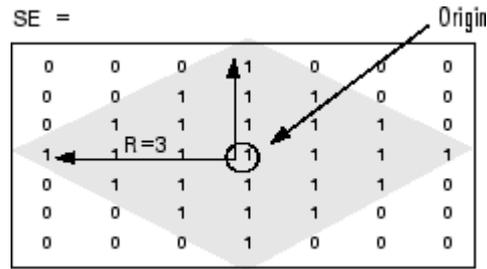


Figura 5 – Elemento estruturante em forma de diamante.

Fonte: <http://www.mathworks.com/>

Na Figura 5, há um modelo de representação de um elemento estruturante com o formato diamante, de tamanho 3. Nota-se visualmente que a partir do *pixel* de origem há um deslocamento no eixo x, e no eixo y, tanto para a direita quanto para a esquerda, para cima e para baixo, gerando assim uma vizinhança de 3 pixels em cada um dos eixos.

Através de operadores morfológicos, que utilizam como parâmetro o elemento estruturante é possível preencher áreas de uma imagem que apresentem falhas, assim como, remover *pixels* que apresentem-se isolados que podem estar compondo ruído em uma imagem final.

Os principais operadores morfológicos são os operadores de erosão e de dilatação. Existem muitos outros operadores, dentre eles é importante dar ênfase aos operadores de abertura e fechamento que são derivações dos operadores de dilatação e erosão.

Em Wangnhein, Izidoro e Almeida (2013) tem-se a definição da operação de erosão como a interseção de dois conjuntos de valores, ou seja, a erosão basicamente encolhe a região de uma imagem.

A dilatação é definida como a combinação de dois conjuntos, e expande a região de uma imagem. A dilatação e a erosão não são necessariamente operações inversas, somente em algumas situações isso ocorre.

Davies (2004) diz que operações morfológicas de fechamento são capazes de eliminar elementos em uma imagem, dentre eles ruídos e pequenos buracos através do fechamento de falhas.

Os operadores de abertura tendem a suavizar uma imagem, são utilizados para remover ruídos, além de serem capazes de eliminar da imagem pequenas saliências, com a criação de buracos.

Abertura e fechamento são operações contrárias na sua formação. Uma erosão, seguida de uma dilatação usando o mesmo elemento estruturante é uma operação de abertura, e uma operação de fechamento é a aplicação de uma dilatação seguida de uma

erosão.

Combinações entre operadores morfológicos de abertura e fechamento permitem que feche-se buracos internos em uma imagem sem perder uma quantidade relevante da informação do contorno pois o que for distorcido pelo fechamento, geralmente, pode ser corrigido pelo operador de abertura. Assim sendo, com o uso sequencial de ambos, são preenchidas lacunas e a estrutura do objeto é mantida o mais próximo possível do seu estado original.

### 3.6 Correlação Cruzada Normalizada (NCC)

Carvalho e Tavares (2005) denotam que correlação cruzada é a comparação de uma imagem com uma sub imagem, e dessa relação exibir a semelhança entre as duas de forma numérica.

O algoritmo de correlação cruzada normalizada NCC, provê diferentes escalas de um mesmo sinal, ou seja, se no quadro do vídeo há uma variação apenas na intensidade da iluminação em determinada área, tal região será possivelmente uma região de sombra. Assim sendo, poderá haver uma variação do sinal em relação à imagem de fundo, porém, fortemente correlacionada. Entretanto, se na região houvesse um objeto de *foreground*, esse possuiria um sinal cuja a correlação entre os quadros do vídeo seria baixa, e tal região não poderia ser caracterizada como sombra.

Em Jacques, Jung e Musse (2005), o algoritmo de NCC é definido pela seguinte fórmula:

$$NCC(i, j) = \frac{ER(i, j)}{E_B(i, j)E_{T_{ij}}},$$

Onde:

$$ER(i, j) = \sum_{n=-N}^N \sum_{m=-N}^N B(i+n, j+m)T_{ij}(n, m),$$

$$E_B(i, j) = \sqrt{\sum_{n=-N}^N \sum_{m=-N}^N B(i+n, j+m)^2}, \text{ e}$$

$$E_{T_{ij}} = \sqrt{\sum_{n=-N}^N \sum_{m=-N}^N T_{ij}(n, m)^2}.$$

Para classificar um *pixel* na posição  $(i, j)$  da imagem como sombra, é calculado sua vizinhança de pixels para um *template*  $T$ . Esse *template* deve possuir correlação alta e a energia da imagem naquela região, definida na fórmula descrita anteriormente por  $E_{T_{ij}}$ , deve ser menor que a energia daquela mesma região na imagem de *background*, representado por  $E_B(i, j)$ .

O resultado do quociente definido pela fórmula, será utilizado para classificação do *pixel*. Para os pixels de sombra o valor resultante do cálculo das energias em uma determinada região será próximo à 1 (uma vez que o algoritmo de NCC vai de -1 à 1, sendo 1 o sinal altamente correlacionado, e -1 o sinal com pouca correlação).

Baseando-se nas informações obtidas pelo resultado da correlação de regiões, Jacques, Jung e Musse (2005) geram uma classificação inicial dos pixels candidatos à sombra.

Após a detecção de regiões de sombras e a remoção das mesmas, geralmente é realizado o refinamento do objeto final obtido, a fim de preencher buracos e remover ruídos. Para tal tarefa, é comum o uso de elementos estruturantes juntamente com morfologia matemática, para aprimorar o resultado.

### 3.7 Trabalhos Relacionados

Devido a importância da remoção de sombras em sequências de vídeo, vários trabalhos descritos na literatura especializada de processamento de imagens abordam esse tema. Essa Seção visa demonstrar alguns dos principais trabalhos na área e a metodologia abordada por cada um para lidar com o problema de sombras.

No trabalho de Jacques, Jung e Musse (2005), é realizada a detecção de sombras, assim como a criação de um modelo de fundo, em sequências de vídeo em escala de cinza.

A abordagem de Jacques, Jung e Musse (2005), primeiramente realiza a modelagem da imagem do fundo. Essa modelagem proposta é uma melhoria ao modelo W4 de Haritaoglu, Harwood e Davis (2000), gerada a partir de um intervalo de tempo do vídeo, onde pixels em movimento são excluídos da imagem final obtida, criando assim uma imagem representando apenas o ambiente.

Para a geração do modelo de fundo, é aplicado em cerca de 100 quadros do vídeo, um filtro de mediana, para separar pixels estáticos, de pixels que pertencem a objetos que estão em movimento no cenário. Somente após o emprego desse filtro os pixels estáticos são processados, e a partir desse processamento forma-se o modelo de fundo.

Uma vez obtido um modelo de fundo, cada quadro do vídeo é comparado individualmente com o mesmo. Os objetos de *foreground* são identificados através de uma subtração entre modelo de fundo e um quadro do vídeo (subtração de *background*).

A Figura 6, de Jacques, Jung e Musse (2005), demonstra a segmentação de objetos



Figura 6 – Exemplo de obtenção de objetos de *foreground*.

Fonte: Jacques, Jung e Musse (2005)

através do método de subtração descrito anteriormente.

Em 6(a) é gerado e exibido o modelo de fundo (*background*), já 6(b) representa um quadro aleatório do vídeo, o qual difere do modelo de fundo uma vez que possui objetos de *foreground*. Pode-se observar em 6(c) a obtenção de um novo quadro, esse, resultante da diferença entre quadro aleatório 6(b) e o fundo obtido em 6(a), e representa os objetos de interesse a serem tratados.

No caso da existência de objetos já segmentados, realiza-se o tratamento das regiões de sombra. Essas regiões são detectadas através da utilização do algoritmo de correlação cruzada normalizada NCC de Jacques, Jung e Musse (2005), citado anteriormente.

Uma vez realizada a detecção das regiões de sombra é proposto um refinamento nessa detecção por meio de uma fórmula matemática, e por fim, é realizado a reconstrução morfológica da imagem. No trabalho de Jacques, Jung e Musse (2005), é possível observar o uso de operadores de abertura e fechamento com a aplicação de um elemento estruturante do tipo diamante de tamanho 5 para esse fim.

A proposta de Wang et al. (2007) é voltada para a aplicação da remoção de regiões de sombras à *Intelligent Transport System (ITS)*. Sistemas inteligentes de transportes, buscam obter uma identificação de veículos como intuito de realizar o seu rastreamento.

As sombras tornam complexo a obtenção do posicionamento real dos objetos contidos em uma sequência de *frames*, uma vez que tal objeto passa a ser o veículo somado à sua sombra.

Para a detecção de sombras, é descrito um modelo combinatório entre a detecção baseada em propriedade e detecção baseada em modelo. Detecção baseada em propriedade engloba métodos de identificação com o uso de informações como forma, brilho e cor. De outra forma, métodos de identificação por modelo necessitam de informações prévias, como por exemplo a direção da iluminação.

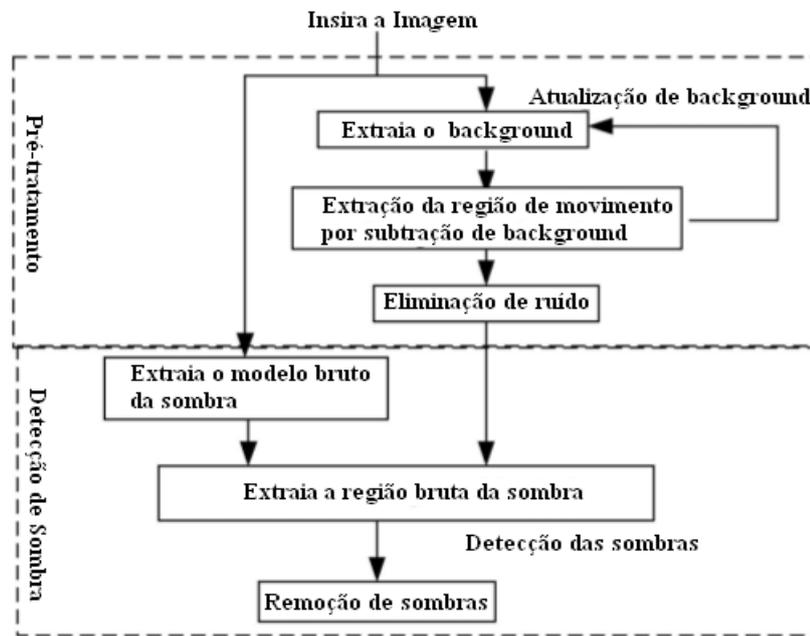


Figura 7 – Fluxograma do algoritmo de remoção de sombras.

Fonte: Wang et al. (2007)

A Figura 7, baseada na imagem de Wang et al. (2007), exibe a sequência de passos executados para alcançar a imagem final pós-tratamento dos objetos. Inicialmente é gerado um modelo de *background* autoadaptativo que consiste em fazer rapidamente o *update* do quadro de fundo, conforme ocorram variações no ambiente.

Para obter a região de movimento é feita a subtração de *background*. A região de movimento é detectada para diminuir a quantidade de cálculos necessários para a obtenção da atualização do *background*. Após é aplicado um cálculo simples para a eliminação de ruído.

$$\text{Sombra}(i, j) = \begin{cases} 1 & Th1 \leq \frac{I_k^v(i, j)}{B_k^v(i, j)} \leq Th2 \\ 0 & \text{Senão} \end{cases}$$

Equação 1

Usando o espaço de cores **HSV** e levando em consideração a informação de luminosidade, é possível definir as regiões da sombra a partir da fórmula descrita na Equação 1.

O *pixel* de *sombra*( $i, j$ ), sendo  $i$  e  $j$  índices de posicionamento do *pixel*, é definido como 1 se o quociente da imagem atual, representado por  $I$ , pela imagem de *background* resultar em valor contido no intervalo dos limiares obtidos durante a detecção de regiões de movimento.

Analogamente à Wang et al. (2007), Sun e Li (2010) focam a detecção e remoção de sombras em sequências de vídeos de monitoramento de tráfego.

A proposta definida por Sun e Li (2010), parte de que a intensidade da cor em áreas sombreadas é inferior a regiões que possuem objetos de *foreground*. Por tal, optou-se pela utilização do modelo de cores *Hue Saturation Intensity* (HSI), pois o mesmo possui os valores de cor e intensidade separados em suas componentes. Primeiramente, apresenta-se uma fórmula matemática a qual realiza a transformação entre os espaços de cores RGB e HSI.

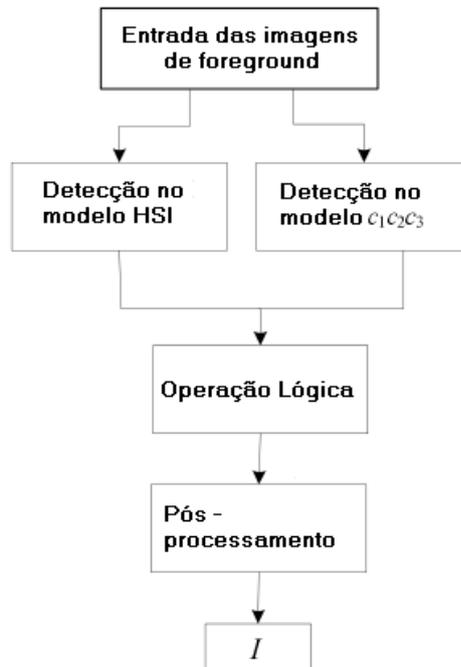


Figura 8 – Fluxograma do algoritmo de detecção de Sun.

Fonte: Sun e Li (2010)

Na Figura 8, baseada no fluxograma de Sun e Li (2010), define-se a maneira através da qual a remoção de sombras é realizada neste modelo. Imagens de *foreground*, são inseridas e então a detecção é obtida através de dois modelos de cores. O modelo HSI, citado anteriormente, e o modelo de cores  $c1c2c3$ .

A fórmula matemática, extraída de Sun e Li (2010), denota a forma de obtenção do modelo

$$\begin{cases} c_1 = \arctan\left[\frac{R}{\max(G, B)}\right] \\ c_2 = \arctan\left[\frac{G}{\max(R, B)}\right], \\ c_3 = \arctan\left[\frac{B}{\max(R, G)}\right] \end{cases}$$

de cores  $c_1c_2c_3$ , onde pra cada tom de cor no modelo RGB, é calculado seu arco-tangente do quociente de uma das bandas pelo valor máximo entre as outras duas.

Através das características espectrais dos modelos de cores obtém-se duas imagens nas quais as regiões de sombra estão definidas, e então aplica-se uma operação lógica de forma à fundir tais imagens em um único objeto.

Para contornar possíveis erros na detecção faz-se um pós-processamento para finalmente obter-se a região de sombra da forma mais apurada possível. Após a realização dos passos definidos obtém-se a imagem  $I$ , que representa a detecção final.

Da mesma forma que os métodos descritos anteriormente, porém com foco na detecção robusta de pessoas em cenários de vigilância, [Sanin, Sanderson e Lovell \(2010\)](#), demonstram um método baseado na informação de cromaticidade e gradiente. Em [Gonzalez e Woods \(2000\)](#), é dado como definição de cromaticidade a junção do matiz com a saturação da cor. Logo, uma cor pode ser caracterizada pelo seu brilho juntamente com a informação de cromaticidade.

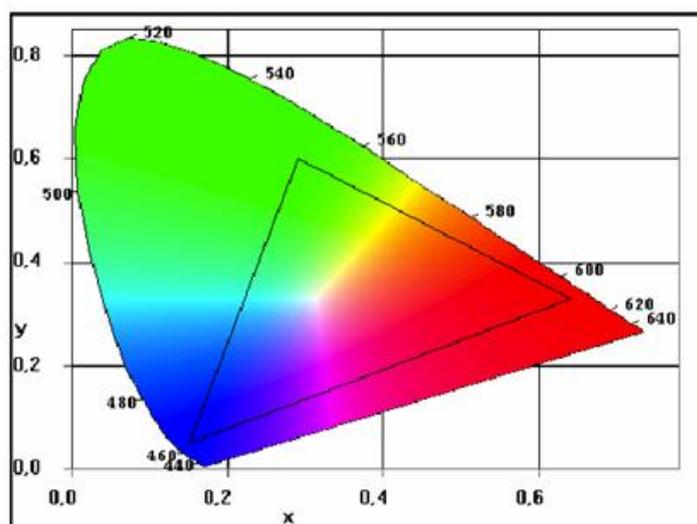


Figura 9 – Diagrama de cromaticidade.

Fonte: [Souto \(2003\)](#)

A Figura 9, de [Souto \(2003\)](#), representa graficamente os chamados coeficientes tricromáticos, o triângulo no interior dessa representação é denominado “*gamut*”, que nada mais é do que um subconjunto de cores, no qual os vértices estão sobre as cores primárias (RGB).

Usando um método de cromaticidade da literatura, [Sanin, Sanderson e Lovell \(2010\)](#) geram uma máscara de pixels candidatos a pertencerem à regiões de sombra. Com o uso da máscara obtida inicialmente são criados os componentes a partir dos quais serão calculados a magnitude do gradiente e a direção do mesmo.

Com o uso das informações de magnitude e direção do gradiente do fundo e do quadro atual do vídeo, [Sanin, Sanderson e Lovell \(2010\)](#), verificam a correlação entre os dois. Uma vez a correlação sendo alta, a região é considerada sombra e é removida da máscara de objetos de *foreground*.

[Sanin, Sanderson e Lovell \(2010\)](#), definem seu método como uma fusão dos métodos baseados em cromaticidade, juntamente com uma versão do método de correlação de direção de gradiente.

Da mesma forma que os autores citados anteriormente [Xuehua et al. \(2008\)](#), utiliza vídeos coloridos de monitoramento de tráfego, e trabalha a utilização do NCC para detectar e remover as sombras.

Partindo do pressuposto que regiões sombras são apenas um decréscimo no valor da luminosidade em cada uma das bandas de cor, é proposto por [Xuehua et al. \(2008\)](#) uma fórmula para o cálculo da porcentagem da variação da tonalidade da cor no espaço de cores RGB.

$$C_R = \frac{R}{R + G + B} \quad C_G = \frac{G}{R + G + B} \quad C_B = \frac{B}{R + G + B}$$

Equação 2

A fórmula descrita pela Equação 2, chamada por [Xuehua et al. \(2008\)](#) de  $C_R, C_G, C_B$ , obtém como resultado a porcentagem das cores vermelho, verde e azul, respectivamente, contidas na imagem de entrada.

Uma vez comparada a cor do quadro definido como *background* através do modelo  $C_R, C_G, C_B$  com algum outro quadro do vídeo, percebe-se que regiões de sombra possuem apenas uma diminuição no valor de RGB, entretanto a porcentagem de RGB é praticamente a mesma. Uma vez havendo objetos de *foreground* na imagem, ao realizar a comparação dos quadros sob esse modelo, observa-se que há uma variação substancial no valor de RGB dessas regiões.

Xuehua et al. (2008) realiza uma subtração simples de *background* para encontrar os objetos de *foreground*. Subtraindo cada imagem de *foreground* da imagem de *background*, para os *pixels* que possuam um valor maior que o valor de luminosidade obtido através do modelo  $C_R, C_G, C_B$  é atribuído a condição de objeto de *foreground* cuja a sombra é removida pela fórmula:

$$\mathbf{Sombra}(i, j) = \begin{cases} 1 & \text{if } \Delta = C(i, j) - C'(i, j) > T_2 \\ 0 & \text{Outros} \end{cases}$$

$$\begin{array}{ll} \text{if } C_R > C_G & C = C_R \\ \text{else} & C = C_G \\ \text{if } C_B > C & C = C_B \end{array}$$

Equação 3

Na Equação 3,  $C(i, j)$  é o valor do pixel no quadro de *foreground*, já o valor  $C'(i, j)$  é o valor do *pixel* obtido no *background*, e  $T$  é um limiar para evitar a perda de informações importantes juntamente com as informações de sombra.

Os parâmetros utilizados descritos previamente na fórmula de Xuehua et al. (2008), visam decidir qual é a melhor banda da imagem  $C$  utilizar para determinado pixel.

Stander, Mech e Ostermann (1999), propõem um método para detecção de sombras em objetos em movimento aplicado à vídeos que utilizam o padrão ISO/MPG4.

Para o trabalho de Stander, Mech e Ostermann (1999) de detecção de sombras, quatro aspectos foram levados em conta para alcançar seu objetivo. O primeiro ponto, diz que a intensidade da luminosidade necessita, obrigatoriamente, ser alta, para gerar uma máscara de detecção de sombras. O segundo ponto, ressalta que, tanto câmera quanto a imagem de fundo utilizadas devem ser estáticos. O próximo item relata que o fundo deve ser plano e a fonte de luz deve estar longe do mesmo, e por último a quantidade de luz e a distância do objeto em relação ao fundo necessitam ser levados em consideração quando compara-se a distância entre a luz e o objeto.

Uma vez possuindo as características descritas anteriormente, Stander, Mech e Ostermann (1999), realizam a detecção de cantos (*edges*) das imagens, através da separação dos cantos entre estáticos (os que delimitam regiões sem movimentos), e cantos em movimento. Tal detecção é utilizada para uma futura integração de imagens com o intuito de tentar obter um melhor *tracking* (rastreamento) dos objetos.

Na Figura 10, Stander, Mech e Ostermann (1999) visam demonstrar de que maneira as sombras são geradas nas sequências de vídeo em que os testes foram realizados. A câmera é fixada em determinado ponto e na representação o círculo denota o objeto de

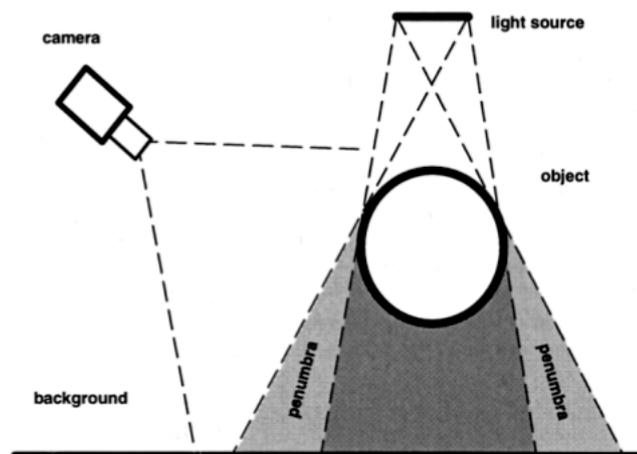


Figura 10 – Formação de sombras.

Fonte: [Stander, Mech e Ostermann \(1999\)](#)

interesse, logo acima do objeto, encontra-se a fonte de luz. A implicação dessa luminosidade sobre o objeto, gera a sombra e as regiões de penumbra (penumbra são regiões com menor luminosidade, entretanto, que não podem ser definidas como sombra).

A detecção das regiões que contém sombras é dada pela divisão entre a luminância do quadro atual pelo o quadro anterior gerando uma taxa de quadros. Então, os pixels que possuem uma variação pequena dentro dessa taxa de quadros medida em uma vizinhança espacial local são selecionados como candidatos à sombra.

Em [Yuan, Yang e Xu \(2010\)](#), nota-se a aplicação da remoção de sombras no mesmo ambiente que [Xuehua et al. \(2008\)](#), ou seja em sistemas inteligentes de transportes – ITS.

Primeiramente, [Yuan, Yang e Xu \(2010\)](#) propuseram a geração de um modelo de *background* em cores. Para tal, utilizou-se de um histograma da imagem para cada *pixel*, em cada um dos seus canais de cores. Os pixels que surgiram um número maior de vezes serviram como base para a geração do fundo, o qual foi finalmente determinado a partir do uso de uma função de avaliação aplicada sobre essa base de pixels.

Do mesmo modo que muitos outros métodos da literatura, [Yuan, Yang e Xu \(2010\)](#) realizam a subtração de *background*, entretanto, segundo os autores, o seu método com o uso de um fundo colorido obtêm resultados mais satisfatórios que os métodos tradicionais em escala de cinza.

Já no âmbito da remoção de sombras, é utilizado o mesmo método de [Jacques, Jung e Musse \(2005\)](#) de correlação cruzada normalizada *NCC*, porém, juntamente com o algoritmo de *NCC*, aplica-se um método sob a diferença de quadros, como forma de complemento, visando alcançar resultados superiores ao uso do *NCC* puro.

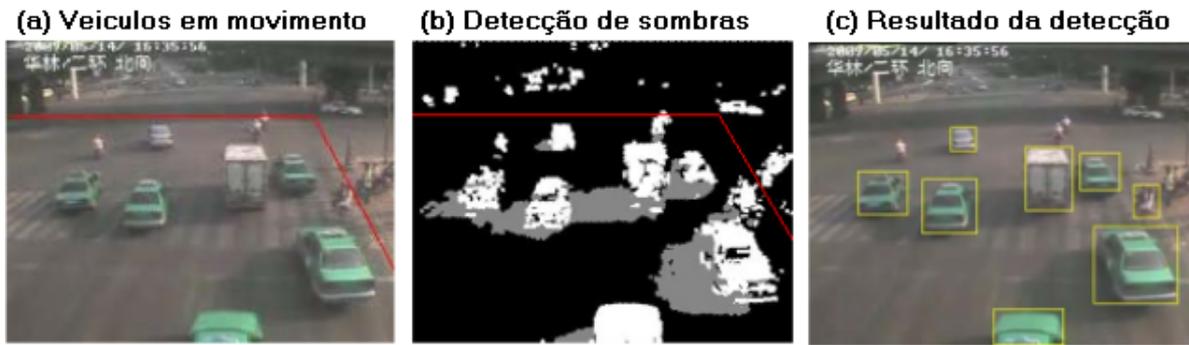


Figura 11 – Detecção de sombras.

Fonte: Yuan, Yang e Xu (2010)

Na Figura 11, é possível visualizar o resultado do método proposto por Yuan, Yang e Xu (2010). Em 11(a), denota-se um grupo de veículos em uma via, já em 11(b) tem-se a detecção de sombras, e por fim em 11(c) observa-se o resultado final da detecção. Por fim, em Yuan, Yang e Xu (2010), é dito que seu método possui bons resultados até mesmo onde há intersecção de via, o que é visto na Figura 11.

Outro método é proposto por Schreer et al. (2002), com o foco na detecção e remoção de sombras em tempo real voltada à sistemas de videoconferência. O método proposto por Schreer et al. (2002) trabalha em vídeos coloridos sob o sistema de cores YUV.

O algoritmo de Schreer et al. (2002) realiza a análise dos pixels de *foreground* em um primeiro momento para verificar quais possuem a mesma cor. Levando-se em consideração o fato, que já fora citado anteriormente, que regiões de sombra são apenas uma redução da intensidade de cor, realiza-se uma subtração entre cada quadro do vídeo com um fundo, sendo essa diferença limitada por um limiar, a fim de evitar a obtenção de valores negativos.

Finalmente, comparando-se o valor de  $U$  e  $V$  decide-se qual o arco-tangente entre eles que será utilizado como limiar para a detecção de sombra. Com os limiares definidos, os pixels que pertencerem à faixa definida pelo arco-tangente serão classificados como sombra e removidos do resultado final.

## 4 Apresentação da Pesquisa e Análise dos Resultados

Para fins de testes em métodos propostos por outros autores, utiliza-se vídeos de monitoramento, obtidos na base de vídeo [CAVIAR \(2013\)](#). Tal base é de acesso público, porém deve ser referênciada uma vez obtidos resultados com seu uso. Assim como realizados testes em vídeos de outras bases de dados disponíveis na Internet e amplamente utilizados na área de processamento de imagens e vídeo para teste de desempenho de novos algoritmos.

De forma mais específica, este trabalho inicialmente aborda a construção de um modelo de fundo do vídeo (*background*), através de algumas técnicas de modelagem. Entre essas técnicas de modelagem de fundo a serem testadas encontram-se algumas das mais conhecidas, como a utilização de modelo de fundo fixo [Yuan, Yang e Xu \(2010\)](#) e um modelo que usa a mediana de imagens [Jacques, Jung e Musse \(2005\)](#).

O primeiro modelo a ser testado utiliza apenas um *frame* (quadro) do vídeo onde sabe-se que não há objetos de *foreground* aplicando-o como fundo. No modelo de fundo fixo depara-se com o problema da variação de iluminação. Tal variação pode ocasionar uma detecção falha por parte do algoritmo uma vez que a própria variação pode ser detectada como um objeto de interesse.

Devido ao problema do método citado anteriormente, realizar-se-á também testes a partir de um método o qual utilize uma mediana de uma sequência de imagens pertencentes ao vídeo. Essa metodologia, explora o fato de que objetos que não pertencem ao cenário, apenas fazem a oclusão de determinada área por um curto período de tempo. Assim sendo, através do cálculo da mediana dos quadros durante um certo intervalo de tempo é viável a obtenção de uma detecção robusta do *background* mesmo que haja transeuntes em determinados momentos da sequência de vídeo.

Uma vez adquirido o modelo de fundo, busca-se a obtenção de uma segmentação robusta dos objetos de interesse contidos na sequência de vídeo a ser trabalhada. Tal segmentação será obtida a partir da subtração de *background* os objetos de interesse na sequência de vídeo.

Com a obtenção de objetos de *foreground* a partir da segmentação descrita previamente será necessário a realização do tratamento das regiões de sombra, pois uma vez segmentados os objetos, os mesmos muitas vezes trazem consigo a informação de sombra.

Para que o objeto de interesse mantenha seu formato com correteza, suas sombras necessitam ser detectadas e removidas para que não haja uma deformidade na detecção

do mesmo. Como já fora citado previamente há muitos problemas agregados à uma má detecção de objetos de interesse.

Tendo em vista os diversos métodos literários para a remoção de sombras descritos na revisão bibliográfica deste trabalho, o mesmo basear-se-á no método de [Jacques, Jung e Musse \(2005\)](#) que é um método que apresenta bons resultados e propõe-se melhorias com base em métodos probabilísticos.

Assim que alcançado uma detecção do formato de um objeto de interesse por meio do uso do algoritmo proposto [Jacques, Jung e Musse \(2005\)](#), aplicar-se-á uma razão de pixels que utilizam as informações da imagem original se a mesma for colorida com o intuito de obter um refinamento na detecção final da sombra pertencente ao objeto.

Por fim, utiliza-se um pós-processamento morfológico para a correção de possíveis falhas no objeto proveniente de alguma variação no ambiente que fora detectado incorretamente como objeto de interesse. A utilização da técnica de morfologia matemática é importante uma vez que é aplicada à fim de reconstruir informações que são relevantes para o resultado final, como remoção de ruído e fechamento de buracos na imagem.

Propõe-se um refinamento baseado na utilização de operadores morfológicos juntamente com elementos estruturantes para obter uma melhor detecção para vídeos que trabalhem apenas com o modelo de cores em escala de cinza.

Para a implementação de todos os métodos descritos nesse projeto, assim como os pós-processamentos morfológicos utilizar-se-á o software [MATLAB \(2010\)](#).

O *software* [MATLAB \(2010\)](#), utiliza-se da linguagem de programação homônima e permite-nos o desenvolvimento de códigos fontes, não obstante, traz agregado a si, uma *toolbox* (conjunto de ferramentas), dentre as quais contém um pacote de processamento de imagens que será utilizado para o tratamento final dos objetos segmentados durante o processo de detecção e remoção de sombras.

Optou-se inicialmente pela utilização de vídeos em escala de cinza para o desenvolvimento do algoritmo inicial. Esta decisão foi tomada baseando-se no fato que tais sequências de vídeo apresentam um tamanho inferior de armazenamento de disco, além de reduzir a carga necessária de processamento.

Um outro fator determinante para optar-se por esse tipo de vídeo é o fato de que grande parte dos cenários de vigilância (os quais são foco desse projeto), utilizam-se de câmeras que capturam vídeos em escala de cinza para realizar o monitoramento de regiões de interesse.

Dentre os métodos estudados na bibliografia, adotou-se como parâmetro o método proposto por [Jacques, Jung e Musse \(2005\)](#), primeiramente por utilizar-se de vídeos em escala de cinza, assim como fora proposto no projeto, e apresentar bons resultados

visualmente.

O método de [Jacques, Jung e Musse \(2005\)](#), como citado anteriormente, utiliza como fundo um quadro obtido com o uso da técnica de modelagem da mediana de *frames* para uma sequência de 100 quadros do vídeo (cerca de pouco mais de 3 segundos).

Com o objetivo de testar a eficiência do método mais rapidamente, no primeiro instante de desenvolvimento utilizou-se de um modelo de fundo fixo para a detecção e remoção das sombras contidas na sequência de vídeo, por tratar-se de uma técnica de modelagem de fundo mais simples.

O primeiro vídeo utilizado pertence ao cenário de testes da base [CAVIAR \(2013\)](#), chamado “*Walk1.mpg*”, que como conteúdo traz uma pessoa andando em linha reta e depois retornando, seguida de uma segunda pessoa que atravessa todo o cenário.

Tabela 1 – Informação técnica do vídeo *Walk1.mpg*

Descrição	Valor
Nome do vídeo:	Walk1
Formato:	.mpg
Tamanho:	6,9MB
Dimensão:	384x288
Duração:	24s
Quadros por segundo:	25fps
Taxa de bits:	200kbps
Formato de cores:	RGB

Fonte: Autor

A Tabela 1 descreve as informações técnicas do vídeo *Walk1.mpg*, o qual foi escolhido pois apresenta poucos objetos de *foreground* a serem detectados e tratados de forma que torna-se fácil obter uma rápida percepção visual dos resultados.



Figura 12 – Representação de alguns *frames* de *Walk1.mpg*.

Fonte: [CAVIAR \(2013\)](#)

O *software* [MATLAB \(2010\)](#), a partir do qual o projeto fora desenvolvido, trabalha apenas com a forma não comprimida dos arquivos no formato Moving Picture Group ([MPG](#)) e Audio Video Interleave ([AVI](#)). Os vídeos disponíveis em bases de dados na internet, são comprimidos para facilitar sua distribuição por meio de *downloads*. Assim sendo, o primeiro passo para trabalhar-se com tais vídeos, é a descompressão dos mesmos.

Para obter uma versão não comprimida do vídeo *Walk1.mpg* optou-se pela utilização do *software* [virtualdub \(2013\)](#), esse software pode ler arquivos no formato [MPG](#), e [AVI](#), e gerar uma versão em formato [AVI](#) não comprimido. Assim sendo, obtêm-se uma nova versão do vídeo agora com o formato [AVI](#) que pode ser trabalhado no *software* [MATLAB \(2010\)](#).

O formato de cores do vídeo selecionado é [RGB](#), e uma vez que fora proposto o desenvolvimento dos algoritmos em escala de cinza, e apenas uma posterior análise à métodos em vídeos em cores é necessário a conversão do mesmo para escala de cinza.

Para realizar a conversão do arquivo, lê-se o mesmo do disco rígido com a utilização do comando *aviread* (“*Walk1.avi*”), do *software* [MATLAB \(2010\)](#) e armazena-se o seu conteúdo em uma variável. Aplica-se então um segundo comando, *aviinfo()*, passando-se o conteúdo da variável obtido no comando anterior como parâmetro, assim é possível saber as informações do arquivo.

A partir das informações obtidas pelos comandos descritos anteriormente, é determinável quantos quadros compõe o arquivo, e então para cada um dos quadros contidos na sequência de vídeo aplica-se o comando *rgb2gray*.

O comando *rgb2gray* da linguagem [MATLAB \(2010\)](#), converte o formato de cores [RGB](#) que compunham o vídeo originalmente para o formato em escala de cinza.

Possui-se então um vídeo, no qual todos os seus quadros estão em tons de cinza. Tais quadros são salvos em formato de imagem *.png*, para serem trabalhados individualmente de forma mais simples.

## 4.1 Abordagem com *Background* Gerado Por Fundo Fixo

Baseando-se no modelo de fundo fixo extrai-se um dos quadros da sequência de vídeo, o qual não possui elementos que possam ser caracterizados como objetos de *foreground* e toma-se tal quadro como *background* do cenário.

A Figura 13 exemplifica o uso do fundo fixo, uma vez que é simplesmente a extração do primeiro quadro do vídeo, o qual não contém nenhum transeunte.

Uma vez que o foco é a detecção e segmentação de sombras em objetos de *foreground* e já possui-se um modelo de fundo o próximo passo é a detecção desses objetos. A Figura 14, é referente ao quadro de número 100 da sequência de vídeo “*Walk1.avi*” e



Figura 13 – *Background* com fundo fixo.

Fonte: [CAVIAR \(2013\)](#)



Figura 14 – Objeto de interesse no cenário.

Fonte: [CAVIAR \(2013\)](#)

representa um objeto de interesse (pessoa no centro do vídeo).

Para obter-se a segmentação dos objetos a técnica mais difundida é a subtração de *background*. Entretanto, a técnica de [Jacques, Jung e Musse \(2005\)](#), baseia-se no modelo de fundo gerado pelo cálculo da mediana para realizar tal subtração, como o método inicial testado é obtido através da utilização de um fundo fixo, é necessário a utilização de outro método para a obtenção dessa segmentação.

Para a obtenção de uma segmentação com o uso da técnica de fundo fixo, calcula-se uma subtração simples entre o modelo de fundo e o quadro atual.

A subtração leva em conta o uso de um limiar. Esse limiar define um limite de diferença tolerável para a variação da tonalidade de um pixel. Uma variação muito pequena pode ocorrer durante a captura do vídeo, imperceptível a olho nu, porém deve-se considerá-la.

Realizou-se testes com diferentes limiares para gerar a segmentação de objetos. Testes com um conjunto de valores, variando de 5 até 50 pixels (5,15,25,35,45,50), mostraram que um valor de diferença em torno 25 no valor do tom de cinza é o suficiente para contornar esse tipo de variação.



Figura 15 – Resultado da subtração de *background* com fundo fixo.

Fonte: Autor

O resultado da subtração entre o fundo fixo (Figura 12) e o *frame* aleatório (Figura 13), pode ser visto na Figura 15. Nessa imagem, observa-se a identificação dos objetos de interesse, que são as pessoas que transitam do lado de fora da janela, e a pessoa localizada no centro do quadro.

Na Figura 15 ainda pode-se notar uma quantidade mediana de ruído, ou seja, uma quantidade de pixels que não chega a alterar o objetos, porém, tais valores ruidosos são indesejáveis na identificação final do objeto. Esse ruído provém da variação de iluminação e ruídos da própria sequência de vídeo. A fim de obter um objeto mais preciso isso será utilizado morfologia matemática juntamente com elementos estruturantes visando removê-los.

A separação dos objetos só foi possível, pois durante a classificação dos pixels, se o *pixel* possuía característica de fundo, era atribuído a ele a cor branca (valor 255 dentro

da escala de cinza), os demais pixels foram mantidos.

Para posteriormente trabalhar com os filtros morfológicos, optou-se por tornar a imagem binária. Isso não interfere na identificação das áreas de sombra, portanto, para todos os pixels tidos como *foreground*, na etapa descrita anteriormente, é atribuído o valor a cor preta (valor 1 na escala de cinza).

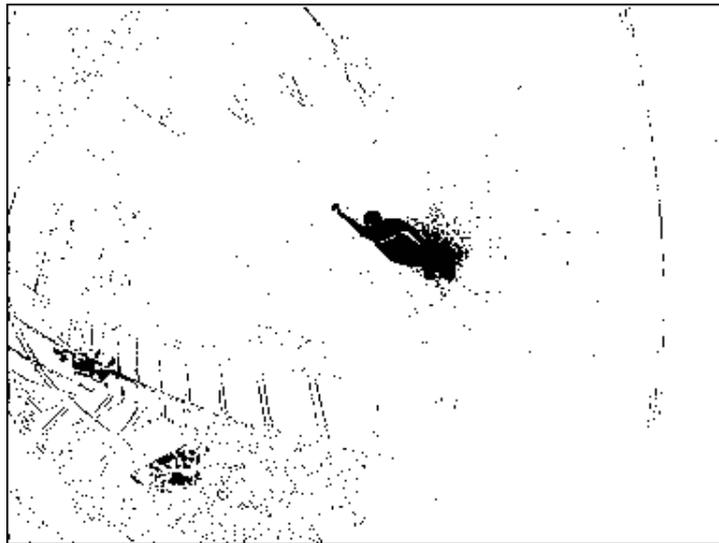


Figura 16 – Imagem de *foreground* binarizada.

Fonte: Autor

O passo seguinte é a identificação das áreas de sombra. Para isso aplica-se o algoritmo de correlação cruzada normalizada [NCC](#) descrito na seção 3.5.

Como parâmetro utiliza-se o valor de limiar 0.98, diferentemente de 0.95 proposto por [Jacques, Jung e Musse \(2005\)](#), ou seja, busca-se uma alta correlação entre as imagens e tal valor obteve melhores resultados visuais nos testes realizados.

O algoritmo de [NCC](#) pode ser aplicado para todos os *pixels* da imagem, porém utiliza-se um *template* (máscara de pixels) aplicando-se somente a pixels que pertencem ao objeto de *foreground*, analogamente ao algoritmo proposto por [Jacques, Jung e Musse \(2005\)](#), meramente por questões de desempenho, uma vez que não implicará em diferenças no processamento da imagem obtida no resultado final.

Gera-se o *template*, para cada *pixel* do objeto de *foreground* (*pixel* cujo valor seja diferente de 255), gerando assim uma vizinhança de  $(2M + 1) \times (2M + 1)$  *pixels* utilizando-se o valor de  $M = 1$ , exceto se tais valores tentem acessar posições inválidas (fora da borda da imagem).

Para maior acurácia na detecção de sombras, assim como a técnica de [NCC](#), aplica-se também um refinamento proposto por [Jacques, Jung e Musse \(2005\)](#). A fórmula a seguir



Figura 17 – Imagem após a remoção da sombra.

Fonte: Autor

demonstra tal refinamento.

$$\text{std}_R \left( \frac{I(i, j)}{B(i, j)} \right) < L_{\text{std}} \quad \text{e} \quad L_{\text{low}} \leq \left( \frac{I(i, j)}{B(i, j)} \right) < 1$$

Equação 4

Na Equação 4  $\text{std}_R$  é o desvio padrão de uma região  $R$  do *foreground* sobre a mesma região de pixels de *background*. Tal região foi definida anteriormente pela vizinhança de pixels de tamanho  $(2M + 1) \times (2M + 1)$ .

Além da região já descrita, são definidos dois limiares,  $L_{\text{low}}$  e  $L_{\text{std}}$ . O valor de  $L_{\text{std}}$  define o desvio padrão máximo da vizinhança de pixels e  $L_{\text{low}}$  é utilizado a fim de evitar que pixels que possuam um valor muito baixo de intensidade sejam erroneamente classificados como regiões de sombra.

Os limiares  $L_{\text{std}}$  e  $L_{\text{low}}$ , foram definidos em [Jacques, Jung e Musse \(2005\)](#), com os valores de 0.05 e 0.5, respectivamente. Segundo os autores, esses são apenas valores experimentais.

Testa-se os valores de pixels e os que se encontrarem dentro da faixa definida na Equação 4 entre os limiares e os valores já definidos na equação devem ser classificados como pixels de sombra.

Nota-se porém na Figura 17 que parte da informação do objeto de *foreground* foi classificado erroneamente como sombra. Para que haja uma minimização das perdas de informações assim como o preenchimento de falhas na imagem, e ainda remoção de ruídos, pode-se utilizar morfologia matemática como descrito na Seção 3.5.

A *priori*, faz-se o uso dos mesmos operadores morfológicos descritos no algoritmo de Jacques, Jung e Musse (2005), a saber: abertura seguido de fechamento. Entretanto, devido a resolução do vídeo ser menor, utiliza-se o elemento estruturante no mesmo formato (diamante), mas com o tamanho 1 ao invés de 5 como descrito no algoritmo desse autor.

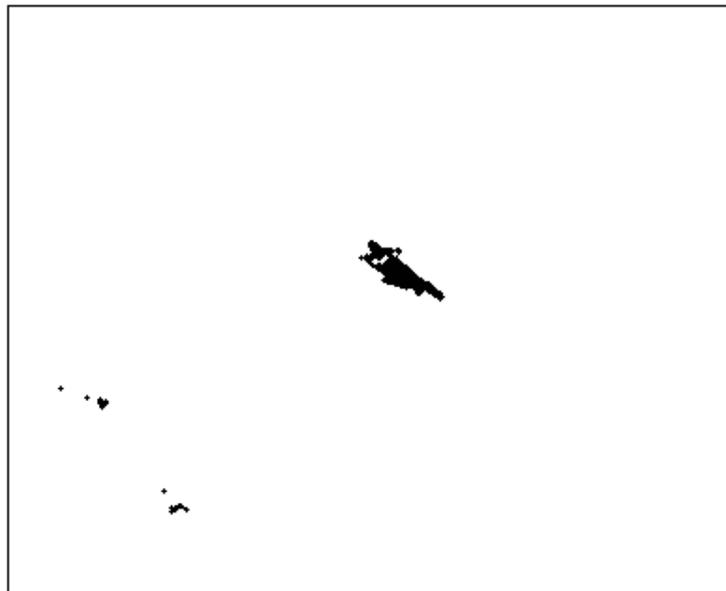


Figura 18 – Imagem após o processamento morfológico.

Fonte: Autor

Na Figura 18, obtém-se a imagem final após o processamento morfológico. Algumas partes do objeto foram perdidas, porém a identificação é mais precisa do que inicialmente com as regiões de sombra agregadas a ela.

No canto inferior esquerdo nota-se 2 pontos, esses são classificados como ruídos, são apenas variações de luminosidade que foram classificados como objetos de interesse por haver diferença com o fundo, porém não são objetos foco da detecção.

Apesar da perda de informações, conforme ilustrado na Figura 18, obtêve-se uma identificação razoável do objeto e este é um algoritmo aplicável.

Propõe-se para essa abordagem um refinamento baseado em morfologia matemática. Com a análise dos resultados dos *frames* obtidos pode-se verificar que ao invés de utilizar apenas o operador de abertura e fechamento, pode-se aplicar sob a imagem uma

sequência de operadores visando um preenchimento de falhas mais preciso do objeto com o uso em sequência de operadores de erosão, abertura, dilatação e por fim fechamento

Para tal, utilizou-se o mesmo elemento estruturante em forma de diamante de tamanho 1 o qual foi o elemento que melhor se adequou aos experimentos realizados. Consegue-se então obter uma detecção mais precisa do objeto de interesse.

A Figura 19, busca demonstrar o resultado da sequência de operadores descritos anteriormente. Comparando-se com a Figura 18, nota-se que houve uma reconstrução mais precisa da imagem.

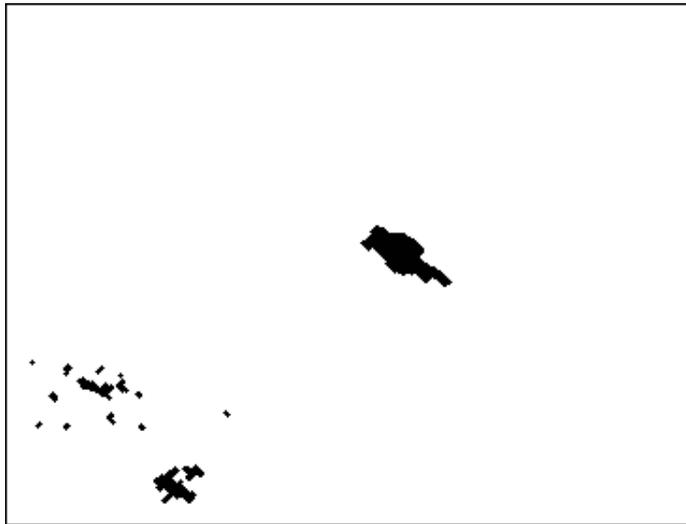


Figura 19 – Refinamento morfológico aprimorado.

Fonte: Autor

## 4.2 Abordagem com *Background* Gerado por Mediana

A abordagem por fundo fixo descrita na seção anterior é viável e apresenta resultados visualmente satisfatórios para algoritmos de remoção de sombra uma vez que nota-se que apenas o objeto foi detectado e as regiões de sombra descartadas. Entretanto é de interesse a verificação de outras técnicas de modelagem e sua aplicabilidade.

Seguindo a mesma proposta do algoritmo que utiliza fundo fixo, gera-se o *background*, entretanto, ao invés de utilizar apenas um quadro, o fundo é gerado a partir de uma sequência de quadros do vídeo. Assim como proposto por Jacques, Jung e Musse (2005), aplica-se um filtro de mediana sobre um intervalo de 100 quadros do vídeo à fim de gerar o *background*.

O intuito de utilizar-se um *background* baseado em uma mediana de quadros é obtenção de um algoritmo que possua menos ruído.



Figura 20 – Resultado da geração de *background* por mediana.

Fonte: [CAVIAR \(2013\)](#) e Autor

Com o fundo já modelado realiza-se novamente a subtração entre o fundo e o quadro atual. Para fins de comparação, utiliza-se o mesmo quadro a que aplicou-se a subtração de *background* no modelo de fundo fixo.



Figura 21 – Resultado da subtração de *background* binarizada obtida por mediana.

Fonte: Autor

Aplicando-se a modelagem de mediana para o fundo, no momento em que realiza-se a subtração de *background*, o problema de ruído é minimizado ao ponto de ser praticamente eliminado. Poucas regiões contendo ruídos são detectadas com o uso dessa técnica.

Essa modelagem de fundo leva à um maior tempo de processamento, entretanto, verifica-se visualmente na prática a obtenção de objetos menos ruidosos. Logo após a obtenção do *foreground*, realiza-se a binarização da imagem, como pode ser visto na Figura 21.

Tendo em vista o quadro obtido na Figura 21, novamente aplica-se o algoritmo de NCC, juntamente com o refinamento de regiões de sombra proposto por Jacques, Jung e Musse (2005). A Figura 22 apresenta o resultado do uso do NCC para remoção de sombras aplicado à imagem 21.

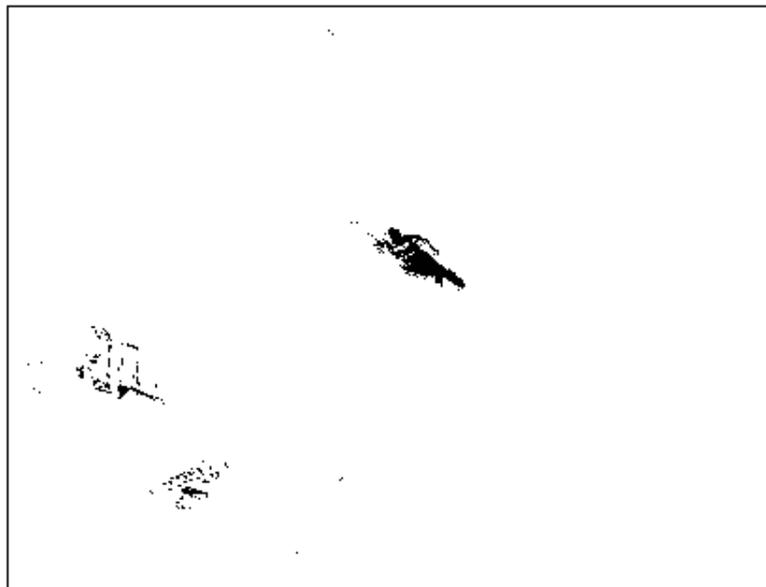


Figura 22 – Remoção de sombra obtida por mediana.

Fonte: Autor

Por fim, assim como descrito anteriormente aplica-se novamente morfologia matemática para a obtenção do objeto final.



Figura 23 – Objeto final obtido por mediana.

Fonte: Autor

Pode-se notar na Figura 23, que há uma pequena melhoria visualmente em relação a imagem obtida na Figura 18 com o uso do fundo fixo, entretanto, tem-se que ainda não é o resultado desejado.

### 4.3 Abordagem com uso de Informações de RGB

Uma vez proposto, realizou-se testes com a utilização de informações de vídeos com o formato de cores RGB. A partir dos mesmos quadros trabalhados anteriormente busca-se a obtenção da subtração de *background* e a identificação de objetos.

Para a modelagem de *background*, utiliza-se a técnica de modelagem de fundo fixo, pois para vídeos coloridos o cálculo da mediana tornar-se-ia mais complexo de implementar, assim como, teria um custo computacional bem maior uma vez que tem-se que trabalhar com os 3 canais de cores.

Utilizando as mesmas imagens descritas na Figura 13 e na Figura 14, porém, com cores, realiza-se a subtração de *background* da mesma forma que para quadros em escala de cinza.

Uma vez segmentado os objetos, como pode ser visualizado na Figura 24, cria-se cópias temporárias da imagem de fundo e da imagem de *foreground* em escala de cinza, a fim de utilizar o algoritmo de NCC para definir as regiões de sombra.

Após realizada a detecção das regiões de sombra aplicada à cópia temporária em escala de cinza da imagem, somente a fim de visualizar os resultados obtidos, utiliza-se



Figura 24 – Objetos detectados em subtração de imagem RGB.

Fonte: Autor

o canal azul do *RGB* para destacar na imagem as regiões de sombra identificadas pelo o algoritmos de *NCC* no objeto de *foreground* como pode visualizado na Figura 25. Tais pixels em destaque são regiões a serem excluídas da imagem, para tal, altera-se o valor desses pixels para branco.

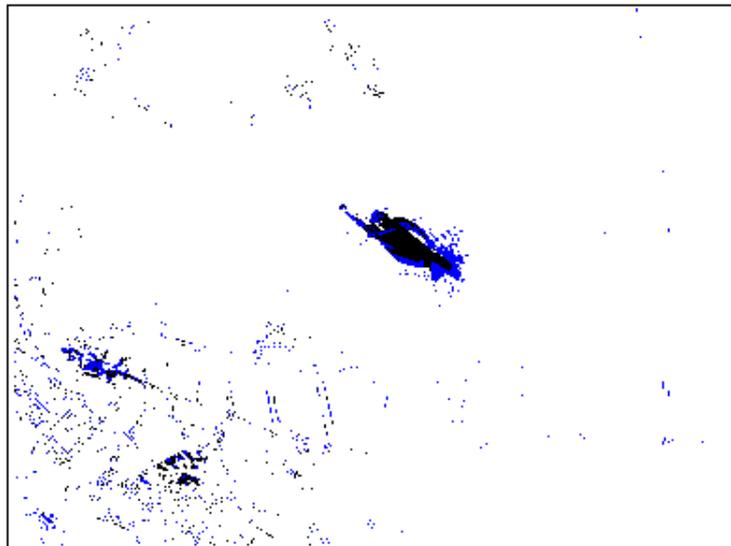


Figura 25 – Regiões de sombra identificadas por *NCC*.

Fonte: Autor

Analogamente ao processo realizado para os quadros das imagens de tom de cinza, aplica-se morfologia matemática para obter uma melhor identificação do objeto final.

Assim como nas figuras que descrevem os modelos finais em escala de cinza, a identificação não é perfeita, apesar de ambas abordagens possuírem bons resultados. A fim de melhorar a identificação dos objetos segmentados propõe-se um método de refinamento baseado em informações probabilísticas, o qual é descrito na Seção 5.4.

## 4.4 Refinamento por Meio de Informação de Probabilidade

Como proposto pelo trabalho, realiza-se testes com o uso de informações probabilísticas sob os vídeos trabalhados. Busca-se viabilizar com o uso de tais informações um refinamento do método descrito anteriormente.

Tendo em vista os métodos que utilizam-se de imagens de câmeras coloridas, sob o sistema de cores **RGB**, aplicou-se o método proposto. Esse método funciona apenas utilizando esse modelo de cores, pois para os cálculos a informação obtida em cada um de seus canais é necessária.

Como fora descrito anteriormente áreas de sombra são apenas variações na luminosidade de certa região, portanto, como descrito no trabalho de [Xuehua et al. \(2008\)](#), são apenas uma pequena variação na intensidade dos pixels **RGB**.

Diferentemente da fórmula  $C_R, C_G, C_B$  proposta por [Xuehua et al. \(2008\)](#), propõe-se um método que combina a aplicação de uma razão entre os tons de cores da imagem de *foreground* e a imagem de *background* e o uso de **NCC**, semelhantemente [Yuan, Yang e Xu \(2010\)](#).

O algoritmo de refinamento de sombras probabilístico proposto, difere do algoritmo de [Yuan, Yang e Xu \(2010\)](#), no âmbito que a razão de imagens é aplicada para apenas um grupo de pixels e é apenas um dos parâmetros do refinamento.

O primeiro passo para gerar o refinamento probabilístico, é criar uma razão entre o níveis de cores dos pixels de *foreground* que forem detectados durante a subtração de *background*. A fórmula pode ser definida como segue:

$$\gamma^R = \frac{R}{R_B}, \gamma^G = \frac{G}{G_B}, \gamma^B = \frac{B}{B_B}$$

Onde tem-se  $R, G, B$  como a intensidade dos pixels de cor vermelha, verde e azul, respectivamente, na imagem de *foreground*, e  $R_B, G_B, B_B$ , como essas mesmas posições de pixels nas cores na imagem de *background* e uma vez que tenha sido realizado o cálculo dessa razão para todos os pixels de *foreground* é realizado novamente o cálculo somente

para os pixels classificados como sombra pelo algoritmo de **NCC**, a razão então é dada pela fórmula:

$$\gamma R = \frac{R_S}{R_B}, \gamma G = \frac{G_S}{G_B}, \gamma B = \frac{B_S}{B_B}$$

Onde nesse quociente  $R_S, G_S, B_S$ , são os valores dos pixels os quais foram definido por meio do **NCC** como sombra, e  $R_B, G_B, B_B$ , são os mesmos pixels na imagem de *background* e  $\gamma R, \gamma G$  e  $\gamma B$  são as razões entre pixels na região de sombra.

A partir do cálculo de tais quocientes, para os pixels classificados na Figura 24, e para os pixels classificados como sombra pela Figura 25 (pixels em tom de azul), obtêm-se os seguintes resultados:

Tabela 2 – Valores de pixels

Banda de Cor	Média dos pixels	Desvio padrão da razão de sombra	Média da razão de pixels
R	0.6868	$\pm 0.2227$	0.6138
G	0.7032	$\pm 0.2280$	0.6436
B	0.7002	$\pm 0.2285$	0.6521

Fonte: Autor

Deve-se levar em conta, entretanto, que alguns valores podem tender a infinito, ou ainda, haver indeterminações do tipo zero sobre zero no momento da obtenção dos valores de média e também na obtenção dos valores do desvio padrão. Os valores desse tipo são desconsiderados no momento da realização do cálculo.

Para contornar o problema de valores inválidos, verifica-se então se o valor realmente existe e é finito, se não for, ele será tratado e desconsiderado do cálculo final.

É proposta então uma fórmula de reclassificação dos pixels de sombra detectados pelo **NCC** a partir do resultado do processo de média e desvio padrão descrito para os três canais de cores.

Na Equação 5, é descrito o cálculo que visa obter uma melhor classificação da região de sombra, levando em conta a relação entre os pixels nos três canais de **RGB**. Cada *pixel* inicialmente classificado como *foreground* é testado em cada um de seus canais de cores.

Pixels cujo o valor está entre a média dos pixels de sombra subtraído do seu desvio padrão e a média da sombra somada a seu desvio padrão possuem uma grande chance de

$$\begin{aligned}
& (mR_S - dR_S \leq R_F(i, j) \leq mR_S + dR_S) \\
& \quad \cap \\
& (mG_S - dG_S \leq G_F(i, j) \leq mG_S + dG_S) \\
& \quad \cap \\
& (mB_S - dB_S \leq B_F(i, j) \leq mB_S + dB_S)
\end{aligned}$$

Equação 5

serem pixels de sombra. Porém, para que a fórmula seja válida é necessário que tanto o canal R, G e B, daquele pixel se enquadrem nessa faixa de valores.

A Figura 26 demonstra o resultado da reclassificação por meio desse refinamento de relações de RGB para o *frame* da Figura 25.

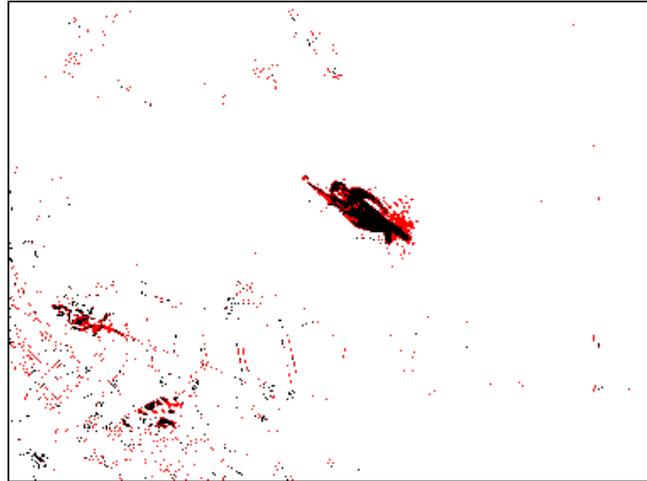


Figura 26 – Reclassificação de pixels.

Fonte: Autor

Os pixels em vermelho denotam os pixels classificados como sombra pelo refinamento proposto. Como pode ser visto numa comparação com a Figura 26, há uma melhora na classificação dos pixels, classificando um número menor de pixels erroneamente.

A execução do algoritmo proposto é demonstrado para outro quadro da sequência “Walk1.mpg”, na Figura 27, e a validação dos algoritmos descritos nesse trabalho podem ser visualizados nas sequências de quadros exibidas no apêndice.

O quadro 27(a), representa a imagem original do *frame* 51 da sequência de vídeo. Já em 27(b) tem-se a segmentação dos objetos de *foreground*, em 27(c) a identificação com o uso do NCC juntamente com o refinamento proposto por Jacques, Jung e Musse (2005), e finalmente em 27(d) é denotada a imagem obtida através do refinamento proposto.

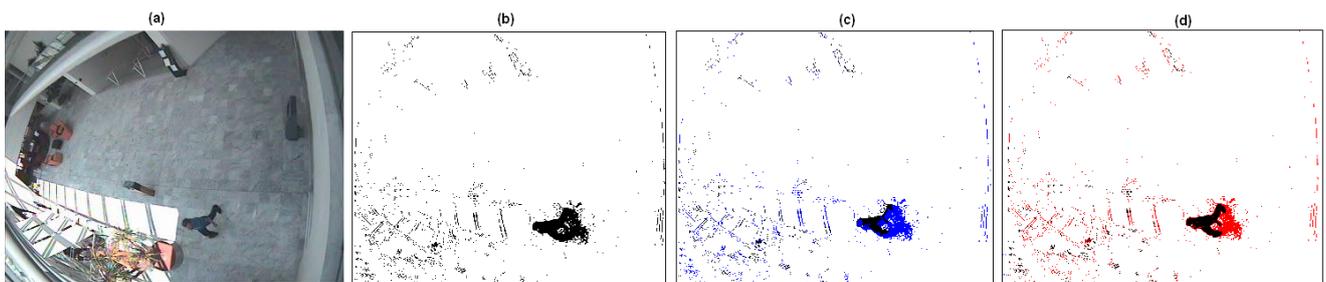


Figura 27 – Sequência de identificação de regiões de sombra

Fonte: [CAVIAR \(2013\)](#) e Autor

## 5 Considerações Finais

O presente estudo focou-se na detecção de objetos de interesse em sequências de vídeos obtidos a partir de câmeras de vigilância e a remoção de áreas contendo sombras. Para tal, inicialmente criou-se diferentes modelos de *background* baseando-se nas técnicas básicas descritas na literatura especializada.

Uma vez possuindo o modelo de *background*, realizou-se a segmentação de objetos de interesse e em tais objetos testou-se diferentes formas de detecção e remoção de sombras.

As técnicas desenvolvidas para remoção das regiões de sombra basearam-se nos métodos descritos na literatura e foram aplicados em vídeos em escala de cinza e vídeos coloridos.

A principal técnica utilizada nos algoritmos baseou-se na correlação cruzada normalizada, que visa detectar sinais de intensidade semelhante, uma vez que é sabido que a sombra é, geralmente, apenas uma variação de intensidade do sinal, e não na forma do mesmo.

A partir dos resultados obtidos, foi possível determinar técnicas objetivando alcançar melhores resultados na segmentação tanto de objetos quanto de regiões de sombra.

Para vídeos em escala de cinza os quais possuíam fundo fixo, aplicou-se testes morfológicos os quais demonstraram uma melhoria na identificação do objeto final reconstruído. Porém, pode-se dizer que a maior contribuição desse trabalho, foi a melhoria alcançada em testes de vídeo coloridos através do uso de cálculos probabilísticos.

O método probabilístico proposto por neste trabalho através do modelo matemático desenvolvido sobre os canais **RGB** para o cálculo de prováveis regiões de sombra, obteve melhores resultados do que o método baseado somente em correlação de regiões.

Como trabalho futuro, pretende-se realizar testes agregando o uso de informações de textura de regiões a fim de obter um resultado final de segmentação ainda mais robusto.



# Referências

- BOVIK, A. *The Essential Guide to Video Processing*. 1. ed. [S.l.]: Elsevier Science, 2009. ISBN 9780080922508. Citado na página 23.
- CARVALHO, F. J. S.; TAVARES, J. M. R. da S. Metodologias para identificação de faces em imagens: introdução e exemplos de resultados. *Congresso Internacional de Métodos Numéricos en Ingeniería*, 2005. Citado na página 35.
- CAVIAR. *Test Case Scenarios/CAVIAR*. 2013. Acesso em: 28 mar. 2012. Disponível em: <<http://homepages.inf.ed.ac.uk/rbf/CAVIARDATA1/>>. Citado 11 vezes nas páginas 45, 47, 49, 55, 62, 69, 70, 71, 72, 73 e 74.
- DAVIES, E. R. *Machine vision: theory, algorithms, practicalities*. 3. ed. [S.l.]: Elsevier, 2004. Citado 2 vezes nas páginas 30 e 34.
- ENVI. *Guia do ENVI em português*. Sulsoft, 2000. Acesso em 10 ago. 2013. Disponível em: <[www.sulsoft.com.br](http://www.sulsoft.com.br)>. Citado 2 vezes nas páginas 31 e 32.
- GONZALEZ, R.; WOODS, R. E. *Processamento de imagens digitais*. 1. ed. [S.l.]: Edgard Blucher, 2000. Citado na página 40.
- HARITAOGLU, I.; HARWOOD, D.; DAVIS, L. S. W4: real-time surveillance of people and their activities. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, IEEE, v. 22, n. 8, p. 809–830, 2000. Citado na página 36.
- JACQUES, J. C. S.; JUNG, C. R.; MUSSE, S. R. Background subtraction and shadow detection in grayscale video sequences. In: IEEE. *18th Brazilian Symposium on Computer Graphics and Image Processing, 2005*. [S.l.]: SIBGRAPI, 2005. p. 189–196. Citado 20 vezes nas páginas 11, 13, 24, 27, 30, 31, 35, 36, 37, 43, 45, 46, 47, 49, 51, 52, 53, 54, 56 e 61.
- MATLAB. *Mathworks/MATLAB*. 2010. Acesso em: 28 mar. 2012. Disponível em: <<http://www.mathworks.com/>>. Citado 2 vezes nas páginas 46 e 48.
- MATSUSHITA, Y. *Shadow Elimination and Interpolation for Computer Vision and Graphics*. Tese (Doutorado) — University of Tokyo, 2002. Acesso em: 22 mar. 2012. Citado na página 24.
- PONCE, J. et al. Computer vision: a modern approach. *Computer*, v. 16, p. 11, 2011. Citado na página 24.
- RUSS, J. C. *The image processing handbook*. [S.l.]: CRC press, 2006. Citado na página 31.
- SANIN, A.; SANDERSON, C.; LOVELL, B. C. Improved shadow removal for robust person tracking in surveillance scenarios. In: *20th International Conference on Pattern Recognition (ICPR), 2010*. [S.l.: s.n.], 2010. p. 141–144. Citado 4 vezes nas páginas 24, 30, 40 e 41.

- SCHREER, O. et al. Fast and robust shadow detection in videoconference applications. In: *International Symposium on Video/Image Processing and Multimedia Communications 4th EURASIP-IEEE Region 8 VIPromCom*. [S.l.: s.n.], 2002. p. 371–375. Citado na página 44.
- SHAPIRO, L.; STOCKMAN, G. *Computer vision*. 1. ed. [S.l.]: Prentice Hall, 2001. ISBN 9780130307965. Citado na página 23.
- SOUTO, R. P. *Segmentação de imagem multiespectral utilizando-se o atributo matiz*. Tese (Doutorado) — Tese. INPE, São José dos Campos, 2003. Acesso em 16 out. 2012. Disponível em: <[www.sulsoft.com.br](http://www.sulsoft.com.br)>. Citado 3 vezes nas páginas 31, 40 e 41.
- STANDER, J.; MECH, R.; OSTERMANN, J. Detection of moving cast shadows for object segmentation. *IEEE Transactions on Multimedia*, v. 1, n. 1, p. 65–76, 1999. Citado 2 vezes nas páginas 42 e 43.
- SUN, B.; LI, S. Moving cast shadow detection of vehicle using combined color models. In: IEEE. *Chinese Conference on Pattern Recognition (CCPR)*. [S.l.], 2010. p. 1–5. Citado 2 vezes nas páginas 24 e 39.
- VIRTUALDUB. *VIRTUALDUB/Virtualdub.org: Proof that I had too much free time in college*. 2013. Acesso em: 17 maio 2012. Disponível em: <<http://www.virtualdub.org/>>. Citado na página 48.
- WANG, S.-K. et al. Fast shadow detection according to the moving region. In: IEEE. *International Conference on Machine Learning and Cybernetics*. [S.l.], 2007. v. 3, p. 1590–1595. Citado 4 vezes nas páginas 24, 37, 38 e 39.
- WANGNHEIN, A. V.; IZIDORO, A.; ALMEIDA, C. Morfologia matemática. In: . [s.n.], 2013. Acesso em: 30 set. 2013. Disponível em: <<http://www.inf.ufsc.br/~visao/morfologia.pdf>>. Citado na página 34.
- WHRIGHT, C. *YUV Colorspace*. 2013. Acesso em 17 jun. 2013. Disponível em: <<http://softpixel.com/~cwright/programming/colorspace/yuv/>>. Citado na página 33.
- XUEHUA, S. et al. Shadow removal of vehicles in a video system based on rgb chroma model. In: IEEE. *International Conference on Computer Science and Software Engineering*. [S.l.], 2008. v. 1, p. 977–980. Citado 4 vezes nas páginas 41, 42, 43 e 59.
- YUAN, C.; YANG, C.; XU, Z. Simple vehicle detection with shadow removal at intersection. In: IEEE. *Second International Conference on Multimedia and Information Technology (MMIT)*. [S.l.], 2010. v. 2, p. 188–191. Citado 6 vezes nas páginas 24, 30, 43, 44, 45 e 59.

# Apêndices



# APÊNDICE A – RESULTADOS

Para elucidar os resultados obtidos pelos algoritmos desenvolvidos o presente capítulo exibe de forma visual os mesmos.

Baseando-se nas imagens de fundo fixo definidas a seguir, exibe-se os resultados obtidos por alguns quadros.



Figura 28 – Fundo fixo em escala de cinza

Fonte: [CAVIAR \(2013\)](#) e Autor



Figura 29 – Fundo fixo RGB

Fonte: [CAVIAR \(2013\)](#)

Nas sequências abaixo, em escala de cinza, os quadros representam, respectivamente: a imagem a ser tratada, o objeto segmentado com sua sombra, imagem com o

refinamento morfológico proposto na literatura e o refinamento para imagens em escala de cinza proposto no presente trabalho.

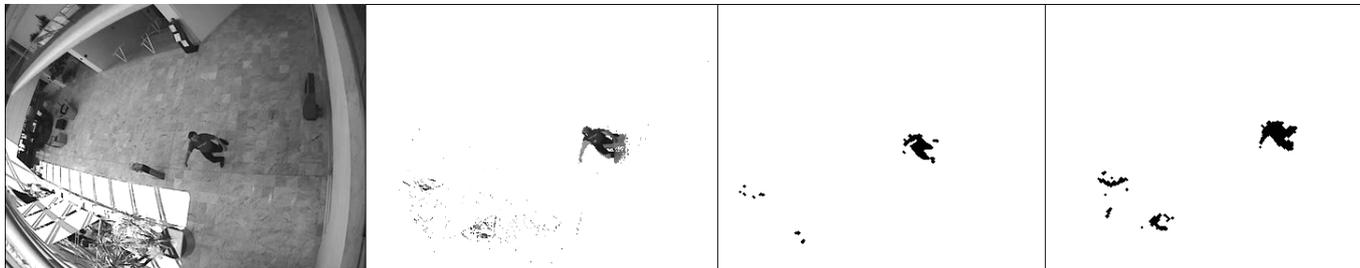


Figura 30 – Sequência em escala de cinza 1 - *Frame 76*

Fonte: [CAVIAR \(2013\)](#) e Autor

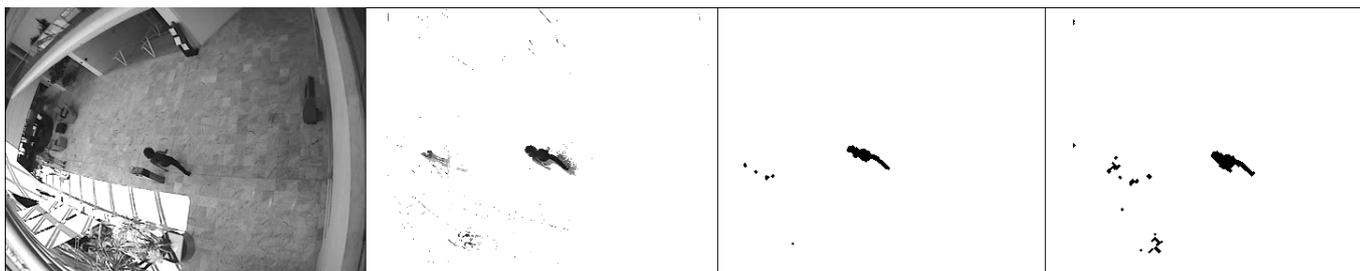


Figura 31 – Sequência em escala de cinza 2 - *Frame 316*

Fonte: [CAVIAR \(2013\)](#) e Autor

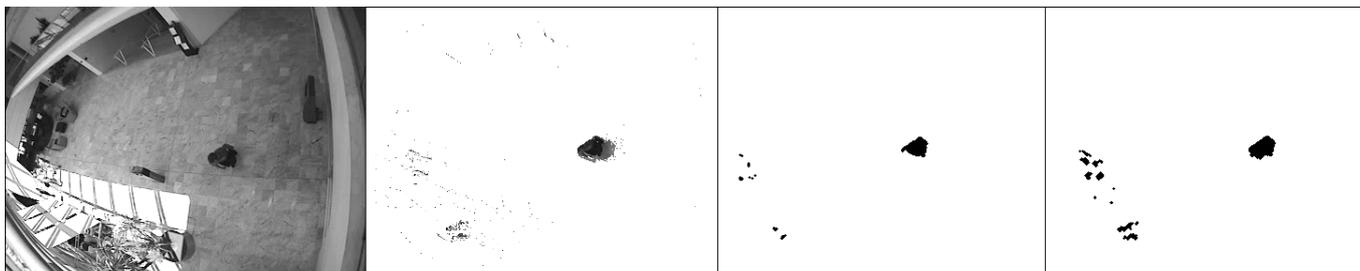


Figura 32 – Sequência em escala de cinza 3 - *Frame 137*

Fonte: [CAVIAR \(2013\)](#) e Autor

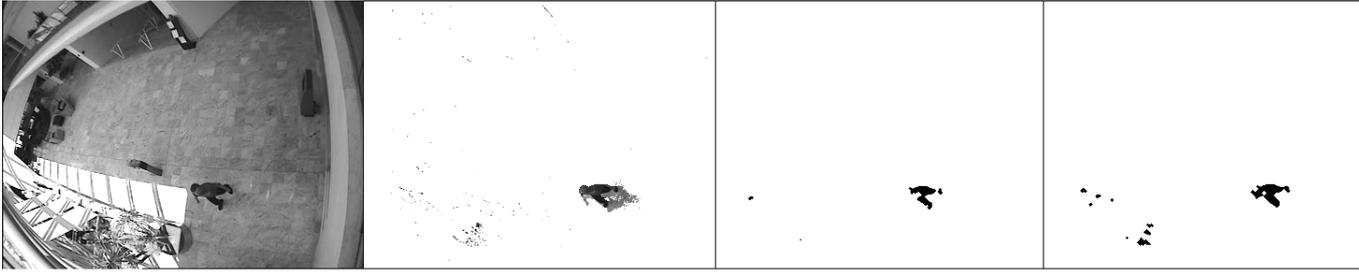


Figura 33 – Sequência escala de cinza 4 - *Frame* 46

Fonte: [CAVIAR \(2013\)](#) e Autor

Já para as demais sequências de imagens coloridas em vídeos [RGB](#), o primeiro conjunto de imagens representa o quadro do vídeo a ser tratado e o mesmo já segmentado, o segundo conjunto de imagens denota a detecção por meio de [NCC](#) (representado em azul) em comparação com a detecção probabilística proposta por esse trabalho (representado em vermelho) e por fim, o terceiro conjunto de imagens representa os operadores morfológicos de abertura e fechamento aplicados às imagens resultantes da remoção de sombra por meio de [NCC](#) e refinamento probabilístico, respectivamente.



Figura 34 – Quadro do vídeo e segmentação de objetos de *foreground*

Fonte: [CAVIAR \(2013\)](#) e Autor

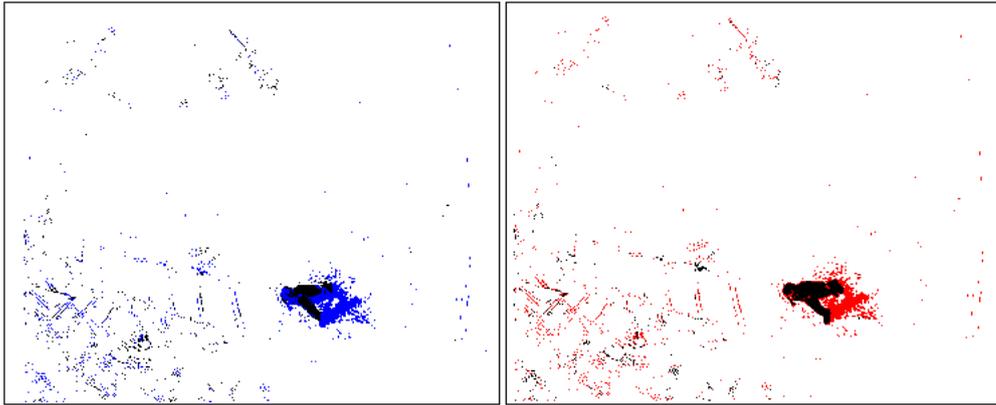


Figura 35 – NCC x Refinamento Probabilístico

Fonte: Autor

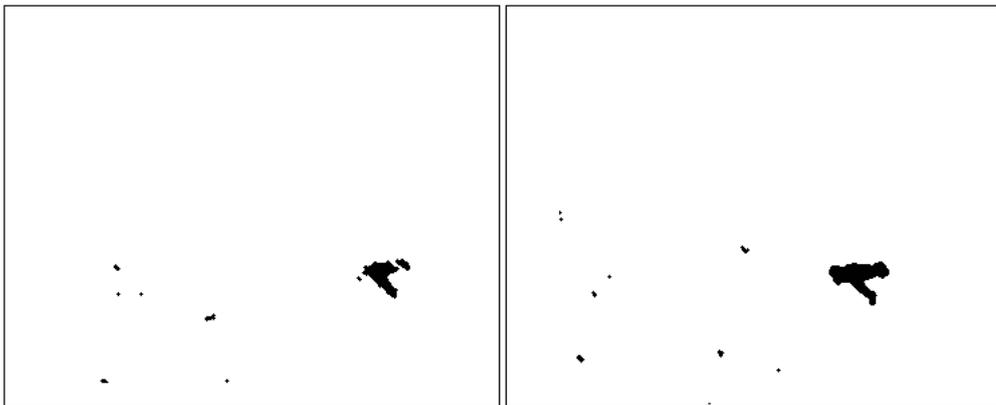


Figura 36 – Objeto final: Operadores morfológicos aplicados à NCC e Refinamento probabilístico

Fonte: Autor



Figura 37 – Quadro do vídeo e segmentação de objetos de *foreground*

Fonte: [CAVIAR \(2013\)](#) e Autor

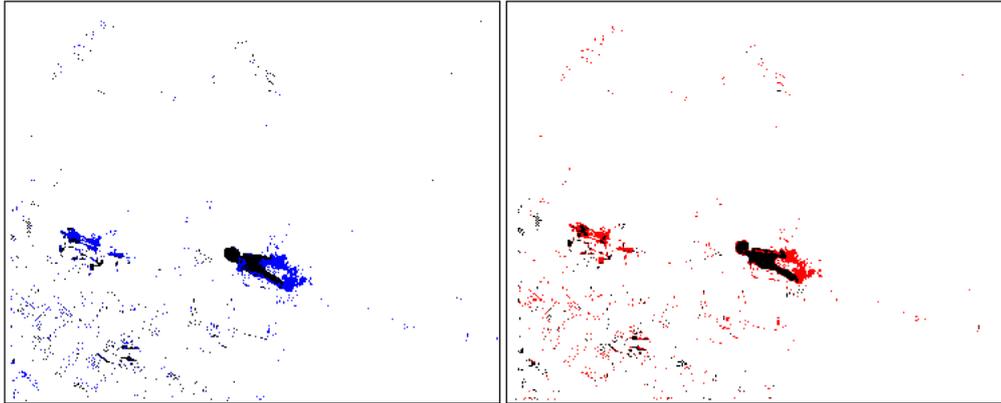


Figura 38 – NCC x Refinamento Probabilístico

Fonte: Autor

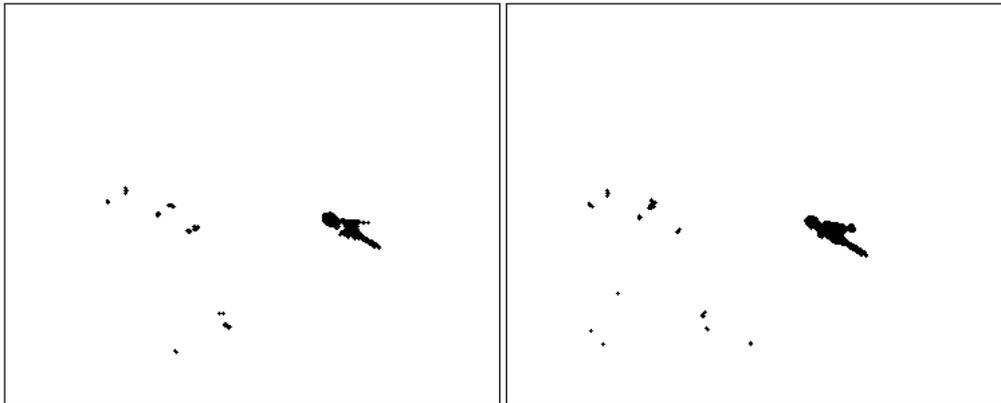


Figura 39 – Objeto final: Operadores morfológicos aplicados à NCC e Refinamento probabilístico

Fonte: Autor



Figura 40 – Quadro do vídeo e segmentação de objetos de *foreground*

Fonte: [CAVIAR \(2013\)](#) e Autor

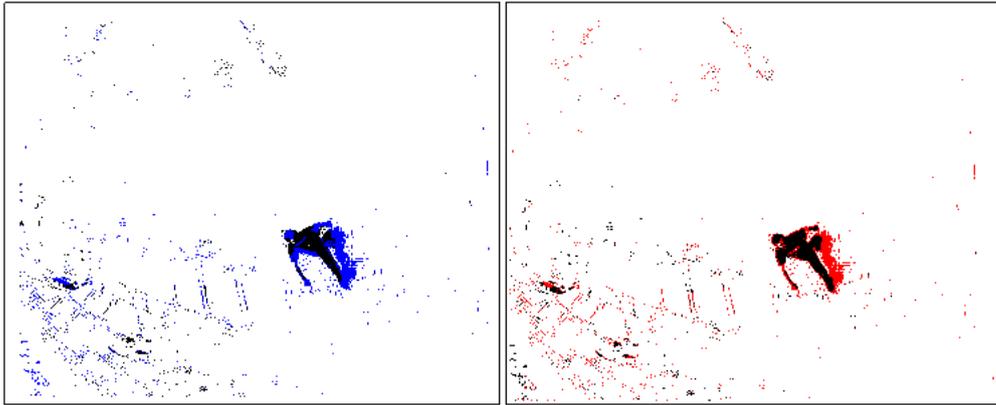


Figura 41 – NCC x Refinamento Probabilístico

Fonte: Autor



Figura 42 – Objeto final: Operadores morfológicos aplicados à NCC e Refinamento probabilístico

Fonte: Autor



Figura 43 – Quadro do vídeo e segmentação de objetos de *foreground*

Fonte: [CAVIAR \(2013\)](#) e Autor

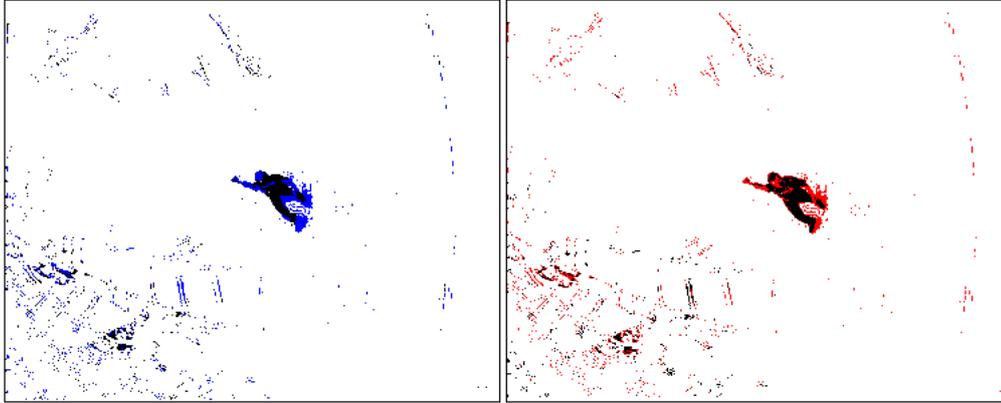


Figura 44 – NCC x Refinamento Probabilístico

Fonte: Autor

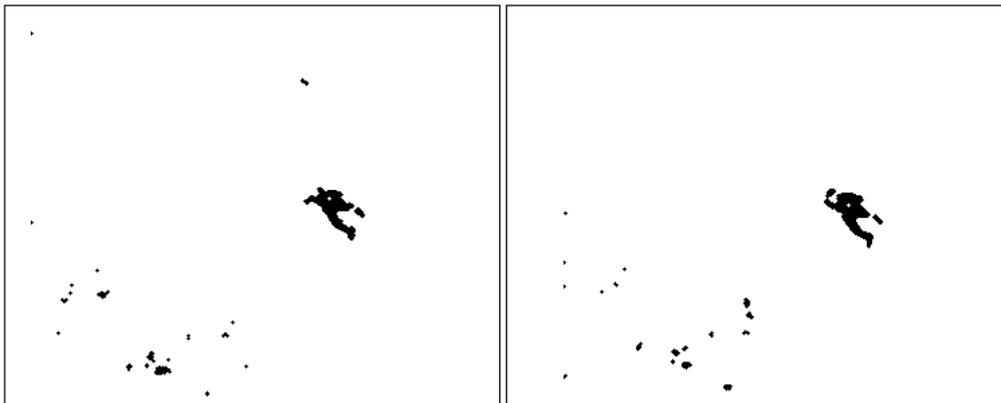


Figura 45 – Objeto final: Operadores morfológicos aplicados à NCC e Refinamento probabilístico

Fonte: Autor



# Índice

AVI, 46

HSI, 37

HSV, 13, 28, 30, 36

ITS, 35, 41

MPG, 46

NCC, 13, 33–35, 39, 41, 49, 54–59, 69

RGB, 13, 19, 28–30, 37–39, 46, 55–59, 61,  
67, 69

YUV, 13, 28, 30, 31, 42