

CENTRO UNIVERSITÁRIO FEI

VICTOR HENRIQUE CONFORTO

**SEGMENTAÇÃO DE IMAGENS COLORIDAS UTILIZANDO ALGORITMOS  
BIOINSPIRADOS**

São Bernardo do Campo

2017

VICTOR HENRIQUE CONFORTO

**SEGMENTAÇÃO DE IMAGENS COLORIDAS UTILIZANDO ALGORITMOS  
BIOINSPIRADOS**

Dissertação de Mestrado, apresentada ao Centro Universitário da FEI para obtenção do título de Mestre em Engenharia Elétrica. Orientado pelo Prof. Dr. Paulo Sergio Rodrigues.

São Bernardo do Campo

2017

Conforto, Victor Henrique.

Segmentação de imagens coloridas utilizando algoritmos bioinspirados  
/ Victor Henrique Conforto. São Bernardo do Campo, 2017.

72 f. : il.

Dissertação - Centro Universitário FEI.

Orientador: Prof. Dr. Paulo Sérgio Rodrigues.

1. Segmentação de imagens. 2. Algoritmos bioinspirados. 3. imagens coloridas. 4. Firefly. I. Rodrigues, Paulo Sérgio, orient. II. Título.

**Aluno:** Victor Henrique Conforto

**Matrícula:** 114125-8

**Título do Trabalho:** Segmentação de imagens coloridas utilizando algoritmos bioinspirados.

**Área de Concentração:** Processamento de Sinais

**Orientador:** Prof. Dr. Paulo Sérgio Silva Rodrigues

**Data da realização da defesa:** 10/02/2017

**ORIGINAL ASSINADA**

**Avaliação da Banca Examinadora:**

---

---

---

---

---

São Bernardo do Campo, 10 / 02 / 2017.

**MEMBROS DA BANCA EXAMINADORA**

Prof. Dr. Paulo Sérgio Silva Rodrigues

Ass.: \_\_\_\_\_

Prof. Dr. Flavio Tonidandel

Ass.: \_\_\_\_\_

Prof. Dr. Francisco de Assis Zampirolli

Ass.: \_\_\_\_\_

A Banca Julgadora acima-assinada atribuiu ao aluno o seguinte resultado:

APROVADO

REPROVADO

**VERSÃO FINAL DA DISSERTAÇÃO**

**APROVO A VERSÃO FINAL DA DISSERTAÇÃO EM QUE  
FORAM INCLUÍDAS AS RECOMENDAÇÕES DA BANCA  
EXAMINADORA**

Aprovação do Coordenador do Programa de Pós-graduação

\_\_\_\_\_  
Prof. Dr. Carlos Eduardo Thomaz

## RESUMO

Segmentação de imagens é uma das áreas mais antigas de visão computacional, com muitos problemas bem definidos e várias soluções propostas bem aceitas. No entanto ainda há muito trabalho a ser feito, sobretudo em segmentação de imagens coloridas, devido à demanda por mais aplicações. Recentemente, duas novas tecnologias têm se destacado na área. O estudo de análise de imagens sob o ponto de vista da estatística não-extensiva e a utilização de algoritmos bio-inspirados para lidar com problemas que demandam multi-limiarização, geralmente computacionalmente inviáveis quando o espaço de busca é histogrâmico. Assim, a proposta desta dissertação é o estudo de um novo método baseado em enxame de partículas, recentemente proposto na literatura e chamado de *Firefly*, juntamente com Kernel entrópico não extensivo para a multilimiarização de imagens espectrais. Os resultados obtidos mostram que o método proposto utilizando o algoritmo *firefly* segmentando a imagem baseado apenas na dimensão H de HSV obteve o melhor resultado dentre os experimentos realizados. Este trabalho aborda ainda a comparação entre o uso de diferentes espaços de cores, parâmetros e filtros para a segmentação de imagens coloridas.

Palavras-chave: Segmentação de imagens. Imagens coloridas. Algoritmos bioinspirados.

## **ABSTRACT**

Image segmentation is one of the oldest areas of computer vision, with many well-defined problems and several well-accepted proposed solutions. However there is still a lot of work to be done, especially in color image segmentation, due to the demand for more applications. Recently, two new technologies have stood out in the area. The study of image analysis from the point of view of non-extensive statistics and the use of bio-inspired algorithms to deal with problems that require multi-thresholding, usually computationally unviable when the search space is histogrammic. Thus, the proposal of this dissertation is the study of a new method based on swarms of particles, recently proposed in the literature and called "Firefly", together with non-extensive entropic kernel for the multilimiarization of spectral images. The results show that the proposed method using the firefly algorithm segmenting the image based on the HSV H-dimension only obtained the best result among the experiments. This work also discusses the comparison between the use of different color spaces, parameters and filters for the segmentation of colored images.

Keywords: Image Segmentation. Colored image. Bio-inspired algorithms

## LISTA DE ILUSTRAÇÕES

Ilustração 1 – Representação do sistema de cores RGB . . . . .	21
Ilustração 2 – Representação do sistema de cores HSV . . . . .	22
Ilustração 3 – Representação do sistema de cores CIELAB . . . . .	24
Ilustração 4 – Exemplo de histograma . . . . .	25
Ilustração 5 – Fluxograma da metodologia proposta . . . . .	40
Ilustração 6 – Exemplo de tratamento da imagem . . . . .	43
Ilustração 7 – Exemplo de redução de espessura de bordas . . . . .	44
Ilustração 8 – Canais de imagem HSV . . . . .	45
Ilustração 9 – Imagem original e imagem segmentada colorida . . . . .	45
Ilustração 10 – Imagem original e imagem segmentada com 2 limiares . . . . .	46
Ilustração 11 – Histogramas e limiares definidos para 1 limiar (à esquerda) e 2 limiares (à direita) . . . . .	46
Ilustração 12 – Extração de bordas pelo método Canny . . . . .	47
Ilustração 13 – Fluxograma do procedimento de avaliação da segmentação de uma imagem . . . . .	47
Ilustração 14 – Exemplo do resultado das projeções em $x$ , $y$ e histograma . . . . .	48
Ilustração 15 – Imagens diferentes nas quais o método $hXhYhZ$ indica similaridade na coordenada X . . . . .	50
Ilustração 16 – Exemplos de gráficos representando resultados da segmentação . . . . .	51
Ilustração 17 – Teste de variância sobre os resultados envolvendo diversos valores de $q$ (eixo horizontal) . . . . .	52
Ilustração 18 – Histograma do resultado utilizando H,S e V e histograma do resultado utilizando apenas a componente H . . . . .	53
Ilustração 19 – Exemplo de segmentação na coordenada H . . . . .	54
Ilustração 20 – Análise de variância dos resultados da segmentação em diferentes espaços de cores . . . . .	55
Ilustração 21 – Resultados entre espaços de cores e coordenada H . . . . .	56
Ilustração 22 – Sobreposição dos histogramas e média dos valores de diferença . . . . .	56
Ilustração 23 – Análise de variância das segmentações em H, S e V . . . . .	57
Ilustração 24 – Análise de variância dos resultados da segmentação com diferentes números de limiares no espaço HSV . . . . .	58
Ilustração 25 – Exemplo de <i>oversegmentation</i> . . . . .	58
Ilustração 26 – Resultados de múltiplos limiares na dimensão H . . . . .	59
Ilustração 27 – Análise de variância dos resultados da segmentação com com K-means e Firefly . . . . .	60
Ilustração 28 – Resultados com filtro passa-baixa . . . . .	61
Ilustração 29 – Resultados com passa-baixa na coordenada H . . . . .	62

## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO</b>	8
1.1	OBJETIVO	9
<b>2</b>	<b>TRABALHOS RELACIONADOS</b>	10
2.1	TÉCNICAS BASEADAS EM ENTROPIA E ALGORITMOS BIOINSPIRADOS	10
2.2	TÉCNICAS BASEADAS EM CLUSTERIZADORES	12
<b>2.2.1</b>	<b>Segmentação com <i>K-means</i></b>	13
<b>2.2.2</b>	<b>Segmentação com Fuzzy c-means</b>	14
<b>2.2.3</b>	<b>Segmentação com Watershed</b>	15
2.3	TÉCNICAS BASEADAS EM ÁREA	16
2.4	TÉCNICAS BASEADAS EM BORDAS	17
2.5	TÉCNICAS BASEADAS EM CARACTERÍSTICAS FÍSICAS	18
2.6	TÉCNICAS BASEADAS EM GRAFOS	18
2.7	TÉCNICAS BASEADAS EM CLASSIFICADORES ESTATÍSTICOS	19
<b>3</b>	<b>CONCEITOS FUNDAMENTAIS</b>	20
3.1	ESPAÇOS DE CORES	20
<b>3.1.1</b>	<b>Espaço de cores RGB</b>	20
<b>3.1.2</b>	<b>Espaço de cores HSV</b>	21
<b>3.1.3</b>	<b>Espaços de cores CIE</b>	23
3.2	SEGMENTAÇÃO DE IMAGENS COLORIDAS	24
<b>3.2.1</b>	<b>Análise de histogramas aplicada à segmentação de imagens</b>	25
<b>3.2.2</b>	<b>Imagens coloridas e imagens espectrais</b>	26
<b>3.2.3</b>	<b>Segmentação no espaço de cores HSI/HSV</b>	27
<b>3.2.4</b>	<b>Segmentação no espaço de cores RGB</b>	28
3.3	ALGORITMOS BIOINSPIRADOS	29
<b>3.3.1</b>	<b>Particle swarm optimization (PSO)</b>	29
<b>3.3.2</b>	<b>Algoritmo Firefly</b>	31
<b>3.3.2.1</b>	<b><i>Firefly aplicado à segmentação</i></b>	33
<b>3.3.3</b>	<b>Colônia de formigas</b>	34
3.4	MÉTODOS ENTRÓPICOS	36
<b>3.4.1</b>	<b>Entropia de Shannon</b>	36
<b>3.4.2</b>	<b>Entropia de Tsallis</b>	37
<b>4</b>	<b>PROPOSTA</b>	39
4.1	DEFINIÇÃO DO PROBLEMA E OBJETIVO	39
4.2	Base de dados	39
4.3	METODOLOGIA	39

<b>5</b>	<b>EXPERIMENTOS E RESULTADOS</b>	42
5.1	MÉTODO DE AVALIAÇÃO	42
5.1.1	Parametrização Padrão	42
5.1.2	Preparação da base de dados	42
5.1.3	Procedimento de segmentação	44
5.1.4	Extração de bordas da imagem segmentada	46
5.1.5	Método de comparação	48
5.2	EXPERIMENTOS	51
5.2.1	Variação do parâmetro $q$ na entropia de Tsallis	51
5.2.2	Segmentação apenas na dimensão $H$	53
5.2.3	Segmentação em diferentes espaços de cores	54
5.2.4	Segmentação com múltiplos limiares	55
5.2.5	Comparação com k-means	59
5.2.6	Experimentos com filtros passa-baixa	60
<b>6</b>	<b>DISCUSSÃO E CONCLUSÃO</b>	63
	<b>REFERÊNCIAS</b>	64

## 1 INTRODUÇÃO

Segmentação (SEZGIN et al., 2004) de imagens tem sido um assunto amplamente abordado em visão computacional. Trata-se de uma tarefa que é aplicada por exemplo na análise de imagens médicas (PHAM; XU; PRINCE, 2000), segurança (FRIEDMAN; RUSSELL, 1997), reconhecimento facial (M.-H. YANG; AHUJA, 1998), análise de imagens de satélite (BHANDARI et al., 2014), entre outros. Este procedimento trata-se do particionamento de uma imagem em múltiplos segmentos, com o objetivo de simplificar ou alterar representação da imagem para uma forma mais facilmente analisável ou mais significativa em termos de análise computacional (L. SHAPIRO; STOCKMAN, 2001).

Existem diversos métodos de segmentação de imagem. Estes podem levar em consideração diferentes critérios, tais como; cor, forma, textura e distribuição espacial. Um dos métodos mais tradicionais de segmentação é a limiarização por histograma, que busca a melhor separação da distribuição de probabilidades das intensidades dos pixels presentes na cena. Esta separação define os diferentes grupos presentes na imagem de acordo com os valores de intensidade/cor dos pixels, permitindo a definição da quantidade de limiares a serem aplicados. Esta técnica não leva em consideração informações espaciais, sendo calculada apenas sobre o histograma da imagem.

Dentre os diferentes critérios que possam ser utilizados para a definição destes limiares, há o uso da entropia não-extensiva (Tsallis) (TSALLIS, 1999). A partir do princípio da pseudo-aditividade da entropia de Tsallis, é possível analisar um histograma avaliando as possíveis posições de limiares em busca da que melhor subdivide a distribuição, definindo assim classes da imagem. Este processo pode ser custoso, principalmente em casos onde devem ser definidos mais de um limiar (multilimiarização). Isto ocorre pois realizando esta análise de forma exaustiva, o tempo de processamento cresce exponencialmente de acordo com o número de limiares desejados, tendo como base constante todo o espectro de intensidades possíveis de uma imagem.

O trabalho de X. Yang (2009) apresenta uma forma de melhorar a performance deste tipo de segmentação, utilizando algoritmos bio-inspirados para a segmentação de imagens em tons de cinza (1 canal). Estes algoritmos são meta-heurísticas responsáveis por aproximar os valores ótimos de solução de um problema, a partir da avaliação e aproximação de resultados, obtendo geralmente melhor performance do que métodos exaustivos que necessitam verificar todas as possibilidades de resultados. Em P. S. Rodrigues et al. (2015) o autor utiliza o algoritmo *firefly* (também um algoritmo bio-inspirado) com a função de avaliação (Kernel) baseada na entropia de Tsallis para a segmentação de imagens em tons de cinza.

Levando em consideração os resultados obtidos em P. S. Rodrigues et al. (2015) e a importância da presença de cores em uma imagem de acordo com o propósito da segmentação, este trabalho aborda o uso deste método de segmentação aplicado a imagens coloridas. Neste caso, diferentemente de uma imagem em tons de cinza, as imagens possuem mais de uma dimensão

a ser analisada e as características destas dimensões podem não corresponder à linearidade de uma escala em tons de cinza.

Dentre diversos espaços de cores possíveis, este trabalho aborda o HSV, o RGB e o CIE(L\*a\*b\*). Porém o foco é a utilização do sistema HSV, devido as suas propriedades apresentadas na Seção 3.1.2.

Primeiramente, são apresentados trabalhos na área onde abordam diferentes técnicas de segmentação citando seus casos de aplicação, vantagens e desvantagens. Então, abordados os conceitos fundamentais de assuntos relacionados ao tema como espaços de cores e algoritmos bio-inspirados para ambientação da apresentação do método proposto. A metodologia define como aplicar os conceitos de segmentação de imagens a partir da limiarização do histograma utilizando algoritmos bio-inspirados e entropia. Finalmente, os resultados obtidos são analisados e avaliados.

## 1.1 OBJETIVO

Estudar e implementar um método de segmentação de imagens coloridas baseado no uso de algoritmos bio-inspirados.

## 2 TRABALHOS RELACIONADOS

A tarefa de segmentação é um dos desafios no processamento digital de imagens, combinando aspectos cognitivos e da área de ciências da computação. A segmentação é a divisão da imagem em diversas regiões (também chamadas áreas de interesse), de acordo com suas características, como cor, forma, textura e relação espacial.

Diversas técnicas de segmentação são abordadas em Sezgin et al. (2004), onde os métodos citados foram classificados entre 6 categorias principais, sendo elas: Análise de formas do histograma, baseados em clusterização, baseados em entropia, baseados em atributos de objetos, métodos espaciais e locais. Dentre estas categorias, este trabalho se enquadra na de técnicas baseadas em entropia para determinação de limiares de segmentação.

Outra vertente em segmentação de imagem é o uso de imagens coloridas que podem conter informações não aproveitadas nos métodos que segmentam baseados em um histograma de tons de cinza.

Skarbek et al. (1994) mostra diferentes técnicas de segmentação de imagens coloridas e as classifica entre as seguintes categorias: técnicas baseadas em pixels, técnicas baseadas em área, técnicas baseadas em bordas e técnicas baseadas em características físicas. A maioria dos métodos utilizados para segmentação de imagens coloridas derivam de métodos utilizados em imagens em tons de cinza, portanto também são válidos para tal aplicação, sendo realizados os ajustes necessários.

Devido à quantidade de informações presentes em uma imagem colorida, os métodos de segmentação são bastante variados visto que cada método pode utilizar como fator principal, diferentes características da imagem (como matiz, textura, bordas, saturação, dispersão de cores entre outros). isto faz da segmentação de imagens coloridas uma área com grande quantidade de abordagens possíveis, o intuito deste capítulo é apresentar algumas destas abordagens descritas em publicações da área.

### 2.1 TÉCNICAS BASEADAS EM ENTROPIA E ALGORITMOS BIOINSPIRADOS

Nesta categoria, o trabalho Chang et al. (2006) disserta sobre o uso da entropia de Shannon e entropia relativa (distância de Kullback-Leibler) para limiarização de dados. O uso de entropia para este fim foi primeiramente proposto por Pun (1981) e aprimorado por Kapur, P. K. Sahoo e Wong (1985). A limiarização baseada em entropia considera o histograma de uma imagem como uma distribuição de probabilidades, e então seleciona o valor ótimo de limiar que maximiza a entropia máxima. Esta análise não leva em consideração a correlação espacial entre os pixels da imagem, sendo assim, diferentes imagens com um mesmo histograma irão gerar o mesmo valor de limiar. O Autor descreve a evolução das pesquisas na área e realiza uma comparação entre diversas técnicas de limiarização com entropia e entropia relativa, apresentando para as imagens utilizadas, qual foi o melhor método para determinar seu limiar.

Em Albuquerque (2004) é explorado o uso da entropia não-extensiva de Tsallis (TSALLIS, 1999) para segmentação de imagens. O autor utiliza a entropia para identificar limiares dentro de um histograma de uma imagem em tons de cinza; tal técnica foi motivada pela presença de informações não-aditivas em algumas classes de imagens. Para analisar a técnica de segmentação proposta foram simulados diferentes histogramas onde o objeto e o plano de fundo eram definidos por dois picos gaussianos na distribuição. O método busca então um valor de intensidade que melhor subdivide o histograma com base na pseudo-aditividade. O uso destes histogramas simulados permite a avaliação dos resultados de acordo com as características dos picos do histograma. Também é abordada a influência que o parâmetro  $q$  tem na segmentação. Os resultados mostram que a definição dos limiares de segmentação pela entropia de Tsallis é eficaz, sendo melhor aplicada em imagens ruidosas onde a distribuição de tons de cinza é normalmente definida por um histograma unimodal. Quanto ao parâmetro  $q$ , foi identificado que possui influência no resultado final da segmentação, mas não faz parte do escopo deste trabalho a definição de um valor ótimo para  $q$  (assunto explorado em P. Rodrigues e Giraldo (2009), de onde foram obtidos os valores de  $q$  para este trabalho).

O trabalho de Yin (2007) apresenta uma forma alternativa de se calcular os limiares para segmentação de imagens em casos de multilimiarização. Utilizando técnicas de *particle swarm optimization* (PSO) combinados com *cross-entropy*, o autor busca encontrar os limiares que define em seu trabalho como *minimum cross entropy thresholding* (MCET) A segmentação com diversos limiares por MCET utilizando a busca exaustiva é computacionalmente pesada e demorada, sendo de complexidade  $O(c^n)$ , onde  $n$  é o número de limiares desejados. O autor propõe calcular o MCET com resultados semelhantes aos obtidos pela busca exaustiva utilizando um algoritmo bioinspirado baseado em PSO.

No algoritmo proposto, cada partícula contém  $N$  parâmetros, sendo  $N$  a quantidade de thresholds desejados para a segmentação da imagem. A função objetivo é minimizar o MCET, utilizada para o cálculo de aptidão de cada partícula. Quando o algoritmo finaliza, a melhor posição visitada por uma partícula qualquer durante o experimento, é considerada a melhor solução.

Os resultados apresentados mostram que o algoritmo proposto baseado em PSO obteve resultados semelhantes aos obtidos pela busca exaustiva em um tempo de processamento bem menor e menos suscetível à quantidade de limiares (não possui complexidade  $O(c^n)$ ).

O trabalho de X. Yang (2009) estuda o uso do firefly comparando-o a outros algoritmos bioinspirados e genéticos para otimização multimodal. O autor utiliza 9 diferentes funções para comparar os resultados entre os algoritmos baseados nas meta-heurísticas testadas. Os resultados da simulação apontam que o firefly é melhor do que o PSO e algoritmos genéticos, em termos de eficiência e taxa de sucesso. Sugere-se também que os resultados podem ser melhores se reduzir o fator de aleatoriedade gradualmente conforme a convergência dos dados através das gerações de *fireflies*.

Utilizando o mesmo conceito de algoritmos bioinspirados, em Horng e Liou (2011) o autor propõe utilizar o *minimum cross entropy thresholding* (MCET) utilizando como kernel o algoritmo firefly (X. YANG, 2009) a fim de calcular limiares para segmentação de imagens.

Assim como em Yin (2007), cada partícula (*FireFly*) representa uma possível solução contendo os possíveis thresholdings a serem utilizados. O uso do FireFly é comparado com outros algoritmos bio-inspirados, sendo eles, *Particle swarm optimization* (PSO), *Quantum particle swarm optimization* (QPSO) e o *Honey bee mating optimization* (HBMO), assim como os resultados também são comparados com a busca exaustiva.

Comparado com a busca exaustiva, para as imagens analisadas, o firefly obteve limiares muito próximos, com tempo de execução consideravelmente menor. Com relação aos algoritmos bioinspirados, utilizando o *peak signal to noise ratio* (PSNR) como avaliação, o firefly apresentou resultados superiores ao PSO e ao QPSO, tanto com relação à qualidade da segmentação como em tempo de processamento. A comparação com o HBMO mostra que a performance do firefly não é significativamente diferente.

Em P. Rodrigues e Giraldi (2009) o autor propõe uma técnica de calcular o valor de  $q$  da entropia de Tsallis aplicado à segmentação de imagens. Partindo da ideia de que o valor ótimo de  $q$  varia não só de imagem para imagem, mas também internamente em uma mesma imagem em diferentes regiões, é apresentado um algoritmo que obtém os valores dos limiares pela entropia de Tsallis de forma recursiva. A imagem é inicialmente segmentada em duas regiões (*background* e *foreground*) utilizando um valor  $q$  inicial (definição a seguir), e então o mesmo procedimento é aplicado a cada uma das regiões segmentadas, gerando diferentes valores para  $q$ . O melhor valor para  $q$  em cada iteração é definido pela minimização da função  $S_q/S_{max}$  onde  $S_q$  é a máxima entropia obtida utilizando  $q$  em  $[0.01, 0.02, \dots, 2.0]$  e  $S_{max}$  a máxima entropia teórica do sistema. Para análise dos resultados, foi utilizada a base de Berkley contendo as imagens e sua segmentação feita por humanos como padrão de referência. Os resultados mostram que a técnica proposta obteve melhores resultados que os apresentados em Albuquerque (2004) e que utilizando a entropia tradicional de *Boltzaman-Gibbs*.

O trabalho de Yu e Jin (2013) também aborda o uso do firefly para determinação de múltiplos limiares na segmentação de imagens utilizando a máxima entropia como função objetivo. Considerando cada firefly como um conjunto de possíveis soluções de limiares, o autor compara a técnica utilizada com a busca por algoritmo genético em imagens em tons de cinza.

Os resultados mostram que o firefly consegue um melhor resultado (cerca de 15% de diferença) de entropia máxima (função objetivo) além de levar metade do tempo de processamento.

## 2.2 TÉCNICAS BASEADAS EM CLUSTERIZADORES

De acordo com Lucchesez e Mitray (2001), assumindo a cor como uma propriedade constante na superfície de um objeto em uma imagem e classificando os pixels de acordo com

sua cor descrita em um espaço de cores, é provável que diferentes objetos em cena sejam representados por diferentes classes de pixels.

Segundo Theiler e Gisler (1997) uma "boa" clusterização é aquela na qual os objetos de uma mesma classe possuem grande semelhanças entre si e se diferem significativamente das demais classes.

As técnicas presentes nesta categoria buscam agrupar pixels de acordo com o nível de semelhança de suas características, não necessariamente levando em consideração sua posição espacial na imagem. Métodos como clusterização utilizando *k-means* (A. K. JAIN, 2010), aplicado à segmentação de imagens se enquadram nesta categoria. Outra técnica que é vastamente aplicada à segmentação de imagens em tons de cinza, mas que pode ser adaptada também para o uso em imagens coloridas é a limiarização de histograma, que consiste em encontrar um ou mais pontos de corte no histograma da imagem, estes que definem os limites entre as classes de pixels observadas.

No caso de limiarização de histograma em imagens coloridas, é necessário adaptar a técnica para combinar a análise de mais de um histograma (um para cada canal do espaço de cores).

### 2.2.1 Segmentação com *K-means*

A segmentação utilizando o *k-means* agrupa os dados calculando uma média para cada classe definida, e então determina a pertinência do pixel a cada classe de acordo com sua distância das medias dos clusteres.

Em Theiler e Gisler (1997) é apresentado um algoritmo baseado no *k-means* que utiliza propriedades espaciais e espectrais da imagem. Os pixels vizinhos têm uma maior tendência de pertencerem a uma mesma classe do que comparando um par aleatório de pixels. Desta forma o autor leva em consideração informações espaciais sobre os pixels, como presença de textura ou homogeneidade. Estas propriedades são importantes na segmentação, pois são informações que quando presentes no contexto da imagem, perceptualmente descrevem a separação de áreas de interesse.

Para levar em consideração estas informações espaciais utilizando um clusterizador como *k-means*, pode-se alterar a representação dos dados de forma que cada pixel também possua informações sobre seus vizinhos, ou então adaptar o próprio algoritmo de clusterização para ponderar a distância espacial entre os pixels como fator de classificação.

O trabalho de T.-W. Chen, Y.-L. Chen e Chien (2008) apresenta uma técnica de segmentação de imagens coloridas no espaço de cores HSV utilizando clusterização por *k-means*. Primeiramente o autor discretiza o espaço HSV em  $30 \times 8 \times 8$  posições e para melhor representação das cores separa os componentes que possuem baixa saturação pela sua falta de significância na determinação da cor em um histograma separado. Isto evita que cores iguais de saturação

O gerem diversas vezes o mesmo tom de cinza repetido no espaço de cores, reduzindo assim o número de componentes.

Após a quantização do espaço de cores, é identificada qual a cor mais presente no espaço para ser definida como o primeiro centroide de um cluster. As demais cores presentes na imagem são analisadas com relação a este centroide. A cor que estiver mais distante é selecionada como um novo centróide de uma nova classe. Este procedimento é repetido até que atinja o número máximo de clusters definido por parâmetro, ou a distância entre os centroides seja menor que um valor também pré estabelecido.

Este conjunto de centroides então é submetido ao *k-means* de forma a classificar os dados e a um pós-processamento que identifica segmentação de ruídos e bordas utilizando filtros de análise de vizinhança dos pixels (informação espacial). O tratamento destas informações espaciais separadamente após a clusterização pelo *k-means* (diferentemente do trabalho de Theiler e Gisler (1997) citado anteriormene) torna a segmentação mais eficiente, porém por outro lado caso a clusterização gere classes baseadas no ruído, este método pode não ser capaz de discernir os falsos clusters pós segmentação.

### 2.2.2 Segmentação com Fuzzy c-means

A clusterização por *Fuzzy c-means* (FCM) é uma técnica não supervisionada que classifica os dados observados de acordo com sua semelhança utilizando conceitos da logica fuzzy. A principal diferença entre o FCM e a clusterização por *k-means* é que em FCM os dados são classificados em clusters fuzzy. A definição de clusters fuzzy que envolve o fator fuzzificador  $m$  que determina o nível de fuzzificação.

O trabalho Liu e Zhou (2014) aborda o uso de mais de uma *feature* da imagem ao mesmo tempo a fim de determinar a segmentação de uma imagem colorida. É abordado o uso de características de cor e de textura, que são extraídos da imagem. Informações sobre textura são obtidas a partir da função de Gabor, que consiste na aplicação de um conjunto de filtros em orientações e escalas diferentes para classificação.

Partindo destas informações, é utilizado o *Fuzzy c-means* (FCM) (BEZDEK; EHR- LICH; FULL, 1984) para particionar a imagem, e em seguida o índice de Xie-Beni (clusterização baseada em simetria) para determinar o número adequado de partições. O resultado são dois conjuntos de clusters, um baseado em cor e outro em textura. Então, estes dois conjuntos de clusters são combinados em uma única imagem segmentada. Se um pixel é pertencente a um mesmo cluster extraído da cor, como da textura, então ele é considerado parte deste grupo, caso contrário é separado em um cluster diferente. O resultado é uma imagem *oversegmented* (muitas áreas, ainda não otimizada).

Esta imagem é tratada utilizando algoritmos evolucionários de multi-objetivos. A partir destes algoritmos obtém-se a aproximação do melhor conjunto de paretos dentre as informações, onde cada solução representa um resultado de segmentação.

O FCM por si só não considera informações espaciais dos pixels se tratando de segmentação de imagens. Chuang et al. (2006) apresenta um algoritmo baseado no FCM que incorpora informações espaciais da imagem para a clusterização dos dados. O uso destas informações faz com que o algoritmo encontre regiões de forma mais homogêneas e lide melhor com imagens ruidosas do que um método baseado apenas no clusterizador.

Uma importante característica de uma imagem é que pixels vizinhos são altamente relacionados, ou seja, tendem a possuir características semelhantes e a probabilidade de pertencerem a um mesmo cluster é alta. O autor utiliza uma janela para varrer a imagem e a componente espacial que influencia a determinação de um pixel pertencer a um cluster  $X$  é maior conforme a quantidade de pixels vizinhos já pertencentes à classe  $X$ . Os resultados demonstram maior tolerância a ruídos da imagem e menor susceptibilidade a criar falsos clusters.

### 2.2.3 Segmentação com Watershed

O método *watershed* é inspirado na forma como a água interage com diferentes níveis de elevação em um terreno. Em visão computacional, especificamente se tratando de segmentação de imagens, este método "inunda" a imagem de forma que as características dos pixels determinam seu relevo e como influem na distribuição do fluxo. Pontos iniciais de distribuição de classes (geralmente nos valores mínimos) são definidos na imagem e, a partir deles, os pixels vizinhos vão sendo classificados de acordo com sua semelhança. A técnica é comumente aplicada sobre a transformação gradiente da imagem, onde as áreas definidas pelos pixels de maior valor costumam representar as bordas entre os objetos, que são identificadas por não serem nomeadas em nenhuma classe pelo *watershed*.

O trabalho de Beucher et al. (1992) aborda o assunto e discute sobre sua aplicabilidade e limitações. Segundo o autor, o *watershed* é bastante sensível a ruídos, pois a análise do gradiente na maioria das vezes resulta em *oversegmentation*, e por este motivo, sua aplicação em segmentação de imagem normalmente ocorre combinada com outros métodos.

Um exemplo disto é o trabalho de Haris et al. (1998), primeiramente o autor utiliza uma técnica de redução de ruído preservando as bordas da imagem para obter seu gradiente. Então uma primeira segmentação da imagem é realizada utilizando o algoritmo *watershed* no gradiente da imagem. Esta segmentação inicial é submetida a um procedimento de *merging* hierárquico, produzindo a segmentação final.

Este último processo utiliza a representação da imagem em um grafo de regiões adjacentes em um procedimento iterativo. A cada passo, os pares de regiões mais semelhantes são determinados e mesclados (*merge*), e o grafo é atualizado para a próxima iteração. Este trabalho contorna as dificuldades de aplicação do *watershed* reduzindo o ruído e unindo regiões da imagem com *oversegmentation*.

Em Grau et al. (2004), o autor apresenta a aplicação da técnica em imagens do segmento médico. Para a segmentação são introduzidas informações previamente conhecidas sobre a

natureza das imagens para serem consideradas durante a segmentação pelo *watershed*. Estas informações são introduzidas no modelo pelo uso de cálculo de probabilidade, e combinam o uso do *watershed* com o registro do mapa de intensidade da imagem, utilizando marcadores posicionados na imagem nas áreas onde deve-se iniciar o inundamento.

### 2.3 TÉCNICAS BASEADAS EM ÁREA

São métodos que tratam a imagem dentro de seu domínio espacial a fim de identificar regiões de interesse (homogêneas). A classificação é fortemente relacionada com a disposição espacial dos pixels. Um destes métodos que é amplamente utilizado é técnica de crescimento de região (ADAMS; BISCHOF, 1994), que consiste em selecionar pontos da imagem como *seeds*. A partir deste ponto os pixels vizinhos são analisados de acordo com sua semelhança. Este procedimento ocorre recursivamente aos pixels vizinhos até que esta comparação não atinja limiar de similaridade, caracterizando o limite da área homogênea detectada. Tremeau e Borel (1997) utiliza esta técnica para segmentação em imagens coloridas.

Um dos trabalhos que aborda o assunto é Deng, Manjunath e Shin (1999), que apresenta um método de segmentação voltado para imagens naturais baseado em cor e textura. Este procedimento se inicia contabilizando as cores presentes na imagem (quantidade de pixels por cor) a fim de identificar quais são as cores mais significativas que serão usadas para a separação das áreas.

Após a contagem, as cores relevantes obtidas são classificadas e substituídas na imagem por rótulos. Esta imagem composta por rótulos é submetida a funções de análise de distância entre os seus componentes no plano cartesiano da imagem a fim de identificar áreas mais suscetíveis a serem regiões de interesse. A partir destas áreas, aplica-se o crescimento de região analisando os pixels em sua volta até a obtenção de um valor de parada (um vale na função que indica probabilidade de borda entre as classes).

O trabalho de Zhu e Yuille (1996) combina o método de crescimento de região com técnicas baseadas na detecção de bordas, criando um *framework* para segmentação de imagens que leva em consideração tanto a homogeneidade das superfícies, quanto os limites da imagem definidos utilizando a técnica de detecção de bordas ACM 2.4.

Outro método que se enquadra nesta categoria é o *split and merge* (HARALICK; L. G. SHAPIRO, 1985), onde o algoritmo inicia analisando uma região não-homogênea, e subdivide-se até encontrar pequenas áreas homogêneas. Então, é feito o processo inverso, são aplicadas técnicas de combinação (*merging*) para unir estas pequenas áreas homogêneas a fim de se obter as superfícies desejadas na segmentação da imagem.

## 2.4 TÉCNICAS BASEADAS EM BORDAS

Estes métodos baseiam-se na detecção de bordas da imagem para a segmentação da mesma. A análise pode ser feita localmente, de acordo com a vizinhança dos pixels, se expandindo recursivamente por toda a imagem, ou globalmente, que realizam otimizações e manipulam grandes áreas da imagem a fim de identificar as bordas em um contexto mais amplo que as técnicas locais. O trabalho de Tao e Huang (1997) clusteriza a imagem baseado em suas características de cor RGB, e então considera a transição desta clusterização como as bordas da imagem. As informações do gradiente de cada camada de cor são calculadas utilizando um filtro Sobel (R. JAIN; KASTURI; SCHUNCK, 1995). Em Perez e Koch (1994), o autor indica que para segmentação de imagens coloridas, a matiz (*hue*) é a característica mais importante a ser observada. Para isso demonstra que, para diversos tipos de sombras e luzes, a matiz continua inalterada, sendo um fator que determina a segmentação mais precisamente em cenas com variação de iluminação. Em seu trabalho, a detecção de bordas é realizada aplicando um filtro Laplaciano sobre o canal de matiz da imagem.

O artigo A. Sahoo e Chandra (2013) trata do uso do método de Active Contour Models (ACMs) juntamente com o *L'evy flight firefly* para segmentação de imagens. *Active Contour Model* é uma técnica de detecção de bordas que utiliza conjuntos de pontos relacionados, também chamados de *snakes*, que interagem com a imagem onde o possível contorno gera uma força de atração (energia), e a própria *snake* tem restrições internas que regram sua deformação para delimitação da borda. O *L'evy flight algorithm* (LFA) consiste na adição de um termo na função de movimento do firefly baseado na distribuição de L'evy. A inclusão deste termo implica em uma maior aleatoriedade na movimentação dos *fireflies*, com relação à direção e tamanho do passo.

De acordo com A. Sahoo e Chandra (2013), o ACM, quando utilizado em pequenas janelas de busca na imagem, resulta em pouca qualidade de convergência da *snake* (contorno ativo), atrelando-se a uma borda que possa não ter significado quando observada a imagem por inteiro. Por outro lado, a busca em áreas maiores da imagem é pouco eficiente e exige pesado processamento. O LFA é introduzido a fim de cobrir estas dificuldades, sendo aplicado em um grande espaço de busca. Cada Active Contour (conjunto de pontos que representam a *snake*) é representado por um firefly. O brilho é determinado pela função de energia da *snake*. O algoritmo proposto converge quando atinge o máximo de gerações ou encontra um contorno que atinge o limiar de energia mínima requerida (conceito do ACM).

Os resultados demonstram que a técnica utilizada é capaz de encontrar os contornos da imagem mais precisamente e com menor esforço computacional do que o método ACM tradicional. Os testes foram realizados com imagens médicas.

## 2.5 TÉCNICAS BASEADAS EM CARACTERÍSTICAS FÍSICAS

Grande parte dos métodos de segmentação de imagens coloridas está sujeita a falhas de segmentação caso os objetos na imagem estejam expostos alterações de luminosidade, que afetam a uniformidade de cores nas superfícies dos objetos apresentados na imagem. Uma forma de contornar este problema é analisar como a luz interage com os objetos e considerar tal análise na segmentação. As técnicas baseadas em características físicas buscam classificar os topos de objetos presentes na imagem antes de realizar a segmentação propriamente dita. De acordo com Lucchese e Mitray (2001) os objetos podem ser divididos entre 3 categorias principais: Dielétricos óticamente não-homogêneos, dielétricos óticamente homogêneos e metais. Esta classificação é conhecida como modelo de reflexão dicromática. Um trabalho que utiliza esta divisão é o de P. W. Tsang e W. Tsang (1996), que realiza detecção de bordas no espaço de cores HSV considerando a classificação de objetos segundo o modelo previamente citado.

Maxwell e Shafer (1997) apresenta um *framework* para segmentação de imagens que analisa hipóteses sobre a composição da imagem, levando em consideração a iluminação, forma e propriedades físicas dos objetos obtidas a partir de uma segmentação inicial. Tais hipóteses são analisadas, mescladas e filtradas até que o algoritmo encontre uma segmentação adequada à cena.

## 2.6 TÉCNICAS BASEADAS EM GRAFOS

A segmentação de imagens baseada em grafos, geralmente representa o problema em termos de um grafo  $G = (V, E)$  onde cada nó  $v_i \in V$  corresponde a um pixel na imagem, e uma aresta  $(v_i, v_j) \in E$  liga os vértices  $v_i$  e  $v_j$ . Um valor é associado a cada aresta baseado na propriedade dos pixels conectados por ela. Esta característica pode variar de acordo com a metodologia, mas de forma geral são características que podem ser usadas para analisar a similaridade dos pixels, como por exemplo a diferença entre a intensidade dos pixels. Também dependendo do método, não necessariamente haverá uma aresta conectando cada par de vértices (FELZENSZWALB; HUTTENLOCHER, 2004).

Um dos trabalhos pioneiros na área é o de Zahn (1971), onde apresenta uma técnica de segmentação baseada na árvore de extensão mínima do grafo observado. O critério utilizado para segmentar é analisar as arestas e eliminar as que possuem maior peso baseado em um limiar fixo definido. Desta forma, eliminando a relação no grafo entre os pixels menos semelhantes entre si. O resultado é um conjunto de subgrafos que representam os clústeres da imagem onde os pixels são mais semelhantes. Este método encontra dificuldades na segmentação em imagens onde há áreas com alta variabilidade de pixels, como texturas por exemplo, segmentando uma área que perceptivelmente seria única, em uma grande quantidade de pequenas áreas.

Felzenszwalb e Huttenlocher (2004) propõe um método de segmentação baseada em grafos onde os grafos são analisados para detectar bordas entre regiões da imagem, e a partir

destas informações, o algoritmo de segmentação faz a análise de características globais e locais a fim de obter a segmentação adequada. No trabalho é definido um predicado para avaliação da presença de características de borda entre duas regiões. Este predicado é baseado na medida de dissimilaridade dos elementos ao longo da borda, relativamente à medida de dissimilaridade entre os elementos vizinhos em cada região. Desta forma é levado em consideração não apenas a relação local pixel a pixel, mas também o contexto da área observada para corresponder à separação perceptual da imagem.

## 2.7 TÉCNICAS BASEADAS EM CLASSIFICADORES ESTATÍSTICOS

Haralick e L. G. Shapiro (1985) sugere que uma boa segmentação deve separar a imagem em regiões simples com características homogêneas. Diversos autores utilizam a técnica de Bayes para segmentar imagens segundo esta premissa. Este método modela a forma de regiões segmentadas considerando o comportamento dos pixels internamente a cada região homogênea, cada região é identificada por um *label* que será utilizada na classificação dos pixels. A segmentação então é realizada estimando o melhor *label* para cada pixel.

Em diversos trabalhos como o de Haralick e L. G. Shapiro (1985) para a definição dos labels, são utilizadas técnicas como *Markov random Fields*(MRF) ou distribuições de Gibbs (BESAG, 1974) pois requerem apenas a especificação de interações locais, de forma que as interações locais resultam em algoritmos de segmentação que também se baseiam apenas em características locais.

Classificadores estatísticos são métodos supervisionados pois requerem treinamento onde referências de segmentação são utilizadas como padrões para a segmentação de novas imagens. Nestes casos o número de classes de texturas da imagem é conhecido, porém nem sempre estão disponíveis informações sobre as propriedades destas classes. Abordando o problema com técnicas Bayesianas, o objetivo é de inferir a distribuição de possíveis segmentações. Devido às características mais relevantes do método serem as texturas observadas, técnicas baseadas em classificadores estatísticos são comumente aplicadas em imagens aéreas e de satélites.

O trabalho de Elia, Poggi e Scarpa (2003) apresenta um algoritmo de segmentação baseado em um modelo de MRF em estrutura de árvore. A imagem é segmentada recursivamente em regiões cada vez menores até que um critério de parada seja atingido. A segmentação é obtida como a solução de um MAP (*maximum a posteriori probability*) com as regiões previamente modeladas como MRF. O uso da estrutura em árvore faz com que o algoritmo proposto seja mais rápido para encontrar a segmentação do que outros que utilizam puramente os MRF.

### 3 CONCEITOS FUNDAMENTAIS

Este capítulo aborda os conceitos relevantes pertinentes ao tema do trabalho.

#### 3.1 ESPAÇOS DE CORES

Em visão computacional, o uso de imagens colorida oferece maior quantidade de informações do que imagens em tons de cinza, pois as cores podem auxiliar discernir contornos, contrastes e texturas que seriam menos perceptíveis se não coloridas. Oferecendo assim uma gama de possibilidades de aplicações. Algumas técnicas de processamento de imagens em tons de cinza podem ser aproveitadas para imagens coloridas, como K-means ou c-fuzzy, porém, necessitam de adaptação pois normalmente as informações de cores são representadas em mais de uma dimensão.

Há diversas formas de representar uma gama de cores, chamados de modelos de cores (ou espaços de cores). Segundo Gonzalez, Woods e Masters (2009), um modelo de cores é a especificação de um sistema de coordenadas e um sub-espaço deste sistema onde cada cor é representada por um único ponto.

Dentre diversos espaços de cores, os mais conhecidos são: RGB (*Red, Green, Blue*), CMYK (*Cyan, Magenta, Yellow, Black*), HSI (*Hue, Saturation, Intensity*) e HSV (*Hue, Saturation, Value*), que será mais abordado neste trabalho.

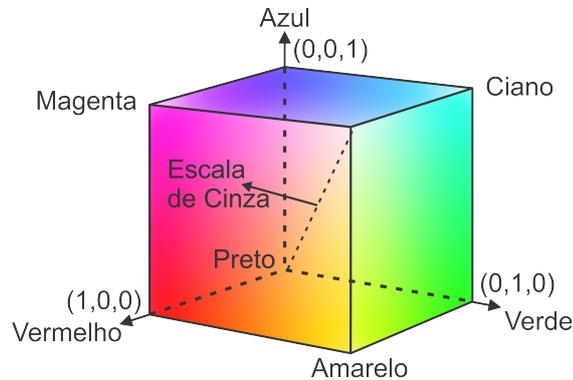
##### 3.1.1 Espaço de cores RGB

O espaço de cores RGB é um modelo que descreve as cores em termos de adição das cores primárias vermelho, verde e azul (*red, green e blue*, iniciais que definem o nome do modelo). O espaço de cores RGB é baseado na teoria de Young-Helmholtz (YOUNG, 1802) que disserta sobre como os fotorreceptores dos olhos humanos funcionam para visualizar cores. A teoria propõe que a retina humana possui três tipos de receptores que captam o comprimento de onda das luzes incidentes na retina: os receptores de curto comprimento (azul), médio (verde) e longo (vermelho). As forças dos sinais recebidos por estes três fotorreceptores são interpretados no cérebro como as cores visíveis. Esta teoria foi posteriormente confirmada pelos estudos Eysenck e Keane (2000) e Svaetichin (1956)

O propósito principal para a utilização do RGB é a representação de imagens em equipamentos eletrônicos, como monitores, televisores e impressoras. Comumente, dispositivos do gênero possuem em seu hardware fontes que geram as três cores primárias do modelo, e representam todo seu espectro de cores combinando os valores de vermelho, verde e azul.

A quantidade de cores possíveis de serem representadas, dependem da quantidade de bits utilizados na representação do pixel RGB, medida conhecida como profundidade do pixel. Em uma imagem onde seus correspondentes canais vermelho, verde e azul possuem dimensão

Figura 1 – Representação do sistema de cores RGB



Fonte: Autor.

de 8-bit, cada pixel RGB tem profundidade 24 (8 bits pelos 3 canais). O número total de cores que uma imagem RGB 24-bit é capaz de representar é de  $(2^8)^3 = 16,777,216$  cores.

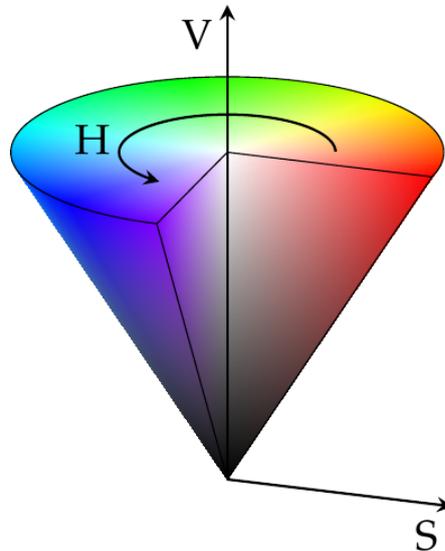
Uma das formas mais usuais de se representar o espaço de cores RGB é no formato de um cubo, onde os três eixos com origem em  $[0,0,0]$  representam as cores primárias do modelo, variando de sua intensidade mínima na origem, à sua intensidade máxima na extremidade oposta. A combinação das três cores em sua intensidade máxima resulta na cor branca (vértice  $[1,1,1]$ , considerando valores normalizados de 0 a 1). Um exemplo desta representação pode ser observada na Figura 1.

Nem todos sistemas possuem capacidade e/ou necessidade de reproduzir imagens com profundidade de 24-bit (imagens também conhecidas como *truecolor*). Por vezes, uma quantidade menor de cores possíveis é mais adequada a aplicações que necessitam mais de performance do que de fidelidade de exibição. Para estas situações é aplicado o conceito de *safe colors* (GONZALEZ; WOODS; MASTERS, 2009), que é uma discretização do espaço de cores RGB em que as três componentes principais podem assumir apenas 6 valores distintos, sendo eles 0, 51, 102, 153, 204 e 255. Desta forma, a combinação das três componentes gera  $(6)^3 = 216$  possíveis cores. Padrão este muito usado na web e comumente representado em 6 dígitos contendo valores hexadecimais, onde os dois primeiros representam o vermelho, os dois dígitos do meio representam o verde, e os dois últimos, o azul.

### 3.1.2 Espaço de cores HSV

Proposto por Smith (1978), o espaço de cores HSV classifica as cores de acordo com sua matiz, saturação e valor, esta forma de representação facilita o isolamento da matiz e o tratamento de iluminação se comparado a um sistema orientado à hardware (impressoras) como RGB. Este espaço é comumente representado na forma de um cone, onde sua dimensão angu-

Figura 2 – Representação do sistema de cores HSV



Fonte: Autor.

lar representa a matiz de cores (*hue*), seu raio determina a saturação (*saturation*) e sua altura corresponde ao valor (*value*), tal representação pode ser observada na Figura 2

A saturação é o grau de pureza da cor e é medida pela distância radial do eixo central do cone, seu valor varia entre 0 e 1. Para  $S = 0$ , a variação da componente V resulta na obtenção dos tons de cinza, desde o preto para  $V = 0$  até o branco onde  $V = 1$ . Assim, qualquer cor do sistema HSV pode ser transformada em um tom de cinza, onde o valor (componente V) determina qual o tom de cinza.

Visto que quando  $S=0$ , a cor resultante é independente de sua matiz, a significância da matiz para a identificação da cor cai conforme a saturação tende a 0, tornando a componente V a principal característica nestes casos.

De forma semelhante, quando a componente V possui valores muito baixos, as demais componentes tem sua significância reduzida, pois independentemente de saturação e matiz, a cor tende ao preto ( $V = 0$ ).

Uma das motivações para o uso do sistema de cores HSV é a intuitividade na obtenção de cores, o modelo é baseado em como as cores são conceituadas na visão humana, em termos de construção da cor. Por exemplo, no sistema de cores RGB, se desejarmos reduzir a saturação de uma cor, será necessário ajustar os valores dos 3 componentes do RGB até chegar no resultado desejado, sendo que as alterações não são evidentes, nem lineares. A mesma operação no HSV é feita de forma mais transparente, apenas alterando a componente S do sistema.

Assim como diversos outros espaços de cores, os componentes do HSV podem ser discretos em valores pré-estabelecidos, a fim de limitar a gama de cores trabalhadas. Geralmente

esta técnica é adotada para reduzir o custo computacional de seu processamento em problemas que exijam tal medida. Um exemplo disto é o HSV 162 (MATHIAS; CONCI, 1998), onde a matiz é discretizada em 18 unidades, o valor em 3 e a saturação também em 3, gerando um espaço de cores com apenas 162 cores ( $18 * 3 * 3$ ).

### 3.1.3 Espaços de cores CIE

Os sistemas de cores CIE (*Commission International de l'Eclairage*) foram desenvolvidos para representar a uniformidade perceptual e características psicofísicas presentes na visualização das cores por um observador humano (CHENG et al., 2001). As primárias de um sistema CIE são as componentes  $X$ ,  $Y$  e  $Z$ , qualquer cor pode ser representada por uma combinação destes componentes. Os valores de  $X$ ,  $Y$  e  $Z$  podem ser obtidos por uma transformação linear das componentes RGB. A matriz de transformação de acordo com o NTSC (*National Television System Commission, United states*) é aplicada segundo a Equação (1).

$$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \begin{pmatrix} 0.607 & 0.174 & 0.200 \\ 0.299 & 0.587 & 0.114 \\ 0.000 & 0.066 & 1.116 \end{pmatrix} \begin{pmatrix} R \\ G \\ B \end{pmatrix} \quad (1)$$

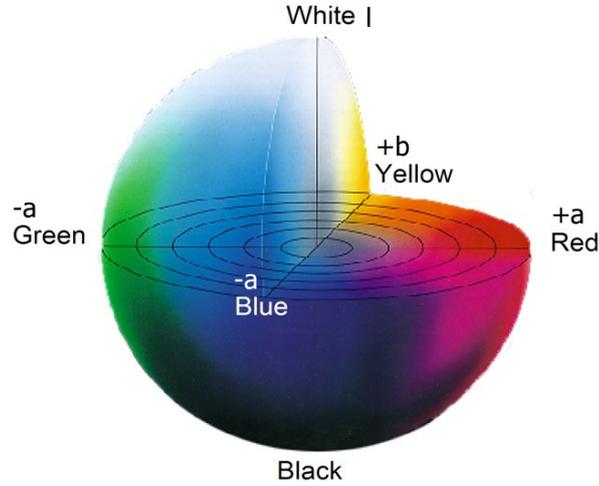
Existem diversos espaços de cores CIE criados a partir das coordenadas  $XYZ$ , dentre os quais, dois bastante conhecidos são os espaços  $CIE(L^*, a^*, b^*)$  e  $CIE(L^*, u^*, v^*)$  (X. ZHANG; WANDELL, 1997). Estes espaços são obtidos através de transformações não-lineares das coordenadas  $XYZ$ . No  $CIE(L^*, a^*, b^*)$  o valor  $L^*$  representa a luminância da cor, sendo  $L^* = 0$  para preto e  $L^* = 100$  para branco, o valor de  $a^*$  é a variação entre o vermelho e o verde (valores negativos tendem ao verde, enquanto valores positivos tendem ao vermelho), e a coordenada  $b^*$  que representa a variação da cor entre o azul e o amarelo (valores negativos tendem ao azul, e positivos ao amarelo). A representação do espaço de cores  $CIE(L^*, a^*, b^*)$  pode ser observada na Figura (3). O sistema de cores  $CIE(L^*, u^*, v^*)$  utiliza as coordenadas de forma semelhante em funcionamento, porém possui transformações cromáticas diferentes (o  $L$  é mantido).

O  $CIE(L^*, a^*, b^*)$  é definido nas Equações (2).

$$\begin{aligned} L^* &= 116 \left( \sqrt[3]{\frac{Y}{Y_0}} \right) - 16, \\ a^* &= 500 \left( \sqrt[3]{\frac{X}{X_0}} - \sqrt[3]{\frac{Y}{Y_0}} \right), \\ b^* &= 200 \left( \sqrt[3]{\frac{Y}{Y_0}} - \sqrt[3]{\frac{Z}{Z_0}} \right) \end{aligned} \quad (2)$$

Onde  $\frac{X}{X_0} > 0.01$ ,  $\frac{Y}{Y_0} > 0.01$  e  $\frac{Z}{Z_0} > 0.01$ , sendo  $X_0$ ,  $Y_0$  e  $Z_0$  as coordenadas do espaço que juntas representam a cor branca.

Figura 3 – Representação do sistema de cores CIELAB



Fonte: Autor.

O sistema  $CIE(L^*, u^*, v^*)$  é dado pelas Equações (3).

$$\begin{aligned} L^* &= 116\left(\sqrt[3]{\frac{Y}{Y_0}}\right) - 16, \\ u^* &= 13L^*\left(\frac{4X}{X + 15Y + 3Z} - u_0\right), \\ v^* &= 13L^*\left(\frac{6Y}{X + 15Y + 3Z} - v_0\right) \end{aligned} \quad (3)$$

Onde  $\frac{X}{X_0} > 0.01$  e  $u_0, v_0$  são os valores para a cor branca definida pelas coordenadas  $(X, Y, Z)$  por  $\frac{4X}{X+15Y+3Z}$  e  $\frac{6Y}{X+15Y+3Z}$  respectivamente.

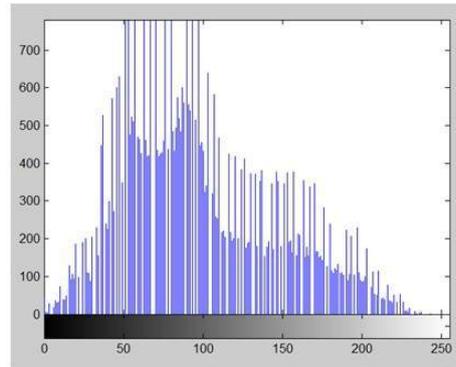
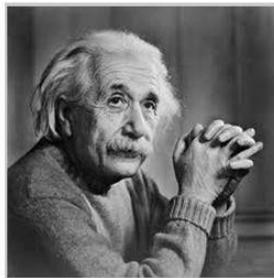
Os pontos do espaço de cores  $CIE(L^*, a^*, b^*)$  e  $CIE(L^*, u^*, v^*)$  são pontos em um espaço tridimensional e suas diferenças podem ser calculadas de acordo com sua distância euclidiana. Esta possibilidade é importante se tratando de segmentação de imagens, juntamente com a característica de representar a diferenciação de cores de acordo com a percepção humana. Característica esta que não está presente no sistema RGB. A distância entre duas cores distintas no espaço  $CIE(L^*, a^*, b^*)$  é dada por  $\Delta E_{ab} = \sqrt{(\Delta L^*)^2 + (\Delta a^*)^2 + (\Delta b^*)^2}$ . Analogamente, a distância no espaço  $CIE(L^*, u^*, v^*)$  é dada por  $\Delta E_{ab} = \sqrt{(\Delta L^*)^2 + (\Delta u^*)^2 + (\Delta v^*)^2}$ .

Esta distância é bastante útil quando segmentando imagens com o objetivo de separar às áreas de forma semelhante à que uma pessoa a faria à mão.

### 3.2 SEGMENTAÇÃO DE IMAGENS COLORIDAS

As cores são elementos fundamentais da visão humana para interpretação do ambiente. Cores são determinadas pelo comprimento de onda das luzes incidentes na retina que são por nós interpretadas. A segmentação de imagens coloridas envolve desafios com relação à forma

Figura 4 – Exemplo de histograma



Fonte:Autor.

imagem (à esquerda) e seu histograma (à direita)

de tratamento das informações, visto que existem diversas formas de representá-las (Seção 3.1), e dentre estas formas, vários métodos de interpretação e análise dos dados.

N. R. Pal e S. K. Pal (1993) apresenta a segmentação de imagens coloridas e considera que se assemelham às imagens multiespectrais, uma vez que uma imagem colorida pode ser descrita pela soma de diversos canais (dependentes do espaço de cores utilizados). Portanto, métodos utilizados para análise de imagens multiespectrais podem também ser aplicados às imagens coloridas. Um conceito fundamental para a segmentação de imagens (coloridas ou não) é a análise de histogramas.

### 3.2.1 Análise de histogramas aplicada à segmentação de imagens

De acordo com Roger (2010) os histogramas podem ser utilizados para distinguir na imagem objetos que se diferem nas características descritas no mesmo. Para facilitar a compreensão do conceito, o assunto será abordado tratando de imagens em tons de cinza, porém, o funcionamento também se aplica às imagens coloridas de forma semelhante.

O histograma, inicialmente proposto por Pearson (1895), é uma representação gráfica da distribuição de dados numéricos. Ele estima a probabilidade de distribuição de variáveis. No caso de segmentação de imagens em tons de cinza, estas variáveis representam a intensidade de cada pixel. Os valores presentes no histograma são a quantidade de pixels (eixo Y) que possuem X valor de intensidade (Valores discretos no eixo X). Desta forma, observando o histograma é possível calcular a probabilidade sobre o valor de intensidade um pixel aleatório selecionado na imagem. Um exemplo de histograma pode ser observado na Figura (4)

Em segmentação de imagens, um valor de intensidade  $f_T$  pode ser determinado a partir do histograma e utilizado como limiar de segmentação, separando possíveis áreas de interesse do restante da imagem, caso a diferença de intensidades entre estas áreas seja suficiente para tal.

A segmentação de uma imagem utilizando apenas um único valor  $f_T$  mapeia todos os pixels com intensidade superior a  $f_T$  como branco, e todos os demais como preto (no caso de uma imagem em tons de cinza). Este tipo de segmentação pode ser utilizado com mais valores de limiares. Quanto maior o número de limiares, mais a imagem segmentada tende a parecer-se com a original.

Geralmente, os clusteres são definidos no histograma por picos (Gaussianas) que enfatizam os principais valores dos pixels presentes na imagem. A segmentação utilizando limiares no histograma tem resultados melhores, caso estes picos estejam bem definidos e separados no histograma. Caso contrário, a sobreposição destes clusteres pode resultar em falhas na segmentação. Outra forma efetiva de utilizar o histograma na segmentação de imagens é classificar os pixels de acordo com os picos do histograma utilizando a "Distância de Mahalanobis".

A separação relativa destes picos do histograma, por vezes é medida pela Distância de Mahalanobis (MAHALANOBIS, 1936). Esta medida é utilizada para determinar a probabilidade de um pixel pertencer a um determinado conjunto (*background* ou *foreground*, por exemplo). Se especificada a "localização" do pixel de acordo com sua intensidade  $f$ , e a média e desvio padrão do pico do histograma referente a uma classe  $n$  de pixels como  $\mu_n$  e  $\sigma_n$  respectivamente, então podemos pensar na distância de Mahalanobis de forma intuitiva. O cálculo da distância normalizada no histograma pode ser definido pela Equação (4).

$$d_n \equiv \frac{f - \mu_n}{\sigma_n} \quad (4)$$

O mínimo entre conjunto de distâncias  $d_n$  representa a qual classe este pixel deve pertencer. (Yong ZHANG et al., 2011) utiliza este conceito da distância de Mahalanobis em conjunto com o PSO (Seção 3.3.1) na segmentação de imagens

### 3.2.2 Imagens coloridas e imagens espectrais

Imagens coloridas são compostas de diversos canais que compõem as propriedades de cada pixel. A quantidade e natureza destes canais varia de acordo com o espaço de cores utilizado (Seção 3.1). Tomando como exemplo uma imagem RGB, a mesma pode ser descrita como  $f[x,y,\lambda_n]$  onde  $n$  representa o índice para os canais vermelho, verde e azul. A imagem pode ser descrita por (5).

$$g[x,y] = \alpha \cdot f[x,y,\lambda_r] + \beta \cdot f[x,y,\lambda_g] + \gamma \cdot f[x,y,\lambda_b] \quad (5)$$

Onde os coeficientes  $[\alpha, \beta, \gamma]$  são constantes determinadas pelo filtro de espectro e sensibilidade do meio de obtenção da imagem. No caso do RGB, a imagem pode ser observada como a junção de 3 imagens, cada uma correspondendo às intensidades dos pixels com relação a um canal. Normalmente, a intensidade de um mesmo pixel, nas três imagens é diferente, mas os valores dos pixels vizinhos são correlacionados. Por exemplo, para um objeto pred-

minantemente verde, a intensidade do canal verde será alta, enquanto dos demais canais será baixa.

Tratando de processamento de imagens coloridas, as abordagens possíveis caem em duas principais categorias (GONZALEZ; WOODS; MASTERS, 2009). Na primeira, cada componente (canal) da imagem é processado separadamente e então combinados para obtenção do resultado final. Na segunda, é trabalhando diretamente com os pixels coloridos. Como imagens coloridas possuem ao menos três componentes, os pixels podem ser considerados como vetores. Utilizando o sistema RGB como exemplo, um ponto de cor no espaço de cores é definido por um vetor, que vai da origem à coordenada no espaço RGB. Um exemplo pode ser observado na Equação (6)

$$c = \begin{bmatrix} c_R \\ c_G \\ c_B \end{bmatrix} = \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (6)$$

Esta equação indica que os componentes de  $c$  são os componentes RGB de um ponto da imagem. Levando em consideração que os componentes de cores são funções das coordenadas  $(x,y)$ , tem-se a Equação (7)

$$c(x,y) = \begin{bmatrix} c_R(x,y) \\ c_G(x,y) \\ c_B(x,y) \end{bmatrix} = \begin{bmatrix} R(x,y) \\ G(x,y) \\ B(x,y) \end{bmatrix} \quad (7)$$

Para uma imagem de tamanho  $M \times N$ , há  $MN$  vetores  $c(x,y)$ , para  $x = 0,1,2,\dots,M - 1$ ;  $y = 0,1,2,\dots,N - 1$ . O fato de os pixels possuírem mais de um componente, deve-se determinar a forma adequada de representar e interpretar cada um destes componentes no espaço  $(x,y)$  da imagem. Um meio simples de fazer isto é tratar cada componente separadamente, sendo assim possível aplicar as mesmas técnicas utilizadas no histograma em tons de cinza, em cada um dos canais que compõem a imagem. Porém, o resultado de analisar cada componente individualmente nem sempre é equivalente ao processamento do pixel como um vetor no espaço de cores.

### 3.2.3 Segmentação no espaço de cores HSI/HSV

No caso da segmentação de imagens utilizando o espaço de cores HSI, Gonzalez, Woods e Masters (2009) afirma que este espaço de cores é conveniente à segmentação pois as cores estão representadas em uma única dimensão (matiz). Normalmente, a saturação é utilizada como um segundo recurso para isolar regiões de interesse dentro do canal da matiz. Por sua vez, a intensidade é menos utilizada para segmentação pois não contém informação de cor. De acordo com Cheng et al. (2001), no sistema HSI, a matiz não pode ser identificada caso a saturação seja zero, e a saturação não pode ser identificada caso a intensidade seja zero.

Para segmentar objetos de diferentes cores, pode-se aplicar os algoritmos de segmentação apenas ao componente da matiz (hue), identificando um limiar que seja capaz de distinguir os objetos em questão. Este método é particularmente eficiente em imagens de iluminação não uniforme, como sombras, visto que a matiz é representada de forma independente à intensidade. E se tratando de esforço computacional, a segmentação baseada na matiz (1 dimensão) exige menor processamento comparado ao RGB (3 dimensões). A desvantagem do uso desta técnica é a inconsistência crescente dos valores da matiz, conforme se aproximam do eixo central do cone que representa o HSI (baixa saturação), causando descontinuidade na representação das cores (CHAPRON, 1992). Além disso, quando os valores de intensidade se aproximam de seus extremos (branco ou preto), a matiz e saturação perdem significância na segmentação.

No quesito de segmentação de imagens, os espaços HSI e HSV trabalham de forma bastante semelhante, pois ambos mantêm sua matiz em um canal separado e portanto é facilmente isolada das demais características. Utilizando como exemplo o HSV, como conforme menor o valor da saturação, menor também a diferenciação de cores baseadas na matiz. Sendo assim uma abordagem possível para o problema é a determinação de um limiar de saturação que define até que ponto a matiz é relevante na classificação (SURAL; QIAN; PRAMANIK, 2002).

### 3.2.4 Segmentação no espaço de cores RGB

Sendo o objetivo da segmentação classificar os pixels de uma imagem RGB em grupos de acordo com sua similaridade, é necessário primeiramente especificar uma forma de medir esta similaridade. Uma das formas de fazer isto é utilizando a distância euclidiana.

Sejam  $a$  e  $b$  dois pontos aleatórios no espaço de cores RGB.  $a$  e  $b$  pertencerão á uma mesma classe caso a distância entre eles não ultrapasse um valor de limiar estabelecido. Esta forma de representação de similaridade permite estabelecer um limiar unidimensional  $D_0$ , pois analisa a unidade de distância entre as cores, ao invés de suas componentes RGB, ainda que a distância  $D(a,b)$  seja  $f(R,G,B)$ , como apresentado na Equação (8) (GONZALEZ; WOODS; MASTERS, 2009).

$$D(a,b) = \sqrt{(a_R - b_R)^2 + (a_G - b_G)^2 + (a_B - b_B)^2} \quad (8)$$

Onde  $R$ ,  $G$  e  $B$  são os componentes RGB de seus respectivos pontos do espaço de cores. A área do espaço de cores em que  $D(a,b) \leq D_0$  é uma esfera de raio  $D_0$  com seu centro definido em  $a$ . Os pontos internos e na superfície desta esfera satisfazem as condições de similaridade com  $a$ , e os externos não. Classificando estes pontos na imagem como duas diferentes classes, o resultado é uma imagem segmentada binária (preto e branco).

Um problema na utilização das coordenadas RGB para determinação de similaridade entre as cores, é que sua distância euclidiana não é perceptualmente uniforme, ou seja, as diferenças entre as cores percebidas pela observação humana não são proporcionais à distância euclidiana entre estas cores no espaço RGB. Para reduzir este problema, foram introduzidos

os modelos de cor CIE( $L^*a^*b^*$ ) e CIE( $L^*u^*v^*$ ), que têm o intuito de que a distância entre as cores em seu espaço seja aproximadamente equivalente às diferenças percebidas pelos seres humanos.

### 3.3 ALGORITMOS BIOINSPIRADOS

De acordo com Binitha e Sathya (2012) as metodologias de solução de problemas, são classificadas em duas principais categorias: Métodos exatos (lógico-matemáticos) e heurísticas. As heurísticas se mostram superiores em problemas complexos de otimização, onde os métodos exatos falham ou se tornam inviáveis devido ao esforço computacional. Os algoritmos bio-inspirados são meta-heurísticas que se baseiam nas estratégias adotadas pela natureza para processos que podem ser vistos como problemas de otimização.

Alguns exemplos de algoritmos bio-inspirados são: *particle swarm optimization* (PSO) *firefly* (FF), *ant colony* (AC), entre outros. De acordo com X.-S. Yang (2010b) a maioria dos algoritmos clássicos são determinísticos, e parte deles são baseados em informações de gradiente. Tais algoritmos funcionam bem com problemas unimodais contínuos, porém, encontram problemas caso haja alguma descontinuidade na função objetivo. Pelo fato dos algoritmos bio-inspirados serem métodos estocásticos (meta-heurísticas), os mesmos não compartilham esta limitação.

#### 3.3.1 Particle swarm optimization (PSO)

*Particle Swarm Optimization* (PSO) (KENNEDY, 2010) é uma técnica de otimização estocástica baseada em população, inspirada no comportamento de pássaros e cardumes de peixes.

O modelo básico de PSO (KENNEDY, 2010) consiste em um conjunto de partículas dispostas em um espaço de busca que representam as possíveis soluções do problema proposto. Estas partículas interagem entre si de acordo com sua função de aptidão a fim de convergir para um resultado adequado.

As partículas são iniciadas com valores aleatórios (posições) e movem-se iterativamente pelo espaço de dimensão  $d$  definida pelo problema em busca da solução. A proximidade da solução representada pela partícula com relação à solução desejada é medida por sua função de aptidão  $f$  também definida pelo problema.

Cada partícula tem sua posição representada pelo vetor  $X_i$  (sendo  $i$  o índice da partícula), e a velocidade representada pelo vetor  $V_i$ . As melhores posições visitadas por cada partícula (definidas pela função de aptidão  $f$ ) são armazenadas em um vetor de  $i$  posições, bem como seus valores  $d$ -dimensionais (pbest).

Durante o tempo de iteração  $t$ , a alteração de sua velocidade anterior para a atual é determinada pela Equação 9

$$V_{i+1} = w * v_i + c_1 * r_1 * (P_g - x_i) + c_2 * r_2 * (P_i - x_i) \quad (9)$$

Onde  $i = 1, 2, \dots, N$ ,  $w$  é o peso da inércia,  $r_1$  e  $r_2$  são números aleatórios definidos no intervalo  $[0, 1]$  a fim de manter a diversidade da população,  $c_1$  é uma constante positiva conhecida como coeficiente do componente social,  $c_2$  é uma constante positiva conhecida como coeficiente do componente de autoconhecimento,  $P_g$  é a melhor posição encontrada até o momento considerando todas as partículas, e  $P_i$  é a melhor posição visitada pela partícula em questão. A nova posição da partícula é definida por 10 (CLERC, 2010)

$$X_{i+1} = x_i + V_{i+1} \quad (10)$$

Na Equação 9, a partícula decide para onde deve se mover levando em consideração sua própria experiência, que é sua melhor posição visitada até o momento ( $P_i$ ), e a experiência do enxame, que é a melhor posição visitada em todo sistema ( $P_g$ ). A fim de guiar as partículas efetivamente na exploração das soluções do sistema, sua movimentação durante uma iteração deve ser limitada por uma velocidade máxima ( $V_{max}$ ) (KENNEDY, 2010)

Ou seja, cada partícula mantém o registro de suas coordenadas no espaço de busca associadas ao melhor resultado encontrado por ela até o momento. Este registro é levado em consideração em conjunto com a análise da aptidão de suas partículas vizinhas e o melhor resultado atingido considerando todo o sistema para a determinação da velocidade e direção do movimento da partícula em questão. Desta forma a partícula tende a se aproximar das áreas mais promissoras do espaço de busca. Na formulação deste movimento há também um fator de aleatoriedade presente que favorece a exploração do espaço de busca evitando que as partículas movam-se linearmente.

O algoritmo pode ser descrito da seguinte forma:

**Entrada:** Parâmetros de inicialização do algoritmo ( $c_1$ ,  $c_2$ ,  $w$ ,  $V_{max}$ , tamanho da população, máximo de iterações,  $r_1$  e  $r_2$ )

**Saída:** A otimização tendo o maior valor de aptidão encontrado.

**Passo 1:** Gerar aleatoriamente partículas e suas respectivas velocidades.

**Passo 2:** Calcular a aptidão de cada partícula.

**Passo 3:** Se a posição atual da partícula é melhor que seu histórico anterior, atualizar as partículas para refletirem este fato.

**Passo 4:** Encontrar a melhor partícula do enxame, atualizar as posições das partículas segundo as Equações 9 e 10.

**Passo 5:** Voltar ao passo 2 até que o número máximo de iterações seja excedido, ou o valor ótimo de aptidão for encontrado.

**Passo 6:** Retornar o melhor valor encontrado e finalizar.

Para a implementação do PSO, há dois passos importantes que definem o funcionamento do algoritmo: A representação da solução (definição do espaço de busca) e a função de avalia-

ção. Estes os que requerem maior atenção pois são fundamentais para a movimentação efetiva das partículas dentro do espaço de busca a fim de guiarem-se às áreas com maior probabilidade de encontrar uma solução ótima.

O PSO, assim como outros algoritmos evolucionários pode ser aplicado para resolver a maioria dos problemas de otimização e que podem ser convertidos em problemas de otimização. O PSO possui aplicações em diversas áreas, como reconhecimento de padrões, segmentação de imagens e otimização multi-objetiva. Outro motivo pelo qual o PSO é adotado é que não há muitos parâmetros complexos que necessitem ser ajustados. Uma configuração com pequenas variações paramétricas pode funcionar bem em uma ampla variedade de aplicações.

Diversos trabalhos como Hu e R. Eberhart (2002), R. C. Eberhart e Shi (2001) e Shi e R. C. Eberhart (2001) exploram o modelo fundamental do PSO e o adaptam para aplicações mais específicas aprimorando sua busca de resultados. Esta maleabilidade do algoritmo gera inúmeras possibilidades de implementações diferentes em diversas áreas de interesse.

### 3.3.2 Algoritmo Firefly

O algoritmo Firefly é um algoritmo de meta-heurística inspirado no comportamento de vagalumes (*fireflies*). Este algoritmo foi introduzido inicialmente por X. Yang (2009) que o descreve a partir das seguintes regras: (1) Todos fireflies são unissexuais, assim qualquer *firefly* pode ser atraído por qualquer outro; (2) A atração é proporcional ao brilho, de forma que um *firefly* de brilho menor tenda a voar em direção a seu vizinho mais brilhante. No caso do *firefly* ser o mais brilhante, o mesmo se moverá aleatoriamente.

O conceito básico do *Firefly* é de modelar problemas de otimização não linear associando as variáveis do problema aos *fireflies*. Cada *firefly* possui um "brilho" que é determinado de acordo com uma função objetivo definida pelo problema que analisa a aptidão de cada *firefly* representar a solução ótima em questão. Então, a cada iteração, os valores das variáveis dos *fireflies* são atualizadas respeitando um conjunto de regras até a convergência. Para cada geração, seguem-se os seguintes procedimentos:

- (1) Avaliação do brilho;
- (2) Cálculo das distâncias entre cada par de *fireflies*;
- (3) Movimentação de atração dos *fireflies* de acordo com seu brilho;
- (4) Armazenamento do melhor resultado;
- (5) Geração aleatória de novas soluções.

A função de avaliação do brilho dos *fireflies* é sua parte principal (Kernel) e é determinada de acordo com o problema em questão. Para problemas de maximização, o brilho pode simplesmente ser proporcional à função objetivo. Outras formas podem ser definidas de forma similar à função de avaliação dos algoritmos genéticos (DEB et al., 2002). O algoritmo Firefly (FA) pode ser observado em (1).

### Algoritmo 1 – Algoritmo Firefly

Função objetivo  $f(x), x = (x_1, \dots, x_d)^T$   
 Gerar população inicial de *fireflies*  $x_i (i = 1, 2, \dots, n)$   
 Brilho  $I_i$  em  $x_i$  é determinado por  $f(x_i)$   
 Definição do coeficiente de absorção de luz  $\gamma$   
**while**  $t < MaxGen$  **do**  
   **for**  $i \leftarrow 1, n$  **do**  
     **for**  $j \leftarrow 1, i$  **do**  
       **if**  $I_j > I_i$  **then**  
         Mover firefly  $i$  em direção de  $j$  em d-dimensões;  
       **end if**  
       Atração varia com a distância  $r$  por  $exp[-r\gamma]$   
       Avaliar novas soluções e atualizar intensidade do brilho  
     **end for**  
**end for**  
 Ranquear os *fireflies* e encontrar o melhor atual  
**end while**

De acordo com X.-S. Yang (2010b), além da função kernel, outros dois pontos importantes são a variação da intensidade de luz e a formulação da atração. Normalmente assume-se que a atração é determinada pelo brilho, que é associado à função objetivo. O brilho de um *firefly* é um valor determinado associado a ele naquela posição, porém, sua atratividade é relativa aos demais. Portanto, irá variar de acordo com a distância  $r_{ij}$  entre os *fireflies*  $i$  e  $j$ . Adicionalmente, a intensidade da luz diminui conforme a distância de sua fonte e a luz é absorvida no meio em que se propaga. Desta forma, a atratividade deve variar de acordo com o grau de absorção. A intensidade da luz  $I(r)$  varia de acordo com a raiz inversa  $I(r_{ij}) = \frac{I_s}{r_{ij}^2}$ , onde  $I_s$  é a intensidade da luz na fonte. Para um coeficiente de absorção fixo  $\gamma$ , a intensidade da luz  $I$  varia com a distância  $r$  como apresentado na Equação 11, onde  $I_0$  é a intensidade de luz original.

$$I = I_0 e^{-\gamma r} \quad (11)$$

A atratividade é proporcional à intensidade da luz vista pelos *fireflies* adjacentes, de forma que o coeficiente  $\beta$  pode ser definido pela Equação (12), onde  $\beta_0$  é a atratividade em  $r_{ij} = 0$

$$\beta = I \beta_0 e^{-\gamma r^2} \quad (12)$$

A distância  $r_{ij}$  entre dois *fireflies*  $i$  e  $j$  nas posições  $x_i$  e  $x_j$ , respectivamente, é definida pela Equação (13).

$$r_{ij} = \|x_i - x_j\| = \sqrt{\sum_{k=1}^d (x_{i,k} - x_{j,k})^2} \quad (13)$$

O movimento de um *firefly*  $i$  atraído por um outro *firefly* mais brilhante  $j$  é determinado pela Equação (14).

$$x_i = x_i + \beta_0 e^{-\gamma r_{ij}^2} (x_j - x_i) + \alpha \epsilon_i \quad (14)$$

Onde  $\beta_0$  é a atratividade na distância  $r = 0$ ,  $\gamma$  é o coeficiente de absorção de luz,  $\alpha$  é o parâmetro de aleatoriedade, e  $\epsilon_i$  é um vetor de número aleatórios; ou seja, o segundo termo da equação é devido à atração, e o terceiro termo é o fator randômico da movimentação.

Basicamente, o parâmetro  $\gamma$  caracteriza a variação da atratividade, e seu valor é importante na determinação da velocidade de convergência e como o algoritmo se comporta (X.-S. YANG, 2010a). Na prática,  $\gamma = O(1)$  é determinado pelo tamanho  $S_k$  do sistema a ser otimizado. Em um extremo, quando  $\gamma \rightarrow 0$  a atratividade é constante  $\beta = \beta_0$ . Desta forma a intensidade da luz não decai com a distância do observador e um *firefly* pode ser observado a partir de todo o domínio fazendo com que o modelo passe a se assemelhar bastante com um caso específico de PSO. Se a iteração interna de  $j$  for removida e  $I_j$  for substituído por um máximo global  $g$ , o FA torna-se o PSO padrão e, conseqüentemente, sua eficiência passa a ser a mesma também. Por outro lado, se  $\gamma \rightarrow \infty$ , temos que  $\beta(r) \rightarrow \delta(r)$ . Isto significa que a atratividade é quase zero vista pelos outros *fireflies*; ou seja, os *fireflies* têm um campo de visão muito curta e sofrem pouca ou nenhuma influência dos demais. Neste caso é como se todos os *fireflies* se movessem de forma aleatória. O que corresponderia ao método de busca aleatória. Então  $\gamma$  parcialmente controla a forma que o algoritmo se comporta.

Em estudos como X. Yang (2009) e Horng e Liou (2011), o Firefly obteve resultados superiores de otimização se comparado á outros algoritmos bioinspirados com o mesmo propósito, obtendo resultados similares à busca intensiva em menor tempo de processamento.

### 3.3.2.1 Firefly aplicado à segmentação

O *firefly* é o algoritmo utilizado neste trabalho para definição dos limiares que compõem a segmentação. Nesta seção é abordado sua aplicação específica a este caso.

Para cada um dos canais HSV da imagem, o *firefly* busca os valores ótimos de separação do histograma segundo sua distribuição, utilizando a função de avaliação parametrizada (*kernel*). Como exemplo inicial será utilizada a entropia não-extensiva de Tsallis. O funcionamento desta função pode ser observado no Algoritmo 2 de P. S. Rodrigues et al. (2015). Neste algoritmo,  $Z$  representa a função de avaliação, ou seja, o *kernel*. A dimensão  $d$  é dada pelo número de limiares desejados, onde cada dimensão do resultado de saída representa o valor de um limiar.

Nesta implementação, a cada iteração  $t$ , a solução  $f_i$  depende principalmente da solução  $f_j$  e a diferença entre seus brilhos  $r_{i,j}$  para todos *fireflies*  $j$ . O brilho de um *firefly*  $i$  é atualizado quando  $i$  é menos brilhante que qualquer outro *firefly*  $j$ . A influência sofrida neste caso é determinada pelo fator de atração  $\beta$ , o coeficiente de absorção  $\gamma$  e o tamanho do passo  $\alpha$ .

### Algoritmo 2 – Firefly aplicado à definição de múltiplos limiares

Entrada:  $n$ : Número de vagalumes em cada geração;  $d$ : dimensão do problema;  $\gamma$ : coeficiente de absorção;  $\alpha$ : passo;  $\beta$ : fator de atração;  $\Gamma$ : número máximo de gerações

Saída: Conjunto de limiares  $f_i^* = \{x_1^i, x_2^i, \dots, x_d^i\}$

iniciar as variáveis:  $t = 0, \alpha = 1.0$

Gerar população inicial aleatória  $\{f_1, f_2, \dots, f_n\}$

Para cada *firefly*  $f_i$ , calcular o brilho a partir da função de avaliação  $Z(f_i)$

**while**  $t < \Gamma$  **do**

**for**  $i = 1 \rightarrow n$  **do**

**for**  $j = 1 \rightarrow n$  **do**

      Calcular a distância  $r_{i,j}$  entre os brilhos  $Z(f_i)$  e  $Z(f_j)$ ;

**end for**

**end for**

**for**  $i = 1 \rightarrow n$  **do**

    Função de avaliação  $Z(f_i)$

**for**  $j = 1 \rightarrow n$  **do**

**if**  $Z(f_i) < Z(f_j)$  **then**

        Mover firefly  $f_i$  em direção de  $f_j$  de acordo com a seguinte regra de atualização:

        Gerar nova solução aleatória  $\mu = \{x_1, x_2, \dots, x_d\}$

$\alpha_t = \alpha \alpha_t$

$\beta_0 = \beta \exp(-\gamma r_{i,j}^2)$

**for**  $k = 1 \rightarrow d$  **do**

$x_k^i = (1 - \beta_0)x_k^i + \beta_0 x_k^j + \alpha_t \mu_t$

**end for**

**end if**

**end for**

**end for**

  Para cada *firefly*  $f_i$ , calcular seu brilho  $Z(f_i)$

  Ordenar os *fireflies* de acordo com seu brilho

  Considerar o *firefly* com maior brilho  $f_i^*$  como o resultado atual

$t = t + 1$

**end while**

O tipo de função de avaliação  $Z$  que determina o brilho tem grande influência no resultado e pode ser alterada de acordo com a necessidade do problema independentemente do *firefly*.

#### 3.3.3 Colônia de formigas

O algoritmo Ant Colony (AC) (DORIGO et al., 2008) é um método de solução para problemas de otimização combinatorial (COP) baseado no comportamento de formigas e na forma que utilizam para guiarem-se através de uma substância que expelem chamada feromônio. O algoritmo de colônia de formigas é aplicado à problemas de otimização combinatorial. Um problema de otimização combinatorial é definido pelo terceto  $(S, \Omega, f)$  (DORIGO, 2007), onde

$S$  é o espaço de busca,  $\Omega$  é o conjunto de restrições do problema e  $f$  é a função objetivo a ser minimizada.

Define-se o espaço de busca  $S$ , é dado um conjunto de variáveis discretas  $X_i, i = 1, \dots, n$ . Os elementos de  $S$  são atribuições em que cada variável  $X_i$  possui um valor  $v_i^j$  de seu domínio  $D_i$ , onde  $i$  e  $j$  representam os índices de identificação do item do espaço de busca em questão. O conjunto de soluções viáveis  $S_\Omega \subset S$  é dado pelos elementos de  $S$  que satisfazem o conjunto de restrições  $\Omega$ . Uma solução  $s^* \in S_\Omega$  é a solução ótima se, e somente se,  $f(s^*) \leq f(s) \forall s \in S_\Omega$ .

O conjunto de todas soluções ótimas é denotado por  $S_\Omega^* \subseteq S_\Omega$ . A solução de um COP requer ao menos um  $s^* \in S_\Omega^*$ .

No AC, "formigas" artificiais constroem uma solução para um COP transitando em um grafo de construção totalmente conectado, chamado de grafo de construção, definido da seguinte forma. Cada variável de decisão  $X_i$  é chamada de componente de solução e denotada por  $c_{ij}$ , o conjunto de todos  $c_{ij}$  constitui  $C$ . O grafo de construção  $G_c(V, E)$  é criado associando os componentes  $C$  com o conjunto de vértices  $V$  ou de arestas  $E$ . Um valor de feromônio  $\tau_{ij}$  é associado a cada  $c_{ij}$ . Estes valores de feromônio são atualizados pelo algoritmo durante a busca pela solução. As formigas se movem pelo grafo de construção influenciadas pelo valor de feromônio encontrado. Cada formiga adiciona um certo valor de feromônio nos vértices e arestas visitados. A quantidade  $\Delta_\tau$  depositada depende da qualidade da solução encontrada. Analogamente, as formigas seguintes utilizam-se desta informação para guiarem-se em direção às regiões mais promissoras do espaço de busca (YASEEN; AL-SLAM, 2008).

A atualização do feromônio tem o objetivo de fazer com que as soluções melhores do espaço de busca e seus vizinhos sejam mais explorados pelas iterações seguintes. Para isto existem dois mecanismos que são usados para garantir que as formigas guiem-se às áreas mais promissoras. A primeira, que é o depósito de feromônio aumenta a quantidade de feromônio da solução dos componentes associados a um conjunto  $S_{upd}$  de boas soluções, fazendo que tais soluções se tornem mais atrativas às formigas na próxima iteração. O segundo mecanismo é a evaporação das trilhas de feromônio, que faz decair com o tempo a quantidade de feromônio previamente depositados pelas formigas. Este mecanismo se faz necessário para evitar a rápida convergência do algoritmo em uma região sub-otimizada do espaço de busca. Implementando assim uma forma útil de "esquecimento" dos caminhos anteriores, favorecendo a exploração de novas áreas. A atualização do feromônio é comumente implementado segundo a Equação (15)

$$\tau_{ij} = (1 - \rho)\tau_{ij} + \sum_{s \in S_{upd} | c_i^j \in s} g(s) \quad (15)$$

Onde  $S_{upd}$  é o conjunto de soluções consideradas promissoras (pela quantidade de feromônio,  $\rho \in (0, 1]$  é a taxa de evaporação,  $g(\cdot) : S \mapsto \mathbb{R}^+$  é uma função tal que  $f(s) < f(s') \Rightarrow g(s) \geq g(s')$ . Esta função determina a qualidade da solução e é chamada de função de avaliação.

Os algoritmos AC normalmente diferem na forma de atualização do feromônio. Diferentes especificações de como determinar  $S_{upd}$  resultam em diferentes formas da Equação (15). Tipicamente  $S_{upd}$  é um subconjunto de  $S_{iter} \cup \{s_{gb}\}$ , onde  $S_{iter}$  é o conjunto de todas soluções construídas na iteração presente, e  $S_{gb}$  é a melhor solução encontrada até o momento (DORIGO; BIRATTARI, 2010).

A construção das soluções é um processo iterativo iniciando com um conjunto vazio de soluções parciais  $s^p = \emptyset$ . A cada iteração,  $S_p$  é estendida adicionando uma possível componente da solução dentre o conjunto de vizinhos possíveis  $N(s^p) \subseteq C$ . O processo de construção de soluções pode ser considerado como um caminho no grafo  $G_c(V, E)$ .

A escolha de um componente de solução de  $N(s^p)$  é feita probabilisticamente a cada passo e guiado por um mecanismo estocástico, onde é influenciado pela quantidade de feromônio presente em cada elemento de  $N(s^p)$ .

As regras para a escolha destes componentes variam de acordo com a aplicação. Uma das regras mais conhecidas segundo (DORIGO; MANIEZZO; COLORNI, 1996) para escolha da solução é apresentada na Equação (16).

$$p(c_{ij}|s^p) = \frac{\tau_{ij}^\alpha * \eta_{ij}^\beta}{\sum_{c_{ij} \in N(s^p)} \tau_{ij}^\alpha * \eta_{ij}^\beta}, \forall c_{ij} \in N(s^p) \quad (16)$$

Onde  $\tau_{ij}$  e  $\eta_{ij}$  são respectivamente o valor do feromônio e o valor da heurística associada com o componente  $c_{ij}$ . E  $\alpha$  e  $\beta$  são valores reais positivos que ponderam  $\tau_{ij}$  e  $\eta_{ij}$  definindo sua importância.

O trabalho inicial no AC foi fundamentado em experimentos com o objetivo de demonstrar que as ideias presentes nesta técnica podem originar algoritmos que funcionem bem.

### 3.4 MÉTODOS ENTRÓPICOS

Em teoria da informação, a distribuição probabilística de eventos juntamente com a quantidade de informação presente em cada evento define uma variável aleatória cujo valor esperado é a média da quantidade de informação (entropia) gerada por sua distribuição. Os métodos entrópicos destinam-se a medida de incerteza de uma variável aleatória. A entropia não depende dos valores assumidos por esta variável aleatória, e sim sobre a probabilidade destes valores ocorrerem.

#### 3.4.1 Entropia de Shannon

A teoria da entropia de Shannon tem sido utilizada em diversas aplicações desde que C. Shannon a propôs em teoria da informação (SHANNON, 2001). Formalmente, seja  $P = \{p_1, p_2, \dots, p_L\}$  uma distribuição de probabilidade, Shannon denota sua entropia  $S(P)$  segundo a Equação (17).

$$S(P) = \sum_{i=1}^L p_i \log(p_i) \quad (17)$$

Pun (1981) aplicou esta idéia para o caso específico de definição de um limiar através das características da entropia de Shannon: sejam  $P_1 = \{p_1^1, p_2^1, \dots, p_t^1\}$  e  $P_2 = \{p_1^2, p_2^2, \dots, p_{t+L}^2\}$  subdivisões de  $P$ . Se assumirmos que  $P_1$  e  $P_2$  são distribuições independentes, a entropia  $S(P)$  pode ser calculada pela regra aditiva, dada pela Equação (18).

$$S(P) = S(P_1) + S(P_2) \quad (18)$$

Sob as restrições  $\sum_i^t p_i^1 = 1$  e  $\sum_{i=t+1}^L p_i^2 = 1$ , após as normalizações apropriadas.

O trabalho de Pun (1981) trata sobre a segmentação de imagens em tons de cinza, realizando testes com imagens naturais e artificiais a fim de observar o comportamento do método em diferentes tipo de histogramas gerados. No caso do uso de apenas um limiar, o limiar ótimo  $t^*$  é aquele que maximiza a Equação (18).

### 3.4.2 Entropia de Tsallis

A entropia de Tsallis (TE) é uma generalização da entropia de Shannon (SE). Uma prova dessa generalização pode ser vista em Tavares (2003). A entropia não-extensiva de Tsallis de uma distribuição  $P$ , denotada por  $S_q(p)$ , de tamanho  $L$ , é definida pela Equação (19)

$$S_q(P) = \frac{1 - \sum_i^L p_i^q}{1 - q} \quad (19)$$

A principal característica observada da Equação (19) é a introdução de um parâmetro real  $q$ , chamado de parâmetro não-extensivo. Em Tsallis (1999) é mostrado que, no limite  $q \rightarrow 1$ , a Equação (19) reduz-se à Equação (17).

O trabalho de Albuquerque (2004) propõe um método de segmentação de imagens semelhante ao apresentado em Kapur, P. K. Sahoo e Wong (1985), porém utilizando a entropia não-extensiva de Tsallis como critério de limiarização. Para isto, o autor considera duas distribuições de probabilidades distintas. Uma para o *background* e outra para o *foreground*. O limiar ótimo considerado é a maximização da soma das entropias destas duas distribuições de probabilidade.

P. S. Rodrigues et al. (2015) apresenta uma versão generalizada do método de Pun (1981) para segmentação com múltiplos limiares utilizando a entropia de Tsallis. Dada a Equação (18), o autor apresenta a Equação (20) que é útil para a definição de múltiplos limiares.

$$S_q(H_1 * \dots * H_d + 1) = S_q(H_1) + \dots + S_q(H_{d+1}) + (1 - q)S_q(H_1)S_q(H_2)\dots S_q(H_d + 1) \quad (20)$$

O autor parte da Equação (20) e define um método de segmentação utilizando o algoritmo *firefly*, método este apresentado na Seção 3.3.2.1.

O trabalho de Yudong Zhang e Wu (2011) utiliza a mesma técnica também para definição de múltiplos limiares baseado na entropia de Tsallis. Juntamente com o algoritmo bio-inspirado *Artificial Bee Colony* (ACB) na segmentação de imagens em tons de cinza. O autor aplica o ACB utilizando a entropia de Tsallis como função de avaliação e assim como P. S. Rodrigues et al. (2015), obtém (devido à meta-heurística) resultados com performance computacional melhor do que os métodos exaustivos.

## 4 PROPOSTA

Este capítulo apresenta a proposta do trabalho, explicando os processos utilizados para a segmentação de imagens.

### 4.1 DEFINIÇÃO DO PROBLEMA E OBJETIVO

A ideia geral deste trabalho é a utilização do algoritmo *Firefly* (Seção 3.3.2) para a solução do problema de particionamento de histogramas de imagens multi-espectrais, visando multi-limiarização.

### 4.2 BASE DE DADOS

Como base de dados para os testes será utilizada a base de Berkeley (MARTIN et al., 2001), disponibilizada pela Universidade de Berkeley. Esta base contém 300 imagens de cenas naturais e suas respectivas segmentações realizadas manualmente para fins comparativos dos resultados. É uma base de dados bastante utilizada nos estudos de segmentação de imagens, o que permite a comparação do método apresentado com outros trabalhos, inclusive com relação à segmentação de imagens coloridas.

### 4.3 METODOLOGIA

A metodologia para segmentação de imagens proposta neste trabalho baseia-se nas técnicas de limiarização de histograma. Dada uma imagem colorida  $I$ , utilizando o espaço de cores HSV, sabe-se que esta imagem é a composição de 3 canais (H, S e V) (Seção 3.2.3, onde cada um possui sua própria distribuição de probabilidades (histograma) independente.

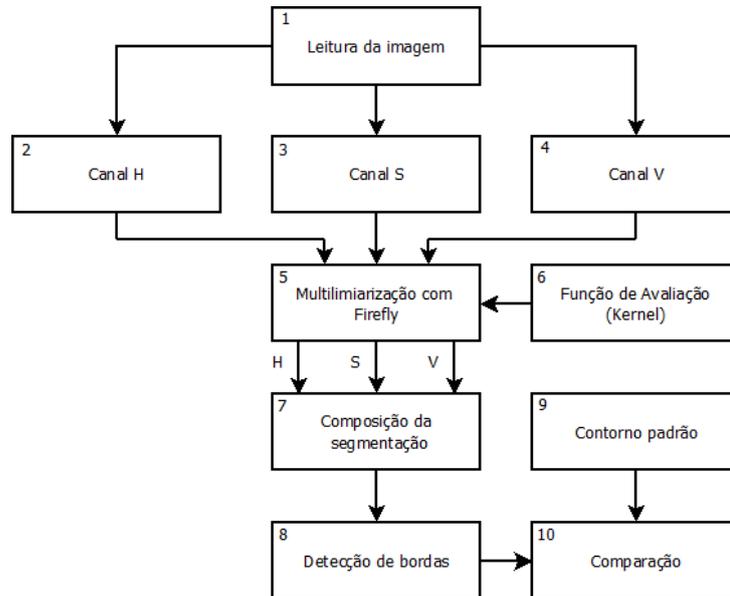
A estratégia principal é realizar a multi-limiarização independente em cada canal de cor, de acordo com a proposta de multi-limiarização descrita na Seção 3.3.2.1, baseada nos trabalhos de P. S. Rodrigues et al. (2015), X. Yang (2009) e Horng e Liou (2011). A proposta será estudada sob Kernel entrópico não-extensivo (Seção 3.4.2).

A base de dados utilizada nos experimentos será a base de Berkeley (Seção 4.2). Uma vez que trata-se de uma base de dados anotada manualmente, o que permite avaliação e comparação dos resultados com outras metodologias já fundamentadas. Por outro lado, a base de Berkeley é a mais conhecida base anotada para segmentação de baseada em cor.

A Figura 5 ilustra a ideia geral da metodologia de segmentação proposta e análise dos resultados.

A imagem obtida no processo 1 da Figura 5 é composta de três canais independentes. Os histogramas destes canais (itens 2, 3 e 4 da Figura 5) são obtidos e submetidos ao método

Figura 5 – Fluxograma da metodologia proposta



Fonte:Autor.

de determinação dos limiares utilizando o *Firefly* apresentado na Seção 3.3.2.1 representado na Figura 5 pelo processo de número 5.

Como parametrização do método, é necessário informar o número de limiares desejado, o tipo de função de avaliação utilizado como *kernel* no *firefly* (item 6 da Figura 5), o número de iterações do *firefly*, número de vagalumes, fator de atração entre os vagalumes, tamanho do passo e o coeficiente de absorção. O resultado é o conjunto de limiares definidos para cada canal.

Estes limiares obtidos serão utilizados não apenas para a segmentação da imagem, mas também para a definição dos valores presentes na composição da imagem segmentada (processo 7 da Figura 5), dadas as características do espaço de cor utilizado.

Para isto, partindo dos histogramas dos canais da imagem e dos limiares definidos pelo Algoritmo 2, são estabelecidos valores representantes de cada classe segundo o Algoritmo 3, que calcula o valor esperado dentro de cada região dividida do histograma.

Assim, os valores  $V$  obtidos fazem o resultado da segmentação preservar características como sua distribuição de cores e intensidades.

Algoritmo 3 – Definição dos valores representantes das classes definidas

Entrada: Número de limiares  $\alpha$ ; Histograma  $H = \{H_1, H_2, \dots, H_n\}$ ; Conjunto de limiares

$L = \{L_1, L_2, \dots, L_{\alpha+2}\}$

Saída: Conjunto de valores representantes  $V = \{V_1, V_2, \dots, V_{\alpha+1}\}$

**for**  $x = 1 \rightarrow \alpha + 1$  **do**

$S = 0$

**for**  $Y = l_\alpha \rightarrow l_{\alpha+1}$  **do**

$S = S + H_y$

**end for**

**for**  $Y = l_\alpha \rightarrow l_{\alpha+1}$  **do**

$H_y = H_y / S$

**end for**

**end for**

**for**  $x = 1 \rightarrow \alpha + 1$  **do**

**for**  $Y = l_\alpha \rightarrow l_{\alpha+1}$  **do**

$V_x = V_x + (Y * H_y)$

**end for**

**end for**

Tabela 1 – Tabela de parametrização padrão dos testes

Firefly	
Número de limiares $d$	1
Número de fireflies $n$	50
Número máximo de iterações $\Gamma$	100
Função de avaliação	Entropia de Tsallis
Coefficiente de absorção $\gamma$	1
Fator de aleatoriedade $\alpha$	0,01
Entropia de Tsallis	
Fator $q$	0,55

## 5 EXPERIMENTOS E RESULTADOS

### 5.1 MÉTODO DE AVALIAÇÃO

Esta seção apresenta o método pelo qual será determinada numericamente a qualidade dos resultados obtidos nos testes. A princípio, todos os testes seguem os padrões apresentados nesta seção, exceto quando apresentadas as alterações no próprio experimento. O método de avaliação também é o mesmo em todos os testes, obtendo-se assim resultados que possam ser comparados entre si.

Os testes serão realizados utilizando o *Matlab* como ferramenta de apoio. Os testes serão efetuados variando os parâmetros de execução e o número de limiares desejados.

Como método de avaliação, serão extraídas as bordas dos resultados da segmentação (procedimento 8 da Figura 5) para comparação. As diferentes segmentações obtidas serão comparadas entre si e com os padrões da base de Berkeley (item 9 da Figura 5) a fim de avaliar a qualidade da segmentação final (procedimento 10 da Figura 5).

#### 5.1.1 Parametrização Padrão

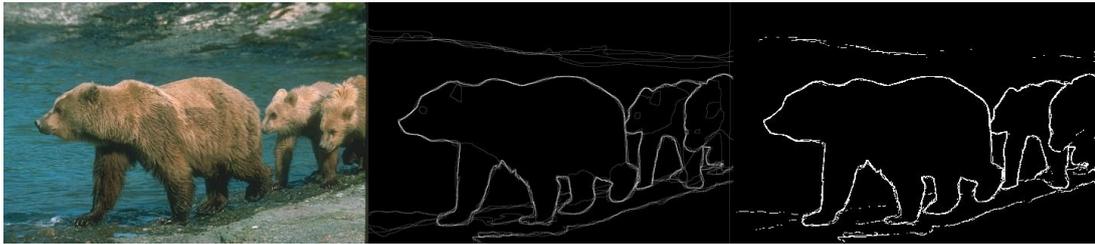
Quando não especificado no experimento de forma diferente, os testes são realizados seguindo os seguintes parâmetros:

Os valores da tabela foram sugeridos na bibliografia em trabalhos como Horng e Liou (2011) e P. S. Rodrigues et al. (2015) ou definidos após os experimentos descritos na Seção 5.2

#### 5.1.2 Preparação da base de dados

Conforme citado na Seção 4.2, a base de dados utilizada neste trabalho é a base de Berkeley (MARTIN et al., 2001), que é composta de 300 imagens naturais e seus respectivos contornos anotados manualmente, que assumiremos como sendo o padrão-ouro para a definição das regiões de segmentação.

Figura 6 – Exemplo de tratamento da imagem



Fonte: Autor.

Da esquerda para a direita: Imagem original, Bordas anotadas da base de Berkeley, Imagem das bordas em binário.

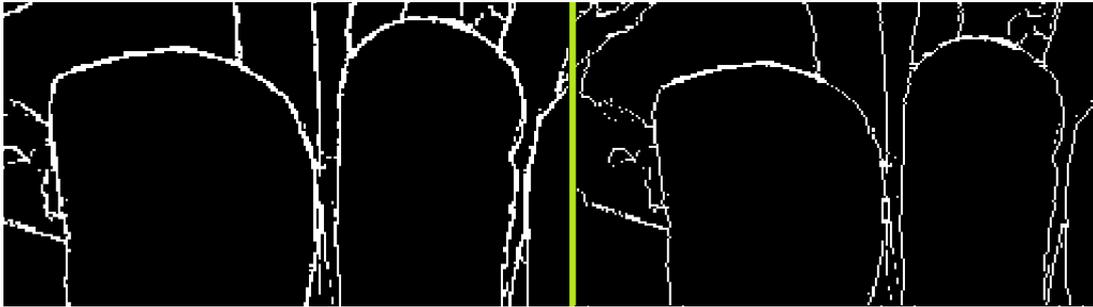
Estes contornos anotados levam em consideração quantas vezes um mesmo contorno foi marcado por diferentes pessoas, representando isto em forma de intensidade. Os contornos que ocorreram com maior frequência são apresentados em maior destaque; ou seja, com mais luminosidade, considerando que estas imagens de contorno estão em tons de cinza. O método de avaliação que será apresentado na Seção 5.1.5 realiza comparação entre duas imagens e mede sua diferença baseado na intensidade e posição dos pixels. Como o intuito do método proposto é a segmentação da imagem, o principal fator a ser considerado é o posicionamento dos pixels que compõem sua borda, e não necessariamente sua intensidade.

Para obter resultados mais precisos e focados na definição de regiões, ambas as imagens de bordas são transformadas em imagens binárias, onde a cor branca representa a presença de borda, e a cor preta a sua ausência. Analisando por amostragem algumas imagens da base de dados e suas respectivas anotações, observa-se que os contornos menos frequentemente marcados possuem intensidade entre 20 e 65. Sendo assim, o procedimento efetuado nas anotações da base de Berkeley foi feito com um limiar de intensidade 70. Este limiar filtra os contornos menos significativos. A Figura 6 exemplifica a transformação do contorno em uma imagem binária para a comparação. Note a supressão das bordas menos significativas como as que representam a região dos olhos e orelhas dos ursos.

Estas imagens binárias comumente apresentam contornos de bordas grossas, traços de 2 ou 3 pixel de espessura. Esta característica impacta consideravelmente no resultado final do método de avaliação utilizado descrito na Seção 5.1.5. Isto ocorre devido ao fato de que tal método baseia-se na quantidade de pixels por valor e o procedimento de extração de bordas descrito em 5.1.4 apenas obtém bordas com espessura de 1 pixel.

Para amenizar este problema, a imagem binária extraída da base de Berkeley é submetida a um procedimento de redução da espessura de suas bordas a 1 pixel. Este procedimento basicamente analisa a imagem em busca dos padrões binários 0110 e 01110, transformando-os em 0100 ou 01000, respectivamente; ou seja, busca pela continuidade de 2 ou 3 pixels brancos delimitados por pixels pretos em suas extremidades nas linhas e colunas da imagem, e trans-

Figura 7 – Exemplo de redução de espessura de bordas



Fonte: Autor.

À esquerda, imagem binária extraída da base de Berkeley. À direita, mesma imagem após o procedimento de redução das bordas.

forma esta continuidade em apenas 1 pixel branco. Desta forma, bordas de 2 ou 3 pixels de espessura são reduzidos a 1.

Os padrões utilizados neste procedimento são apenas de 2 e 3 pixels para evitar alteração na estrutura fundamental da imagem. Continuidades maiores que 3 pixels já são considerados traços verticais ou horizontais do contorno original. Um exemplo do resultado da aplicação deste procedimento pode ser observado na Figura 7

Como exemplo da importância deste procedimento, uma base segmentada das 300 imagens apresenta uma média de diferença de 4,1 se analisada segundo as imagens binárias sem redução de bordas. Esta mesma base quando analisada com os padrões que possuem as bordas reduzidas, apresenta a média de diferença de apenas 0,801. O teste t de Student indica que os dados destas duas amostras são estatisticamente diferentes, com um valor de  $P$  de  $7,2E-274$ .

### 5.1.3 Procedimento de segmentação

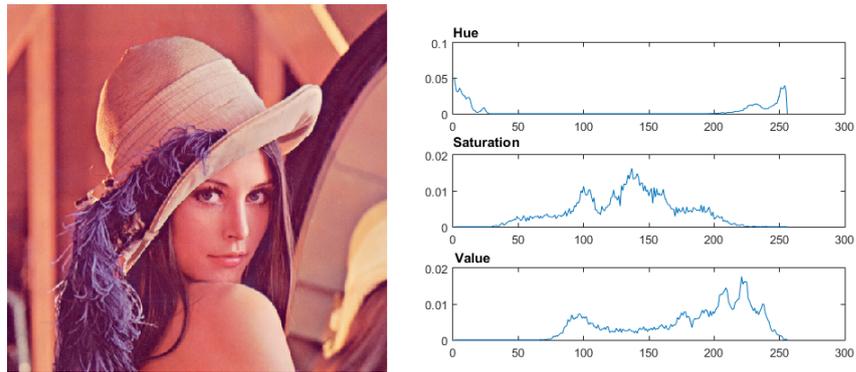
A imagem utilizada originalmente em RGB é convertida para o espaço de cores HSV, resultando em uma matriz tridimensional  $M[A,L,C]$  onde cada coordenada corresponde à altura ( $A$ ), largura ( $L$ ) e canal do espaço de cor ( $C$ ).

São gerados histogramas normalizados separados para cada um dos canais do espaço HSV, como observado na Figura 8. Estes histogramas são submetidos à função de avaliação descrita na Seção 4 que determina quais limiares melhor separam as classes.

Neste exemplo, esta função utiliza o algoritmo *Firefly* com população de 50 vagalumes, máximo número de iterações 100, e Kernel baseado na entropia de Tsallis, conforme descrito na Tabela 1. O resultado da segmentação utilizando um limiar pode ser observado na Figura (9)

A segmentação utilizando um limiar em cada canal pode segmentar a imagem em até 8 diferentes classes, pois, seja  $n$  o número de limiares e  $c$  o número de canais presentes na imagem, o número de possíveis classificações é de  $(n + 1)^c$ . No caso de um único limiar por canal,  $2^3$ , sendo o espaço de cores HSV composto por 3 canais.

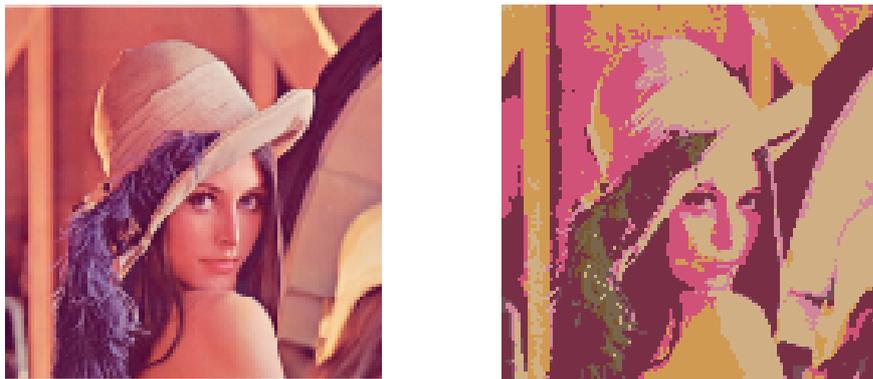
Figura 8 – Canais de imagem HSV



Fonte: Autor.

Imagem HSV (à esquerda) e histogramas de seus respectivos canais (à direita)

Figura 9 – Imagem original e imagem segmentada colorida



Fonte: Autor

O resultado do teste utilizando a segmentação com 2 limiares por canal pode ser observado na Figura 10

Observa-se que utilizando dois limiares o resultado da segmentação aproxima-se mais da imagem original, pois o número de classes definidas é de  $3^3 = 27$ , que permite uma maior aproximação dos valores originais nos 3 canais.

Os limiares definidos pelo método descrito na Seção 3.3.2.1 podem ser observados nos histogramas apresentados na Figura 11, para os testes com um e dois limiares, respectivamente.

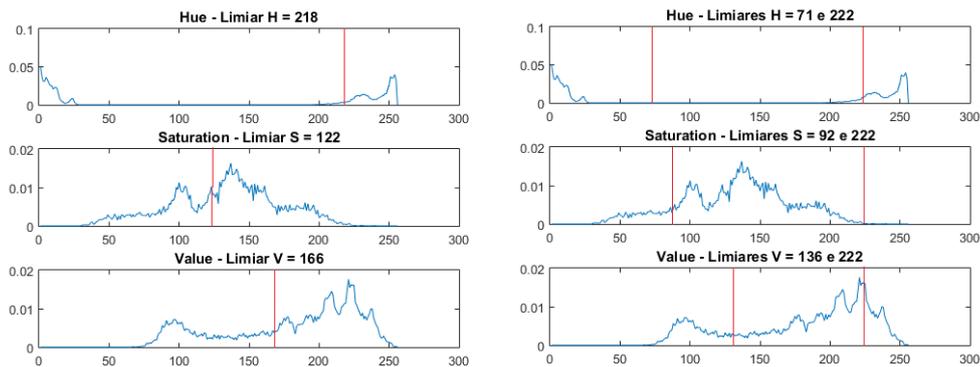
Os limiares representados na Figura 11 tendem a dividir as diferentes classes presentes nos histogramas por serem definidos pela entropia de Tsallis. Esta divisão procura manter em uma mesma classe áreas do histograma que possuem similaridades, como regiões de alta frequência de incidência ou regiões estáveis até que haja uma alteração mais brusca no padrão.

Figura 10 – Imagem original e imagem segmentada com 2 limiares



Fonte: Autor

Figura 11 – Histogramas e limiares definidos para 1 limiar (à esquerda) e 2 limiares (à direita)



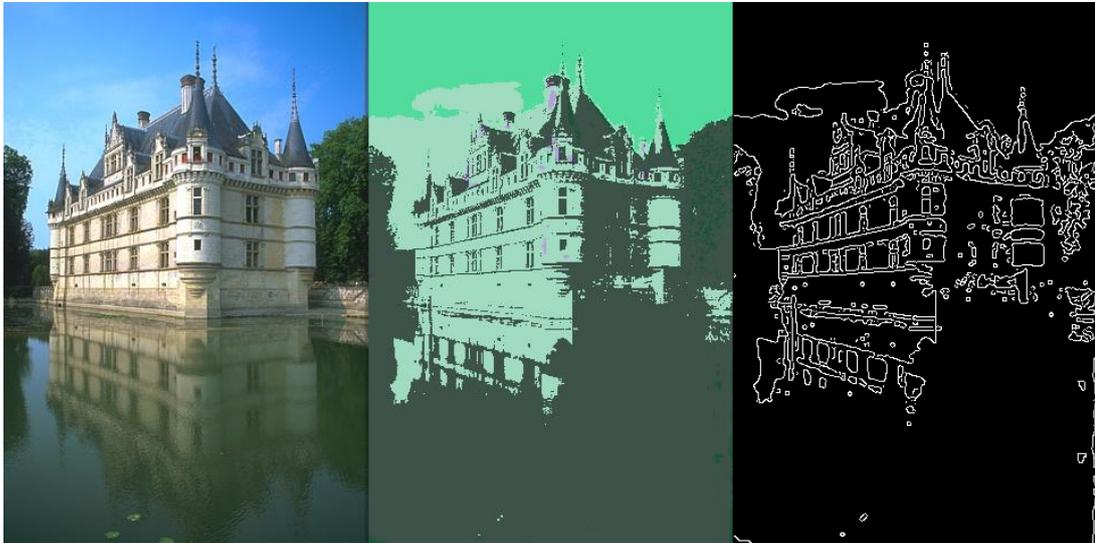
Fonte: Autor.

#### 5.1.4 Extração de bordas da imagem segmentada

A partir das imagens segmentadas, são extraídas as bordas através do método de Canny (CANNY, 1986), originalmente definida para a aplicação em uma dimensão. No caso deste trabalho que trata sobre imagens coloridas, a imagem segmentada é convertida para o espaço de cores YCbCr (PHUNG; BOUZERDOUM; CHAI, 2002) e o método então é aplicado sobre a coordenada Y (luma) que representa a luminosidade da cor. Um exemplo do resultado deste método de extração de bordas pode ser observado na Figura 12

O resultado é uma imagem binária contendo as bordas extraídas da imagem segmentada. Esta imagem posteriormente será utilizada na comparação com as imagens binárias da base de Berkeley descritas da Seção 5.1.2 a partir do método descrito na Seção 5.1.5. A Figura 13 apresenta o procedimento de avaliação de forma geral.

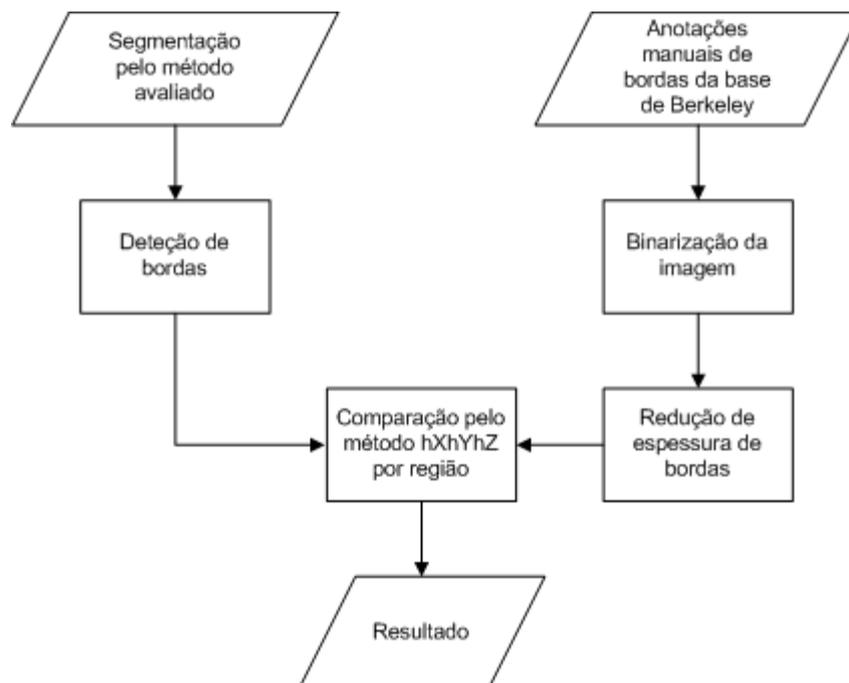
Figura 12 – Extração de bordas pelo metodo Canny



Fonte: Autor

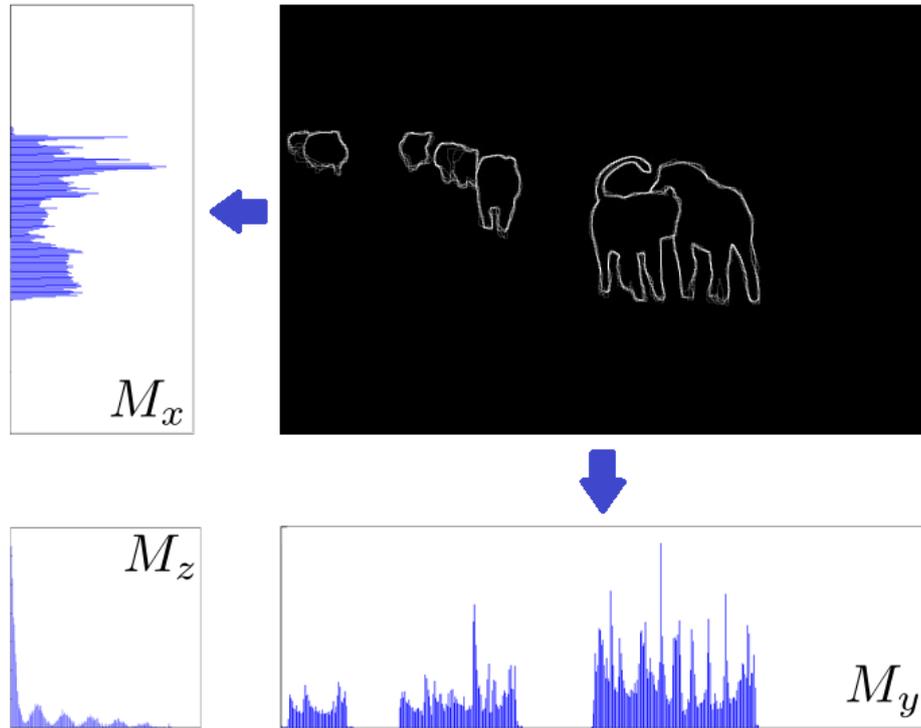
Da esquerda para a direita: Imagem original da base e Berkeley, Imagem segmentada, Bordas extraídas da imagem segmentada.

Figura 13 – Fluxograma do procedimento de avaliação da segmentação de uma imagem



Fonte: Autor.

Figura 14 – Exemplo do resultado das projeções em  $x$ ,  $y$  e histograma



Fonte: (ERDMANN et al., 2015)

### 5.1.5 Método de comparação

Sejam duas imagens binárias  $I_1$  e  $I_2$ , onde  $I_1$  é o mapa de bordas resultante da extração de bordas da segmentação pelo método de Canny, e  $I_2$  a anotação de contorno da base de Berkeley, modificada conforme descrito na Seção 5.1.2, ambas referentes a uma mesma imagem da base de Berkeley. Deseja-se comparar a similaridade entre as duas para avaliação do método de segmentação.

Para esta comparação direta entre  $I_1$  e  $I_2$  é utilizada uma adaptação do método hXhYhZ (ERDMANN et al., 2015). Segundo este método, a diferença de informação entre  $I_1$  e  $I_2$  é calculada pela distância euclidiana entre os pixels das mesmas e por suas respectivas intensidades. Seja  $H_x$  a projeção vertical de  $I_1$  e  $M_x$  o correspondente para  $I_2$ , pode-se calcular a diferença entre as imagens na dimensão  $x$  subtraindo  $H_x$  de  $M_x$ . Analogamente realiza-se o mesmo procedimento para a coordenada  $y$ , obtendo suas projeções  $H_y$  e  $M_y$ . Por fim as distribuições de intensidades  $H_z$  e  $M_z$  (consideradas a coordenada  $Z$ ) são os histogramas de  $I_1$  e  $I_2$ . Um exemplo destes dados extraídos pode ser observado na Figura 14 extraída de Erdmann et al. (2015).

Considerando as imagens  $I_1$  e  $I_2$  possuindo o mesmo tamanho  $A \times L$ , o cálculo da diferença na coordenada  $x$  é dado pela Equação (21).

$$Sim_x(I_1, I_2) = \sum_{i=1}^L (M_x(i) - H_x(i))^2 \quad (21)$$

Onde  $i$  representa cada unidade de largura da imagem onde a projeção de  $x$  está representada. Da mesma forma, calcula-se a diferença entre as projeções no eixo  $y$ , conforme a Equação (22)

$$Sim_y(I_1, I_2) = \sum_{i=1}^A (M_y(i) - H_y(i))^2 \quad (22)$$

Quanto às intensidades das imagens, uma vez que  $I_1$  e  $I_2$  são mapas de bordas binários, esta parte do cálculo apenas reforça a diferença quantitativa de pixels já observada nas coordenadas  $x$  e  $y$ . Seja  $K = [0, \dots, 255]$  os níveis de intensidade possíveis na imagem, esta componente do método é dada pela Equação (23).

$$Sim_z(I_1, I_2) = \sum_{i=1}^K (M_z(i) - H_z(i))^2 \quad (23)$$

Por fim, as 3 componentes são somadas resultando em um número que representa a diferença de informações entre  $I_1$  e  $I_2$  conforme Equação 24.

$$Sim(I_1, I_2) = -\log(Sim_x + Sim_y + Sim_z) \quad (24)$$

Este método de avaliação quando aplicado à comparação de imagens binárias de bordas apresenta uma limitação que o torna inviável para tal. As imagens de bordas geradas são imagens binárias onde poucos pixels possuem valor diferente de 0 (no caso, 255); ou seja, se aplicado diretamente à imagem toda para avaliação, o único quesito levado em consideração será a quantidade de pixels nas coordenadas  $x$  e  $y$ , visto que não são consideradas características espaciais da imagem.

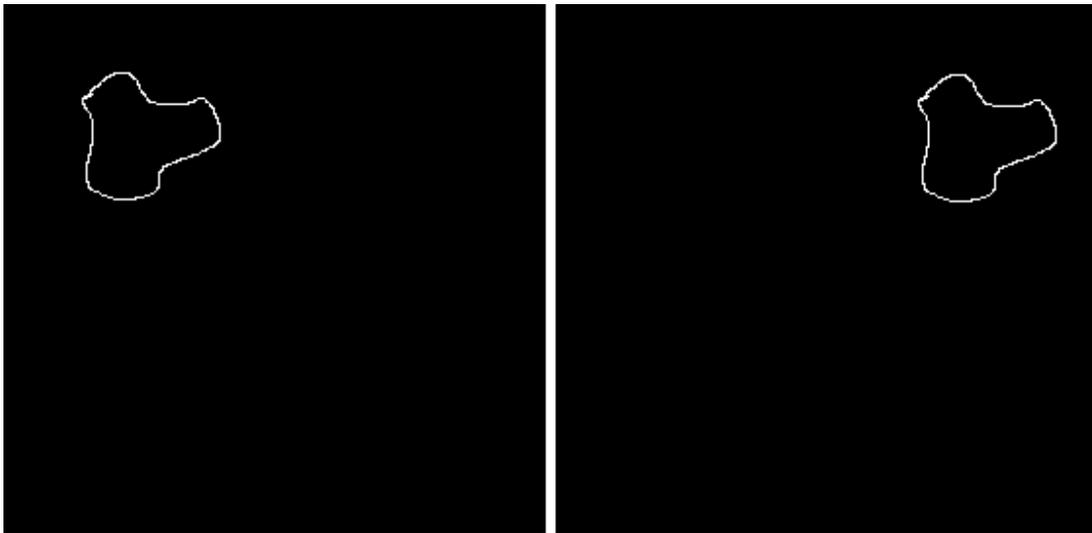
As características espaciais da imagem são importantes neste caso para determinar a qualidade da segmentação. Casos como demonstrado na Figura 15 apresentam total semelhança utilizando o método hXhYhZ apesar de representarem o contorno de áreas diferentes.

Para solucionar esta questão, este trabalho utiliza o método hXhYhZ aplicado iterativamente em pequenas áreas da imagem. O resultado de cada área é então somado e dividido pelo número de áreas avaliadas, obtendo-se uma média como valor de avaliação final. Desta forma soluciona-se a questão de levar-se em consideração as características espaciais da imagem, pois as semelhanças são medidas apenas internamente a cada um destes pequenos espaços definidos na imagem.

Para os experimentos deste trabalho, esta área da imagem foi definida com um tamanho de  $20 \times 20$  pixels, de forma que as áreas sejam grandes o suficiente para tolerar pequenas variações espaciais, e pequenas o suficiente para evitar a limitação acima citada e demonstrada na Figura 15.

Portanto, seja  $T$  o número de áreas avaliadas na comparação entre duas imagens  $I_1$  e  $I_2$ , considerando as variáveis do método hXhYhZ, a diferença entre tais imagens é dada pela

Figura 15 – Imagens diferentes nas quais o método hXhYhZ indica similaridade na coordenada X



Fonte: Autor

Equação (25), onde  $Sim_{x_t}$ ,  $Sim_{y_t}$  e  $Sim_{z_t}$  são referentes à área em questão  $i$ .

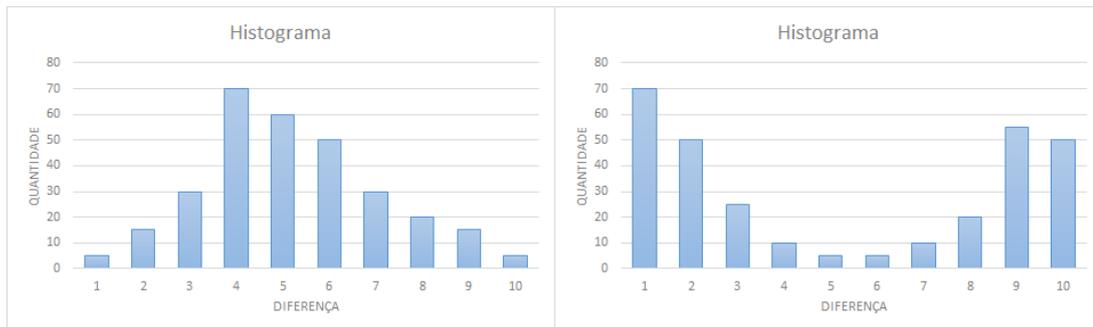
$$Sim(I_1, I_2) = \left( \sum_{t=1}^T -\log(Sim_{x_t} + Sim_{y_t} + Sim_{z_t}) \right) / T \quad (25)$$

Quanto menor for este resultado, mais semelhantes são os mapas de bordas comparadas. No caso deste trabalho onde a comparação é feita entre o mapa de bordas de uma imagem segmentada e as anotações manuais ajustadas da base de Berkeley, quanto menor o valor resultante do método, melhor foi a segmentação, considerando como parâmetro ideal as anotações da base de Berkeley.

Como os testes são efetuados utilizando as 300 imagens da base de Berkeley a fim de avaliar a eficácia da segmentação em diversos cenários, é necessário unificar os resultados de cada imagem individual em um valor que represente a qualidade do método aplicado. Seja  $S = S_1, \dots, S_n$  o conjunto de resultados de similaridades. O valor  $S_t$  que representa a eficácia do método de segmentação aplicado é dado pela média simples do conjunto  $S$ .

Os resultados também poder ser apresentados de forma mais visual em gráficos, onde o eixo  $x$  representa os diferentes níveis de similaridade, e o eixo  $y$  a quantidade acumulada de imagens que atingiram o respectivo nível de similaridade. Esta representação gráfica é útil para identificar a homogeneidade dos valores que resultam na média obtida, pois uma mesma média pode ser derivada de quantidades distribuídas entre níveis intermediários de similaridade (onde o método foi razoável na maioria dos casos), ou de picos nos extremos (onde parte das imagens obteve excelentes resultados, e outra parte resultados ruins). Um exemplo destes casos pode ser observado na Figura 16 onde a média resultante é próxima em ambos, mas representam situações diferentes.

Figura 16 – Exemplos de gráficos representando resultados da segmentação



Fonte: Autor.

Ambos os resultados possuem média aproximada de 5, apesar de sua diferente distribuição.

## 5.2 EXPERIMENTOS

Esta Seção destina-se à apresentação dos experimentos realizados e seus respectivos resultados.

### 5.2.1 Variação do parâmetro $q$ na entropia de Tsallis

A fim de avaliar a influência que o parâmetro  $q$  da entropia de Tsallis têm no resultado final da segmentação e definir um valor como base do experimento, foram realizados testes com o método proposto variando os valores do parâmetro  $q$  da entropia de Tsallis.

Foram utilizados como valores de testes: 0,15; 0,35; 0,55; 0,75; 1,1; 1,5; 2; 2,5 e 3. Os valores apresentados são referentes a aos valores obtidos nas 300 imagens segundo o método de avaliação descrito na Seção 5.1.5. A Tabela 2 apresenta os resultados dos testes realizados. O valor que está sendo considerado para caracterizar a qualidade do resultado é a média obtida pelo método de comparação; quanto menor a média, melhor a segmentação, pois apresentou menor diferença entre os padrões de referência da base de Berkeley. Outros dados apresentados são os valores máximos (piores casos) e mínimos (melhores casos) dentre as 300 imagens analisadas, e a quantidade de vezes em que o experimento em questão superou todos os outros (obteve o menor valor de avaliação dentre os outros para uma imagem).

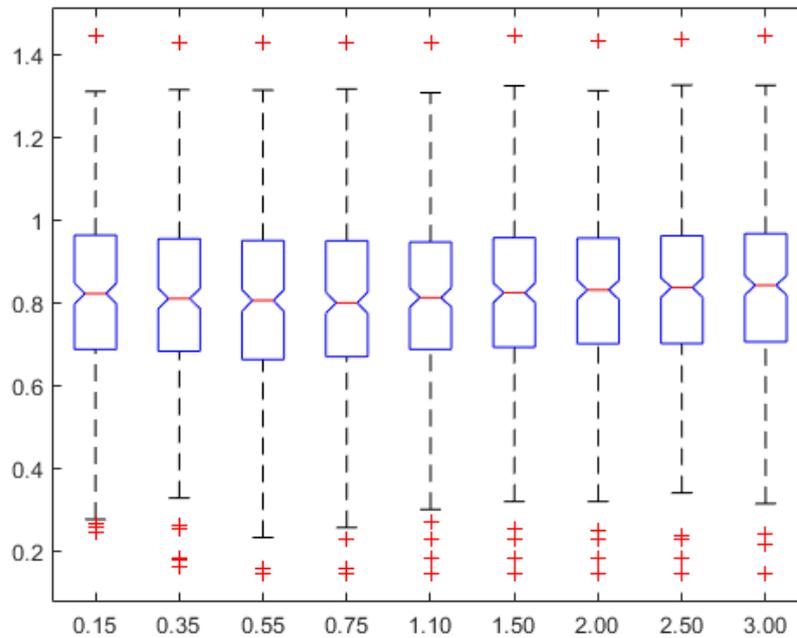
Como observado na Tabela 2, utilizando o valor do parâmetro  $q$  de 0,15 obteve-se o melhor resultado em 55 das 300 imagens avaliadas. Apesar disto, a média geral foi melhor utilizando-se o valor  $q=0,55$ . Porém, ao realizar a análise de variância sobre os dados (ANOVA) o resultado indica que não se pode desconsiderar a hipótese de que os dados não são estatisticamente diferentes. O valor de  $p$  obtido pela análise de variância foi de 0,753, superior aos 0,05 de confiança comumente adotado como parâmetro. Os resultados da análise são apresentados na Figura 17;

Tabela 2 – Resultados dos testes variando o parâmetro  $q$  da entropia de Tsallis

$q$	Média	Máximo	Mínimo	Melhores casos
0,15	0,817	1,444	0,246	55
0,35	0,807	1,428	0,160	40
0,55	0,802	1,430	0,145	47
0,75	0,802	1,430	0,145	29
1,1	0,809	1,428	0,145	23
1,5	0,813	1,444	0,145	30
2	0,817	1,431	0,145	24
2,5	0,824	1,437	0,145	31
3	0,829	1,445	0,145	21

Fonte: Autor

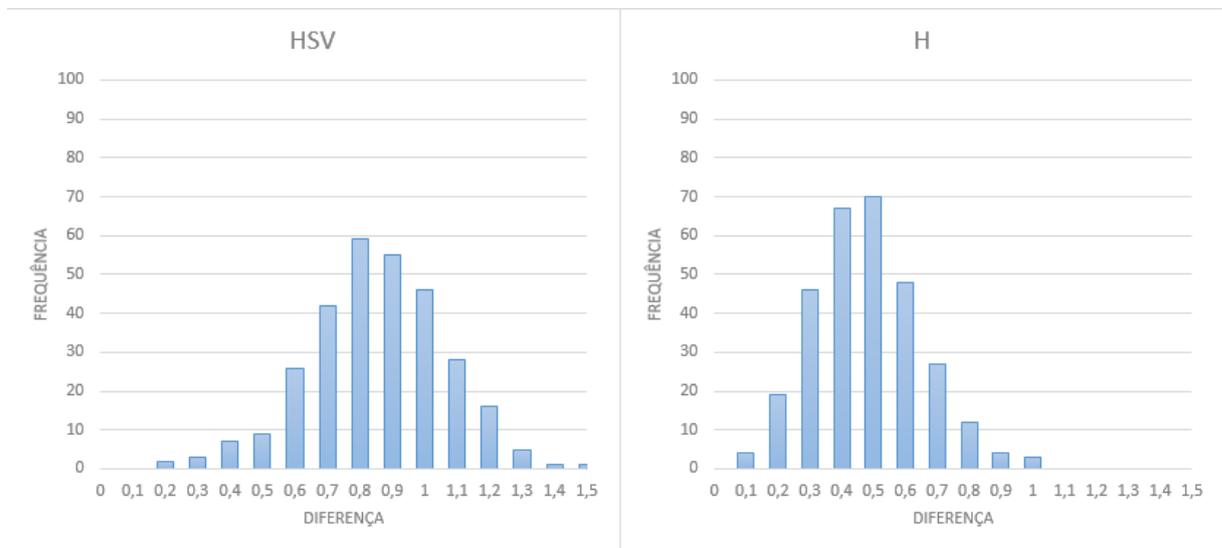
Figura 17 – Teste de variância sobre os resultados envolvendo diversos valores de  $q$  (eixo horizontal)



Fonte: Autor

Na Figura 17, para cada teste realizado, os traços centrais horizontais representam as medianas, os limites dos chanfros definem os limites da media com 95% de confiança. Neste caso como as regiões destes chanfros se sobrepõem se projetadas no eixo Y, não é possível afirmar que os dados são estatisticamente diferentes.

Figura 18 – Histograma do resultado utilizando H,S e V e histograma do resultado utilizando apenas a componente H



Fonte: Autor.

## 5.2.2 Segmentação apenas na dimensão H

Tratando-se de segmentação de imagens coloridas, uma das abordagens é de determinar a segmentação utilizando apenas a matiz da imagem, pressupondo-se que cada objeto possui sua cor (matiz) própria.

Os resultados apresentados nos gráficos na Figura 18

A média dos resultados das segmentações em HSV e H foram, respectivamente, 0,801 e 0,429. A utilização apenas da componente matiz obteve resultados melhores em 297 das 300 imagens. O uso da matiz como critério de segmentação apresenta maior eficácia devido à tolerância de variância de iluminação. Tratando-se de objetos na cena compostos por uma matiz predominante, a saturação e o valor tendem a dividir a classificação da região quando levados em consideração para a segmentação. Um exemplo pode ser observado na Figura 19 onde a segmentação apenas com a matiz obteve resultado de 0,553 contra 0,974 do HSV. O teste t de Student indica que as amostras são estatisticamente diferentes, com um valor P de 4,23E-88.

Este experimento de segmentação em uma dimensão não é possível nos espaços de cores RGB e  $CIE(L^*, a^*, b^*)$  também abordados neste trabalho, pois nestes espaços de cores citados não há como separar as componentes das cores com base em seu significado (como matiz ou luminância). De forma que o isolamento de apenas uma dimensão (R por exemplo, do RGB) não faz sentido para o conjunto de imagens como um todo.

Um fator que favorece a segmentação apenas na dimensão H, é que os limiares utilizados não se multiplicam de acordo com a quantidade de canais, evitando assim casos de *oversegmentation*. Ou seja, se definido um limiar, a imagem segmenta-se em apenas duas regiões distintas utilizando somente a componente H. Por outro lado, se forem utilizadas todas as

Figura 19 – Exemplo de segmentação na coordenada H



Fonte: Autor

Imagem onde a segmentação considerando apenas a matiz foi significativamente superior à utilizando H, S e V. Nota-se que os ursos possuem variação de intensidades em suas cores devido às sombras e à iluminação do sol, mas não de sua matiz. o mesmo ocorre com a água do plano de fundo.

componentes de HSV, o resultado da segmentação teria 8 regiões distintas, sendo duas classes de H multiplicando duas classes de S e multiplicando duas classes de V.

Os resultados deste experimento utilizando apenas a matiz para segmentação com Fire-Fly foram os melhores dentre todos os demais experimentos considerando como parâmetro de qualidade as médias dos resultados.

### 5.2.3 Segmentação em diferentes espaços de cores

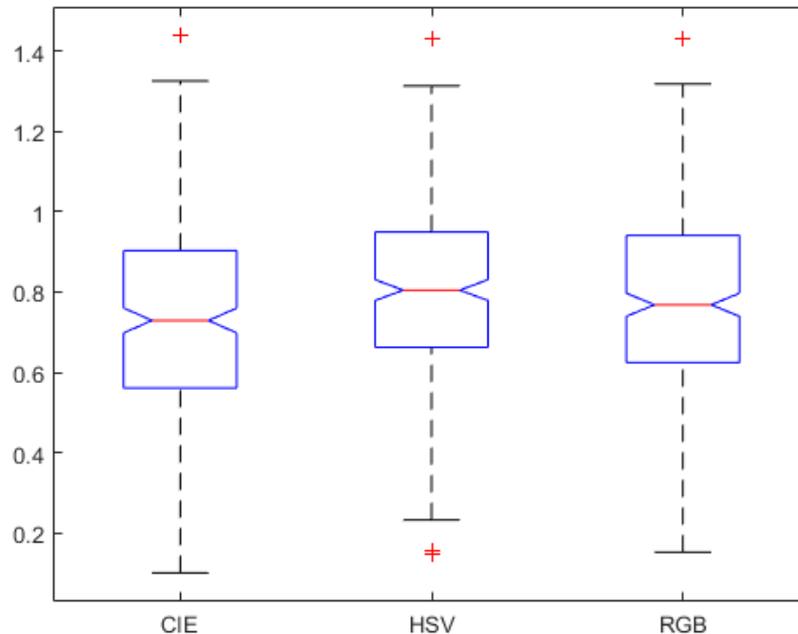
Para fins comparativos sobre qual espaço de cores utilizar com o método proposto, foram realizados testes com três distintos espaços de cores: HSV, RGB e  $CIE(L^*, a^*, b^*)$ , todos descritos na Seção 3.1. A única diferença no procedimento de segmentação apresentado nesta Seção foi o espaço de cores utilizado. Os resultados são exibidos na Figura 20.

Os resultados apresentados apontam que o espaço de cores  $CIE(L^*, a^*, b^*)$  obteve resultados superiores em relação ao HSV e ao RGB. A análise de variância indica que pode-se descartar a hipótese de que os dados não sejam estatisticamente diferentes, com um valor P de 0,0007. Dentre os três espaços de cores avaliados, o HSV obteve o pior resultado.

Se na comparação for incluso o teste realizado na Seção 5.2.2, que é o uso da componente H do espaço de cores HSV, observa-se que o resultado da segmentação baseada na matiz é superior a todos os demais.

Observando a sobreposição dos histogramas nota-se que os valores entre os espaços de cores HSV, RGB e  $CIE(L^*, a^*, b^*)$  foram bastante aproximados, porém o  $CIE(L^*, a^*, b^*)$  foi mais eficaz na determinação da segmentação dentre os três. Com uma média de 0,731 comparado com 0,801 e 0,775 do HSV e RGB, respectivamente. Porém, se comparados com o uso apenas do componente H do espaço de cores HSV, este produz resultados significativamente melhores, com uma média de 0,429 conforme descrito na Seção 5.2.2. Dentre as 300 imagens da base de

Figura 20 – Análise de variância dos resultados da segmentação em diferentes espaços de cores



Fonte: Autor

Berkeley utilizadas, o uso apenas do canal H de HSV obteve o melhor resultado de segmentação em 279 delas, o CIE( $L^*, a^*, b^*$ ) em 16, o RGB em 4 e o HSV em apenas uma imagem.

A análise estatística de variância dos dados (ANOVA) apresenta que há diferença significativa entre os testes realizados. O valor de  $P$  da análise de variância neste caso é de  $8.89E-104$ . a Figura 21 apresenta os resultados da análise de variância, os padrões deste gráfico são os mesmos do descrito na Seção 5.2.1.

Os gráficos na Figura 22 exibem os histogramas dos resultados de cada espaço de cores e o comparativo entre as medias obtidas.

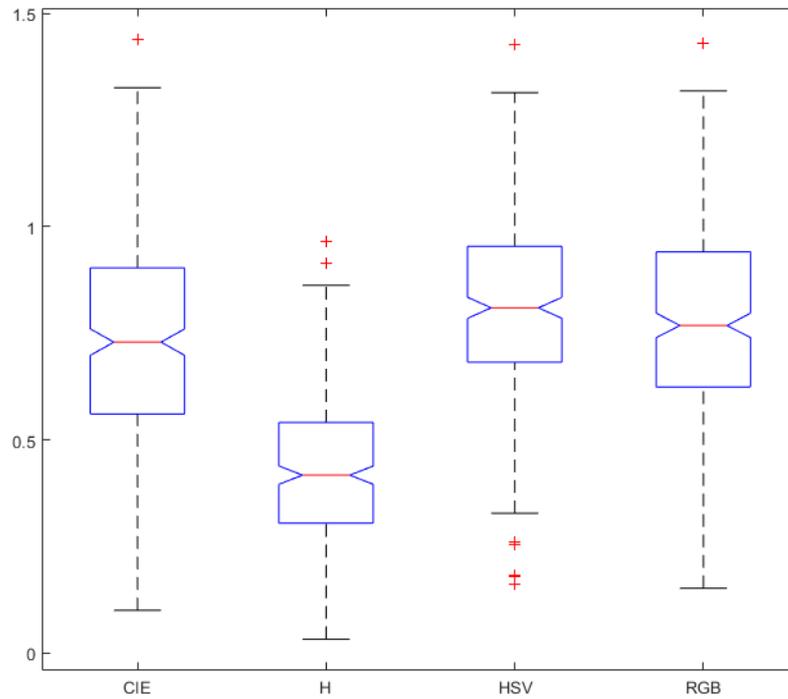
Também foi realizado o teste comparando a segmentação apenas na dimensão H, com a segmentação apenas na dimensão S e a segmentação na dimensão V. Os resultados são exibidos na Figura 23

As segmentações baseadas na matiz se saíram melhor que as demais. segundo a análise de variância os dados são estatisticamente diferentes com um valor de  $1.92E-53$ .

#### 5.2.4 Segmentação com múltiplos limiares

Em segmentação de imagens, a utilização de múltiplos limiares normalmente é ajustada de acordo com a necessidade de seu uso, quanto maior o número de limiares, em mais regiões a imagem é dividida. Tratando-se da utilização de imagens coloridas no método proposto quando

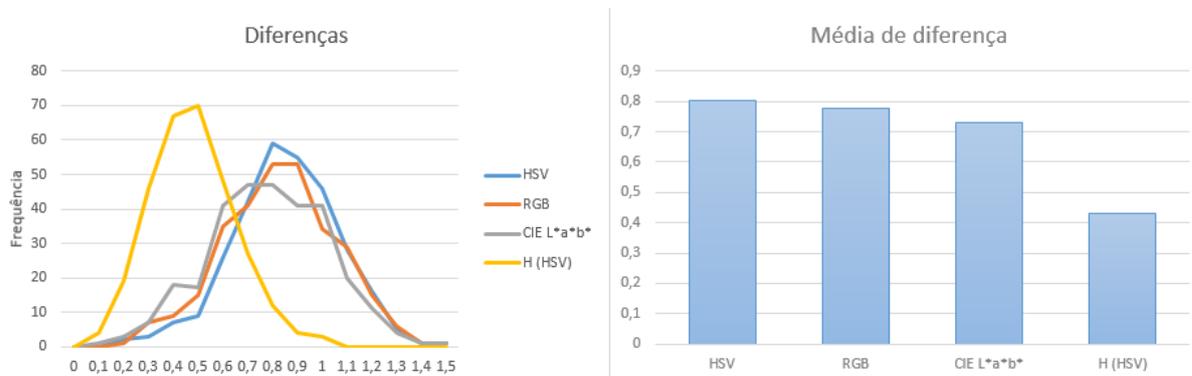
Figura 21 – Resultados entre espaços de cores e coordenada H



fonte: Autor.

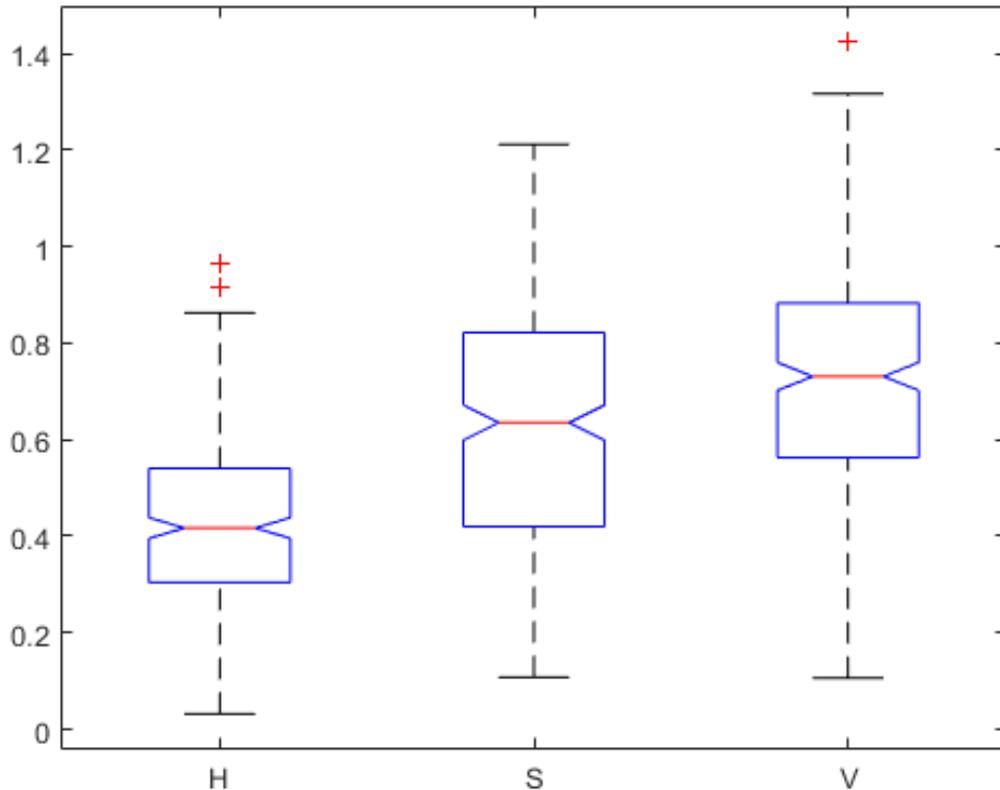
Análise de variância dos resultados da segmentação em diferentes espaços de cores, incluindo a segmentação em H de HSV.

Figura 22 – Sobreposição dos histogramas e média dos valores de diferença



Fonte: Autor.

Figura 23 – Análise de variância das segmentações em H, S e V



Fonte: Autor.

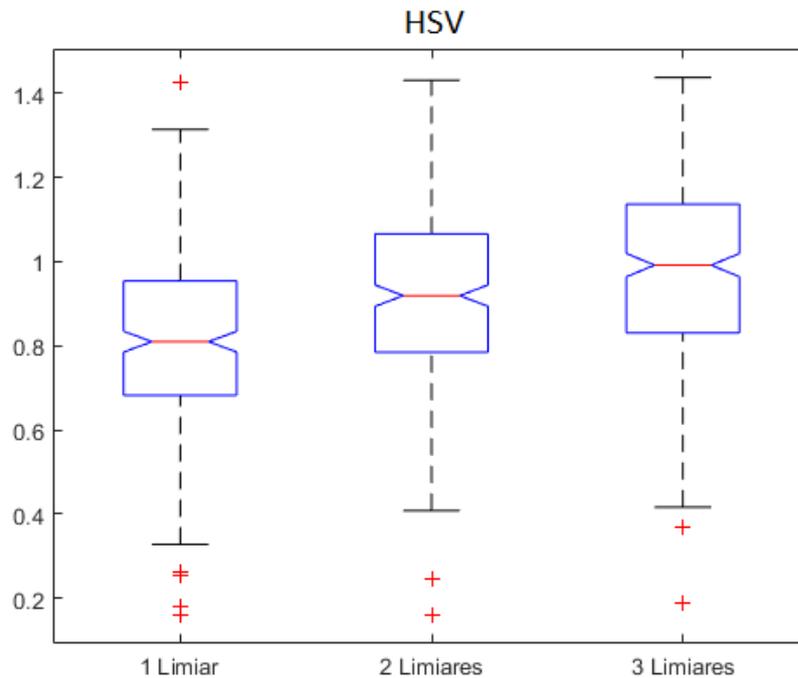
aplicado sobre o espaço de cores HSV (utilizando os 3 canais) a segmentação com 1 limiar classifica a segmentação resultante em até 8 possíveis regiões. Pois segmentando-se em mais de um canal ao mesmo tempo define até  $(l+1)^c$  classes, onde  $l$  é o número de limiares utilizados e  $c$  é a quantidade de canais.

Este conceito é importante ao analisar os resultados dos experimento com múltiplos limiares, pois os testes foram feitos utilizando tanto os 3 canais HSV quanto apenas a dimensão H para comparação. Quando consideradas as 3 coordenadas, o acréscimo de um limiar tem um impacto maior no número de regiões segmentadas se comparado ao acréscimo de um limiar apenas na coordenada H.

Foram realizados testes utilizando 1, 2 e 3 limiares utilizando o espaço de cores HSV completo e apenas a dimensão H. Os resultados obtidos a partir do espaço de cores HSV são apresentados na Figura 24

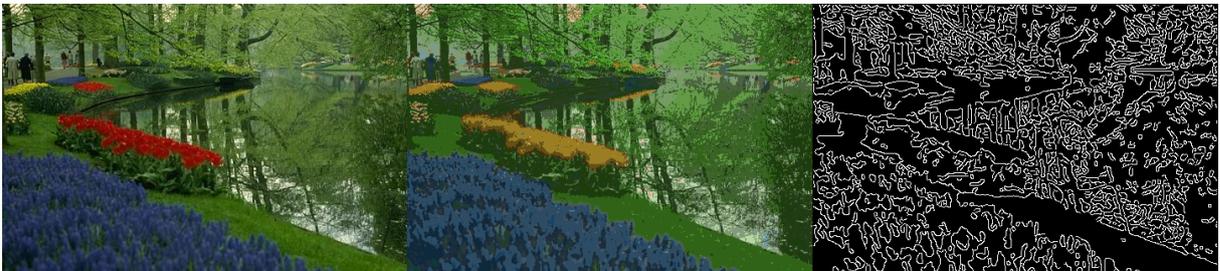
O resultado de P do análise de variância apresentada na Figura 24 foi de 4.84E-20, indicando que os resultados são estatisticamente diferentes. De acordo com os resultados obtidos, a melhor segmentação foi obtida utilizando-se apenas 1 limiar. Os resultados utilizando

Figura 24 – Análise de variância dos resultados da segmentação com diferentes números de limiares no espaço HSV



Fonte: Autor.

Figura 25 – Exemplo de *oversegmentation*



Fonte: Autor.

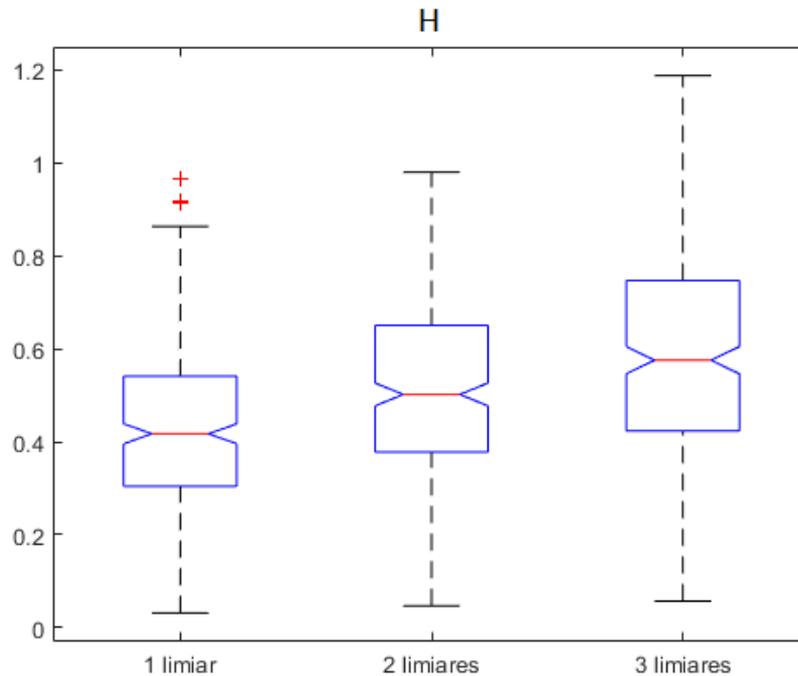
Imagem original, imagem segmentada em HSV com 3 limiares e bordas extraídas da segmentação, respectivamente. O excesso de regiões segmentadas faz com que o resultado da segmentação seja próximo à imagem segmentada, perdendo seu propósito de uso.

maiores quantidades de limiares tendem a causar *oversegmentation* conforme exemplo exibido na Figura 25

Os dados resultantes da segmentação apenas na dimensão H do HSV seguem a mesma tendência, onde os experimentos com apenas 1 limiar obtiveram resultados melhores. Os resultados são demonstrados na Figura 26

No caso da segmentação apenas da dimensão H do espaço de cores HSV, a análise de variância também indica que os resultados são significativamente diferentes, com um valor P de 1.34E-20. Todos os testes realizados na componente H obtiveram resultados superiores aos

Figura 26 – Resultados de múltiplos limiares na dimensão H



Fonte: Autor.

Análise de variância dos resultados da segmentação com diferentes números de limiares na dimensão H de HSV.

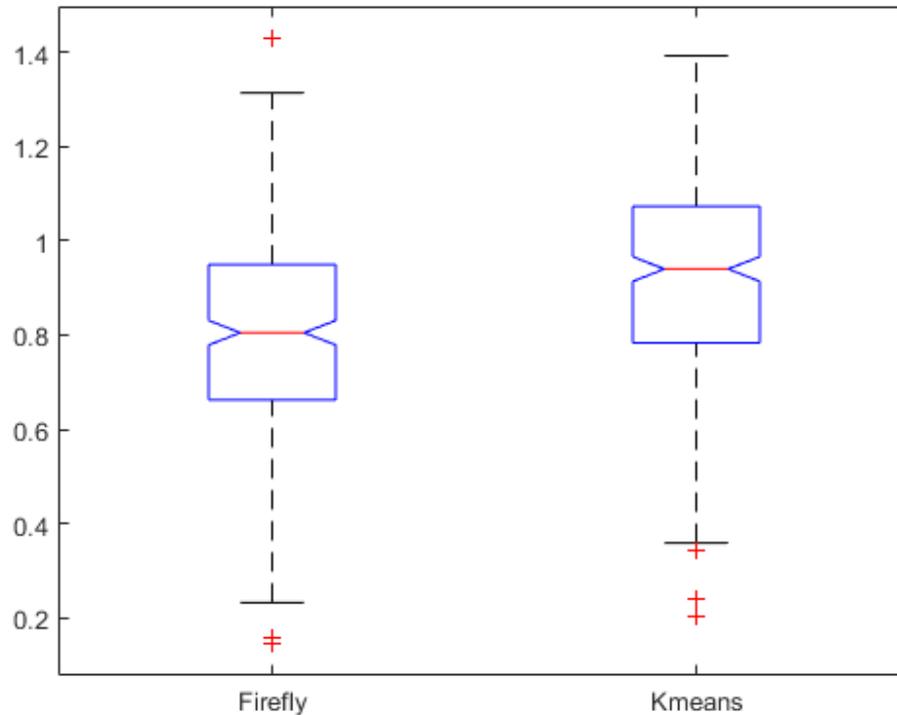
realizados em HSV, mesmo quando comparados os resultados da segmentação em H com 3 limiares com os resultados da segmentação em HSV com 1 limiar. Ou seja, de acordo com os experimentos realizados o aumento de limiares em uma única dimensão é mais eficaz do que aplicar limiares em outras dimensões do espaço HSV (visto que ambos os casos resultam em 3 limiares).

### 5.2.5 Comparação com k-means

O K-means é um método de segmentação baseado em clusterização bastante conhecido e utilizado, portanto é um bom parâmetro comparativo para a eficácia do método de segmentação apresentado neste trabalho. Os testes foram realizados no espaço de cores HSV. O resultado dos testes é apresentado na Figura 27.

O teste T de student para duas amostras pareadas indica que os dados dos resultados são estatisticamente diferentes com um valor de P de  $9.62E-11$ . O firefly obteve resultados melhores em 236 dentre as 300 imagens segmentadas. A média dos resultados do firefly foi de 0.801, enquanto a do k-means foi de 0.918.

Figura 27 – Análise de variância dos resultados da segmentação com com K-means e Firefly



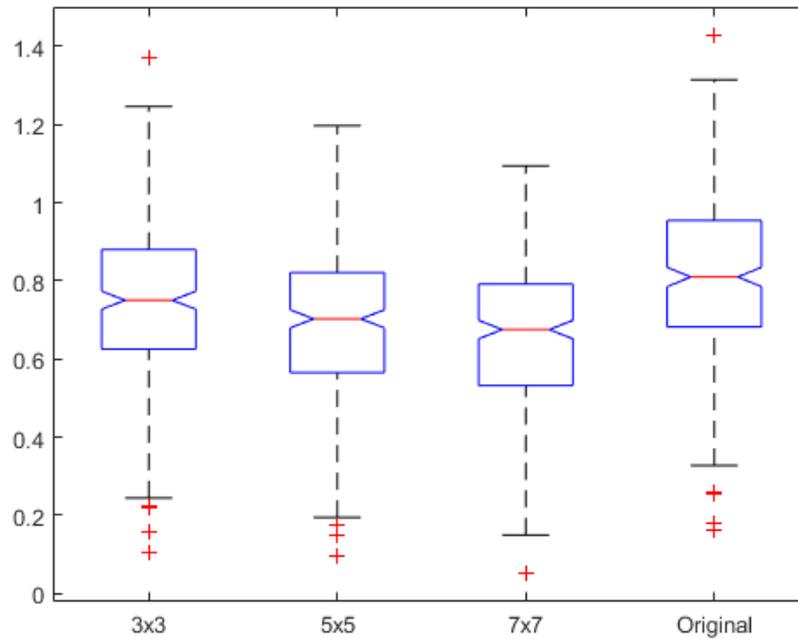
Fonte: Autor.

### 5.2.6 Experimentos com filtros passa-baixa

A utilização de filtros passa-baixa antes da segmentação de uma imagem tem como objetivo evitar a *oversegmentation* de regiões da imagem que possuam alta frequência (como texturas por exemplo). Os testes foram realizados utilizando filtros gaussianos passa baixa nas dimensões  $3 \times 3$ ,  $5 \times 5$  e  $7 \times 7$ , sendo avaliadas segmentações em HSV e também apenas na componente H do espaço de cores. A Figura 28 exibe os resultados referentes ao uso destes filtros no espaço de cores HSV.

A análise de variância deste experimento indica que as médias dos dados são estatisticamente diferentes, com um  $p$  de  $3.26E-20$ . Observa-se que o uso dos filtros melhorou os resultados finais, as médias dos resultados foram 0.806 para as imagens originais, 0.736 para as imagens com filtro  $3 \times 3$ , 0.681 para as imagens com filtro  $5 \times 5$  e 0.653 para as imagens com filtro  $7 \times 7$ . Estes dados apontam que regiões de alta frequência das imagens causam dificuldade na segmentação e a amenização das mesmas pelo filtros passa baixa faz com que as imagens sejam melhores segmentadas segundo os critérios de avaliação deste trabalho (no espaço de cores HSV).

Figura 28 – Resultados com filtro passa-baixa



Fonte: Autor.

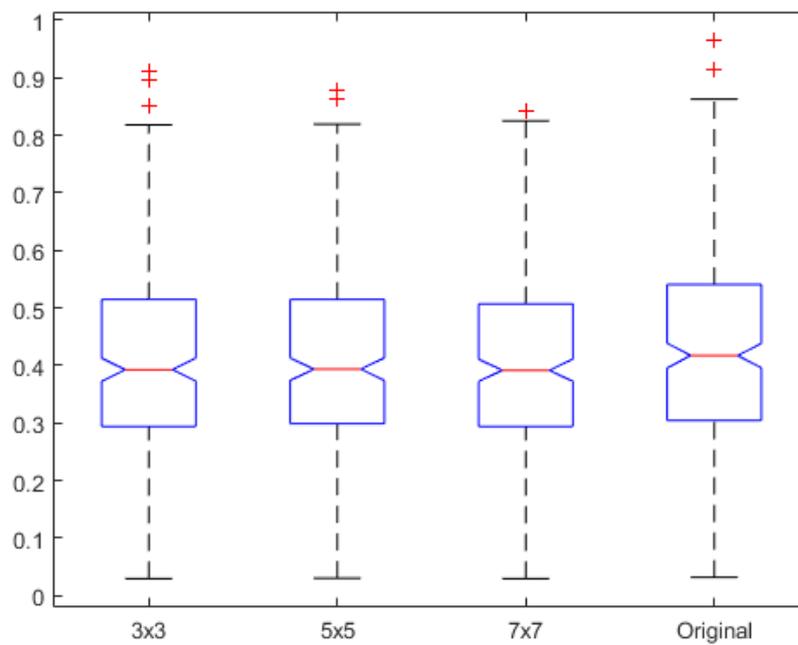
Análise de variância dos resultados da segmentação utilizando filtros passa-baixa.

O mesmo teste foi realizado utilizando apenas a coordenada H do HSV e os resultados são apresentados na Figura 29

A análise de variância neste caso indica que não é possível descartar a hipótese de que as médias dos dados são estatisticamente iguais considerando-se 95% de confiança, com um valor P de 0.25.

Portanto no caso da segmentação utilizando a coordenada H do espaço de cores HSV, as altas frequências das imagens não teve grande influencia no resultado final da segmentação. Fato este que indica que boa parte das altas frequências presentes nas imagens que dificultam a segmentação, encontram-se nos canais S e V.

Figura 29 – Resultados com passa-baixa na coordenada H



Fonte: Autor.

Análise de variância dos resultados da segmentação utilizando filtros passa-baixa (segmentação na coordenada H).

## 6 DISCUSSÃO E CONCLUSÃO

O método proposto neste trabalho mostrou-se eficaz em seu propósito de segmentar imagens coloridas conforme os resultados apresentados na Seção 5.2. Apesar do espaço de cores HSV não obter resultados tão bons quanto o CIE( $L^*$ ,  $a^*$ ,  $b^*$ ) e o RGB, o uso apenas da coordenada H superou ambos e obteve os melhores resultados dentre os testes realizados. Isto ocorre pois a limiarização de apenas um canal tende a sofrer menos *oversegmentation* do que se aplicados limiares a 3 canais (aumentando consideravelmente o número de regiões segmentadas). Outra característica que favorece o uso apenas da dimensão H (matiz) é a tendência das regiões anotadas da base de Berkeley possuírem distinção entre suas matizes, visto que comumente representam objetos distintos.

Ainda com relação ao espaço de cores, os testes utilizando apenas as coordenadas S e V não obtiveram resultados superiores à segmentação baseada na matiz. Esta informação comprova que a qualidade da segmentação utilizando a coordenada H não deve-se apenas pela redução do número de limiares, mas também pela significância do canal na composição das regiões da imagem.

O uso de apenas um limiar para a segmentação também provou-se mais eficaz do que o uso de 2 ou 3 limiares tanto no espaço HSV quanto apenas na coordenada H. Assim como citado no parágrafo anterior os resultados melhores são provenientes de segmentações com um menor número de regiões definidas.

O uso de filtros passa-baixa melhorou os resultados obtidos no espaço de cores HSV, porém não foi significativo quando considerada apenas a coordenada H. Estes resultados indicam que os canais S e V possuem maior influência de altas frequências na segmentação do que o H, pois os filtros passa baixa amenizaram estas altas frequências evitando a *oversegmentation*.

Por fim o método proposto obteve melhores resultados para segmentação em imagens coloridas quando comparado ao K-means, que é um método bastante popular nesta aplicação.

## REFERÊNCIAS

- ADAMS, Rolf; BISCHOF, Leanne. Seeded region growing. **Pattern Analysis and Machine Intelligence, IEEE Transactions on**, IEEE, v. 16, n. 6, p. 641–647, 1994.
- ALBUQUERQUE, M. Image thresholding using Tsallis entropy. **Pattern Recognition Letters** 25, Elsevier, 1059fffd1065, 2004.
- BESAG, Julian. Spatial interaction and the statistical analysis of lattice systems. **Journal of the Royal Statistical Society. Series B (Methodological)**, JSTOR, p. 192–236, 1974.
- BEUCHER, Serge et al. The watershed transformation applied to image segmentation. **SCANNING MICROSCOPY-SUPPLEMENT-**, Scanning Microscopy International, p. 299–299, 1992.
- BEZDEK, James C; EHRLICH, Robert; FULL, William. FCM: The fuzzy c-means clustering algorithm. **Computers & Geosciences**, Elsevier, v. 10, n. 2, p. 191–203, 1984.
- BHANDARI, Ashish Kumar et al. Cuckoo search algorithm and wind driven optimization based study of satellite image segmentation for multilevel thresholding using Kapur entropy. **Expert Systems with Applications**, Elsevier, v. 41, n. 7, p. 3538–3560, 2014.
- BINITHA, S; SATHYA, S Siva. A survey of bio inspired optimization algorithms. **International Journal of Soft Computing and Engineering**, v. 2, n. 2, p. 137–151, 2012.
- CANNY, John. A computational approach to edge detection. **Pattern Analysis and Machine Intelligence, IEEE Transactions on**, IEEE, n. 6, p. 679–698, 1986.
- CHANG, C-I et al. Survey and comparative analysis of entropy and relative entropy thresholding techniques. In: IET, 6. VISION, Image and Signal Processing, IEE Proceedings-. [S.l.: s.n.], 2006. v. 153, p. 837–850.
- CHAPRON, Michel. A new chromatic edge detector used for color image segmentation. In: IEEE. PATTERN Recognition, 1992. Vol. III. Conference C: Image, Speech and Signal Analysis, Proceedings., 11th IAPR International Conference on. [S.l.: s.n.], 1992. p. 311–314.
- CHEN, Tse-Wei; CHEN, Yi-Ling; CHIEN, Shao-Yi. Fast image segmentation based on K-Means clustering with histograms in HSV color space. In: IEEE. MULTIMEDIA Signal Processing, 2008 IEEE 10th Workshop on. [S.l.: s.n.], 2008. p. 322–325.
- CHENG, Heng-Da et al. Color image segmentation: advances and prospects. **Pattern recognition**, Elsevier, v. 34, n. 12, p. 2259–2281, 2001.

- CHUANG, Keh-Shih et al. Fuzzy c-means clustering with spatial information for image segmentation. **computerized medical imaging and graphics**, Elsevier, v. 30, n. 1, p. 9–15, 2006.
- CLERC, Maurice. **Particle swarm optimization**. [S.l.]: John Wiley & Sons, 2010. v. 93.
- DEB, Kalyanmoy et al. A fast and elitist multiobjective genetic algorithm: NSGA-II. **Evolutionary Computation, IEEE Transactions on**, IEEE, v. 6, n. 2, p. 182–197, 2002.
- DENG, Yining; MANJUNATH, B Shin; SHIN, Hyundoo. Color image segmentation. In: IEEE. **COMPUTER Vision and Pattern Recognition**, 1999. IEEE Computer Society Conference on. [S.l.: s.n.], 1999. v. 2.
- DORIGO, Marco. Ant colony optimization. **Scholarpedia**, v. 2, n. 3, p. 1461, 2007.
- DORIGO, Marco; BIRATTARI, Mauro. Ant colony optimization. In: **ENCYCLOPEDIA of Machine Learning**. [S.l.]: Springer, 2010. p. 36–39.
- DORIGO, Marco; MANIEZZO, Vittorio; COLORNI, Alberto. Ant system: optimization by a colony of cooperating agents. **Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on**, IEEE, v. 26, n. 1, p. 29–41, 1996.
- DORIGO, Marco et al. **Ant Colony Optimization and Swarm Intelligence: 6th International Conference, ANTS 2008, Brussels, Belgium, September 22-24, 2008, Proceedings**. [S.l.]: Springer, 2008. v. 5217.
- EBERHART, Russell C; SHI, Yuhui. Particle swarm optimization: developments, applications and resources. In: IEEE. **EVOLUTIONARY Computation**, 2001. Proceedings of the 2001 Congress on. [S.l.: s.n.], 2001. v. 1, p. 81–86.
- ELIA, Ciro D; POGGI, Giovanni; SCARPA, Giuseppe. A tree-structured Markov random field model for Bayesian image segmentation. **Image Processing, IEEE Transactions on**, IEEE, v. 12, n. 10, p. 1259–1273, 2003.
- ERDMANN, H et al. A study of a firefly meta-heuristics for multithreshold image segmentation. In: **DEVELOPMENTS in Medical Image Processing and Computational Vision**. [S.l.]: Springer, 2015. p. 279–295.
- EYSENCK, Michael W; KEANE, Mark T. **Cognitive psychology: A student's handbook**. [S.l.]: Taylor & Francis, 2000.

- FELZENSZWALB, Pedro F; HUTTENLOCHER, Daniel P. Efficient graph-based image segmentation. **International Journal of Computer Vision**, Springer, v. 59, n. 2, p. 167–181, 2004.
- FRIEDMAN, Nir; RUSSELL, Stuart. Image segmentation in video sequences: A probabilistic approach. In: MORGAN KAUFMANN PUBLISHERS INC. PROCEEDINGS of the Thirteenth conference on Uncertainty in artificial intelligence. [S.l.: s.n.], 1997. p. 175–181.
- GONZALEZ, Rafael C; WOODS, Richard E; MASTERS, Barry R. Digital Image Processing. **Journal of Biomedical Optics**, v. 14, n. 2, p. 9901, 2009.
- GRAU, Vicente et al. Improved watershed transform for medical image segmentation using prior information. **Medical Imaging, IEEE Transactions on**, IEEE, v. 23, n. 4, p. 447–458, 2004.
- HARALICK, Robert M; SHAPIRO, Linda G. Image segmentation techniques. In: INTERNATIONAL SOCIETY FOR OPTICS e PHOTONICS. 1985 Technical Symposium East. [S.l.: s.n.], 1985. p. 2–9.
- HARIS, Kostas et al. Hybrid image segmentation using watersheds and fast region merging. **Image Processing, IEEE Transactions on**, IEEE, v. 7, n. 12, p. 1684–1699, 1998.
- HORNG, M.; LIOU, R. Multilevel minimum cross entropy threshold selection based on the firefly algorithm. **Expert Systems with Applications 38**, Science Direct, p. 14805–14811, 2011.
- HU, Xiaohui; EBERHART, Russell. Multiobjective optimization using dynamic neighborhood particle swarm optimization. In: IEEE. WCCI. [S.l.: s.n.], 2002. p. 1677–1681.
- JAIN, Anil K. Data clustering: 50 years beyond K-means. **Pattern recognition letters**, Elsevier, v. 31, n. 8, p. 651–666, 2010.
- JAIN, Ramesh; KASTURI, Rangachar; SCHUNCK, Brian G. **Machine vision**. [S.l.]: McGraw-Hill New York, 1995. v. 5.
- KAPUR, Jagat Narain; SAHOO, Prasanna K; WONG, Andrew KC. A new method for gray-level picture thresholding using the entropy of the histogram. **Computer vision, graphics, and image processing**, Elsevier, v. 29, n. 3, p. 273–285, 1985.
- KENNEDY, James. Particle swarm optimization. In: ENCYCLOPEDIA of Machine Learning. [S.l.]: Springer, 2010. p. 760–766.

LIU, C.; ZHOU, A. Image segmentation framework based on multiple feature spaces. **IET Image Processing**, IET, p. 271–279, 2014.

LUCCHESSEY, L.; MITRAY, SK. Color image segmentation: A state-of-the-art survey. **Proceedings of the Indian National Science Academy (INSA-A)**, v. 67, n. 2, p. 207–221, 2001.

MAHALANOBIS, Prasanta Chandra. On the generalized distance in statistics. **Proceedings of the National Institute of Sciences (Calcutta)**, v. 2, p. 49–55, 1936.

MARTIN, D. et al. A Database of Human Segmented Natural Images and its Application to Evaluating Segmentation Algorithms and Measuring Ecological Statistics. In: PROC. 8th Int'l Conf. Computer Vision. [S.l.: s.n.], jul. 2001. v. 2, p. 416–423.

MATHIAS, Everest; CONCI, Aura. Comparing the influence of color spaces and metrics in content-based image retrieval. In: IEEE. COMPUTER Graphics, Image Processing, and Vision, 1998. Proceedings. SIBGRAP'98. International Symposium on. [S.l.: s.n.], 1998. p. 371–378.

MAXWELL, Bruce A; SHAFER, Steven A. Physics-based segmentation of complex objects using multiple hypotheses of image formation. **Computer Vision and Image Understanding**, Elsevier, v. 65, n. 2, p. 269–295, 1997.

PAL, Nikhil R; PAL, Sankar K. A review on image segmentation techniques. **Pattern recognition**, Elsevier, v. 26, n. 9, p. 1277–1294, 1993.

PEARSON, Karl. Contributions to the mathematical theory of evolution. II. Skew variation in homogeneous material. **Philosophical Transactions of the Royal Society of London. A**, JSTOR, p. 343–414, 1895.

PEREZ, Frank; KOCH, Christof. Toward color image segmentation in analog VLSI: algorithm and hardware. **International Journal of Computer Vision**, Springer, v. 12, n. 1, p. 17–42, 1994.

PHAM, Dzung L; XU, Chenyang; PRINCE, Jerry L. Current methods in medical image segmentation 1. **Annual review of biomedical engineering**, Annual Reviews 4139 El Camino Way, PO Box 10139, Palo Alto, CA 94303-0139, USA, v. 2, n. 1, p. 315–337, 2000.

PHUNG, Son Lam; BOUZERDOUM, Abdesselam; CHAI, Douglas. A novel skin color model in YCbCr color space and its application to human face detection. In: IEEE. IMAGE Processing. 2002. Proceedings. 2002 International Conference on. [S.l.: s.n.], 2002. v. 1, p. i–289.

PUN, Thierry. Entropic thresholding, a new approach. **Computer Graphics and Image Processing**, Elsevier, v. 16, n. 3, p. 210–239, 1981.

RODRIGUES, P; GIRALDI, G. Computing the q-index for Tsallis Nonextensive Image Segmentation. **Computer Graphics and Image Processing (SIBGRAPI), 2009 XXII Brazilian Symposium on**, Sibgraph, p. 232–237, 2009.

RODRIGUES, Paulo S et al. Improving a firefly meta-heuristic for multilevel image segmentation using Tsallis entropy. **Pattern Analysis and Applications**, Springer, p. 1–20, 2015.

ROGER, Easton. **Fundamentals of Digital image processing**. [S.l.: s.n.], 2010. v. 1.

SAHOO, A.; CHANDRA, S. Lffdevy-Flight Firefly Algorithm based Active Contour Model for Medical Image Segmentation. **Contemporary Computing**, IEEE, p. 159–162, 2013.

SEZGIN, Mehmet et al. Survey over image thresholding techniques and quantitative performance evaluation. **Journal of Electronic imaging**, International Society for Optics e Photonics, v. 13, n. 1, p. 146–168, 2004.

SHANNON, Claude Elwood. A mathematical theory of communication. **ACM SIGMOBILE Mobile Computing and Communications Review**, ACM, v. 5, n. 1, p. 3–55, 2001.

SHAPIRO, Linda; STOCKMAN, George C. Computer vision. 2001. **ed: Prentice Hall**, 2001.

SHI, Yuhui; EBERHART, Russell C. Fuzzy adaptive particle swarm optimization. In: IEEE. **EVOLUTIONARY Computation**, 2001. Proceedings of the 2001 Congress on. [S.l.: s.n.], 2001. v. 1, p. 101–106.

SKARBEEK, Wladyslaw et al. Colour image segmentation-a survey. Citeseer, 1994.

SMITH, Alvy Ray. Color gamut transform pairs. In: ACM, 3. **ACM Siggraph Computer Graphics**. [S.l.: s.n.], 1978. v. 12, p. 12–19.

SURAL, Shamik; QIAN, Gang; PRAMANIK, Sakti. Segmentation and histogram generation using the HSV color space for image retrieval. In: IEEE. **IMAGE Processing**. 2002. Proceedings. 2002 International Conference on. [S.l.: s.n.], 2002. v. 2, p. ii–589.

SVAETICHIN, G. **Spectral Response Curves from Single Cones**. [S.l.]: acta physiologica, 1956. (Acta Physiologica Scandinavica. Vol. 39, Supplement 134:2).

TAO, Hai; HUANG, Thomas S. Color image edge detection using cluster analysis. In: IEEE. **IMAGE Processing**, 1997. Proceedings., International Conference on. [S.l.: s.n.], 1997. v. 1, p. 834–836.

TAVARES, Ana Helena Marques de Pinho. Aspectos matemáticos da entropia. Universidade de Aveiro, 2003.

THEILER, James P; GISLER, Galen. Contiguity-enhanced k-means clustering algorithm for unsupervised multispectral image segmentation. In: INTERNATIONAL SOCIETY FOR OPTICS e PHOTONICS. OPTICAL Science, Engineering and Instrumentation'97. [S.l.: s.n.], 1997. p. 108–118.

TREMEAU, Alain; BOREL, Nathalie. A region growing and merging algorithm to color segmentation. **Pattern recognition**, Elsevier, v. 30, n. 7, p. 1191–1203, 1997.

TSALLIS, Constantino. Nonextensive statistics: theoretical, experimental and computational evidences and connections. **Brazilian Journal of Physics**, SciELO Brasil, v. 29, n. 1, p. 1–35, 1999.

TSANG, Peter WM; TSANG, WH. Edge detection on object color. In: IEEE. IMAGE Processing, 1996. Proceedings., International Conference on. [S.l.: s.n.], 1996. v. 3, p. 1049–1052.

YANG, Ming-Hsuan; AHUJA, Narendra. Detecting human faces in color images. In: IEEE. IMAGE Processing, 1998. ICIP 98. Proceedings. 1998 International Conference on. [S.l.: s.n.], 1998. v. 1, p. 127–130.

YANG, X. Firefly Algorithms for Multimodal Optimization. **Stochastic algorithms: foundations and applications**, Springer, p. 169–178, 2009.

YANG, Xin-She. Firefly algorithm, stochastic test functions and design optimisation. **International Journal of Bio-Inspired Computation**, Inderscience Publishers, v. 2, n. 2, p. 78–84, 2010.

\_\_\_\_\_. **Nature-inspired metaheuristic algorithms**. [S.l.]: Luniver press, 2010.

YASEEN, Saad Ghaleb; AL-SLAMAY, Nada MA. Ant colony optimization. **IJCSNS**, v. 8, n. 6, p. 351, 2008.

YIN, P. Multilevel minimum cross entropy threshold selection based on particle swarm optimization. **Applied Mathematics and Computation** **184**, Science Direct, p. 506–513, 2007.

YOUNG, Thomas. The Bakerian lecture: On the theory of light and colours. **Philosophical transactions of the Royal Society of London**, JSTOR, p. 12–48, 1802.

YU, C.; JIN, B. Multi-threshold image segmentation based on Firefly Algorithm. **Ninth International Conference on Intelligent Information Hiding and Multimedia Signal Processing**, IEEE, p. 415–419, 2013.

ZAHN, Charles T. Graph-theoretical methods for detecting and describing gestalt clusters. **Computers, IEEE Transactions on**, IEEE, v. 100, n. 1, p. 68–86, 1971.

ZHANG, Xuemei; WANDELL, Brian A. A spatial extension of CIELAB for digital color-image reproduction. **Journal of the Society for Information Display**, Wiley Online Library, v. 5, n. 1, p. 61–63, 1997.

ZHANG, Yong et al. Image segmentation using PSO and PCM with Mahalanobis distance. **Expert Systems with Applications**, Elsevier, v. 38, n. 7, p. 9036–9040, 2011.

ZHANG, Yudong; WU, Lenan. Optimal multi-level thresholding based on maximum Tsallis entropy via an artificial bee colony approach. **Entropy**, Molecular Diversity Preservation International, v. 13, n. 4, p. 841–859, 2011.

ZHU, Song Chun; YUILLE, Alan. Region competition: Unifying snakes, region growing, and Bayes/MDL for multiband image segmentation. **Pattern Analysis and Machine Intelligence, IEEE Transactions on**, IEEE, v. 18, n. 9, p. 884–900, 1996.