



# UFBA

UNIVERSIDADE FEDERAL DA BAHIA  
ESCOLA POLITÉCNICA  
PROGRAMA DE PÓS GRADUAÇÃO EM  
ENGENHARIA INDUSTRIAL - PEI

DOUTORADO EM ENGENHARIA INDUSTRIAL

GABRIEL JESUS ALVES DE MELO

Métodos para segmentação de células somáticas  
do leite bovino e caprino



**SALVADOR**  
**2015**

Gabriel Jesus Alves de Melo

**Métodos para segmentação de células somáticas do leite bovino e caprino**



Tese apresentada, como requisito parcial para obtenção do título de Doutor em Engenharia Industrial, do Programa de Pós-Graduação em Engenharia Industrial, da Universidade do Federal da Bahia.

Orientadores

Prof. Dr. Antonio Cezar de Castro Lima

Prof. Dr. Luiz Alberto Luz de Almeida

Salvador

2015

---

M528 Melo, Gabriel Jesus Alves de.  
Métodos para segmentação de células somáticas do leite  
bovino e caprino / Gabriel Jesus Alves de Melo. – Salvador, 2015.

105 f.: il. color.

Orientador: Prof. Dr. Antonio Cezar de Castro Lima  
Co-orientador: Prof. Dr. Luiz Alberto Luz de Almeida.

Tese (Doutorado) – Universidade Federal da Bahia. Escola  
Politécnica, 2015.

1. Leite bovino – Células somáticas. 2. Leite de cabra –  
Qualidade. 3. Caprino - Doenças. 4. Bovino de leite – Doenças. 5.  
Mastite caprino – Diagnóstico. 6. Processamento de imagens. I.  
Lima, Antonio Cezar de Castro. II. Almeida, Luiz Alberto Luz de.  
III. Universidade Federal da Bahia. IV. Título.

---

CDD: 664

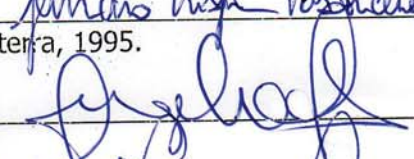
**MÉTODOS PARA SEGMENTAÇÃO DE CÉLULAS SOMÁTICAS DO LEITE  
BOVINO E CAPRINO**

**GABRIEL JESUS ALVES DE MELO**

Tese submetida ao corpo docente do programa de pós-graduação em Engenharia Industrial da Universidade Federal da Bahia como parte dos requisitos necessários para a obtenção do grau de doutor em Engenharia Industrial.

Examinada por:

Prof. Germano Crispim Vasconcelos   
Doutor em Electronic Engineering, Inglaterra, 1995.

Prof. Ângelo Amâncio Duarte   
Doutor em Informática, Espanha, 2007.

Prof. Karina Medici Madureira   
Doutor em Ciências, Brasil, 2012.

Prof. Márcio Fontana   
Doutor em Engenharia Elétrica, Brasil, 2004.

Prof. Luiz Alberto Luz de Almeida \_\_\_\_\_  
Doutor em Engenharia Elétrica, Brasil, 2004.

Prof. Antonio Cezar de Castro Lima   
Doutor em Electronic Engineering, Inglaterra, 1994.

Salvador, BA - BRASIL  
julho/2013

**Conforme resolução do Programa, o conjunto de orientadores teve a  
representação de 1 (um) único voto no parecer final da banca examinadora.**

## AGRADECIMENTO

Ao meu Senhor Deus que tanto fez e tanto faz por mim e por minha família.

A meus pais, Railda e Bianco, pela bravura de toda uma vida. Obrigado pelas orações, pela educação, amor, carinho, sabedoria, otimismo, enfim, obrigado por me fazerem existir. Obrigado, mãe, por todo o empenho para me ensinar a ler, entender, escrever e estudar.

A meu irmão Neto, pela preocupação e incentivo à pesquisa desde o princípio da graduação. Obrigado por me receber e compartilhar comigo momentos inesquecíveis em Itapetinga, durante a primeira fase longe da proteção de casa. Você é meu exemplo.

À minha irmã Milena, por tudo que ela é e representa para mim. Com ela aprendi, cresci e compartilhei toda minha infância. Hoje, muito do que sei agradeço a você, minha irmã. Você é minha segurança. Você é minha segunda mãe.

A todos os meus professores das séries iniciais, do ensino fundamental, do ensino médio, do ensino superior e da pós-graduação.

Ao professor Jânio da UESB, pela compreensão e por me ajudar a voltar ao curso de Matemática em um momento decisivo da minha vida.

À minha noiva Karine, por estar ao meu lado e por participar dessa vitória junto comigo.

A todos os colegas de graduação, pós-graduação e amigos que moraram comigo durante esse percurso. Obrigado por todos os momentos e por compartilharem comigo conhecimento, alegrias e dificuldades.

Aos meus orientadores Antonio Cezar e Luiz Alberto, pela orientação, paciência e compreensão.

À funcionária do PEI, Tatiane, sempre competente, educada e profissional.

Aos membros da Banca de exame da proposta de tese, pelas valiosas orientações.

*A persistência é o caminho do êxito.*

Charles Chaplin

## RESUMO

A mastite é uma doença muito comum e que causa diversos prejuízos às indústrias de laticínio e aos produtores de leite. Essa doença está associada ao aumento de células somáticas no leite e sua detecção tem sido objeto de preocupação e estudos. Este trabalho apresenta dois métodos para segmentação e contagem de células somáticas do leite bovino e caprino em imagens microscópicas capturadas através do microscópio óptico. Os métodos propostos são fundamentalmente baseados no espaço de cor CIELAB, no algoritmo de clusterização K-means, na transformada de Watershed e na proposição de dois novos métodos para limiarização unimodal e segmentação da cor. Os resultados experimentais alcançados para um grande número de imagens indicam que os novos métodos são uma solução promissora para a contagem de células somáticas, através de sistema de vídeo microscopia.

Palavras-chave: Clusterização, Limiarização, Processamento de imagens.

## ABSTRACT

The mastitis is a very common disease that causes several damages to dairy industry and to the milk producers. This disease is related to the increase of somatic cells in milk and its detection has been a subject of concern and study. This paper presents two methods for segmentation and counting of somatic cells from goat and bovine milk, on microscopic images captured by an optical microscope. The proposed methods are fundamentally based on the color space CIELAB, on the clustering algorithm K-means, on the Watershed transform and on two new methods proposed for unimodal thresholding and color segmentation. The experimental results obtained for a large number of images indicates that the new methods proposed are a promising solution for counting somatic cells, through a video microscopy system.

Keywords: Clustering, Thresholding, Image processing.



## LISTA DE FIGURAS

Figura 1 - Imagens com células somáticas do leite bovino. ....	32
Figura 2 - Imagens com células somáticas do leite caprino: (a) e (b) 9 células; (c) 29 células e (d) 9 células. ....	33
Figura 3 - Histograma da imagem: (a) Imagem com flores e (b) Histograma correspondente. ....	36
Figura 4 - (a) Exemplo de <i>stretch</i> linear e (b) Exemplo de <i>stretch</i> logaritmo. Fonte: BRYS (2008). ....	37
Figura 5 - (a) Imagem sem equalização e (b) Imagem equalizada. ....	39
Figura 6 - Limiarização da imagem : (a) representa o histograma da imagem (b); a limiarização da imagem para os valores de $T = 60$ e $T = 100$ são apresentadas nas figuras (c) e (d), respectivamente. ....	40
Figura 7 - Histograma de imagens tons de cinza: (a) Unimodal e (b) Bimodal. ....	42
Figura 8 - (a) Imagem tons de cinza microscópica do pão e (b) Histograma de frequência unimodal correspondente. Fonte: Gonzales-Barron e Butler (2006). ....	43
Figura 9 - (a) Imagem em tons de cinza com microscópica confocal e (b) Histograma de frequência unimodal correspondente. Fonte: Baradez et al. (2004). ....	43
Figura 10- Método de limiarização de Rosin. ....	45
Figura 11- Procedimento para determinação do limiar de OTSU. ....	48
Figura 12- Histograma de intensidades relacionadas ao primeiro plano. ....	48
Figura 13- Histograma com três intensidades. ....	49
Figura 14- Resultado da limiarização de Otsu. ....	49
Figura 15- Transformada de Watershed: (a) Mínimos regionais localizados por setas; (b) Formação do primeiro dique e (c)-(d) As linhas de Watershed. ....	53
Figura 16- (a) Imagem colorida; (b) Imagem P&B e (c) Resultado da TD. ....	54
Figura 17- (a) Imagem P&B $3 \times 3$ ; (b) <i>Pixel</i> central preto representado por 0 e (c) Cálculo das distâncias. ....	55

Figura 18- (a) Duas moedas em sobreposição; (b) Imagem em tons de cinza; (c) imagem após limiarização; (d) TD da imagem P&B; (e) As regiões resultantes após o algoritmo de Watershed e (f) Imagem original com as linhas de Watershed. ....	56
Figura 19- (a) Vizinhança-4 e (b) Vizinhança-8 de um <i>pixel</i> .....	57
Figura 20- Algoritmo de rotulagem de componentes conectados: (a) Imagem P&B; (b) Rótulo único para cada <i>pixel</i> codificado com o valor 1; (c) Avaliação da vizinhança-8 e (d) Resultado da operação de mínimo. ....	58
Figura 21 - Continuação do algoritmo iterativo de rotulagem apresentado na Figura 20. ....	58
Figura 22- Coordenadas do sistema de cor CIELAB.....	59
Figura 23- Esquema para aquisição das imagens.....	62
Figura 24- Fluxograma da metodologia proposta para segmentação e contagem de células somáticas do leite bovino. ....	63
Figura 25- Fluxograma da metodologia proposta para segmentação e contagem de células somáticas do leite caprino. ....	65
Figura 26- Procedimento para a determinação do limiar de intensidade num histograma unimodal. ....	67
Figura 27- Correlação $C_n$ em função do parâmetro de ponderação $L$ . ....	68
Figura 28- (a) Imagem em tons de cinza de bovino. Resultados da limiarização das imagens utilizando: (b) Otsu; (c) Rosin; (d) Limiar proposto e (e) Limiar correspondente a cada método no histograma da imagem. ....	71
Figura 29- Distribuição de frequência dos tons de cinza com dois limiares para extração de detritos. ....	72
Figura 30- (a) Imagem colorida com células somáticas e detritos; (b) Imagem correspondente em tons de cinza e (c) Distribuição de frequência da imagem tons cinza e três limiares para segmentação das células e extração dos detritos. ....	73
Figura 31- Limiarização da imagem tons de cinza utilizando três faixas de limiar para extração de detritos e segmentação das células. ....	74

Figura 32- Imagem sem tratamento (Figura 32a). Conversão de cinza (Figura 32b). Conversão para CIELAB (Figura 32c). Clusterização K-means (Figura 32d e Figura 32e). Imagem clusterizada (Figura 32f). Imagem cluste- rizada no formato tons de cinza (Figura 32g). Imagem limiarizada sem o K-means e com detritos (Figura 32i). Imagem sem detrito com limiar proposto (Figura 32h). . . . .	75
Figura 33- Exemplos de imagens com células em aglomeração: exemplo 1 e 2 cor- respondem às imagens (a-b-c-d) e (e-f-g-h), respectivamente. . . . .	77
Figura 34- Imagem com células somáticas do leite bovino com baixo contraste: (a) Imagem original; (b) Imagem original após a clusterização K-means; (c) Resultado da limiarização proposta e (d) Resultado após a Transformada de Watershed. . . . .	78
Figura 35- Distribuição de frequência $N_p - N_e$ . . . . .	81
Figura 36- (a) e (c) Imagem original RGB; (b) e (d) Imagens correspondentes no espaço CIELAB. . . . .	81
Figura 37- Extração dos canais $L^*$ , $a^*$ e $b^*$ na imagem CIELAB. . . . .	82
Figura 38- Método proposto para segmentação de imagens com células somáticas do leite caprino: (a) Imagem original; (b) Imagem no espaço CIELAB; (c) Imagem CIELAB após operação do algoritmo proposto; (d) Imagem após o filtro de partículas; (e) Transformada da Distância e (f) Resultado da segmentação Watershed. . . . .	83
Figura 39- Ampliação do recorte da aglomeração de células da Figura 38a. . . . .	84
Figura 40- Imagem RGB com 10 células somáticas do leite caprino. . . . .	84
Figura 41- (a) Imagem no formato CIELAB. (b) Imagem após operação proposta. (c) Imagem P&B. (d) Imagem P&B após o filtro de partículas. (e) Trans- formada da distância e (f) Resultado da segmentação Watershed. . . . .	85
Figura 42- Imagem com fundo complexo: (a) Imagem original; (b) Imagem no for- mato CIELAB; (c) Imagem CIELAB após operação proposta para classi- ficação de <i>pixels</i> e (d) Resultado da segmentação Watershed. . . . .	86
Figura 43- Imagem com fundo complexo: (a) Ampliação referente a Figura 42a e (b) Resultado após a Transformada de Watershed. . . . .	86

## LISTA DE TABELAS

Tabela 1 - Limiar de OTSU ( $T = 100$ ) que minimiza a variância entre as classes.....	49
Tabela 2 - Comparação entre os métodos de limiarização no algoritmo proposto. ....	79
Tabela 3 - Comparação entre os métodos de limiarização no algoritmo proposto. ....	79

## LISTA DE ABREVIATURAS

IN51	Instrução Normativa 51
CCS	Contagem de Células Somáticas
MAPA	Ministério da Agricultura, Pecuária e Abastecimento
FAO	<i>Foods and Agriculture Organization</i>
DMSCC	<i>Direct Microscopy Somatic Cell Count</i>
PDI	Processamento Digital de Imagens
SCN	<i>Staphylococcus coagulase-negativa</i>
CPP	Contagem Padrão em Placas
CMT	<i>California Mastitis Test</i>
WMT	<i>Wisconsin Mastitis Test</i>
CE	Condutividade Elétrica
DNA	Ácido Desoxirribonucleico
IA	Inteligência Artificial
PCA	<i>Principal Component Analysis</i>
IDF	<i>International Dairy Federation</i>
VD	<i>Vídeo Microscopy</i>
ROC	<i>Receiver Operating Characteristic</i>
TD	Transformada da Distância
TDE	Transformada de Distância Euclidiana
UFC	Unidades Formadoras de Colônias
CIE	<i>Commission International de L'Eclairage</i>
P&B	Preto e Branco

## LISTA DE SÍMBOLOS

$\sigma_W^2(t)$	Variância intra-classe
$\omega_1(t)$	Probabilidade de ocorrência da classe 1
$\mu_1$	Média da classe 1
$\omega_2(t)$	Probabilidade de ocorrência da classe 2
$\sigma_1^2$	Variância da classe 1
$\mu_2$	Média da classe 2
$\sigma_2^2$	Variância da classe 2
$P(i)$	Função distribuição de probabilidade
$w_k$	Centros do K-means
$C_j$	Classes associadas aos centros
$N_8(p)$	Oitava vizinhança de um pixel
$p_k$	Função histograma
$r_k$	$k$ -ésimo nível de cinza
$\alpha$	Inclinação da reta
$\beta$	Fator de incremento
$\rho$	Fator de ajuste para os níveis
$f(x)$	Função de transformação da imagem
$T$	Valor de limiar
$k_0$	Intensidade com frequência máxima
$H_0$	Frequência com intensidade máxima
$k_m$	Intensidade com frequência nula
$k_n$	Valor de intensidade de maximização da reta
$L$	Parâmetro de ponderação
$L_{\max}$	Parâmetro de ponderação ajustado
$C_n$	Correlação
$k_m$	Frequência nula
$L^*$	Camada de luminância do espaço CIELAB
$a^*$	Gama de cor verde para vermelho
$b^*$	Gama de cor azul para amarelo

## SUMÁRIO

	<b>INTRODUÇÃO</b> .....	17
1	<b>REVISÃO DE LITERATURA</b> .....	20
1.1	<b>Células Somáticas do Leite Bovino</b> .....	20
1.1.1	Fatores que influenciam a quantidade de células somáticas .....	20
1.1.2	Efeitos da contagem de células somáticas sobre a qualidade do leite .....	22
1.1.3	Perdas econômicas decorrentes da mastite .....	24
1.2	<b>Células Somáticas do Leite Caprino</b> .....	25
1.3	<b>Detecção da Mastite na Espécie Bovina e Caprina</b> .....	26
1.3.1	Métodos indiretos para avaliação da celularidade do leite .....	26
1.3.2	Métodos diretos .....	28
1.3.3	Métodos baseados em processamento de imagens .....	30
1.4	<b>Características das Imagens Microscópicas</b> .....	31
1.4.1	Imagens do leite bovino .....	31
1.4.2	Imagens do leite caprino .....	32
1.5	<b>Técnicas de Processamento das Imagens</b> .....	33
1.5.1	Introdução .....	33
1.5.2	Histograma .....	35
1.5.3	<i>Stretching</i> .....	36
1.5.4	Equalização .....	38
1.5.5	Limiarização .....	38
1.5.6	Limiarização Unimodal .....	42
1.5.7	Limiarização de Rosin .....	45
1.5.8	Limiarização de Otsu .....	46
1.5.9	Algoritmo de Agrupamento K-means .....	50
1.5.10	Transformada de Watershed .....	51
1.5.11	Transformada da Distância .....	53
1.5.12	Watershed de Imagens Preto e Branco .....	54
1.5.13	Vizinhança de um <i>pixel</i> .....	56
1.5.14	Algoritmo de contagem .....	57

1.5.15	Espaço de Cor CIELAB .....	58
2	<b>METODOLOGIA PROPOSTA</b> .....	61
2.1	Aquisição das amostras de leite .....	61
2.2	Aquisição das imagens .....	61
2.3	Metodologia proposta para bovinos.....	62
2.4	Metodologia proposta para caprinos.....	64
3	<b>PROPOSIÇÃO DE NOVOS ALGORITMOS</b> .....	66
3.1	Proposição de novo limiar adaptativo .....	66
3.2	Proposição para segmentação da cor .....	68
3.3	Proposição para contagem das células em imagens preto e branco.....	69
4	<b>RESULTADOS E DISCUSSÃO</b> .....	70
4.1	Proposição de novo limiar: imagem sem detritos .....	70
4.2	Proposição de novo limiar: imagem com detritos.....	72
4.3	Segmentação de células do leite bovino.....	74
4.4	Imagem com fundo complexo .....	77
4.5	Resultados para imagens simples do leite bovino .....	79
4.6	Resultados para imagens reais do leite bovino .....	79
4.7	Segmentação de células do leite caprino .....	80
4.7.1	Extração dos canais $L^*$ , $a^*$ e $b^*$ da imagem CIELAB .....	82
4.7.2	Fundo com tom vermelho .....	82
4.7.3	Fundo e células com tons azuis .....	84
4.8	Imagem com fundo complexo .....	85
5	<b>CONCLUSÃO</b> .....	87
5.1	Sugestões para Trabalhos Futuros .....	88
	<b>LISTA DE ARTIGOS PUBLICADOS</b> .....	90
	<b>REFERÊNCIAS</b> .....	91



## INTRODUÇÃO

A agroindústria do leite e seus derivados desempenha, devido à sua enorme importância social e econômica, uma função relevante para a geração de empregos, fornecimento de alimentos e renda para a população.

Segundo informações da EMBRAPA, no ano de 2011, a produção brasileira atingiu 23 bilhões de litros de leite anual, sendo que 78% da produção total concentra-se nos Estados de Minas Gerais, Goiás, Rio Grande do Sul, Paraná, São Paulo, Santa Catarina e Bahia (EMBRAPA, 2011).

Dados do IBGE indicam que o rebanho bovino brasileiro atingiu, em 2009, a marca de aproximadamente 210 milhões de cabeças sendo 22.440,516 de vacas leiteiras com a produtividade de 1.297 L/vaca/ano. O Nordeste contribui com a produção de 3.819,693 L de leite por ano, com a produtividade de 795.000 L/vaca/ano e a Bahia tem sua produção em torno de 1182.019,00 L (IBGE, 2009).

No ano de 2008, a Foods and Agriculture Organization (FAO) indicou que o Brasil foi o 6º produtor mundial de leite, ficando atrás dos Estados Unidos, Índia, China, Rússia e Alemanha.

No Brasil, a Instrução Normativa 51/2002 (IN 51) do Ministério da Agricultura, Pecuária e Abastecimento, que entrou em vigor em julho de 2005, instituiu critérios técnicos e requisitos mínimos de higiene sanitária para a obtenção, coleta, produção, comercialização e qualidade do leite, permeando, assim, os seus níveis de qualidade BRASIL (2002).

Um critério mundialmente utilizado por indústrias, produtores e entidades governamentais para avaliar a qualidade do leite é a Contagem de Células Somáticas (CCS). Tais células são parte do mecanismo natural de defesa do animal e um aumento em seu número é um indicativo do estado imunológico do animal frente a alguma infecção.

Uma infecção muito comum que está associada ao aumento da CCS é a mastite clínica ou subclínica. No caso da mastite clínica, os sinais clássicos da inflamação como dor, inchaço, eritema, calor e perda de função estão presentes. Por outro lado, a mastite subclínica pode ser definida como uma infecção da glândula mamária, sem os sinais clínicos visíveis, podendo passar despercebida, a menos que as ferramentas de diagnóstico adicionais sejam usadas. Dessa forma, o aperfeiçoamento de técnicas para detecção da

mastite tem sido objeto de preocupação e de estudos.

A detecção dessa infecção pode ser feita indiretamente, avaliando-se a viscosidade do leite após a adição de um reagente próprio, ou diretamente, contando-se o número de células por meio do microscópio ou equipamentos eletrônicos. No entanto, embora sejam acessíveis, os métodos indiretos são interpretados subjetivamente e apresentam apenas indicativos de mastite subclínica visível.

Por outro lado, os equipamentos eletrônicos apresentam características relacionadas à grande rapidez e precisão das análises. Entretanto, devido ao custo da aquisição desses equipamentos, o seu uso é restrito aos poucos laboratórios credenciados e autorizados pela Rede Brasileira de Controle da Qualidade do Leite (Ministério da Agricultura, Pecuária e Abastecimento - MAPA).

A IN 51 exige que a contagem de células somáticas do leite cru seja efetuada, no mínimo, uma vez por mês. A fim de cumprir essa exigência, produtores de todos os Estados enviam amostras de leite para os poucos laboratórios credenciados. O controle, no entanto, tem diversas variáveis que resultam na dificuldade do cumprimento da Legislação. No Estado da Bahia, por exemplo, 100% das amostras são analisadas fora do Estado, agregando à produção o custo de transporte e o acondicionamento do material.

Por outro lado, é possível detectar a mastite e a sua gravidade em amostras de leite através da CCS, realizada com auxílio do microscópio (Direct Microscopy Somatic Cell Count - DMSCC). Embora outros testes possam ser utilizados para avaliar a CCS, o método padrão referenciado pelo MAPA e pela Federação Internacional do Leite é o microscópico. Nesse método, esfregaços de leite são preparados em lâminas, corados e examinados no microscópio.

Embora o DMSCC seja utilizado até mesmo para a calibração dos equipamentos eletrônicos, ele ainda é considerado uma tarefa tediosa, sujeita a erros de contagem e requer preparação da lâmina e mais uma ou duas horas para sua visualização no microscópio. Por outro lado, como um número considerável de produtores de várias regiões do mundo não pode pagar o equipamento de contagem eletrônica, o DMSCC ainda se faz necessário.

Assim, a introdução de técnicas de Processamento Digital de Imagens (PDI) para CCS através do microscópio e sistemas de vídeo microscopia são consideradas promissoras e adequadas. No entanto, apesar dos diversos avanços na área da microscopia digital em muitas áreas do conhecimento, poucos trabalhos na literatura têm feito referências à

utilização de Técnicas de Processamento de Imagens (PDI) para CCS em leite de bovinos. Até o presente momento, nenhum trabalho da literatura referencia técnicas de PDI para CCS em leite de caprinos. Portanto, esses são temas novos de grande relevância teórica e prática.

Dessa maneira, o objetivo geral desta tese é desenvolver uma abordagem para auxílio e diagnóstico da contagem de células somáticas em imagens do leite bovino e caprino, aplicando técnicas de processamento de imagens. Para tanto, este estudo tem quatro objetivos específicos. O primeiro consiste em investigar quais técnicas de processamento de imagens podem ser utilizadas para extrair detritos em imagens do leite bovino e caprino. O segundo objetivo busca determinar um método que permita separar as células somáticas do leite bovino e caprino em imagens com aglomerados de células. O terceiro consiste em determinar quais métodos de limiarização são os adequados para imagens tons de cinza com células somáticas do leite bovino e caprino. Por fim, o quarto objetivo específico busca elaborar um esquema para a contagem de células somáticas em imagens capturadas do microscópio óptico e com variação de cor, tonalidade e contraste.

Uma das principais contribuições deste trabalho é a proposição de dois novos métodos para processamento de imagens: 1. um algoritmo de limiarização para imagens com histogramas unimodal e 2. um método para segmentação da cor em imagens CIELAB. Outra contribuição importante é a utilização do algoritmo K-means, como estratégia para retirar previamente os detritos das imagens. Além disso, vale destacar que a utilização do espaço CIELAB permitiu uniformizar as imagens de bovinos e caprinos por meio de uma representação que permite controlar as variações nas imagens.

## 1 REVISÃO DE LITERATURA

### 1.1 Células Somáticas do Leite Bovino

O número de células dentro do leite pode ser avaliado e essa medida é conhecida como CCS (WEST, 2013). Esse é um indicador geral que está claramente relacionado com a saúde das glândulas mamárias e a qualidade do leite (NORMAN et al., 2011; STERRETT, 2013).

As células somáticas são representadas, basicamente, por leucócitos (glóbulos brancos do sangue) e células epiteliais que são eliminadas no leite durante o curso normal da lactação (GALIERO et al., 2000).

Do total das células somáticas, de 75 a 98% correspondem às células de defesa, que têm a função de englobar e fagocitar os microrganismos invasores da glândula mamária e reparar o tecido secretório mamário (LANGONI, 2006).

As células epiteliais são oriundas da descamação normal do tecido de revestimento e secretor interno da glândula mamária e compreendem de 2 a 25% do total de células somáticas (PHILPOT e NICKERSON, 1991).

A proporção normal de células somáticas no leite de vacas não infectadas tem sido relatada como sendo 80% de macrófagos, 16% de linfócitos, 3% leucócitos polimorfonucleares (Sharma et al., 2011). Além dessas, as células epiteliais podem estar presentes em pequeno número (1 a 2%), embora alguns autores tenham registrado percentuais de 10 e até 20% delas no leite normal (RIBAS, 1994; VOLTOLINI et al., 2008).

Como a finalidade dos glóbulos brancos do sangue é combater infecções, o número dessas células aumenta após uma invasão bacteriana no úbere do animal (NORMAN et al., 2011).

Em princípio, quando as bactérias patogênicas entram no úbere (glândula mamária), o sistema de defesa envia um grande número de leucócitos para o leite, com o objetivo de remover as bactérias patogênicas (BLOWEY, 2010).

#### 1.1.1 Fatores que influenciam a quantidade de células somáticas

Dentre todos os fatores, a mastite é o que mais pode influenciar o aumento do número de células somáticas (HARMON, 1994; WELLNITZ et al., 2009). A mastite é

a doença mais comum vista em bovinos e caprinos de leite em todo o mundo, tornando-se um problema significativo, em termos de saúde da fêmea e da produtividade agrícola (WEST, 2013). De acordo com Medrano-Galarza et al. (2012) a doença pode comprometer o bem-estar dos animais, devido, muitas vezes, ao desconforto e à dor.

A mastite é um processo inflamatório da glândula mamária, devido a uma infecção que pode ser causada por várias espécies de microrganismos, tais como bactérias, fungos, algas, vírus e protozoários (OLIVER, 2012).

Durante algum processo inflamatório da glândula mamária, resultante da introdução e multiplicação de microrganismos patogênicos, uma série complexa de eventos reduz a atividade sintética da glândula, provoca mudanças na composição do leite e eleva a contagem de células somáticas (HARMON, 1994).

Mudanças significativas do leite que podem ser observados na mastite bovina são a presença de coágulos, descoloração e um elevado número de leucócitos. Além disso, os sinais clínicos aparentes na mastite bovina compreendem o inchaço, calor e dor no úbere.

Esta doença pode estar presente na forma clínica e subclínica e, ambas as formas, estão associadas com o aumento da CCS (PANDEY et al., 2005).

Na forma subclínica não se observam sinais clínicos visíveis da inflamação na glândula mamária. Por outro lado, na forma clínica os sinais são evidentes, tais como: edema, dor e endurecimento na úbere, aumento da temperatura, inchaço, grumos, pus ou qualquer alteração das características do leite (MANSOR, 2012).

O tipo de mastite que mais preocupa é a subclínica, pois é difícil de ser identificada e não apresenta sintomas visíveis, tornando-se responsável por 70% das perdas econômicas dela decorrentes.

Mais de 135 tipos de microrganismos foram registrados na literatura científica como sendo uma causa de mastite bovina (BLOWEY, 2010; LAM, 2011; WEST, 2013).

Além da mastite, outros fatores podem influenciar a CCS no leite, dentre eles: estresse, idade e estágio de lactação, fatores ambientais, práticas de manejo, alimentação e intervalo entre as ordenhas (LAM, 2011). O estresse de vários tipos, tais como cio, a doença, a vacinação e administração de drogas (BLOWEY, 2010) e estresse por calor também podem afetar a CCS das vacas.

Assim, por exemplo, o aumento da incidência da mastite clínica no verão, em países temperados, está associada ao ambiente quente e úmido que aumenta a exposição

dos animais aos agentes patogênicos (HILLERTON, 2004).

No que se refere às idades das vacas, as mais velhas, geralmente, apresentam uma elevada CCS, fato que se deve às maiores chances de infecção pela prolongada exposição aos patógenos envolvidos, e também à maior descamação de células epiteliais, ao longo de várias lactações (LANGONI, 2006).

Com relação à estação do ano, Magalhães et al. (2006) observaram uma CCS mais alta no verão (janeiro a março). Nesse período, ocorre um aumento na umidade e maior estresse térmico, que aumentam a susceptibilidade do animal a infecções e o número de patógenos aos quais estaria exposto, favorecendo a incidência de mastite nesses meses (HARMON, 1993).

Todos os requisitos são de extrema importância e podem influenciar a contagem de células somáticas. No entanto, para efeito deste trabalho será abordada a contagem de células somáticas (CCS), pois este é um parâmetro que afeta diretamente a qualidade do leite, causando perdas irreparáveis aos produtores, à indústria e, conseqüentemente, ao produto final.

### 1.1.2 Efeitos da contagem de células somáticas sobre a qualidade do leite

A demanda dos consumidores nacionais e internacionais por produtos lácteos de alta qualidade e a preocupação dos órgãos de saúde com relação à qualidade dos alimentos vêm aumentando nos últimos anos.

Atualmente, todos os países que possuem uma indústria leiteira desenvolvida usam a CCS como medida da qualidade do leite para avaliar a saúde do úbere com relação à qualidade e higiene do leite (HARMON, 1994).

Segundo Riqueira (2011) o leite com elevadas contagens de células somáticas tem a sua qualidade comprometida, ocorrendo alterações na sua composição e no rendimento de seus derivados. Na produção de queijos, o elevado número de células somáticas é indesejável, pois acarreta uma diminuição no rendimento, maiores perdas de proteínas no soro, aumento da taxa de proteólise, o que pode favorecer alterações no perfil e teores de amins biogênicas.

De acordo com Nielsen et al. (2005) o aumento de CCS no leite pode causar alteração no teor de gordura, lactose e proteína desse leite. O declínio do teor de gordura durante o processo é devido à capacidade sintética e secretora reduzida da glândula

mamária.

Os ácidos graxos livres no leite com mastite podem aumentar, em consequência de inflamação, causada, provavelmente, por aumento da atividade da enzima lipase (LAM, 2011). Esse processo biológico também altera as características sensoriais dos produtos lácteos fabricados (GOMES et al., 2006) e provoca perdas econômicas para a indústria de laticínios (CAVERO et al., 2007).

Muitos países têm estabelecido limites máximos de CCS do leite de rebanho para assegurar a CCS do leite. Os Estados Unidos e Canadá estabeleceram um limite de 500.000 e 750.000 CS/ml, respectivamente, enquanto na União Europeia, Nova Zelândia e Austrália, o valor é de 400.000 CS/ml (SANTOS, 2002).

No Brasil, visando estabelecer padrões de qualidade para o leite, em 18 de setembro de 2002 o Ministério da Agricultura, Pecuária e Abastecimento publicou a Instrução Normativa nº 51 (IN51), que regulamenta a produção e os padrões de identidade e qualidade para o leite tipos A, B e C, do leite pasteurizado e do leite cru refrigerado, bem como os procedimentos para coleta do leite cru e de seu transporte a granel.

Os parâmetros estabelecidos pela IN51 contemplam, para o leite cru refrigerado, a CCS, a contagem padrão em placas (CPP), a pesquisa de resíduos de antibióticos, a análise de composição centesimal e as temperaturas máximas de referência para conservação do leite cru (BRASIL, 2002).

Nos últimos anos, várias pesquisas foram realizadas em diferentes regiões do Brasil com intuito de avaliar a qualidade do leite com relação à CCS. Por exemplo, na região sudeste Souza et al. (2008) avaliaram a qualidade do leite em relação à CCS, de 2005 a 2008, e verificaram uma variação de 78,4 a 80,4% das amostras com contagem inferior a 750.000 CS/ml. Aproximadamente 20% das amostras estavam acima de 750.000 CS/ml, porém, a partir de julho de 2011 o limite seria reduzido para 400.000 CS/ml, e o percentual de amostras acima desse limite aumentará para 50%.

Não existem evidências de que a CCS no leite, por si só, tenha algum efeito sobre a saúde humana; entretanto, a presença elevada de CCS no leite pressupõe risco de o leite cru estar contaminado por patógenos e resíduos de antibióticos, que podem indiretamente representar um risco à saúde humana (SMITH, 2002).

### 1.1.3 Perdas econômicas decorrentes da mastite

A mastite é um tema importante na indústria de laticínios, em parte porque o leite não pode ser vendido a partir de vacas tratadas com antibióticos, o que muitas vezes ocorre pela infecção de mastite. O leite de vacas tratadas normalmente é descartado ou fornecido aos bezerros (DAIRY, 2010).

Segundo Schukken et al. (2008), em todo o mundo a mastite é uma preocupação, até mesmo para rebanhos bem manejados. Bar et al. (2008) afirmaram que a mastite é a doença de maior importância econômica para a indústria de laticínios.

As perdas econômicas decorrentes da mastite subclínica e clínica incluem: diminuição da produção de leite, custos do tratamento, aumento de trabalho, descarte de leite, taxas veterinárias, preço do leite reduzido por causa da alta CCS, aumento do risco de mastite posterior, aumento do risco de abate ou morte (HANSSON et al., 2011), risco aumentado de outras doenças, e os custos de cuidados preventivos.

Segundo Seegers et al. (2003) a diminuição persistente na produção de leite é o principal efeito negativo que contribui para o impacto econômico da mastite.

Como existe uma perda crescente da produção de leite à medida que os níveis de CCS aumentam, a principal razão com a preocupação em manter baixos os níveis de CCS nos rebanhos é econômica.

De acordo com Bhikane e Kawitkar (2000), essa é a doença mais importante em bovinos e pode causar até 70% na redução da produção de leite, 9% de descarte após o tratamento, 7% do custo dos serviços veterinários e de 14% do abate prematuro.

Tradicionalmente, o controle de mastite tem envolvido a utilização de desinfetantes químicos para as tetas, antissépticos e terapia com antibióticos. Os antibióticos vêm sendo utilizados em controle de mastite por cerca de 50 anos. No entanto, foi sugerido que a terapia com antibiótico não estava, na verdade, ajudando a reduzir a incidência da doença. De fato, vários problemas surgem a partir da utilização de antibióticos, como o desenvolvimento de resistência a antibiótico, a questionável eficácia da droga e a presença de resíduos de antibióticos no leite. Assim, abordagens alternativas têm sido testadas para controle da mastite, a fim de aumentar a imunidade da glândula mamária, usando-se micronutrientes, vacinas e citocinas (SHARMA et al., 2003).

Segundo Sterrett (2013) o diagnóstico precoce da mastite clínica pode reduzir as perdas de produção e melhorar as perspectivas de recuperação tornando-se muito impor-



tante para a indústria de laticínios. Ações proativas também podem reduzir a utilização de antibióticos, o que pode diminuir a possibilidade de resíduos de antibióticos no tanque (OLIVER, 2012).

## 1.2 Células Somáticas do Leite Caprino

As cabras estão entre os primeiros animais de criação doméstica e seu leite tem sido consumido por seres humanos há milhares de anos. Cerca de 2% do total de leite em todo o mundo é produzido por cabras (SHANGGUAN, 2010; PETERS et al., 2013).

Em uma escala global, a produção total do leite caprino chega a 15.510,411 de toneladas por ano, das quais 80% são produzidas em países em desenvolvimento, onde seu leite tem um papel importante na vida de centenas de milhões de seres humanos (SILANIKOVE et al., 2010; CARVALHO, 2010; SILANIKOVE et al., 2014).

De acordo com a Empresa Brasileira de Pesquisa Agropecuária, a produção mundial de leite caprino é de, aproximadamente, 12.435 mil toneladas, sendo 7.900,000 litros produzidos no Brasil. O Nordeste apresenta a maior produção, contribuindo com 90% da produção nacional. Entretanto, a baixa produtividade dos rebanhos caprinos, causada pela falta de tecnologias, aliada aos produtos de baixa qualidade e a desarticulação da cadeia produtiva, são fatores de entraves da caprinocultura leiteira em todo o nordeste brasileiro (BRASIL, 2005; MELO NETO, 2007).

De acordo com Zhao et al. (2014) o interesse por produtos derivados do leite caprino está aumentando, devido à demanda crescente por alimentos mais saudáveis. Para Park e Haenlein (2008) o leite caprino tem qualidades nutritivas semelhantes às do leite humano, no entanto, é menos alergênico para consumo humano do que o leite de vaca.

Por outro lado, assim como em vacas, estudos apontam que o fator mais importante e que determina o rendimento e qualidade em produtos lácteos de derivados do leite caprino são os efeitos prejudiciais do número elevado de células somáticas associados com a mastite subclínica (LEITNER et al., 2006; SILANIKOVE et al., 2010; MARTI De OLIVES et al., 2011; SILANIKOVE et al., 2014).

Mastites em cabras podem ser causadas por certo número de agentes patogênicos, mas o gênero mais importante é o de bactérias *Staphylococcus*, normalmente dividido em *Staphylococcus aureus* e *Staphylococci coagulase-negativa* (CNS).

Esses microrganismos induzem a resposta imune no tecido mamário, que resulta na

migração de leucócitos da circulação para a mama e aumento na CCS do leite (GOMES et al., 2014). O *Staphylococcus aureus* pode provocar a mastite subclínica, mas também a mastite clínica aguda ou mastite gangrenosa (WHITE, 1999; CONTRERAS et al., 2003), que é uma forma superaguda de mastite, caracterizada por necrose do tecido do úbere, causada por alfa-toxinas (SMITH, 2009).

Em caprinos, a mastite tem sido muitas vezes vista como um problema insignificante, isso porque, nessa espécie, a mastite clínica é rara e a subclínica é muitas vezes ignorada por falta de familiaridade com a interpretação dos resultados dos testes de diagnóstico disponíveis. Assim, a mastite subclínica passa despercebida, a menos que técnicas de diagnóstico adicionais sejam usadas (KOOP, 2012).

### 1.3 Detecção da Mastite na Espécie Bovina e Caprina

Nesta seção é apresentada uma revisão sobre os métodos indiretos e diretos para detecção da mastite. São revisados os métodos mais simples, baseados na inspeção visual, até os métodos mais sofisticados baseados em sistemas inteligentes para ordenha automática. Uma revisão sobre as técnicas de diagnósticos da mastite baseadas no processamento de imagens para CCS é apresentada na seção 1.3.3.

#### 1.3.1 Métodos indiretos para avaliação da celularidade do leite

Os métodos que não utilizam a CCS para diagnóstico da mastite são denominados de métodos indiretos e os que a utilizam denominam-se métodos diretos.

Os métodos indiretos, como não realizam a contagem das células somáticas, apresentam apenas indicativos da mastite. Exemplos dos métodos indiretos são o *California Mastitis Test* (CMT) e o *Wisconsin Mastitis Test* (WMT).

Nesses métodos, o leite é misturado a um detergente que provoca uma coagulação tornando a mistura viscosa. O WMT verifica o grau de viscosidade através da velocidade da passagem da mistura por um orifício, ao passo que no CMT o grau de viscosidade é estimado visualmente (Schalm e Noorlander, 1957) .

De acordo com Brito (2015) o WMT não é apropriado para grande número de amostras, por não permitir automação técnica e não diferenciar contagens baixas de células. Por outro lado, embora o CMT seja um dispositivo prático e barato que pode ser aplicado

ao lado do animal, a sua interpretação é subjetiva e pode variar entre os avaliadores.

Além da viscosidade, outros indicadores como a condutividade elétrica (CE), contagem de neutrófilos com base na detecção de  $O_2$ , lactato desidrogenase (LDH) e NAGase foram introduzidos nas últimas décadas (CHAGUNDA et al., 2006; OKADA et al., 2009; VIGUIER et al., 2009; ANKINAKATTE et al., 2013).

O CE do leite foi introduzido em 1940, como um indicador da mastite (DAVIS, 1947; NORBERG, 2005). Desde então, vários estudos têm sido realizados para avaliar a precisão da CE e prever o estado de infecção (HILLERTON, 1999; SHOSHANI et al., 2000; MELE et al., 2001; NORBERG et al., 2004; NORBERG, 2005; JANZEKOVIC et al., 2009; FERRERO et al., 2014).

Por outro lado, outros estudos têm sido realizados para desenvolver sistemas de detecção indireta da mastite. Por exemplo, Eriksson et al. (2005) propôs um sistema baseado no gás-sensor e Wu et al. (2005) propôs a detecção da mastite através de um sensor capaz de detectar o ácido desoxirribonucleico (DNA).

Sistemas como esses, em conjunto com algoritmos de inteligência artificial (IA), fornecem dados que permitem a detecção precoce de mastite. Por exemplo, Cavero et al. (2007) e Miekley et al. (2012) implementaram um sistema de detecção baseado na análise de wavelet, enquanto o Kamphuis et al. (2010) aplicou árvores de decisão para o problema. Cavero et al. (2008) e Ankinakatte et al. (2013) também desenvolveram sistemas baseados em redes neurais, enquanto o De Mol e Woldt (2001), Cavero et al. (2006) e Kramer et al. (2009) conceberam um sistema baseado em lógica *fuzzy*. Finalmente, Miekley et al. (2013) propôs outro baseado em PCA (Análise de Componentes Principais).

No entanto, além de todas as tentativas de utilização de IA e outros algoritmos, tais sistemas baseados em sensores são vulneráveis ao ruído, o que tem uma influência considerável sobre as características do sinal a ser processado e a técnica de detecção para ser aplicado (KAMPHUIS et al., 2010). Portanto, ainda há uma forte necessidade de melhores métodos de detecção que traduzam os dados do sensor em uma informação confiável para ser processada.

Por outro lado, devido a fatores biológicos, para caprinos estes métodos necessitam de padronização (PEIXOTO et al., 2010). Por exemplo, para Andrade et al. (2001) o método indireto WMT aplicado para fêmeas caprinas pode subestimar a CCS, enquanto que para Madureira et al. (2010) o método CMT, quando positivo, pode não ser indicativo

de processos infecciosos da glândula mamária.

### 1.3.2 Métodos diretos

Ao contrário dos métodos indiretos, que fornecem apenas uma indicação de problemas potenciais associados com alta CCS (FERRERO et al., 2014), métodos diretos são mais precisos e recomendados pela *Food and Drug Administration* (FDA).

Nas abordagens diretas, os níveis de mastite são determinados pela CCS por mililitro de leite (JANZEKOVIC et al., 2009).

A CCS provou ser uma das técnicas mais úteis de diagnóstico para detectar a presença e ocorrência de mastite bovina, especialmente a mastite subclínica (SCHUKKEN et al., 2003; NORMAN et al., 2011).

Uma das abordagens mais utilizadas para essa verificação é realizada através do equipamento eletrônico “*Coulter counter*” que pode ser calibrado para reconhecer e contar determinados tipos de células. O equipamento apresenta limitações por não ser informatizado, sendo pouco eficiente para um número grande de amostras.

Por outro lado, nos últimos anos ganhou grande aceitação o sistema que usa o princípio da fluorescência ótica (equipamento *Fossomatic*, fabricado na Dinamarca). Esse equipamento conta células que têm intensidade mínima de fluorescência, devido à coloração do DNA nucleico e é muito utilizado em países desenvolvidos.

Outros sistemas como *Somascope* (fabricado na Holanda), *Somacount* (Estados Unidos) e ANADIS (França) que usam como princípio de contagem a citometria de fluxo são também encontrados no mercado. Tais equipamentos são muito utilizados para esse fim, devido ao seu alto rendimento e precisão, tendo sido o seu uso padronizado e recomendado por organizações internacionais de lácteos (IDF/FIL) (GONZALO et al., 2004; GARCIA-CORDERO et al., 2010).

No entanto, dependendo do país ou da região, esses equipamentos eletrônicos são caros, difíceis de operar e apresentam valor alto de manutenção, características essas que impedem os pequenos agricultores de adotarem para monitorar a saúde de rotina de seus rebanhos.

Alternativamente, como os pequenos agricultores não podem adquirir os equipamentos eletrônicos, eles enviam amostras de leite para as agências governamentais ou a laboratórios centralizados privados para testá-las por CCS, juntamente com outros testes

(GARCIA-CORDERO et al., 2010).

Em contrapartida, segundo Madureira et al. (2010) e Zeng et al. (1999) as características biológicas da glândula mamária e do leite caprino podem levar a erros de interpretação das técnicas convencionais padronizadas para leite bovino, quando essas são aplicadas para o leite caprino. Isso porque, quando comparado ao leite bovino, o leite de caprino apresenta secreção láctea do tipo apócrina, diferente da secreção merócrina da vaca. Ou seja, na espécie bovina, a glândula mamária é classificada como merócrina, pois somente o produto (leite) sintetizado pela glândula é secretado. Na espécie caprina, a glândula é classificada como apócrina, sendo o produto de secreção eliminado juntamente com pequena parte da célula (POUTREL, 1983; ZENG, 1996; GOMES et al., 2010).

Diversos esforços e estudos têm sido realizados com o propósito de avaliar e adequar técnicas diretas e indiretas de detecção da mastite bovina para o diagnóstico da mastite em cabras (POUTREL, 1983; ZENG et al., 1999; WHITE, 1999; TANGORRA et al., 2010; ROMERO et al., 2012). No entanto, ainda não existem técnicas e equipamentos eletrônicos padronizados para CCS e detecção da mastite em caprinos.

Por exemplo, os estudos são divergentes quanto à utilização do equipamento eletrônico *Somacount*, calibrado para espécie bovina, sendo encontrada tanto correlação positiva com a microscopia direta (ANDRADE et al., 2001), quanto o inverso, com superestimação da CCS (ZENG, 1996). Por outro lado, o DMSCC é o método quantitativo mais confiável para CCS do leite caprino e, assim como em bovinos, é adotado pela FDI como o método de referência para CCS em caprinos.

De acordo com Zeng et al. (1999), Madureira et al. (2010) e Gomes et al. (2010), a contagem manual de células somáticas via microscópica direta apresentará resultados mais confiáveis, se durante o processo de preparação das lâminas forem utilizados corantes específicos de DNA, como o verde de metil e a pironina-Y.

Por outro lado, é possível detectar os níveis de CCS no leite, pela contagem manual, realizada por meio do microscópio (DMSCC). Nesse método, os esfregaços do leite são preparados em lâminas microscópicas, levados ao microscópio e examinados pelo técnico responsável.

### 1.3.3 Métodos baseados em processamento de imagens

O processamento de imagem é amplamente utilizado para o reconhecimento de diferentes células nas áreas da biologia e da medicina, utilizando técnicas ou algoritmos já padronizados. No entanto, aplicado para contagem de células somáticas é um tema novo e pouco discutido na literatura.

Existem diversos *softwares* disponíveis no mercado que utilizam processamento de imagem para a contagem de diferentes tipos de objetos, partículas e células (Abràmoff et al., 2004). No entanto, esses tipos de *softwares* têm funções específicas para fins de processamento muito particulares, e cada aplicação prática requer o desenvolvimento de novos algoritmos ou a personalização dos já existentes. Por exemplo, esses *softwares* geralmente têm funções específicas para a segmentação e contagem de células biológicas em um sentido geral, mas as limitações esperadas levaram Kong et al. (2011), Pan et al. (2012) e Dorini et al. (2013) a desenvolverem novos algoritmos para lidar com as células de sangue humano.

Apesar dos diversos avanços nessa área, poucos trabalhos na literatura reportaram à utilização de técnicas de processamento de imagens para segmentação e contagem de células somáticas do leite de bovino. Baro et al. (2005) apresentaram uma descrição detalhada dos métodos de *hardware* para contagem de células somáticas de bovinos através de sistema automatizado vídeo microscopia para DMSCC (referido como Vídeo Microscopy - VM). Embora este sistema VM seja destinado à CCS no leite de vaca, a obra não revela detalhes sobre os algoritmos que são utilizados para a segmentação das células.

Trabalhos mais recentes de segmentação células somáticas do leite bovino foram propostos por Xue et al. (2008), Na e Heru (2009), Xue et al. (2009), Na e Heru (2010) e Wang e Xue (2010). Esses trabalhos relatados são, no entanto, limitados, pois deixam muitas questões em aberto. Por exemplo, nenhum deles tem a investigação necessária para a segmentação em imagens com presença de detritos e variações de tonalidade e cor.

Como a técnica DMSCC manual está sujeita a diferentes procedimentos de manipulação, as imagens obtidas das lâminas sofrem muitas variações em suas características, tais como saturação, variação de contraste, luminosidade e brilho, bem como a presença de detritos. Tais variações são inerentes ao processo que antecede a captura de imagens e dependem do período de lactação do animal, das condições de higiene durante a aquisição da amostra, colorações e tipos de corantes na preparação da lâmina, iluminação do mi-

croscópio e deslize das regiões, a partir de onde as imagens foram feitas. Além disso, as variações dos padrões morfológicos de células somáticas podem deteriorar o desempenho de técnicas de segmentação tradicionais, uma vez que eles podem ser eficazes para um grupo de imagens, mas não para outro.

Quanto à utilização de métodos de processamento de imagens para CCS em caprinos, até o presente momento, nenhum trabalho na literatura tem referenciado essa aplicação. Portanto, esse é um tema novo e de grande relevância teórica e prática.

Assim, como o DMSCC é o método mais confiável para detecção da mastite em caprinos e, ao mesmo tempo tedioso e sujeito a erros, a utilização de PDI em imagem capturadas do microscópio pode contribuir com rapidez e eficiência para o diagnóstico da mastite em caprinos.

#### 1.4 Características das Imagens Microscópicas

Nesta seção são apresentados os padrões morfológicos das imagens microscópicas que contêm células somáticas do leite bovino.

##### 1.4.1 Imagens do leite bovino

As células somáticas têm citoplasma e núcleo, com tamanhos que variam de 4 a 8  $\mu\text{m}$  (ISO 13366-1:2008 IDF 148-1:2008). Em geral, o núcleo e citoplasma são corados em azul ou azul claro, respectivamente (Figura 1).

A massa nuclear geralmente não tem um formato pré-definido fixo e pode aparecer parcialmente destruído (Figura 1a, célula 3 e Figura 1f, célula 9). Normalmente, o corpo nuclear aparece como uma unidade sólida azul ou numa forma granular (Figura 1e, célula 3 e Figura 1f, célula 1 e 8) e, por vezes, podem ser observados em outras formas peculiares.

O uso de diferentes técnicas de coloração pode levar a cores distintas para núcleo e citoplasma, sendo o primeiro mais escuro do que o outro (Figura 1a-f). As imagens têm variações de contraste e de tonalidade (Figura 1a e um Figura 1b) e também a presença de detritos (Figura 1c e Figura 1f).

Além dos fatores citados, as regiões das lâminas onde foram registradas as fotografias também podem modificar o padrão estatístico do citoplasma nas imagens. No

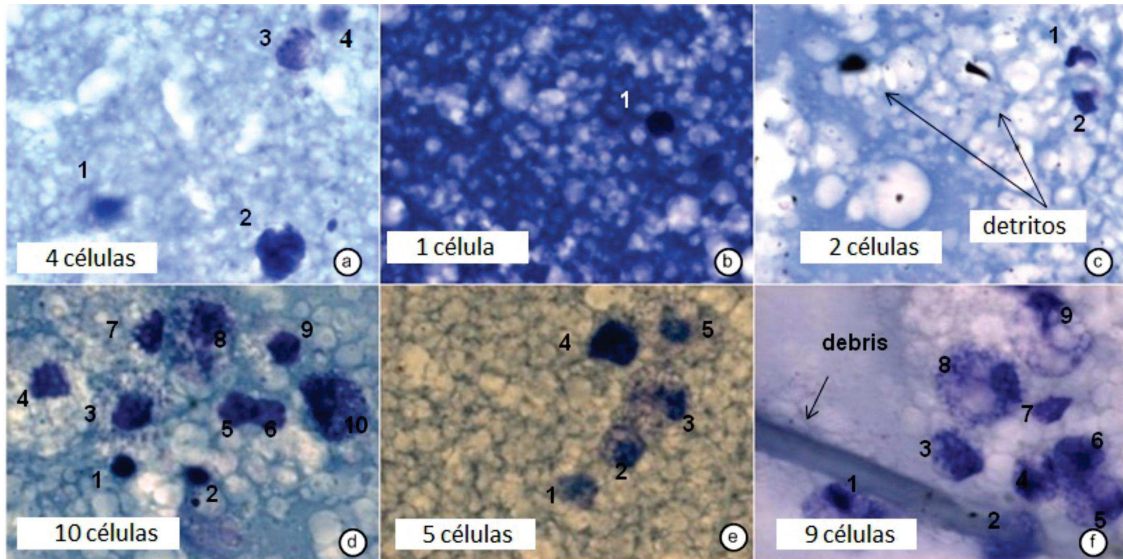


Figura 1 - Imagens com células somáticas do leite bovino.

entanto, o núcleo sempre terá tonalidade mais intensa que o fundo. Essa característica particular é fundamento para as técnicas de PDI adotadas nesse trabalho.

#### 1.4.2 Imagens do leite caprino

Essas células são identificadas nas imagens originais (RGB) como uma unidade sólida, azul ou vermelha com intensidade mais intensa em relação ao fundo (Figura 2a-f). As células podem apresentar alguma distorção indefinida adicional (Figura 2b e d). O fundo não possui forma definida, sendo predominantemente avermelhado e (ou) azulado com manchas brancas (glóbulos de gordura). Em geral, o padrão indefinido das células e do fundo, juntamente com as aglomerações ou sobreposição das células (Figura 2c), tornam a tarefa de contagem das células do leite caprino por microscopia direta uma tarefa laboriosa.



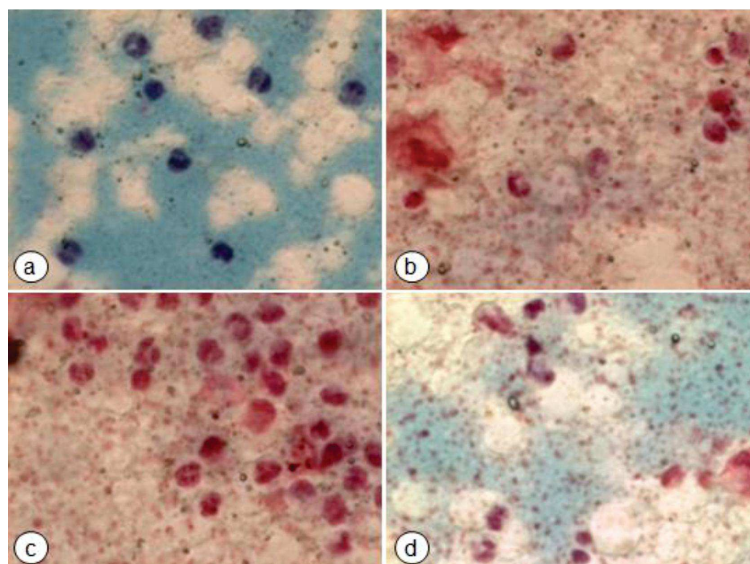


Figura 2 - Imagens com células somáticas do leite caprino: (a) e (b) 9 células; (c) 29 células e (d) 9 células.

## 1.5 Técnicas de Processamento das Imagens

Nesta seção são apresentados os conceitos, as aplicações e as descrições das principais técnicas de processamento de imagens utilizadas nesta tese. A fundamentação teórica sobre os métodos de limiarização é apresentada na seção 1.5.5. Uma revisão sobre a limiarização unimodal e, em particular, a limiarização de Rosin são apresentadas nas seções 1.5.6 e 1.5.7, respectivamente.

### 1.5.1 Introdução

As primeiras aplicações com processamento de imagens datam do início do século XX, na busca de maneiras de aprimorar a qualidade de impressão de imagens digitalizadas transmitidas através do sistema Bartlane (GONZALEZ e WOODS, 2000).

Desde então, essa área vem apresentando crescimento expressivo e suas aplicações permeiam quase todos os ramos da atividade humana, como em medicina, biologia, geografia, geoprocessamento, meteorologia, automação industrial, entre outras.

Dessa forma, o número de áreas usando imagens digitais como ferramenta para avaliação e análise é cada vez maior. Constantemente, a melhoria da tecnologia para sistemas de geração de imagem e computadores mais baratos e melhores, constituem as principais razões subjacentes. Esse crescimento, por sua vez, dá origem à procura e ao

desejo para a análise automática e extração de informações captadas das imagens.

A ideia de usar computadores para a realização de estudos quantitativos e objetivos de informações presentes em imagens, já existe há algum tempo. O primeiro livro sobre o assunto foi publicado no final dos anos sessenta por Rosenfeld (1969) e, desde então, a demanda e o interesse para a análise de imagem digital tem sido cada vez maior (SINTORN, 2005).

Na Medicina, o uso de imagens para o diagnóstico médico tornou-se rotineiro e os avanços em processamento de imagens vêm permitindo tanto o desenvolvimento de novos equipamentos quanto a maior facilidade de interpretação de imagens produzidas por equipamentos mais antigos.

As imagens digitais para esfregaços de sangue periférico em hematologia são utilizados na identificação de tipos de células e morfologia celular, auxiliando nos processos de diagnóstico e tratamento clínico (VANVRANKEN, 2012).

Muitos avanços têm sido feitos no campo de imagens digitais e microscopia virtual em uma variedade de especialidades médicas. Muitas subdivisões em patologia começaram também a aproveitar as vantagens da utilização das imagens digitais, incluindo histologia, radiologia, citologia e dermatologia (CHEN et al., 2011; PATTERSON et al., 2011).

Em biologia e biomedicina, a capacidade de processar automaticamente imagens obtidas de microscópios contribui com a contagem e identificação de células, resíduos, vírus, bactérias, fungos e parasitas, facilitando a execução de tarefas laboratoriais com alto grau de precisão e repetitividade.

Como exemplo de tais aplicações, Garcia-Armesto et al. (1993) e Corkidi et al. (1998) propuseram diferentes métodos para automatização ou semi-automatização do processo de contagem de bactérias em placas de Petri. O método utiliza placas de Petri como meio de cultura para o desenvolvimento das chamadas unidades formadoras de colônias (UFC). A principal vantagem desse processo é a sua simplicidade em termos de equipamentos e, por conseguinte, um menor custo.

Em outra aplicação, Koya (2001) apresentou um sistema de contagem automática de células sanguíneas, baseado em imagens adquiridas por microscópio, como alternativa à contagem manual, que é muito laboriosa e cujo erro estatístico é muito alto. Uma metodologia para identificação automática de ovos helmintos em amostras de fezes microscópicas foi desenvolvido por Yang et al. (2001), um modelo para contagem de colônias

de bactérias por Osowsky e Gamba (2001), uma técnica para reconhecimento e contagem de reticulócitos por Guimarães (2002) e um procedimento para a detecção de células de bactérias nos alimentos por Schönholzer et al. (2002).

Um sistema para o reconhecimento de esporos de fungos microscópicos usando PDI também é apresentado por Perner et al. (2004), modelos para reconhecer partículas biológicas como grãos de pólen e partículas da urina foram apresentado por Ranzato et al. (2007), Vezey e Skvarla (1990), Langford et al. (1990) e France et al. (1997). Finalmente, uma metodologia para o reconhecimento de formas biológicas para o diagnóstico automático de parasitas foi apresentado por Castañon (2006).

### 1.5.2 Histograma

Em geral, o processamento de imagens envolve as etapas de aquisição, análise e transformação das imagens através do computador, visando melhorar suas características visuais e extrair regiões de interesse para a representação de objetos e interpretação das imagens (ALBA-FLORES, 2005).

Após a aquisição de uma imagem digital, o próximo passo consiste em pré-processar a imagem. A função-chave no pré-processamento é melhorar a imagem de forma a aumentar as chances de sucesso dos processos seguintes (GONZALES e WOODS, 1993).

Na etapa de pré-processamento, podem ser utilizados filtros no domínio espacial ou domínio de frequência com a finalidade de suavizar e eliminar ruídos contidos nas imagens. Os ruídos podem ser adicionados à imagem durante sua aquisição, por dispositivos como scanners, máquinas fotográficas, câmera de vídeo entre outros (FARIA, 2005).

Uma imagem pode ser representada através de uma matriz  $n \times m$  onde cada elemento corresponde à intensidade ou ao nível de cinza  $f(x, y)$  em um determinado ponto da imagem. Esses elementos de imagens digitais são os *pixels* que, na prática, possuem uma intensidade que representa uma cor em um determinado ponto da imagem. Para imagens coloridas, em geral o valor de cada *pixel* é dado pela combinação das três cores primárias: vermelho, verde e azul (R, G, B) onde cada cor tem intensidade que varia de 0 a 255. Para imagens em Preto e Branco (P&B) os *pixels* podem assumir os valores 0 e 1, enquanto que para imagens tons de cinza esses valores podem variar de 0 a 255.

O histograma é uma das formas mais comuns de se representar a distribuição de valores de *pixel* de uma imagem tons de cinza. Ao se observá-lo, tem-se uma noção ins-

tantânea sobre as características da imagem, sendo possível inferir informações relevantes, tais como, intensidade média e distribuição dos valores de *pixel* (Figura 3). Dessa forma, é possível, através do histograma, identificar o quanto uma imagem é escura, clara, ou se ela está nos níveis intermediários. O histograma é uma função discreta que pode ser representado por:

$$p_k(r_k) = \frac{n_k}{n} \quad (1)$$

onde

$r_k = k$  -ésimo nível de cinza;

$n_k$  = número de *pixels* na imagem cujo o nível de cinza corresponde a  $r_k$ ;

$n$  = número total de *pixels* na imagem;

$k = 0, 1, 2, \dots, L-1$ , onde  $L$  é o número de valores de *pixels* da imagem digitalizada;

$p_k$  = resultado da função discreta representada por  $r_k$ .

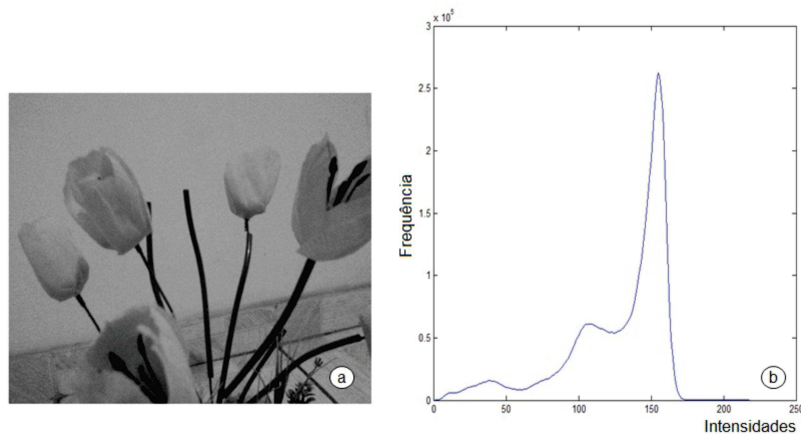


Figura 3 - Histograma da imagem: (a) Imagem com flores e (b) Histograma correspondente.

### 1.5.3 *Stretching*

É frequente a aquisição de imagens com baixo contraste em situações que envolvem iluminação não uniforme, baixa intensidade ou ainda devido a deficiências do sensor de visão. Nesses casos, operações de manipulação do histograma como *stretching* e equalização do histograma podem ser úteis.

A operação de *stretching*, ou espalhamento de contraste, visa uniformizar a distribuição das raias de um histograma de forma que elas preencham toda a faixa do espectro de cinza (BRYS, 2008).

A função de mapeamento pode ser uma função linear, quadrática, raiz quadrada, logaritmo, linear inversa ou ainda outra função definida pelo usuário:

$$Y = \alpha X + \beta$$

$$Y = \phi\sqrt{X}$$

$$Y = \phi X^2$$

$$Y = \phi \log(X + 1)$$

$$Y = -(\alpha X + \beta)$$

onde,

$Y$  = novo valor de contador digital;

$X$  = valor original de contador digital;

$\alpha$  = inclinação da reta;

$\beta$  = fator de incremento, definido pelos limites mínimo e máximo fornecidos pelo usuário;

$\phi$  = fator de ajuste para os níveis de saída ficarem entre 0 e 255;

O aumento de contraste por uma transformação linear é a forma mais simples de *stretch*. A função de transferência é uma reta e apenas dois parâmetros são controlados: a inclinação da reta e o ponto de intersecção com o eixo X (Figura 4). A inclinação controla a quantidade de aumento de contraste e o ponto de intersecção com o eixo X controla a intensidade média da imagem final (BRYS, 2008).

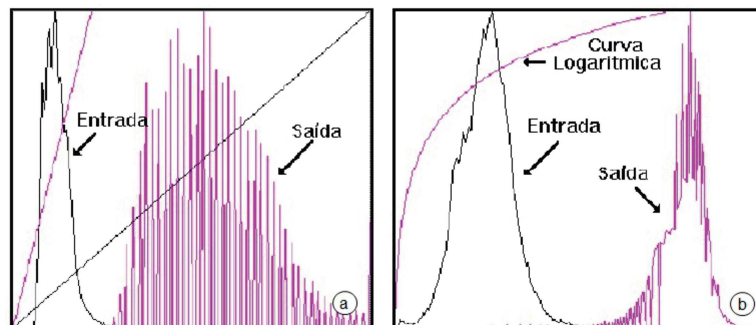


Figura 4 - (a) Exemplo de *stretch* linear e (b) Exemplo de *stretch* logaritmo. Fonte: BRYS (2008).

#### 1.5.4 Equalização

A equalização de histogramas é um método de ajuste de contraste que utiliza o histograma da imagem. Esse método normalmente aumenta o contraste local de muitas imagens, especialmente quando seus dados utilizáveis são representados por valores próximos de contraste. Através desse ajuste, as intensidades podem ser mais bem distribuídas no histograma. Isso permite que as áreas com menor contraste local obtenham um contraste mais elevado sem afetar o contraste global (Figura 5).

De forma sucinta, a equalização adaptativa de histograma consiste em realçar o contraste da imagem, operando em regiões pequenas de  $8 \times 8$  *pixels*.

Esse método consegue, efetivamente, a melhoria do contraste, espalhando os valores de intensidade mais frequentes. A equalização de histograma mostra-se bastante útil em imagens onde os objetos e o fundo se encontram, ambos, na mesma intensidade de luz (claros ou escuros).

A Equação 2 expressa a função para equalização do histograma (Richards e Richards, 1999):

$$y = f(x) = \frac{L-1}{N} \int h_i(x) \quad (2)$$

onde,

$y$  = contador digital na imagem de saída;

$f(x)$  = função de transformação da imagem original para a imagem equalizada;

$L$  = número total de CD's disponíveis;

$N$  = número total de *pixels* na imagem.

#### 1.5.5 Limiarização

A segmentação de imagens desempenha uma regra fundamental em muitas aplicações de visão computacional. Refere-se à técnica de separar os *pixels* da imagem formando diferentes regiões, identificando um fator de semelhança entre as suas características.

Na maioria dos casos, devido a vários fatores, é difícil segmentar o objeto na imagem (ENSAFI, 2011). Em termos gerais, a segmentação é uma operação que divide uma imagem de entrada em partes ou objetos constituintes (GONZALES e WOODS, 1993). Segundo Bengtsson et al. (2004), muitas vezes, essa é a etapa mais importante e



Figura 5 - (a) Imagem sem equalização e (b) Imagem equalizada.

mais difícil do processamento de imagens.

A tarefa de segmentação é altamente dependente do tipo de amostra a ser analisada. Por exemplo, se o objetivo é estudar as características nucleares das células em amostras de citologia, onde as células nas imagens são isoladas, com fundo limpo e os núcleos são bem corados, então um simples método de limiarização automática pode ser suficiente. No entanto, se o objetivo é separar objetos não bem definidos, que apresentam sobreposição e variação de contraste, como acontece normalmente nas imagens com células somáticas, então a tarefa pode ser complexa.

A limiarização é um método de classificação baseado nas intensidades de *pixels* para converter uma imagem de tons cinza para P&B (ENSAFI, 2011). Esse método é eficiente para aplicações de processamento de imagem onde os níveis de cinza dos *pixels* que pertencem ao objeto são substancialmente diferentes dos níveis de cinza que pertencem ao fundo. Segundo Gonzalez e Eddins (2004), a limiarização é o estágio primário em problemas de segmentação de imagens e, devido à simplicidade dos algoritmos, outras técnicas devem ser tentadas quando essa não apresentar resultados satisfatórios.

A limiarização é uma operação de baixo nível que é essencial para realizar outras tarefas de nível mais elevado. O resultado da operação de limiarização é uma imagem P&B que irá indicar o estado dos objetos em primeiro plano. Esses objetos de primeiro plano poderiam ser, por exemplo, um texto impresso ou uma célula, enquanto que o estado complementar, o fundo da imagem.

Os métodos de limiarização são baseados nas características do histograma de



intensidades dos *pixels* da imagem. A função que realiza esta técnica é descrita pela equação a seguir:

$$g(x, y) = \begin{cases} 1, & \text{se } f(x, y) \geq T \\ 0, & \text{se } f(x, y) < T \end{cases} \quad (3)$$

em que  $g(x, y)$  é a função que representa a imagem limiarizada,  $f(x, y)$  é a função que representa a imagem original e  $T$  é um valor de nível de cinza denominado limiar.

É difícil escolher a melhor forma de se definir o limiar, pois isso depende das condições da imagem (contraste, iluminação, complexidade da cena) e também do resultado que se espera através da limiarização (Figura 6).

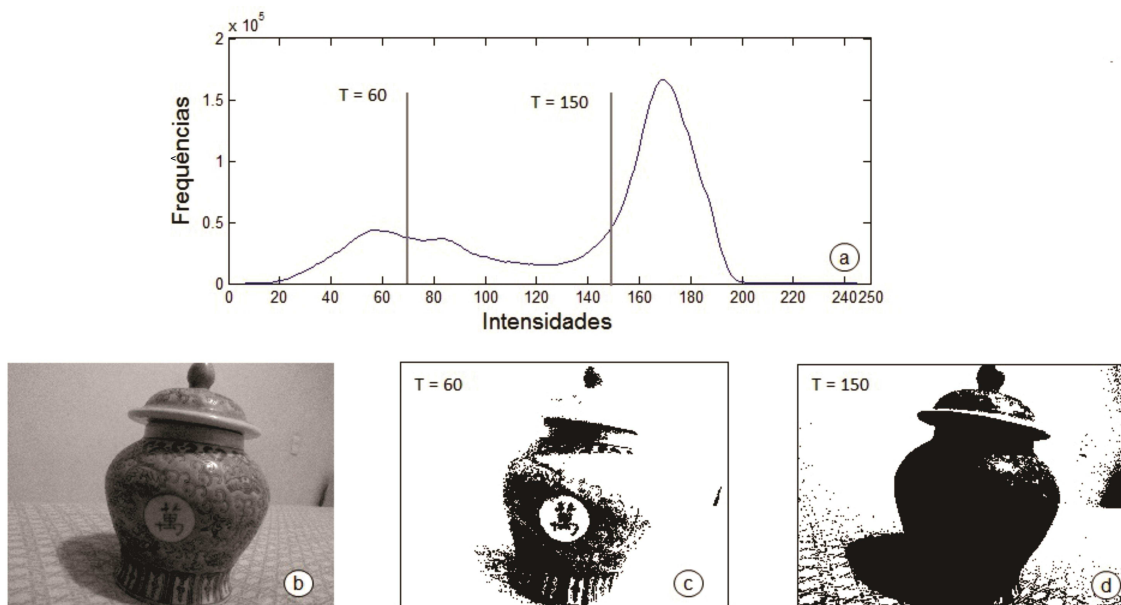


Figura 6 - Limiarização da imagem : (a) representa o histograma da imagem (b); a limiarização da imagem para os valores de  $T = 60$  e  $T = 100$  são apresentadas nas figuras (c) e (d), respectivamente.

Devido à eficiência para grandes números de aplicações, os métodos baseados em limiarização têm sido largamente utilizados na literatura. Por exemplo, Chang et al. (1997) utilizaram a limiarização para análise de imagens de documentos, Russ e Russ (1987), Sieracki et al. (1989), Wu et al. (2000) e Qian et al. (2009) para segmentação de células, Wu et al. (2000), Qian et al. (2009) e Tian et al. (2001) para microscopia de varredura e Oh e Lindquist (1999) para imagens de tomografia computadorizada.

Sezgin et al. (2004) listaram mais de 40 técnicas diferentes de limiar que foram



relatados nos últimos 34 anos. Devido ao grande número de algoritmos já propostos, é comum a divergência a respeito da classificação dos métodos de limiarização. No entanto, as mais utilizadas na literatura são as classificações apresentadas por Weszka (1978) e Sahoo et al. (1988).

Para Weszka (1978) os métodos de limiarização são classificados em três grupos: global, local e dinâmico. O critério usado nessa definição é a dependência do operador de limiar. Se for determinado usando apenas as intensidades é classificado como global. Se usar as intensidades e alguma informação espacial da imagem é classificado como local. Se usar ambas as informações mais as coordenadas espaciais é classificado como dinâmico.

Por outro lado, para Sahoo et al. (1988) as técnicas de limiarização se dividem apenas em dois grandes grupos: globais e locais. Os limiares globais são aqueles que dividem toda a imagem usando apenas um limiar. Os locais dividem a imagem em “subimagens” e para cada “subimagem” é atribuído um limiar. Esta técnica também é conhecida como limiar multinível.

Segundo Tsai (1995) os algoritmos de limiarização global ainda podem ser classificados como paramétricos e não paramétricos. Nas abordagens paramétricas a distribuição de nível de cinza de cada classe é assumida para se ter uma função densidade de probabilidade que geralmente é assumida como sendo uma distribuição Gaussiana (WESZKA, 1979; SNYDER et al., 1990). Dessa maneira, busca-se uma estimativa dos parâmetros da distribuição que melhor atende os dados do histograma. O resultado é tipicamente um problema de otimização não linear que é computacionalmente dispendioso e de solução demorada. Nas abordagens não paramétricas, buscam-se os limites que separam as regiões de nível de cinza de uma imagem de maneira ótima e de acordo com alguns critérios, tais como a variação entre classe ou a entropia.

Embora o critério de limiarização baseado na variação entre classe seja mais amplamente aplicado, devido à sua simplicidade e coerência do desempenho, seu desempenho, no entanto, pode ser não confiável sob a condição de tamanhos de população altamente desiguais (TSAI, 1995). Lee et al. (1990) também descobriram que os métodos de variância entre classe e entropia não funcionam corretamente se objetos forem de tamanhos muito pequenos.

Em comparação com os métodos paramétricos as abordagens não paramétricas são computacionalmente eficientes e simples de programar. Por esse motivo, diversos autores

têm proposto métodos para encontrar limiar global em histogramas com mais de um pico predominante (PUN, 1981; JOHANNSEN, 1982; KAPUR et al., 1985; LEE et al., 1990; CHI et al., 1996; LI, 1993).

Métodos mais recentes para determinar limiares em imagens com histogramas multimodal foram apresentados por Pham e Crane (2006), Medina-Carnicer (2008), Medina-carnicer e Carmona-Poyato (2011), Jang et al. (2013), Alves de Melo et al. (2014), Mahapatra et al. (2014) e Mishra et al. (2014).

### 1.5.6 Limiarização Unimodal

Na Figura 7 é apresentado um exemplo de histograma bimodal e de outro unimodal com as intensidades variando entre 0 (preto) e 255 (branco). No histograma bimodal é possível observar dois picos: o primeiro representa o objeto (escuro) e o segundo o fundo da imagem (claro). Eles estão separados por um vale que, nesse exemplo, seria o local mais apropriado para estabelecer o limiar (FIGUEIREDO, 2003).

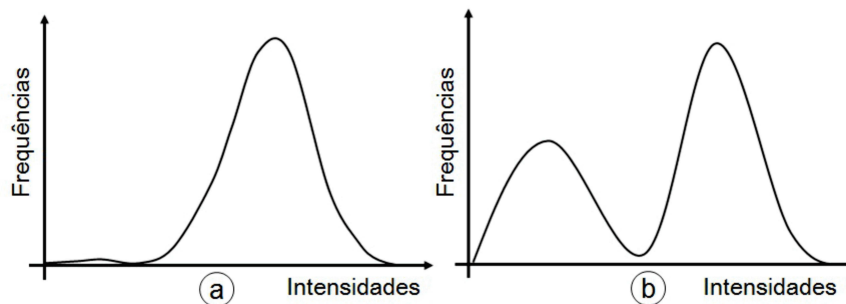


Figura 7 - Histograma de imagens tons de cinza: (a) Unimodal e (b) Bimodal.

Após determinar um limiar  $T$ , a imagem inteira fica dividida em duas regiões: o objeto, que recebe a cor preta (0), e o fundo da imagem, que recebe a cor branca (255), ou vice-versa.

Geralmente, os algoritmos de limiarização são destinados para histogramas multimodais. No entanto, em muitas aplicações práticas como detecção de bordas, detecção de canto, fluxo óptico, textura e aproximação poligonal a limiarização unimodal é necessária.

Por exemplo, Gonzales-Barron e Butler notaram forte relação entre imagens microscópicas do pão e a forma unimodal correspondente dos histogramas (Figura 8). Essa relação também foi notada por Wiltschi et al. (2000) em imagens microscópicas do aço, por Sifre - Maunier et al. (2006) em imagens do tecido conjuntivo intramuscular bovino,

por Medina-Carnicer (2008) em imagens do músculo bovino e por Baradez et al. (2004) em imagens adquiridas através de microscopia confocal (Figura 9).

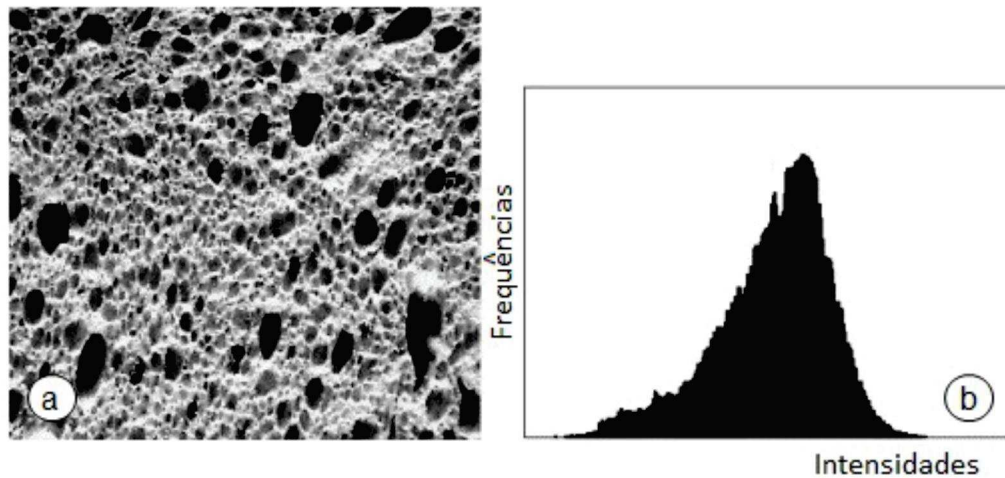


Figura 8 - (a) Imagem tons de cinza microscópica do pão e (b) Histograma de frequência unimodal correspondente. Fonte: Gonzales-Barron e Butler (2006).

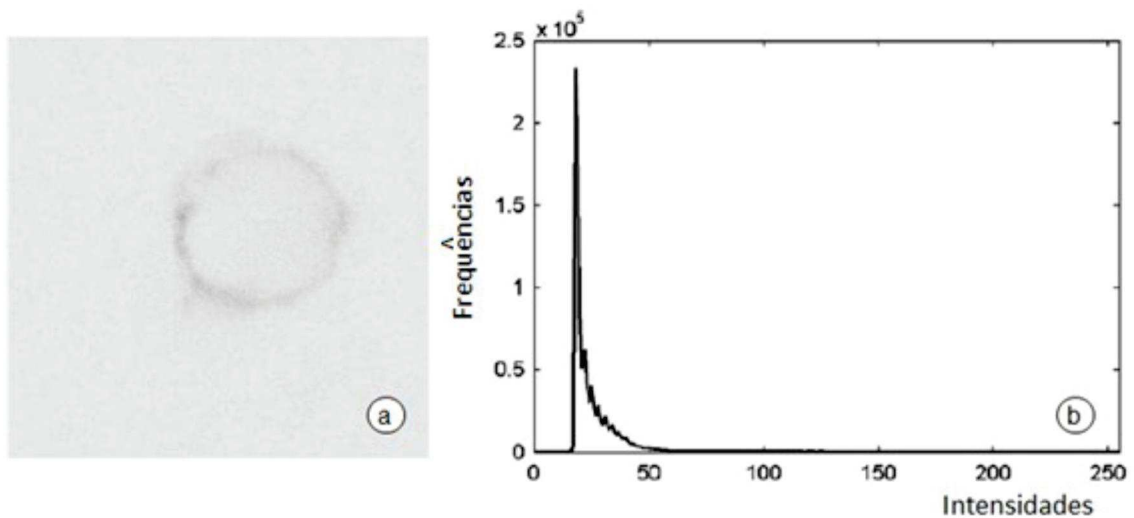


Figura 9 - (a) Imagem em tons de cinza com microscópica confocal e (b) Histograma de frequência unimodal correspondente. Fonte: Baradez et al. (2004).

Segundo Baradez et al. (2004) métodos destinados à limiarização de histogramas bimodais podem falhar se o histograma for unimodal, pois, nesse caso, encontrar o vale que determina o limiar não é tão óbvio. Medina-Carnicer (2008) esclarece que as técnicas de limiarização multimodal não são estáveis quando aplicadas a histogramas unimodais. Por esse motivo, diversos algoritmos têm sido propostos com o intuito de obter limiar de histogramas desse tipo.

Bhanu e Faugeras (1982) introduziu o problema de limiarização unimodal apresentando uma abordagem para rotulagem em histograma unimodal e Tsai (1995) propôs um método baseado na curvatura local para detectar pontos de descontinuidade em tais histogramas.

Um algoritmo de limiarização baseado na forma unimodal do histograma foi apresentado por Rosin (2001); uma técnica unimodal, utilizando a variância entre classe e mapas probabilísticos por Baradez et al. (2004) e Sifre-Maunier et al. (2006), respectivamente.

Baseado na relação entre as imagens obtidas por sistemas de detecção de bordas e forma unimodal dos histogramas, Medina-Carnicer (2008) propôs um método não paramétrico, utilizando o espaço ROC (receiver operating characteristic). A mesma relação foi notada por Medina-Carnicer e Carmona-Poyato (2011) que propôs um algoritmo de transformação do histograma para melhorar o desempenho de Rosin e Otsu. Nesse trabalho, a transformação é obtida aumentando o nível de cinza de alguns *pixels* da população dominante.

No entanto, segundo Sezgin et al. (2004) não existem critérios ou medidas claras para avaliar e classificar a eficiência dos algoritmos de limiarização. Dessa maneira, todos estes métodos apresentam vantagens e desvantagens e o sucesso de um método específico depende exclusivamente do grupo de imagens que se deseja tratar.

Por exemplo, o método de Rosin apresenta eficiência comprovada em aplicações gerais; no entanto, para detecção de bordas Medina-Carnicer e Carmona-Poyato (2011) concluiu que esse método apresenta resultados muito ruidosos. Baradez et al. (2004) esclarece também que nenhum dos métodos de limiarização unimodal é adequado para imagens de microscopia confocal.

Por outro lado, devido à eficiência e à simplicidade de implementação para aplicações gerais, não relacionadas com a detecção de bordas, a limiarização unimodal de Rosin (2001) tem sido largamente utilizada, referenciada e comumente utilizada para validação de novos algoritmos de limiarização (SIFRE-MAUNIER et al., 2006; MEDINACARNICER, 2008; MEDINA-CARNICER e CARMONA-POYATO, 2011).

### 1.5.7 Limiarização de Rosin

Alguns tipos de imagens possuem histogramas essencialmente unimodais, pois uma porção maior de apenas uma classe de *pixels* (por exemplo, o fundo) está presente na imagem e domina o histograma (MEDINA-CARNICER, 2008).

Em tais circunstâncias, muitos dos algoritmos de limiarização poderão falhar. Desta maneira, o método de limiarização de Rosin foi proposto para encontrar o limiar em imagens com distribuição de frequência unimodal (ROSIN, 2001).

A Figura 10 mostra os principais parâmetros associados à limiarização de Rosin. O método parte da premissa de que existe um pico principal único de frequência  $H_0$  para o tom de cinza  $k_0$ , desprezando-se *outliers* próximos dele.

O primeiro valor de tom para frequência nula é definido como  $k_m$ . Para obter o limiar de Rosin, uma reta é traçada entre o ponto  $(k_0, H_0)$  até o ponto  $(k_m, 0)$ . O limiar escolhido por esse método é o valor de tom  $k_n$  que maximiza a distância  $d$  perpendicular à reta. Ou seja, o limiar, então, é o valor do histograma que maximiza a distância perpendicular entre a curva que representa distribuição de frequência dos níveis de cinza da imagem e a reta que passa pelo valor mínimo  $(k_m, 0)$  e máximo do histograma  $(k_0, H_0)$ .

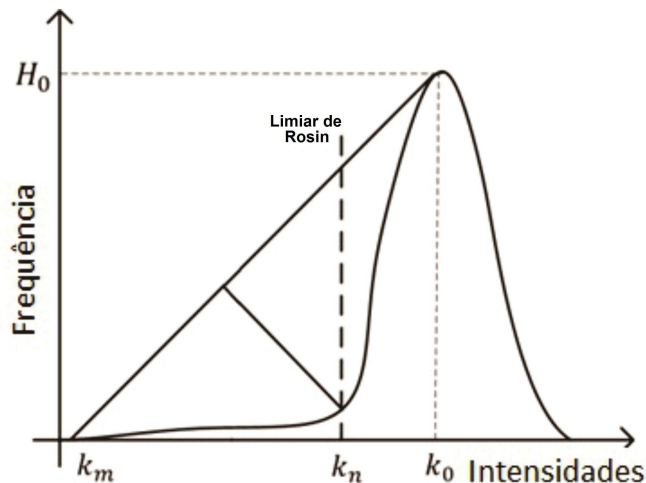


Figura 10 - Método de limiarização de Rosin.

Segundo Coudray et al. (2010), esse método resulta em muitos casos adequados, mas tende a ser sensível a parâmetros como: flutuações estatísticas do histograma e posição do ponto final (o mais alto nível de cinza).

### 1.5.8 Limiarização de Otsu

Esse é o método de limiarização global histograma mais popular e tem sido amplamente utilizado em aplicações nas áreas da medicina (GARGOURI et al., 2014; PRIYA et al., 2013), em análise de imagens de satélite (DE-LU, 2007), contagem de veículos em rodovias (ALBA-FLORES, 2005), contagem de células sanguíneas (YAO, 2009; BUXTON et al., 2007), caracterização de materiais (GOMES, 2001), segmentação de frutas (MIZUSHIMA, 2013), dentre outras.

Uma das vantagens desse método é que ele não se restringe a apenas um tipo de histograma da imagem, podendo, portanto, ser aplicado em histogramas unimodais, bimodais ou multimodais.

De acordo com Fan e Lei (2012), esse método pode obter bons resultados de segmentação para muitos casos. Sua principal desvantagem está no desempenho sensível à presença de ruídos na imagem (LEE et al., 1990). Por outro lado, segundo Nashat et al. (2012) o método funciona excepcionalmente bem somente se o histograma for bimodal. Além disso, Petrou e Bosdogianni (1999), Sezgin et al. (2004) e Nashat et al. (2012) afirmam que Otsu possui melhor desempenho para imagens com maior variância de intensidade, podendo falhar, entretanto, se o histograma for unimodal ou perto de unimodal.

Esse método é fundamentado no fato de que o histograma de uma imagem é composto por duas classes, uma representando os *pixels* do objeto e outra representando os *pixels* do fundo. Cada classe possui características próprias, ou seja, média e desvio padrão próprios (PEDRINI, 2008).

O algoritmo testa todos os possíveis limiares e calcula uma medida de variância das intensidades de cinza dos *pixels* para cada uma das classes (fundo e primeiro plano). O objetivo é achar o limiar que minimiza variância intraclasse que é dada pela equação

$$\sigma_W^2(t) = \omega_1(t)\sigma_1^2(t) + \omega_2(t)\sigma_2^2(t) \quad (4)$$

Se os *pixels* de uma dada imagem forem representados por  $L$  níveis de cinza, então o número de *pixels* com nível  $i$  é descrito por  $n_i$  e o número total de *pixels* por  $N = n_1 + n_2 + \dots + n_L$ .

O histograma normalizado é uma função de distribuição de probabilidade (Equação

5). As Equações 6 e 7 mostram as probabilidade de ocorrência, Equações 8 e 9 a média e Equações 10 e 11 variância por classe, respectivamente.

$$P(i) = \frac{n_i}{N} \quad (5)$$

$$\omega_1(t) = \sum_{i=1}^t P(i) \quad (6)$$

$$\omega_2(t) = \sum_{i=t+1}^L P(i) \quad (7)$$

$$\mu_1(t) = \sum_{i=1}^t \frac{P(i)}{\omega_1(t)} \quad (8)$$

$$\mu_2(t) = \sum_{i=t+1}^L \frac{P(i)}{\omega_2(t)} \quad (9)$$

$$\sigma_1^2(t) = \sum_{i=1}^t \frac{(1 - \mu_1)^2 P(i)}{\omega_1(t)} \quad (10)$$

$$\sigma_2^2(t) = \sum_{i=1}^t \frac{(1 - \mu_2)^2 P(i)}{\omega_2(t)} \quad (11)$$

O método de OTSU será demonstrado abaixo utilizando uma imagem  $I$  (Figura 11) com dimensão  $6 \times 6$ .

Os cálculos dos pesos, médias e variância serão apresentados para um provável limiar ( $T = 50$ ). Os demais resultados são apresentados na Tabela 3.

### PRIMEIRO PLANO

Peso:  $\omega_1 = \frac{7+5+8}{36} = 0,55$ .

Observe que o número total de *pixels* do histograma (Figura 11) é igual a 36.

Média:  $\mu_1 = \frac{0(7)+25(5)+50(8)}{0,55} = \frac{0(7)+25(5)+50(8)}{20} = 26,25$ .

Observe que é soma total dos *pixels* do histograma da Figura 12 é igual 20.

Variância:  $\sigma_1^2 = \frac{(0-26,25)^2(7)+(25-26,25)^2(5)+(50-26,25)^2(8)}{20} = 467,18$

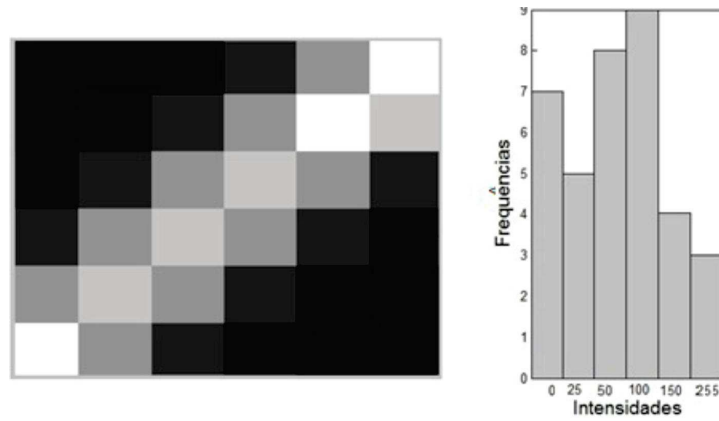


Figura 11 - Procedimento para determinação do limiar de OTSU.

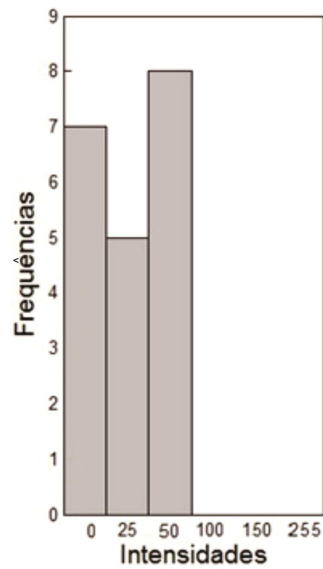


Figura 12 - Histograma de intensidades relacionadas ao primeiro plano.

### FUNDO

Peso:  $\omega_1 = \frac{9+4+3}{36} = 0,44$ .

Observe que o número total de *pixels* do histograma (Figura 11) é igual a 36.

Média:  $\mu_2 = \frac{100(9)+150(4)+255(3)}{36} = \frac{100(9)+150(4)+255(3)}{16} = 141,56$ .

Observe que a soma total dos *pixels* do histograma da Figura 12 é igual a 20.

Variância:  $\sigma_1^2 = \frac{((100-141,56)^2(9)+(150-141,56)^2(4)+(255-141,56)^2(3))}{16} = 3402,24$



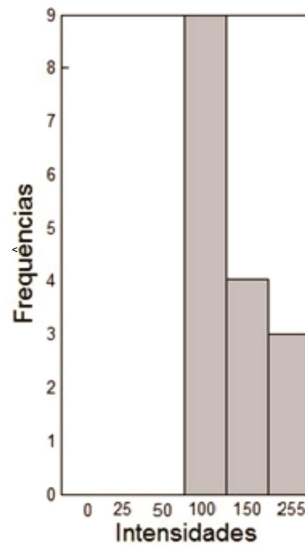


Figura 13 - Histograma com três intensidades.

Limiar	$T = 25$	$T = 50$	$T = 100$	$T = 150$
Peso: $\omega_1$	0,33	0,55	0,80	0,91
Média: $\mu_1$	10,41	26,25	49,13	61,36
Variância: $\sigma_1^2$	151,90	467,18	1486,32	2389,80
Peso: $\omega_2$	0,66	0,44	0,19	0,08
Média: $\mu_2$	111,04	141,56	195,00	255,00
Variância: $\sigma_2^2$	4131,20	3402,24	2700,00	0,00
Variância entre classes: $\sigma_W^2$	2804,77	1771,65	<b>1722,31</b>	2190,65

Tabela 1 - Limiar de OTSU ( $T = 100$ ) que minimiza a variância entre as classes.

Portanto, o valor de limiar que minimiza a variância entre classe é  $T = 100$  e resultado final da limiarização de Otsu é apresentado na Figura 14.

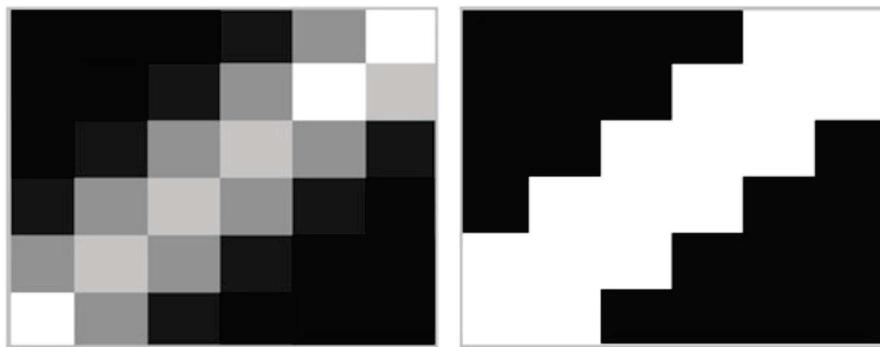


Figura 14 - Resultado da limiarização de Otsu.

### 1.5.9 Algoritmo de Agrupamento K-means

Vários fatores, tais como o ruído não estacionário, iluminação ambiente, contraste inadequado e tamanho do objeto não compatível com a cena, podem complicar a operação de segmentação através técnicas baseadas somente em limiares (SEZGIN et al., 2004).

Dessa maneira, é comum em problemas de reconhecimento de padrões que os algoritmos de limiarização sejam utilizados em conjuntos com outras técnicas de segmentação.

Técnica de agrupamento classifica os objetos em grupos diferentes, ou mais especificamente, particiona um conjunto em aglomerados (subconjuntos) de dados, de modo a que os dados em cada *cluster* compartilhe algum traço comum, muitas vezes de acordo com alguma medida de distância definida.

O particionamento de dados é uma técnica comum para a análise estatística dos dados, sendo utilizado em muitas áreas, incluindo a aprendizagem de máquina, de análise de imagem, reconhecimento de padrões, bioinformática e mineração de dados.

A tarefa computacional de particionar o conjunto de dados em  $K$  subconjuntos é muitas vezes referida como aprendizagem não supervisionada (DUBEY et al., 2013). O K-means é um algoritmo muito popular de agrupamento de dados abstratos (ex. um *pixel* ou grupo de *pixels*) que possuem características definidas por  $n$  parâmetros (HARTIGAN, 1979).

Através de uma métrica que estabelece a distância entre tais objetos, o K-means busca classificar padrões existentes em  $p$  agrupamentos distintos e particiona o conjunto de dados de entrada em  $K$  *clusters*. Cada grupo é representado por um centro (também chamado centro de *cluster*, a partir de alguns valores iniciais chamados pontos de sementes. Neste trabalho, o K-means foi utilizado para dividir a imagem original em dois *clusters* conforme descrito por Dubey et al. (2013).

#### Descrição do algoritmo:

Passo 1 - Inicialize os  $K$  centros  $(w_1, \dots, w_k)$ , que representam as classes.

Passo 2 - Cada classe  $C_j$  é associada a um centro  $w_j$ .

Passo 3 - Para cada dado de entrada  $p_l$ , assinale-o como pertencente à classe  $C_j$  (caso  $w_j$  seja o centro mais próximo a esse dado).

Passo 4 - Para cada classe  $C_j$ , atualize os centros  $w_j$  através da média de todas as

amostras de dados assinaladas como pertencentes a cada classe  $C_j$ .

Passo 5 - Calcular a função de erro  $E$ :

$$E = \sum_{j=1}^k \sum_{p_i \in C_j} |p_i - w_j|^2 \quad (12)$$

Passo 6 - Repetir 3, 4 e 5 até que o erro  $E$  não mude significativamente, ou até o término das  $N$  iterações.

Esse algoritmo tem sido aplicado em diversos problemas de segmentação de imagens, tais como o apresentado por Deoni et al. (2007) para segmentação de núcleos em imagens de ressonância magnética, Yao et al. (2013) para extração de contornos em peixes, Date e Akarte (2013) para segmentação de imagens do cérebro e seleção de frutas com defeitos (DUBEY et al., 2013).

Uma das principais limitações do algoritmo reside no fato de que o número de classes deve ser pré-determinado (ZALIK, 2008). Sendo assim, o algoritmo produzirá uma solução baseada na escolha inicial dos centroides que representam cada classe.

#### 1.5.10 Transformada de Watershed

A Transformada de Watershed foi originalmente sugerida por Digabel e Lantuéjoul (1978) e estendida para um quadro mais geral por Beucher et al. (1992) em aplicação para detecção de contornos.

O método pode ser aplicado para diferentes tipos de informação de imagem, tais como o nível de cinza, a distância ou o gradiente (SINTORN, 2005).

A transformada Watershed propõe uma abordagem morfológica para o problema de segmentação de imagens, interpretando as imagens como superfícies, onde cada *pixel* corresponde a uma posição e os níveis de cinza determinam as altitudes. Essa é uma das técnicas mais eficientes para segmentação e tem sido amplamente aplicada em muitos campos do processamento de imagem para a separação dos objetos conectados, como apresentado por Tulsani (2013), Pan et al. (2012), Chen et al. (2006), Soille e Vogt (2009) e Duan et al. (2011).

Devido à sua propriedade em separar formas semelhantes e que se tocam, esse método tem sido usualmente utilizado em aplicações voltadas ao processamento de ima-

gens microscópicas. Allalou e Wahlby (2009), por exemplo, a utiliza para segmentação de células em imagem capturadas por meio de microscopia de fluorescência, Pan et al. (2006) e Arslan et al. (2014) para segmentação de células do sangue humano.

Uma imagem em escala de cinza pode ser compreendida como uma superfície topográfica, onde os níveis de cinza representam a altitude da superfície. Assim, as áreas claras e escuras da imagem representam, respectivamente, as regiões mais altas e mais baixas da superfície. Ou seja, os valores altos (regiões claras) na imagem correspondem a montanhas ou colinas na paisagem, e valores baixos (regiões escuras) correspondem aos lagos ou vales.

Intuitivamente, considere que esta imagem tons de cinza, representada como uma superfície, passe a ser lentamente imersa em água, onde cada mínimo local (bacias) tenha um pequeno furo no centro. Intuitivamente, a água que entra por dois buracos diferentes tendem se encontrar. Desta forma, assim que elas se encontram, são construídas “barragens”. Essas “barragens” são as linhas da transformada Watershed.

Uma breve descrição do método, aplicado a uma imagem com duas dimensões é dada na Figura 15. Observe que o perfil da superfície topográfica apresentada nesta ilustração contém 4 mínimos regionais. Em cada um dos mínimos são feitos furos (Figura 15a). Esses furos são localizados na Figura 15a por setas. As águas então penetram pelos diferentes furos e, aos poucos, a superfície passa a ser inundada. Em seguida, constrói-se o primeiro dique (ver “barreira”) para separar as águas (Figura 15b). O segundo dique é apresentado na Figura 15c. Por fim, cada uma das bacias são separadas por linhas situadas acima de cada um dos diques (Figura 15c). Essas linhas são denominadas, linhas de Watershed.

No entanto, devido o grande número de diferentes aplicações, os métodos de segmentação baseados na Transformada de Watershed são sujeitos ao efeito de super-segmentação. Por esta razão, diversos aperfeiçoamentos do método têm sido apresentados na literatura (ANGULO, 2007; NG et al., 2008; WANG et al., 2011; XU et al., 2014). Uma revisão completa sobre definições, algoritmos e implementação de segmentação das bacias hidrográficas é apresentado por Roerdink (2000).

Neste trabalho, foi escolhido o método de Watershed para imagens P&B apresentado por Duan et al. (2011). Foi feita esta escolha, pois o objetivo em uma das etapas deste trabalho é separar *cluster* de células presentes em imagens P&B. Além disso, o método de

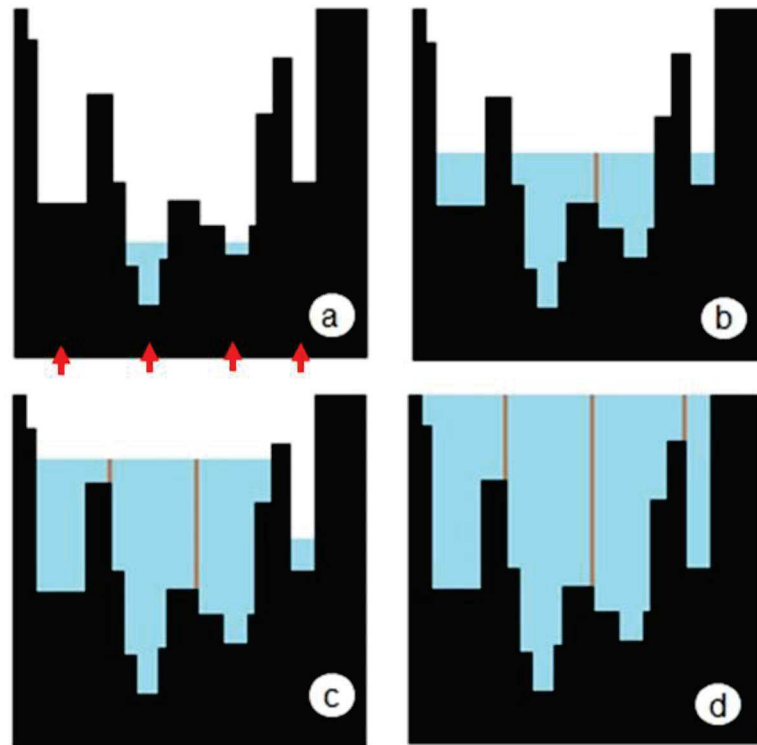


Figura 15 - Transformada de Watershed: (a) Mínimos regionais localizados por setas; (b) Formação do primeiro dique e (c)-(d) As linhas de Watershed.

Watershed para imagens P&B tem se mostrado adequado para diversas aplicações voltadas ao reconhecimento de padrões em imagens microscópicas. Por exemplo, Malpica et al. (1997) utilizaram este algoritmo para segmentação de núcleos interfásicos, Nilsson e Heyden (2005) e Grau et al. (2004) realizaram a segmentação de grupos de células sanguíneas, Pan et al. (2006) para segmentação de *cluster* de leucócitos e Karvelis et al. (2010) utilizou o mesmo princípio para segmentação de cromossomos sobrepostos. A Transformada de Watershed para imagens P&B é apresentada na seção 1.5.12.

#### 1.5.11 Transformada da Distância

Medir distâncias é útil em muitas aplicações de análise de imagem para o cálculo de medidas descritivas ou para posterior utilização da informação.

A TD produz um mapa de distância a partir de uma imagem P&B. Para cada *pixel* no interior dos objetos na imagem P&B, o *pixel* correspondente no mapa-distância tem um valor igual à distância mínima para o fundo. Ou seja, cada *pixel* do primeiro plano da imagem P&B, é substituído com um valor de cinza igual à distância do *pixel* ao *pixel*

de fundo mais próximo.

A TD é uma operação básica, aplicável em muitos problemas de processamento de imagens, como a compressão de dados, seleção de regiões de interesse, e diagramas digitais (EGGERS, 1998).

A distância entre os pontos de uma imagem pode ser calculada com diferentes métricas. A métrica mais natural, e também a mais utilizada em processamento de imagens, é a euclidiana, principalmente porque ela é invariante à rotação (ZAMPIROLI, 2003). A transformada de distância que utiliza a métrica euclidiana é chamada Transformada de Distância Euclidiana (TDE).

A TDE é amplamente utilizada em diversas áreas do processamento de imagens, tais como, dilatação e erosão (RUSS, 2010), diagramas de Voronoi (YE, 1988), esqueletização (SHIH e PU, 1995), Watershed de imagens P&B (SOILLE e VOGT, 2009), dimensão fractal (BRUNO et al., 2008) e navegação robótica (ZELINSKY, 1992).

A imagem resultante da TD é uma imagem nível cinza (Figura 16) onde cada elemento da imagem tem um valor de intensidade igual à distância mais próxima do fundo (YE, 1988; SINTORN, 2005). Um funcionamento da TD para uma imagem  $3 \times 3$  é apresentado na seção seguinte.

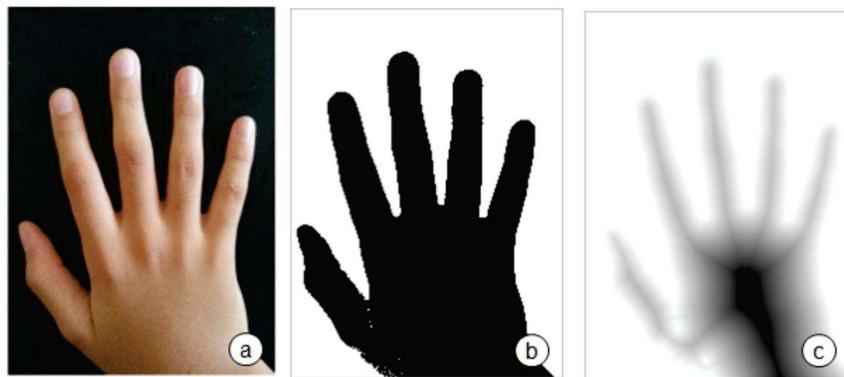


Figura 16 - (a) Imagem colorida; (b) Imagem P&B e (c) Resultado da TD.

#### 1.5.12 Watershed de Imagens Preto e Branco

A segmentação Watershed produz bons resultados para imagens em tons de cinza e com diferentes mínimos e captação de bacias. Para imagens P&B, no entanto, como existem apenas dois níveis de cinza 0 e 255 (preto e branco), antes de aplicar a Transformada

de Watershed é necessário usar a Transformada da Distância (TD) para pré-processar a imagem e torná-la adequada para a Transformada de Watershed.

O método de Watershed para Imagens P&B (DUAN et al., 2011) define a TD de uma imagem P&B como a distância Euclidiana de cada *pixel* preto do objeto para o *pixel* branco mais próximo (Equação 13).

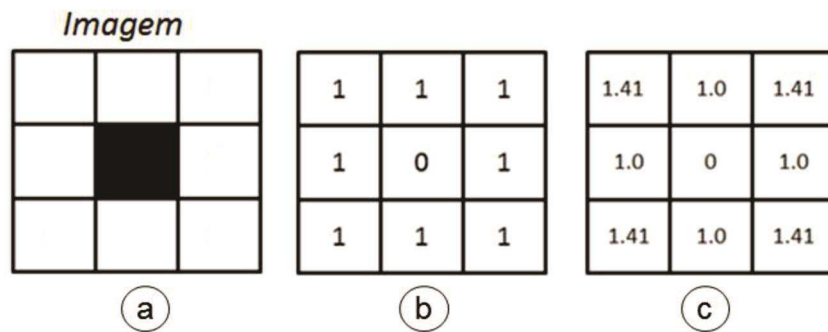


Figura 17 - (a) Imagem P&B  $3 \times 3$ ; (b) *Pixel* central preto representado por 0 e (c) Cálculo das distâncias.

Dessa forma, dada uma imagem P&B a distância Euclidiana entre dois *pixels*  $p_1$  e  $p_2$ , com coordenadas,  $(x_1, y_1)$  e  $(x_2, y_2)$ , respectivamente, é dada por:

$$D = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} \quad (13)$$

Há muitas maneiras diferentes de definir a distância entre dois *pixels*  $p_1$  e  $p_2$ , com coordenadas  $(x_1, y_1)$  e  $(x_2, y_2)$ . Os efeitos dessa escolha em relação às linhas do Watershed são apresentados em (CHEN et al., 2004).

Várias abordagens para a segmentação de imagens P&B, usando bacias hidrográficas e TD foram encontrados na literatura (ZHANG et al., 2014; WANG e PALIWAL, 2006; FABBRI et al., 2008).

Uma vez que a segmentação da imagem em objetos e fundo tenha sido realizada, o Watershed para imagens P&B pode ser usado para dividir objetos em partes menores. Este método é geralmente desejável, pois, muitas vezes, os objetos nas imagens encontram-se ligados uns aos outros em aglomerações.

O Watershed para imagens P&B aplicado após a TD pode servir como uma forma eficiente de dividir a aglomeração em objetos individuais (ver Figura 18).

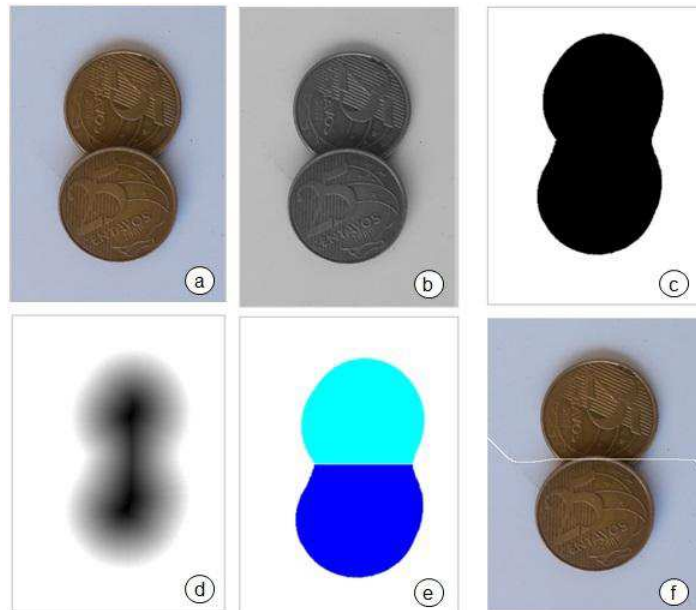


Figura 18 - (a) Duas moedas em sobreposição; (b) Imagem em tons de cinza; (c) imagem após limiarização; (d) TD da imagem P&B; (e) As regiões resultantes após o algoritmo de Watershed e (f) Imagem original com as linhas de Watershed.

#### 1.5.13 Vizinhança de um *pixel*

Um *pixel*  $p$  de coordenadas  $(x, y)$  possui 4 vizinhos nas direções vertical e horizontal. As coordenadas deste *pixel* é dada por

$$(x - 1, y), (x + 1, y), (x, y - 1), (x, y + 1). \quad (14)$$

Esse conjunto de *pixels* é chamado de vizinhança-4 de  $p$ , e é representado por  $N_4(p)$  (Figura 19a).

Os quatro vizinhos mais próximos nas direções diagonais de  $p$  são os *pixels* de coordenadas

$$(x - 1, y - 1), (x - 1, y + 1), (x + 1, y - 1), (x + 1, y + 1). \quad (15)$$

Esses *pixels* constituem o conjunto  $N_d(p)$  (Figura 19b). Portanto, a vizinhança-8 de  $p$  é definida como:

$$N_8(p) = N_4(p) \cup N_d(p) \quad (16)$$



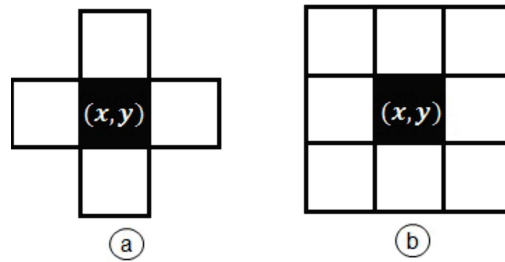


Figura 19 - (a) Vizinhança-4 e (b) Vizinhança-8 de um *pixel*.

#### 1.5.14 Algoritmo de contagem

No reconhecimento de padrões, a rotulagem dos componentes conectados em imagens P&B é uma das operações mais fundamentais (HU, et al., 2005). Os pontos em um componente conectado formam uma região candidata para representar um objeto.

Para a rotulagem bidimensional de imagens (2D), vários algoritmos foram propostos (LUMIA et al., 1983; SAMET e TAMMINEN, 1988; HARALOCK e SHAPIRO, 1991; BHATTACHARYA, 1996; SHAN et al., 2002; SUZUKI et al., 2003).

No entanto, segundo Suzuki et al. (2003) o algoritmo de Haralock e Shapiro (1991) é o representante mais utilizado em trabalhos para comparação de outros algoritmos. Esse algoritmo não utiliza nenhum armazenamento auxiliar para produzir a imagem rotulada.

O processo de rotulagem consiste em uma etapa de inicialização, mais uma sequência de propagação de rótulo da esquerda para direita, seguindo uma sequência de cima para baixo e baixo para cima. O processo ocorre até que nenhuma mudança de rótulo ocorra. A Figura 20 mostra uma imagem  $7 \times 6$ , onde o conjunto de *pixel* codificados com o valor 1 indica os objetos na imagem. A próxima etapa consiste em codificar cada *pixel* por meio de um rótulo único, conforme apresentado na Figura 20b e c. Em seguida, o valor de cada *pixel* não zero é substituído pelo valor mínimo dos seus vizinhos não nulos (Figura 20d). O algoritmo então continua avaliando o mínimo da vizinhança-8 de cada *pixel* não zero conforme apresentado na Figura 21.

A Figura 21 apresenta os resultados após a primeira passagem de cima para baixo (um ciclo), em que o valor de cada *pixel* não zero é substituído pelo valor mínimo dos seus vizinhos não nulos, em forma recursiva indo da esquerda para a direita e de cima para baixo. Observe, neste exemplo, que após um ciclo (Figura 21f), todos os *pixel* conectados passam ser rotulados por meio de um rótulo único (1, 6 e 5).

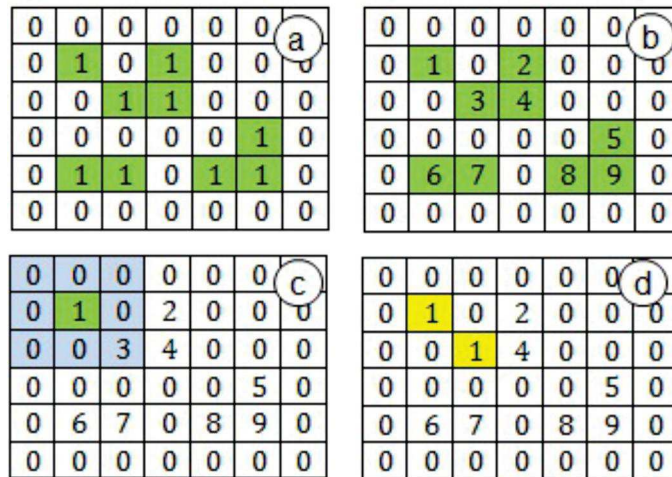


Figura 20 - Algoritmo de rotulagem de componentes conectados: (a) Imagem P&B; (b) Rótulo único para cada *pixel* codificado com o valor 1; (c) Avaliação da vizinhança-8 e (d) Resultado da operação de mínimo.

Para contar as formas rotuladas semelhantes às apresentadas na Figura 21f basta contar os algarismos diferente presentes na matriz correspondente a imagem. Neste caso, três formas são identificadas por meio dos três algarismos referentes aos rótulos 1, 6 e 5.

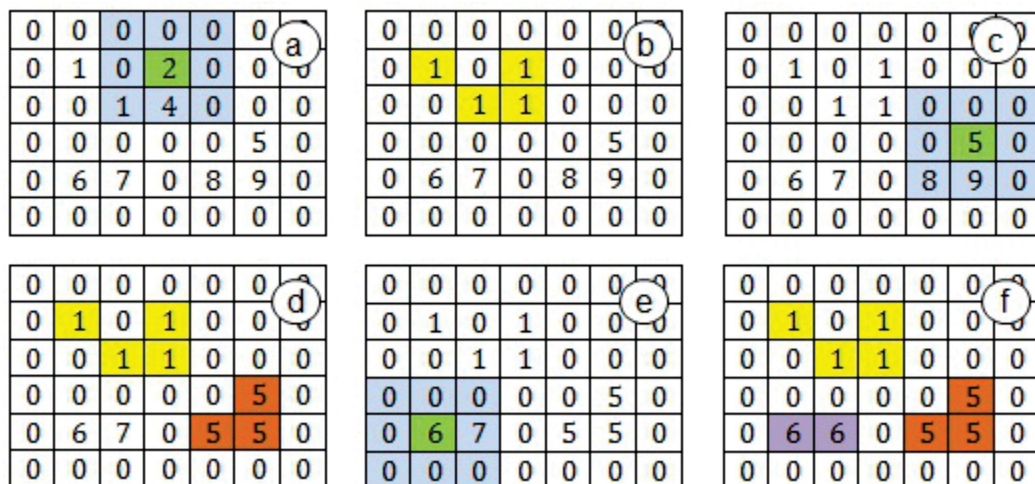


Figura 21 - Continuação do algoritmo iterativo de rotulagem apresentado na Figura 20.

### 1.5.15 Espaço de Cor CIELAB

Com o objetivo de desenvolver um espaço de cores perceptualmente uniforme e que atende às necessidades psicofísicas do homem, a CIE (Commission International de l'Eclairage) desenvolveu os espaços de cores  $L^*a^*b^*$  (C.I.E., 1986). Entende-se por espaço

de cores perceptualmente uniforme, o espaço onde a distância entre duas cores representa a diferença percebida pelo cérebro humano entre essas duas cores (CHENG et al., 2001; LUCCHESSEY e MITRAY, 2001).

O CIE  $L^*a^*b^*$  (CIELAB) é considerado o espaço de cor mais completo especificado pela Comissão Internacional de Iluminação (PASCHOS, 2001). Ele descreve todas as cores visíveis para o olho humano (Figura 22) e permite definir uma cor independentemente dos periféricos utilizados.

Neste espaço, três valores são atribuídos a cada uma das cores. O primeiro é a camada de luminância  $L^*$ , a qual varia de zero a 100 (0 para preto e 100 para o branco) e, os outros dois são camadas  $a^*$  e  $b^*$ , os quais representam duas gamas de cores que variam de -120 a +120 (YAM, 2004). A camada de cor  $a^*$  varia de verde para vermelho, e a camada de  $b^*$  varia de azul para amarelo (Figura 22).

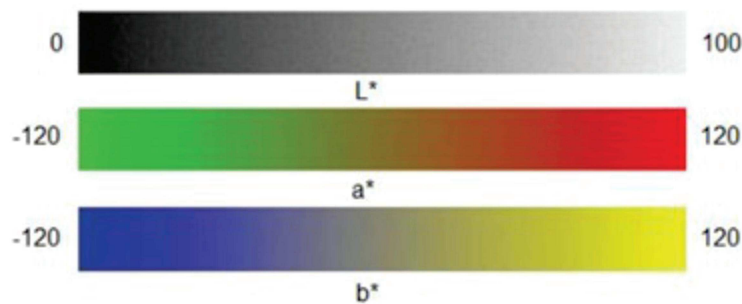


Figura 22 - Coordenadas do sistema de cor CIELAB.

A codificação das componentes  $a^*$  e  $b^*$  é baseada na teoria da cor oponente, o que indica que uma cor não pode ser vermelho e verde, ou azul e amarelo, ao mesmo tempo (SANCHEZ-CUEVAS et al., 2013).

Para obter os valores do  $L^*$ ,  $a^*$  e  $b^*$  é necessário converter inicialmente o espaço RGB para o XYZ através da equação 17, que foi estabelecida pela National Television System Committee (NTSC) (Cheng et al., 2001).

$$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \begin{pmatrix} 0,607 & 0,174 & 0,200 \\ 0,299 & 0,587 & 0,114 \\ 0,000 & 0,066 & 1,116 \end{pmatrix} \cdot \begin{pmatrix} R \\ G \\ B \end{pmatrix} \quad (17)$$

Em seguida, as equações 18, 19 e 20 convertem as cores do XYZ para o  $L^*a^*b^*$ , onde  $\frac{Y}{Y_0} > 0,01$ ,  $\frac{X}{X_0} > 0,01$ ,  $\frac{Z}{Z_0} > 0,01$  e  $(X_0, Y_0, Z_0)$  é o branco padrão.

$$L^* = 116 \left( \sqrt[3]{\frac{Y}{Y_0}} \right) - 16 \quad (18)$$

$$a^* = 500 \left( \sqrt[3]{\frac{X}{X_0}} - \sqrt[3]{\frac{Y}{Y_0}} \right) \quad (19)$$

$$b^* = 200 \left( \sqrt[3]{\frac{Y}{Y_0}} - \sqrt[3]{\frac{Z}{Z_0}} \right) \quad (20)$$

Estudos baseados em aplicações do espaço CIELAB têm sido propostos na literatura por Li e Yuen (2000), Kim et al. (2004), Paschos (2001) e Yue et al. (2011).

Segundo Hernandez-Gomez et al. (2009), a característica da cor pode facilitar o processo de segmentação de imagens, quando cores semelhantes precisam ser segmentadas em um mesmo objeto. Neste trabalho, o espaço de cor CIELAB foi utilizado em uma das etapas do algoritmo para segmentação de células somáticas do leite bovino e caprino.

## 2 METODOLOGIA PROPOSTA

A metodologia proposta foi baseada em duas etapas fundamentais. A primeira destinou-se à aquisição das amostras de leite e captura das imagens, enquanto a segunda consistiu no processamento das imagens através dos algoritmos propostos.

As imagens utilizadas nesse estudo foram obtidas por outra equipe no estado de São Paulo através da parceria estabelecida entre a Universidade Federal da Bahia e a Faculdade de Medicina Veterinária e Zootecnia - USP.

### 2.1 Aquisição das amostras de leite

No estudo realizado, procedeu-se à análise de 1.000 imagens provenientes de amostras do leite bovino e 200 imagens provenientes do leite caprino. As vacas da raça Holandesa e cabras da raça SAANEN são oriundas de propriedades localizadas no Estado de São Paulo (Brasil).

Inicialmente, foi realizada a higienização das tetas com solução clorada, secagem com papel toalha individual e antisepsia com álcool 70%. Após o descarte dos primeiros jatos de leite, as amostras foram colhidas em tubos de vidro com capacidade para 5 ml, imersos em gelo triturado e encaminhados ao laboratório em caixa de material isotérmico.

A preparação da lâmina de microscopia (esfregaços) seguiu o método padrão de Prescott e Breed (1910) que consiste na disseminação de 1  $\mu\text{L}$  de leite em uma área de um centímetro 1  $\text{cm}^2$ . Em seguida, deixa-se secar à temperatura ambiente para ser corada em seguida.

Os esfregaços lácteos dos bovinos foram preparados pela técnica de Rosenfeld, e os dos caprinos foram preparados utilizando-se os corantes verde de metil e pironina-Y. Após essa etapa, as lâminas foram examinadas e fotografadas via microscopia óptica, de acordo com os procedimentos descritos (BERRY e BROUGHAN, 2007).

### 2.2 Aquisição das imagens

O sistema para captura manual das imagens e contagem de células somáticas presentes nos esfregaços lácteos (Figura 23) foi constituído por uma câmera fotográfica digital (CoolSnap Color, Media Cybernetics<sup>TM</sup>, USA) acoplada ao microscópio óptico (Eclipse

800, Nikon<sup>®</sup>, Japan) no aumento de 1000 × e um computador que executa os algoritmos desenvolvido no ambiente Matlab<sup>®</sup>, versão 7.8 (licença concedida ao Instituto Federal de Educação da Bahia).

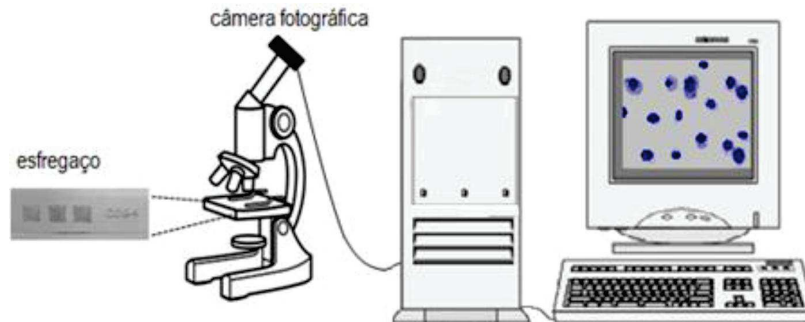


Figura 23 - Esquema para aquisição das imagens.

### 2.3 Metodologia proposta para bovinos

A Figura 24 ilustra esquematicamente o processo de segmentação e contagem das células somáticas do leite bovino. A metodologia proposta para a segmentação e contagem de células somáticas do leite de bovinos é baseada na combinação do algoritmo de clusterização K-means (HARTIGAN, 1979) juntamente com a transformada Watershed (ROERDINK, 2000) e uma nova proposta de limiarização de histogramas. Os resultados obtidos através da limiarização proposta foram comparados com os métodos de Rosin e Otsu. Esses métodos são comumente utilizados na literatura para validação de novos algoritmos de limiarização (SIFRE-MAUNIER et al., 2006; MEDINA-CARNICER, 2008; MEDINA-CARNICER e CARMONA-POYATO, 2011).

Inicialmente, a imagem RGB é convertida para o espaço de cor CIELAB (WYSZECKI, 1982; DE MELO et al., 2014) e, em seguida, agrupadas usando K-means (DUBEY et al., 2013). Esse algoritmo é definido para separar o conteúdo de imagem em dois grandes grupos. Em um deles, o *cluster 2*, toda a informação de imagem desejada do fundo e detritos são aglomerados. O outro, *cluster 1*, mantém as informações separadas e as células ligadas. A imagem do *cluster 2* é descartada e a imagem do *cluster 1* é convertida em tons de cinza, sendo, em seguida, transformada para P&B através da aplicação de um limiar de nível de cinza. O limiar é obtido diretamente a partir do histograma da imagem original no formato em tons de cinza e convertido para o formato P&B como detalhado

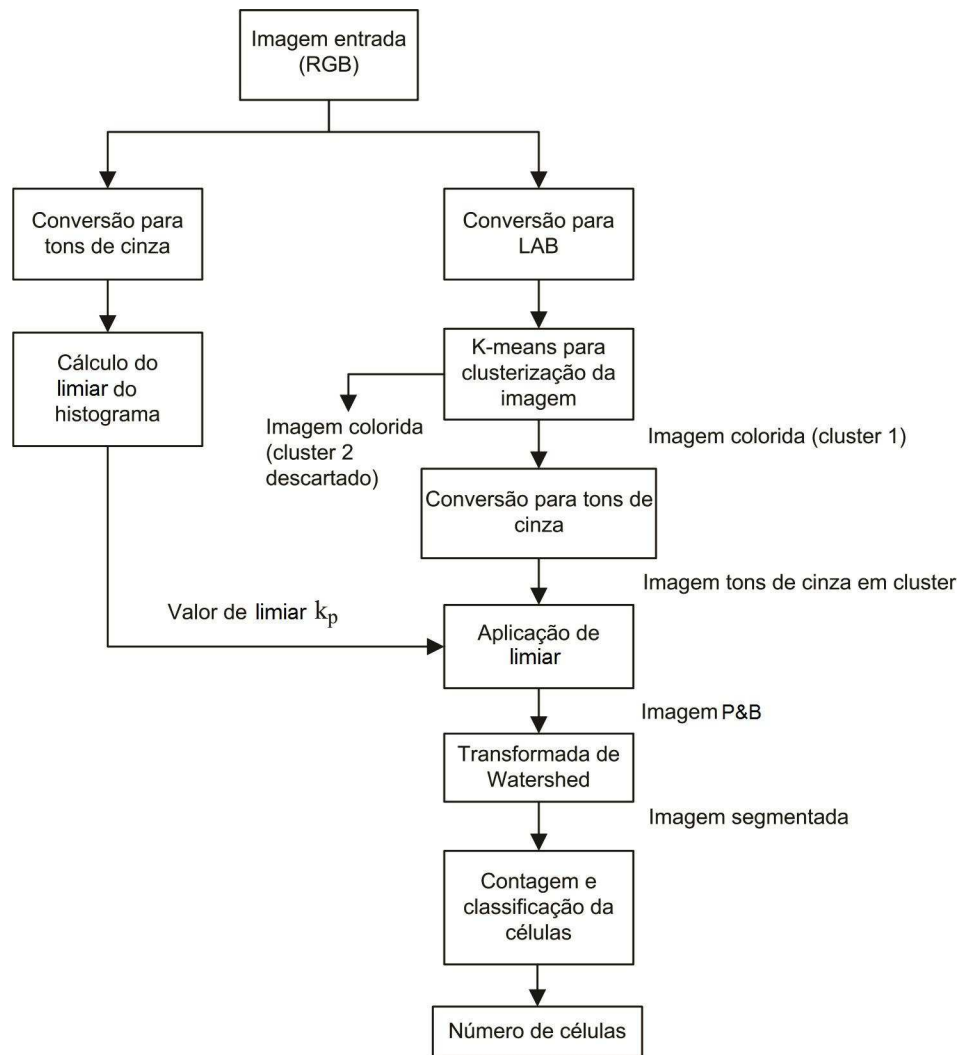


Figura 24 - Fluxograma da metodologia proposta para segmentação e contagem de células somáticas do leite bovino.

na seção 3.1.

Em seguida, utiliza-se a transformada da distância para mudar a imagem P&B para um mapa da distância em tons de cinza. Essa estratégia permite a aplicação da Transformada de Watershed para separar objetos conectados.

Após essa etapa, conforme descrito nas seções 1.5.14 e 3.3, as formas restantes são rotuladas, contadas e classificadas em dois grupos: células somáticas ou fragmentos de células.

## 2.4 Metodologia proposta para caprinos

A metodologia proposta para segmentação e contagem de células somáticas do leite de caprinos é baseada, fundamentalmente, no espaço de cor CIELAB, em um novo algoritmo para segmentação da cor e na Transformada de Watershed. Esta abordagem destina-se à segmentação de células do leite caprino. A Figura 25 mostra esquematicamente o processo de segmentação e contagem das células somáticas.

Inicialmente, a imagem de entrada RGB é convertida para o formato CIELAB. Esse espaço representa a imagem original através de dois grupos diferentes. O primeiro grupo são as células com tonalidade predominantemente verde e, o segundo, o fundo com tonalidade predominantemente vermelho. Depois de concluir esse procedimento, as células são inicialmente segmentadas por meio do algoritmo proposto para classificação *pixel* de cor (secção 3.2). Em seguida, a imagem resultante dessa operação é convertida para o formato P&B. Nessa fase, a fim de realçar a imagem, utiliza-se um filtro de partículas para remover pequenos objetos a partir da imagem P&B. As formas restantes são células isoladas, aglomeração de células ou fragmentos de células.

Os espaços vazios (buracos) nas formas identificadas são removidos e completados com *pixels* pretos, como pode ser visto em detalhe na seção 4.7. Em seguida, assim como mostrado para bovinos, utiliza-se a Transformada da Distância e a Transformada de Watershed.

Finalmente, conforme descrito nas seções 1.5.14 e 3.3, as formas restantes são rotuladas, contadas e classificadas em dois grupos: células somáticas ou fragmentos de células.

O resultado da CCS realizada através do processamento das imagens para bovinos e caprinos foi comparado com a contagem manual realizada por dois especialistas.



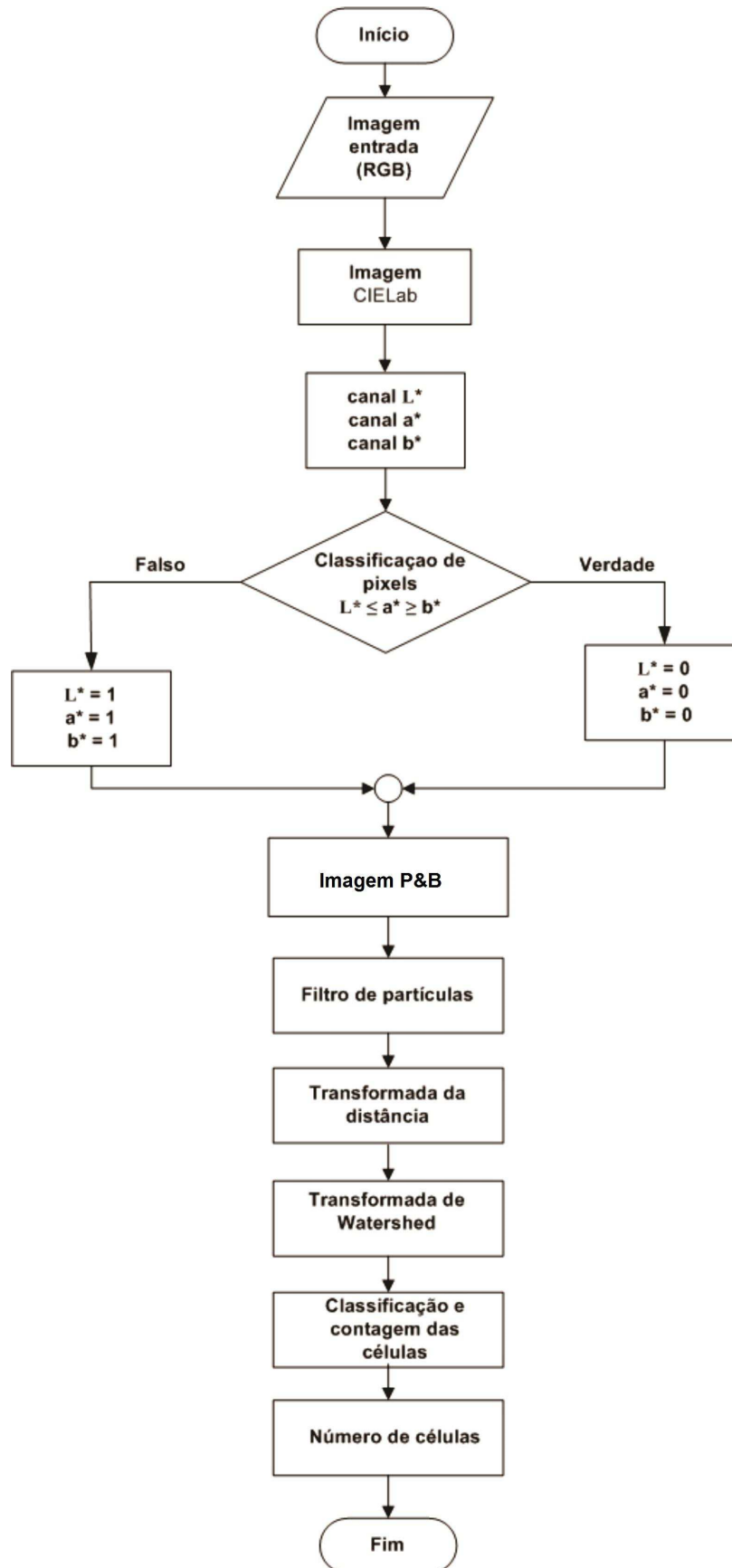


Figura 25 - Fluxograma da metodologia proposta para segmentação e contagem de células somáticas do leite caprino.

### 3 PROPOSIÇÃO DE NOVOS ALGORITMOS

Nesta seção são apresentados dois novos algoritmos de processamento de imagens. O primeiro consiste em uma nova proposta de limiarização de histogramas e é utilizado para a metodologia de segmentação de células somáticas do leite bovino. O segundo algoritmo consiste na classificação da cor baseado espaço no CIELAB e destina-se à segmentação de células somáticas do leite caprino.

#### 3.1 Proposição de novo limiar adaptativo

Assim como os demais métodos de limiarização unimodal, o método proposto assume que o histograma é unimodal (pico principal único). Além disso, o método considera premissas semelhantes às de Rosin e assume a região em torno do ponto  $(k_0, H_0)$  como relacionada ao fundo, enquanto que as informações das células e detritos correspondem às frequências de baixa intensidade. O nível de limiar está previsto para ser o melhor ponto possível de separação entre esses dois grupos, no sentido de obter o menor erro de classificação.

No entanto, observou-se, a partir de uma avaliação empírica de milhares de imagens de leite, que o nível de limiar tem uma intensidade mais baixa quando comparado ao de Rosin.

Com base nesta observação, propõe-se um nível de limiar  $k_p \in [k_m, k_o]$ , definido como um valor médio ponderado

$$k_p = \frac{k_m + k_0}{L} \quad (21)$$

onde  $L \in [b, a]$  é o parâmetro de ponderação limitado por  $a = \frac{k_m + k_0}{k_m}$  e  $b = \frac{k_m + k_0}{k_o}$ , como se mostra na Figura 26.

O valor ótimo de  $k_p$  é aquele que produz o melhor resultado em termos de concordância entre a contagem manual e a contagem automatizada. Isso ocorre para a condição  $L = L_{\max}$ . Para obter o valor máximo  $L_{\max}$  de ponderação do parâmetro  $L$ , nós calculamos  $k_p(L_n)$  em valores discretos  $L_1, L_2, \dots, L_N$ . Por simplicidade, assumimos que as etapas são igualmente espaçadas:

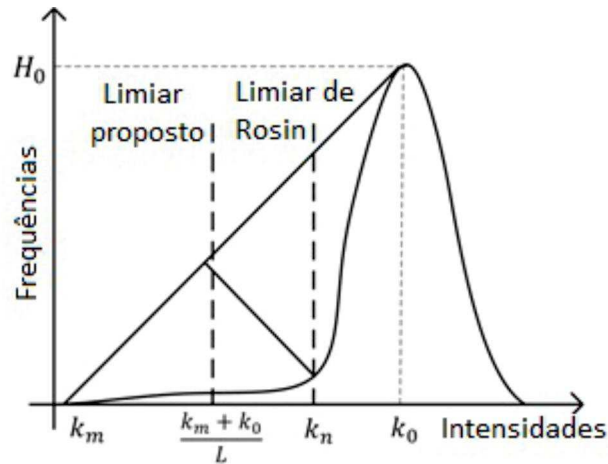


Figura 26 - Procedimento para a determinação do limiar de intensidade num histograma unimodal.

$$h = L_n - L_{n-1}, \quad n = 1, 2, \dots, N.$$

Para  $h = 0, 1$  no intervalo de  $L \in [0, 10]$ , nós calculamos uma sequência  $k_p(L_n)$ , com uma centena de pontos.

Para cada ponto dessa sequência, obtém-se um único valor de  $k_p$ , aplicando o método proposto para todo o conjunto de imagens e calculando o coeficiente de correlação  $C_n$  para a correspondente contagem manual. Assim, para cada  $L_n$  obtemos um coeficiente de correlação correspondente  $C_n$  registrado na Figura 27, e utilizando a inspeção obtemos o valor  $L_{\max}$  para máximo valor de  $C_n$ .

A Figura 27 mostra a correlação em função de  $L$  para um grande conjunto de imagens. O coeficiente de correlação é o ponto máximo  $L_{\max}$  e o melhor valor de  $k_p$  é calculado como se segue

$$k_p = \frac{k_m + k_0}{L_{\max}} \quad (22)$$

onde  $L_{\max} = 2, 4$  é o valor computado para o presente trabalho. Por outro lado, para outros grupos de imagens, em aplicações diferentes, o valor  $L_{\max}$  encontrado poderá ser diferente. Portanto, o método proposto pode se adaptar a grupos diferentes de imagens.

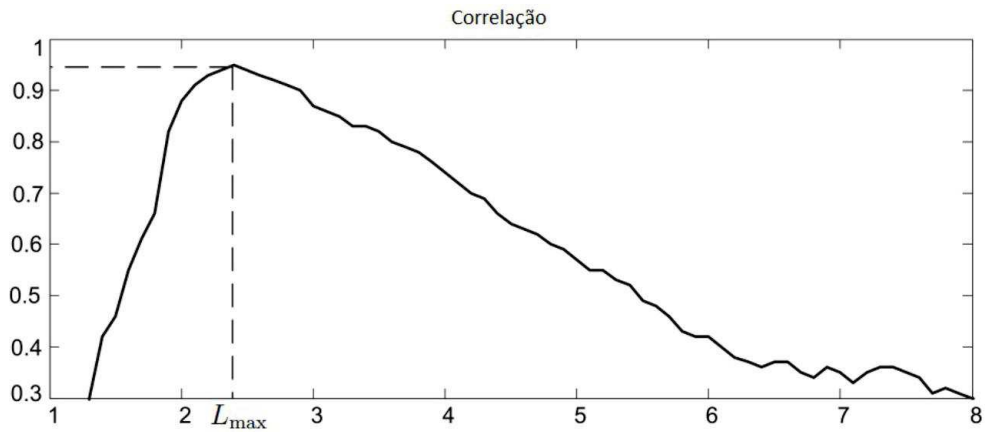


Figura 27 - Correlação  $C_n$  em função do parâmetro de ponderação  $L$ .

### 3.2 Proposição para segmentação da cor

Na transformação de RGB para CIELAB, as células somáticas são mapeadas em tons predominantemente verdes. O método proposto visa separar esses padrões do resto da imagem, de uma forma eficiente.

Para qualquer imagem dada, um *pixel* é percebido como verde, se a sua intensidade na coordenada G for maior do que aqueles em coordenadas R e B. Quanto maior for a intensidade do *pixel* de coordenadas em L, em relação a R e B, mais visualmente perceptível é a tonalidade verde.

De modo a separar os padrões que representam as células somáticas (tonalidade verde), o método proposto utiliza a imagem CIELAB e compara a intensidade de cada *pixel* da coordenada  $a^*$  com aqueles das coordenadas  $L^*$  e  $b^*$ . Cada *pixel* cuja intensidade de coordenadas  $a^*$  for maior do que a intensidade correspondente na coordenadas  $L^*$  e  $b^*$  será classificado como célula somática.

Assim, dada a imagem com dimensão  $n \times m$  o método proposto classifica como célula cada *pixel* que pertence à regra

$$\begin{cases} a_{ij} > b_{ij} \\ a_{ij} > L_{ij} \end{cases} \quad (23)$$

onde  $i = 1, 2, 3, \dots, n$  e  $j = 1, 2, 3, \dots, m$ .

### 3.3 Proposição para contagem das células em imagens preto e branco

O algoritmo de rotulagem de Haralock e Shapiro (1991) foi utilizado para contagem das formas (células somáticas) nas imagens P&B. Conforme apresentado na seção 1.5.14 esse algoritmo transforma uma imagem P&B para outra imagem correspondente, onde é atribuído um rótulo único a cada componente conectado. Após essa etapa, verifica-se o tamanho espacial de cada uma das formas (quantidade de *pixel*). Assim, formas com quantidade de *pixels* equivalente ao tamanho de 4 a 8 microns são classificadas como células somáticas.

Para determinar a faixa de *pixel* foi fixado o tamanho das imagens em  $278 \times 208$ . Em seguida, analisou-se manualmente um grupo de imagens para determinar a quantidade de *pixels* que seria utilizado para representar o tamanho das células. Assim, as formas dentro da faixa estabelecida foram classificadas como células e, as demais, classificadas como fragmentos de células.

## 4 RESULTADOS E DISCUSSÃO

A robustez dos métodos foi avaliada, comprovando-se a capacidade em processar grupos de imagens de bovinos e caprinos com variações de tonalidade, cor, contraste e presença de detritos, sem a necessidade de utilizar filtros de suavização, correção de contraste e de ruídos. Na literatura acessada, em trabalhos semelhantes, o volume da quantidade de imagens utilizadas para avaliar os métodos de processamento de imagens é menor que estabelecido neste trabalho (XUE et al., 2008; NA e HERU, 2009; XUE et al., 2009; NA e HERU, 2010 e WANG e XUE, 2010). Conforme abordagens semelhantes para contagens de células, apresentadas por Meas-Yedid et al. (2002), Refai et al. (2003), Bengtsson et al. (2004), Nilsson e Heyden (2005) e Nguyen et al. (2011) as métricas utilizadas para analisar a robustez dos métodos propostos foram a taxa de acerto entre a contagem automática e manual, a correlação Pearson, o erro absoluto médio e o erro relativo médio (seções 4.4 e 4.7).

### 4.1 Proposição de novo limiar: imagem sem detritos

A Figura 28 apresenta o resultado da limiarização proposta em comparação com os métodos de Rosin e Otsu. Observe que, nesse exemplo, quatro células podem ser claramente identificadas por inspeção visual humana. A Figura 28b e a Figura 28c apresentam a imagem resultante para o limiar de Otsu e de Rosin, respectivamente. Em ambos os casos, pode-se notar que as imagens em P&B ainda preservam formas indesejáveis. Por outro lado, o método proposto produz imagem mais evidente onde as células somáticas são mais fáceis de serem classificadas (Figura 28d). A Figura 28e mostra o histograma da presente imagem e os limites calculados por métodos diferentes.

Através do método de limiarização proposto é possível realizar a segmentação completa das células somáticas desde que estas sejam claramente separadas em imagens sem detritos.

Assim como os demais métodos de limiarização multimodal e unimodal apresentados na literatura, os métodos de Rosin e Otsu buscam o limiar utilizando informações individuais da distribuição de frequência de cada imagem. Ou seja, Otsu utiliza um fator que maximiza a variância interclasse e, Rosin, um fator que maximiza a distância perpendicular entre a curva que representa a distribuição de frequência dos níveis de cinza da

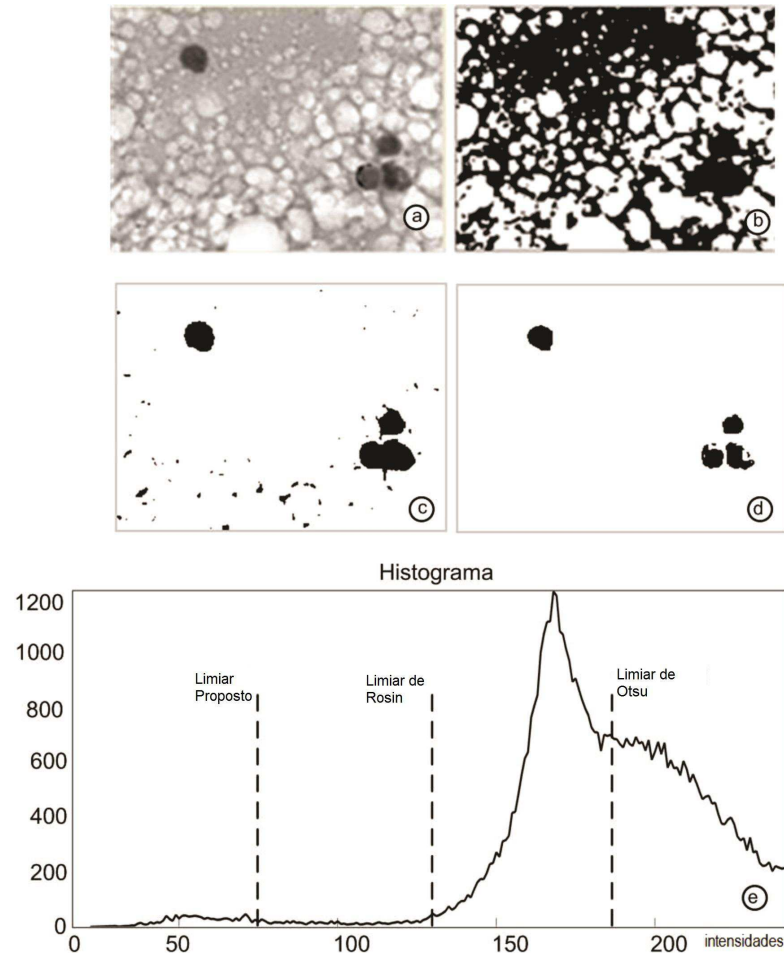


Figura 28 - (a) Imagem em tons de cinza de bovino. Resultados da limiarização das imagens utilizando: (b) Otsu; (c) Rosin; (d) Limiar proposto e (e) Limiar correspondente a cada método no histograma da imagem.

imagem e a reta que passa pela origem. Por outro lado, a limiarização proposta utiliza como fator de maximização, o parâmetro  $L$ . Esse parâmetro é ajustado utilizando-se todas as imagens do grupo, e reflete o melhor valor no denominador do modelo que maximiza a correlação entre o resultado da contagem automática e resultado esperado pelo especialista.

Em aplicações reais, por exemplo, em segmentação da retina, é comum processar grandes grupos de imagens com características muito semelhantes. No entanto, variações biológicas ou físicas inerentes às etapas que antecedem ou não as capturas das imagens podem modificar o padrão geral e individual esperado para as imagens. Dessa maneira, os histogramas das imagens serão sempre semelhantes, porém, nunca iguais.

Portanto, é adequado que os algoritmos de limiarização de histogramas utilizem modelos que melhor se ajustem a todas as imagens de determinado grupo e, não somente, a

cada imagem. Caso contrário, o método proposto poderá ser muito eficiente para algumas imagens e pouco eficiente para outras imagens do mesmo grupo.

#### 4.2 Proposição de novo limiar: imagem com detritos

A primeira impressão que se tem ao analisar imagens com células somáticas é que os detritos apresentam sempre cor semelhante ao preto. Dessa maneira, em imagens com tons de cinza os detritos seriam mais escuros que as células. Portanto, a utilização de dois limiares (faixa de intensidades) poderia, alternativamente, separar detritos, células e fundo. A Figura 29 ilustra um exemplo de histograma de células somáticas e a indicação de dois possíveis limiares.

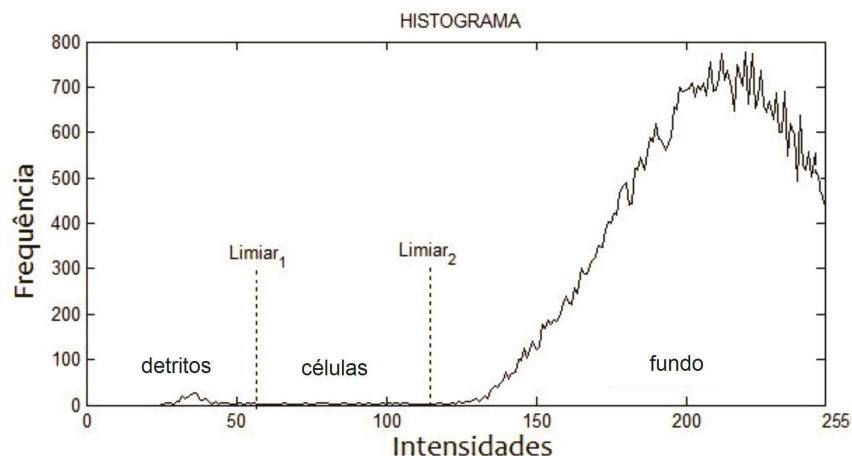


Figura 29 - Distribuição de frequência dos tons de cinza com dois limiares para extração de detritos.

Nesse caso, as células são representadas pela faixa de intensidades que vai do  $\text{Limiar}_1$  até  $\text{Limiar}_2$ . Ou seja, de acordo com o histograma as intensidades de zero até o  $\text{Limiar}_1$  representam detritos (tons mais escuros), as intensidades do  $\text{Limiar}_1$  até o  $\text{Limiar}_2$  as células (tons intermediários) e do  $\text{Limiar}_2$  até 255 o fundo.

Embora pareça razoável, essa abordagem não é bem sucedida, pois para alguns grupos de imagens deste trabalho, os tons que representam os detritos e células são muito semelhantes. Para esclarecer melhor essas observações considere o exemplo apresentado na Figura 30. Neste exemplo, para segmentar as células são testadas três faixas de intensidades.



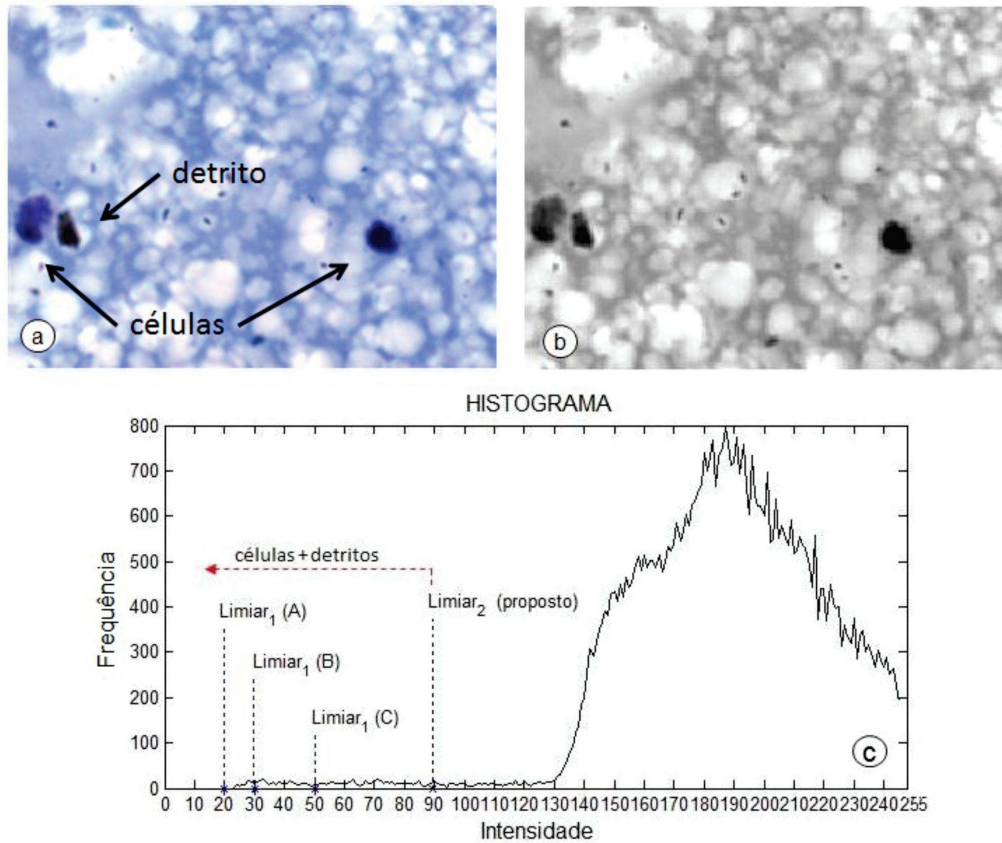


Figura 30 - (a) Imagem colorida com células somáticas e detritos; (b) Imagem correspondente em tons de cinza e (c) Distribuição de frequência da imagem tons cinza e três limiares para segmentação das células e extração dos detritos.

Faixa 1: Limiar<sub>1</sub>(A) - Limiar<sub>2</sub>

Faixa 2: Limiar<sub>1</sub>(B) - Limiar<sub>2</sub>

Faixa 3: Limiar<sub>1</sub>(C) - Limiar<sub>2</sub>

A limiarização da Figura 30b utilizando cada uma das faixas de intensidades (1, 2 e 3) é apresentada na Figura 31. Observe que nenhuma das faixas testadas realiza a segmentação das células simultaneamente a extração dos detritos. Esta observação reforça a necessidade de extrair os detritos antes de se utilizar métodos de limiarização.

Note que a estratégia de encontrar uma faixa de limiarização que permita, ao mesmo tempo, segmentar as células e extrair os detritos não é bem sucedida. Observe que a escolha de faixa diferentes de limiares implica em extração dos detritos acompanhada com destruição das células. Observe que a faixa escolhida na Figura 31b mantém a integralidade das células e dos detritos, porém não os diferenciam. Observe que as faixas

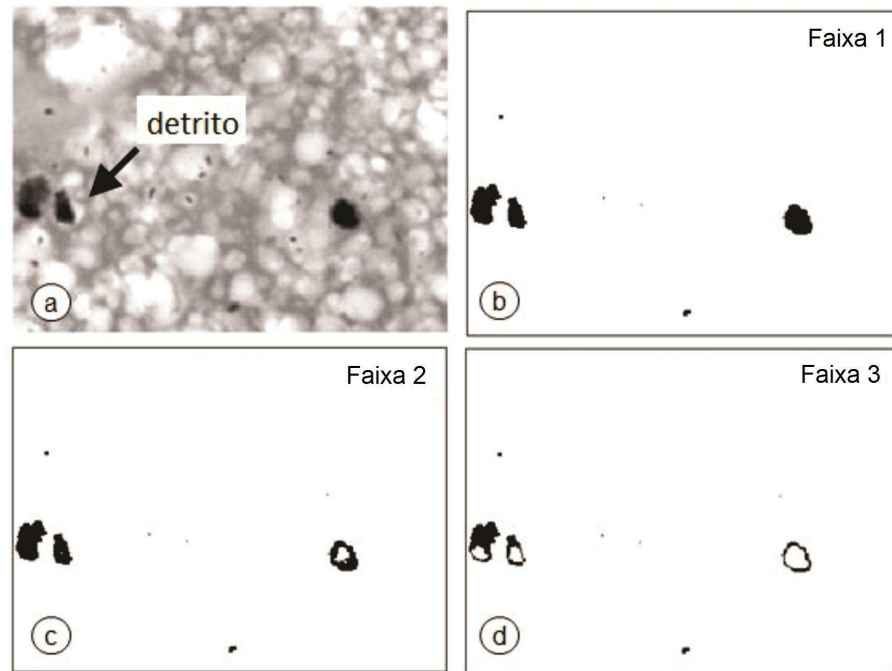


Figura 31 - Limiarização da imagem tons de cinza utilizando três faixas de limiar para extração de detritos e segmentação das células.

referentes às Figura 31c e Figura 31d apenas “destroem” algumas informações referentes a células ou a detritos.

Portanto, devido à semelhança dos tons que representam detritos e células somáticas nas imagens, metodologias que se baseiam na escolha de faixa de limiares são inadequadas para a extração de detritos e segmentação de células somáticas. Além disso, esse tipo de abordagem, baseado na escolha de dois limiares, acrescentaria outro limiar que deveria também ser ajustado automaticamente.

#### 4.3 Segmentação de células do leite bovino

Para demonstrar a robustez do algoritmo de segmentação proposto, será apresentada no exemplo seguinte uma imagem contendo células e detritos (Figura 32). A Figura 32a mostra a imagem RGB original, enquanto que as Figura 32b e Figura 32c mostram, respectivamente, a conversão da imagem de entrada para o formato tons de cinza e para o espaço CIELAB. A imagem tons de cinza é utilizada para determinar o limiar proposto  $k_p$  enquanto que a imagem CIEALB é utilizada como entrada do algoritmo K-means.

Quando comparado ao RGB, a representação do fundo e das células somáticas no espaço CIELAB é mais uniforme e propensa a menos variações. Ou seja, na imagem

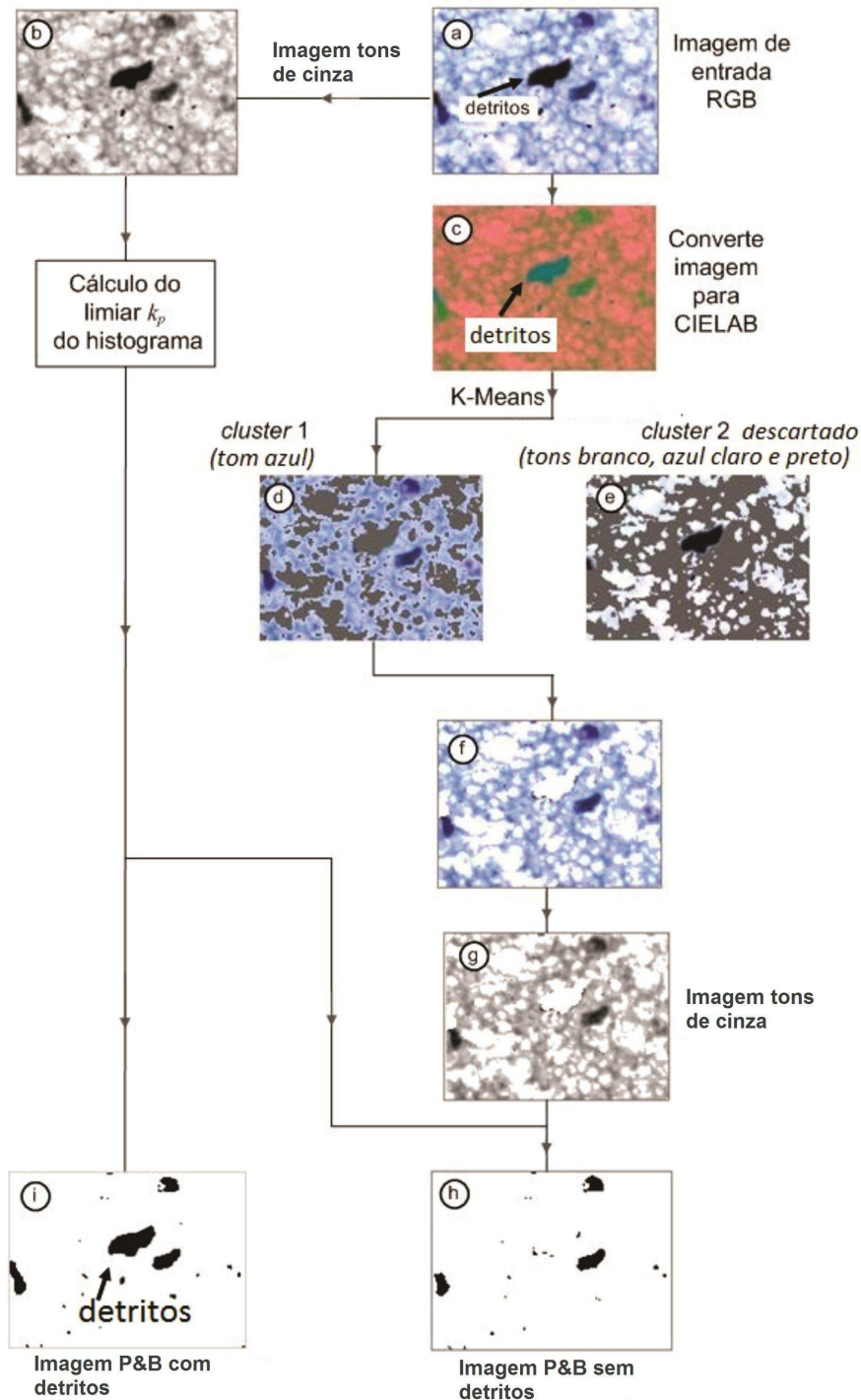


Figura 32 - Imagem sem tratamento (Figura 32a). Conversão de cinza (Figura 32b). Conversão para CIELAB (Figura 32c). Clusterização K-means (Figura 32d e Figura 32e). Imagem clusterizada (Figura 32f). Imagem clusterizada no formato tons de cinza (Figura 32g). Imagem limiarizada sem o K-means e com detritos (Figura 32i). Imagem sem detrito com limiar proposto (Figura 32h).

CIELAB é possível perceber visualmente uma quantidade menor de cores. Nesse espaço, as informações do brilho de cada cor são armazenadas na camada  $L^*$  e as informações

sobre a cor nas camadas  $a^*$  e  $b^*$ . De acordo com Sanchez-Cuevas et al. (2013), o espaço RGB é adequado para a detecção e exposição, mas não é uma boa escolha para realizar a segmentação de cor, devido ao fato de que ele tem uma elevada redundância. A alta correlação entre as coordenadas do espaço RGB pode torná-lo inadequado para a segmentação de cor.

A conversão do espaço RGB para espaço CIELAB é apresentada na Figura 32c e a clusterização K-means para  $p = 2$  padrões (*cluster 1* e *cluster 2*) encontra-se nas Figuras 32d e 32e.

A cor cinza do fundo nas Figuras 32d e 32e é proposital e tem a finalidade de facilitar a identificação dos *clusters* pelo leitor. O *cluster 1* é representado na Figura 32d com o tom azul. Esse *cluster* retém informações relacionadas com as células somáticas e de partes do fundo. O *cluster 2* é representado na Figura 32e com o tom branco, azul claro e preto. Esse *cluster* retém informações relacionadas com o restante do fundo, detritos e glóbulos de gordura. De acordo com a Figura 32f é possível observar a capacidade de K-means para eliminar fundo da imagem. Finalmente, a Figura 32f é então convertida para tons de cinza e limiarizada, utilizando limiar proposto calculado a partir da Figura 32b.

A limiarização direta, sem K-means, resulta em detritos contados como células somáticas o que pode aumentar o erro de contagem (Figura 32i). Para manter a relevância estatística adequada para o algoritmo proposto, o limiar  $k_p$  deve ser sempre obtido a partir da imagem original em tons de cinza (Figura 32b). Caso contrário, se a imagem de *cluster* em escala de cinza for usada (Figura 32g), a condição necessária da distribuição unimodal não pode ser garantida para todo o conjunto de imagens, reduzindo a precisão do método.

Para transformar imagens P&B semelhantes à Figura 32h em uma representação que faça sentido para a presente aplicação, devemos aplicar a segmentação de Watershed, que pode ser aplicada diretamente sobre as imagens em tons de cinza com diferentes mínimos e bacias hidrográficas. Em vez disso, para imagens P&B, precisamos usar primeiro a transformada da distância para produzir uma representação adequada das imagens P&B para a segmentação de Watershed. Este resultado está apresentado na Figura 33 para duas imagens com aglomerações de células (Figura 33a e Figura 33e). Observe na Figura 33c e Figura 33g o mapa das distâncias correspondentes as imagens da Figura 33b e Figura 33f.

As formas segmentadas na Figura 33d e h são células somáticas individuais, fragmentos de células ou fragmentos do fundo. Por meio do critério de tamanho, apresentado na seção 1.4, estas formas são então classificadas contando-se apenas as células individuais.

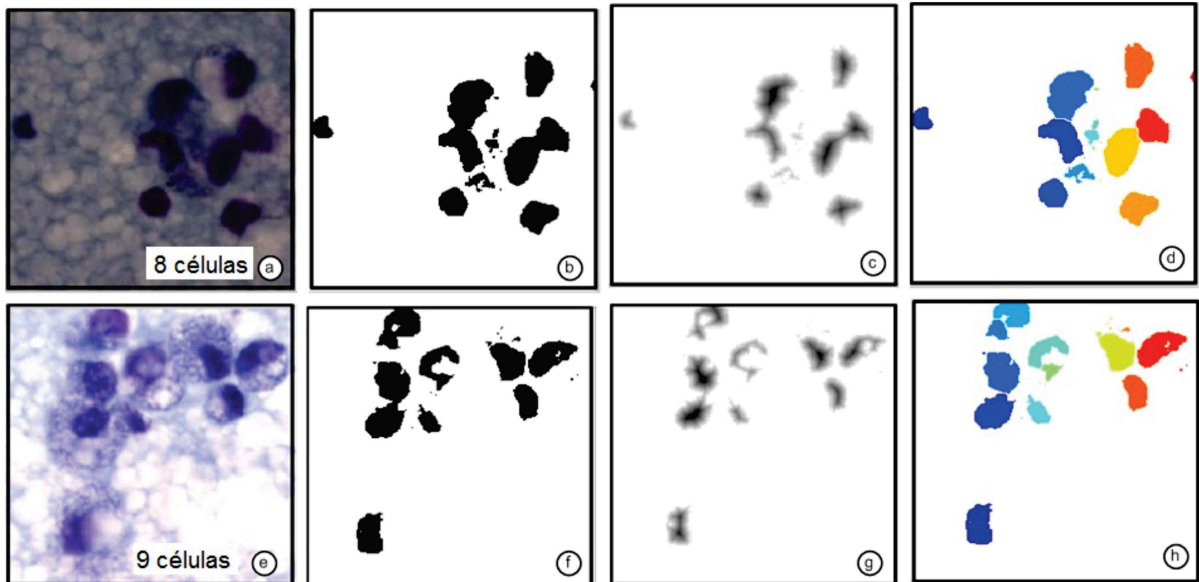


Figura 33 - Exemplos de imagens com células em aglomeração: exemplo 1 e 2 correspondem às imagens (a-b-c-d) e (e-f-g-h), respectivamente.

#### 4.4 Imagem com fundo complexo

A Figura 34 mostra o método proposto aplicado para uma imagem escura, com baixo contraste e, portanto, de difícil segmentação devido a complexidade do fundo. Segundo Zhang et al. (2014) imagens de fundo complexo são aquelas que apresentam fundo com áreas que são visualmente diferentes. Por outro lado, para Jung et al. (2004) a complexidade está relacionada com o baixo contraste, as variações de tamanho, alinhamento, estilo e orientação das formas nas imagens. Para Yang e Yan (2000) imagens complexas são aquelas com intensidade de fundo variável causada por iluminação não uniforme, sombra, mancha ou manchas e contraste muito baixo.

Neste trabalho, as imagens com fundo complexo são aquelas que apresentam principalmente baixo contraste do fundo em relação às células. Além disso, outras características adicionais como a presença de detritos, sujeiras, restos de corantes e variação de iluminação, cor e tonalidade também podem estar presentes nestas imagens.

Nesse caso, operações de manipulação do histograma como *stretching* ou equa-

lização para correção do contraste podem ser úteis. No entanto, devido à alta variabilidade das imagens analisadas, a aplicação dessas técnicas permitiu aprimorar as imagens mais escuras, mas deteriorou as imagens com baixo contraste, tornando-as exageradamente mais claras e inadequadas para as etapas seguintes.

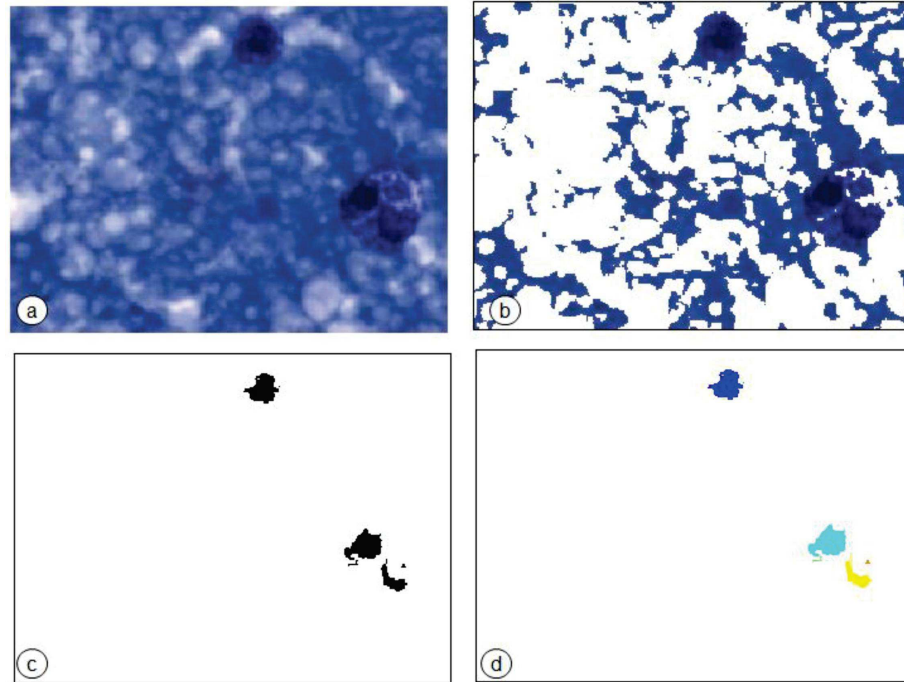


Figura 34 - Imagem com células somáticas do leite bovino com baixo contraste: (a) Imagem original; (b) Imagem original após a clusterização K-means; (c) Resultado da limiarização proposta e (d) Resultado após a Transformada de Watershed.



#### 4.5 Resultados para imagens simples do leite bovino

A Tabela 2 mostra os resultados para a contagem de células somáticas em 400 imagens simples, ou seja, imagens sem detritos e com alto contraste do fundo em relação as células. O esquema proposto na seção 2.3 foi testado através de quatro métodos de limiar: o proposto, Rosin (2001), Otsu (1975) e Melo et al. (2014). A acurácia é medida como uma relação entre a contagem manual e automática como  $\frac{100N_p}{N_e}$ , onde  $N_e$  representa o número total de células somáticas contadas pelo especialista e  $N_p$  o número de células contadas automaticamente.

Tabela 2 - Comparação entre os métodos de limiarização no algoritmo proposto.

Métodos	Células	Correlação	Acurária
Humano	1.321	1,00	100,0
Rosin	2.034	0,71	78,2
Otsu	4.302	0,12	0,0
Melo et al.(2014)	1.318	0,98	99,8
Proposto	1.319	0,98	99,8

#### 4.6 Resultados para imagens reais do leite bovino

A Tabela 3 mostra os resultados para a contagem de células somáticas, usando o algoritmo proposto, para 1000 imagens com variações, ou seja, imagens com fundo complexo, variações de contraste e detritos.

Tabela 3 - Comparação entre os métodos de limiarização no algoritmo proposto.

Métodos	Células	Correlação	Acurária
Humano	2.686	1,00	100,0
Rosin	2.118	0,67	78,9
Otsu	5.650	-0,13	0,0
Melo et al.(2014)	2.431	0,93	90,5
Proposto	2.678	0,98	99,7

Observe que a acurácia do método de Melo et al. (2014) para imagens com detritos (Tabela 3) é menor que para as imagens sem detritos (Tabela 2). Essa variação é justificada, pois Melo et al.(2014) utilizam na Equação 22 o fator de ponderação  $L = 2$ , enquanto o método proposto utiliza o valor  $L = 2, 4$  (ver seção 3.1).

É possível observar a diferença entre o total de 2.686 células contadas pelo perito humano e os obtidos com técnicas automáticas. O algoritmo apresentado na seção 2.3 (Figura 24) com o limiar proposto produz o melhor resultado em termos de acurácia para imagens com variação (Tabela 3). Quando se utiliza o limiar proposto no âmbito do algoritmo proposto, a acurácia e a correlação são mais elevadas do que os métodos apresentados por Wang e Xue (2010) e Na e Heru (2010).

De acordo com Ng (2006), a técnica de limiar de Otsu é muito útil e proporciona resultados satisfatórios ao limiar da imagem com histograma de distribuição bimodal, mas pode falhar se o histograma for unimodal ou perto de unimodal. Isso explica o baixo desempenho do limiar da Otsu no presente trabalho. Embora o limiar de Rosin sozinho tenha alto desempenho de segmentação no histograma unimodal, quando utilizado no âmbito do algoritmo proposto mostrou acurácia inferior ao limiar da abordagem proposta.

#### 4.7 Segmentação de células do leite caprino

Neste experimento, 200 imagens microscópicas com células somáticas do leite caprino foram processadas. A acurácia é medida como uma relação entre a contagem manual e a automática como  $\frac{100N_p}{N_e}$ , onde  $N_e$  representa o número total de células somáticas contadas pelo especialista e  $N_p$  o número de células contadas automaticamente.

O número de células contadas pelo perito humano foi 1.896 e as células contadas pelo método proposto foram 1.892 células com uma acurácia de 99.8%. A correlação de Pearson encontrada foi de 0,99, o erro absoluto médio de 0,02 e erro relativo médio de 2,4. A acurácia obtida é superior ao obtido por Na e Heru (2009), Xue et al. (2009) e Wang e Xue (2010) em aplicações semelhantes para a contagem de células somáticas no leite bovino. O gráfico da distribuição de frequência pode ser verificado na Figura 35.

A Figura 36 mostra duas imagens originais (Figura 36a e Figura 36c). O resultado da conversão destas imagens para CIELAB é mostrado nas Figura 36b e Figura 36d. É possível perceber que na transformação de RGB para CIELAB, todas as células somáticas, independente de sua cor e tonalidade nas imagens originais, são mapeadas para tons de verde nas imagens CIELAB. Isso significa que até mesmo as células azuis e vermelhas são mapeadas para verde. Pode-se observar que o fundo azul e branco (Figura 36a) e o fundo vermelho rosado (Figura 36c) são transformados para fundo vermelho na imagem CIELAB (Figura 36b e Figura 36d). Portanto, quando comparado ao RGB, a representação do



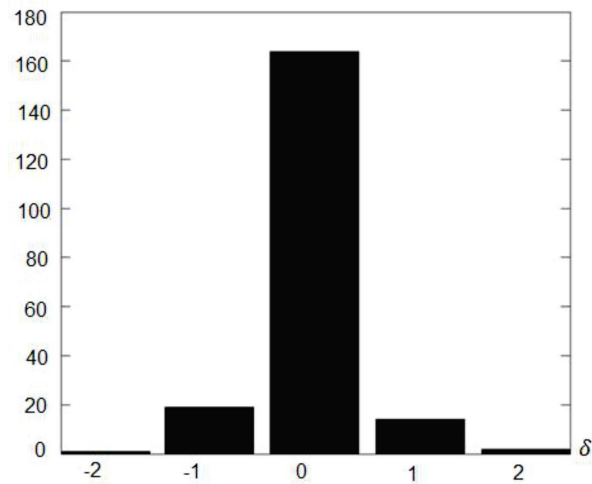


Figura 35 - Distribuição de frequência  $N_p - N_e$ .

fundo e das células somáticas no espaço CIELAB é mais uniforme e propenso a menos variações.

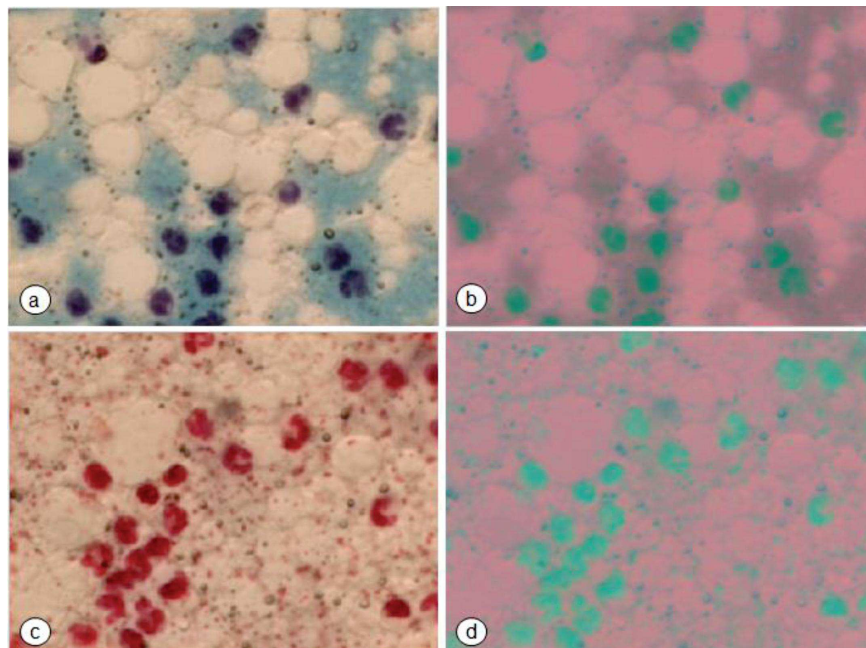


Figura 36 - (a) e (c) Imagem original RGB; (b) e (d) Imagens correspondentes no espaço CIELAB.

De acordo com Sanchez-Cuevas et al. (2013), o modelo RGB é adequado para a detecção e exposição, mas não é uma boa escolha para realizar a segmentação de cor devido ao fato de que ele tem uma elevada redundância. A alta correlação entre as coordenadas em espaços RGB torna inadequada a segmentação de cor.

#### 4.7.1 Extração dos canais $L^*$ , $a^*$ e $b^*$ da imagem CIELAB

A extração em tons de cinza (256 níveis) das coordenadas  $L^*$ ,  $a^*$  e  $b^*$  da imagem CIELAB é apresentada na Figura 37. Note que as células são mais visíveis nas Figuras 37b e f, e o fundo mais homogêneo nas Figuras 37c-g-d e h. Através dessa representação, pode-se observar que a conversão para CIELAB não fornece a segmentação adequada das células somáticas em nenhuma das coordenadas.

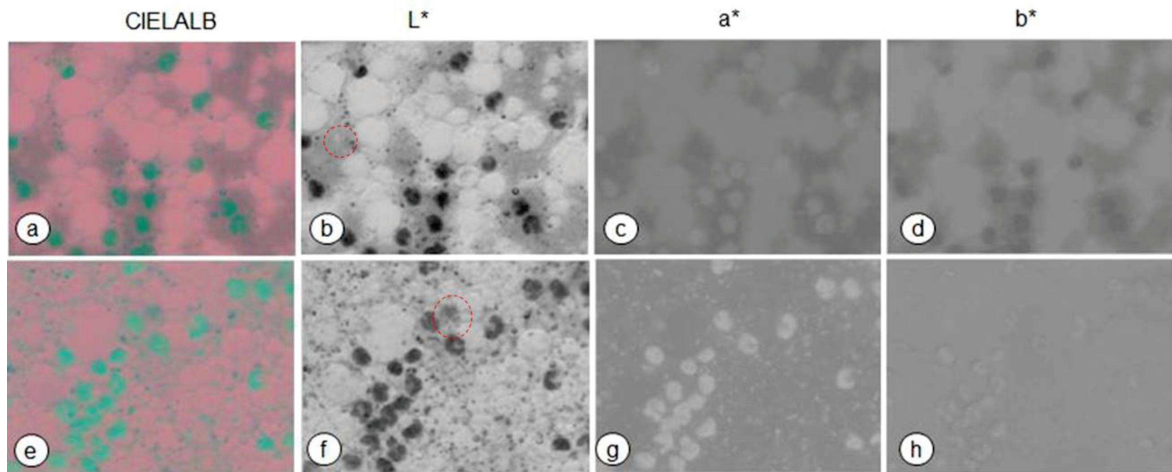


Figura 37 - Extração dos canais  $L^*$ ,  $a^*$  e  $b^*$  na imagem CIELAB.

Alternativamente, técnicas de limiar podem ser utilizadas para as imagens em tons de cinza (Figura 37b, Figura 37c, Figura 37d, Figura 37f, Figura 37g e Figura 37h) para a segmentação das células. No entanto, para alguns grupos de imagens, o padrão estatístico das células se sobrepõe ao do fundo (Figura 37b e Figura 37f).

Observe que o mesmo tom fundo na Figura 37b (sinalizado em vermelho) pode representar o tom de algumas células na Figura 37f. Dessa maneira, um limiar pode ser adequado para uma imagem e não adequado para outras. Assim, a remoção do fundo é uma condição necessária para aplicar a segmentação através de limiarização.

#### 4.7.2 Fundo com tom vermelho

A Figura 38 mostra um exemplo de imagem RGB (Figura 38a) com fundo vermelho e células vermelhas. O mapeamento da imagem RGB para CIELAB é apresentado na Figura 38b, enquanto que o resultado da operação proposta para a segmentação das células somáticas aplicando classificação *pixel* é apresentado na Figura 38c. Essa operação re-

move o fundo da imagem e resulta em células parcialmente segmentadas. A representação P&B da Figura 38c é apresentada na Figura 38d. A cor preta indica a existência de células isoladas, fragmentos de células somáticas, ou aglomeração de células com sobreposição. Nessa fase, a limiarização é realizada considerando cada *pixel* diferente de zero como iguais a zero (preto) e os demais iguais a 255 (branco). A transformada da distância é apresentada na Figura 38e. Pode-se notar que algumas células formam um único componente conectado, devido ao toque ou à superposição. Elas devem ser separadas de uma maneira consistente com a integridade das células ligadas. Para isso, utiliza-se a transformada da distância para produzir uma representação adequada à segmentação de Watershed (Figura 38f).

Uma ampliação referente a área tracejada da Figura 38 é apresentada na Figura 39. Por meio dessa ilustração, é possível notar a sobreposição das células, o mapa das distâncias dos *pixels* (Figura 39e) referente à imagem P&B (Figura 39d), assim como, as linhas de separação das bacias da Transformada de Watershed (Figura 39f).

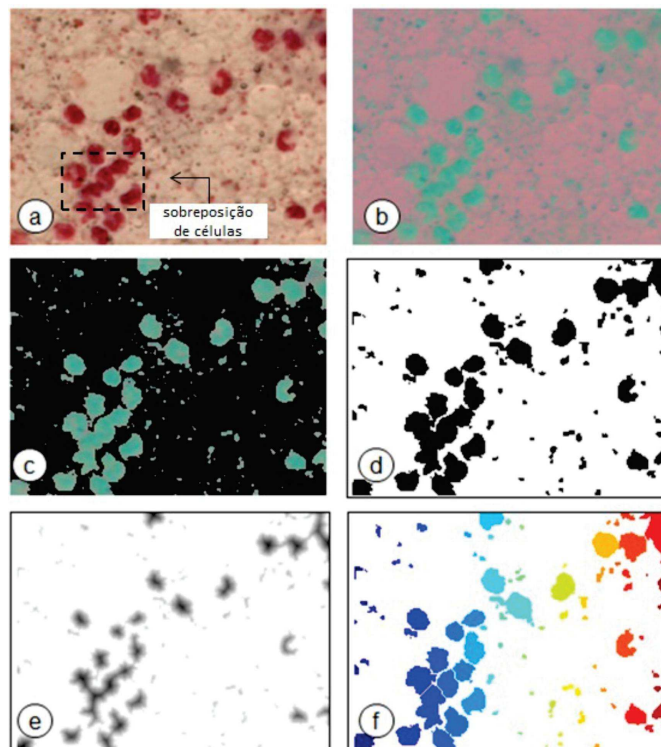


Figura 38 - Método proposto para segmentação de imagens com células somáticas do leite caprino: (a) Imagem original; (b) Imagem no espaço CIELAB; (c) Imagem CIELAB após operação do algoritmo proposto; (d) Imagem após o filtro de partículas; (e) Transformada da Distância e (f) Resultado da segmentação Watershed.

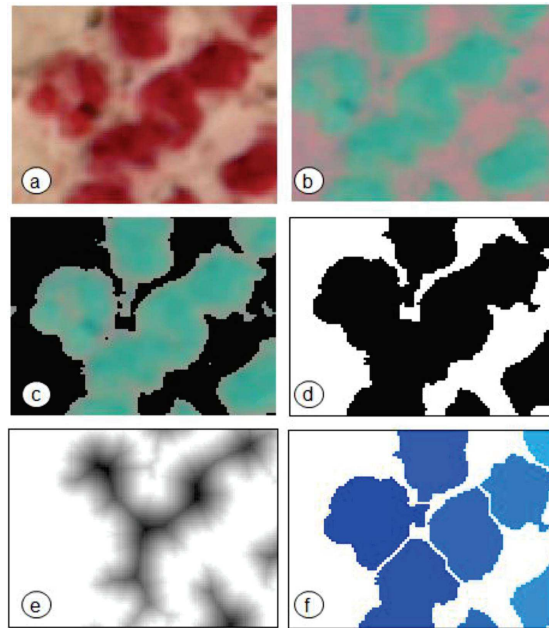


Figura 39 - Ampliação do recorte da aglomeração de células da Figura 38a.

#### 4.7.3 Fundo e células com tons azuis

A Figura 40 mostra uma imagem RGB com tons azuis de fundo e células azuis isoladas.

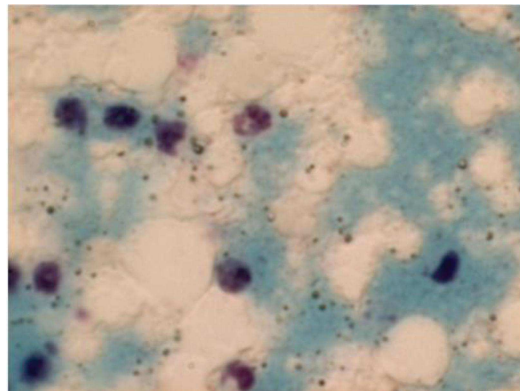


Figura 40 - Imagem RGB com 10 células somáticas do leite caprino.

A conversão para CIELAB da imagem referente a Figura 40 é apresentada na Figura 41a. É possível observar na Figura 41b a remoção do fundo através do algoritmo proposto para classificação de *pixels*. Por outro lado, a Figura 41c apresenta pequenas partículas e buracos na imagem P&B.

A remoção dessas formas é apresentada Figura 41d. A estratégia utilizada para

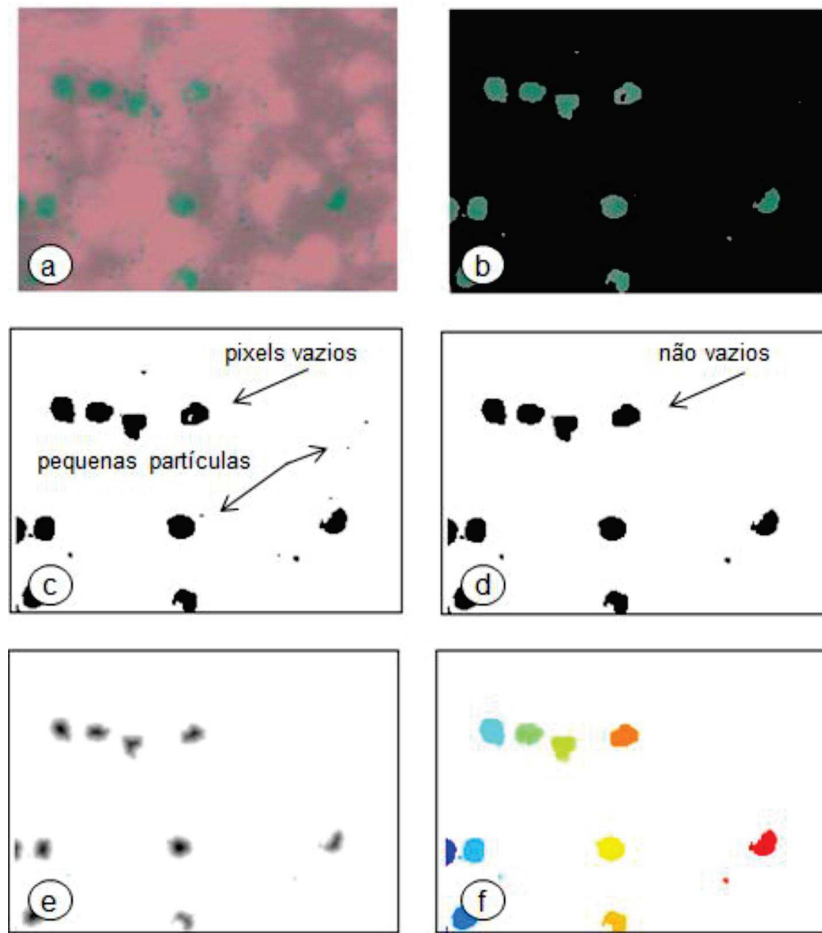


Figura 41 - (a) Imagem no formato CIELAB. (b) Imagem após operação proposta. (c) Imagem P&B. (d) Imagem P&B após o filtro de partículas. (e) Transformada da distância e (f) Resultado da segmentação Watershed.

remover “furos” nas formas, recupera os *pixels* que pertencem às células somáticas (Figura 41c) e, eventualmente, não foram classificados na etapa do algoritmo de classificação de *pixels*.

#### 4.8 Imagem com fundo complexo

Para demonstrar a robustez do método proposto é apresentada na Figura 42 uma imagem com fundo complexo, de segmentação difícil e que poderia elevar a contagem do número de falso-positivos. É possível observar na Figura 42a a baixa luminosidade da imagem, a presença de detritos e tonalidade do fundo e das células variando entre os tons azuis e vermelhos. Note também a menor intensidade do brilho das células em relação às imagens que foram apresentadas nas Figura 36 e Figura 40. A ampliação referente à



área tracejada da Figura 42a é apresentada na Figura 43. Por meio dessa ampliação, é possível observar a eficiência do método proposto para remoção do fundo, extração de detritos (Figura 42b-c) e isolamento das células (Figura 42d) em imagens complexas.

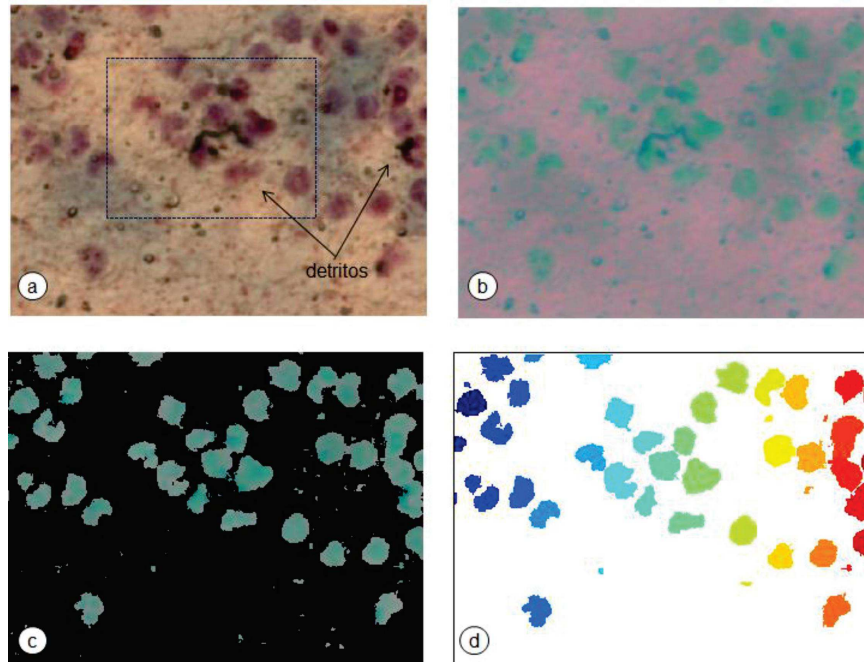


Figura 42 - Imagem com fundo complexo: (a) Imagem original; (b) Imagem no formato CIELAB; (c) Imagem CIELAB após operação proposta para classificação de *pixels* e (d) Resultado da segmentação Watershed.

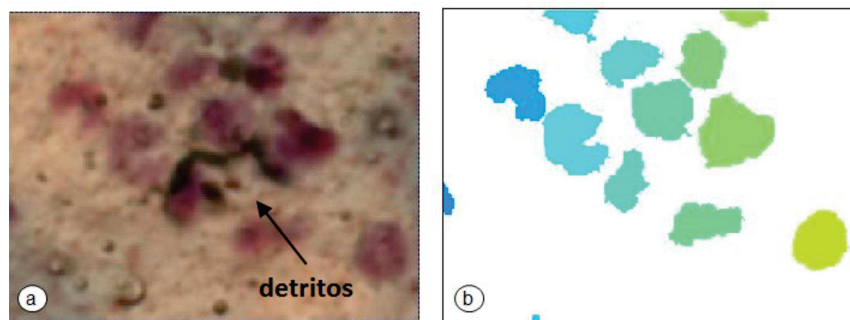


Figura 43 - Imagem com fundo complexo: (a) Ampliação referente a Figura 42a e (b) Resultado após a Transformada de Watershed.

## 5 CONCLUSÃO

Este trabalho teve como objetivo principal desenvolver uma abordagem para auxílio da contagem de células somáticas no leite bovino e caprino. Uma grande quantidade de imagens de células, capturadas a partir de diferentes processos de preparação da lâmina, foi usado para validar os resultados experimentais. O esquema proposto para segmentação de células somáticas do leite bovino baseia-se em técnicas de clusterização e limiarização. Por sua vez, o esquema o proposto para o leite de cabra baseia-se no espaço de cor CIE-LAB.

A revisão de literatura efetuada revelou que os trabalhos de Baro et al. (2005), Xue et al. (2008), Na e Heru (2009), Xue et al. (2009), Na e Heru (2010) e Wang e Xue (2010) não discutem os problemas dos detritos, variação de contrastes, cor e tonalidade das imagens. Como apresentado nos resultados deste trabalho, tais características são comuns em imagens microscópicas do leite e são inerentes ao processo de preparação das lâminas, captura das imagens e manejo do leite.

Vale, no entanto, ressaltar que não foram encontrados trabalhos que referenciem a utilização de técnicas de PDI para CCS do leite caprino, portanto, este é um tema novo e, também, de grande relevância teórica e prática.

Observa-se que a proposição do novo limiar para histogramas unimodais, produz melhores resultados do que os métodos tradicionais de limiarização de Otsu e Rosin. É possível notar que a escolha do limiar é a etapa fundamental da metodologia adotada para bovinos. Isso porque a eficiência dos outros algoritmos nas etapas seguintes, como o algoritmo da transformada de Watershed, está condicionada ao sucesso na etapa de limiarização.

É importante ressaltar, também, que o método proposto para limiarização em histogramas unimodais difere dos métodos tradicionais conhecidos na literatura, e isso porque os métodos tradicionais encontram o limiar do histograma utilizando informações individuais da distribuição de frequência de cada imagem. Por outro lado, o algoritmo proposto define um fator de maximização que é ajustado utilizando todas as imagens de um grupo. Assim, a depender da aplicação escolhida, esse fator pode ser melhorado e ajustado automaticamente, e seu resultado é muito importante porque, diferente dos demais métodos apresentados, a limiarização unimodal proposta é adaptativa e tem possibilidades de ser

aplicada para diversos grupos de imagens.

Outro ponto importante a ser considerado é o fato de que a utilização do espaço CIELAB permitiu uniformizar as imagens de bovinos e caprinos por meio de uma representação que possibilita controlar as variações da cor e tonalidade comuns nos espaços RGB. Portanto, a utilização das estratégias apresentadas permitem que todas as imagens (de bovinos e caprinos) sejam processadas, sem a necessidade prévia de algoritmos comumente utilizados para correção de contraste, suavização e filtragem. Essas características agregam robustez aos métodos e permitem o processamento de imagens de difícil interpretação.

Além disso, verificou-se que o espaço CIELAB, juntamente com o algoritmo K-means, contornam os problemas das imagens de bovinos relacionados com a variação de contrastes e presença de detritos. Como apresentado nos resultados, se os detritos não forem retirados previamente das imagens, o número de células somáticas pode ser aumentado incorretamente.

A robustez dos métodos foi avaliada, comprovando-se a capacidade em processar grupos de imagens com presença de detritos, variações de tonalidade e contraste. Assim, os algoritmos propostos são uma solução promissora para automatizar sistemas de detecção de mastite por meio da microscopia óptica e com possibilidade de uso em outras aplicações semelhantes.

## 5.1 Sugestões para Trabalhos Futuros

A próxima etapa deste trabalho consiste em criar um *software* que se comunique com o microscópio, permitindo que o usuário obtenha rapidamente o processamento das imagens capturadas pelo técnico do laboratório. Com isso, pretende-se assistir diversos laboratórios que, por muitos motivos, necessitam aperfeiçoar a CCS feita no microscópio.

Outra etapa do trabalho consiste em elaborar um esquema de automação do microscópio para que a captura das imagens seja feita sem sobreposição e de forma automática. Em seguida, deve-se estabelecer o embasamento experimental e estatístico para validação do método, junto ao Ministério da Agricultura, Pecuária e Abastecimento.

Além dessas linhas de trabalho, é importante definir outras que poderão ser abordadas futuramente por diferentes pesquisadores. Entre elas, verificar a eficiência dos algoritmos propostos para imagens capturadas no microscópio com aumento inferior a



1000×. Outra possibilidade é propor ajustes no algoritmo K-means, para que ele seja utilizado em conjunto com espaço CIELAB para segmentação de células de caprinos.

Sugere-se, ainda, que novas pesquisas sejam desenvolvidas com o propósito de adequar o algoritmo K-means para segmentação de células do leite caprino. Por fim, será interessante que os algoritmos de limiarização e segmentação da cor, propostos neste trabalho, sejam testados em novas simulações com grupos e quantidades diferentes de imagens. Assim, será possível verificar as potencialidades e limitações dos métodos propostos em relações a outras técnicas tradicionais da literatura.

## LISTA DE ARTIGOS PUBLICADOS

**Artigo aceito para publicação** (ver Apêndice)

1. *A robust segmentation method for counting bovine milk somatic cells in microscope slide images*. ELSEVIER - **Computers and Electronics in Agriculture**, ISSN: 168 – 1699, (A1 - ENG III).

**Artigos publicados em periódicos** (ver Apêndice)

2. *Segmentation of Somatic Cells in Goat Milk Using Color Space CIELAB* - (doi: 10.17265/2161-6256/2014.10.009). **Journal of Agricultural Science and Technology**, ISSN 939 – 1250, (B1 - EGN III)

3. *Método para contagem de células somáticas do leite bovino para imagens capturadas do microscópio óptico*. **Zootecnia Tropical**, ISSN: 0798 – 7269, (B2 - EGN III).

4. *Método de limiarização automática para a contagem de células somáticas em imagens microscópicas* - (d.o.i: 10.7198/S2237 – 0722201400030021).

**GEINTEC - Gestão, Inovação e Tecnologias**, ISSN: 2237 – 0722, (B3 - EGN III)

## Referências

- ABRÀMOFF, M. D.; MAGALHÃES, P. J.; RAM, S. J. Image processing with ImageJ. **Biophotonics International**, v. 11, n. 7, p. 36-43, 2004.
- ALBA-FLORES, R. Evaluation of the use of high-resolution satellite imagery in transportation applications. 2005.
- ALLALOU, A; WÄHLBY, C. BlobFinder, a tool for fluorescence microscopy image cytometry. **Computer Methods and Programs in Biomedicine**, v. 94, n. 1, p. 58-65, 2009.
- ALVES DE MELO, G. J. et al. Método de limiarização automática para a contagem de células somáticas em imagens microscópicas. **GEINTEC- Gestão, Inovação e Tecnologias**, v. 4, n. 3, p. 1283-1291, 2014.
- ANDERSON, D. E. et al. Enfermidades da glândula mamária. **PUGH, DG Clínica de Ovinos e Caprinos. Editora Roca: São Paulo**, 2005.
- ANDRADE, PV D. et al. Contagem de células somáticas em leite de cabra. **Arq. Bras. Med. Vet. Zootec**, v. 53, n. 3, p. 396-400, 2001.
- ANGULO, J.; JEULIN, D.. Stochastic watershed segmentation. In: **Proc. of the 8th International Symposium on Mathematical Morphology**. 2007. p. 265-276.
- ANKINAKATTE, S., et al. Predicting mastitis in dairy cows using neural networks and generalized additive models: A comparison. **Computers and Electronics in Agriculture**, v. 99, p. 1-6, 2013.
- ARSLAN, S.; OZYUREK, E.; GUNDUZ-DEMIR, C. A color and shape based algorithm for segmentation of white blood cells in peripheral blood and bone marrow images. **Cytometry Part A**, v. 85, n. 6, p. 480-490, 2014.
- BAR, D. et al. The cost of generic clinical mastitis in dairy cows as estimated by using dynamic programming. **Journal of Dairy Science**, v. 91, n. 6, p. 2205-2214, 2008.
- BARADEZ, M.-O. et al. Robust and automated unimodal histogram thresholding and potential applications. **Pattern Recognition**, v. 37, n. 6, p. 1131-1148, 2004.
- BARO, J. A., et al. Video microscopy as an alternative method for somatic cell count in milk. **Journal of Dairy Research**, v. 72, n. 01, p. 93-100, 2005.
- BRASIL. Ministério da Agricultura, Pecuária e Abastecimento. In: Embrapa Gado de Leite. *Estatísticas do leite*: Disponível em: <<http://www.cnp.gl.embrapa.br/producao/dados2002/producao/tabela0210.php>>, acesso em: dezembro de 2005.
- BECK, H. S.; WISE, W. S.; DODD, F. H. Cost benefit analysis of bovine mastitis in the UK. **Journal of Dairy Research**, v. 59, n. 04, p. 449-460, 1992.

- BENGTSSON, E.; WAHLBY, C.; LINDBLAD, J. Robust cell image segmentation methods. **Pattern Recognition and Image Analysis C/C Of Raspoznavaniye Obrazov I Analiz Izobrazhenii.**, v. 14, n. 2, p. 157-167, 2004.
- BERRY, E.; BROUGHAN, J. Use of the DeLaval cell counter (DCC) on goats' milk. **Journal of Dairy Research**, v. 74, n. 03, p. 345-348, 2007.
- BEUCHER, S., et al. The watershed transformation applied to image segmentation. **Scanning Microscopy-Supplement**, p. 299-299, 1992.
- BHANU, B.; FAUGERAS, O. D. Segmentation of images having unimodal distributions. **Pattern Analysis and Machine Intelligence, IEEE Transactions on**, n. 4, p. 408-419, 1982.
- BHATTACHARYA, P. Connected component labeling for binary images on a reconfigurable mesh architecture. **Journal of Systems Architecture**, v. 42, n. 4, p. 309-313, 1996.
- BHIKANE, A. V.; KAWITKAR, S. B. **Hand Book for Veterinary Clinician**. 2000.
- BLOWEY, R. W.; EDMONDSON, Peter. **Mastitis Control in Dairy Herds**. Cabi, 2010.
- BRASIL, 2002. Ministério da agricultura, pecuária e abastecimento. Instrução Normativa nº 51 de 18 de setembro de 2002. **Diário Oficial da União**, 20 de Setembro de 2002, seção 1, página 13.
- Brito, J. R. F., 2015. **Células somáticas no leite**. Disponível em: <[http://www.cnpqgl.embrapa.br/totem/conteúdo/Qualidade de leite e mastite/Outraspublicacoes/Celulassomaticas no leite.pdf](http://www.cnpqgl.embrapa.br/totem/conteúdo/Qualidade%20de%20leite%20e%20mastite/Outraspublicacoes/Celulassomaticas%20no%20leite.pdf)>. Acesso em: 25 mar 2015.
- BRUNO, O. M., et al. Fractal dimension applied to plant identification. **Information Sciences**, v. 178, n. 12, p. 2722-2733, 2008.
- BRYS, L. M. **Página dinâmica para aprendizado do sensoriamento remoto**. 2008. Tese de Doutorado. Universidade Federal do Rio Grande do Sul.
- BUXTON, B. F., et al. Development of an extension of the Otsu algorithm for multidimensional image segmentation of thin-film blood slides. In: **Computing: Theory and Applications, 2007. ICCTA'07. International Conference on**. IEEE, 2007. p. 552-562.
- CARVALHO, P. C. D. F., 2010. **Country pasture/forage resource profiles**. Disponível em: <<http://www.test.org/doi/>>. Acesso em: 28 de fev. 2015.
- CASTAÑON, C. A. B. **Análise e reconhecimento digital de formas biológicas para o diagnóstico automático de parasitas do gênero Eimeria**. 2006. Tese de Doutorado. Universidade de São Paulo.
- CAVERO, D., et al. Mastitis detection in dairy cows by application of fuzzy logic. **Livestock Science**, v. 105, n. 1, p. 207-213, 2006.
- CAVERO, D., et al. Mastitis detection in dairy cows by application of neural networks. **Livestock Science**, v. 114, n. 2, p. 280-286, 2008.

- CAVERO, D., et al. Analysing serial data for mastitis detection by means of local regression. **Livestock Science**, v. 110, n. 1, p. 101-110, 2007.
- CHAGUNDA, M. G., et al. L-lactate dehydrogenase and N-acetyl- $\beta$ -D-glucosaminidase activities in bovine milk as indicators of non-specific mastitis. **Journal of Dairy Research**, v. 73, n. 04, p. 431-440, 2006.
- CHANG, F.; CHEN, C.; LU, C. A linear-time component-labeling algorithm using contour tracing technique. **Computer Vision and Image Understanding**, v. 93, n. 2, p. 206-220, 2004.
- CHANG, J., et al. New automatic multi-level thresholding technique for segmentation of thermal images. **Image and Vision Computing**, v. 15, n. 1, p. 23-34, 1997.
- CHEN, Q., et al. Isolating individual trees in a savanna woodland using small footprint lidar data. **Photogrammetric Engineering & Remote Sensing**, v. 72, n. 8, p. 923-932, 2006.
- CHEN, Q.; YANG, X.; PETRIU, E. M. Watershed segmentation for binary images with different distance transforms. In: **Haptic, Audio and Visual Environments and Their Applications, 2004. HAVE 2004. Proceedings. The 3rd IEEE International Workshop**. 2004. p. 111-116.
- CHEN, X.; ZHENG, B.; LIU, H. Optical and digital microscopic imaging techniques and applications in pathology. **Analytical Cellular Pathology**, v. 34, n. 1, p. 5-18, 2011.
- CHENG, Heng-Da et al. Color image segmentation: advances and prospects. **Pattern recognition**, v. 34, n. 12, p. 2259-2281, 2001.
- CHI, Z.; YAN, H.; PHAM, T. **Fuzzy algorithms: with applications to image processing and pattern recognition**. World Scientific, 1996.
- CONTRERAS, A., et al. The role of intramammary pathogens in dairy goats. **Livestock Production Science**, v. 79, n. 2, p. 273-283, 2003.
- CORKIDI, G., et al. COVASIAM: an image analysis method that allows detection of confluent microbial colonies and colonies of various sizes for automated counting. **Applied and Environmental Microbiology**, v. 64, n. 4, p. 1400-1404, 1998.
- COSTA, E. O. et al. Índices de mastite bovina clínica e subclínica nos Estados de São Paulo e Minas Gerais. **Revista Brasileira de Medicina Veterinária**, v. 17, n. 5, p. 215-217, 1995.
- COUDRAY, N.; BUESSLER, J.; URBAN, J. A robust thresholding algorithm for unimodal image histograms. **Pattern Recognition Letters**, n. 31, p. 1010-1019, 2010.
- DAIRY, U. S. D. A. Heifer calf health and management practices on US dairy operations, 2007. **Fort Collins (CO): USDA: APHIS: VS, CEAH**, v. 550, 2010.
- DATE, M. K.; AKARTE, Mr SP. Brain image segmentation algorithm using K-means clustering. **International Journal Of Computer Science And Applications**, v. 6, n. 2, 2013.
- DAVIS, J. G. The rapid abnormality indicator. **Dairy Ind**, v. 12, p. 35, 1947.

ZAMPIROLI, F. A. Transformada de distância por morfologia matemática. **Doutor, Universidade Estadual de Campinas, Faculdade de Engenharia Elétrica e de Computação, Campinas, SP, Brasil, 2003.**

DE-LU, W. H. P.; DE-SHEN, X. I. A. State Key Laboratory of Satellite Ocean Environment Dynamics, Second Institute of Oceanography, State Oceanic Administration, Hangzhou 310012 2. Department of Computer, Nanjing University of Science and Technology, Nanjing 210094; A Fast Algorithm for Two-dimensional Otsu Adaptive Threshold Algorithm [J]. **Acta Automatica Sinica**, v. 9, 2007.

DE MELO, G. J. A., Gomes, V., B., C. C., de Almeida, L. A. L., de Castro Lima, A. C., 2014. Segmentation of somatic cells in goat milk using color space cielab. **Journal of Agricultural Science and Technology A**, n. 4 , 865–873.

DE MOL, R. M.; WOLDT, W. E. Application of fuzzy logic in automated cow status monitoring. **Journal of Dairy Science**, v. 84, n. 2, p. 400-410, 2001.

DEONI, S. C., et al. Segmentation of thalamic nuclei using a modified k-means clustering algorithm and high-resolution quantitative magnetic resonance imaging at 1.5 T. **Neuroimage**, v. 34, n. 1, p. 117-126, 2007.

DIGABEL, H.; LANTUÉJOUL, C. Iterative algorithms. In: **Proc. 2nd European Symp. Quantitative Analysis of Microstructures in Material Science, Biology and Medicine**. Stuttgart, West Germany: Riederer Verlag, 1978. p. 8.

DORINI, L. B.; MINETTO, R.; LEITE, N. J. Semiautomatic white blood cell segmentation based on multiscale analysis. **Biomedical and Health Informatics, IEEE Journal of**, v. 17, n. 1, p. 250-256, 2013.

DUAN, L., et al. Fast discrimination and counting of filled/unfilled rice spikelets based on bi-modal imaging. **Computers and Electronics in Agriculture**, v. 75, n. 1, p. 196-203, 2011.

DUBEY, S. R., et al. Infected fruit part detection using K-means clustering segmentation technique. **IJIMAI**, v. 2, n. 2, p. 65-72, 2013.

EGGERS, H. Two fast Euclidean distance transformations in  $Z^2$  based on sufficient propagation. **Computer Vision and Image Understanding**, v. 69, n. 1, p. 106-116, 1998.

EMBRAPA, 2011. **Empresa Brasileira de Pesquisa Agropecuária**. Produção de leite.

ENSAFI, P. Weighted Opposition-Based Fuzzy Thresholding. 2011.

ERIKSSON, A., et al. Detection of mastitic milk using a gas-sensor array system (electronic nose). **International Dairy Journal**, v. 15, n. 12, p. 1193-1201, 2005.

FABBRI, R., et al. 2D Euclidean distance transform algorithms: A comparative survey. **ACM Computing Surveys (CSUR)**, v. 40, n. 1, p. 2, 2008.

FAN, J.; LEI, B. A modified valley-emphasis method for automatic thresholding. **Pattern Recognition Letters**, v. 33, n. 6, p. 703-708, 2012.

- FARIA, D. R. **Reconhecimento de impressões digitais com baixo custo computacional para um sistema de controle de acesso**. 2005. Tese de Doutorado. Universidade Federal do Paraná.
- FERRERO, F. J.; VALLEDOR, M.; CAMPO, J. C. Screening method for early detection of mastitis in cows. **Measurement**, v. 47, p. 855-860, 2014.
- FIGUEIREDO, M. F. **Um modelo perceptivo de limiarização de imagens digitais**. 2003. Tese de Doutorado. Universidade Estadual de Maringá.
- FRANCE, I. et al. A comparative study of approaches to automatic pollen identification. In: **BMVC**. 1997.
- GALIERO, G., et al. The meaning of the somatic cell count in buffalo milk. **Bubalus bubalis**, v. 6, n. 4, p. 26-27, 2000.
- GARCIA-ARMESTO, M. R., et al. Modern microbiological methods for foods: colony count and direct count methods. a review. **Microbiologia**, v. 9, n. 1, p. 1-13, 1993.
- GARCIA-CORDERO, Jose L. et al. Microfluidic sedimentation cytometer for milk quality and bovine mastitis monitoring. **Biomedical Microdevices**, v. 12, n. 6, p. 1051-1059, 2010.
- GARGOURI, F., et al. Automatic localization methodology dedicated to brain tumors based on ICP matching by using axial MRI symmetry. In: **Advanced Technologies for Signal and Image Processing (ATSIP), 2014 1st International Conference on**. IEEE, 2014. p. 209-213.
- GEORGE, Y. M., et al. Automated cell nuclei segmentation for breast fine needle aspiration cytology. **Signal Processing**, v. 93, n. 10, p. 2804-2816, 2013.
- GOMES, O. F. M. **Processamento e Análise de Imagens Aplicados à Caracterização Automática de Materiais**. 2001. Tese de Doutorado. Pontifícia Universidade Católica do Rio de Janeiro Rio de Janeiro.
- GOMES, V., et al. Contagem automática e microscópica direta das células somáticas do leite de ovelhas da raça lacaune, utilizando como corantes o rosenfeld e verde de metil e pironina-y. **Ciência Animal Brasileira**, v. 11, n. 1, p. 162-167, 2010.
- GOMES, V., et al. Effect of the stage of lactation on somatic cell counts in healthy goats (*Caprae hircus*) breed in Brazil. **Small Ruminant Research**, v. 64, n. 1, p. 30-34, 2006.
- GOMES, V., et al. Etiologia e fatores de risco para a infecção mamária de cabras leiteiras em seis propriedades do Estado de São Paulo. **Semina: Ciências Agrárias**, v. 35, n. 5, p. 2551-2562, 2014.
- GONZALES, R. C.; WOODS, R. E. **Digital image processing**, 1993
- GONZALES-BARRON, U.; BUTLER, F. A comparison of seven thresholding techniques with the k-means clustering algorithm for measurement of bread-crumbs features by digital image analysis. **Journal of Food Engineering**, v. 74, n. 2, p. 268-278, 2006.
- GONZALEZ, R. C.; WOODS, R. E. **Processamento de imagens digitais**. Edgard Blucher, 2000.

- GONZALEZ, R. C.; WOODS, R. E.; EDDINS, S. L. **Digital image processing using MATLAB**. Pearson Education India, 2004.
- GONZALO, C., et al. Evaluation of rapid somatic cell counters under different analytical conditions in ovine milk. **Journal of Dairy Science**, v. 87, n. 11, p. 3623-3628, 2004.
- GRAU, V., et al. Improved watershed transform for medical image segmentation using prior information. **Medical Imaging, IEEE Transactions on**, v. 23, n. 4, p. 447-458, 2004.
- GUIMARÃES, L. V. Segmentation and Decomposition Techniques Applied to Blood Cell Images. **Doctor Thesis**, 2002.
- HANSSON, H.; SZCZENZA-RUNDBERG, M.; NIELSEN, C. Which preventive measures against mastitis can increase the technical efficiency of dairy farms. **Animal**, v. 5, n. 04, p. 632-640, 2011.
- HARALOCK, R. M.; SHAPIRO, L. G. **Computer and Robot Vision**. Addison-Wesley Longman Publishing Co., Inc., 1991.
- HARMON, R. J. Physiology of mastitis and factors affecting somatic cell counts. **Journal of Dairy Science**, v. 77, n. 7, p. 2103-2112, 1994.
- HARTIGAN, John A.; WONG, Manchek A. Algorithm AS 136: A k-means clustering algorithm. **Applied Statistics**, p. 100-108, 1979.
- HERNANDEZ-GOMEZ, G., et al. Natural image segmentation using the CIELab space. In: **Electrical, Communications, and Computers, 2009. CONIELECOMP 2009. International Conference on**. IEEE, 2009. p. 107-110.
- HILLERTON, J. E.; BERRY, E. A. Quality of the milk supply: European regulations versus practice. In: **NMC Annual Meeting Proceedings**. 2004. p. 207-214.
- HILLERTON, J. E.; SEMMENS, J. E. Comparison of treatment of mastitis by oxytocin or antibiotics following detection according to changes in milk electrical conductivity prior to visible signs. **Journal of Dairy Science**, v. 82, n. 1, p. 93-98, 1999.
- HOGAN, J.; SMITH, K. L.. Coliform mastitis. **Veterinary research**, v. 34, n. 5, p. 507-519, 2003.
- HOGVEEN, H., et al. Sensors and clinical mastitis—The quest for the perfect alert. **Sensors**, v. 10, n. 9, p. 7991-8009, 2010.
- HU, Q.; QIAN, G.; NOWINSKI, W. L. Fast connected-component labelling in three-dimensional binary images based on iterative recursion. **Computer Vision and Image Understanding**, v. 99, n. 3, p. 414-434, 2005.
- JANG, J. et al. Global thresholding algorithm based on boundary selection. In: **Control, Automation and Systems (ICCAS), 2013 13th International Conference on**. IEEE, 2013. p. 704-706.
- JANZEKOVIC, M. et al. Mastitis detection based on electric conductivity of milk. **Journal of**



**Achievements in Materials and Manufacturing Engineering**, v. 34, n. 1, p. 39-46, 2009.

JOHANNSEN, G.; BILLE, J. A threshold selection method using information measures. In: **ICPR**. 1982. p. 140-143.

KAMPHUIS, C. et al. Decision-tree induction to detect clinical mastitis with automatic milking. **Computers and Electronics in Agriculture**, v. 70, n. 1, p. 60-68, 2010.

KAPUR, J. N.; SAHOO, P. K.; WONG, A. KC. A new method for gray-level picture thresholding using the entropy of the histogram. **Computer Vision, graphics, and image processing**, v. 29, n. 3, p. 273-285, 1985.

KARVELIS, P.; LIKAS, A.; FOTIADIS, D. I. Identifying touching and overlapping chromosomes using the watershed transform and gradient paths. **Pattern Recognition Letters**, v. 31, n. 16, p. 2474-2488, 2010.

KIM, D.; LEE, K. H.; LEE, D. A novel initialization scheme for the fuzzy c-means algorithm for color clustering. **Pattern Recognition Letters**, v. 25, n. 2, p. 227-237, 2004.

KONG, H.; GURCAN, M.; BELKACEM-BOUSSAID, K. Partitioning histopathological images: an integrated framework for supervised color-texture segmentation and cell splitting. **Medical Imaging, IEEE Transactions on**, v. 30, n. 9, p. 1661-1677, 2011.

KOOP, G. Udder health in dairy goats. 2012.

KOYA, T. K. **Sistema Automático de Contagem de Células Sangüíneas**. 2001. Tese de Doutorado. Dissertação de Mestrado. Departamento de Engenharia Elétrica da Escola de Engenharia da UFRGS, Porto Alegre.

KRAMER, E., et al. Mastitis and lameness detection in dairy cows by application of fuzzy logic. **Livestock Science**, v. 125, n. 1, p. 92-96, 2009.

JUNG, K.; KIM, K. I.; JAIN, A. K. Text information extraction in images and video: a survey. **Pattern Recognition**, v. 37, n. 5, p. 977-997, 2004.

LAGONI, H. Estudo Sobre Contagem de Células Somáticas (CCS) no Brasil—Uma Revisão. **MESQUITA, AJ; DÜRR, JW; COELHO, KO Perspectivas e Avanços da Qualidade do Leite no Brasil**. Goiânia: Talento, p. 199-208, 2006.

LAM, V. **Milk production on smallholder dairy cattle farms in Southern Vietnam**. 2011.

LANGFORD, M.; TAYLOR, G. E.; FLENLEY, J. R. Computerized identification of pollen grains by texture analysis. **Review of Palaeobotany and Palynology**, v. 64, n. 1, p. 197-203, 1990.

LEE, S. U.; CHUNG, S. Y.; PARK, R. H. A comparative performance study of several global thresholding techniques for segmentation. **Computer Vision, Graphics, and Image Processing**, v. 52, n. 2, p. 171-190, 1990.

LEITNER, G., et al. Interactions between bacteria type, proteolysis of casein and physico-chemical properties of bovine milk. **International Dairy Journal**, v. 16, n. 6, p. 648-654, 2006.

- LI, C. H.; LEE, C. K. Minimum cross entropy thresholding. **Pattern Recognition**, v. 26, n. 4, p. 617-625, 1993.
- LI, C. H.; YUEN, P. C. Regularized color clustering in medical image database. **Medical Imaging, IEEE Transactions on**, v. 19, n. 11, p. 1150-1155, 2000.
- LIU, C., et al. A pectoral muscle segmentation algorithm for digital mammograms using Otsu thresholding and multiple regression analysis. **Computers & Mathematics with Applications**, v. 64, n. 5, p. 1100-1107, 2012.
- LUCCHESSEY, L.; MITRAY, S. K. 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.
- LUMIA, R.; SHAPIRO, L.; ZUNIGA, O. A new connected components algorithm for virtual memory computers. **Computer Vision, Graphics, and Image Processing**, v. 22, n. 2, p. 287-300, 1983.
- MADUREIRA, K. M., et al. Análise das metodologias diretas e indiretas para a contagem de células somáticas no leite de cabras híbridas. **Pesquisa Veterinária Brasileira**, v. 30, n. 4, p. 311-316, 2010.
- MAGALHÃES, H. R., et al. Influência de fatores de ambiente sobre a contagem de células somáticas e sua relação com perdas na produção de leite de vacas da raça Holandesa. **Revista Brasileira de Zootecnia**, v. 35, n. 2, p. 415-421, 2006.
- MAHAPATRA, P. K., et al. Improved thresholding based on negative selection algorithm (NSA). **Evolutionary Intelligence**, v. 6, n. 3, p. 157-170, 2014.
- MALPICA, N., et al. Applying watershed algorithms to the segmentation of clustered nuclei. 1997.
- MANSOR, R. **Proteomic and metabolomic studies on milk during bovine mastitis**. 2012. Tese de Doutorado. University of Glasgow.
- MARTÍ-DE OLIVES, A., et al. Short communication: Effect of subclinical mastitis on proteolysis in ovine milk. **Journal of Dairy Science**, v. 94, n. 11, p. 5369-5374, 2011.
- MEAS-YEDID, V.; TILIE, S.; OLIVO-MARIN, J. C. Color image segmentation based on markov random field clustering for histological image analysis. In: **Pattern Recognition, 2002. Proceedings. 16th International Conference on. IEEE**, 2002. p. 796-799.
- MEDINA-CARNICER, R.; MADRID-CUEVAS, F. J. Unimodal thresholding for edge detection. **Pattern Recognition**, v. 41, n. 7, p. 2337-2346, 2008.
- MEDINA-CARNICER, R., et al. A novel histogram transformation to improve the performance of thresholding methods in edge detection. **Pattern Recognition Letters**, v. 32, n. 5, p. 676-693, 2011.
- MEDRANO-GALARZA, C., et al. Behavioral changes in dairy cows with mastitis. **Journal of**

**Dairy Science**, v. 95, n. 12, p. 6994-7002, 2012.

MELE, M., et al. Application of the 'tracking signal' method to the monitoring of udder health and oestrus in dairy cows. **Livestock Production Science**, v. 72, n. 3, p. 279-284, 2001.

MELO, G. A. Jesus., et al. MÉTODO DE LIMIARIZAÇÃO AUTOMÁTICA PARA A CONTAGEM DE CÉLULAS SOMÁTICAS EM IMAGENS MICROSCÓPICAS. **GEINTEC-Gestão, Inovação e Tecnologias**, v. 4, n. 3, p. 1283-1291, 2014.

MELO NETO, B. A. **Aproveitamento de soro de leite de cabra na elaboração de pão de forma**. 2007. Dissertação (Mestrado em Ciência e Tecnologia de Alimentos). Universidade Federal da Paraíba, João Pessoa. 2007.

MIEKLEY, B.; TRAULSEN, I.; KRIETER, J. Principal component analysis for the early detection of mastitis and lameness in dairy cows. **Journal of Dairy Research**, v. 80, n. 03, p. 335-343, 2013.

MISHRA, D., et al. A Multilevel Image Thresholding Using Particle Swarm Optimization. **International Journal of Engineering & Technology (0975-4024)**, v. 6, n. 2, 2014.

MIZUSHIMA, A.; LU, R. An image segmentation method for apple sorting and grading using support vector machine and Otsu's method. **Computers and Electronics in Agriculture**, v. 94, p. 29-37, 2013.

MOTA, R. A. Aspectos epidemiológicos, diagnóstico e controle das mastites em caprinos e ovinos. **Tecnol. Ciênc. Agropec**, v. 2, n. 3, p. 57-61, 2008.

NA, S.; HERU, X. The segmentation of overlapping milk somatic cells based on improved watershed algorithm. In: **Artificial Intelligence and Computational Intelligence, 2009. AICI'09. International Conference on**. IEEE, 2009. p. 563-566.

NA, S.; HERU, X. Studies on discrimination and segmentation method of touching milk somatic cell. In: **Bioinformatics and Biomedical Technology (ICBBT), 2010 International Conference on**. IEEE, 2010. p. 390-393.

NASHAT, S.; ABDULLAH, A.; ABDULLAH, M. Z. Unimodal thresholding for Laplacian-based Canny–Deriche filter. **Pattern Recognition Letters**, v. 33, n. 10, p. 1269-1286, 2012.

NG, H. Automatic thresholding for defect detection. **Pattern recognition letters**, v. 27, n. 14, p. 1644-1649, 2006.

NGUYEN, Ngoc-Tung; DUONG, Anh-Duc; VU, Hai-Quan. Cell splitting with high degree of overlapping in peripheral blood smear. **International Journal of Computer Theory and Engineering**, v. 3, n. 3, p. 473-478, 2011.

NG, H. P., et al. Masseter segmentation using an improved watershed algorithm with unsupervised classification. **Computers in Biology and Medicine**, v. 38, n. 2, p. 171-184, 2008.

NIELSEN, N. I., et al. Quarter health, milking interval, and sampling time during milking

- affect the concentration of milk constituents. **Journal of Dairy Science**, v. 88, n. 9, p. 3186-3200, 2005.
- NILSSON, B.; HEYDEN, A. Segmentation of complex cell clusters in microscopic images: Application to bone marrow samples. **Cytometry part A**, v. 66, n. 1, p. 24-31, 2005
- NORBERG, E. Electrical conductivity of milk as a phenotypic and genetic indicator of bovine mastitis: A review. **Livestock Production Science**, v. 96, n. 2, p. 129-139, 2005.
- NORBERG, Elise et al. Electrical conductivity of milk: ability to predict mastitis status. **Journal of Dairy Science**, v. 87, n. 4, p. 1099-1107, 2004.
- NORMAN, H. D., et al. Consequence of alternative standards for bulk tank somatic cell count of dairy herds in the United States. **Journal of Dairy Science**, v. 94, n. 12, p. 6243-6256, 2011.
- NUNES, G. R., et al. Avaliação de indicadores inflamatórios no diagnóstico da mamite ovina. **Arquivos do Instituto Biológico**, v. 75, n. 3, p. 271-278, 2008.
- OH, W; LINDQUIST, W. B. Image thresholding by indicator kriging. **Pattern Analysis and Machine Intelligence, IEEE Transactions on**, v. 21, n. 7, p. 590-602, 1999.
- OKADA, K., et al. Rapid diagnostic device for mastitis based on electrochemical detection of superoxide produced from neutrophils in fresh milk. In: **Sensors, 2009 IEEE**. IEEE, 2009. p. 1772-1775.
- OLIVER, S. P; MURINDA, S. E. Antimicrobial resistance of mastitis pathogens. **Veterinary Clinics of North America: Food Animal Practice**, v. 28, n. 2, p. 165-185, 2012.
- OSOWSKY, J.; GAMBA, H. R. Sistema automático para contagem de colônias em placas de Petri. **Rev. bras. eng. biomed**, v. 17, n. 3, p. 131-139, 2001.
- OTSU, N. A threshold selection method from gray-level histograms. **Automatica**, v. 11, n. 285-296, p. 23-27, 1975.
- PAN, C., et al. Robust segmentation for low quality cell images from blood and bone marrow. **International Journal of Control Automation and System**, v. 4, n. 5, p. 637-644, 2006.
- PAN, C., et al. Leukocyte image segmentation using simulated visual attention. **Expert Systems with Applications**, v. 39, n. 8, p. 7479-7494, 2012.
- PANDEY, P; PANDEY, V.; SINGH, J. Somatic cell count (scc) in relation to milk production, milk composition and subclinical mastitis: A review. **International Journal of Cow Science**, v. 1, n. 1, p. 27-32, 2005.
- PARK, Y. W.; HAENLEIN, G. **Handbook of milk of non-bovine mammals**. John Wiley & Sons, 2008.
- PASCHOS, G. Perceptually uniform color spaces for color texture analysis: an empirical evaluation. **Image Processing, IEEE Transactions on**, v. 10, n. 6, p. 932-937, 2001.

- PATTERSON, E. S., et al. Barriers and facilitators to adoption of soft copy interpretation from the user perspective: Lessons learned from filmless radiology for slideless pathology. **Journal of Pathology Informatics**, v. 2, n. 1, p. 1, 2011.
- PEDRINI, H.; SCHWARTZ, W. R. **Análise de imagens digitais: princípios, algoritmos e aplicações**. Thomson Learning, 2008.
- PEIXOTO, R. M.; MOTA, R. A.; COSTA, M. M. Mastite em pequenos ruminantes no Brasil. **Pesq. Vet. Bras**, v. 30, n. 9, p. 754-762, 2010.
- PERNER, P., et al. Recognition of airborne fungi spores in digital microscopic images. In: **Pattern Recognition, 2004. ICPR 2004. Proceedings of the 17th International Conference on**. IEEE, 2004. p. 566-569.
- PETERS, J., et al. The Origins and Spread of Domestic Animals in Southwest Asia and Europe, eds Colledge S, et al. 2013.
- PETROU, M.; BOSDOGIANNI, P. **Image processing: The fundamentals** John Wiley and Sons. **New York**, 1999.
- PHAM, T. D.; CRANE, D. I. Segmentation of neuronal-cell images from stained fields and monomodal histograms. In: **Engineering in Medicine and Biology Society, 2005. IEEE-EMBS 2005. 27th Annual International Conference of the**. IEEE, 2006. p. 6289-6292.
- PHILPOT, W. N.; NICKERSON, S. C. Mastitis: counter attack. 1991.
- POUTREL, B.; LERONDELLE, C. Cell content of goat milk: California Mastitis Test, Coulter Counter, and Fossomatic for predicting half infection. **Journal of Dairy Science**, v. 66, n. 12, p. 2575-2579, 1983.
- PRIYA, R. K., et al. Fuzzy entropy-based MR brain image segmentation using modified particle swarm optimization. **International Journal of Imaging Systems and Technology**, v. 23, n. 4, p. 281-288, 2013.
- PRESCOTT, S. C.; BREEED, R. S. The determination of the number of body cells in milk by a direct method. **Journal of Infection Disease**. n.7, p.632, 1910.
- PUN, T. Entropic thresholding, a new approach. **Computer Graphics and Image Processing**, v. 16, n. 3, p. 210-239, 1981.
- QIAN, X., et al. Self-organizing map-based multi-thresholding on neural stem cells images. **Medical & biological engineering & computing**, v. 47, n. 7, p. 801-808, 2009.
- RANZATO, M., et al. Automatic recognition of biological particles in microscopic images. **Pattern Recognition Letters**, v. 28, n. 1, p. 31-39, 2007.
- REFAI, Hazem et al. Automatic count of hepatocytes in microscopic images. In: **Image Processing, 2003. ICIP 2003. Proceedings. 2003 International Conference on**. IEEE, 2003. p. II-1101-4 vol. 3.
- RIBAS, N. P. Análise do leite. **Gado Holandês**, v. 57, n. 10, p. 92-94, 1994.

RICHARDS, J. A.; RICHARDS, J. A. **Remote sensing digital image analysis**. Berlin et al.: Springer, 1999.

RIQUEIRA, J. **Influência da contagem de células Somáticas no perfil e teores de aminos Bioativas e na qualidade de leite cru e De queijo mussarela**. 2010. 147p. 2011. Tese de Doutorado. Tese (Doutorado em Ciência de Alimentos)–Faculdade de Farmácia, Universidade Federal de Minas Gerais, Belo Horizonte.

ROERDINK, J. B.; MEIJSTER, A. The watershed transform: Definitions, algorithms and parallelization strategies. **Fundamenta informaticae**, v. 41, n. 1, p. 187-228, 2000.

ROMERO, G., et al. Analysis of the electrical conductivity in milking fractions as a mean for detecting and characterizing mastitis in goats. **Small Ruminant Research**, v. 107, n. 2, p. 157-163, 2012.

ROSENFELD, A. Picture processing by computer. **ACM Computing Surveys (CSUR)**, v. 1, n. 3, p. 147-176, 1969.

ROSIN, P. L. Unimodal thresholding. **Pattern Recognition**, v. 34, n. 11, p. 2083-2096, 2001.

RUSS, J. C.; WOODS, R. P. The image processing handbook. **Journal of Computer Assisted Tomography**, v. 19, n. 6, p. 979-981, 1995.

SAHOO, P. K.; SOLTANI, S. A. K. C.; WONG, A. K. A survey of thresholding techniques. **Computer vision, graphics, and image processing**, v. 41, n. 2, p. 233-260, 1988.

SAMET, H.; TAMMINEN, M. Efficient component labeling of images of arbitrary dimension represented by linear bintrees. **Pattern Analysis and Machine Intelligence, IEEE Transactions on**, v. 10, n. 4, p. 579-586, 1988.

SANCHEZ-CUEVAS, M. C.; AGUILAR-PONCE, R. M.; TECPANECATL-XIHUITL, J. L. A Comparison of Color Models for Color Face Segmentation. **Procedia Technology**, v. 7, p. 134-141, 2013.

SANTOS, M. V. Efeito da mastite sobre a qualidade do leite e derivados lácteos. In: **Congresso panamericano de qualidade do leite e controle da mastite**. 2002. p. 179-188.

SCHÖNHOLZER, F., et al. Automated image analysis and in situ hybridization as tools to study bacterial populations in food resources, gut and cast of *Lumbricus terrestris* L. **Journal of Microbiological Methods**, v. 48, n. 1, p. 53-68, 2002.

SCHUKKEN, Y. H., et al. Improving udder health on well managed farms: mitigating the “perfect storm.”. In: **International Conference on the Mastitis Control from Science to Practice**. The Haque. 2008. p. 21-35.

SCHUKKEN, Y. H., et al. Monitoring udder health and milk quality using somatic cell counts. **Veterinary research**, v. 34, n. 5, p. 579-596, 2003.

SEEGERS, H.; FOURICHON, C.; BEAUDEAU, F. Production effects related to mastitis and mastitis economics in dairy cattle herds. **Veterinary Research**, v. 34, n. 5, p. 475-491, 2003.

SEZGIN, M., et al. Survey over image thresholding techniques and quantitative performance evaluation. **Journal of Electronic Imaging**, v. 13, n. 1, p. 146-168, 2004.

SHAN, Z. Y.; YUE, G. H.; LIU, J. Z. Automated histogram-based brain segmentation in T1-weighted three-dimensional magnetic resonance head images. **NeuroImage**, v. 17, n. 3, p. 1587-1598, 2002.

SHANGGUAN, R. **Effect of Subclinical Mastitis and Stage of Lactation on Somatic Cell Count, Milk Composition and Plasmin Activity in Goat Milk**. 2010. Tese de Doutorado. Oklahoma State University.

SHARMA, N.; SINGH, N. K.; BHADWAL, M. S. Relationship of somatic cell count and mastitis: An overview. **Asian-Aust. J. Anim. Sci**, v. 24, n. 3, p. 429-438, 2011.

SHIH, Frank Y.; PU, Christopher C. A skeletonization algorithm by maxima tracking on Euclidean distance transform. **Pattern Recognition**, v. 28, n. 3, p. 331-341, 1995.

SHOSHANI, E., et al. Mammary infection with *Staphylococcus aureus* in cows: progress from inoculation to chronic infection and its detection. **Journal of Dairy research**, v. 67, n. 02, p. 155-169, 2000.

SIERACKI, M. E.; REICHENBACH, S. E.; WEBB, K. L. Evaluation of automated threshold selection methods for accurately sizing microscopic fluorescent cells by image analysis. **Applied and Environmental Microbiology**, v. 55, n. 11, p. 2762-2772, 1989.

SIFRE-MAUNIER, Laurence., et al. A global unimodal thresholding based on probabilistic reference maps for the segmentation of muscle images. **Image and Vision Computing**, v. 24, n. 10, p. 1080-1089, 2006.

SILANIKOVE, N., et al. Recent advances in exploiting goat's milk: quality, safety and production aspects. **Small Ruminant Research**, v. 89, n. 2, p. 110-124, 2010.

SILANIKOVE, N.; MERIN, U.; LEITNER, G. On effects of subclinical mastitis and stage of lactation on milk quality in goats. **Small Ruminant Research**, v. 122, n. 1, p. 76-82, 2014.

SINTORN, I. **Segmentation methods and shape descriptions in digital images**. 2005.

SMITH, M. C.; SHERMAN, D. M. Mammary gland and milk production. **Goat Medicine, Second Edition**, p. 647-689, 2009.

SMITH, K. L. A discussion of normal and abnormal milk based on somatic cell count and clinical mastitis. **Bulletin-International Dairy Federation**, n. 372, p. 43-45, 2002.

SNYDER, W., et al. Optimal thresholding—a new approach. **Pattern Recognition Letters**, v. 11, n. 12, p. 803-809, 1990.

SOILLE, P.; VOGT, P. Morphological segmentation of binary patterns. **Pattern Recognition Letters**, v. 30, n. 4, p. 456-459, 2009.

SOUZA, G. N., et al. Qualidade do leite de rebanhos bovinos localizados na região sudeste: Espírito Santo, Minas Gerais, Rio de Janeiro, Janeiro/2007 a junho/2008. In: **CONGRESSO**

**BRASILEIRO DE QUALIDADE DO LEITE**. 2008. p. 187-189.

STERRETT, A. E. Management and technology solutions for improving milk quality. 2013.

SUZUKI, K.; HORIBA, I.; SUGIE, N. Linear-time connected-component labeling based on sequential local operations. **Computer Vision and Image Understanding**, v. 89, n. 1, p. 1-23, 2003.

TANGORRA, F. M., et al. Milk electrical conductivity and mastitis status in dairy goats: Results from a pilot study. **Small Ruminant Research**, v. 90, n. 1, p. 109-113, 2010.

TIAN, H.; SRIKANTHAN, T.; ASARI, K. V. Automatic segmentation algorithm for the extraction of lumen region and boundary from endoscopic images. **Medical and Biological Engineering and Computing**, v. 39, n. 1, p. 8-14, 2001.

TRIPALDI, C., et al. Effects of mastitis on buffalo milk quality. **Asian-Australasian Journal of Animal Sciences**, v. 23, n. 10, p. 1319-1324, 2010.

TSAI, D. A fast thresholding selection procedure for multimodal and unimodal histograms. **Pattern Recognition Letters**, v. 16, n. 6, p. 653-666, 1995.

TULSANI, H.; SAXENA, S.; YADAV, N. Segmentation using morphological watershed transformation for counting blood cells. **International Journal of Computer Applications & Information Technology**, v. 2, n. 3, p. 28-36, 2013.

VANVRANKEN, S. J. **Utilization of Digital Images in Clinical Hematology-Evaluating the Perceived Benefits and Limitations With the Use of Digital Image Software for Peripheral Blood Differentials**. 2012. Tese de Doutorado. The Ohio State University.

VEZEY, E. L.; SKVARLA, J. J. Computerized feature analysis of exine sculpture patterns. **Review of Palaeobotany and Palynology**, v. 64, n. 1, p. 187-196, 1990.

VIGUIER, C.; et al. Mastitis detection: current trends and future perspectives. **Trends in Biotechnology**, v. 27, n. 8, p. 486-493, 2009.

VOLTOLINI, T., et al. Influência dos estádios de lactação sobre a contagem de células somáticas do leite de vacas da raça holandesa e identificação de patógenos causadores de mastite no rebanho. **Acta Scientiarum. Animal Sciences**, v. 23, p. 961-966, 2008.

WANG, H., et al. Modified watershed technique and post-processing for segmentation of skin lesions in dermoscopy images. **Computerized Medical Imaging and Graphics**, v. 35, n. 2, p. 116-120, 2011.

WANG, W.; PALIWAL, J. Separation and identification of touching kernels and dockage components in digital images. **Canadian Biosystems Engineering**, v. 48, p. 7, 2006.

WANG, Y.; XUE, H. The color image processing technology of the milk somatic cells. In: **Computer Engineering and Technology (ICCET), 2010 2nd International Conference on**. IEEE, 2010. p. V4-280-V4-283.

WELLNITZ, O., et al. Prediction of total quarter milk somatic cell counts based on foremilk



- sampling. **Journal of Dairy Research**, v. 76, n. 03, p. 326-330, 2009.
- WEST, E. D. Assessment of Antimicrobial Resistance in Pathogens Responsible for Causing Bovine Mastitis in Kentucky. 2013.
- WESZKA, J. S.; ROSENFELD, A. Histogram modification for threshold selection. **IEEE Transactions on Systems Man and Cybernetics**, v. 9, n. 1, p. 38-52, 1979.
- WESZKA, J. S. A survey of threshold selection techniques. **Computer Graphics and Image Processing**, v. 7, n. 2, p. 259-265, 1978.
- WHITE, E. C.; HINCKLEY, L. S. Prevalence of mastitis pathogens in goat milk. **Small Ruminant Research**, v. 33, n. 2, p. 117-121, 1999.
- WILTSCHI, K.; PINZ, A.; LINDEBERG, T. An automatic assessment scheme for steel quality inspection. **Machine Vision and Applications**, v. 12, n. 3, p. 113-128, 2000.
- WU, H. S.; BERBA, J.; GIL, J. Iterative thresholding for segmentation of cells from noisy images. **Journal of Microscopy**, v. 197, n. 3, p. 296-304, 2000.
- WU, J. Y., et al. Deoxyribonucleic acid sensor for the detection of somatic cells in bovine milk. **Biosystems Engineering**, v. 90, n. 2, p. 143-151, 2005.
- WYSZECKI, G.; STILES, W. S. **Color science**. New York: Wiley, 1982.
- XU, L.; LU, H.; ZHANG, M. Automatic segmentation of clustered quantum dots based on improved watershed transformation. **Digital Signal Processing**, v. 34, p. 108-115, 2014.
- YANG, Y.; YAN, H. An adaptive logical method for binarization of degraded document images. **Pattern Recognition**, v. 33, n. 5, p. 787-807, 2000.
- ZHANG, Y., et al. Salient region detection for complex background images using integrated features. **Information Sciences**, v. 281, p. 586-600, 2014.

**UFBA**  
**UNIVERSIDADE FEDERAL DA BAHIA**  
**ESCOLA POLITÉCNICA**

**PROGRAMA DE PÓS GRADUAÇÃO EM ENGENHARIA INDUSTRIAL - PEI**

Rua Aristides Novis, 02, 6º andar, Federação, Salvador BA  
CEP: 40.210-630  
Telefone: (71) 3283-9800  
E-mail: [pei@ufba.br](mailto:pei@ufba.br)  
Home page: <http://www.pei.ufba.br>

