

ARLETE TERESINHA BEUREN

**ABORDAGEM DE SEPARAÇÃO ENTRE CÉU E
CHÃO EM IMAGENS DIGITAIS POR
APRENDIZAGEM DE MÁQUINA**

CURITIBA-PR

2018

ARLETE TERESINHA BEUREN

**ABORDAGEM DE SEPARAÇÃO ENTRE CÉU E
CHÃO EM IMAGENS DIGITAIS POR
APRENDIZAGEM DE MÁQUINA**

Tese de Doutorado apresentada ao Programa de Pós-Graduação em Informática da Pontifícia Universidade Católica do Paraná – PUCPR como requisito parcial para obtenção do título de Doutora em Informática.

Área de Concentração: Ciência da Computação

Orientadores: Prof. Dr. Alceu de Souza Britto Jr.
Prof. Dr. Jacques Facon

CURITIBA-PR

2018

Dedico esta pesquisa a minha família que sempre me motivou para que eu conseguisse atingir mais essa etapa.

“Descobrir consiste em olhar para o que todo mundo está vendo e pensar uma coisa diferente”. (Roger Von Oech)

AGRADECIMENTOS

Um agradecimento especial aos meus orientadores ao longo destes três anos de pesquisa, Prof. Dr. Alceu de Souza Britto Jr. e Prof. Dr. Jacques Facon, pela disponibilidade e paciência. Pessoas que admiro muito pelo profissionalismo e dedicação com a profissão.

À minha família, meus pais, esposo e filho pelo incentivo e compreensão nos momentos de ausência, durante todo este período.

Aos meus amigos de doutorado pelo apoio e momentos de estudo compartilhados.

A CAPES pelo apoio financeiro.

E a todos que direta ou indiretamente colaboraram na execução desta pesquisa.

SUMÁRIO

LISTA DE FIGURAS	VII
LISTA DE TABELAS	IX
RESUMO	XI
CAPÍTULO 1	1
INTRODUÇÃO	1
1.1 Desafio – Questão da Pesquisa	2
1.2 Motivação	4
1.3 Proposta	5
1.4 Objetivos da Pesquisa	5
1.4.1 Objetivo Geral.....	5
1.4.2 Objetivos Específicos	5
1.5 Contribuição	6
1.6 Organização do Trabalho	7
CAPÍTULO 2	8
FUNDAMENTAÇÃO TEÓRICA E ESTADO DA ARTE	8
2.1 Segmentação de Imagens.....	8
2.1.1 Segmentação por Cor	8
2.1.1.1 Segmentação por Índices de Branquidão (<i>Whiteness</i>).....	8
2.1.1.2 Segmentação por Índices de Azul (<i>Blueness</i>).....	13
2.1.2 Segmentação por Textura	15
2.1.2.1 Algoritmo LBP (<i>Local Binary Pattern</i>)	15
2.1.2.2 Perfil Morfológico Estendido (<i>Extended Morphological Profile</i>).....	16
2.1.3 Segmentação com características de Média/Variância/Obliquidade/ Curtose/Energia/Entropia	17
2.2 Classificador SVM (Support Vector Machine)	18
2.3 Trabalhos Relacionados.....	19
2.3.1 Abordagens usando classificadores	19
2.3.2 Abordagens em níveis de cinza/cor	25
2.3.3 Abordagens usando Detecção de Bordas	27
2.4 Análise dos Trabalhos Relacionados	32
CAPÍTULO 3	39

MÉTODO PROPOSTO	39
3.1 Base de Dados.....	39
3.2 Extração de Características para Segmentação	43
3.2.1 Características Baseadas em Índices de Branquidão	43
3.2.2 Características Baseadas Índices de Azul	43
3.2.3 Características Baseadas em LBP (<i>Local Binary Pattern</i>)	44
3.2.4 Características Baseadas em Perfil morfológico estendido (<i>Extended Morphological Profile</i>)	44
3.2.5 Outras Características	45
3.3 Classificação	46
3.3.1 Uso de Classificadores Monolíticos	46
3.3.2 Uso de Múltiplos Classificadores	46
3.4 Avaliação da Abordagem	47
CAPÍTULO 4	49
EXPERIMENTOS.....	49
4.1 Protocolo Experimental	49
4.2 Experimentos – Segmentação Baseada em Espaços de Cor	51
4.3 Experimentos – Segmentação Baseada em Índices de Branquidão	53
4.4 Experimentos – Segmentação Baseada em Índices de Azul	58
4.5 Experimentos – Segmentação Baseada em LBP (<i>Local Binary Pattern</i>)	62
4.6 Experimentos – Segmentação Baseada em Perfil Morfológico Estendido	62
4.7 Experimentos - Classificadores Monolíticos	63
4.8 Experimentos – Múltiplos Classificadores	68
4.9 Discussão dos Resultados	74
CAPÍTULO 5	76
CONSIDERAÇÕES FINAIS	76
REFERÊNCIAS.....	79

LISTA DE FIGURAS

Figura 1.1: Exemplo de imagem com a linha do horizonte destacada.	1
Figura 1.2 – Exemplo de imagens com variações de céu.	3
Figura 1.3: Exemplo de imagens com variações de chão.	3
Figura 1.4: Exemplo de imagem com contexto cidade (a) com sua respectiva <i>groundtruth</i> (b).....	4
Figura 2.1: Processo de construção de um código LBP.....	16
Figura 2.2: Classificação entre duas classes w_1 e w_2 usando hiperplanos.	19
Figura 3.1: Exemplo das categorias de imagens com variações de céu e chão e suas respectivas <i>groundtruth</i> contemplando a Base Variações.	41
Figura 3.2: Imagens representando as quatro categorias e suas respectivas <i>groundtruth</i> contemplando a Base Web.	42
Figura 3.3: Tela da aplicação java para seleção dos índices de branquidão.	43
Figura 3.4: Tela da aplicação java para seleção dos índices de azul.....	44
Figura 3.5: Exemplo de tela da aplicação java para seleção de características.....	45
Figura 3.6: Exemplo de tela da aplicação java para gerar o arquivo de treino.	47
Figura 4.1: Protocolo dos experimentos para características individuais/agrupadas.	50
Figura 4.2: Protocolo dos experimentos com combinação de classificadores.	50
Figura 4.3: Exemplo de processamento de imagem para cada componente R, G, B, H, S, V, Y, Cr, Cb.	52
Figura 4.4: Resultados individuais dos índices de branquidão para a Base Variações.	55
Figura 4.5: Resultados individuais dos índices de branquidão para a Base Web.	58
Figura 4.6: Resultados individuais dos índices de azul para a Base Variações.....	59
Figura 4.7: Resultados individuais dos índices de azul para a Base Web.	61
Figura 4.8: As imagens (a) e (c) representam a segmentação ideal e as imagens (b) e (d) correspondem as respectivas imagens segmentadas.	64
Figura 4.9: Matriz de confusão do índice $W_{Stephansen}$	69
Figura 4.10: Matriz de confusão do índice $W_{Stephansen}$	69

Figura 4.11: As imagens (a) e (c) representam a segmentação ideal e as imagens (b) e (d) correspondem as respectivas imagens segmentadas.	70
Figura 4.12: As imagens (a) e (c) representam a segmentação ideal e as imagens (b) e (d) correspondem as respectivas imagens segmentadas.	71
Figura 4.13: As imagens (a) e (c) representam a segmentação ideal e as imagens (b) e (d) correspondem as respectivas imagens segmentadas.	72
Figura 4.14: As imagens (a) e (c) representam a segmentação ideal e as imagens (b) e (d) correspondem as respectivas imagens segmentadas.	73

LISTA DE TABELAS

Tabela 2.1: Tabela de valores de a e b .	9
Tabela 2.2: Tabela de valores para x_0 e y_0 .	10
Tabela 2.3: Cálculo de erros médios absolutos na comparação dos métodos.	21
Tabela 2.4: Erros absolutos médios entre os horizontes detectados e <i>groundtruth</i> para várias opções de custos nodais.	24
Tabela 2.5: Informação de RGB na área VERDE.	27
Tabela 2.6: Informação de RGB na área AZUL.	27
Tabela 2.7: Taxas de acerto da comparação das bases de imagens.	30
Tabela 2.8: Comparação dos métodos.	33
Tabela 4.1: Avaliação da separação de céu e chão para os componentes R , G , B , H , S , V , Y , Cr e Cb .	51
Tabela 4.2: Avaliação da separação de céu e não céu variando tamanho da janela.	53
Tabela 4.3: Avaliação individual para os índices de branquidão para a Base Variações.	55
Tabela 4.4: Avaliação individual para os índices de branquidão para a Base Web.	58
Tabela 4.5: Avaliação individual dos índices de azul para a Base Variações.	60
Tabela 4.6: Avaliação individual dos índices de azul para a Base Web.	61
Tabela 4.7: Resultados do algoritmo LBP para as Bases de Imagens.	62
Tabela 4.8: Resultados com perfil morfológico estendido para as Bases de Imagens.	62
Tabela 4.9: Melhores resultados para cada vetor de característica.	63
Tabela 4.10: Resultados para Perfil morfológico estendido para a Base Variações.	64
Tabela 4.11: Resultados para Perfil morfológico estendido para a Base Web.	65
Tabela 4.12: Resultados usando a combinação de características de branquidão e azul.	65
Tabela 4.13: Resultados usando a combinação de características de branquidão e azul com LBP.	66
Tabela 4.14: Resultados usando a combinação de características de branquidão e azul com perfil morfológico estendido.	66

Tabela 4.15: Resultados usando a combinação de votos para a Base Variações com diferentes tamanhos de janelas.....	67
Tabela 4.16: Resultados usando a combinação de votos para a Base Web com diferentes tamanhos de janelas.....	67
Tabela 4.17: Resultados usando a combinação de classificadores para a Base Variações.	68
Tabela 4.18: Resultados usando a combinação de classificadores para a Base Web.	69
Tabela 4.19: Resultados usando a combinação de classificadores para a categoria de Cidade.....	70
Tabela 4.20: Resultados usando a combinação de classificadores para a categoria de Mar/Porto.....	71
Tabela 4.21: Resultados usando a combinação de classificadores para a categoria de Natureza/Montanha.	72
Tabela 4.22: Resultados usando a combinação de classificadores para a categoria de Rodovia/Estrada.....	73
Tabela 4.23: Resultados para a Base Variações.	75
Tabela 4.24: Resultados para a Base Web.....	75

RESUMO

A segmentação de céu e chão em imagens tem sido alvo de várias pesquisas no âmbito acadêmico para posterior detecção da linha do horizonte. Nesse contexto, diferentes soluções computacionais têm sido apresentadas na literatura, no entanto, a maioria dos trabalhos não utiliza as características para classificação e uma avaliação numérica como as citadas nesta pesquisa. O objetivo proposto neste trabalho consiste em definir uma abordagem de segmentação de céu e chão em imagens digitais com o uso do algoritmo supervisionado SVM (Máquina de Vetor de Suporte) usando diferentes características. O estudo conta com duas bases de 1.200 imagens, cada uma contendo imagens com diferentes contextos de céu e chão. Os experimentos foram classificados para segmentar as imagens em duas classes com diversas características, entre elas, índices de branquidão e perfil morfológico estendido, não usadas na literatura para esse tipo de abordagem. Esta pesquisa pode ser dividida em duas etapas: a primeira em que as melhores características são obtidas com a criação de classificadores monolíticos através da técnica de validação cruzada e a segunda que contempla duas abordagens distintas para segmentar céu e chão com combinação de classificadores: a primeira que contém imagens misturadas para a cada base e a segunda abordagem que consiste na pré-classificação das imagens em categorias (Cidade; Mar/Porto; Natureza/Montanha; Rodovia/Estrada). Na primeira abordagem o melhor índice identificado é $W_{Stephansen}$ com *F-measure* de 93% para a primeira base e 89% para a segunda base e, na segunda abordagem os melhores resultados foram os índices de $W_{Stephansen}$ para as categorias de Cidade com 97% e Rodovia/Estrada com 92% e $W_{HunterLAB}$ para as categorias de Mar/Porto com 95% e Natureza/Montanha com 92% de taxa de acerto para a primeira base. Para a segunda base os melhores índices são $W_{Stephansen}$ para as categorias de Cidade com 91%, Natureza/Montanha com 92% e Rodovia/Estrada com 90% e o índice $W_{Taubé}$ para a categoria de Mar/Porto com 84% de taxa de acerto.

Palavras-chave: imagens digitais; características; segmentação; avaliação.

ABSTRACT

The segmentation of sky and ground in images has been the object of several researches in the academic scope for later detection of the horizon line. In this context, different computational solutions have been presented in the literature, however, most of the works do not use the classification characteristics and a numerical evaluation as the ones cited in this research. The goal of this work is to define a sky and ground segmentation approach in digital images using the supervised algorithm SVM (Support Vector Machine) using different characteristics. The study has two bases of 1,200 images, both containing images with different contexts of sky and ground. The experiments were classified to segment the images into two classes with several characteristics. Among them, whiteness indexes and extended morphological profile, both not used in the literature for this type of approach. This research can be divided into two stages. In the first, the best characteristics are obtained by the creation of monolithic classifiers through the cross-validation technique. The second stage considers two different approaches to segment sky and ground with a combination of classifiers: the first combination holds mixed images for each base, and the second approach carries a pre-classification of images into categories (City, Sea / Harbor, Nature / Mountain, Highway / Road). In the first approach, the best index identified is $W_{Stephansen}$, with F-measure of 93% for the first base, and 89% for the second base. In the second approach, the best results were the index $W_{Stephansen}$ for the categories City, with 97%, and Highway / Road, with 92%, and the index $W_{HunterLAB}$, for the categories Sea / Harbor, with 95%, and Nature / Mountain, with 92% of hit rate for the first base; for the second base, the best indexes were $W_{Stephansen}$, for the categories City, with 91%, Nature / Mountain, with 92%, and Highway / Road, with 90%, and the index W_{Taube} for the Sea / Harbor category, with 84% of success rate.

Keywords: digital images; characteristics; segmentation; evaluation.

CAPÍTULO 1

INTRODUÇÃO

O horizonte é uma palavra derivada do grego antigo significando limiar e é definido como a linha aparente aonde o céu parece encontrar o chão. Esta linha varia dependendo da altura e posição do observador. Quanto mais alta a linha do horizonte se situar, maior é o espaço de visão (LIE et al., 2005). A Figura 1.1 exemplifica uma imagem com a linha do horizonte destacada.



Figura 1.1: Exemplo de imagem com a linha do horizonte destacada.

No trabalho de Lie et al. (2005), a linha do horizonte ou simplesmente o horizonte representa uma informação valiosa em imagens aéreas. No caso de imagens digitais e vídeos, a linha do horizonte é usada para diferentes fins, tais como guiar a navegação marítima e aérea no sentido de pilotar veículos marinhos, ajustar planos de voos, e facilitar a vigilância para evitar acidentes, principalmente colisões com obstáculos.

Definir uma abordagem automática de segmentação do céu e chão em imagens digitais não é uma tarefa fácil já que o céu pode se encontrar com variações de luminosidade e clima e o chão pode ser mar, gelo, montanhas,

florestas, cidades, por exemplo, e assim gerando contextos diferentes de cor, texturas e etc.

De acordo com Bao, Xiong e Zhou (2005), muitos estudos supõem que a linha do horizonte na imagem é uma linha reta com características claras da borda, mas esta suposição não é verdadeira sob condições meteorológicas e ambientais diferentes. Por exemplo, quando está nublado, o contraste e o brilho da imagem se tornam mais fracos e o horizonte não é tão óbvio e os recursos de borda da linha do horizonte, por exemplo, são difíceis de serem encontrados.

1.1 Desafio – Questão da Pesquisa

A segmentação precisa do céu e chão para gerar posteriormente a linha do horizonte é essencial para um grande número de aplicações, tais como segurança portuária, voos e controle de navegação, gerenciamento de fluxo etc. A localização da linha do horizonte pode ser vista como a linha que separa a imagem em duas regiões: o céu e o chão, sendo o chão uma floresta, mar, cidade, rodovia, etc. Nas imagens de cidade, torna-se mais difícil a separação de céu e chão levando em conta o grande número de obstáculos como, por exemplo, antenas, prédios, linhas de energia, postes, entre outros.

Várias abordagens já foram propostas na literatura, entre elas, abordagens usando classificadores e detecção de bordas. A maioria das técnicas propostas usam classificadores para a separação de céu e chão, a dificuldade encontrada nessas técnicas é que a maioria informa taxas de acerto referentes à velocidade e comparação de desempenho de algoritmos e, muitas vezes, não informam como a avaliação foi realizada, além de apresentarem bases muito pequenas (em torno de 10 a 50 imagens), sem muita diversificação de céu e chão. Bebis et al. (2013) apresentam a taxa de erro absoluto médio com o cálculo de TP (verdadeiro positivo), TN (verdadeiro negativo), FP (falso positivo) e FN (falso negativo) de 4,29% para 80 imagens montanhosas extraídas da web, ou seja, imagens de um mesmo tipo.

Nesse contexto, o presente estudo propõe o uso de novas características com o classificador SVM para bases de imagens maiores e com grande variação de céu e chão, conforme Figuras 1.2 e 1.3, comparadas com as citadas na literatura.



Figura 1.2 – Exemplo de imagens com variações de céu.

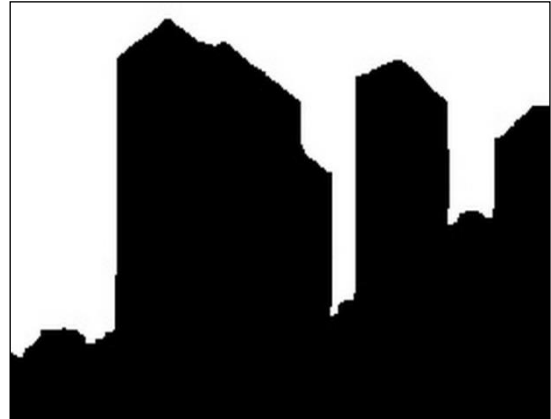


Figura 1.3: Exemplo de imagens com variações de chão.

A avaliação proposta nesta abordagem usa a métrica de *F-measure* como taxa de acerto para analisar a técnica empregada com a sua respectiva *groundtruth*. A Figura 1.4 apresenta uma imagem com sua *groundtruth*.



(a) Imagem de cidade



(b) Imagem *groundtruth*

Figura 1.4: Exemplo de imagem com contexto cidade (a) com sua respectiva *groundtruth* (b).

As duas bases de imagens criadas para este estudo contemplam imagens coloridas que apresentam diferentes contextos de céu (nublado, limpo, claro/escuro) e de chão como, cidade, mar/porto, natureza/montanha e rodovia/estrada. Estes diferentes contextos despertam o interesse em saber se as características devem ser diferentes, ou não, dependendo do tipo das imagens, o que não foi abordado na literatura até o momento. Por isso, esta pesquisa apresenta duas grandes etapas: a primeira com classificadores monolíticos que analisam quais as melhores características para a segmentação entre céu e chão para as bases de imagens apresentadas e a segunda etapa que possui duas abordagens com a combinação de classificadores. A primeira abordagem com todas as imagens misturadas para cada base e a outra apresenta a divisão das imagens das bases em categorias sendo Cidade, Mar/Porto, Natureza/Montanha e Rodovia/Estrada.

1.2 Motivação

Há uma grande variedade entre os tipos de imagens como, por exemplo, imagens com diferentes tipos de céu (nublado, chuvoso, limpo, claro/escuro) mais chão (cidade, mar/porto, montanha/floresta, rodovia/estrada), o que dificulta a segmentação do céu e chão.

Portanto, busca-se nessa pesquisa criar duas bases de imagens com grande diversidade de céu e chão com suas respectivas *groundtruth* para que sejam disponibilizadas para pesquisas futuras; índices de branquidão, índices de azul e perfil morfológico estendido não usados na literatura para esse fim; características citadas na literatura para esse tipo de abordagem; gerar um vetor de características; avaliar numericamente as imagens geradas pela abordagem com suas *groundtruth* através de métricas para avaliação numérica.

1.3 Proposta

Desenvolver mecanismos automáticos de segmentação de céu e chão, usando bases que contemplam imagens diversificadas, com o uso de características como cor, média, etc., índices de branquidão, algoritmo LBP (*Local Binary Pattern*), índices de azul e perfil morfológico estendido para gerar o vetor de características usado no classificador supervisionado SVM (Máquina de Vetor de Suporte).

1.4 Objetivos da Pesquisa

1.4.1 Objetivo Geral

O objetivo deste trabalho consiste em definir uma abordagem automática de segmentação de céu e chão, considerando diferentes contextos de céu e chão, utilizando diferentes características apresentando uma avaliação numérica entre a imagem binária e sua *groundtruth* para bases que contemplam imagens com diferentes contextos de céu e chão.

1.4.2 Objetivos Específicos

Para definir mecanismos de segmentação de céu e chão faz-se necessário a realização de etapas, são elas:

- a) Realizar uma revisão sistemática da literatura sobre a segmentação de céu/chão e detecção da linha do horizonte em imagens digitais, analisando trabalhos já publicados sobre o assunto;
- b) Criar duas bases de imagens: uma contendo imagens apresentando variações de luminosidade, clima e estações do ano, como representado na Figura 1.2 e outra contendo imagens com diferentes contextos de chão coletadas na web, conforme Figura 1.3;
- c) Criar as bases de imagens correspondentes com a segmentação *groundtruth*;
- d) Detectar regiões de segmentação de céu e chão, qualquer que seja o contexto de céu (neblina, chuva, pouca ou muita luz) e chão (cidade, mar/porto, natureza/montanha, rodovia/estrada);
- e) Utilizar características como a cor, média, algoritmo LBP, Perfil Morfológico Estendido (*Extended Morphological Profile*), índices de azul e índices de branquidão, este último não usado na literatura para esse fim;
- f) Gerar um vetor de características para alimentar o classificador SVM. A escolha pelo SVM se deu por ser uma abordagem de classificação usada em muitos estudos descritos na literatura;
- g) Treinar um classificador para cada característica; Treinar um classificador para características agrupadas; Aplicar o esquema de combinação de votos;
- h) Medir a abordagem proposta através de métrica de segmentação que define taxas de acerto.

1.5 Contribuição

Nesta pesquisa espera-se obter como resultado a revisão sistemática da literatura analisando os trabalhos já publicados sobre o assunto; a criação de bases de imagens aéreas a partir de bases de dados disponíveis e coletadas da web com a segmentação *groundtruth*; a segmentação de céu/chão qualquer que seja o

contexto de céu (neblina, chuva, pouca ou muita luz) e de chão (cidade, mar/porto, natureza/montanha, rodovia/estrada); uso de características para gerar o vetor de características; treinar um classificador para cada característica e para características agrupadas; aplicar a combinação de votos; medir a eficiência da abordagem através de uma métrica de comparação.

No contexto geral, a contribuição da pesquisa está em avaliar as características implementadas em relação à variação de luminosidade e diferentes cenários (contextos), por isso a necessidade da criação de duas diferentes bases de imagens: uma com grande variação no contexto de céu e outra com grande variação no contexto de chão. O estudo apresenta como nova característica o uso dos índices de branquidão, ainda não citados na literatura para a segmentação entre céu e chão. Estes índices foram comparados com demais características citadas na literatura e agrupados para análise de sua eficiência. O estudo apresenta ainda a combinação de votos, não usada nesse tipo de abordagem, aplicada aos melhores índices de branquidão para uma análise dos classificadores gerados. Além dos índices de branquidão o estudo apresenta o uso de índices de azul e o perfil morfológicos estendido, não aplicados neste tipo de abordagem.

Parte deste estudo foi publicado no MICAI (*16th Mexican International Conference on Artificial Intelligence*).

1.6 Organização do Trabalho

Este trabalho está organizado da seguinte forma: O Capítulo 2 apresenta a fundamentação teórica e os trabalhos relacionados, tais pesquisas foram fundamentais para a realização de uma análise crítica do método proposto neste trabalho, bem como para avaliação dos resultados obtidos. O Capítulo 3 apresenta a metodologia para o desenvolvimento deste trabalho: o método proposto e o protocolo de experimentos. O Capítulo 4 aborda o detalhamento dos resultados obtidos durante esta pesquisa. Finalmente, no Capítulo 5 são apresentadas as considerações finais e trabalhos futuros desta pesquisa.

CAPÍTULO 2

FUNDAMENTAÇÃO TEÓRICA E ESTADO DA ARTE

2.1 Segmentação de Imagens

A segmentação de céu e chão para então gerar a linha do horizonte representa uma informação valiosa em imagens digitais. No caso de imagens digitais e vídeos, a linha do horizonte é usada para diferentes fins, tais como guiar a navegação marítima e aérea no sentido de pilotar veículos marinhos, ajustar planos de voos, e facilitar a vigilância para evitar acidentes, principalmente colisões com obstáculos.

Esta Seção descreve a segmentação por cor utilizando espaços de cores para índices de branquidão e índices de azul, algoritmo LBP de textura e perfil morfológico estendido.

2.1.1 Segmentação por Cor

Estudos experimentais e teóricos permitiram expressar a natureza física das cores através de modelos matemáticos que resultaram no desenvolvimento de espaços de cores. Um espaço de cor é um modelo matemático utilizado para representar uma gama de cores. Existem vários espaços de cores, como RGB, HSI, HSV, YCrCb, Luv, Lab, XYZ, entre outros (SANGWINE E HORNE, 1998). Alguns espaços de cores são usados para o cálculo dos índices de branquidão e de azul.

2.1.1.1 Segmentação por Índices de Branquidão (*Whiteness*)

O conceito de branquidão é um índice que mede o grau relativo de branco publicados na literatura (X-RITE, 2012, p 31). Esses índices são usados no clareamento de próteses dentárias (JOINER et al., 2008, p 2-7), (PUEBLA, 2006, p

64), na indústria para análise de branquidão em materiais como plástico e papel (LIN, 2013, p 366).

Os índices de branquidão encontrados nos estudos realizados são:

- $W_{Astme313}$, (LIN, 2013, p 366), (X-RITE, 2012, p 31): O índice de branquidão é calculado de acordo com a equação 1:

$$W_{Astme313} = 3.388 * Z - 3 * Y \quad (1)$$

Com a conversão do espaço de cor RGB para o espaço de cor XYZ sendo apresentado na equação 2:

$$\begin{Bmatrix} x \\ y \\ z \end{Bmatrix} = \begin{Bmatrix} 0.412453 & 0.357580 & 0.180423 \\ 0.212671 & 0.715160 & 0.072169 \\ 0.019334 & 0.119193 & 0.950227 \end{Bmatrix} * \begin{Bmatrix} R \\ G \\ B \end{Bmatrix} \quad (2)$$

- W_{Berger} , (LIN, 2013, p 366), (PUEBLA, 2006, p 64): Conversão do espaço de cor RGB para o espaço de cor XYZ. Em seguida, o índice de branquidão é calculado de acordo com a equação 3:

$$W_{Berger} = Y + a * Z - b * X \quad (3)$$

Com a e b sendo coeficientes do observador tabelados da seguinte forma:

Tabela 2.1: Tabela de valores de a e b .

	a	b
2º observador	3.440	3.895
10º observador	3.448	3.904

Fonte: Adaptado de (LIN, 2013, p 366).

- W_{Cie} , (JOINER et al., 2008, p 2-7), (PUEBLA, 2006, p 64): Conversão do espaço de cor RGB para o espaço de cor XYZ. Em seguida, o índice de branquidão é calculado conforme a equação 4:

$$W_{Cie} = Y + 800 * (x_0 - x) + 1700 * (y_0 - y) \quad (4)$$

Com x e y sendo coordenadas de cromaticidade normalizadas definidas de acordo com a equação 5:

$$x = \frac{X}{X+Y+Z} \quad y = \frac{Y}{X+Y+Z} \quad (5)$$

E x_0 e y_0 sendo as coordenadas de cromaticidade para a luminância do observador de origem, sendo os valores tabelados de acordo com a tabela 2.2:

Tabela 2.2: Tabela de valores para x_0 e y_0 .

Value	C/2°	D50/2°	D65/2°	C/10°	D50/10°	D65/10°
x_0	0.3101	0.3457	0.3127	0.3104	0.3477	0.3138
y_0	0.3161	0.3585	0.3290	0.3191	0.3595	0.3310

Valor C/2°, seja $x_0 = y_0 = 0.3101$.

Fonte: Adaptado de (JOINER, 2008, p 2-7).

- W_{Ganz} , (LIN, 2013, p 366): Conversão do espaço de cor RGB para o espaço de cor XYZ. Em seguida, o índice de branquidão é calculado conforme equação 6:

$$W_{Ganz} = Y + P * x + Q * y + C \quad (6)$$

X, Y, Z : Valores encontrados a partir da conversão.

Com x e y sendo coordenadas de cromaticidade normalizadas definidas a partir da equação 7:

$$x = \frac{X}{X+Y+Z} \quad y = \frac{Y}{X+Y+Z} \quad (7)$$

e P, Q, C sendo coeficientes nominais dados para a luminância D65/10, tabelados da seguinte maneira: $P = 1868,322$; $Q = 3695,690$; $C = 1809.441$.

- $W_{Harrison}$ (PUEBLA, 2006): Com R e B sendo os componentes de vermelho e azul do pixel conforme a equação 8.

$$W_{Harrison} = 100 - (R - B) \quad (8)$$

- W_{Hunter} , (X-RITE, 2012, p 31), de acordo com a equação 9:

$$W_{Hunter} = 100 - \left\{ \left[220 * \frac{(G - B)}{(G + 0.242 * B)^2} \right] + \left[\left(100 - \frac{G}{2} \right) \right]^2 \right\}^{1/2} \quad (9)$$

Com G e B sendo os componentes de verde e azul do pixel.

- $W_{Stensby}$, (X-RITE, 2012, p 31): Conversão do espaço de cor RGB para o espaço de cor XYZ. Em seguida, o espaço de cor XYZ é convertido para o espaço de cor L*a*b*. O índice de branquidão é calculado como mostra a equação 10:

$$W_{Stensby} = L^* - 3b^* + 3a^* \quad (10)$$

- $W_{Stephansen}$, (LIN, 2013, p 366): Com R e B sendo os componentes de vermelho e azul do pixel de acordo com a equação 11.

$$W_{Stephansen} = 2B - R \quad (11)$$

- W_{Taube} , (LIN, 2013, p 366): Com G e B sendo os componentes de verde e azul do pixel conforme a equação 12.

$$W_{Taube} = 4B - 3G \quad (12)$$

- $W_{HunterLAB}$, (LIN, 2013, p 366), (JOINER, 2008, p 2-7): Aqui se utiliza a conversão do espaço de cor RGB para o espaço de cor espaço de cor XYZ. Em seguida, o espaço de cor XYZ é convertido para o espaço de cor L*a*b*. O índice de branquidão é calculado conforme a equação 13:

$$W_{HunterLAB} = L - 3b \quad (13)$$

A conversão do espaço XYZ para o espaço de cor L*a*b* é feito de acordo com a equação 14:

$$\begin{aligned} L &= 100 \cdot \sqrt{\frac{Y}{Y_n}} \\ a &= 175 \cdot \sqrt{\frac{0.0102 \cdot X_n}{Y/Y_n}} \cdot \left(\frac{X}{X_n} - \frac{Y}{Y_n} \right) \\ b &= 70 \cdot \sqrt{\frac{0.00847 \cdot Z_n}{Y/Y_n}} \cdot \left(\frac{X}{X_n} - \frac{Z}{Z_n} \right) \end{aligned} \quad (14)$$

Sob a luminância D65/10° os valores de X_n , Y_n , Z_n são:

$$X_n = 0.95047; Y_n = 1.00000; Z_n = 1.08883.$$

- W_{Lanier} , (JOINER, 2008, p 2-7): Utiliza a conversão do espaço de cor RGB para o espaço de cor XYZ. Em seguida, o espaço de cor XYZ é convertido para o espaço de cor L*a*b*. O índice de branquidão é calculado como mostra a equação 15:

$$W_{Lanier} = 100 - \sqrt{(100 - L^*)^2 + a^{*2} + b^{*2}} \quad (15)$$

Com L^* , a^* , b^* , sendo definidos pela conversão mostrada na equação 16:

$$\begin{aligned} L^* &= 116f(Y/Y_n) - 16 \\ a^* &= 500[f(X/X_n) - f(Y/Y_n)] \\ b^* &= 200[f(Y/Y_n) - f(Z/Z_n)] \end{aligned} \quad (16)$$

Sob a luminância D65/10°, os valores de X_n , Y_n , Z_n são:

$$X_n = 0.95047; Y_n = 1.00000; Z_n = 1.08883.$$

2.1.1.2 Segmentação por Índices de Azul (*Blueness*)

O conceito de azul é um índice que mede o grau relativo de azul. Esses índices são usados em aplicações que medem a quantidade de azul encontrada nos mares e oceanos (WATANABE et al., 2016). Os vários índices de azul descritos na literatura foram:

- B_{SI} , (YAMASHITA e YOSHIMURA, 2012, p 138), conforme equação 17:

$$B_{SI} = \frac{DN_B - DN_R}{DN_B + DN_R} \quad (17)$$

Que R e B são os canais do espaço de cor RGB.

E que DN_B e DN_R representam o número de pixels que tem informações nos canais azul e vermelho, respectivamente.

Os valores SI tem intervalo entre -1.0 e 1.0.

- B_{BI} , (YAMASHITA e YOSHIMURA, 2012, p 138), conforme equação 18:

$$B_{BI} = \frac{DN_B + DN_G + DB_R}{(2^n - 1) * 3} \quad (18)$$

Onde DN_B = Número digital do canal azul;

DN_G = Número digital do canal verde;

DN_R = Número digital do canal vermelho;

n = bit número de nível de quantização.

Que R, G e B são os canais do espaço de cor RGB.

E que DN_B , DN_G e DN_R representam o número de pixels que tem informações nos canais azul, verde e vermelho, respectivamente.

Os valores SI tem intervalor entre -1.0 e 1.0.

- B_{YUV} , (CASTRO, 2015): Conversão do espaço RGB para o espaço YCrCb, de acordo com a equação 19:

$$\begin{aligned}
Y &= (0.257 * R) + (0.504 * G) + (0.098 * B) + 16 \\
Cr &= (0.439 * R) - (0.368 * G) - (0.071 * B) + 128 \\
Cb &= -(0.148 * R) - (0.291 * G) + (0.439 * B) + 128
\end{aligned}
\tag{19}$$

Usando a conversão do espaço RGB para o espaço **YUV**, cuja formulação matemática é apresentada na equação 20.

$$\begin{aligned}
Y &= W_r * R + W_g * G + W_b * B \\
U &= 0.436 * (B - Y) / (1 - W_b) \\
V &= 0.615 * (R - Y) / (1 - W_r)
\end{aligned}
\tag{20}$$

Em que $W_r = 0.299$; $W_b = 0.114$; $W_g = 1 - W_r - W_b = 0.587$.

O índice de Azul SI usando os espaços de cor YUV e YCrCb é representado na equação 21:

$$SI = U/Cb \tag{21}$$

- $B_{Threshold}$, (LOUI E SAVAKIS, 2000, p 3), conforme equação 22:

$$B_{Threshold} = \begin{cases} 1 & \text{se } B > \max(R, Tb) \\ 1 & \text{se } G : \max(R, Tb) \\ 0 & \text{senão} \end{cases}
\tag{22}$$

Em que Tb é um valor de limiar para assegurar que a cor de Azul não seja preta.

- $B_{ExcessoAzul}$, (CASTRO, 2015, p 5):

$$B_{ExcessoAzul} = 1.4B - G \tag{23}$$

- $B_{SkyIndex}$, (YAMASHITA e YOSHIMURA, 2012, p 138):

$$B_{SkyIndex} = \frac{B - R}{B + R'} \quad (24)$$

- B_{Index} , (MISHRA et al., 2012, p 69):

$$B_{Index} = \frac{\sum_{i=216}^{i=256} Bi}{\sum_{j=236} B_j} \quad (25)$$

Onde R_i , G_i , B_i é o total de números de pixels em cada intensidade e R_j , G_j , B_j é o valor da intensidade.

2.1.2 Segmentação por Textura

2.1.2.1 Algoritmo LBP (*Local Binary Pattern*)

Na segmentação entre céu e chão várias abordagens existentes na literatura usam características de texturas (BEBIS et al. (2014); CROON et al. (2011)). Para esta pesquisa optou-se pelo algoritmo LBP (*Local Binary Pattern*), método de descrição de textura em que a comparação do ponto central de uma região delimitada (normalmente 3x3) com seus vizinhos gera um código para substituir o valor do ponto central (OJALA; PIETIKAINEN; MAENPAA, 2002). Esse processo depende de um limiar, de acordo com a equação 26.

$$s | x | = \begin{cases} 1, & Px > P \\ 0, & \text{senão} \end{cases} \quad (26)$$

onde Px corresponde ao ponto central e P ao ponto vizinho.

O método consiste em:

- Comparar cada *pixel* (chamado de *pixel* central) na imagem com seus 8-vizinhos;
- Se o *pixel* central for maior que seu vizinho no círculo escreva '1', senão, escreva '0';

- c. Com isso, tem-se um código com 8 números '0' ou '1', ou seja, um número binário de 8 bits;
- d. Transformando este número binário para decimal obtemos um valor entre 0 e 255;
- e. O descritor é então composto por estes valores decimais atribuídos a cada *pixel* formando a nova imagem.

A representação do método é mostrada na Figura 2.1.

104	101	96	1	1	0
97	100	6	0		0
102	110	54	1	1	0

Figura 2.1: Processo de construção de um código LBP.

2.1.2.2 Perfil Morfológico Estendido (*Extended Morphological Profile*)

O perfil morfológico estendido (*Extended Morphological Profile*) por LIANG, ZHOU e GAO (2016) consiste em aplicar abertura e fechamento como elemento estruturante de tamanho crescente para cada pixel também para gerar um vetor de características, de acordo com as equações 27 e 28:

$$\Omega^{(n)}(I) = [\gamma^{(n)}(I), \dots, \gamma^{(1)}(I), I, \phi^{(1)}(I), \dots, \phi^{(n)}(I)] \quad (27)$$

Onde $\gamma^{(n)}(I)$ e $\gamma^{(1)}(I)$ são a abertura e o fechamento com um elemento estruturante em forma de disco de tamanho n . Em seguida, os perfis morfológicos são obtidos em cada um dos componentes primários p conforme a equação 28:

$$\hat{\Omega}_p^{(n)}(I) = [\Omega_1^{(n)}(I), \Omega_2^{(n)}(I), \dots, \Omega_p^{(n)}(I)] \quad (28)$$

No último passo, os perfis morfológicos são empilhados com a resposta espectral para formar a característica espectral-espacial.

2.1.3 Segmentação com características de Média/Variância/Obliquidade/Curtose/Energia/Entropia

Características encontradas na literatura para este tipo de abordagem também foram utilizadas nesta pesquisa, são elas:

- a) A **Média** (μ) permite caracterizar o comportamento local e reduzir a influência do ruído. Nos casos de homogeneidade absoluta da vizinhança, a média é igual ao valor do pixel, sendo N o número de pixels da região. A Equação 29 apresenta o cálculo da média μ para cada canal R , G e B .

$$\mu = \frac{1}{N} \sum_{i=1}^N x_i \quad (29)$$

- b) A **Variância** (σ) fornece a noção da distribuição ao redor do pixel. Permite a caracterização do aspecto homogêneo, junto com a detecção da influência do ruído. A Equação 30 apresenta o cálculo da variância σ para cada um dos nove canais $j=R, G, B, H, S, V, Y, Cr, Cb$.

$$\sigma_j^2 = \frac{1}{N} \sum_{i=1}^N (x_{(i,j)} - \mu_j)^2 \quad (30)$$

- c) O **Skewness** (γ) é definido como o quociente do momento de terceira ordem e a terceira potência do desvio padrão. A Equação 31 apresenta o cálculo do *Skewness* γ para cada um dos nove canais $j=R, G, B, H, S, V, Y, Cr, Cb$.

$$\gamma_j = \frac{1}{N\sigma^3} \sum_{i=1}^N (x_{(i,j)} - \mu_j)^3 \quad (31)$$

- d) A **Curtose** (Cu) é definida como o quociente do momento de quarta ordem em torno da média μ e a quarta potência do desvio padrão (σ). A Equação 32 apresenta o cálculo da curtose **Cu** para cada um dos nove canais $j= R, G, B, H, S, V, Y, Cr, Cb$.

$$cu_j = \frac{1}{N\sigma^4} \sum_{i=1}^N (x_{(i,j)} - \mu_j)^4 \quad (32)$$

- e) A **Energia** mostra como o nível de cinza é distribuído. O valor máximo é 1 e diminui na medida que o valor do pixel tende ao cinzento. A energia é definida como mostra a equação 33:

$$Energia = \sum_{g=0}^{l-1} (p(g))^2 \quad (33)$$

- f) A **Entropia** é usada para reduzir o número de bits necessários para armazenar as imagens e é definida pela fórmula apresentada na equação 34:

$$Entropia = - \sum_{g=0}^{l-1} P(g) \log_2[P(g)] \quad (34)$$

2.2 Classificador SVM (*Support Vector Machine*)

Em função das melhores taxas de acerto apresentadas pelo classificador SVM e por este ser uma abordagem de classificação usada em muitos estudos descritos na literatura, SVM foi o classificador escolhido para o método proposto.

O SVM (Máquina de Vetor de Suporte), de acordo com Scarpel (2005), é um procedimento construtivo universal de aprendizagem baseado na teoria de aprendizagem estatística.

SVMs minimizam o risco estrutural, ou seja, a probabilidade de classificar de forma errada padrões ainda não vistos por uma distribuição de probabilidade dos dados (LIMA. 2002).

Segundo Osuna, Freud e Girosi (1997), o objetivo é separar as classes com superfícies que maximizem a margem entre elas. Um conjunto de dados de duas classes sendo w_1 e w_2 , linearmente separáveis com margem máxima δ ou também chamado de hiperplano de separação ótima, como mostra a Figura 2.2.

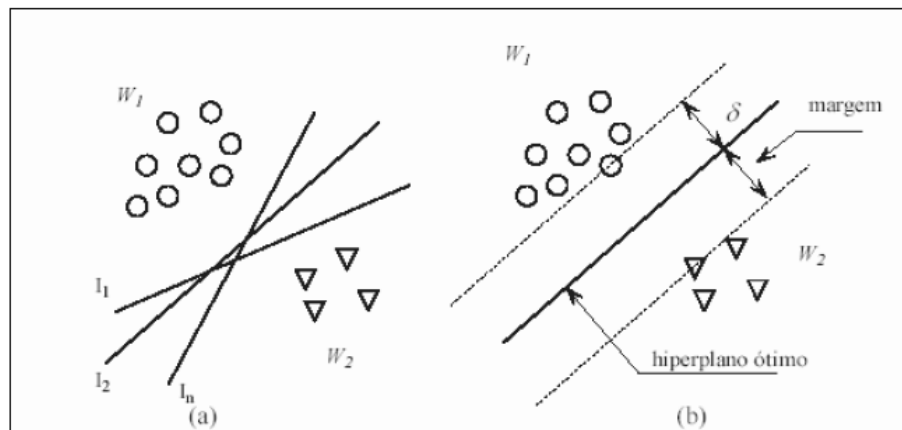


Figura 2.2: Classificação entre duas classes w_1 e w_2 usando hiperplanos.

FONTE: Adaptado de Osuna, Freud e Girosi (1997).

Na Figura 2.2 (a) as classes w_1 e w_2 são classificadas usando hiperplanos arbitrários enquanto na Figura 2.2 (b) é usado o hiperplano com separação ótima.

Para se chegar à superfície de decisão ótima o algoritmo de treinamento SVM separa da melhor forma possível os pontos dos dados de ambas as classes. Os pontos mais próximos dos limites entre as duas classes são selecionados para ajudar a definir a melhor superfície de decisão que outros pontos.

2.3 Trabalhos Relacionados

A literatura revisada relacionada ao tema apontou abordagens distintas, separando estas em três principais: uma que utiliza classificadores para o auxílio na separação de céu e chão, outra baseada em imagens de cinza/cor e, por fim, uma abordagem usando detecção de bordas.

2.3.1 Abordagens usando classificadores

Em Croon et al. (2011) é apresentada uma abordagem baseada na segmentação através do algoritmo J48. Os autores usaram uma base de dados pública – labelME¹ que contempla 7.456 imagens contendo céu. A base inclui

¹ B.C. Russell, A. Torralba, K.P. Murphy, and W.T. Freeman. Labelme: a database and web-based tool for image annotation. International Journal of Computer Vision, 77(1–3):157–173, 2008.

imagens urbanas, de mar e de floresta, todas misturadas, 10% das imagens foram usadas para teste e os outros 90% para treinamento. Para cada imagem são extraídas 45 amostras. Esses dados são extraídos em uma matriz $250.000 \times n$, sendo n o número de características. Em seguida, essa matriz é usada no Weka (kit de ferramentas de aprendizagem de máquina) aplicando o algoritmo J48. Os autores usam 34 características gradiente e texturas extraídas para cada pixel durante a fase de treinamento. Os métodos de segmentação aplicados no conjunto de testes das imagens são a classificação do pixel usando valores dos canais H , S e V , a segmentação de Otsu no canal B da imagem RGB , o filtro *Leung-Malik* para calcular a localização da linha do horizonte e vários algoritmos J48 que foram modificados para a abordagem chamados de *HSV-tree*, *full-tree*, *tree without y-coordinate* e *BF-tree*. Nos resultados da segmentação, os autores relatam que o método que obteve o máximo de precisão segue a ordem de aplicação dos algoritmos: árvore com *y-coordinate* seguida da árvore *BF-tree/HSV-tree*. Os autores não fornecem a taxa de acerto da abordagem e o método de avaliação.

Os autores Bebis et al. (2013) apresentam um método para a segmentação de céu através da fusão do *K-means* com Redes Neurais. A saída é 1 para céu e 0 para não céu. Esses resultados são então pós-processados e fundidos com os resultados obtidos usando o algoritmo *k-means*. Após essa fusão, uma segunda etapa de pós-processamento é realizada produzindo uma máscara válida em que 1 é a região de céu e o restante é marcado como 0. As redes neurais usadas nessa abordagem consistem em duas camadas de rede *back-propagation* com 16 entradas, 20 nós na camada oculta e um nó de saída. A rede neural foi treinada usando o método de gradiente descendente com janela 9×9 para imagens selecionadas aleatoriamente a partir do conjunto de dados. O treinamento usa intensidades de pixel e 15 características de textura, entre elas, entropia, energia, variância e etc. Na etapa de fusão, o algoritmo verifica se o tamanho da intersecção entre cada região obtida utilizando *K-means* e a região correspondente determinado como o céu nos resultados pós-processados da rede neural constitui uma determinada parte das regiões do *K-means*. Se esta condição se mantém, o algoritmo marca toda a região como o céu. O algoritmo utiliza um limiar para avaliar o tamanho relativo da intersecção. O valor padrão é $th = 0,5$. No trabalho, os autores usam dois conjuntos de dados: 45 imagens em tons de cinza da base NASA

e 80 imagens montanhosas extraídas aleatoriamente da web. A abordagem é comparada com Lie, Lin e Hung (2005). Nos dois casos, as linhas obtidas são comparadas com as ideais calculando a distância absoluta S pela fórmula:

$$S = \frac{1}{N} \sum_{j=1}^N |L_{d(j)} - L_{g(j)}| \quad (35)$$

A média e o desvio padrão são calculados para todas as imagens do conjunto de dados, conforme tabela 2.3, mostrando o valor de erro médio menor para a abordagem proposta.

Tabela 2.3: Cálculo de erros médios absolutos na comparação dos métodos.

Abordagem	NASA Basalt Hill		Web	
	Média	Desv. Pad.	Média	Desv. Pad.
Lie, Lin e Hung (2005).	5.5548	9.4599	9.1500	17.9195
Proposta	1.6077	2.8493	4.2964	6.1581

FONTE: Adaptado de Bebis et al. (2013).

Na abordagem proposta por Verbickas e Whitehead (2014), as redes neurais convolucionais (CNN) foram treinadas para detectar regiões de céu e não céu, bem como a linha do horizonte em frames de vídeos. O método proposto usa CNN para realizar a detecção de céu gerando uma imagem binária onde os pixels brancos são classificados como céu e os pixels pretos como chão. Uma base de um total de 13.687 imagens foi construída baseada em frames de vídeos especificamente para esta finalidade. No trabalho os autores usam a medida do erro quadrado comum para a função de perda de rede e também o cálculo de aproximação diagonal de *Gauss-Newton* para a matriz *Hessian*, a fim de calcular as taxas de aprendizagem mais eficientes para gradiente descendente. Finalmente, considera-se a imagem de entrada como tendo um ou mais canais de entrada em algum espaço de cor, enquanto que imagens de saída são binários. Os autores usam o erro quadrado como métrica para a condução de treinamento da rede. Para realizar a avaliação da rede após o treinamento, incluindo testes, usam uma matriz de confusão. A estrutura implementada para a construção da CNN foi uma GPU CUDA. Os autores observaram que as CNNs, na diferença média, superam SVMs de 1,27%.

Os autores Porzi, Buló e Ricci (2016) introduzem uma nova abordagem para a detecção de linha, que adere a um paradigma de aprendizagem e exploram o poder de representação de arquiteturas profundas para melhorar a precisão de detecção de linha de horizonte. Os autores exploram uma nova arquitetura deconvolucional, que introduz níveis intermediários de supervisão para apoiar o processo de aprendizagem. No método dos autores a linha do horizonte é detectada pela função $h: \mathcal{X} \rightarrow \mathcal{Y}$ que torna uma imagem $X \in \mathcal{X}$ e retorna uma máscara que representa a probabilidade $Y \in \mathcal{Y}$ de pertencer à linha do horizonte com o mesmo tamanho da imagem de entrada. Consequentemente, o espaço de entrada \mathcal{X} consiste da $h \times w$ RGB de imagens, *i.e.* $\mathcal{X} = \mathbb{R}^{h \times w \times 3}$, enquanto o espaço de saída é dado pela máscara de probabilidade $h \times w$, *i.e.* $\mathcal{Y} = [0,1]^{h \times w}$. Neste trabalho os autores propõem modelar o detector h como uma rede neural profunda que combina camadas convolucionais com camadas deconvolucionais para produzir toda a máscara de detecção em saída dada a imagem como entrada. Em vez de considerar uma única camada de saída, os autores adicionam saídas intermediárias no bloco deconvolucional a fim de proporcionar uma supervisão mais profunda durante a fase de treinamento. A rede é treinada a partir de imagens rotuladas $\{(X_i, Y_i)\}_{i=1}^n$, onde $X_i \in \mathcal{X}$ e a i -ésima imagem com a correspondente *groundtruth* dada por $Y_i \in \mathcal{Y}$. Os autores usaram o conjunto de dados CH1² publicamente disponível que compreende 203 imagens de paisagens montanhosas. Nos experimentos, as imagens foram divididas aleatoriamente o conjunto de dados em 60% para treinamento e 40% para teste. O método proposto pelos autores são comparados com outras abordagens que utilizam aprendizado e, segundo os autores, superam os métodos tradicionais com erro médio = 1.53. Os autores não informam as métricas que determinam as taxas de acerto da abordagem.

Os autores Boroujeni, Etemad e Whitehead (2012), propõem uma técnica para a detecção da linha do horizonte baseado na existência de um único campo de luz. A abordagem emprega a segmentação da imagem e a análise dos segmentos da imagem para a extração da linha do horizonte. O método utiliza gradientes de intensidade de luz próximos à linha do horizonte para detectar o caminho exato da

² Disponível em < <https://cvg.ethz.ch/research/mountain-localization/> >

região onde o céu se encontra com o chão. Para detectar as regiões descritas perto do caminho exato da linha do horizonte em imagens, um algoritmo de agrupamento é usado como ferramenta principal. Antes do agrupamento, um pré-processamento da imagem em tons de cinza é realizado pelo filtro de *Gauss* a fim de minimizar os efeitos do ruído, variações de intensidade de pixel, e remoção de bordas falsas e inválidas. A próxima etapa é usar um algoritmo de agrupamento para detectar a região de interesses (ROI). Dois tipos de agrupamento são empregados para este propósito: um agrupamento baseado na intensidade que é realizada através da divisão de todo o espectro de intensidade em um número menor de valores através da quantização e o agrupamento por *k-means* em que a classificação baseia-se na semelhança de cada objeto para com o ponto central de cada cluster. O resultado dos dois métodos é semelhante em termos de produto final. A técnica com agrupamento por intensidade foi aplicada a 160 quadros de vídeos que representam imagens diferentes (céu e mar/floresta) a uma taxa de erro de 1,2%. Em geral, o tempo computacional médio de agrupamento baseado em intensidade é menos de 1,5 segundos, enquanto para *k-means* é de aproximadamente 10 segundos. No entanto, os autores não esclarecem de que forma avaliaram os resultados em termos de percentuais.

Mihail et al. (2016) apresentam uma avaliação de três métodos existentes com um conjunto de dados do mundo real. Para construir o conjunto de dados, os autores aproveitam imagens coletadas de câmeras estáticas ao ar livre durante longos períodos de tempo, aumentando assim a probabilidade de exposição a diferentes condições atmosféricas. Usando este conjunto de dados, vários experimentos foram realizados para analisar o impacto que o clima e o tempo têm sobre cada desempenho dos métodos. As imagens foram coletadas de 53 câmeras a partir de muitas cenas ao ar livre. Nos experimentos, cada imagem do conjunto de dados foi processada através dos métodos de Hoiem, Efros e Hebert (2005), Lu et al. (2014), Tighe e Lazebnik (2010), usando o código fonte fornecido pelos respectivos autores. Os parâmetros ajustáveis utilizado nesta avaliação foram os mesmos que os autores originais usaram nas suas próprias avaliações. Os autores avaliam os métodos com uma métrica de desempenho por pixel, a taxa de erro de classificação (MCR). Os dados foram divididos em um conjunto de treinamento de

40 câmeras e 13 câmeras de teste. No geral, o conjunto de dados supera os três métodos avaliados com média de MCR de 12.96%.

Outra abordagem proposta por Ahmad et al.(2014) usa várias características de textura locais como SIFT (*Scale Invariant Feature Transform*), LBP (*Local Binary Patterns*), HOG (*Histogram of Oriented Gradients*) e suas combinações SIFT-LBP, SIFTHOG, LBP-HOG e peneira-LBP-HOG prevendo 128, 58 e 31 vetores dimensionais para treinar o classificador SVM. Usando programação dinâmica, os autores propõe um uso duplo do classificador: primeiro distinguir entre bordas de horizonte e não horizonte e segundo para proporcionar uma confiança sobre e pixel da borda de horizonte. O trabalho apresenta a criação manual das imagens *groundtruth* para o conjunto de 45 imagens *Basalt Hills*. Descrições de textura e suas combinações como características para treinar o classificador SVM. A validação é feita com 9 imagens para treinamento e 36 imagens para teste.

Para comparar as linhas do horizonte detectadas por cada método com as *groundtruth* calcula-se o erro médio absoluto, pixel a pixel, de acordo com a tabela 2.4.

Tabela 2.4: Erros absolutos médios entre os horizontes detectados e *groundtruth* para várias opções de custos nodais.

Custo nodal	Erro Absoluto	
	Média	Desv. Padrão
Lie, Lin e Hung (Edges)	5,554	9,459
Informações de gradiente	3,990	6,353
Bordas SIFT+HOG	0,578	1,022
Pontuações SIFT+HOG	0,412	0,812
Pontuações SIFT+HOG + Informações de gradiente	0,435	0,9812

FONTE: Adaptado de Ahmad, Bebis, Regentova e Nefian (2014).

Os autores Ahmad et al. (2014) concluem dizendo que o uso dos custos nodais propostos com informações de gradiente contribui mais que a simples detecção de bordas.

2.3.2 Abordagens em níveis de cinza/cor

Segundo os autores Hugues, Cieutat e Guitton (2013), o objetivo é filtrar as imagens para remoção dos ruídos para depois detectar o horizonte. Em seguida, apresentam uma avaliação dos algoritmos de desempenho no que diz respeito à taxa de detecção do horizonte e o impacto da resolução da imagem no tempo de execução. A aquisição das imagens é realizada por uma câmera de vídeo clássica ou térmica. A análise inclui a normalização de histograma, filtro de convolução de *Gauss*, fechamento morfológico, laplaciano, transformada de *Hough* para detectar as linhas da imagem. A fase seguinte consiste em analisar estatisticamente as linhas detectadas pelo método descrito. Calcula-se então, a média e o desvio padrão para a posição de cada extremidade dos segmentos de grupo, segmentos próximos ou justapostos. Finalmente, os segmentos detectados são filtrados de acordo com a inclinação na imagem. Todos os segmentos cuja inclinação na imagem tem uma diferença de mais de 5% do valor são eliminados. No artigo 13.062 imagens são extraídas a partir de sequências de vídeo e testadas. O algoritmo detecta a linha do horizonte em cerca de 98,45% dos casos. O cálculo de tempo para a rotina de extração horizonte atinge uma média de 6,5 milissegundos para uma resolução de 640x480. Os autores não informam como a avaliação desse percentual de 98,45% foi realizada, se foi de forma visual ou com aplicação de critérios matemáticos.

Herdtwack e Wallraven (2010) focaram seus estudos em dois experimentos: um perceptivo e outro computacional. O experimento perceptivo se baseia na capacidade humana de analisar imagens diversas e detectar visualmente a linha do horizonte. No experimento computacional foi analisado o desempenho de vários algoritmos sendo que o melhor resultado foi o obtido através do método de gradiente. Usando as mesmas imagens da experiência perceptiva (quantidade de imagens não informada), na análise computacional, as imagens foram decompostas para o espaço de cor CIE $L^*a^*b^*$. Após os experimentos com os algoritmos implementados, os autores informam que nenhum algoritmo computacional foi capaz de coincidir totalmente com o desempenho humano, eles estavam fora por pelo menos 4 a 5% da altura da imagem com as suas estimativas comparadas com as anotações humanas.

Na abordagem de Ananth (2014), o autor propõe a segmentação baseada em OWT (*stationary wavelet transform*) para imagens coloridas. OWT extrai as características *wavelet* (leves ondulações) que dão uma boa separação dos diferentes padrões. Além disso, o algoritmo proposto utiliza operadores morfológicos para a segmentação eficaz. Na segmentação baseada em OWT cada imagem original é dividida em quatro subimagens e, em seguida, submetido à transformada *wavelet* em que as imagens de saída são do mesmo tamanho que a imagem original. O algoritmo proposto é aplicado sobre a variedade de imagens coloridas e testado em vários espaços de cores. Inicialmente a cor da imagem é dividida em três planos (R, G, B) e histograma 2D RG, RB, GB são calculados. Em seguida, o histograma é alisado por filtro de *Gauss* com um desvio padrão de 0,625. A erosão morfológica é aplicada sobre o histograma alisado que extrai diretamente o cluster centroides, isto é, picos dominantes no histograma 2D. Estes centroides são rotulados e a transformação do histograma 2D é aplicada para fornecer o histograma agrupado. A partir do histograma agrupado o mapa de segmentação é obtido a partir de mapeamento simples. Em seguida, a transformação de concordância é realizada de forma a uniformizar os mapas de segmentação. Esses mapas são então filtrados para gerar o resultado final da segmentação. A principal vantagem do método segundo os autores é o uso de parâmetro único e também muito mais rápido. Ao se comparar com cinco espaços de cor, o esquema de segmentação produz resultados visivelmente melhores no espaço de cor RGB em comparação com todos os outros espaços de cores. Os autores não informam quais os outros espaços de cores que foram testados, como também não citam as taxas de acertos e taxas de erros.

Kulkarni (2012) apresenta uma técnica de limiar por cor se baseia na adaptação e ligeira modificação do algoritmo de limiar por nível de cinza. As propriedades dos pixels RGB são estudadas para extrair as características importantes a partir da imagem, por exemplo, se o interesse é áreas verdes (chamado de Floresta), com base na informação de cor, o algoritmo de limiar por cor deve ser capaz de extrair os pixels de cor verde e rejeitar os pixels de outros objetos. Foi calculado a gama de intensidades RGB para cores verdes e azuis, e depois se aplicou o algoritmo de limiar. Para que o limiar exato possa ser obtido, os valores mínimos e máximos para cada um dos componentes RGB são obtidos para

a área azul e verde. Esses valores mínimo e máximo são usados para extrair características importantes dos valores RGB que são convertidos em valores de limiar, como mostram as tabelas 2.5 e 2.6.

Tabela 2.5: Informação de RGB na área VERDE.

	Mínimo	Máximo
Red	0	173
Green	102	255
Blue	0	173

FONTE: Adaptado de Kulkarni (2012).

Tabela 2.6: Informação de RGB na área AZUL.

	Mínimo	Máximo
Red	0	90
Green	127	255
Blue	127	255

FONTE: Adaptado de Kulkarni (2012).

Uma vez que a determinação do melhor resultado só pode ser feita por observação humana, várias tentativas têm sido feitas de modo que os resultados podem ser comparados para selecionar os melhores valores para o algoritmo de limiar. A fim de determinar se o método de limiar que foi realizado é bem sucedido ou não, os autores se baseiam unicamente na intervenção humana.

2.3.3 Abordagens usando Detecção de Bordas

Lipschutz, Gershikov e Milgrom (2013), propõem um algoritmo baseado na detecção de bordas e na transformada de *Hough* (H-EDHT): o novo algoritmo baseia-se na ideia de segmentar a imagem em duas regiões (céu e mar), comparando as funções regionais de distribuição de probabilidade (PDFs) ou histogramas. A linha de maximização de uma distância significativa entre estas funções regionais PDF é escolhida. Assim, passa-se a introduzir dois novos métodos que combinam a detecção básica de borda e transformada de *Hough* (EDHT) para detectar várias linhas candidatas à imagem e um critério estatístico para encontrar a linha ideal entre elas. O critério pode ser escolhido com base em covariâncias

regionais ou na distância entre as funções regionais PDF. Mostra-se que a seleção da primeira opção proporciona o melhor desempenho entre os métodos testados, pelo menos no que diz respeito ao erro angular, e produz um algoritmo rápido e de baixa complexidade. Os métodos foram analisados não só quantitativamente pela sua precisão média e velocidade relativa, mas também visualmente para várias imagens de teste. O estudo foi aplicado para um grupo de 10 imagens infravermelhas marinhas. Os autores não informam as taxas de acerto e as taxas de erro.

O trabalho proposto em Ahmad et al.(2013), usa a aprendizagem de máquina para detecção da linha do horizonte em imagens em escala de cinza. A ideia é extrair *Maximally Stable Extremal Edges* (MSEE) inspirada na ideia de extração de regiões de bordas maximamente estáveis. As bordas das imagens em escala de cinza são obtidas através do detector de bordas de *Canny* com o parâmetro sigma a ser fixado variando o limiar. Os resultados são a geração de N imagens binárias assumindo N combinações de valores dos parâmetros denominados I_1 para I_N . A localização de um pixel (x, y) da aresta é considerada estável se for detectado como um pixel de borda para k valores de limiar consecutivos. A imagem formada por essas arestas estáveis se referem a borda da imagem denotada por E . Matematicamente mostrada pela equação 36:

$$E(x, y) = \begin{cases} 1, se \sum_{i=1}^N I(x, y) > k. \\ 0, senão. \end{cases} \quad (36)$$

FONTE: Adaptado de Ahmad et al. (2013).

Para treinar o classificador SVM rotularam-se manualmente os pixels da linha do horizonte nas imagens de treinamento utilizando o MSEEs. O classificador foi usado no Weka e foram selecionados pixels de MSEE para treinamento aleatoriamente para cada imagem. O conjunto de dados usado na abordagem proposta é composta de 10 imagens em escala de cinza com variação de brilho e textura. O uso do MSEE, como pré-processamento, reduz em até 92% as arestas não horizonte. Para cada imagem em escala de cinza do conjunto de dados aplicou-

se o detector de bordas de *Canny*. Porém, os autores citam uma desvantagem da abordagem em relação ao pressuposto de que a linha do horizonte é reta.

Gershikov (2014) apresenta em seu estudo dois grupos de métodos e compara o desempenho de algoritmos baseados em cor ao de métodos baseados em luminância. Os métodos baseados em intensidade podem ser divididos em duas categorias: métodos usando características locais, tais como bordas, média de valores locais/mediana e gradientes e métodos usando características globais, tais como as estatísticas regionais. O método baseado em características locais baseia-se em uma fase de pré-processamento, por exemplo, utilizando a detecção de borda por *Canny* e a transformada de *Hough*. Esse método é chamado de GEDHT, que significa "Detecção de Bordas em tons de cinza e Transformada de *Hough*". Os algoritmos que pertencem aos métodos de características globais geralmente usam um critério para calcular a otimização para diferentes posições de linhas candidatas usando toda a imagem ou grande parte dela. Em seguida, a linha do horizonte candidata é escolhida pelo valor ótimo do critério utilizado. Os algoritmos descritos foram aplicados para um grupo de 21 imagens coloridas, a maioria delas imagens de mar. As imagens *groundtruth* foram geradas manualmente, em seguida, as seguintes medidas de precisão foram definidas a partir da equação 37:

$$DesviodeAltura(\%) = \frac{|AlturaDetectada - AlturaReal|}{NumLinhasIg} \quad (37)$$

O algoritmo mais preciso é o CMVDHT, baseado na detecção de borda cor usando estatísticas de vetores de ordem e a transformada de *Hough* (onde *C* representa a cor e *HT* significa Transformada de *Hough*).

Os autores Sassi, Amar e Miguet (2016) propõem uma descrição geométrica do horizonte extraído das paisagens naturais e de cidade. A primeira abordagem proposta foi a classificação das linhas retas para diferenciar as paisagens urbanas e naturais da linha do horizonte. A segunda é a análise de curvatura usando um descritor de espaço de escala de curvatura. Os edifícios e torres em paisagens são caracterizados por longas linhas retas, especialmente quando há uma torre alta. Assim, a ideia é quebrar o horizonte em segmentos através de um algoritmo de aproximação poligonal. O resultado deste processo é um histograma que apresenta o número de segmentos em cada categoria de segmentos de comprimento. Os

resultados obtidos na abordagem proposta não foram comparados com nenhuma abordagem existente e as taxas de acerto não foram informadas.

No algoritmo proposto por Ayadi et al. (2016), a linha do horizonte pode ser extraída para contextos diferentes, dependendo do interesse do usuário (horizonte distante, edifícios intermediários, construções próximas e etc.), e requer assim uma interação humana. Os principais passos do método são: aplicação de bordas de Canny para extrair bordas e permitir ao usuário interagir com os parâmetros. Um gráfico é então construído e um algoritmo de caminho mais curto é usado para ligar descontinuidades. Visão geral do método proposto: a imagem original é transformada em escala de cinza e aplicado o filtro de Canny. Os autores usam inicialmente duas bases de dados (nomeadamente CH1 e CH2). Outro banco de imagens usado contém imagens originais e *groundtruth*. As imagens das três bases que corresponde a 320 imagens de natureza e cidade são usadas para testar o algoritmo proposto pelos autores. Para a avaliação da linha do horizonte extraída, inicialmente, define-se um índice de descontinuidade, que se refere a distância do horizonte extraído em relação a sua *groundtruth* usando a formulação apresentada na equação 38:

$$dist = \frac{1}{w} \sum_{i=0}^w |Sg(k) - Se(k)| \quad (38)$$

Onde w é igual a largura da imagem, Sg é a *groundtruth* e Se é a linha extraída.

Em seguida, calcula-se a média de *Precision* e *Recall*, como mostra a tabela 2.7, em que a precisão varia entre 96 e 99%.

Tabela 2.7: Taxas de acerto da comparação das bases de imagens.

	CH1	CH2	Barcelona
Número de imagens	175	80	104
Precisão (média)	0,999	0,988	0,968
Revocação (média)	0,99	0,989	0,75

FONTE: Adaptado de Ayadi, Suta, Scuturici, Miguet e Amar (2016).

Os autores Chiu et al. (2016) propõem um algoritmo para detectar o contorno de uma linha do horizonte curva em uma imagem sob diferentes brilhos e contrastes. Os experimentos são realizados com 40 imagens de diferentes condições climáticas e ambientais. O método proposto pelos autores consiste em dois módulos: a linha do horizonte candidata por análise de ponto (SCPA) e a linha do horizonte candidata por ponto de conexão (SCPC). O módulo SCPA analisa os pixels em uma coluna vertical definindo amostras de colunas da imagem. Para identificar as características de linha do horizonte sob diferentes condições de brilho e contraste, o algoritmo analisa o gradiente de intensidade das colunas da amostra. Depois de extrair os pontos candidatos de linha do horizonte em cada coluna de amostra, o módulo SCPC compara o gradiente de intensidade entre os candidatos. Segundo os autores, o algoritmo detecta a linha do horizonte para imagens em diferentes condições ambientais e climáticas, mas não informam as taxas de acerto e erro. Também não utilizam métricas para avaliar a abordagem proposta.

Liang et al. (2015) propõem um método que consiste em duas etapas: A primeira consiste em transformar a imagem original em escala de cinza e a região da linha de margem é localizada usando a matriz de co-ocorrência baseada no contraste da característica de textura. Na segunda etapa, os limiares de segmentação adaptativos são obtidos pelo algoritmo de Otsu, que produz um conjunto de pontos candidatos à linha do horizonte. Depois um método de agrupamento simples é usado para selecionar pontos apropriados e transformá-los em linha reta. O conjunto de imagens construído para a abordagem consiste em 200 imagens de mar divididas em quatro categorias: céu com nuvens, mar com ondas, céu com muita luz e imagens consideradas normais, com céu limpo. As imagens que contemplam céu com nuvens e grandes variações de luz tem uma precisão menor do que as imagens consideradas normais. A abordagem proposta pelos autores desconsideram os obstáculos e definem somente um horizonte reto. A precisão é de 93% segundo os autores.

Outra abordagem apresentada por Williams e Howard (2011) usa um esquema de detecção de linha do horizonte utilizando várias dicas visuais para classificar linhas candidatas do horizonte. Primeiro, um conjunto de dados estatísticos de ordem não paramétricos é gerado dentro de uma região S construindo um histograma de intensidade de pixel por região, a intensidade média,

os valores dos quartis, e um intervalo de 95% ao redor do centro podem ser eficientemente calculados. Um detector de bordas de *Canny* é usado para encontrar bordas dominantes da imagem. O comprimento do segmento mínimo é aplicado para remover o ruído da imagem. Os segmentos de linha candidatos que exibem sinais visuais fracos servem para reforçar o caminho dos segmentos fortes, enquanto que os segmentos com sinais visuais fortes têm a capacidade de redirecionar a linha do horizonte. O segmento com o menor custo torna-se a linha do horizonte detectada, e o processo se repete até que a borda da imagem seja alcançada. Para testar o tempo de execução de cada algoritmo, uma única imagem de 640x480 foi carregada para a memória. Cada algoritmo é então processado 1000 vezes para a imagem de teste e o tempo de processamento foi gravado. Os autores não fornecem o tamanho exato da base e também não citam nenhuma métrica computacional.

2.4 Análise dos Trabalhos Relacionados

As abordagens existentes na literatura apresentam vários algoritmos supervisionados e não supervisionados, de detecção de bordas usados na detecção da linha do horizonte, porém, várias delas usam imagens em níveis de cinza para classificação ou agrupamento e não a característica de cor. As avaliações dos métodos apresentados referem-se, na maioria, em termos de velocidade e desempenho. Na maioria das abordagens, não é apresentada as taxas de acerto e as taxas de erro, e também, o detalhamento da avaliação utilizada, o que torna difícil uma análise de comparação entre os métodos. As bases usadas em vários estudos são, em geral, muito pequenas e que não torna a comparação relevante. As abordagens revisadas são apresentadas na Tabela 2.8.

Nesse contexto, o método proposto nesse trabalho busca aplicar características ainda não usadas para esse tipo de abordagem, a partir de imagens coloridas em uma base de imagens maior e com grande variação de céu e chão, e avaliar o método através da métrica *F-measure*.

Tabela 2.8: Comparação dos métodos.

Artigo	Base de Dados	Tipo de imagens	Pré-processamento de imagens	Segmentação de imagens	Classificadores/ Algoritmo	Avaliação dos métodos	Resultados
Abordagens usando classificadores							
CROON, G.C.H.E., DE WAGTER, C., REMES, B.D.W., RUIJSINK, R., "Sky Segmentation Approach to Obstacle Avoidance.", IEEE Aerospace Conference 2011, p. 1 - 31, 2011.	Base de dados publica – labelME contendo 7.456 imagens contendo céu. A base inclui imagens urbanas, de mar, floresta e etc. 10% das imagens foram usadas para teste e os outros 90% para treinamento. Usa base <i>ground-truth</i> labelME.	Imagens em níveis de cinza.	-	- Segmentação de Otsu; - Filtro Leung-Malik para calcular a localização da linha do horizonte; - Árvore J48 com o nome HSV-tree treinada com características HSV; - Árvore J48 com o nome full-tree treinada com as características extraídas das imagens de treinamento; - Árvore J48 chamada de tree without y-coordinate; - Árvore J48 BF-tree treinados com características extraídas de imagens de YCbCr.	Algoritmo J48 no Weka.	Características extraídas não avaliadas a um nível de pixel, mas ao nível de segmentos de cor na imagem.	Nos resultados da segmentação o método que obteve o máximo de precisão segue a ordem: Árvore com y-coordinate → BF-tree/HSV-tree. Os autores não fornecem as taxas de acertos e erros, bem como, o método de avaliação.
BEBIS, George; YAZDANPANA, Ali Pour; REGENTOVA, Emma E.; MANDAVA, Ajay Kumar; AHMAD, Touqeer. Sky Segmentation by Fusing Clustering with Neural Networks. (Eds.): ISVC 2013, Part II, LNCS 8034, pp. 663–672, 2013.	Base de imagens da NASA com 45 imagens + 80 imagens montanhosas extraídas da web.	Imagens em níveis de cinza.	-	As imagens de entrada para classificação utilizando redes neurais incluem valores de pixel e características com janela 9x9. Pós-processamento + fusão com k-means. Depois, uma segunda etapa de pós-processamento.	Redes neurais com método de gradiente. K-means.	A etapa de fusão verifica a interSeção das regiões obtidas entre k-means e redes neurais.	As taxas de erros em comparação com a abordagem de Lie et al. (2005) para a abordagem proposta é de 1.60 para 5,55.
VERBICKAS, Rytis; WHITEHEAD, Anthony. Sky and Ground Detection Using Convolutional Neural Networks. Proceedings of the International Conference on Machine Vision and Machine Learning Prague, Czech Republic, August 14-15, 2014.	Base de 13.687 imagens baseadas em frames de vídeos construídos para o trabalho.	Imagens coloridas.	-	Redes Neurais Convolucionais que segmentação a imagem separando o céu (pixels brancos) do chão (pixels pretos).	Redes Neurais Convolucionais com partilha de pesos.	- Ajustes de parâmetros para a formação das redes neurais convolucionais incluindo duração no treinamento, profundidade da rede e contagem do mapa de características.	As redes neurais convolucionais superam as SVMs em 1,27%. A análise é realizada através da matriz de confusão.

<p>PORZI, Lorenzo; BULÒ, Samuel Rota; RICCI, Elisa. A Deeply-Supervised Deconvolutional Network for Horizon Line Detection. MM '16, October 15-19, 2016, Amsterdam, Netherlands c 2016 ACM. ISBN 978-1-4503-3603-1/16/10. DOI: http://dx.doi.org/10.1145/2964284.2967198.</p>	<p>Base CH1 com 203 imagens de paisagens montanhosas.</p>	<p>Imagens coloridas.</p>	<p>As imagens foram redimensionadas para uma resolução de 640 x 480.</p>	<p>Para a Rede Deconvolucional Supervisionada: o conjunto de dados foi dividido aleatoriamente em 60% para treinamento e 40% para teste</p>	<p>Rede Deconvolucional Supervisionada</p>	<p>A rede neural profunda que combina camadas convolucionais com camadas deconvolucionais para produzir toda a máscara de detecção em saída dada a imagem como entrada.</p>	<p>O método proposto pelos autores são comparados com outras abordagens que utilizam aprendizado e, segundo os autores, superam os métodos tradicionais com erro médio = 1.53. Os autores não informam métricas que determinam as taxas de acerto da abordagem.</p>
<p>BOROJENI, Nasim Sepehri; ETEMAD S. Ali; WHITEHEAD Anthony. "Robust Horizon Detection using Segmentation for UAV Applications" Ninth Conference on Computer and Robot Vision, pp 346 – 352, 2012.</p>	<p>Várias sequências de vídeos totalizando 160 quadros.</p>	<p>Imagens em níveis de cinza.</p>	<p>Borramento da imagem em tons de cinza pelo filtro de Gauss.</p>	<p>- Segmentação baseada na intensidade; - Segmentação baseada no agrupamento por k-means.</p>	<p>Método k-means.</p>	<p>Avaliação em termos de velocidade.</p>	<p>O tempo computacional médio por intensidade é menos de 1,5 segundos e baseado em k-means é de aproximadamente 10 segundos.</p> <p>O agrupamento baseado em intensidade é mais rápido.</p> <p>Os autores não informam o método de avaliação das taxas de acerto em nível de segmentação.</p>
<p>MIHAIL, R.P.; WORKMAN, S; BESSINGER, Z; JACOBS, N. Sky Segmentation in the Wild: An Empirical Study. 2016 IEEE Winter Conference on Applications of Computer Vision (WACV). 2016.</p>	<p>Total de 53 webcams com cerca de 1.500 imagens.</p>	<p>Imagens coloridas.</p>	<p>Comparação através dos métodos de Hoiem, Efros e Hebert (2005), Lu, Lin, Jia e Tang (2014), Tighe e Lazebnik (2010)</p>	<p>-</p>	<p>- Métodos de Hoiem, Efros e Hebert (2005), Lu, Lin, Jia e Tang (2014), Tighe e Lazebnik (2010); - Rede neural convolucional recorrente composta por três camadas convolucionais (Os dados foram divididos em um conjunto de treinamento de 40 câmeras e 13</p>	<p>- Os métodos são avaliados com uma métrica de desempenho por pixel, a taxa de erro de classificação (MCR).</p>	<p>O desempenho do algoritmo varia significativamente com as condições de iluminação.</p> <p>O método usando rCNN tem um desempenho melhor com MCR de 12.96%;</p>

					câmeras de teste).		
AHMAD, Touqeer; BEBIS, George; REGENTOVA, Emma; NEFIAN, Ara. An Experimental Evaluation of Different Features and Nodal Costs for Horizon Line Detection. ISVC 2014, Part I, LNCS 8887, pp. 193–205, 2014.	Total de 45 imagens do conjunto de dados Basalt Hills	Imagens coloridas.	Compara abordagens de detecção de linha do horizonte.	-	1ª abordagem: - calcula os pontos máximos extremos de borda estáveis (MSEE) para todas as imagens; - compara o número de arestas após a obtenção dos MSEE com o número de arestas encontrado pelo detector de bordas Canny. 2ª abordagem: - usa características de textura e suas combinações para treinar o classificador SVM. A validação é feita com 9 imagens para treinamento e 36 imagens para teste.	Avaliação referente ao custo nodal detectado por cada método comparado com as imagens <i>groudtruth</i> calculando-se o erro médio absoluto, pixel a pixel.	Os autores concluem dizendo que o uso dos custos nodais propostos com informações de gradiente contribui mais que a simples detecção de bordas com um erro médio absoluto de 3,99%.
Abordagens em níveis de cinza ou por espaços de cor							
HUGUES Olivier, CIEUTAT Jean-Marc, GUITTON Pascal. Merging Video and Charts for Augmented Reality by Vision and Inertial Fusion to Perform Horizon Tracking. Hal-00787072, version 1 – 11 Feb 2013.	Aquisição das imagens: vídeo de uma câmera clássica ou por câmera térmica. Total de 13.062 imagens.	Imagens em níveis de cinza.	- Normalização de histograma; - Filtro de convolução de Gauss; - Fechamento morfológico; - Laplaciano; - Transformada de Hough.	-	-	Avaliação de vários algoritmos em termos de desempenho; Calculo da média e o desvio padrão para a posição de cada extremidade dos segmentos de grupo.	Taxa de acerto: 98,45%. Cálculo de tempo para a rotina de extração horizonte: 6,5 milissegundos para uma resolução de 640x480.
HERDTWECK C, WALLRAVEN C, 2010, "Horizon estimation: Perceptual and computational experiments", APGV Symposium on Applied Perception in Graphics and Visualization (APGV), pages 49-56, 2010.	Tamanho da base não informada.	Conversão das imagens para o espaço de cor CIE L*a*b*. A análise seguinte concentra apenas os resultados do canal de luminância desconsiderando os efeitos de cor.	-	Algoritmos: - [Lin] - [Gab] - [Div] - [-Div] - [Van] - [GST]	-	Comparação computacional do percentual dos dados calculando o desvio absoluto de cada estimativa computacional com a estimativa perceptiva de cada participante, chegando ao número de resultados igual ao (número de Algoritmos)X(número de Participantes) para cada imagem.	Técnica baseada no gradiente vertical [lin] obteve um bom resultado; Nenhum algoritmo coincidiu com o desempenho humano; Erro de pelo menos 4 a 5% da altura da imagem com as suas estimativas comparadas com as anotações humanas;

							Os autores não informam a taxa de acerto da abordagem.
ANANTH, Christo; et al. Overlap Wavelet Transform for Image Segmentation. <i>International Journal of Electronics Communication and Computer Technology (IJEECT)</i> Volume 4 Issue 3 (May 2014).	Tamanho da base não informada.	Imagens coloridas.	-	-Segmentação baseada em OWT, imagem de entrada dividida em 4 subimagens ; -Transformada de Wavelet; -Agrupamento de histograma 2D; - Filtro de Gauss; -Operadores morfológicos.	-	Comparação entre vários espaços de cor (além do RGB, os autores não informam os demais).	Melhores resultados para RGB; Nenhuma avaliação numérica é realizada.
KULKARNI, Nilima. Color Thresholding Method for Image Segmentation of Natural Images. <i>I.J. Image, Graphics and Signal Processing</i> , 2012, 1, 28-34 Published Online February 2012.	Tamanho da base não informada.	Imagens coloridas.	-	Proposta de algoritmo de segmentação multicamadas para obter informação de cor RGB do objeto.	-	Avaliação baseada na percepção humana.	A técnica de segmentação é capaz de separar o objeto a partir do fundo.
Abordagens por detecção de bordas							
LIPSCHUTZ Ilan, GERSHIKOV Evgeny, MILGROM Benjamin, New Methods for Horizon Line Detection in Infrared and Visible Sea Images. <i>International Journal Of Computational Engineering Research (ijceronline.com)</i> Vol. 3 Issue. 3. pp 226 - 233, 2013.	Grupo de 10 imagens infravermelhas marinhas e 10 marinhas imagens de luz visível. No formato bmp (24 bits) sem compressão e com resoluções utilizadas variando de 300x200 até 1024x768 pixels.	Imagens em níveis de cinza.	- Erosão morfológica com elemento estruturante circular; - Filtro passa-baixa da média; - Detector de bordas; - Transformada de Hough.	Distância Bhattacharyya ou a distância Hellinge.	-	Avaliação e comparação dos algoritmos: Precisão medida pelo ângulo detectado (em graus) em relação a uma linha horizontal, bem como a altura da linha acima do centro da imagem (em pixels). Velocidade dos algoritmos medida em tempo de execução.	Os algoritmos H-COV (61.1 segundos) e H-His (48.9 segundos) significativamente mais lentos do que o algoritmo H-EDHT (0.3 segundos).
AHMAD, Touqeer; BEBIS, George; REGENTOVA, Emma; NEFIAN, Ara. "A Machine Learning Approach to Horizon Line Detection Using Local Features" 9th International Symposium on Visual Computing, Advances in Visual Computing, Volume 8033 of the series Lecture Notes in Computer Science p. 181-193, 2013.	Conjunto de 10 imagens contendo linha do horizonte. Separação dessas imagens em dois grupos: 5 para treinamento e 5 para testes e vice-versa.	Imagens em níveis de cinza.	Localização das arestas externas maximamente estáveis (MSEE) pelo detector de bordas de Canny.	-	Classificador SVM para classificar pixels que pertencem a linha do horizonte e os que não pertencem. Os pixels que não pertencem a linha do horizonte são descartados enquanto que os outros são pós-processados usando programação dinâmica baseado no algoritmo de menor caminho. O classificador retorna 1 ou -1. O valor binário	Avaliação referente ao pré-processamento.	O uso do MSEE, pré-processamento, reduz em até 92% das arestas que não são horizonte.

					positivo se refere a presença de pixels de horizonte.		
GERSHIKOV, Evgeny, Is Color Important for Horizon Line Detection?. The 2014 International Conference on Advanced Technologies for Communications (ATC'14), PP 262- 267, 2014.	Total de 21 imagens coloridas, a maioria delas imagens de mar com linha de horizonte linear.	Imagens coloridas e em escala de cinza.	- Filtragem morfológica para o pré-processamento das imagens.	- Borda de Canny; - Transformada de Hough; - Encontrar os picos e as linhas correspondentes; - Encontrar o segmento de linha mais longa; - Encontrar a equação linear deste segmento; - Marcá-lo na imagem.	- Métodos de características locais e globais para luminância e cor.	- As imagens <i>groundtruth</i> foram geradas manualmente; - A comparação dos métodos é calculada pelo desvio de altura das linhas detectadas e reais da imagem.	- O algoritmo mais preciso é o CMVDHT, baseado na detecção de borda por cor usando estatísticas de vetores de ordem e a transformada de Hough; - A adição de filtros de cor pode melhorar o seu desempenho.
SASSI, Ameni; AMAR, Chokri Ben; MIGUET, Serge. Skyline-based Approach for Natural Scene Identification. 13th ACS/IEEE International Conference on Computer Systems and Applications (AICCSA 2016), Nov 2016, Agadir, Morocco. 2016, < http://www.aiccsa.net/AICCSA2016/ >. <hal-01385537>.	2 a 4 imagens de paisagens naturais e de cidade.	Imagens coloridas.	-	- classificação das linhas retas; - análise de curvatura.	- classificação de retas a partir da aproximação poligonal gerando um histograma. - análise da curvatura ao longo da linha gerando um gráfico.	A classificação das linhas retas deu uma ideia sobre a linha do horizonte inteira. Para melhorar essa classificação, os autores usaram o descritor de espaço de escala de curvatura.	Os resultados obtidos na abordagem proposta não foram comparados com nenhuma abordagem existente. Os autores não apresentam as taxas de acerto da abordagem.
AYADI, Mehdi; SUTA, Loreta; SCUTURICI, Mihaela; MIGUET, Serge; AMAR, Chokri Bem. A Parametric Algorithm for Skyline Extraction. Springer International Publishing. Advanced Concepts for Intelligent Vision Systems: 17th International Conference, ACIVS 2016, Lecce, Italy, October 24-27, 2016, Proceedings, Oct 2016, Lecce, Italy. 10016, pp.604-615, 2016, Lecture Notes in Computer Science. < http://acivs.org/acivs2016/ >. <10.1007/978-3-319-48680-2_53>. <hal-01385542>.	320 imagens de natureza e cidade.	Imagem em cinza.	-	- filtro de Canny; - etapa de descontinuidade de conexão; - Através do gráfico de descontinuidade aplica-se um algoritmo de caminho mais curto para liga-las.	-	O método requer intervenção humana ao aplicar o filtro de Canny indicando parâmetros. O método usado é melhor, segundo os autores, do que os publicados anteriormente.	A precisão média entre as bases varia entre 96 e 99%.
CHIU, Chung-Cheng; LIU, Yuh-Jiun; CHIU, Sheng-Yi; CHANG, Hsing-Chien; HSU, Chia-Lun. 978-1-4673-9985-2/16/\$31.00 ©2016 IEEE.	40 imagens.	Imagens coloridas.	-	Método aplicado em dois módulos: 1º) linha do horizonte candidata por análise de ponto (direção vertical); 2º) linha do horizonte candidata por ponto de conexão (direção horizontal).	-	Comparação entre o gradiente de intensidade entre os pixels candidatos a linha do horizonte.	Os autores não informam as taxas de acerto e erro. Também não utilizam métricas para avaliar a abordagem proposta.

<p>LIANG, Dong; ZHANG, Weigang; HUANG, Qiaozhong; YANG, Fei. Robust Sea-sky-line Detection for Complex Sea Background. National Natural Science Foundation of China: 61303153. 978-1-4673-9088-0/15 ©2015 IEEE.</p>	<p>Base com 200 imagens de mar divididas em: - Céu com nuvens; - Mar com ondas; - Céu com muita luz; - Imagens consideradas normais, com céu limpo.</p>	<p>Imagem de cinza.</p>	<p>-</p>	<p>- Algoritmo de otsu; - Método de agrupamento.</p>	<p>-</p>	<p>Os autores propõem um método que consiste em duas etapas: A primeira consistem em transformar a imagem original em escala de cinza e a região da linha de margem é localizada usando a matriz de co-ocorrência baseada no contraste da característica de textura.</p>	<p>Taxa média de acerto de 93% para as categorias mencionadas no estudo considerando uma linha de horizonte reta, desconsiderando os obstáculos.</p>
<p>Stephen Williams and Ayanna M. Howard. "Horizon Line Estimation In Glacial Environments Using Multiple Visual Cues", IEEE International Conference on Robotics and Automation, ICRA 2011, Shanghai, China, pp 5887—5892, May 2011.</p>	<p>Conjuntos de dados coletados em duas geleiras diferentes no Alasca. Quantidade de imagens não informada.</p>	<p>Imagens em níveis de cinza.</p>	<p>-</p>	<p>- Detector de bordas de Canny; - Limiar histograma adaptativo (AHT); - Crescimento de Região (RG); - Modelo de mistura gaussiana (GMM); - Fusão estatística de Regiões (MRE).</p>	<p>-</p>	<p>Avaliação em termos de velocidade; Uma única imagem de 640x480 foi carregada para a memória; Cada algoritmo processado 1000 vezes para a imagem de teste; Tempo de processamento gravado.</p>	<p>O desempenho do sistema foi testado em termos de tempo e desempenho da classificação, para algoritmos de segmentação representativos de diversas classes diferentes. Resultados – tempo médio por quadro: -Limiar histograma adaptativo (AHT) = 0.0074 segundos; - Crescimento por Região (RG) = 0.0229 segundos; - Modelo de mistura gaussiana (GMM) = 0.2677 segundos; - Fusão estatística de Regiões (MRE) = 0.2803 segundos. - Extração da linha do horizonte em 0.0296 segundos.</p>

CAPÍTULO 3

MÉTODO PROPOSTO

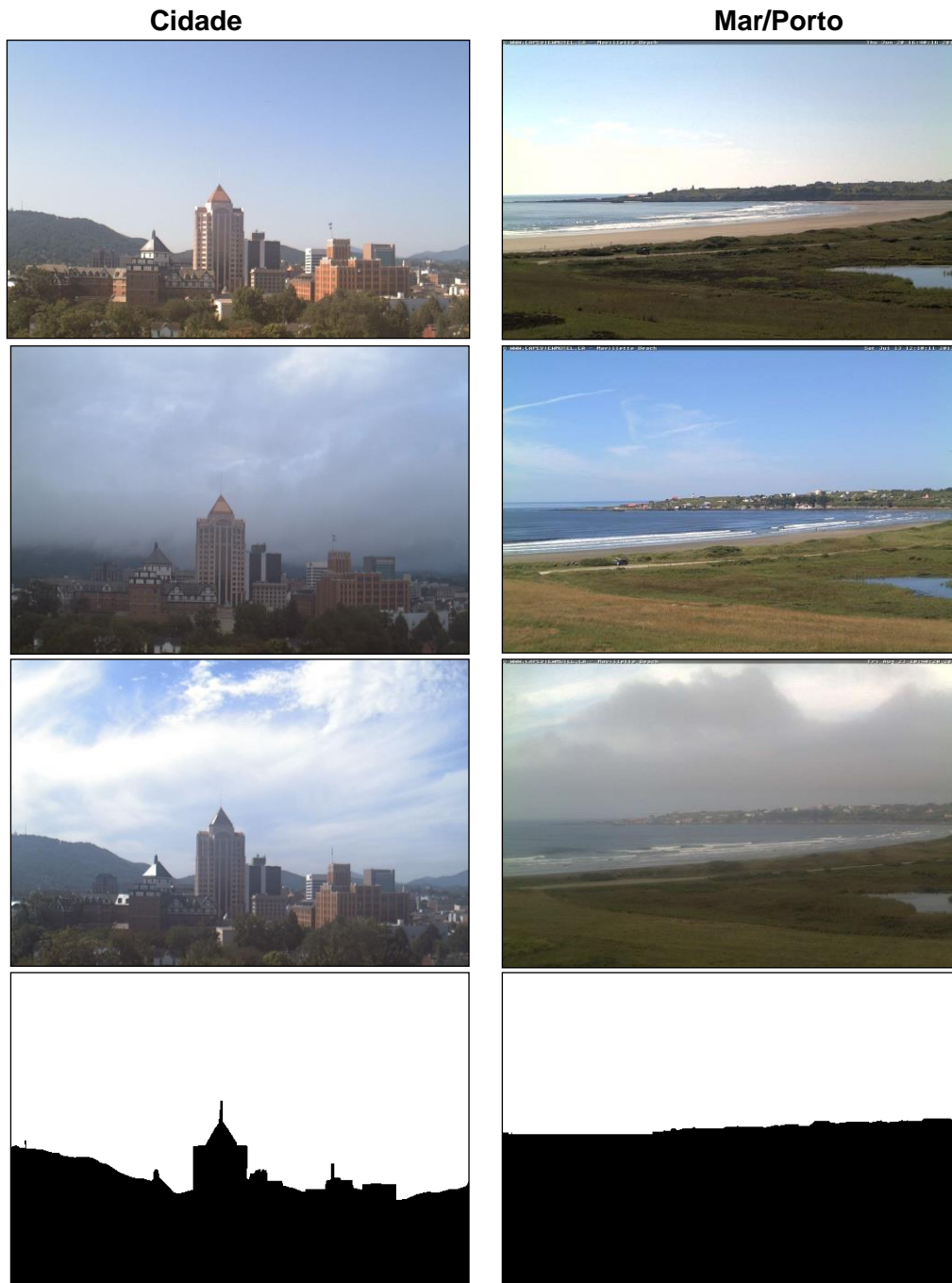
Neste capítulo são apresentados todos os elementos utilizados para o desenvolvimento deste trabalho. A Seção 3.1 descreve as bases de imagens utilizadas. Na Seção 3.2 são descritas as características para segmentação, a Seção 3.3 a classificação e, por fim, a Seção 3.4 a avaliação da abordagem.

3.1 Base de Dados

Durante os experimentos duas bases de imagens foram criadas, as duas contendo 1.200 imagens coloridas cada, apresentando diferentes tipos de imagens: cidade, mar/porto, natureza/montanha e rodovia/estrada. Para as duas bases, as imagens foram padronizadas com tamanho 1280 x 720 do tipo png (*Portable Network Graphic*).

A primeira base de imagens chamada de **Base Variações** foi selecionada através da base disponibilizada pelo professor Radu Paul Mihail usadas em seus experimentos (MIHAIL et al., 2016), contendo uma base de imagens capturadas através de frames de vídeos em diversos horários do dia em diferentes estações do ano contendo grande variação de luminosidade e clima e suas respectivas *groundtruth*. Esta base foi modificada selecionando quatro tipos de imagens e suas variações para contemplar o mesmo tipo de imagem proposta na segunda base (cidade, mar/porto, natureza/montanha, rodovia/estrada). Cada câmera de vídeo foi disponibilizada com mais de 2.000 imagens cada e a seleção dos quatro tipos de imagens foi definida através do contexto de chão que cada imagem representa e a escolha das 300 imagens (por categoria) deste total foi realizada através de um protocolo automático de seleção de imagens com um limiar de 90% de similaridade usando com parâmetros r, g e b de RGB desenvolvido em java para este fim.

A Figura 3.1 mostra as quatro imagens selecionadas com algumas de suas respectivas variações e as *groundtruth* correspondentes.



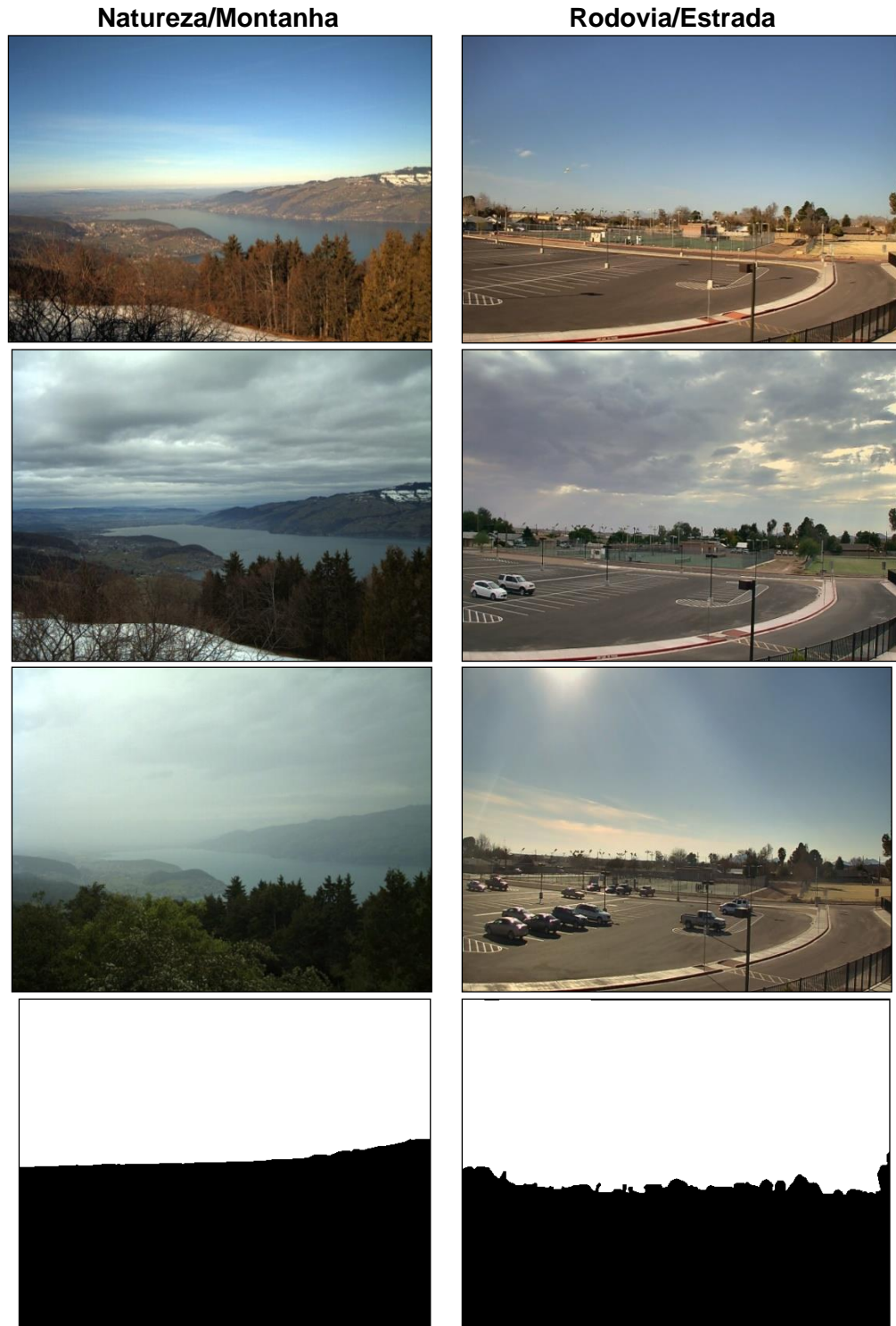


Figura 3.1: Exemplo das categorias de imagens com variações de céu e chão e suas respectivas *groundtruth* contemplando a Base Variações.

A segunda base de imagens chamada **Base Web** foi selecionada aleatoriamente da web levando em conta a maior concentração de chão, ou seja, cidade, mar/porto, natureza/montanha e rodovia/estrada e suas *groundtruth* que

foram obtidas manualmente por um especialista na área. A Figura 3.2 apresenta uma imagem de cada tipo e sua respectiva *groundtruth*.

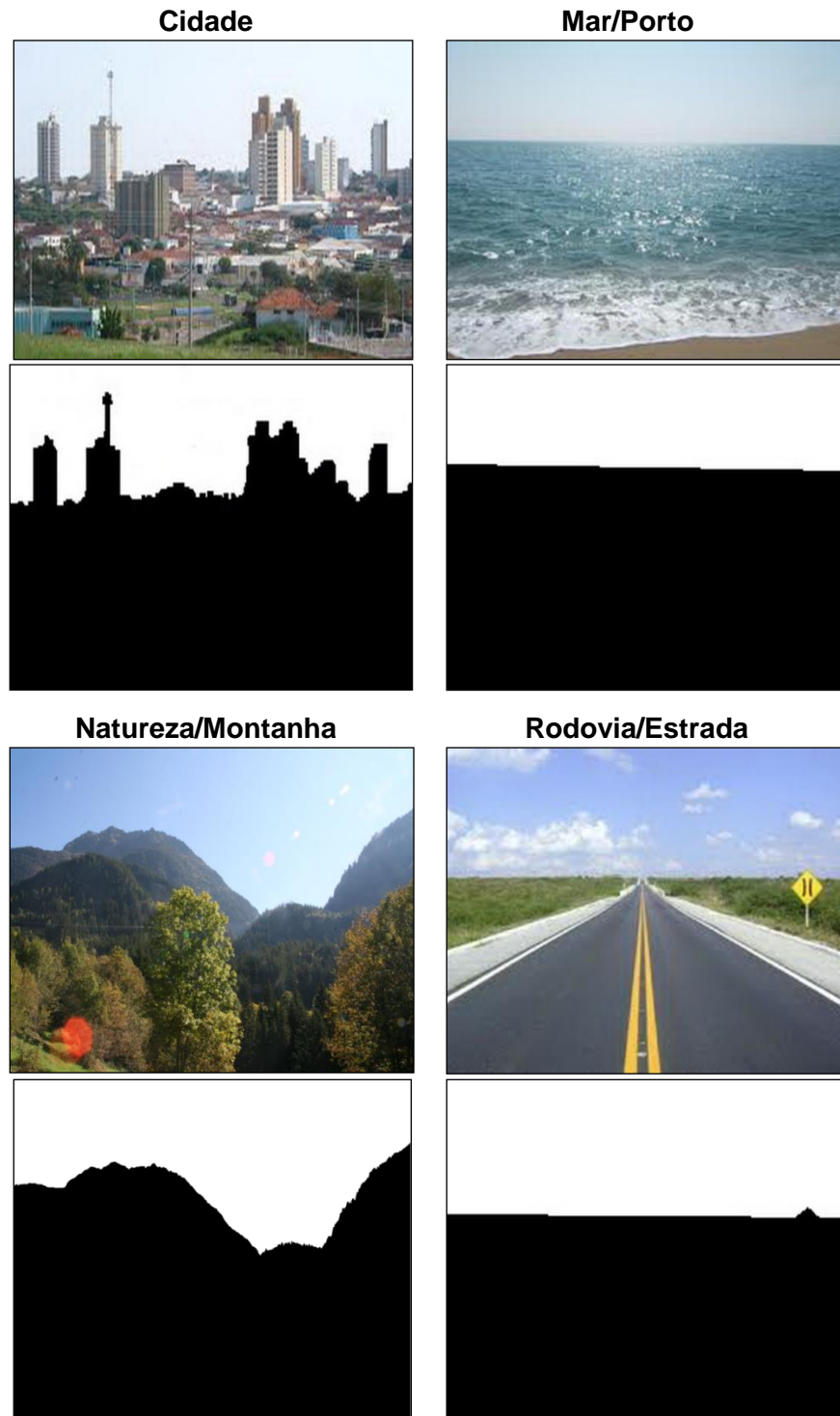


Figura 3.2: Imagens representando as quatro categorias e suas respectivas *groundtruth* contemplando a Base Web.

3.2 Extração de Características para Segmentação

O método em estudo consiste na separação entre céu e chão em imagens coloridas através da extração de características descritas a seguir.

3.2.1 Características Baseadas em Índices de Branquidão

Os vários índices de branquidão descritos na Seção 2.1.1.1 foram computados pixel a pixel em uma janela 3x3 para as imagens das duas bases. Destes calculados foram geradas imagens em níveis de cinza normalizadas entre 0 e 255. A Figura 3.3 mostra a tela da aplicação java usada para a seleção dos índices de branquidão.

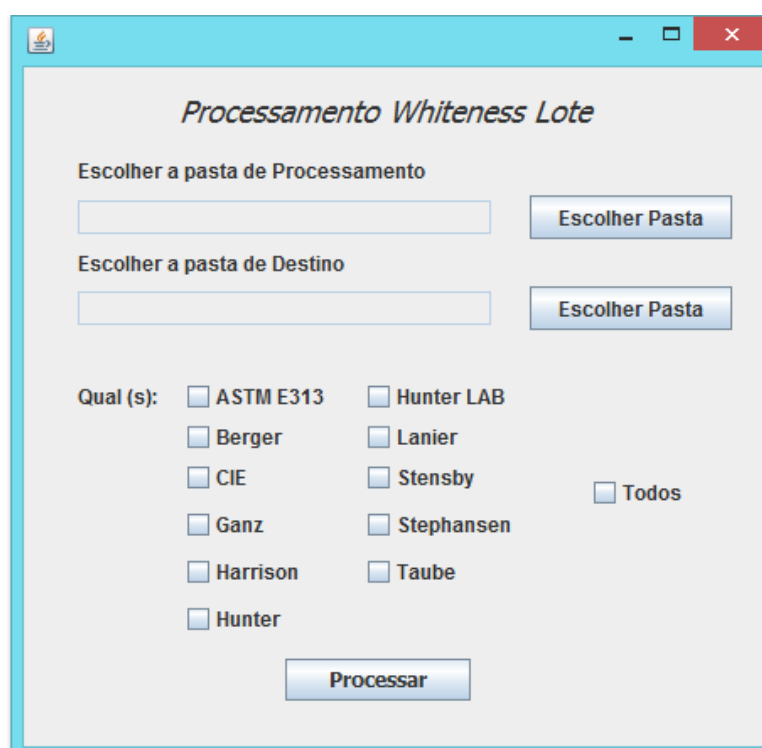


Figura 3.3: Tela da aplicação java para seleção dos índices de branquidão.

3.2.2 Características Baseadas Índices de Azul

O conceito de azul é um índice que mede o grau relativo de azul, como apresentado na Seção 2.1.1.2. Esses índices são usados em aplicações que medem a quantidade de azul encontrada nos mares e oceanos (WATANABE et al., 2016).

Os vários índices de azul foram computados pixel a pixel em uma janela 3x3 para as imagens das duas bases. Destes foram geradas imagens em níveis de cinza normalizadas entre 0 e 255.

A Figura 3.4 mostra a tela da aplicação Java usada para a seleção dos índices de azul.

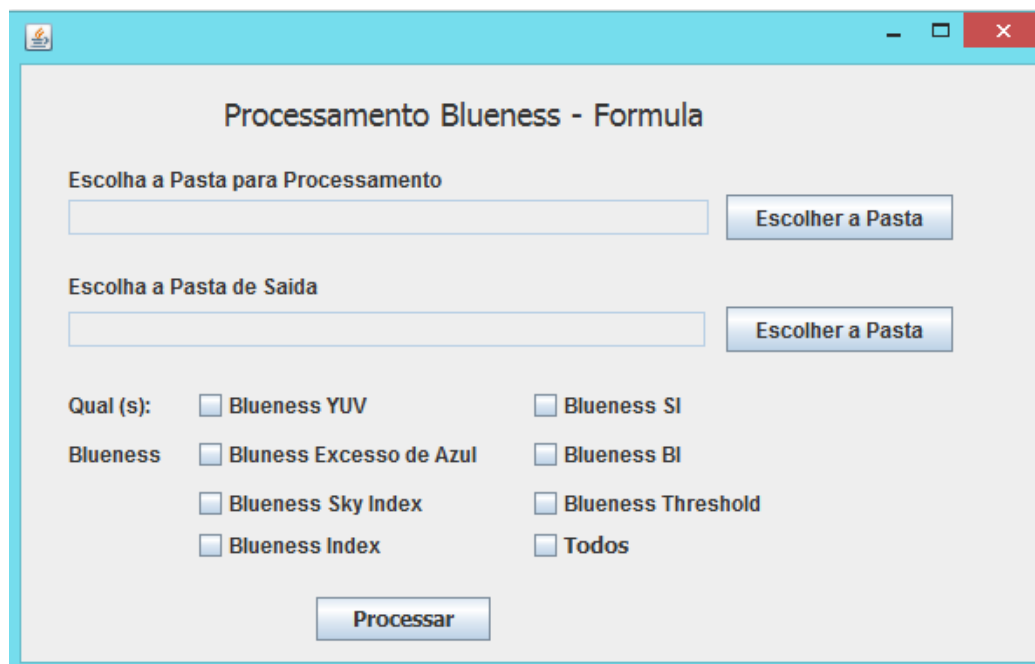


Figura 3.4: Tela da aplicação java para seleção dos índices de azul.

3.2.3 Características Baseadas em LBP (*Local Binary Pattern*)

O algoritmo LBP descrito na Seção 2.1.2.1 foi usado nesta abordagem com a configuração uniforme LBP (8, 1) que relaciona, os 8 vizinhos mais próximos do pixel central em uma matriz 3x3 gerando um vetor de 59 características.

3.2.4 Características Baseadas em Perfil morfológico estendido (*Extended Morphological Profile*)

O perfil morfológico estendido apresentado na Seção 2.1.2.2 significa a sucessão de abertura e fechamento com elemento estruturante de tamanho crescente (em cruz) para cada pixel a partir de imagens coloridas.

A configuração segue a seguinte sequência: Abertura → 1 iteração; Abertura → 2 iterações; Abertura → 3 iterações; Fechamento → 1 iteração; Fechamento → 2

iterações; Fechamento → 3 iterações. A cada configuração as imagens são normalizadas entre 0 e 255 e armazenadas como característica para o vetor que contempla 6 características ao final com espaço de cor HSI (Matiz, Saturação, Intensidade) utilizando a sequência ISH para cada iteração realizada, para então, gerar o vetor de características.

3.2.5 Outras Características

A seleção destas características descritas na Seção 2.1.3 são realizadas por uma aplicação java desenvolvida para este propósito, que, por sua vez, gera o arquivo arff necessário para a etapa de classificação, conforme Figura 3.5.

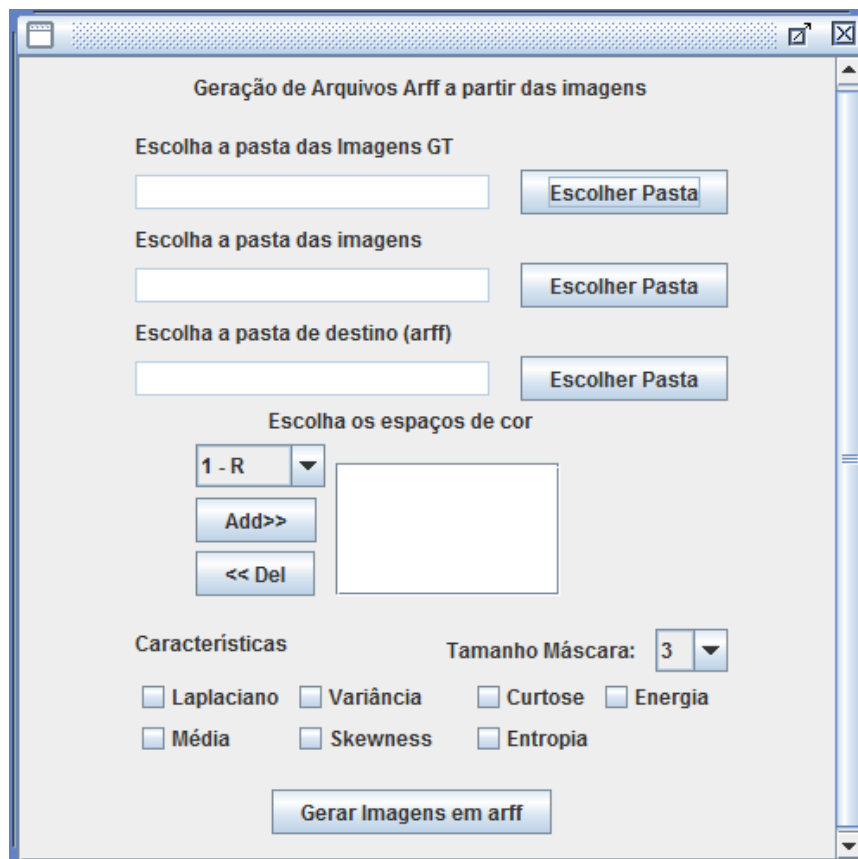


Figura 3.5: Exemplo de tela da aplicação java para seleção de características.

Estas características, citadas na literatura, foram combinadas com as características citadas nas seções anteriores a fim de analisar as taxas de acerto.

3.3 Classificação

Em função das melhores taxas de acerto apresentadas na literatura pelo classificador SVM e por este ser uma abordagem de classificação usada em muitos estudos, o SVM foi o classificador escolhido para o método proposto.

Para este estudo, o classificador SVM segue a seguinte configuração: kernel RBF (*Radial Basis Function*) com valores de $c = 10$ e $\text{gamma} = 0,0001$ no weka (*Waikato Environment for Knowledge Analysis*). A escolha pelos melhores parâmetros usa a pesquisa em grade com validação cruzada definida a partir de um script em *python*.

3.3.1 Uso de Classificadores Monolíticos

Para avaliação da capacidade de generalização das características individuais e agrupadas a técnica de validação cruzada foi empregada com *5-folds*: Primeiro o conjunto de dados é dividido em 5 subconjuntos de tamanhos iguais, depois cada subconjunto é usado para teste e os demais para treinamento. A escolha desse método de avaliação evita a superposição dos conjuntos de teste (YU-QUAN et.al, 2011). Dessa forma, cria-se um classificador para cada vetor de características. Após a avaliação individual, as características com os melhores resultados foram combinadas em um mesmo vetor para análise de seu *F-measure*.

3.3.2 Uso de Múltiplos Classificadores

Esta etapa consiste na combinação da saída dos classificadores descritos na Seção 3.3.1. As melhores características de branquidão foram combinadas através do voto majoritário, ou seja, são consideradas as possíveis combinações entre os classificadores avaliados por meio da construção de conjuntos (*ensembles*) de classificadores. Para utilizar totalmente as informações obtidas de um grupo de classificadores, a saída de cada classificador pode ser combinada com as demais, permitindo uma decisão, aquele que tiver mais votos é o vencedor (BREIMAN, 1996).

Essa técnica foi empregada em duas abordagens para que sejam comparadas posteriormente. Na primeira abordagem, para cada base, todas as imagens foram misturadas e separadas em 60% para treinamento, gerando classificadores SVM para as melhores características individuais e testadas com 40% das imagens. Na segunda abordagem, a mesma técnica é empregada, mas para cada categoria separadamente.

Todos os arquivos de treino foram divididos em duas fatias verticais e, para cada fatia, a soma dos pixels de céu/chão é dividida pela quantidade de amostras, o resultado é o intervalo de pixels a ser selecionado para gerar o arquivo de treino, por exemplo, se o número de amostras informado é de 400 e a soma total de pixels de céu é 12.000, têm-se, $12.000/400=30$, então a cada 30 pixels um é selecionado para o arquivo de treino. Esse cálculo é realizado para os pixels de céu e chão considerando as duas fatias. A Figura 3.6 mostra a tela da aplicação java usada para gerar o arquivo de treino.

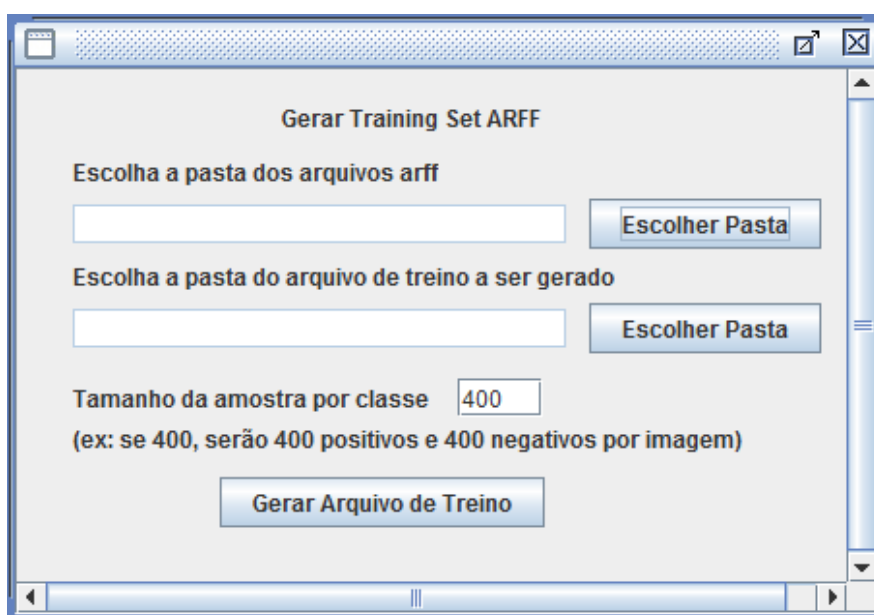


Figura 3.6: Exemplo de tela da aplicação java para gerar o arquivo de treino.

3.4 Avaliação da Abordagem

A métrica utilizada para a avaliação da segmentação é o *F-measure* descrita em (BIMBO, 1999), (GATOS, NTIROGIANNIS E PRATIKAKIS, 2009), (SEZGIN E SANKUR, 2001), conforme equação 39:

- *F-Measure (FM)* retorna 1 para segmentação perfeita:

$$FM = \frac{2 * P * R}{P + R} \quad (39)$$

$$\text{onde } P = \frac{T^+}{T^+ + F^+} \text{ e } R = \frac{T^+}{T^+ + F^-}$$

em que T^+ , T^- , F^+ e F^- é o verdadeiro positivo, o verdadeiro negativo, falso positivo e falso negativo, respectivamente.

A justificativa pela escolha da métrica *F-measure* está em apresentar uma média harmônica entre precisão e *recall*.

CAPÍTULO 4

EXPERIMENTOS

Este capítulo apresenta os principais resultados obtidos durante os experimentos realizados para o método proposto. Dessa forma, a Seção 4.1 apresenta o protocolo experimental, a Seção 4.2 a 4.6 apresentam os resultados da segmentação das imagens pelo classificador supervisionado SVM para as características individuais, os resultados com características agrupadas na Seção 4.7, a Seção 4.8 apresenta os experimentos com a combinação de classificadores e a discussão dos resultados na Seção 4.9.

4.1 Protocolo Experimental

Para validar os resultados obtidos com o método apresentado neste trabalho é necessária a realização de experimentos. Tais experimentos devem seguir um protocolo para que os mesmos possam ser reutilizados em diferentes contextos. Variáveis como o tipo da imagem afetam os resultados desta pesquisa.

A Figura 4.1 apresenta uma visão geral do protocolo seguido ao longo dos experimentos realizados na primeira etapa para as características individuais e para as características agrupadas em que a validação cruzada foi empregada com 5 *folds*. A Figura 4.2 apresenta o protocolo usado na segunda etapa dos experimentos com combinação de classificadores para obtenção do voto da maioria aplicado em duas abordagens: a primeira abordagem para 60% das imagens de cada base para treinamento e 40% das imagens para teste; a segunda abordagem prevê a divisão das imagens em categorias usando o mesmo percentual para treinamento e teste.

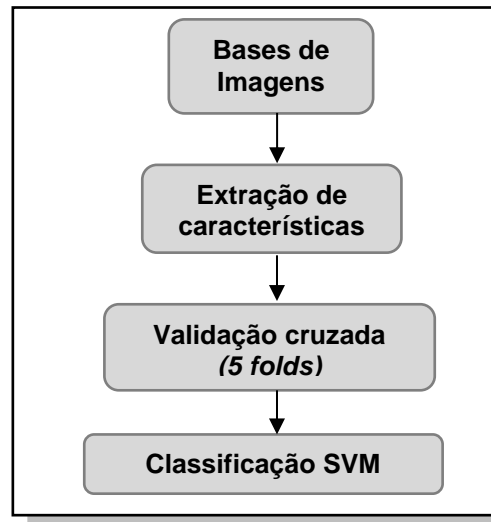


Figura 4.1: Protocolo dos experimentos para características individuais/agrupadas.

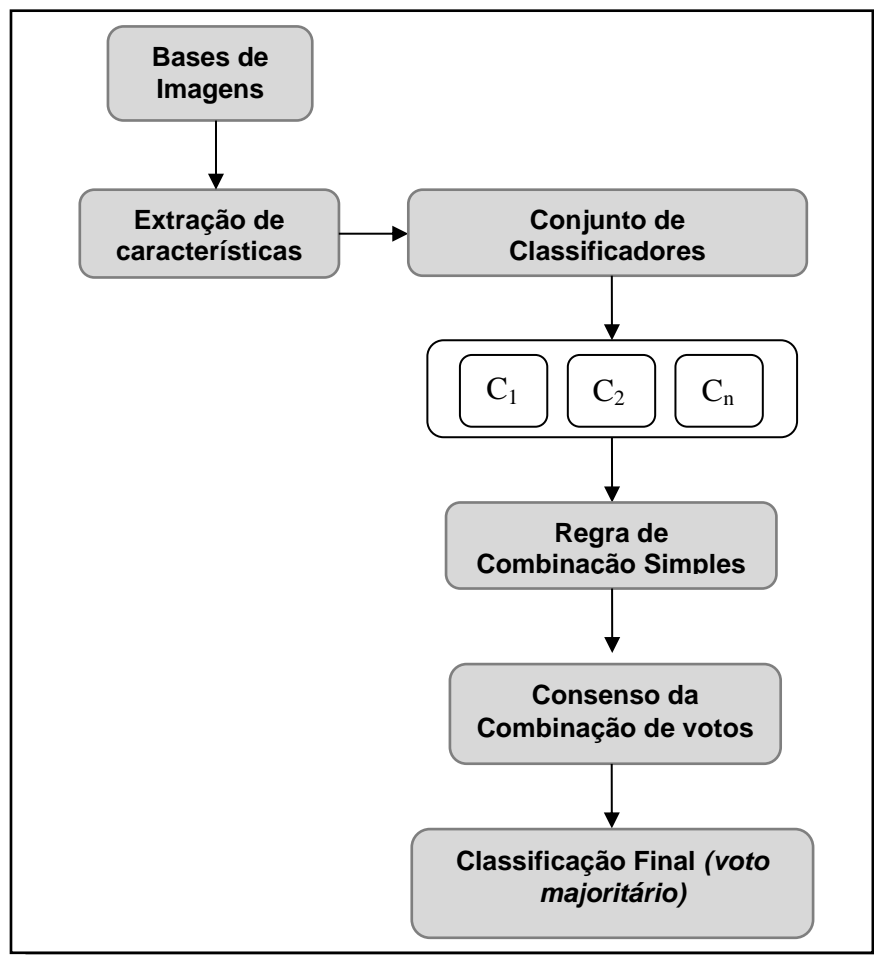


Figura 4.2: Protocolo dos experimentos com combinação de classificadores.

4.2 Experimentos – Segmentação Baseada em Espaços de Cor

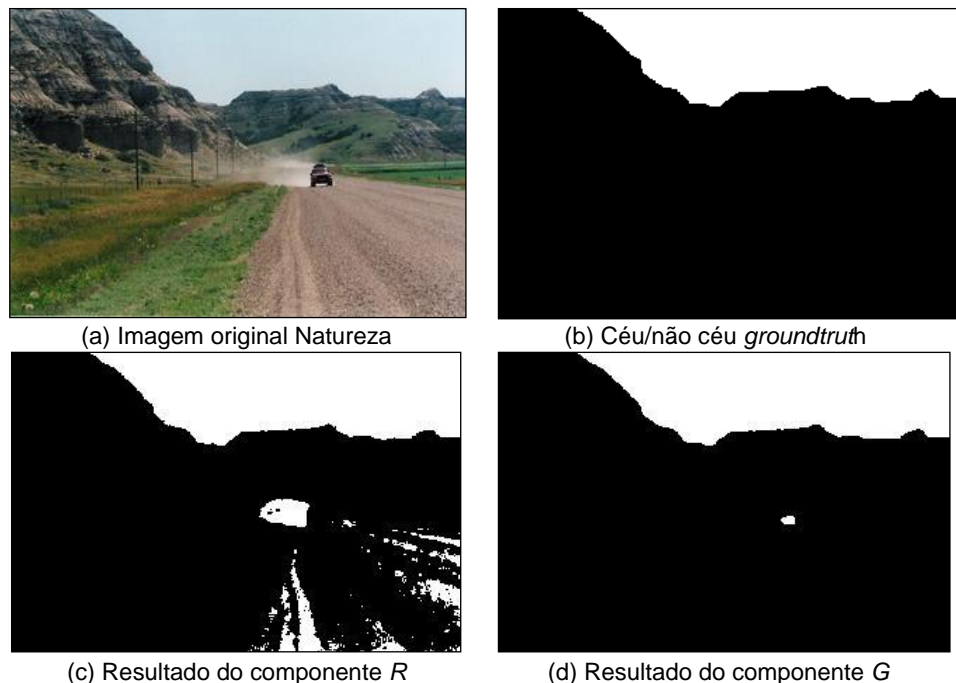
Os resultados iniciais foram realizados para cada canal de cor separadamente R , G , B , Y , Cb , Cr , H , S , V , mais média usando o classificador SVM para 1.200 imagens distintas, Base Web, como mostra a tabela 4.1.

Tabela 4.1: Avaliação da separação de céu e chão para os componentes R , G , B , H , S , V , Y , Cr e Cb .

	R	G	B	Y	Cr	Cb	H	S	V
FM	81,78%	82,91%	87,17%	83,21%	76,38%	76,87%	74,53%	75,68%	86,33%

Avaliando os resultados apresentados na tabela 4.1 observa-se que, para todos os tipos de imagens, os canais B , Y e V apresentam, de forma geral, maior percentual de acerto para a métrica F -measure.

A Figura 4.3 apresenta uma imagem segmentada para cada canal de cor R , B , B , Y , Cr , Cb , H , S , V separadamente.



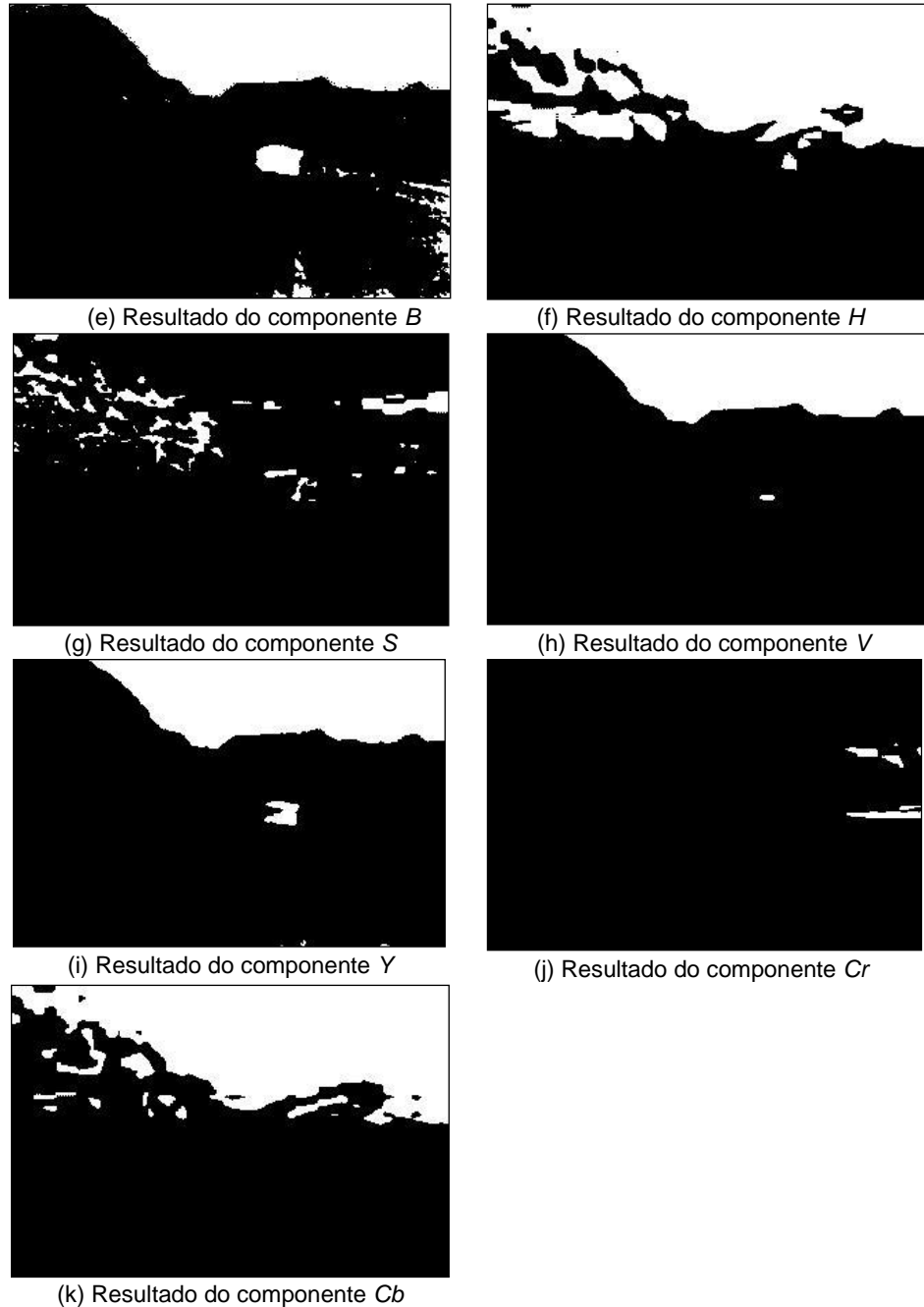


Figura 4.3: Exemplo de processamento de imagem para cada componente R, G, B, H, S, V, Y, Cr, Cb.

A partir destes experimentos, os canais que apresentaram melhores resultados com o classificador SVM foram combinados considerando a taxa de acerto F -measure para cada canal. Portanto, uniu-se os canais de cor B , V e Y com as características de cor e média ($C+\mu$) gerando os seguintes resultados com variação do tamanho da janela, como mostra a tabela 4.2.

Tabela 4.2: Avaliação da separação de céu e não céu variando tamanho da janela.

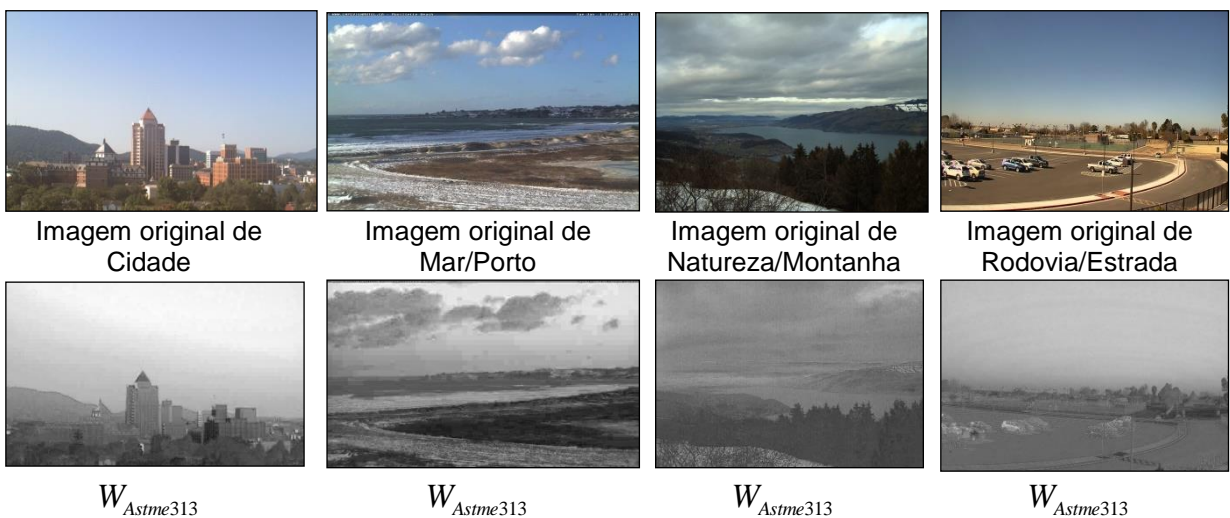
	Janela 3x3	Janela 5x5	Janela 7x7	Janela 9x9
<i>FM</i>	90,72%	88,53%	87,61%	87,22%

Além da variação do tamanho de janela que mostrou-se eficiente somente com janela 3x3, experimentos foram realizados com cor e média mais *skewness*, variância, curtose, energia e entropia, mas as taxas de *F-measure* diminuíram consideravelmente, em torno de 9 e 10 pontos percentuais.

Esses experimentos iniciais servem principalmente para analisar se os índices de branquidão, ainda não usados na literatura para este fim, tem resultados promissores.

4.3 Experimentos – Segmentação Baseada em Índices de Branquidão

Inicialmente a característica de branquidão foi aplicada às duas bases de imagens criadas para esta pesquisa. Os índices de branquidão contemplam onze fórmulas distintas, usadas exclusivamente na indústria. A Figura 4.4 e a tabela 4.3 apresentam os resultados dos índices para a Base Variações contemplando os tipos de imagens. Da mesma forma, a Figura 4.5 e tabela 4.4 apresentam os resultados para a Base Web contemplando também os tipos de imagens encontradas.





W_{Berger}



W_{Berger}



W_{Berger}



W_{Berger}



W_{Cie}



W_{Cie}



W_{Cie}



W_{Cie}



W_{Ganz}



W_{Ganz}



W_{Ganz}



W_{Ganz}



$W_{Harrison}$



$W_{Harrison}$



$W_{Harrison}$



$W_{Harrison}$



W_{Hunter}



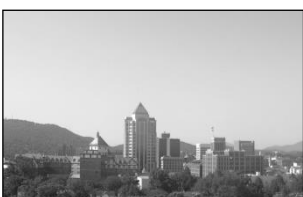
W_{Hunter}



W_{Hunter}



W_{Hunter}



$W_{HunterLAB}$



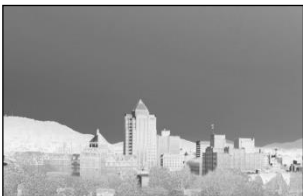
$W_{HunterLAB}$



$W_{HunterLAB}$



$W_{HunterLAB}$



W_{Lanier}



W_{Lanier}



W_{Lanier}



W_{Lanier}

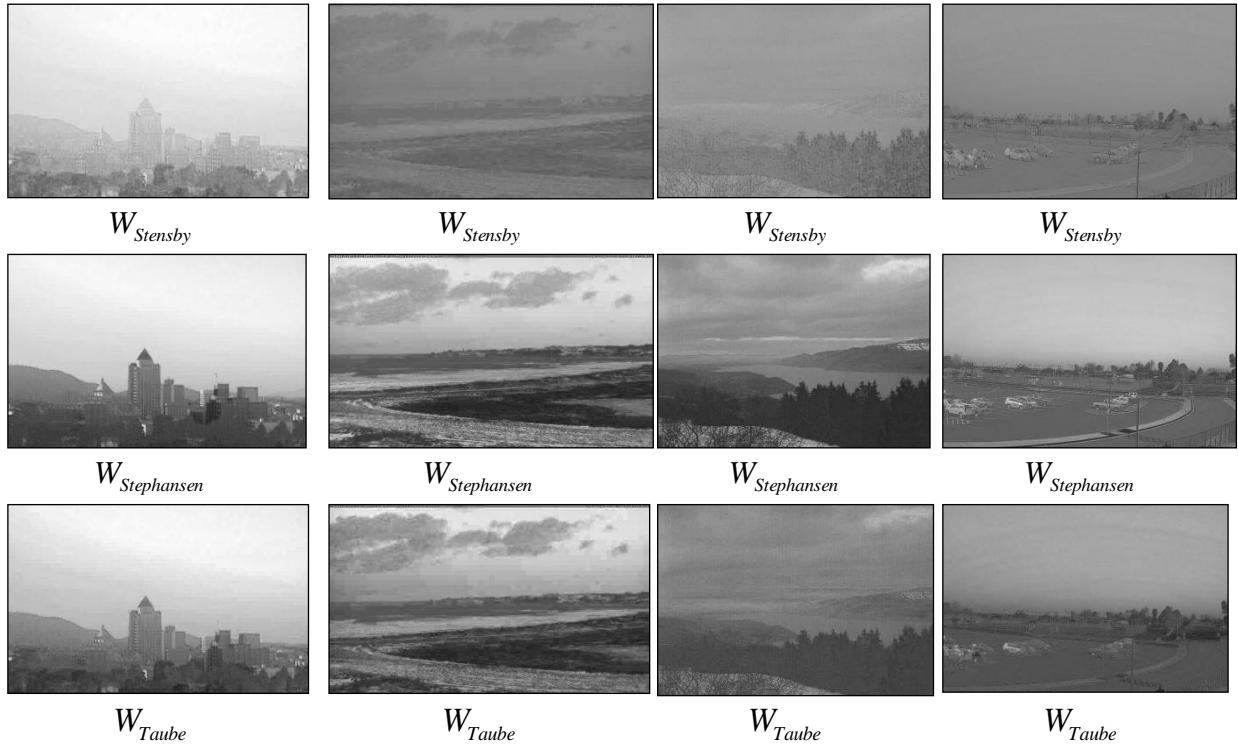


Figura 4.4: Resultados individuais dos índices de branquidão para a Base Variações.

Tabela 4.3: Avaliação individual para os índices de branquidão para a Base Variações.

Índices de branquidão	<i>F-measure</i>
$W_{Astme313}$	89,00%
W_{Berger}	92,00%
W_{Cie}	65,50%
W_{Ganz}	61,50%
$W_{Harrison}$	64,75%
W_{Hunter}	64,00%
$W_{HunterLAB}$	92,50%
W_{Lanier}	70,75%
$W_{Stensby}$	66,25%
$W_{Stephansen}$	92,80%
W_{Taube}	90,00%

Conforme Tabela 4.3 observar-se que os índices de branquidão que apresentaram os melhores resultados para *F-measure*, levando em consideração uma variação de 6 pontos percentuais, na Base Variações são os índices $W_{Astme313}$,

W_{Berger} , $W_{HunterLAB}$, $W_{Stephansen}$ e W_{Taube} com 89%, 92%, 93%, 93% e 90%, respectivamente.



Imagem original de Cidade



Imagem original de Mar



Imagem original de Natureza



Imagem original de Rodovia/ Estrada



$W_{Astme313}$



$W_{Astme313}$



$W_{Astme313}$



$W_{Astme313}$



W_{Berger}



W_{Berger}



W_{Berger}



W_{Berger}



W_{Cie}



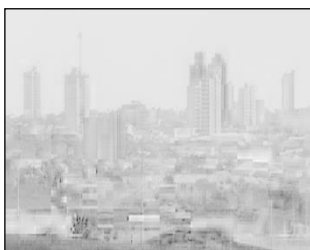
W_{Cie}



W_{Cie}



W_{Cie}



W_{Ganz}



W_{Ganz}



W_{Ganz}



W_{Ganz}



$W_{Harrison}$



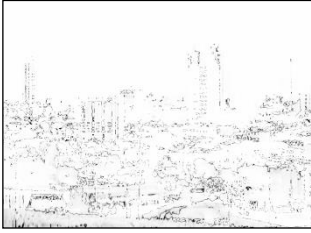
$W_{Harrison}$



$W_{Harrison}$



$W_{Harrison}$



W_{Hunter}



W_{Hunter}



W_{Hunter}



W_{Hunter}



$W_{HunterLAB}$



$W_{HunterLAB}$



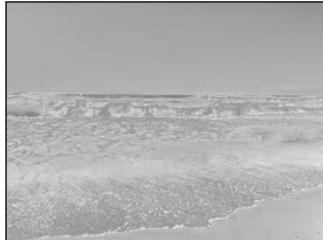
$W_{HunterLAB}$



$W_{HunterLAB}$



W_{Lanier}



W_{Lanier}



W_{Lanier}



W_{Lanier}



$W_{Stensby}$



$W_{Stensby}$



$W_{Stensby}$



$W_{Stensby}$



$W_{Stephansen}$



$W_{Stephansen}$



$W_{Stephansen}$



$W_{Stephansen}$

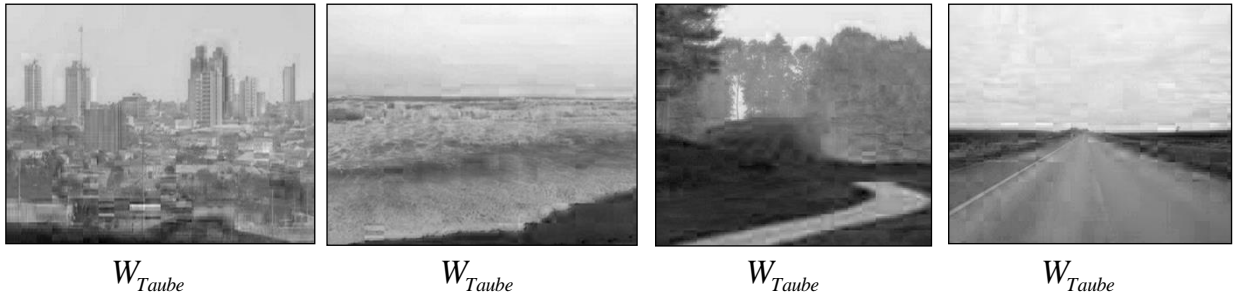


Figura 4.5: Resultados individuais dos índices de branquidão para a Base Web.

Tabela 4.4: Avaliação individual para os índices de branquidão para a Base Web.

Índices de Branquidão	<i>F-measure</i>
$W_{Astme313}$	80,75%
W_{Berger}	82,00%
W_{Cie}	58,50%
W_{Ganz}	55,00%
$W_{Harrison}$	57,50%
W_{Hunter}	69,75%
$W_{HunterLAB}$	77,50%
W_{Lanier}	66,50%
$W_{Stensby}$	57,75%
$W_{Stephansen}$	89,10%
W_{Taube}	82,25%

A partir dos resultados individuais obtidos pode-se detectar cinco melhores índices, $W_{Astme313}$, W_{Berger} , $W_{HunterLAB}$, $W_{Stephansen}$ e W_{Taube} . É importante ressaltar que, para as duas bases de imagens, o melhor *F-measure* foi o índice $W_{Stephansen}$.

4.4 Experimentos – Segmentação Baseada em Índices de Azul

Os índices de azul contemplam sete fórmulas distintas para o cálculo do índice de azul, usados quase que exclusivamente na indústria. A Figura 4.6 e a Tabela 4.5 apresentam os resultados dos índices para a Base Variações e a Figura 4.7 e a Tabela 4.6 apresentam os resultados para a Base Web. Observa-se que algumas imagens retornam pouca ou nenhuma informação de azul.






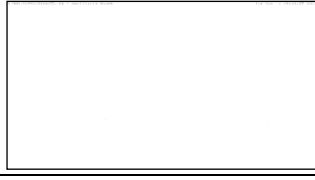
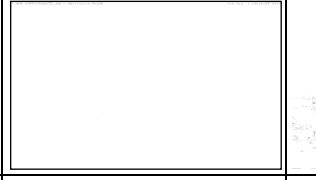





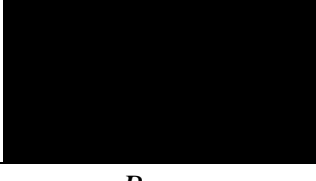
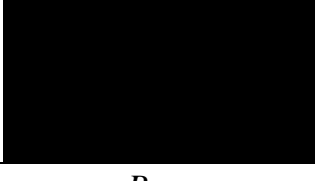


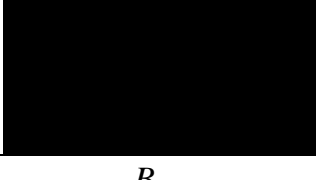




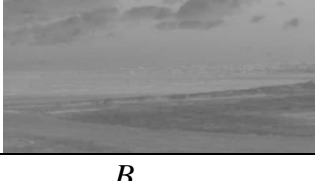


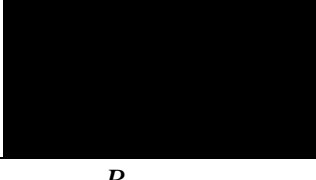
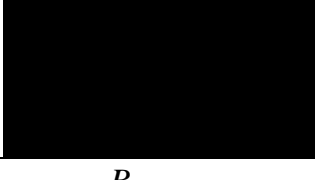










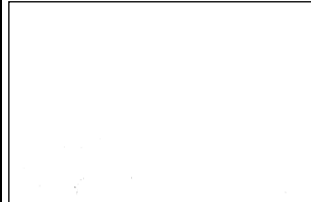




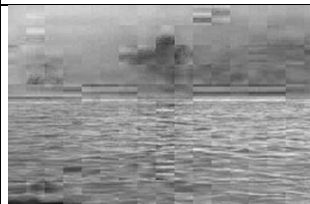


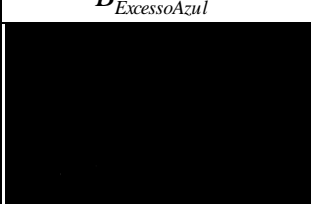



			
Imagem original de Cidade	Imagem original de Mar	Imagem original de Natureza	Imagem original de Rodovia/Estrada
			
B_{BI}	B_{BI}	B_{BI}	B_{BI}
			
$B_{ExcessoAzul}$	$B_{ExcessoAzul}$	$B_{ExcessoAzul}$	$B_{ExcessoAzul}$
			
B_{Index}	B_{Index}	B_{Index}	B_{Index}
			
B_{SI}	B_{SI}	B_{SI}	B_{SI}
			
$B_{SkyIndex}$	$B_{SkyIndex}$	$B_{SkyIndex}$	$B_{SkyIndex}$
			
$B_{Threshold}$	$B_{Threshold}$	$B_{Threshold}$	$B_{Threshold}$
			
B_{YUV}	B_{YUV}	B_{YUV}	B_{YUV}

Figura 4.6: Resultados individuais dos índices de azul para a Base Variações.

Tabela 4.5: Avaliação individual dos índices de azul para a Base Variações.

Índices de Azul	<i>F-measure</i>
B_{BI}	43,00%
$B_{ExcessoAzul}$	91,25%
B_{Index}	38,50%
B_{SI}	45,50%
$B_{SkyIndex}$	78,50%
$B_{Threshold}$	38,50%
B_{YUV}	74,00%

			
Imagem original de Cidade	Imagem original de Mar	Imagem original de Natureza	Imagem original de Rodovia/Estrada
			
B_{BI}	B_{BI}	B_{BI}	B_{BI}
			
$B_{ExcessoAzul}$	$B_{ExcessoAzul}$	$B_{ExcessoAzul}$	$B_{ExcessoAzul}$
			
B_{Index}	B_{Index}	B_{Index}	B_{Index}

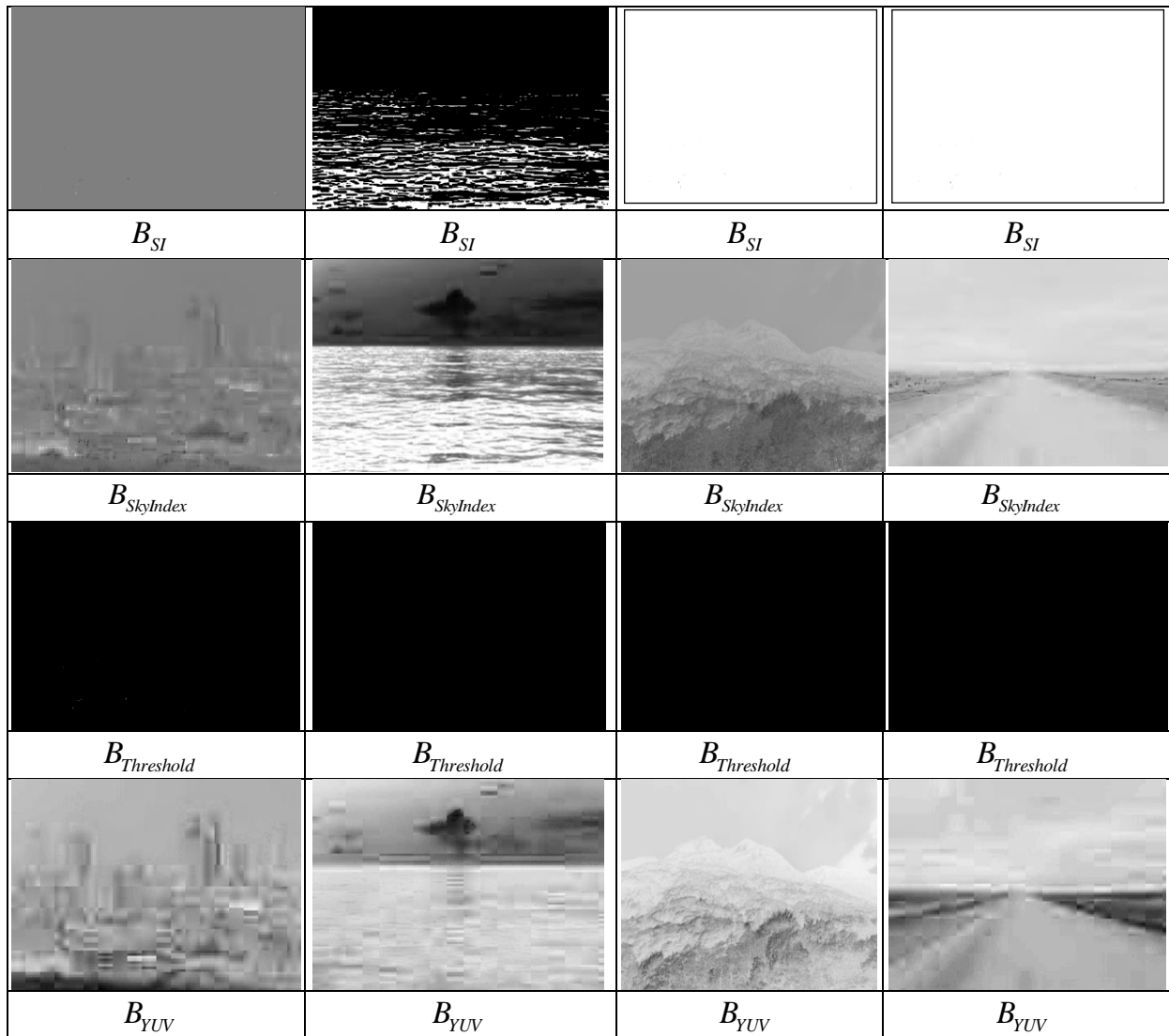


Figura 4.7: Resultados individuais dos índices de azul para a Base Web.

Tabela 4.6: Avaliação individual dos índices de azul para a Base Web.

Índices de Azul'	<i>F-measure</i>
B_{BI}	39,75%
$B_{ExcessoAzul}$	82,50%
B_{Index}	34,75%
B_{SI}	43,50%
$B_{SkyIndex}$	62,50%
$B_{Threshold}$	34,75%
B_{YUV}	67,75%

As duas bases apresentam o mesmo índice de azul, $B_{ExcessoAzul}$, como o melhor resultado, mas não supera o resultado do melhor índice de branquidão.

4.5 Experimentos – Segmentação Baseada em LBP (*Local Binary Pattern*)

A Tabela 4.7 apresenta o *F-measure* do algoritmo LBP para a Base Variações e Base Web.

Tabela 4.7: Resultados do algoritmo LBP para as Bases de Imagens.

	Característica	<i>F-measure</i>
Base Variações	LBP	92,20%
Base Web		88,50%

É interessante ressaltar que o *F-measure* para a Base Web, que contém diferentes contextos de chão, obteve um resultado interessante com 82,20% para a Base Variações e 88,50% para a Base Web.

4.6 Experimentos – Segmentação Baseada em Perfil Morfológico Estendido

Os resultados baseados em perfil morfológico estendido foram obtidos através da configuração descrita na Seção 3.2.4: Abertura → 1 iteração; Abertura → 2 iterações; Abertura → 3 iterações; Fechamento → 1 iteração; Fechamento → 2 iterações; Fechamento → 3 iterações. A Tabela 4.8 apresenta os resultados com perfil morfológico estendido para as duas bases de imagens.

Tabela 4.8: Resultados com perfil morfológico estendido para as Bases de Imagens.

Bases de Imagens	<i>F-measure</i>
Base Variações	77,10%
Base Web	75,15%

A Tabela mostra que o uso do perfil morfológico estendido para este tipo de abordagem não demonstra um resultado melhor do que os índices de branquidão, azul e LBP.

4.7 Experimentos - Classificadores Monolíticos

Os experimentos descritos nesta Seção representam a criação de um classificador para cada vetor de características e a combinação das características em um mesmo vetor (características agrupadas) através da técnica de validação cruzada com 5-*folds* descrita na Seção 3.3.1.

A Tabela 4.9 mostra os classificadores monolíticos criados para cada vetor de características considerando os melhores resultados.

Tabela 4.9: Melhores resultados para cada vetor de característica.

		Base Variações		Base Web	
		<i>F-measure</i>	Desvio Padrão	<i>F-measure</i>	Desvio Padrão
Índices de Branquidão	$W_{Astme313}$	89,00%	4,29%	80,75%	5,30%
	W_{Berger}	92,00%	3,29%	82,00%	5,97%
	$W_{HunterLAB}$	92,50%	2,57%	77,50%	7,07%
	$W_{Stephansen}$	92,80%	2,64%	89,10%	4,60%
	W_{Taube}	90,00%	4,10%	82,25%	4,68%
Índice de Azul	$B_{ExcessoAzul}$	91,25%	4,02%	82,50%	4,61%
LBP		92,20%	2,74%	88,50%	4,80%
Perfil morfológico estendido		77,10%	3,03%	75,15%	3,38%

A figura 4.8 apresenta duas imagens em que a melhor característica individual, $W_{Stephansen}$, foi aplicada: a primeira é uma imagem com contexto de cidade que apresenta a sua respectiva *groundtruth* (a) e a imagem segmentada após a aplicação do método proposto (b); a segunda imagem representa o contexto de natureza com sua *groundtruth* (c) e a imagem segmentada (d).

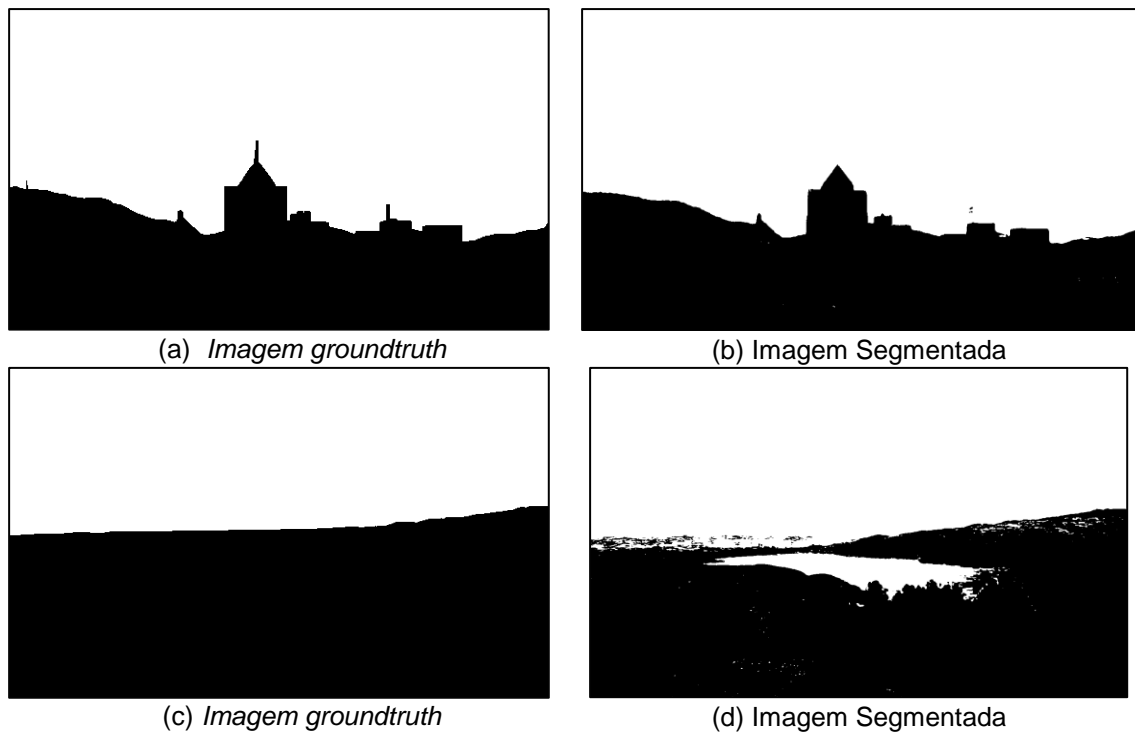


Figura 4.8: As imagens (a) e (c) representam a segmentação ideal e as imagens (b) e (d) correspondem as respectivas imagens segmentadas.

Observa-se na imagem (b) uma segmentação quase perfeita em relação a sua segmentação ideal e na imagem (c) um exemplo de uma segmentação ruim, o que justifica as taxas de acerto.

Na sequência algumas características foram agrupadas em um mesmo vetor para análise do *F-measure*.

Na Tabela 4.10 e 4.11 uniram-se os três melhores canais de cores e média encontrados nos experimentos com perfil morfológico estendido.

Tabela 4.10: Resultados para Perfil morfológico estendido para a Base Variações.

Característica	<i>F-measure</i>
Perfil morfológico estendido + BYV + Média	78,00%
Perfil morfológico estendido	77,10%

Tabela 4.11: Resultados para Perfil morfológico estendido para a Base Web.

Característica	F-measure
Perfil morfológico estendido + BYV + Média	76,00%
Perfil morfológico estendido	75,15%

Para as duas bases de imagens o agrupamento BYV + média + perfil morfológico estendido melhorou um pouco o *F-measure* em relação ao perfil morfológico estendido somente. Na sequência, a Tabela 4.12 combina as melhores características de branquidão e azul encontradas com o intuito de verificar se as taxas de acerto aumentam em relação ao melhor índice de branquidão.

Tabela 4.12: Resultados usando a combinação de características de branquidão e azul.

Bases de Imagens	F-measure		
	$W_{Stephansen} + B_{ExcessoAzul}$	$W_{Astme313} + W_{Berger} + W_{HunterLAB} + W_{Stephansen} + W_{Taube} + B_{ExcessoAzul}$	$W_{Stephansen}$
Base Variações	92,10%	89,30%	92,80%
Base Web	85,90%	80,70%	89,10%

Observa-se que a combinação do índice dos índices de branquidão com o índice de azul não superaram o índice $W_{Stephansen}$, que foi o melhor índice encontrado para as duas bases de imagens.

Os resultados somente com LBP foram menores que o índice de $W_{Stephansen}$, dessa forma, uniu-se o LBP com o melhor índice e com os cinco melhores índices na tentativa de obter uma melhor taxa de *F-measure*, conforme Tabela 4.13.

Tabela 4.13: Resultados usando a combinação de características de branquidão e azul com LBP.

Bases de Imagens	<i>F-measure</i>				
	LBP + $W_{Stephansen}$	LBP + $W_{Stephansen} + B_{ExcessoAzul}$	LBP + $W_{Astme313} + W_{Berger} + W_{HunterLAB} + W_{Stephansen} + W_{Taube}$	LBP + $W_{Astme313} + W_{Berger} + W_{HunterLAB} + W_{Stephansen} + W_{Taube} + B_{ExcessoAzul}$	$W_{Stephansen}$
Base Variações	92,60%	91,93%	91,10%	87,00%	92,80%
Base Web	88,90%	85,70%	86,80%	89,40%	89,10%

O algoritmo LBP com o melhor índice de branquidão tem um resultado aproximado do índice $W_{Stephansen}$ para a Base Variações com 91,93% e para a Base Web o melhor *F-measure* é LBP + $W_{Astme313} + W_{Berger} + W_{HunterLAB} + W_{Stephansen} + W_{Taube} + B_{ExcessoAzul}$ com 89,40%, porém o seu desvio padrão é 5,17% em relação a 4,60% do índice $W_{Stephansen}$.

Unindo o Perfil Morfológico Estendido (PME) com os índices de branquidão e de azul obtêm-se os resultados observados na Tabela 4.14.

Tabela 4.14: Resultados usando a combinação de características de branquidão e azul com perfil morfológico estendido.

Bases de Imagens	<i>F-measure</i>				
	PME + $W_{Stephansen}$	PME + $W_{Stephansen} + B_{ExcessoAzul}$	PME + $W_{Astme313} + W_{Berger} + W_{HunterLAB} + W_{Stephansen} + W_{Taube}$	PME + $W_{Astme313} + W_{Berger} + W_{HunterLAB} + W_{Stephansen} + W_{Taube} + B_{ExcessoAzul}$	$W_{Stephansen}$
Base Variações	85,40%	88,40%	88,78%	90,50%	92,80%
Base Web	83,05%	82,78%	84,60%	83,60%	89,10%

Observa-se que unindo o perfil morfológico estendido com os cinco índices de branquidão mais o melhor índice de azul obtêm-se um resultado melhor para a Base Variações com 90,50% de acerto e para a Base Web o perfil morfológico estendido

com os cinco melhores índices de branquidão com 84,60% de *F-measure*. Mas, nos dois casos, não superam a taxa de acerto do melhor índice de branquidão $W_{Stephansen}$ com 92,80% para a Base Variações e 89,10% para a Base Web.

Além do agrupamento de características na tentativa de aumentar as taxas de acerto, os cinco melhores índices foram testados com tamanhos de janelas diferentes, como mostra as Tabelas 4.15 e 4.16.

Tabela 4.15: Resultados usando a combinação de votos para a Base Variações com diferentes tamanhos de janelas.

Índice de Branquidão	<i>F-measure</i>			
	Janela 3x3	Janela 5x5	Janela 7x7	Janela 9x9
$W_{Astme313}$	88,30%	88,30%	88,30%	88,30%
W_{Berger}	91,50%	91,50%	91,40%	91,40%
$W_{HunterLAB}$	91,90%	91,90%	91,90%	91,90%
$W_{Stephansen}$	92,80%	92,80%	92,80%	92,80%
W_{Taube}	89,60%	89,60%	89,60%	89,60%

Tabela 4.16: Resultados usando a combinação de votos para a Base Web com diferentes tamanhos de janelas.

Índice de Branquidão	<i>F-measure</i>			
	Janela 3x3	Janela 5x5	Janela 7x7	Janela 9x9
$W_{Astme313}$	85,40%	85,40%	85,60%	85,60%
W_{Berger}	87,40%	87,40%	85,70%	85,70%
$W_{HunterLAB}$	82,60%	82,60%	82,60%	82,60%
$W_{Stephansen}$	89,10%	89,10%	89,10%	89,10%
W_{Taube}	86,60%	86,60%	86,60%	86,60%

Para as duas bases de imagens o *F-measure* se mantém com diferentes tamanhos de janela para o melhor índice de branquidão com janela 3x3.

4.8 Experimentos – Múltiplos Classificadores

Esta Seção apresenta duas abordagens que consistem na criação de classificadores para as melhores características e combinação de voto. A Seção 4.8.1 apresenta os melhores resultados para a primeira abordagem que consiste em treinar com 60% e testar com 40% das imagens misturadas, ou seja, sem a divisão das imagens em categorias, e a Seção 4.8.2 que apresenta os resultados para a segunda abordagem que segue o mesmo protocolo (60% treino e 40% teste) para as imagens divididas em categorias.

4.8.1 Primeira abordagem - Análise dos resultados com combinação de classificadores para a junção das Bases de Imagens

Para cada base, as imagens foram misturadas e, posteriormente separadas em 60% das imagens para treinamento, gerando um classificador SVM para cada uma das cinco melhores características de branquidão. As imagens de treino foram divididas em duas fatias com 400 pixels de amostra, a descrição detalhada encontra-se na Seção 3.3.2. Na sequência, cada classificador gerado foi testado com um conjunto de 40% das imagens. A partir dos resultados obteve-se o voto majoritário ou voto da maioria. As Tabelas 4.17 e 4.18 e as Figuras 4.9 e 4.10 apresentam os experimentos realizados para a base de imagens chamada Base Variações e Base Web, respectivamente.

Tabela 4.17: Resultados usando a combinação de classificadores para a Base Variações.

Índice de Branquidão	<i>F-measure</i>
$W_{Astme313}$	88,30%
W_{Berger}	91,50%
$W_{HunterLAB}$	91,90%
$W_{Stephansen}$	92,80%
$W_{Taubé}$	89,60%
Combinação de voto	91,11%

```

=== Confusion Matrix ===
      a      b  <-- classified as
183318  9162 |      a = 0
18618 173862 |      b = 1

```

Figura 4.9: Matriz de confusão do índice $W_{Stephansen}$.

Tabela 4.18: Resultados usando a combinação de classificadores para a Base Web.

Índice de Branquidão	<i>F-measure</i>
$W_{Astme313}$	85,40%
W_{Berger}	87,40%
$W_{HunterLAB}$	82,60%
$W_{Stephansen}$	89,10%
W_{Taube}	86,60%
Combinação de voto	78,16%

```

=== Confusion Matrix ===
      a      b  <-- classified as
172972 19752 |      a = 0
22384 170095 |      b = 1

```

Figura 4.10: Matriz de confusão do índice $W_{Stephansen}$.

Para as duas bases de imagens observa-se que o resultado com o esquema de combinação de votos diminuiu em relação a alguns índices de branquidão. No caso da Base Variações o melhor resultado foi o índice $W_{Stephansen}$, bem próximo do resultado com combinação de votos e para a Base Web o melhor resultado também foi o índice de $W_{Stephansen}$, com uma diferença maior em relação ao esquema de votos.

4.8.2 Segunda abordagem - Análise dos resultados com a pré-classificação das imagens em categorias

Esta segunda abordagem segue o mesmo protocolo realizado na primeira abordagem, ou seja, criando para os cinco melhores índices de branquidão um classificador para usar o esquema de combinação de classificadores, mas agora

mantendo a divisão das imagens em categorias para cada base separadamente, pois as imagens possuem diferentes contextos.

As Tabelas 4.19 a 4.22 e as Figuras 4.11 a 4.14 apresentam os experimentos realizados para a Base Variações e Base Web para as categorias de Cidade, Mar/Porto, Natureza/Montanha e Rodovia/Estrada, respectivamente. A escolha das quatro categorias foi sugerida a partir das imagens usadas na literatura para abordagens de segmentação de linha do horizonte.

Tabela 4.19: Resultados usando a combinação de classificadores para a categoria de Cidade.

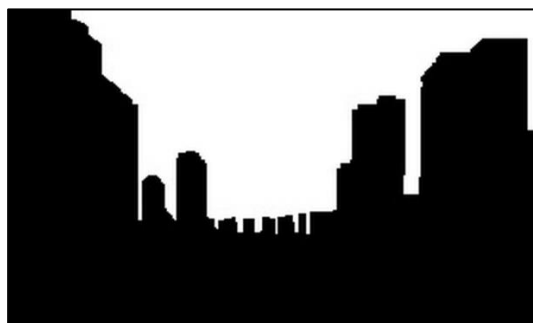
Índice de Branquidão	Base Variações	Base Web
$W_{Astme313}$	95,40%	85,40%
W_{Berger}	96,90%	89,50%
$W_{HunterLAB}$	96,60%	86,00%
$W_{Stephansen}$	97,10%	91,31%
W_{Taube}	95,90%	86,31%
Combinação de voto	96,42%	87,94%



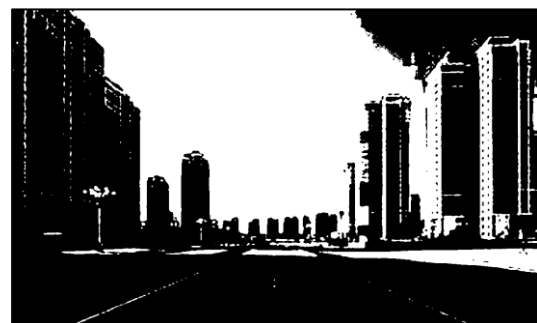
(a) Imagem *groundtruth* - Base Variações



(b) Imagem Segmentada ($W_{Stephansen}$)



(c) Imagem *groundtruth* - Base Web



(d) Imagem Segmentada ($W_{Stephansen}$)

Figura 4.11: As imagens (a) e (c) representam a segmentação ideal e as imagens (b) e (d) correspondem as respectivas imagens segmentadas.

Tabela 4.20: Resultados usando a combinação de classificadores para a categoria de Mar/Porto.

Índice de Branquidão	Base Variações	Base Web
$W_{Asme313}$	89,50%	81,80%
W_{Berger}	92,00%	80,11%
$W_{HunterLAB}$	95,00%	77,40%
$W_{Stephansen}$	93,20%	82,80%
W_{Taube}	90,20%	83,70%
Combinação de voto	91,82%	82,24%



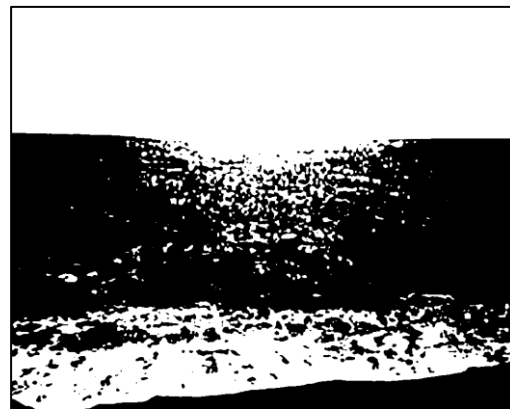
(a) Imagem *groundtruth* - Base Variações



(b) Imagem Segmentada ($W_{HunterLAB}$)



(c) Imagem *groundtruth* - Base Web



(d) Imagem Segmentada (W_{Taube})

Figura 4.12: As imagens (a) e (c) representam a segmentação ideal e as imagens (b) e (d) correspondem as respectivas imagens segmentadas.

Tabela 4.21: Resultados usando a combinação de classificadores para a categoria de Natureza/Montanha.

Índice de Branquidão	Base Variações	Base Web
$W_{Astme313}$	83,80%	91,20%
W_{Berger}	87,40%	91,80%
$W_{HunterLAB}$	92,00%	80,00%
$W_{Stephansen}$	89,40%	92,00%
W_{Taube}	85,80%	90,40%
Combinação de voto	87,98%	88,60%



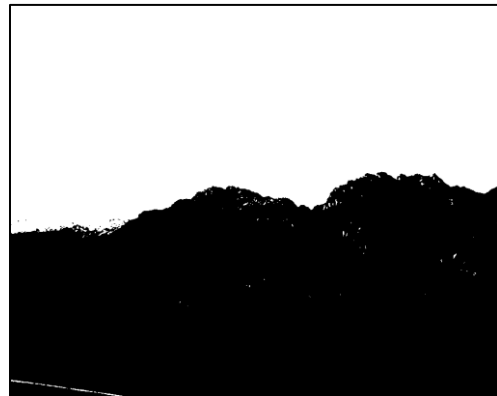
(a) Imagem *groundtruth* – Base Variações



(b) Imagem Segmentada ($W_{HunterLAB}$)



(c) Imagem *groundtruth* – Base Web



(d) Imagem Segmentada ($W_{Stephansen}$)

Figura 4.13: As imagens (a) e (c) representam a segmentação ideal e as imagens (b) e (d) correspondem as respectivas imagens segmentadas.

Tabela 4.22: Resultados usando a combinação de classificadores para a categoria de Rodovia/Estrada.

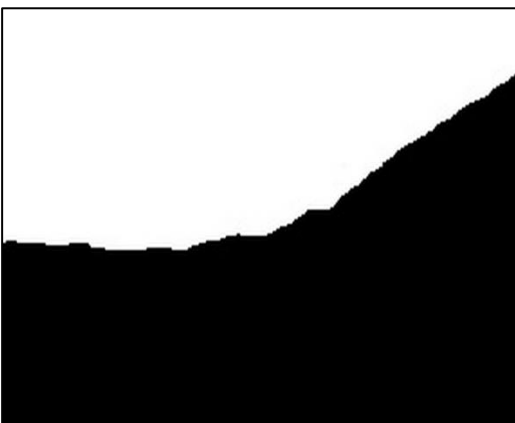
Índice de Branquidão	Base Variações	Base Web
$W_{Astme313}$	88,50%	85,30%
W_{Berger}	90,10%	84,80%
$W_{HunterLAB}$	90,20%	87,50%
$W_{Stephansen}$	91,60%	89,70%
W_{Taube}	88,30%	87,00%
Combinação de voto	90,04%	86,43%



(a) Imagem *groundtruth* – Base Variações



(b) Imagem Segmentada ($W_{Stephansen}$)



(c) Imagem *groundtruth* – Base Web



(d) Imagem Segmentada ($W_{Stephansen}$)

Figura 4.14: As imagens (a) e (c) representam a segmentação ideal e as imagens (b) e (d) correspondem as respectivas imagens segmentadas.

Os melhores resultados levando em consideração a divisão da Base Variações em categorias são o índice de $W_{Stephansen}$ para a categoria de Cidade e

Rodovia/Estrada, o índice de $W_{HunterLAB}$ para as categorias de Mar/Porto e Natureza/Montanha.

Para a Base Web, os melhores resultados levando em consideração a divisão desta base em categorias são o índice de W_{Taube} para a categoria de Mar/Porto e o índice de $W_{Stephansen}$ para as demais categorias.

Levando em consideração que a diferença entre os melhores índices de branquidão de cada categoria é muito pequena, mas existe, pode-se concluir, nesse caso, que o índice $W_{Stephansen}$ pode ser usado para cada categoria separadamente ou para todas as categorias misturadas.

4.9 Discussão dos Resultados

As melhores características individuais testados com validação cruzada foram os índices de $W_{Astme313}$, W_{Berger} , $W_{HunterLAB}$, $W_{Stephansen}$ e W_{Taube} . A partir da identificação desses cinco melhores índices aposta-se na união do(s) melhor(es) índice(s) de branquidão com características citadas na literatura. O agrupamento de características com índices de branquidão em um vetor com *F-measure* maior que o melhor índice de branquidão encontrado ($W_{Stephansen}$) corresponde à sequência LBP + $W_{Astme313} + W_{Berger} + W_{HunterLAB} + W_{Stephansen} + W_{Taube} + B_{ExcessoAzul}$ para a Base Web com um *F-measure* de 89,40% (Tabela 4.13).

A partir dos resultados obtidos aposta-se em duas abordagens: uma com todas as imagens de cada base misturadas e uma segunda abordagem com a divisão destas imagens em categorias. A primeira abordagem é geralmente adotada na literatura, em que o melhor resultado foi o índice de $W_{Stephansen}$ para as duas bases de imagens. A segunda abordagem apresenta resultados interessantes e distintos para as duas bases de imagens e que permitem observar a importância de dividir as imagens em categorias, pois há alteração do contexto de chão, o que leva a resultados diferentes e com maior taxa de acerto, levando em consideração que a diferença é em média 2 e 3 pontos percentuais, conforme Tabelas 4.23 e 4.24.

Tabela 4.23: Resultados para a Base Variações.

Categoria	Índice de Branquidão	F-measure	Desvio Padrão
Cidade	$W_{Stephansen}$	97,10%	0,60%
Mar/Porto	$W_{HunterLAB}$	95,00%	0,78%
Natureza/Montanha	$W_{HunterLAB}$	92,00%	0,77%
Rodovia/Estrada	$W_{Stephansen}$	91,60%	0,62%

Tabela 4.24: Resultados para a Base Web.

Categoria	Índice de Branquidão	F-measure	Desvio Padrão
Cidade	$W_{Stephansen}$	91,31%	1,38%
Mar/Porto	W_{Taube}	83,70%	1,03%
Natureza/Montanha	$W_{Stephansen}$	92,00%	1,36%
Rodovia/Estrada	$W_{Stephansen}$	89,70%	1,34%

Nas duas abordagens apresentadas o resultado com a combinação de votos não apresentou um resultado melhor que os índices testados separadamente. Em alguns casos, o resultado foi bem próximo ao melhor índice, mas não o superou.

CAPÍTULO 5

CONSIDERAÇÕES FINAIS

A detecção da linha do horizonte em imagens digitais é uma tarefa difícil, pois há uma variação de luminosidade nas imagens, obstáculos como nuvens, chuva, neblina e diferentes contextos de chão, como prédios, árvores, antenas e etc.

Esta pesquisa pode ser dividida em duas etapas: a primeira que contempla classificadores monolíticos através da técnica de validação cruzada e a segunda que contempla duas abordagens diferentes para segmentar céu e chão com combinação (fusão) de classificadores.

Na primeira etapa, várias características foram testadas, algumas já citadas na literatura para segmentação entre céu e chão. Estes experimentos foram usados, principalmente, para verificar a eficiência dos índices de branquidão ainda não aplicados nesse tipo de abordagem. Onze índices de branquidão foram usados nos experimentos e os que apresentaram os melhores resultados para as duas bases foram $W_{Astme313}$, W_{Berger} , $W_{HunterLAB}$, $W_{Stephansen}$ e W_{Taube} , com *F-measure* de 89%, 92%, 93%, 93% e 90% para cada índice da Base Variações, e 81%, 82%, 78%, 83% e 82% para cada índice da Base Web, respectivamente, conforme Tabelas 4.3 e 4.4. Ainda nesta etapa, algumas características foram agrupadas a fim de aumentar o *F-measure*. Destes agrupamentos, o único que apresentou uma melhor taxa descreve a sequencia $LBP + W_{Astme313} + W_{Berger} + W_{HunterLAB} + W_{Stephansen} + W_{Taube} + B_{ExcessoAzul}$ para a Base Web com um *F-measure* de 89,40%, como mostra a Tabela 4.13.

A segunda etapa contempla duas abordagens, a primeira que tem as imagens misturadas para a Base Variações e Base Web. Esta abordagem é comum na literatura, cria-se um classificador avaliando diferentes características de branquidão e o esquema de votos. Para as duas bases de imagens o melhor índice é $W_{Stephansen}$,

com *F-measure* de 92,80% para a Base Variações e 89,10% para a Base Web. É importante destacar que a fórmula $W_{Stephansen}$ consiste em $2*B-R$ para cada pixel, ou seja, 2 vezes o canal azul, por isso seu melhor desempenho em relação aos outros índices de branquidão.

A segunda abordagem consiste na pré-classificação das imagens em categorias (Cidade; Mar/Porto; Natureza/Montanha; Rodovia/Estrada), para então, testar as imagens da categoria de Cidade para o melhor classificador de Cidade, testar as imagens da categoria de Mar/Porto para o melhor classificador de Mar, testar as imagens da categoria de Natureza/Montanha para o melhor classificador para a categoria e, por fim, testar as imagens da categoria de Rodovia/Estrada com o melhor classificador para esta categoria. A escolha das quatro categorias foi sugerida a partir das categorias citadas na literatura para abordagens de segmentação de linha do horizonte. Nesta abordagem obtiveram-se como melhores resultados, conforme Tabelas 4.19 a 4.22, os índices de $W_{Stephansen}$ para as categorias de Cidade com 97% e Rodovia/Estrada com 92% e $W_{HunterLAB}$ para as categorias de Mar/Porto com 95% e Natureza/Montanha com 92% de taxa de acerto para a Base Variações. Para a Base Web os melhores índices são $W_{Stephansen}$ para as categorias de Cidade com 91%, Natureza/Montanha com 92% e Rodovia/Estrada com 90% e o índice W_{Taube} para a categoria de Mar/Porto com 84% de taxa de acerto. Aqui vale destacar a importância de separar as imagens em categorias, pois apresentam diferentes contextos de chão no caso da Base Web, apresentando taxas de *F-measure* variadas.

Para as duas abordagens o *F-measure* com o esquema de combinação de votos diminuiu em relação a alguns índices de branquidão. Em alguns casos, o melhor resultado do respectivo índice foi bem próximo do resultado com combinação de votos. Dessa forma, tem-se a taxa geral do sistema para cada base (Variações e Web) e a taxa individual de acerto por categoria que apresenta uma abordagem promissora com os índices de branquidão ainda não citados na literatura para este tipo de abordagem de segmentação entre céu e chão.

Vale destacar ainda que, analisando os experimentos com os conjuntos de dados construídos para esta pesquisa disponível no link³, o clima e o tempo tem sobre o desempenho dos métodos.

O tempo computacional varia de acordo com os contextos de céu/chão, no entanto, o tempo médio varia de três a cinco segundos por imagem.

Como pesquisas futuras sugere-se aumentar o número de imagens para cada base e aplicar outros classificadores para análise e comparação dos resultados.

³ <<https://drive.google.com/open?id=1s9zIKU-mPutFJMw2vVMYmQ-Y9EInWleU>>

REFERÊNCIAS

AHMAD, Touqeer; BEBIS, George; REGENTOVA, Emma; NEFIAN, Ara. **A Machine Learning Approach to Horizon Line Detection Using Local Features**. 9th International Symposium on Visual Computing, Advances in Visual Computing, Volume 8033 of the series Lecture Notes in Computer Science p. 181-193, 2013.

AHMAD, Touqeer; BEBIS, George; REGENTOVA, Emma; NEFIAN, Ara. **An Experimental Evaluation of Different Features and Nodal Costs for Horizon Line Detection**. ISVC 2014, Part I, LNCS 8887, pp. 193–205, 2014.

ANANTH, Christo; SENTHILKANI, A. S.; PRAGHASH, .k; RAJA, Chakka; JOHN, Jerrin; ANNADURAI, I. **Overlap Wavelet Transform for Image Segmentation**. International Journal of Electronics Communication and Computer Technology (IJECCCT) Volume 4 Issue 3. 2014.

AYADI, Mehdi; SUTA, Loreta; SCUTURICI, Mihaela; MIGUET, Serge; AMAR, Chokri Bem. **A Parametric Algorithm for Skyline Extraction**. Springer International Publishing. Advanced Concepts for Intelligent Vision Systems: 17th International Conference, ACIVS 2016, Lecce, Italy, October 24-27, 2016, Proceedings, Oct 2016, Lecce, Italy. 10016, pp.604-615, 2016, Lecture Notes in Computer Science. <<http://acivs.org/acivs2016/>>. <10.1007/978-3-319-48680-2_53>. <hal-01385542>.

BAO, G. Q.; XIONG, S. S.; ZHOU Z. Y. **Vision-based horizon extraction for micro air vehicle flight control**, IEEE Transactions on Instrumentation and Measurement, vol. 54, no. 3, 2005, pp. 1067-1072.

BEBIS, George; YAZDANPANA, Ali Pour; REGENTOVA, Emma E.; MANDAVA, Ajay Kumar; AHMAD, Touqeer. **Sky Segmentation by Fusing Clustering with Neural Networks**. (Eds.): ISVC 2013, Part II, LNCS 8034, pp. 663–672, 2013.

BIMBO, Del, A. **Visual Information Retrieval**, Morgan Kaufmann (1999).

BOROUJENI, Nasim Sepehri; ETEMAD S. Ali; WHITEHEAD Anthony. **Robust Horizon Detection using Segmentation for UAV Application**. Ninth Conference on Computer and Robot Vision, pp 346 – 352, 2012.

BRAND, J.; MASON, J. **A comparative assessment of three approaches to pixel-level human skin detection**. International Conference on Pattern Recognition, Volume 1: pp 1056-1059, 2000.

BREIMAN, Leo. Bagging Predictors. **Machine Learning**, v. 24, p. 123-140, 1996.

CASTRO, Ana I. de. EHSANI, Reza; PLOETZ, Randy C.; CRANE, Jonathan H.; BUCHANON, Sherrie. **Detection of Laurel Wilt Disease in Avocado Using Low Altitude Aerial Imaging**. PLoS ONE 10(4), pp.1-13, 2015.

CHIU, Chung-Cheng; LIU, Yuh-Jiun; CHIU, Sheng-Yi; CHANG, Hsing-Chien; HSU, Chia-Lun. **A Skyline Detection Algorithm for Use in Different**. 978-1-4673-9985-2/16/\$31.00 ©2016 IEEE.

CROON, G.C.H.E., DE WAGTER, C., REMES, B.D.W., RUIJSINK, R. **Sky Segmentation Approach to Obstacle Avoidance**. IEEE Aerospace Conference 2011, p. 1 - 31, 2011.

FACON, J. **Mathematical Morphology: theory and practice** (In Portuguese). Editor Facon Jacques, 1996.

FACON, Jacques; BEUREN, Arlete Teresinha; MATTOS, Flávia de; BRITTO JR., Alceu De Souza; SOUZA, Bruno Miguel Nogueira de. **Supervised Approach to Sky and Ground Classification using Whiteness-based Features**. 16th Mexican International Conference on Artificial Intelligence, 2017. Disponível em: <<https://www.micai.org/2017/pre-print/LNAI/99990578.pdf>>.

FACON, Jacques; BEUREN, Arlete Teresinha; MATTOS, Flávia de. **Unsupervised approach to segment Sky/Ground Regions using whiteness indexes**. Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications. 22nd Iberoamerican Congress, CIARP 2017, Valparaíso, Chile, November 7–10, 2017, Proceedings.

GATOS, B.; NTIROGIANNIS, K.; PRATIKAKIS, I. **ICDAR 2009 Document Image Binarization Contest** (DIBCO 2009). International Conference on Document Analysis and Recognition, pp 1375–1382, 2009.

GERSHIKOV, Evgeny, **Is Color Important for Horizon Line Detection?**. The 2014 International Conference on Advanced Technologies for Communications (ATC'14), PP 262- 267, 2014.

GESTEL, Tony Van; SUYKENS, Johan A. K.; BAESSENS, Bart; VIAENE, Stijn; VANTHIENEN, Jan; DEDENE, Guido; MOOR Bart De; VANDEWALLE, Joos. **Benchmarking least squares support vector machine classifiers**, Mach. Learn. 54 (2004), no. 1, 5–32.

HOIEM, D.; EFROS, A. A., HEBERT, M. **Geometric context from a single image**. In IEEE International Conference on Computer Vision, 2005.

HERDTWECK C, WALLRAVEN C. **Horizon estimation: Perceptual and computational experiments**. APGV Symposium on Applied Perception in Graphics and Visualization (APGV), pages 49-56, 2010.

HUGUES Olivier, CIEUTAT Jean-Marc, GUITTON Pascal. **Merging Video and Charts for Augmented Reality by Vision and Inertial Fusion to Perform Horizon Tracking**. Hal-00787072, version 1 – 11, Feb 2013.

JOINER, Andrew; HOPKINSON, Ian; DENG, Yan; WESTLAND, Stephen. **A review of tooth color and whiteness**. Journal of Dentistry Volume 36, pp. s2-s7, 2008.

KULKARNI, Nilima. **Color Thresholding Method for Image Segmentation of Natural Images**. I.J. Image, Graphics and Signal Processing, 2012, 1, 28-34 Published Online February 2012 in MECS (<http://www.mecspress.org/>) DOI: 10.5815/ijigsp.2012.01.04.

LIANG, Jie; ZHOU, Jun; GAO, Yongsheng. **Tensor Morphological Profile for Hyperspectral Image Classification**. Conference: 2016 IEEE International Conference on Image Processing (ICIP), 2016.

LIANG, Dong; ZHANG, Weigang; HUANG, Qiaozhong; YANG, Fei. **Robust Sea-sky-line Detection for Complex Sea Background**. National Natural Science Foundation of China: 61303153. 978-1-4673-9088-0/15 ©2015 IEEE.

LIAO T.W. **Clustering of time series data—a survey**. Pattern Recognition Vol 38, No 11, pp 1857-1874, 2005.

LIE, W.-N., LIN, T.C.-I., LIN, T.-C., HUNG, K.-S. **A robust dynamic programming algorithm to extract skyline in images for navigation**. Pattern Recognition Letters 26, 221–230, 2005.

LIMA, A. R. G. **Máquinas de vetores suporte na classificação de impressões digitais**. 2002, 81p. Dissertação (Mestrado em Ciência da Computação) - Universidade Federal do Ceará, Fortaleza, Ceará.

LIN, Juan. **Factors Affecting the Perception and Measurement of Optically Brightened White Textiles**. PhD Thesis, Graduate Faculty of North Carolina State University Raleigh, North Carolina, p. 366, 2013.

LIPSCHUTZ Ilan, GERSHIKOV Evgeny, MILGROM Benjamin. **New Methods for Horizon Line Detection in Infrared and Visible Sea Images**. International Journal Of Computational Engineering Research (ijceronline.com) Vol. 3 Issue. 3. pp 226 - 233, 2013.

LOUI, Alexander C.; SAVAKIS, Andreas E. **Automatic Image Event Segmentation and Quality Screening for Albuming Applications**. IEEE International Conference on Multimedia and Expo, New York City, New York, pp 1-4, July 2000.

LU, C.; LIN, D.; JIA, J.; TANG, C.K. **Two-class weather classification**. In Computer Vision and Pattern Recognition, 2014.

MIHAIL, Radu P.; WORKMAN, Scott; BESSINGER, Zach . **Sky segmentation in the wild: An empirical study**. *Applications of Computer Vision (WACV), 2016 IEEE Winter Conference on*. IEEE, 2016.

MISHRA, Pavan Kumar; MAURYA, Sanjay Kumar; SINGH, Ravindra Kumar, MISRA, Arun Kumar. **A semi automatic plant identification based on digital leaf and flower Images**. IEEE-International Conference On Advances In Engineering, Science And Management (ICAESM -2012), pp 68-73, 2012.

OSUNA, E.; FREUD, R.; GIROSI, F. **Support Vector Machines: training and applications**. MIT Artificial Intelligence Memo 1602; MIT A. I.Lab,1997.

OJALA, T.; PIETIKAINEN, M.; MAENPAA, T. **Multiresolution gray-scale and rotation invariant texture classification with local binary patterns**. Pattern Analysis and Machine Intelligence, IEEE Transactions on, [S.l.], v.24, n.7, p.971–987, 2002.

OTSU N. **A Threshold Selection Method from Gray-level Histograms.** IEEE Transactions on Systems, Man and Cybernetics, v. SMC 9, no 1, pp.62-66, 1979.

PORZI, Lorenzo; BULÒ, Samuel Rota; RICCI, Elisa. **A Deeply-Supervised Deconvolutional Network for Horizon Line Detection.** MM '16, October 15-19, 2016, Amsterdam, Netherlands c 2016 ACM. ISBN 978-1-4503-3603-1/16/10. DOI: <http://dx.doi.org/10.1145/2964284.2967198>.

PUEBLA Claudio. **Whiteness Assessment: A Primer.** Axiphos GmbH. Germany, p. 64, 2006.

SANGWINE, S.; HORNE, R. **The colour image processing handbook.** London Chapman Hall, 1998.

SANKUR B.; SEZGIN, M. **Selection of Thresholding methods for non-destructive testing applications.** International Conference On Image Processing, ICIP2001, pp 764-767, 2001.

SASSI, Ameni; AMAR, Chokri Ben; MIGUET, Serge. **Skyline-based Approach for Natural Scene Identification.** 13th ACS/IEEE International Conference on Computer Systems and Applications (AICCSA 2016), Nov 2016, Agadir, Morocco. 2016, <<http://www.aiccsa.net/AICCSA2016/>>. <hal-01385537>.

SEZGIN, M.; SANKUR, B. **Selection of thresholding methods for non destructive testing application.** ICIP2001, v. 3, p. 764 - 767, 2001.

SCARPEL, R. A. **Utilização de Support Vector Machine em previsão de insolvência de empresas.** In: Anais do XXXVII Simpósio Brasileiro de Pesquisa Operacional. Gramado, RS, 2005.

TIGHE, J.; LAZEBNIK, S. **Superparsing: scalable nonparametric image parsing with superpixels.** In European Conference on Computer Vision, 2010.

VERBICKAS, Rytis; WHITEHEAD, Anthony. **Sky and Ground Detection Using Convolutional Neural Networks.** Proceedings of the International Conference on Machine Vision and Machine Learning Prague, Czech Republic, August 14-15, 2014.

WATANABE, S.; VINCENT, W. F.; REUTER, J.; HOOK, S. J.; SCHLADOW, S. G. **A quantitative blueness index for oligotrophic waters: Application to Lake Tahoe, California–Nevada.** Limnol. Oceanogr. Methods, 14: 100–109. doi:10.1002/lom3.10074. 2016.

WILLIAMS, Stephen; HOWARD, Ayanna M. **Horizon Line Estimation In Glacial Environments Using Multiple Visual Cues.** IEEE International Conference on Robotics and Automation, ICRA 2011, Shanghai, China, pp 5887—5892, May 2011.

X-RITE. **Color iQC and Color iMatch Color Calculations Guide.** Version 8.0, p. 31, July 2012.

YAMASHITA, M.; YOSHIMURA, M. **Ground-based cloud observation for satellite-based cloud discrimination and its validation.** International Archives of the

Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XXXIX-B8, pp. 137- 140, 2012.

YU-QUAN, Zhu; GENG, Chen; HAI-PING, Yu. **Dynamic weighting ensemble classifiers based on cross-validation**. Neural Computing & Applications, v. 20, n. 3, p. 309-317, 2011.

ZAHEDI, Adel; Beheshti, Ali Asghar. **Horizon Detection Using A New Fuzzy Algorithm**. Journal: IET Image Processing Manuscript ID: IPR-2008-0191. 20-oct-2008, p. 1-19.