Universidade Federal do Piauí
Centro de Ciências da Natureza
Programa de Pós-Graduação em Ciência da Computação

Método adaptativo com limiarização local para segmentação de veículos

Kalyf Abdalla Buzar Lima

Número de Ordem PPGCC: M001

Teresina-PI, 01 de março de 2016

Kalyf Abdalla Buzar Lima

Método adaptativo com limiarização local para segmentação de veículos

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Ciência da Computação da UFPI (área de concentração: Sistemas de Computação), como parte dos requisitos necessários para a obtenção do Título de Mestre em Ciência da Computação.

Universidade Federal do Piauí – UFPI

Centro de Ciências da Natureza

Programa de Pós-Graduação em Ciência da Computação

Orientador: Prof. Dr. Kelson Rômulo Teixeira Aires

Teresina-PI

01 de março de 2016

FICHA CATALOGRÁFICA
Serviço de Processamento Técnico da Universidade Federal do Piauí
Biblioteca Setorial do CCN

L732m Lima, Kalyf Abdalla Buzar.
Método adaptativo com limiarização local para segmentação de veículos / Kalyf Abdalla Buzar Lima. – Teresina, 2016.
106f.: il. color

Dissertação (Mestrado) – Universidade Federal do Piauí, Centro de Ciências da Natureza, Pós-Graduação em Ciência da Computação, 2016.
Orientador: Prof. Dr. Kelson Rômulo Teixeira Aires.

1. Processamento de Imagens. 2. Segmentação de Veiculos. 3. Modelo de Mistura de Gaussianas. I. Título.

CDD 621.367

Método adaptativo com limiarização local para segmentação de veículos

KALYF ABDALLA BUZAR LIMA

Dissertação apresentada ao Programa de Pós–Graduação em Ciência da Computação do Centro de Ciências da Natureza da Universidade Federal do Piauí, como parte integrante dos requisitos necessários para obtenção do grau de Mestre em Ciência da Computação.

Aprovado por:

Prof. Kelson Rômulo Teixeira Aires

(Presidente da Banca Examinadora)

Prof. Iális Cavalcante de Paula Júnior

(Examinador Externo)

Prof. Ricardo de Andrade Lira Rabêlo

(Examinador Interno)

Prof. Rodrigo Melo de Souza Véras

(Examinador Interno)

Teresina – PI, 01 de março de 2016

Agradecimentos

Agradeço a Deus, Alá, Zeus, Rá, ..., porque independente de rótulos, continuam sendo a Força que desperta em todos nós.

A minha mãe Maria de Fátima Beliche Buzar, por me mostrar que anistias não são fáceis, mas os olhos devem ser direcionados para o futuro e não para o passado.

Ao meu pai Luis Henrique Gomes Lima, por me mostrar que erros são parte do processo de aprendizado da vida e que há duas trilhas: conviver com eles ou esforçar-se para repará-los.

Agradeço a Juliana Raquel Rodrigues da Silva Andrade, por me ensinar mais lições que eu pude aprender, que caminhos se entrelaçam para que essas lições possam ser vividas.

Agradeço ao meu orientador, Doutor Kelson Rômulo Teixeira Aires, pelos anos de orientação, as oportunidades dadas e por apresentar a área na qual dedico-me até hoje.

Aos meus amigos, que na medida do possível, me alegraram nas aflições, me ouviram nas frustrações e me ergueram nas quedas.

"Não creio que haja uma emoção,
mais intensa para um inventor
do que ver suas criações funcionando.
Essa emoção faz você esquecer de comer,
de dormir, de tudo."
(Nikola Tesla)

Resumo

A segmentação de veículos é um problema não linear que tem sido atacado na literatura utilizando técnicas de extração de plano de fundo. Nessas técnicas, os dados são classificados em duas classes: plano de fundo e primeiro plano. Os veículos a serem segmentados, neste caso, compõem parte do primeiro plano. Dois métodos muito utilizados para estimar o plano de fundo em ambientes não-controlados são o modelo de mistura de gaussianas de Stauffer e Grimson, e o modelo não-paramétrico de Horprasert. Embora esses métodos apresentem resultados satisfatórios, no problema da segmentação de veículos, a classificação é deficiente e ocorrem erros nesse processo. O sistema proposto neste trabalho atua na redução dessas deficiências, entre elas, os erros de classificação nas regiões internas dos veículos. Para aprimorar o processo de segmentação e reduzir as falhas, o sistema proposto conta com duas etapas: a filtragem e a estimação de limiar. Enquanto a filtragem busca corrigir e detectar os erros de classificação, a estimação de limiar realiza uma realimentação na próxima iteração da classificação. A realimentação é realizada com o objetivo de melhorar a classificação do primeiro plano, atuando unicamente nessas áreas. Para avaliar a proposta, quatro bases foram criadas. Essas bases avaliam os métodos testados considerando as mudanças e os fatores de variação ao longo do dia. Dessa forma, os resultados mostraram que o sistema proposto é uma solução prática viável para ser utilizada. Os resultados obtidos na avaliação foram maiores do que os outros métodos. No entanto, ainda mais importante que os acertos é a estabilização do comportamento da proposta sob variação dos fatores ambientais. A estabilidade de comportamento permite a definição de limiares fixos que funcionem em todos os contextos, o que não é indicado para os outros métodos testados.

Palavras-chaves: Segmentação de veículos. Modelo de mistura de gaussiana. Limiarização.

Abstract

Vehicle segmentation is a non-linear problem that has been tackled by background subtraction techniques. In these techniques, the pixels are classified into two cluster: the background and the foreground, so vehicles are part of the foreground. Gaussian mixture models, as Stauffer and Grimson's method, and non-parametric approaches, as Horprasert's method, are two popular methods used to estimate the background on non-controlled environments. Although these methods present satisfactory results, the segmentation process results on miclassification, mainly on the vehicle regions. The misclassification occurs due to illumination changes that affect the background estimation. To improve the segmentation results on theses regions, the proposed system includes two steps: the filtering and the threshold estimation. While the filtering fixes noises and detects misclassification, While the filtering fixes noises and detects misclassification, the threshold estimation selects values to improve the vehicle segmentation on next iterations. The threshold estimation works with a feedback approach in the classification step. As the goal of the feedback is the improvement of the vehicle regions, the threshold estimation is limited only to foreground regions. This work created four test bases to evaluate. These bases are a set of videos, image of selected regions and ground truth outputs. They includes illumination changes and other variation factors throughout the day. The experiments allow to verify the viability of the system on real environments. The feedback approach presents positive impacts on probabilistic classification and it stabilizes the behavior of the classification. The behavior stabilization allows the selection of a fixed value for all test bases. This is not possible for the other methods because they present different behaviors, as well as different regions of optimal solutions.

Keywords: Vehicle segmentation. Gaussian Mixture Model. Thresholding.

Lista de ilustrações

Lista de tabelas

Lista de abreviaturas e siglas

ABESE	Associação Brasileira das Empresas de Sistemas Eletrônicos de Segurança
CDF	*Cumulative Density Function*
ECR	*Edge Change Ratio*
EM	*Expectation Maximization*
FN	Falso Negativo
FP	Falso Positivo
GMM	*Gaussian Mixture Model*
IBGE	Instituto Brasileiro de Geografia e Estatísticas
ITS	*Intelligent Transportation System*
MHT	*Multiple Hypothesis Tracking*
MLE	*Maximum Likelihood Estimation*
PDF	*Probability Density Funciton*
PMF	*Probability Mass Function*
PMM	*Poisson Mixture Model*
ROC	*Receiver Operating Characteristic*
SRM	*Statistical Region Merging*
VN	Verdadeiro Negativo
VPN	Valor Preditivo Negativo
VPP	Valor Preditivo Positivo
VP	Verdadeiro Positivo

Lista de símbolos

α	Taxa de atualização dos parâmetros do modelo estatístico
β	Distorção de brilho do *pixel*
Γ_1	Conjunto de valores ótimos para o limiar l_1
Γ_2	Conjunto de valores ótimos para o limiar l_2
$\hat{\vec{\mu}}$	Vetor de média da amostra
$\hat{\pi}_m$	Peso de uma distribuição estatística
$\hat{\sigma}$	Desvio-padrão da amostra
$\mathcal{N}$	Função de densidade de probabilidade gaussiana
$\mathcal{X}_T$	Conjunto de evidências da amostra até o instante
μ_B	Média da componente azul na amostra
μ_G	Média da componente verde na amostra
μ_R	Média da componente vermelha na amostra
$\vec{x}_t$	Vetor do *pixel* em um instante
σ_0	Desvio-padrão inicial pré-definido
τ_{CD}	Limiar da distorção de cromaticidade
τ_{lo}	Limiar da distorção de brilho
a_r	Área calculada dos contornos mais externos das regiões segmentadas
BG	Estado *background* da variável aleatória de classificação
c_{pp}	Proporção das probabilidades a priori de primeiro plano e plano de fundo
c_f	Limiar para o somatório de pesos das distribuições
C_T	Proporção de evidências na amostra.
CD	Distorção de cromaticidade do *pixel*
d_{prob}	Limiar de classificação que considera p_{thr} e c_{pp}

d_{thr}	Limiar correspondente à d_{prob} na distância Mahalanobis
e_r	Contornos mais externos na hierarquia das regiões segmentadas
FG	Estado *foreground* da variável aleatória de classificação
hu_n	Fechos convexos dos contornos e_r
Θ	Conjunto composto por todos os *pixels* localizados dentro dos polígonos hu_n
I_B	Matriz da componente azul da imagem
I_G	Matriz da componente verde da imagem
I_R	Matriz da componente vermelha da imagem
l_1	Limiar inicial da abordagem proposta
l_2	Limiar restritivo da abordagem proposta
$o_m^{(t)}$	Variável de seleção na atualização do modelo
p_{thr}	Distribuição constante da probabilidade a priori do primeiro plano
c_{thr}	Limiar correspondente à p_{thr} na distância Mahalanobis
S_f	Imagem segmentada após a etapa de filtragem
S_v	Imagem segmentada após a etapa de classificação
X_t	Variável aleatória de estado da classificação

Sumário

Introdução

Um estudo realizado pelo Instituto Brasileiro de Pesquisas e Estatísticas (IBGE) mostrou que o número de veículos no Brasil aumentou na última década (IBGE, 2009). A relação de habitantes por veículos era de aproximadamente 10 no ano de 2000. Esse número reduziu para aproximadamente $6,5$ na pesquisa realizada no ano de 2009. Os setores de saúde, segurança e infraestrutura são bastante afetados pelo aumento de veículos. Esses setores frequentemente geram estatísticas evidenciando a relação de seus problemas com o aumento de veículos e a imprudência dos motoristas. Por meio dessas análises, o Estado planeja os investimentos e estima os custos que esses setores geram aos cofres públicos.

Os dados fornecidos pela Secretaria de Segurança Pública do estado de São Paulo mostram que os furtos de veículos aumentaram nesse estado. Foram registrados 116.784 casos nessa categoria de furtos em 2013. No ano seguinte, o número de casos aumentou para 122.593. Considerando os impactos causados na saúde, o governo brasileiro divulgou em 2010 que 7.160 pessoas faleceram em acidentes de trânsito somente no estado de São Paulo (Portal da Saúde, 2011). Em comparação aos outros países, o Brasil encontra-se na 5ª posição no *ranking* de países com mais mortes por acidentes de trânsito em 2015 (Portal da Saúde, 2015).

Esses dados estatísticos evidenciam a necessidade de investimento em infraestrutura nos setores de segurança e de transportes. As câmeras de vigilância, uma das tecnologias utilizadas nesses setores, têm se tornado comum nos centros urbanos. De acordo com a ABESE (Associação Brasileira das Empresas de Sistemas Eletrônicos de Segurança), em 2012, São Paulo contava com cerca de três mil câmeras instaladas pelo poder público (ABESE, 2012). Essa associação prevê o aumento desse número nos anos seguintes.

A análise das imagens dessas câmeras é feita manualmente por técnicos em Centros de Operações de cada localidade. Desse modo, é necessário não só investimentos em equipamentos, mas também em recursos humanos para operar esses dispositivos. Para reduzir a dependência dos operadores, os Sistemas Inteligentes de Transporte (ITS-*Intelligent Transportation System*) automatizam algumas funções antes designadas aos operadores, como a detecção dos veículos e das placas. Um dos processos presentes nesses sistemas é a segmentação dos veículos.

Caracterização do problema

A segmentação dos veículos tem por objetivo detectar e extrair os automóveis presentes na cena. De forma geral, a segmentação é definida como um problema de

classificação (GONZALEZ; WOODS, 2006). Os dados a serem classificados, neste caso os *pixels*, são agrupados, por suas similaridades, em classes. Um modelo simplificado do problema de segmentação de veículos é a separação dos dados em dois grupos distintos: o plano de fundo e o primeiro plano (ZIVKOVIC, 2004). Esse modelo segmenta cenas dinâmicas, ou seja, ele não detecta somente os veículos, mas todos os objetos na cena e efeitos causados por eles, como reflexos e sombras.

O resultado de uma segmentação por meio de uma modelagem de duas classes pode ser observado na Figura 1. Nessa figura, a imagem original é segmentada em duas regiões: o plano de fundo, que é representado pela região escura, e o primeiro plano, representado pela região clara. Esse modelo é sensível aos efeitos que os objetos geram na cena. Erros de classificação comuns são os ruídos que são classificados como objetos e mudanças de iluminação, como sombras e reflexos.

(a) (b)

Figura 1 – Classificação por meio da modelagem simplificada. (a) Imagem original da cena. (b) Imagem segmentada em duas classes.

Modelagens robustas utilizam mais de dois agrupamentos para minimizar esses erros de classificação (PRATI et al., 2003). Esses modelos detectam estimações inconclusivas que são agrupadas nas classes adicionadas. Um modelo de três agrupamentos, por exemplo, pode ser utilizado para identificar mudanças de iluminação (HORPRASERT; HARWOOD; DAVIS, 1999). Uma das classes é utilizada para representar as regiões de sombra e realce. Dessa forma, os sistemas que utilizam essa modelagem oferecem uma robustez maior.

Embora o uso de mais classes reduza os erros de classificação, esses erros ainda são frequentes nas regiões de transição das classes dos modelos já mencionados. Como pode ser observado na Figura 1, essas modelagens apresentam dificuldades para classificar corretamente objetos escuros como veículos e para-brisas Essa dificuldade se deve à similaridade de cor com a via de tráfego. A classificação correta desses padrões é essencial para que o processo de segmentação apresente resultados que possam ser utilizados por processos cognitivos em ITSs mais complexos.

Para compreender a deficiência da abordagem probabilística, tomemos a Figura 2.

Essa figura mostra uma função gaussiana descrevendo um padrão de plano de fundo. A função de densidade de probabilidade (PDF-*Probability Density Function*) do primeiro plano é descrita por uma função constante. Esse valor constante, denominado limiar, separa os dados da amostra em dois grupos distintos. Nessa figura podemos observar duas regiões preenchidas: a primeira, identificada pela cor cinza, mostra o intervalo de valores onde ocorre erros de classificação de primeiro plano. Erros comuns nesse intervalo são a segmentação de veículos escuros e para-brisas A segunda região, identificada pela cor preta, é o intervalo responsável por erros em mudanças de iluminação e sombras.

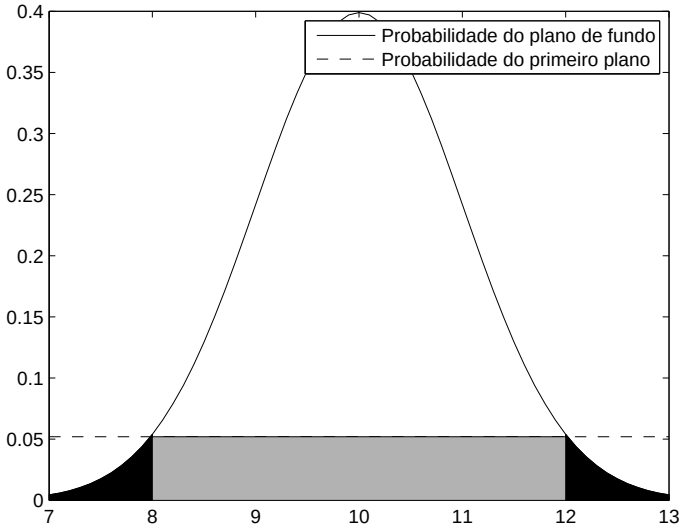


Figura 2 – Classificação por estimação de plano de fundo.

Nos ITS em geral, o limiar que separa as classes é um valor fixo ao longo de toda execução do sistema (ZIVKOVIC, 2004; KEMOUCHE; AOUF, 2009; LUO; ZHU, 2010). Esse valor é descrito por uma probabilidade mínima para aceitação da hipótese do *pixel* ser parte do plano de fundo. A abordagem de limiar constante para a classificação do *pixel* é ineficaz quando as variações são rápidas e frequentes. Métodos de limiarização adaptativa já se provaram eficazes em muitos problemas de segmentação (SEZGIN; SANKUR, 2004; CHANG et al., 2006). Esses métodos analisam a distribuição dos *pixels* e buscam um valor ótimo de acordo com a abordagem do método para maximizar a divisão dos dados.

Trabalhos relacionados

Diversos métodos para segmentar cenas dinâmicas já foram propostos na literatura. Dentre eles, uma abordagem popular é a subtração de plano de fundo (ZIVKOVIC, 2004). Essa abordagem classifica a imagem em duas regiões: plano de fundo e primeiro plano, pela estimação do fundo da imagem. Para estimar o plano de fundo, métodos de aprendizado estatístico são largamente utilizados na literatura. Esses métodos são capazes

de estimar as classes do problema de forma não-supervisionada com eficácia, e permitem a atualização constante dos padrões, dessa forma, o modelo se adapta às mudanças no ambiente rapidamente.

Stauffer e Grimson (1999) propuseram um modelo adaptativo de mistura de gaussianas (GMM-*Gaussian Mixture Model*) para estimar o plano de fundo de vídeos e extrair os objetos da cena. Este é um dos métodos mais referenciados na literatura devido aos resultados que apresenta para diversos problemas de segmentação. Considerando o problema da segmentação de veículos, esse método é limitado, pois não classifica com eficácia os padrões da transição do plano de fundo e do primeiro plano. Na literatura, o método de Stauffer e Grimson (1999) é bastante utilizado na elaboração de métodos mais sofisticados e complexos.

Em contrapartida às modelagens paramétricas que seguem um comportamento específico, como a distribuição gaussiana (SOONG, 1986), as abordagens não-paramétricas utilizam distribuições livres que não restringem o comportamento do sistema (PRATI et al., 2003). O método de Horprasert, Harwood e Davis (1999) é caracterizado como abordagem não-paramétrica para estimação de plano de fundo que se popularizou na literatura. Esse método estima as distorções de cromaticidade e brilho pelo histórico dos *pixels*. Utilizando esses valores, esse método classifica os dados em quatro agrupamentos: plano de fundo, primeiro plano, sombras e realces. As regiões de sombras e de realces, que são áreas onde ocorreram mudanças de iluminação, são descritas como regiões de transição entre o plano de fundo e o primeiro plano. Contudo, o método de Horprasert, Harwood e Davis (1999) não classifica o primeiro plano com a mesma eficácia dos métodos paramétricos.

Withagen, Schutte e Groen (2002) utilizaram o método de Stauffer e Grimson (1999) como modelo de plano de fundo de um ITS que detecta e rastreia os veículos em movimento. Utilizando o modelo de cor RGB, os autores eliminaram a componente de cor B e descrevem que o uso das componentes R e B é suficiente para realizar a segmentação com acurácia. A redução na dimensionalidade resultou em uma melhoria na performance do sistema. Para estimar os parâmetros do modelo, a abordagem de Estimação de Máxima Verossimilhança (MLE-Maximum Likelihood Estimation) é utilizada por meio do método *Expectation-Maximization* (EM) (SOONG, 1986).

Um estudo comparando diversas abordagens paramétricas e não-paramétricas foi realizado por Prati et al. (2003). Os autores avaliaram os resultados dos métodos de extração de plano de fundo sob perspectivas diferentes. Foi constatado que os métodos paramétricos apresentam a tendência de classificar o primeiro plano, enquanto os não-paramétricos, o plano de fundo. A dificuldade em classificar corretamente as sutis diferenças dos dois planos compromete a segmentação dos veículos e dificultam o processo de rastreamento. Nos métodos não-paramétricos, esses erros são minimizados, contudo, os métodos apresentam

deficiência na classificação de veículos com cores similiares à da via de tráfego.

Zivkovic (2004) propôs um GMM adaptativo com o método EM para a segmentação de cenas dinâmicas. O autor considerou as análises de Prati et al. (2003) sobre as vantagens e problemas de cada abordagem e definiu um método com ambas as abordagens. A classificação é composta por dois estágios, no primeiro, os *pixels* são segmentados por um método paramétrico. No segundo, as regiões de primeiro plano são reclassificadas por um método não-paramétrico. O método paramétrico escolhido foi o GMM de Stauffer e Grimson (1999), e o não-paramétrico foi o método de Horprasert, Harwood e Davis (1999). O método EM atualiza os parâmetros destes dois métodos, onde o primeiro foca-se na classificação do primeiro plano e o segundo na detecção das regiões de sombras e realces provocadas por mudanças de iluminação. Em virtude dos resultados que esse método apresenta para ambientes externos, bibliotecas de processamento de imagens, como a OpenCV, incluíram esse método em seus repositórios (BRADSKI, 2000).

Lei et al. (2008) propuseram um ITS para contagem de veículos em uma cena de tráfego. Os autores utilizaram uma modelagem de plano de fundo baseada no canal de luminância. Esse canal é calculado por meio das componentes de cor RGB e define a intensidade da luz refletida por cada *pixel*. Os experimentos mostraram que sombras e realces, em geral, não afetam significamente o canal de saturação do modelo de cor. A aplicação de um filtro *band-stop* na saturação do *pixel* permite a identificação de ruídos e remoção das sombras. Esse processo valida a classificação realizada por meio da luminância de maneira similar ao método de Zivkovic (2004). Embora o ITS proposto cumpra seu objetivo, o processo de segmentação não é eficaz. A fragmentação da região dos veículos ocorre com frequência. Em consequência disso, a aplicação desse método de segmentação é limitada. Em sistemas mais complexos que possuem etapas de reconhecimento, é necessário que os resultados da extração sejam mais eficazes do que os obtidos por Lei et al. (2008).

Um ITS para rastreamento de veículos foi proposto por Kemouche e Aouf (2009). Esse sistema utiliza o método GMM de Stauffer e Grimson (1999) para estimar o plano de fundo e segmentar as imagens. O rastreamento é realizado por meio de fluxo óptico (TAO et al., 2012). Na abordagem proposta por esses autores, a etapa de rastreamento fornece uma realimentação para o processo de estimação do plano de fundo. Desse modo, os parâmetros de atualização do método GMM são ajustados nas regiões onde foi detectado movimento. Entretanto, Kemouche e Aouf (2009) não apresentaram testes de desempenho para avaliar a viabilidade do rastreamento por fluxo óptico. O custo computacional desta técnica pode limitar o sistema a aplicações *off-line*.

Um sistema para análise de informações de tráfego de veículos foi proposto por Shao et al. (2010). A modelagem de plano de fundo desse ITS é realizada pelo modelo de Stauffer e Grimson (1999) e o rastreamento é obtido com o filtro de Kalman (KALMAN, 1960). A abordagem de rastreamento de múltiplas hipóteses (MHT-*Multiple Hypothesis*

Tracking) (BLACKMAN, 2004) detecta infrações de trânsito como travessia fora da faixa de pedestres, avanço de sinal vermelho e veículos transitando em sentido contrário ao da via. Shao et al. (2010) realizaram testes de performance e mostraram que essa abordagem de rastreamento é eficaz e o custo computacional necessário é baixo.

Similar ao método de Kemouche e Aouf (2009) que utiliza uma abordagem de realimentação, Luo e Zhu (2010) melhoraram a estimação do método GMM por meio de uma análise espacial nas estimações dos modelos estatísticos. Os autores utilizaram histogramas locais para cada região segmentada e assim, detectaram as mudanças mais significativas. De forma similar à feita por Kemouche e Aouf (2009), os parâmetros de atualização do plano de fundo são ajustados para aprimorar o processo de classificação. A ausência de mais experimentos impede uma análise mais profunda desse sistema.

Li e Zhong (2009) proporam um ITS para detecção de veículos e estruturas da via de tráfego. Os autores utilizaram o GMM de Stauffer e Grimson (1999) para gerar as estimações de plano de fundo e extrair características invariantes à cor por meio das informações do modelo RGB. Essas características são utilizadas para identificar as distorções causadas por ruídos e assim, eliminar as sombras dos veículos. Embora esse trabalho não realize rastreamento, há uma estimação global de movimento. Essa estimação, denominada *Edge Change Ratio* (ECR) (ZABIH; MILLER; MAI, 1999), utiliza os contornos encontrados em imagens consecutivas. Para detectar estruturas, limites da via e sinalizações horizontais, o sistema utiliza a transformada de Hough (HOUGH, 1962). Esse método de detecção de linhas é utilizado para calibrar a câmera e estimar alguns limiares do sistema, como a área mínima por contorno.

Apesar dos trabalhos existentes na literatura utilizarem com mais frequência distribuições gaussianas, outras funções de distribuição de probabilidade tem sido testadas para o problema da segmentação de veículos. Faro, Giordano e Spampinato (2011) utilizaram um modelo de mistura de Poisson (PMM-*Poisson Mixture Model*) para estimar o plano de fundo e segmentar os veículos da via de tráfego. Os autores aumentaram a acurácia desse ITS por meio de sensores de iluminação que determinam as mudanças de luz com mais precisão. Esse sistema, no entanto, requer uma etapa de rastreamento. Em virtude do modelo de Poisson (LIPSCHUTZ, 1993) exigir a frequência esperada dos dados, a velocidade dos veículos é necessária.

Um ITS para detecção e rastreamento de veículos por meio de modelos tridimensionais do ambiente foi proposto por (ALVAREZ et al., 2012). Neste trabalho um GMM é utilizado para gerar segmentação dos veículos. Esse sistema fundamenta-se na suposição de que mudanças de iluminação afetam as cores de objetos, mas não sua superfície. Dessa forma, a detecção de sombra é feita analisando a textura das regiões segmentadas. Os autores analisam os pontos estáticos entre imagens consecutivas para determinar as instabilidades da posição da câmera e corrigir a segmentação. A detecção de oclusão é

realizada por meio de detecção de linhas. O modelo gera uma aproximação tridimensional das imagens e separa objetos realizando aproximações com o paralelepípedo que envolve os veículos.

Como pode ser observado, os últimos trabalhos na literatura focam-se em modelos probabilísticos auxiliados por estimação de outras *features* de cada problema abordado. Em geral, as características utilizadas são baseadas em texturas, histograma e rastreamento dos objetos. Esses modelos com realimentação e auxílio, em geral, podem focar-se na correção de deficiências presentes nos métodos mais simples e assim, propor extensões que melhorem a classificação de um problema específico.

Objetivos

Este trabalho tem como objetivo propor um sistema com o modelo de segmentação de cenas dinâmicas de Zivkovic (2004) estendido especificamente para o problema de segmentação de veículos que reduza os erros de classificação do plano de fundo. A escolha do modelo de Zivkovic (2004) se deve ao fato deste método ser consolidado e amplamente utilizado por outros sistemas de segmentação e estar inserido na OpenCV (BRADSKI, 2000). O sistema proposto irá auxiliar o método de Zivkovic (2004) analisando seus resultados e fornecendo informações para a seleção de um limiar mais adequado à amostra.

A escassez de bases de vídeos para experimentos é um problema bastante comum no desenvolvimento de ITS. Após a conclusão dos testes, disponibilizaremos as bases criadas para reduzir esse problema da literatura, assim como permitir a reprodução dos resultados em outros trabalhos e a comparação de novos sistemas. As bases criadas neste trabalho tem como objetivo a análise dos métodos sob diferentes contextos do mesmo ambiente. Com esse experimento esperamos analisar se métodos apresentam comportamentos diferentes considerando fatores como horário do dia.

Em geral, a classificação por meio de abordagens probabilísticas utiliza limiares constantes. O escopo deste trabalho inclui testes de limiarização adaptativa em modelos probabilístico, neste caso, nos métodos de Stauffer e Grimson (1999) e Horprasert, Harwood e Davis (1999) que são utilizados por Zivkovic (2004). A análise da limiarização adaptiva nesses métodos permite também verificar a contribuição que cada componente induz no resultado da classificação.

Objetivos específicos

É necessário detalhar as metas pretendidas. O sistema proposto neste trabalho tem por finalidade auxiliar e aprimorar a classificação do primeiro plano no método de Zivkovic (2004). Entretanto, essa melhoria não deve afetar negativamente a classificação

do plano de fundo. A avaliação do sistema sob essa perspectiva deve ser feita utilizando métricas que considerem os acertos relativos ao plano analisado.

Embora não tenha sido definido um limite máximo de tempo aceitável para as iterações do sistema, as etapas encarregadas de realizar a detecção e correção dos erros de classificação não podem interferir na acurácia do sistema. Em virtude do aumento na complexidade causado por métodos de rastreamento, optou-se por remover essa etapa do sistema. A simplificação do modelo proposto diminui os recursos necessários, direcionando-os para a melhoria da classificação.

A classificação da região de um veículo pode apresentar variações nas métricas de análise ao longo do tempo. Isso ocorre devido às mudanças de iluminação, os parâmetros dos modelos estatísticos da região e o momento em que a amostra foi obtida. A análise da estabilidade de classificações sucessivas indica se o modelo de estimação e o classificador se adaptam às mudanças do ambiente com facilidade. Para que o sistema proposto cumpra seu objetivo como uma melhoria do método de Zivkovic (2004), ele deve se mostrar estável em segmentações sucessivas.

A realimentação do modelo é dada pela análise espacial de uma classificação para estimar limiares que maximizem a classificação de regiões previamente segmentadas. No entanto, existem outros métodos que utilizam abordagem de limiarização adaptativa que podem ser analisados. Esses métodos analisam os dados sob várias perspectivas como entropia, *clustering*, vizinhança, geometria, histograma, entre outras. Esses métodos devem ser explorados para determinar se a abordagem proposta é válida.

Em suma, os objetivos deste trabalho podem ser sintetizados como:

- Aprimorar a classificação pela redução dos erros do plano de fundo;
- Aumentar a estabilidade da segmentação em uma sequência de imagens;
- Analisar outros estimadores de limiar para o problema.

Contribuições

Os métodos encontrados na literatura fornecem resultados satisfatórios para problemas genéricos. Porém, em casos específicos como o da segmentação de veículos, eles apresentam deficiências e frequentes falhas na classificação do primeiro plano. O método proposto analisa as classificações e corrige alguns dessas falhas utilizando a geometria dos veículos. Embora o escopo dos experimentos esteja limitado ao problema da segmentação de veículos, esse sistema pode ser utilizado em outros problemas semelhantes.

Os ITS atuais tem como cerne e foco a estimação da localização, rastreamento e reconhecimento dos veículos. Os processos de localização e rastreamento, em geral, são

mais flexíveis com o processo de segmentação. Para sistemas de reconhecimento, entretanto, a precisão da etapa de segmentação é fundamental e afeta o reconhecimento dos padrões dos veículos. O sistema proposto consiste de uma solução *on-line* com custo computacional baixo para ser utilizado em ITS de reconhecimento. Contudo, é interessante destacar que este trabalho empenha-se apenas na segmentação e extração dos veículos.

Muitas das obras analisadas na literatura realizam experimentos sem analisar a diversidade de situações que podem ocorrer no problema da segmentação de veículos. Dentre essas diferentes condições, destacamos a correlação entre os problemas ocasionados pelas variações de luz e os horários do dia. Embora um método apresente bons resultados em um horário determinado, diferentes problemas de iluminação, como reflexos e sombras, podem interferir na precisão dos testes. Como consequência disso, os experimentos acabam por mostrar um comportamento isolado. Nos testes deste trabalho, submetemos os métodos de modelagem do plano de fundo e os de limiarização a diferentes bases de vídeos. Isso permite que resultados relatados em outros trabalhos sejam validados.

As abordagens de realimentação dos métodos recentes utilizam a estimação de movimento. Essa é uma abordagem pertinente, entretanto, pode limitar o desempenho do sistema e restringir sua aplicação em tempo-real. O rastreamento e a detecção de movimento são essenciais quando a performance de quadros por segundo é baixa. No entanto, a variação de localização é baixa e previsível em sistemas com iterações rápidas. O sistema proposto nesse trabalho se favorece dessa característica para gerar realimentações sem o uso de estimações de localização e movimento.

Organização do Trabalho

Neste capítulo introdutório, o problema da segmentação de veículos é contextualizado. Algumas motivações para o desenvolvimento de sistemas inteligentes de transporte são expostas, assim como deficiências que os atuais métodos possuem e que precisam ser melhoradas. São relatados alguns sistemas propostos na literatura que atuam no problema da segmentação de veículos e desempenham outras funcionalidades. As metas do sistema proposto são especificadas, assim como as deficiências que estão sendo atacadas.

O método proposto é detalhado no Capítulo 2. Neste capítulo são especificadas todas as etapas do sistema proposto. Os processos de classificação, filtragem e estimação de limiar são os mais detalhados por serem o diferencial do sistema proposto. Nesse capítulo são descritos o modelo de classificação utilizado e como a estimação de limiar auxilia essa classificação. São expostos também os métodos de filtragem e que erros específicos eles buscam corrigir após o processo de segmentação.

As métricas de avaliação e as bases utilizadas nos experimentos são detalhadas no Capítulo 2. Essas métricas permitem a comparação do sistema proposto a outros métodos

existentes na literatura. Cada uma das bases utilizadas é apresentada e as diferenças entre elas são expostas. O entendimento dessas sutis variações dos padrões nas regiões de interesse do problema permitem análises e discussões mais profundas.

O Capítulo 3 descreve os experimentos e resultados dos testes para discuti-los e expor as vantagens e deficiências observadas. Entre esses experimentos destacamos: a análise de estabilidade, o estudo das características utilizadas pelo processo de segmentação e classificação por limiarização adaptativa. O método proposto é analisado nesses experimentos para verificar suas qualidades e suas deficiências. Por fim, são expostos alguns estudos que podem ser realizados em trabalhos futuros.

Após os resultados, as conclusões deste trabalho são detalhadas em um capítulo não-enumerado que indexa e justifica como as contribuições alcançadas cumprem os objetivos deste trabalho. Esse capítulo final descreve as vantagens oferecidas pelo sistema proposto considerando os métodos avaliados, as limitações encontradas e trabalhos que podem dar continuidade a pesquisa realizada.

1 Referencial Teórico

1.1 Aprendizado de Máquina

Aprendizado de máquina ou aprendizagem automática é uma subárea da inteligência artificial (RUSSELL; NORVIG, 2009). Essa subárea estuda técnicas que permitem ao computador aprender conceitos generalizando exemplos e, por meio dessas generalizações, inferir sobre situações não previstas. A tomada de decisões dessas técnicas de sistemas de aprendizado é feita pela construção de conhecimento por acúmulo de experiências passadas. De forma geral, muitos dos métodos de aprendizado de máquina buscam identificar características relevantes para a categorização dos dados em grupos, mediante a similaridade deles (DOWNEY, 2012).

Existem diversos paradigmas para classificar dados: simbólico, genético, estatístico, conexionista, evolucionista, entre outros (RUSSELL; NORVIG, 2009). Esses paradigmas são intrinsecamente relacionados a forma de aquisição, representação e raciocínio do conhecimento. Na última década, muitos avanços foram feitos no paradigma estatístico. Nesse paradigma, os agrupamentos são descritos por modelos estatísticos que estimam a probabilidade de ocorrência de determinadas características em uma classe (PEDRINI; SCHWARTZ, 2008). A subárea que estuda esse paradigma específico é chamada de aprendizado estatístico.

1.1.1 Aprendizado estatístico

Os métodos de aprendizado estatístico estimam a probabilidade de ocorrência de características, denominadas evidências (RUSSELL; NORVIG, 2009). Cada agrupamento, denominado estado, possui suas funções de distribuição de probabilidade que generalizam o comportamento das evidências analisadas (RYAN, 2009). Esses métodos são bastante abrangentes e possuem formas de representação e raciocínio naturalmente compreensíveis. Uma modelagem probabilística simples e bastante utilizada é a independência de evidências para características diferentes. Essa modelagem reduz a complexidade do modelo, pois, as dimensões da tabela de probabilidade crescem exponencialmente (RUSSELL; NORVIG, 2009).

A probabilidade a priori dos estados é descrita como a frequência de uma classe sem observação das evidências. Essa informação é importante na predição da variável de estado pois mostra o comportamento de uma amostra e a constância de uma classe. Contudo, essa estimativa não é suficiente para que o classificador categorize precisamente os dados de uma base. Em sistemas reais, diversas informações sobre o mundo podem ser extraídas, elas

são chamadas de variáveis de evidências. A correlação dessas informações extraídas com as classes de um problema é descrita como probabilidade a posteriori. Análises envolvendo a probabilidade a posteriori permitem a concepção de modelos mais exatos.

1.2 Processos estocásticos

O aprendizado estatístico generaliza o comportamento dos dados em modelos probabilísticos. Em geral, esses modelos descrevem processos estocásticos que são regidos por leis que não se alteram ao longo do tempo (RUSSELL; NORVIG, 2009), contudo, possuem certo grau de imprevisibilidade. Essas leis, representadas pelas funções de probabilidade, podem ser reduzidas a parâmetros de funções pré-determinadas ou não.

Existem diversas técnicas para modelar processos estocásticos na literatura. Entre elas, podemos citar Modelos Markovianos (MARKOV, 1971), processos Gaussianos (RASMUSSEN; WILLIAMS, 2005), processos Brownianos (LI et al., 2010) e processos de Poisson (POISSON, 1837). Cada uma delas, reduz a amostra de dados sob uma perspectiva diferente. Dentre os processos citados, este trabalho se concentra nos modelos de Markov e nos processos Gaussianos.

1.2.1 Modelos de Markov

Os processos Markovianos buscam gerar uma correlação com o histórico de variáveis aleatórias para a previsão do próximo estado dessa variável (DOWNEY, 2012). Formalmente, um modelo de Markov é descrito pela probabilidade $p(X_t|X_{t-1}, X_{t-2}, ..., X_0)$, onde o estado de X no tempo t é estimado em função do histórico $(X_{t-1}, X_{t-2}, ..., X_0)$.

Facilmente podemos observar que a dimensão do histórico se torna um problema na definição do modelo de Markov. Á medida que a amostra do histórico aumenta, torna-se custoso computacionalmente gerar predições para as variáveis aleatórias. Desse modo, uma abordagem que tem se tornado comum é definir uma janela móvel que limita o número de evidências que são consideradas na estimação. Essa abordagem é descrita como, $p(X_t|X_{t-1}, ..., X_{t-T})$, de modo que T é o tamanho da janela de tempo que é utilizada na estimação da variável aleatória.

O uso do estado anterior e dos dois últimos estados são bastante utilizados em modelos Markovianos. Esses dois casos específicos são descritos como modelo de Markov de primeira ordem e de segunda ordem (RUSSELL; NORVIG, 2009), respectivamente. Embora esses dois modelos sejam simplificações do processo completo, eles são suficientemente precisos para serem utilizados em problemas reais. Além de que eles podem ser modelados facilmente como matrizes de transição de estados.

1.2.2 Estatística paramétrica

A estatística paramétrica impõe restrições ao comportamento dos dados. Nesse paradigma, presume-se que os dados sigam o Teorema Central do Limite (SOONG, 1986; LIPSCHUTZ, 1993), e sigam um comportamento determinado por uma função de probabilidade conhecida. Desse modo, as probabilidades dos eventos são facilmente estimadas por meio dos parâmetros dessa função. Uma das funções mais utilizadas na literatura é a distribuição gaussiana.

Processos gaussianos são descritos por distribuições gaussianas. Essa distribuição é uma função de densidade de probabilidade contínua e simétrica cujo conjunto de parâmetro θ é composto por dois valores. A média μ descreve o valor de convergência pelo Teorema Central do Limite, e o desvio-padrão σ o erro médio considerando toda a amostra. A distribuição gaussiana para vetores de características, que é calculada como:

$$\mathcal{N}(\vec{x}_t, \hat{\vec{\mu}}, \hat{\sigma}) = \frac{1}{\hat{\sigma}\sqrt{2\pi}} e^{-\left(\vec{x}_t - \hat{\vec{\mu}}\right)^2 / 2\hat{\sigma}^2}, \tag{1.1}$$

assume o parâmetro $\hat{\vec{\mu}}$ como o vetor de médias. O parâmetro $\hat{\sigma}$, calculado através da média, é o erro médio entre as medições $\vec{x}_t$ e o parâmetro $\hat{\vec{\mu}}$.

A distribuição gaussiana, também chamada de normal, é a mais popular dentre as paramétricas. Um exemplo dessa distribuição pode ser observada na Figura 3. Nessa figura, a distribuição normal estimada generaliza uma amostra de dados aleatórios. Embora ela seja a mais utilizada, existem outras distribuições específicas para determinados casos. Uma que tem mostrado relevância em muitos estudos é a distribuição Poisson (POISSON, 1837). Essa distribuição tem sido utilizada para estimar eventos temporais (DOWNEY, 2012).

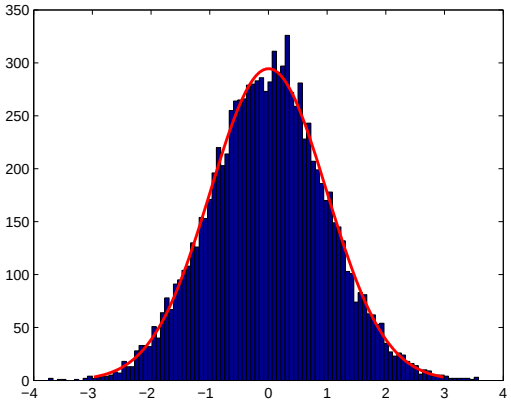


Figura 3 – Distribuição gaussiana estimada a partir de uma amostra.

1.2.3 Estatística não-paramétrica

A estatística paramétrica é bastante eficaz quando o comportamento da amostra é conhecido, contudo, nem sempre isso é possível. O uso de métodos paramétricos nos casos em que os dados não convergem para o modelo normal apresenta vários erros de estimação. O paradigma não-paramétrico é utilizado quando o comportamento não é conhecido previamente. As funções de densidade de probabilidade desse paradigma são também conhecidas como distribuições livres por não impor um padrão de comportamento na modelagem.

Vários métodos não-paramétricos são encontrados na literatura (PRATI et al., 2003). O histograma, uma distribuição de frequência dos dados da amostra, é considerado um dos mais simples modelos desse paradigma. O método de Horprasert, Harwood e Davis (1999) é uma abordagem não-paramétrica que se popularizou na área de segmentação de imagens para detectar mudanças de iluminação.

1.2.4 Inferência Bayesiana

As estimativas geradas pelos modelos estatísticos identificam a pertinência dos dados em cada classe. A decisão do estado que melhor representa as evidências é feita por um processo de inferência Bayesiana. Esse processo é fundamentado por meio do Teorema de Bayes (BEKMAN; NETO, 1980) que considera a correlação de variáveis aleatórias e suas estimações isoladas. Assim, esse teorema descreve de maneira eficaz as relações de causa e efeito de um problema. O teorema de Bayes é definido pela equação,

$$p(causa, efeito) = \frac{p(causa|efeito)}{p(causa)} = \frac{p(efeito|causa)}{p(efeito)}. \tag{1.2}$$

Apesar de sua simplicidade, na prática, o teorema Bayesiano apresenta bons resultados. Ele também possui uma complexidade baixa devido à probabilidade a priori e a posteriori serem facilmente calculadas com as informações das evidências e_i de estados passados.

A decisão Bayesiana avalia os estados x e y e confronta as hipóteses de cada um deles representar o mundo. Essa avaliação utiliza o conjunto de evidências e_i conhecidas. A relação T que julga os estados é descrita pela equação,

$$T = \frac{p(x|e_i)}{p(y|e_i)} = \frac{p(e_i|x)p(x)}{p(e_i|y)p(y)}. \tag{1.3}$$

Na segmentação de cenas dinâmicas, a decisão Bayesiana é aplicada utilizando os *pixels* $\vec{x}$ como evidência para classificação do plano. A relação R confronta as hipóteses do *pixel* pertencer ao plano de fundo ou ao primeiro plano, também chamados respectivamente de *background* (BG) e *foreground* (FG). Essa relação é escrita como:

$$R = \frac{p(BG|\vec{x})}{p(FG|\vec{x})} = \frac{p(\vec{x}|BG)p(BG)}{p(\vec{x}|FG)p(FG)}. \tag{1.4}$$

A probabilidade a posteriori $p(\vec{x}|BG)$ é calculada utilizando modelos estatísticos com base nos valores estimados do *pixel*. Há diversos paradigmas que podem ser utilizados nessa estimação. Eles dependem do contexto do problema e do comportamento dos dados da amostra. De forma simplificada, os métodos estatísticos são categorizados em paramétricos e não-paramétricos. Enquanto os métodos da primeira categoria aplicam restrições no comportamento dos dados, os da segunda permitem que os dados apresentem comportamento variado e modelado por funções que não seguem o Teorema Central do Limite.

1.3 Modelo de Mistura

O modelo de mistura é uma técnica de decomposição de distribuições não-paramétricas, que são mais complexas, em várias distribuições paramétricas mais simples. Esse método consegue generalizar padrões complexos de forma satisfatória com baixo custo computacional. A decomposição de padrões tem sido utilizada não só padrões complexos, como também na transição gradual de padrões.

A ineficácia dos métodos paramétricos para estimar padrões complexos é resolvida pela decomposição da função de densidade de probabilidade em mais de uma componente. O resultado da decomposição de uma amostra hipotética é mostrado na Figura 4. A flexibilidade desse modelo se deve a não serem impostas restrições na escolha das distribuições e a quantidade delas. Consequentemente, há um grande número de combinações para representar o modelo estatístico.

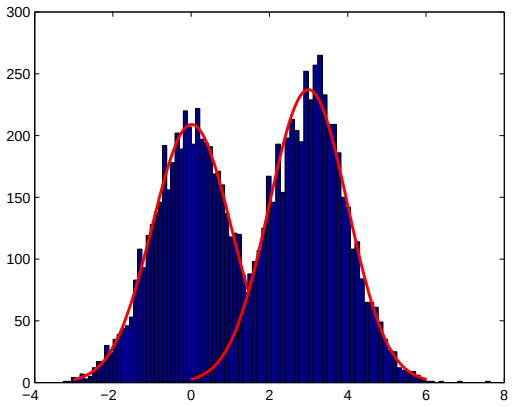


Figura 4 – Decomposição de histograma em distribuições paramétricas por modelo de mistura.

O Modelo de Mistura de Gaussianas (GMM-*Gaussian Mixture Model*) tem sido largamente utilizado em modelagens estatísticas para classificação de dados. Muitas vezes, os sistemas não dispõem de todos os dados e possuem apenas um pequeno histórico de

ocorrências. Os parâmetros desses modelos estatísticos são estimados através dessa amostra limitada. Uma abordagem para estimar valores é a máxima verossimilhança.

1.3.1 Modelo de Mistura de Gaussianas de Stauffer e Grimson

O modelo de mistura de gaussianas (GMM) é um método probabilístico que utiliza múltiplas distribuições normais para as estimações das variáveis aleatórias. Conforme esse conceito, a probabilidade de um evento $\vec{x}$ dado um modelo de uma variável aleatória A é estimada como um somatório de distribuições gaussianas com diferentes parâmetros:

$$p(\vec{x}|A) = \mathcal{N}(\vec{x}, \widehat{\vec{\mu_1}}, \hat{\sigma}_1^2) + \mathcal{N}(\vec{x}, \widehat{\vec{\mu_2}}, \hat{\sigma}_2^2) + ... + \mathcal{N}(\vec{x}, \widehat{\vec{\mu_i}}, \hat{\sigma}_i^2), \tag{1.5}$$

onde $\mathcal{N}$ é descrito pela Equação 1.1.

Uma das vantagens da modelagem por mistura é que a classificação pode ser suavizada em determinadas regiões. As componentes são ponderadas de acordo com sua influência na estimação. Seguindo essa abordagem, a estimação da probabilidade é reescrita como:

$$p(\vec{x}|A) = \sum_{m=1}^{M} \hat{\pi}_m \mathcal{N}(\vec{x}, \widehat{\vec{\mu}}_m, \hat{\sigma}_m^2), \tag{1.6}$$

onde M é o número de componentes e $\hat{\pi}_m$ é o peso de cada distribuição no modelo.

Definido o modelo estatístico, é preciso estimar os parâmetros de cada componente do modelo. Dado as características do problema de segmentação de veículos, o método *Expectation-Maximization* é um dos mais utilizados nestes casos. A atualização iterativa dos parâmetros em contexto local torna esse método ideal para a seleção desses parâmetros.

O método EM utiliza evidências que são obtidas para estimar os parâmetros das distribuições. Essas evidências formam o conjunto $\mathcal{X}_T$ e possuem informações do agrupamento A, e eventualmente de um agrupamento B. Durante a estimação dos parâmetros através do método EM, é inevitável que dados do agrupamento B sejam incorporados ao modelo de A conforme:

$$p(\vec{x}|\mathcal{X}_T, A, B) = \sum_{m=1}^{M} \hat{\pi}_m \mathcal{N}(\vec{x}, \widehat{\vec{\mu}}_m, \hat{\sigma}_m^2). \tag{1.7}$$

Isso ocorre eventualmente devido às componentes obsoletas ou de baixa relevância do modelo da variável X.

Considerando as componentes ordenadas de forma que $\hat{\pi}_m > \hat{\pi}_{m+1}$ ocorra, as últimas componentes da ordem são as de menor relevância. Essas componentes de baixo peso são ignoradas por meio da estimação aproximada,

$$p(\vec{x}|\mathcal{X}_T, X) \sim \sum_{m=1}^{K} \hat{\pi}_m \mathcal{N}(\vec{x}; \widehat{\vec{\mu}}_m, \hat{\sigma}_m^2), \tag{1.8}$$

onde K é o limiar utilizado para ignorar as distribuições menos representativas.

O limiar K é calculado dinamicamente através dos pesos das componentes. Isso é realizado por meio da função de minimização:

$$K = \underset{b}{\operatorname{argmin}} \left(\sum_{m=1}^{b} \hat{\pi}_m > (1 - c_f) \right), \tag{1.9}$$

onde c_f é a porção máxima de dados que são classificados no agrupamento B.

1.4 Estimação de Parâmetros de Modelos Probabilísticos

Os modelos probabilísticos por si descrevem apenas o comportamento de uma população. Esses modelos possuem parâmetros específicos que são utilizados em suas funções de densidade de probabilidade. Esses parâmetros devem ser estimados utilizando uma amostra dos dados. Ryan (2009) descreve três métodos para estimar os parâmetros de um modelo: método da máxima verossimilhança, método dos momentos e método dos mínimos quadrados.

1.4.1 Método de Máxima Verossimilhança

A Estimação de Máxima Verossimilhança (MLE-*Maximum-Likelihood Estimation*) é um método para estimação dos parâmetros de modelos probabilísticos utilizando uma amostra da população (RYAN, 2009). Esse método realiza uma estimativa de similaridade para cada dado da amostra na clusterização. Quando essa amostra atinge uma dimensão grande, essa abordagem se torna inviável devido ao custo necessário para realizar esse processo (RUSSELL; NORVIG, 2009). De acordo com (CASELLA; BERGER, 1990), a derivação dos parâmetros estatísticos por meio da MLE é uma das técnicas mais utilizadas.

Dada uma amostra $(x_1, ..., x_n)$ de evidências da variável aleatória X. A função,

$$\mathcal{L}(x_1, ..., x_n; \theta) = f(x_1; \theta) f(x_2; \theta) ... f(x_n; \theta), \tag{1.10}$$

descreve a verossimilhança com o parâmetro θ de uma PDF ou uma *Probability Mass Function* (PMF) f, uma função de densidade discreta. Uma prática comum na MLE é o uso do logaritmo da função de verossimilhança. O método MLE utiliza técnicas de cálculo diferencial (LEITHOLD, 1994) para maximizar a verossimilhança. Sendo assim o parâmetro θ é estimado por meio da equação,

$$\frac{\delta}{\delta\theta} ln\, \mathcal{L}(x_1, ..., x_n; \theta). \tag{1.11}$$

1.4.2 Método dos Momentos

O método dos momentos é outra técnica de estimação dos parâmetros de uma PDF e PMF. Esse método iguala momentos amostrais ao da população e gera estimações

independentes de função de probabilidade (RYAN, 2009). Esse método é uma opção viável quando a MLE é intratável devido a dimensão da amostra. Embora esse método específico tenha sido bastante utilizado em áreas da engenharia, o método MLE e dos Mínimos Quadrados são mais utilizados em problemas de classificação.

1.4.3 Método dos Mínimos Quadrados

O método dos Mínimos Quadrados é bastante utilizado em modelos de regressão (RYAN, 2009). Nesse método, os parâmetros de um modelo probabilístico são estimados minimizando a soma dos desvios de cada observação na amostra. Em problemas de aprendizado de máquina, esse método é bastante utilizado no rastreamento de objetos por fluxo óptico. O método dos Mínimos Quadrados busca a minimização do erro quadrático entre as observações e um parâmetro específico. Seja μ o parâmetro média que descreve uma amostra $(x_1, ..., x_n)$, o método dos Mínimos Quadrados tem como função objetivo,

$$\min \sum_{i=1}^{I} (x_i - \mu)^2. \tag{1.12}$$

1.4.4 Expectation-Maximization

O método *Expectation-Maximization* (EM) compensa a deficiência do MLE em processos temporais utilizando uma abordagem iterativa. Esse método estima os parâmetros em um contexto local por meio de uma janela. Conforme novos dados são obtidos, essa janela se desloca para que novas medidas sejam utilizadas no cálculo dos parâmetros e descartando dados obsoletos.

O maior proveito que pode ser obtido pelo método EM não é a estimação dos parâmetros. Por estimar padrões em um contexto local, o modelo probabilístico evolui e se adapta as novas condições com facilidade. Em um problema como a segmentação de cenas dinâmicas, essa característica é fundamental para que o modelo se ajuste às várias condições de ambiente possíveis.

O método EM é composto basicamente de duas etapas que são executadas conforme os eventos do conjunto $\mathcal{X}_T$ ocorrem. A primeira, chamada *Expectation*, seleciona a distribuição que representa determinada evidência. Enquanto a segunda, chamada *Maximization*, atualiza os parâmetros da distribuição e ajusta os pesos das componentes. Duas variações desse método aplicado no GMM são descritas nesta seção: a de Stauffer e Grimson (1999) e a de Zivkovic (2004).

As duas implementações do EM são semelhantes, diferindo na atualização dos pesos na etapa *Maximization*. As distribuições são escolhidas através da probabilidade de cada componente em relação ao evento ocorrido. Isso é realizado por meio da função $\mathcal{N}$ definida pela Equação 1.1. Devido a complexidade do cálculo, Zivkovic (2004) optou por

utilizar uma segunda abordagem para comparar as distribuições. Similar ao *Score* padrão, mais conhecido como *Standard score*, a distância Mahalanobis é utilizada nessa etapa para melhorar o desempenho do sistema.

Devido a dinamicidade do ambiente, alguns padrões ficam mais frequentes enquanto outros se tornam obsoletos. Para inserir e eliminar esses padrões de comportamento, componentes são criadas e removidas na etapa *Expectation*. A criação de componentes ocorre quando nenhuma distribuição é selecionada para representar uma ocorrência. A remoção ocorre quando o peso de uma componente chega a valores demasiadamente baixos. Os parâmetros iniciais da nova componente $M+1$ são definidos como:

$$\hat{\pi}_{M+1} = \alpha, \tag{1.13}$$

$$\hat{\vec{\mu}}_{M+1} = \vec{x}^{(t)}, \text{ e} \tag{1.14}$$

$$\hat{\sigma}_{M+1} = \sigma_0, \tag{1.15}$$

onde α é a taxa de atualização e σ_0 é o desvio-padrão inicial.

Na etapa *Maximization*, os parâmetros, $\hat{\vec{\mu}}$, $\hat{\sigma}$ e $\hat{\pi}$ das componentes m devem ser atualizados através das equações:

$$\hat{\mu}_m \leftarrow \hat{\mu}_m + o_m^{(t)}(\alpha/\hat{\pi}_m)\vec{\delta}_m, \text{ e} \tag{1.16}$$

$$\hat{\sigma}_m^2 \leftarrow \hat{\sigma}_m^2 + o_m^{(t)}(\alpha/\hat{\pi}_m)(\vec{\delta}_m^T\vec{\delta}_m - \hat{\sigma}_m^2). \tag{1.17}$$

A diferença entre o vetor de médias $\vec{\mu}_m$ e a nova informação $\vec{x}$ é dada por δ_m. A variável $o_m^{(t)}$ é utilizada para selecionar qual distribuição é atualizada. Ela assume o valor 1 na distribuição mais representativa, e atualiza essa componente. Nas outras distribuições, essa variável assume o valor 0 impedindo a atualização dos parâmetros.

O algoritmo EM utilizado por Stauffer e Grimson possui a seguinte atualização para o peso $\hat{\pi}$:

$$\hat{\pi}_m \leftarrow \hat{\pi}_m + \alpha(o_m^{(t)} - \hat{\pi}_m). \tag{1.18}$$

O peso da componente mais representativa é incrementado, ao passo que ele é reduzido nas outras distribuições. Após o ajuste de peso, os valores são normalizados seguindo a regra:

$$\sum_{m=1}^{M} \hat{\pi}_m = 1. \tag{1.19}$$

Zivkovic (2004) estimou a os pesos de cada distribuição considerando a proporção de evidências de cada classe. Ele propôs que a atualização dos pesos $\hat{\pi}_m$ seja realizada pela equação:

$$\hat{\pi}_m \leftarrow \hat{\pi}_m + \alpha(o_m^{(t)} - \hat{\pi}_m) - \alpha C_T, \tag{1.20}$$

onde $C_T = c/T$ representa a evidência a priori c pela quantidade de exemplos do conjunto $\mathcal{X}_T$. Como resultado, a proporção de evidências da classe desacelera o aumento de relevância, ao passo que permite que as distribuições se tornem obsoletas mais rapidamente.

1.5 Distorções em modelos estatísticos

Problemas de segmentação, em abordagens probabilísticas, são modelados utilizando apenas duas variáveis aleatórias representando os agrupamentos. Esses agrupamentos são o plano de fundo e o primeiro plano, representados pelos estados BG e FG respectivamente. Os métodos paramétricos são utilizados para estimar uma dessas variáveis aleatórias. Na maioria dos casos, é escolhido o estado BG devido à tendência do plano de fundo se manter constante.

Os erros de segmentação nessa modelagem ocorrem na intersecção dos agrupamentos. Uma forma de melhorar a segmentação dessas regiões é a inclusão de um terceiro agrupamento. Esse agrupamento é utilizando para estimar os padrões na transição entre os dois agrupamentos principais. No problema da segmentação de veículos, esse novo estado se chama *shadow* ou SH. Essa variável aleatória é utilizada para estimar as mudanças de iluminação que foram classificadas no primeiro plano: as sombras e os realces. Entre os métodos de detecção de sombra, o método de Horprasert (HORPRASERT; HARWOOD; DAVIS, 1999) é apontado como um dos mais eficazes nos estudos realizados por Prati et al. (2003).

Na modelagem lambertiana descrita Prati et al. (2003), a cor é percebida como o produto da iluminação e a reflectância da superfície do material. Com base nesse conceito, Horprasert descreveu um modelo composto por distorção de brilho e distorção de cor. Seja $\vec{x}$ os dados do *pixel* e $\hat{\vec{\mu}}$ o vetor de médias, a diferença desses vetores é decomposta nas distorções do modelo.

Mesmo em cenas estáticas, a sensibilidade dos sensores resulta em mudanças nos valores dos *pixels*. Por esse motivo, o vetor das variações $\vec{\sigma}$ devem ser consideradas na estimação das distorções. A distorção de cor, cromaticidade ou crominância, é calculada para o modelo de cor RGB através da equação,

$$CD = \sqrt{\frac{I_R - \beta\mu_R}{\sigma_R}^2 + \frac{I_G - \beta\mu_G}{\sigma_G}^2 + \frac{I_B - \beta\mu_B}{\sigma_B}^2}, \tag{1.21}$$

onde β é o valor estimado da distorção de brilho, $\vec{x} = \begin{bmatrix} I_R & I_G & I_B \end{bmatrix}$, sendo os valores de intensidade do *pixel*, $\hat{\vec{\mu}} = \begin{bmatrix} \mu_R & \mu_G & \mu_B \end{bmatrix}$ e $\hat{\vec{\sigma}} = \begin{bmatrix} \sigma_R & \sigma_G & \sigma_B \end{bmatrix}$ os parâmetros do modelo para cada canal. A distorção de brilho é calculada observando a luminância do *pixel*. Ela é

estimada como:

$$\beta_i = \frac{\frac{I_R\mu_R}{\sigma_R^2} + \frac{I_G\mu_G}{\sigma_G^2} + \frac{I_B\mu_B}{\sigma_B^2}}{\left(\frac{\mu_R}{\sigma_R}\right)^2 + \left(\frac{\mu_G}{\sigma_G}\right)^2 + \left(\frac{\mu_B}{\sigma_B}\right)^2}. \tag{1.22}$$

A classificação, tal como ocorre com o método GMM, é realizada através da limiarização dessas distorções.

1.6 Medidas de similaridade de modelos estatísticos

A medida de similaridade entre os dados e as distribuições gaussianas permite a avaliação da representatividade dos modelos. A escolha da distribuição que melhor representa uma evidência e a estimação do estado ocorrem por meio desse processo. As funções de densidade de probabilidade naturalmente permitem a comparação de seus resultados. Entretanto, devido a complexidade do cálculo, muitas modelagens evitam o uso dessas funções.

Métodos de comparação menos complexos, mas igualmente eficazes, são utilizados quando as comparações são realizadas constantemente. As métricas de distância são frequentemente utilizadas em problemas envolvendo padrões de cores. A distância Mahalanobis (MAHALANOBIS, 1936) é uma medida de distância eficaz e que apresenta resultados melhores do que a distância euclidiana em problemas de segmentação por considerar o erro médio da distribuição (ZIVKOVIC, 2004). O método de Mahalanobis, tal como o *Standard score*, normaliza a distância por meio do desvio-padrão da distribuição. Como consequência, em posse dos parâmetros do modelo, essa distância substitui completamente o cálculo da função de probabilidade.

Essa distância abrange modelos multimodais [1] em oposição ao *Standard score* que é utilizado em problemas unidimensionais. Esse método multimodal utiliza o vetor de média $\vec{\mu}$ e desvio-padrão $\vec{\sigma}$ das distribuições para calcular uma distância invariante. Esse cálculo é feito utilizando a equação,

$$d(\vec{x}, \vec{\mu}, \vec{\sigma}) = \sum_{i=1}^{p} \sqrt{\frac{(x_i - \mu_i)^2}{\sigma_i^2}}, \tag{1.23}$$

onde x_i é o dado, μ_i é a média e σ_i é o desvio-padrão da característica i. Por ser normalizada, a distância d é medida em unidades de desvio-padrão σ.

A distância Mahalanobis substitui os cálculos na decisão de hipóteses da inferência Bayesiana. O teorema de Bayes,

$$R = \frac{p(A|\vec{x})}{p(B|\vec{x})} = \frac{p(\vec{x}|A)p(A)}{p(\vec{x}|B)p(B)}, \tag{1.24}$$

[1] Modelos multimodais são modelos mistos formados por composição.

é reescrito utilizando a distância Mahalanobis. Seja $p(\vec{x}|A)$ estimado utilizando o método GMM, essa probabilidade a posteriori é substituída por $d(\vec{x}, \vec{\mu}, \vec{\sigma})$, sendo $\vec{\mu}$ e $\vec{\sigma}$ os parâmetros do modelo. Zivkovic (2004) considera em seu trabalho que $p(\vec{x}|B)$ é constante sendo substituído por c_{thr}.

1.7 Limiarização

A limiarização assume um papel central na segmentação de imagens devido sua simplicidade e seu baixo custo computacional (GONZALEZ; WOODS, 2006). Nesta seção são discutidas técnicas de divisão de regiões que utilizam os valores de intensidade do *pixel*. As técnicas de limiarização buscam encontrar limiares que separem faixas de valores do histograma minimizando os erros de classificação. A limiarização mais simples é descrita pela equação:

$$threshold(x, y) = \begin{cases} 1, & \text{se } I(x, y) > T \\ 0, & \text{se } I(x, y) \leq T \end{cases}. \tag{1.25}$$

Embora essa abordagem de classificação seja geralmente utilizada para dividir a imagem em duas regiões, a limiarização pode ser estendida para separar mais de dois intervalos de valores.

Há diferentes abordagens que são utilizadas em técnicas de limiarização (CHANG et al., 2006). Gonzalez e Woods (2006) as categorizam em limiarização global e variável. Embora essas denominações possam variar na literatura, caracteriza-se limiarização global quando o limiar T é constante para todos os *pixels* da imagem. As técnicas onde o limiar T é mutável ao longo da imagem são classificadas como limiarização variável (PEDRINI; SCHWARTZ, 2008). Pedrini e Schwartz (2008) define que técnicas que utilizam valores de uma vizinhança sejam chamadas de limiarização local ou regional. Enquanto as que utilizam a coordenadas espaciais sejam chamadas de limiarização dinâmica ou adaptativa.

1.7.1 Limiarização global

A limiarização global é utilizada quando as distribuições de intensidade dos *pixels* apresentam faixa de valores distinguíveis para as classes do problema. Dessa forma, é possível utilizar limiares constantes para realizar a divisão de regiões em toda a imagem. Essa situação pode ser observada na Figura 5 onde a dissimilaridade entre as duas distribuições é suficiente para realizar essa separação.

Em situações como na segmentação de veículos, mudanças de iluminação ocorrem nos pontos ao longo de uma imagem. Técnicas de limiarização global apresentam deficiência para segmentar corretamente imagens nesse contexto. Gonzalez e Woods (2006) defendem que há três abordagens básicas para proceder nessa situação específica: corrigir diretamente

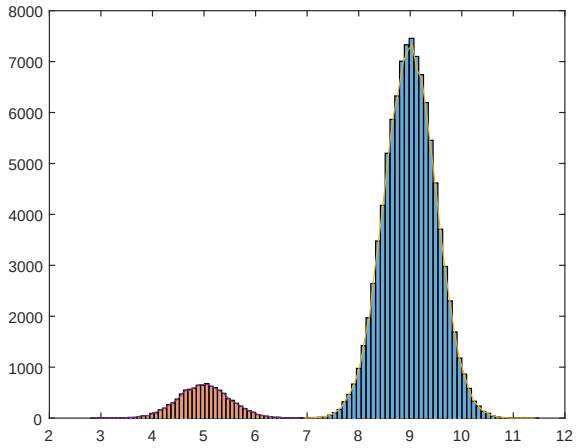


Figura 5 – Histograma com duas distribuições distinguíveis.

o padrão de sombreamento, corrigí-lo por meio de processamento ou utilizar técnicas de limiarização variável.

1.7.2 Limiarização variável

A limiarização variável permite que os valores se adaptem às condições da imagem. Desse modo, as técnicas dessa categoria corrigem mudanças de iluminação e reduzem a influência dos ruídos no processo de segmentação. Existem várias abordagens que podem ser utilizadas, cada uma delas busca analisar os dados sob uma perspectiva diferente (SEZGIN; SANKUR, 2004). Embora isso pareça ter um custo computacional elevado, algoritmos modernos e o *hardware* dos computadores atuais permitem análise de vizinhança por meio de operações lógicas e aritméticas (GONZALEZ; WOODS, 2006).

Embora existam abordagens complexas, técnicas de limiarização variável frequentemente se baseiam em uma das abordagens elementares: particionamento da imagem em sub-regiões ou análise de características locais (GONZALEZ; WOODS, 2006). No particionamento da imagem, a técnica de limiarização é aplicada em cada sub-região da imagem. Enquanto na análise de características locais, os algoritmos consideram a vizinhança do *pixel* para estimar o limiar.

1.8 Limiarização por entropia

Entropia é um conceito físico muito utilizado em estudo de partículas e termodinâmica (ATKINS; PAULA, 2002). Ela é definida como uma grandeza que mede a desordem das partículas em um sistema. Técnicas de limiarização que utilizam as informações de entropia realizam buscas avaliando a desordem do histograma com a divisão dos dados. O entendimento desse conceito básico é fundamental para a compreensão das técnicas que

analisam a disposição das probabilidades do histograma.

A Figura 6 mostra a aplicação das técnicas baseadas em entropia para análise de limiares. Essa figura mostra os dados sendo agrupados por meio de um limiar T. Ao realizar a segmentação, o problema é subdividido em dois subsistemas. Nesse exemplo específico, quaisquer valores no intervalo $]T_0, T_2[$ realizam a segmentação com mesma acurácia. Contudo, o limiar T_1 é o que maximiza o grau de desordem em cada subsistema isolado. Consequentemente, esse limiar é o escolhido pela abordagem da entropia.

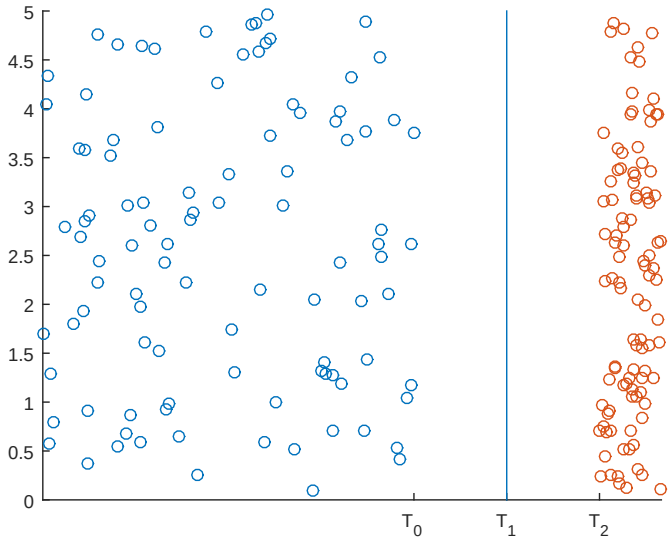


Figura 6 – Entropia de um sistema na divisão do histograma.

Técnicas de limiarização que calculam a entropia dos *pixels* não são tão recentes. Diversas pesquisas nessa área foram propostas na década de 90. Entre os métodos propostos na literatura que utilizam essa abordagem estão os de Pun (1980), Kapur, Sahoo e Wong (1985), Li e Lee (1993), Yen, Chang e Chang (1995), Pal (1996), e Sahoo, Wilkins e Yeager (1997). Estes e outros métodos foram avaliados por Sezgin e Sankur (2004), onde foram submetidos a alguns problemas de segmentação. Entre os métodos baseados em entropia citados, os métodos de Kapur, Sahoo e Wong (1985), e Yen, Chang e Chang (1995) foram bem avaliados e serão utilizados neste trabalho.

1.8.1 Limiar de Kapur

Limiar de Kapur, como é conhecido o método de Kapur, Sahoo e Wong (1985), é um estimador de limiar. Como os métodos baseados em entropia, o limiar de Kapur utiliza a frequência dos *pixels* para determinar a desordem dos subsistemas. Esses subsistemas são criados a partir da divisão dos dados por meio do limiar T no histograma da imagem.

Seja x_i os valores de intensidade dos *pixels* de uma imagem, definimos $p(x_i)$ a

frequência estimada para x_i. $P(x_i)$, calculada pela equação,

$$P(x_i) = \sum_{i=0}^{x_i} p(i), \tag{1.26}$$

é conhecida como Função de Distribuição Acumulada (CDF-*Cumulative Density Function*). Essa função é utilizada para gerar a estimação acumulada de cada bipartição da PDF ou PMF. As estimações $p(x_i)$ e $P(x_i)$ descritas na Equação 1.26 estão presentes nas outras abordagens de limiarização a seguir. Deste modo, torna-se desnecessário a definições redundantes.

O limiar de Kapur é um método de busca onde o objetivo é maximizar a entropia relativa à classificação do primeiro plano e fundo da imagem. Sendo assim, esse método é descrito pela equação,

$$T_{opt} = \underset{T}{\operatorname{argmax}}[H_f(T) + H_b(T)], \tag{1.27}$$

onde $H_f(T)$ e $H_b(T)$ são as entropias de cada subsistema. Essas entropias são calculadas por meio das equações,

$$H_b(T) = -\sum_{t=0}^{T} \frac{p(t)}{P(T)} ln \frac{p(t)}{P(T)}, \text{ e} \tag{1.28}$$

$$H_f(T) = -\sum_{t=T+1}^{L} \frac{p(t)}{1-P(T)} ln \frac{p(t)}{1-P(T)}. \tag{1.29}$$

1.8.2 Limiar de Yen

Outro método de estimação de limiar que utiliza entropia dos dados é o método de Yen, Chang e Chang (1995). Mais conhecido somente como limiar de Yen, esse método é bastante similar ao de Kapur (KAPUR; SAHOO; WONG, 1985). A diferença entre esses métodos se dá no cálculo das entropias. O limiar de Yen utiliza uma equação de maximização definida como,

$$T_{opt} = \underset{T}{\operatorname{argmax}}[C_f(T) + C_b(T)], \tag{1.30}$$

onde $C_f(T)$ e $C_b(T)$ são descritos como,

$$C_b(T) = -\sum_{t=0}^{T} \left(\frac{p(t)}{P(T)} \right)^2, \text{ e} \tag{1.31}$$

$$C_f(T) = -\sum_{t=T+1}^{L} \left(\frac{p(t)}{1-P(T)} \right)^2. \tag{1.32}$$

1.9 Limiarização baseada em agrupamento

Algumas limiarizações baseadas em agrupamento analisadas por Sezgin e Sankur (2004) são o limiar de Otsu (OTSU, 1979), o limiar de Kittler (SEZAN, 1990a) e o limiar de Riddler (OLIVO, 1994). Selecionamos os limiares de Otsu e Kittler para avaliar essa abordagem de limiarização na segmentação dos veículos pelo *ranking* realizado por Sezgin e Sankur (2004).

Os métodos de limiarização de agrupamento consideram histogramas bimodais, ou seja, histogramas que podem ser decompostos em duas distribuições (PEDRINI; SCHWARTZ, 2008). Cada uma dessas distribuições descreve o comportamento dos pontos de fundo e primeiro plano. Para modelar esses métodos, tem sido utilizado Misturas de Gaussianas com duas componentes fixas (SEZGIN; SANKUR, 2004). Essa modelagem por ser observada na Figura 4. Nela, a limiarização utiliza os parâmetros das componentes para determinar o valor que separa as classes minimizando os erros na intersecção das distribuições.

1.9.1 Limiar de Otsu

O limiar de Otsu (1979) se dá pela maximização de função objetivo tal como na limiarização por entropia. O limiar de Otsu (1979) é estimado por meio da equação,

$$T_{opt} = \underset{T}{\operatorname{argmax}} \left[\frac{P(T)[1 - P(T)][\mu_f(T) - \mu_b(T)]^2}{P(T)\sigma_f^2(T) + [1 - P(T)\sigma_b^2(T)]} \right], \tag{1.33}$$

onde μ_b, μ_f, σ_b e σ_f são as médias e desvios resultantes da bipartição do histograma para as classes de plano de fundo e primeiro plano, respectivamente.

1.9.2 Limiar de Kittler

Outro método de limiarização baseado em agrupamento é o limiar de Kittler (SEZAN, 1990a). Esse limiar é semelhante ao limiar de Otsu (1979), entretanto, ele busca minimizar os desvios das classes. A função objetivo desse método é descrita pela equação,

$$T_{opt} = \underset{T}{\operatorname{argmin}}[P(T)\log\sigma_f(T) + [1-P(T)]\log\sigma_b(T) - P(T)\log P(T) - [1-P(T)]\log[1-P(T)]], \tag{1.34}$$

onde σ_b e σ_f são os desvios das classes de fundo e primeiro plano. É importante destacar que o limiar de Kittler não utiliza as médias dessas classes.

1.10 Limiarização com abordagem espacial

Os métodos de limiarização descritos até o momento, em geral, utilizam apenas a frequência dos valores dos *pixels*. Entretanto, podemos analisar a disposição espacial

desses valores, Sezgin e Sankur (2004) analisa alguns deste métodos de limiarização. Entre esses métodos podemos citar os métodos de Pal e Pal (1989) e de Abutaleb (1989). Pal e Pal (1989) propuseram duas funções objetivo para a estimação de um limiar.

Pal e Pal (1989), Lie (1993) e Abutaleb (1989) se baseiam na coocorrência de valores dos *pixels* da imagem para a estimação da entropia bidimensional. Essa abordagem atua nos eixos horizontal e vertical da imagem, sendo descritas por meio da equação,

$$t_{i,j} = \sum_{l=0}^{L} \sum_{k=0}^{L} q_p, \tag{1.35}$$

onde L é a dimensionalidade dos valores que os *pixels* podem assumir. O valor q_p por sua vez é determinado por meio da seleção,

$$q_p = \begin{cases} 1, & \text{se} \begin{cases} I_{l,k} = i \quad \text{e } I_{l,k+1} = j \\ \text{ou} \\ I_{l,k} = i \quad \text{e } I_{l+1,k} = j \end{cases} \\ 0, & \text{caso contrário.} \end{cases} \tag{1.36}$$

A matriz de ocorrência dada por $C_{i,j}$ é calculada em função das transições $t_{i,j}$ como,

$$C_{i,j} = \frac{t_{i,j}}{\sum_{i=0}^{L} \sum_{j=0}^{L} t_{i,j}}. \tag{1.37}$$

Os valores $C_{i,j}$ permitem uma análise da correlação espacial entre valores específicos dos *pixels* e como eles estão distribuídos na imagem. Métodos de limiarização como Pal e Pal (1989) analisam essa correlação para maximizar a segmentação de padrões que, em geral, são mais frequentes.

1.10.1 Limiar de Pal

Pal e Pal (1989) dividem a matriz de ocorrência em quatro regiões e analisam seus valores na estimação de um limiar global. Essas regiões podem ser observadas na Figura 7, onde a matriz está dividida nas regiões A, B, C e D. Esse método estima as entropias para cada uma dessas regiões. Semelhante aos métodos anteriores, é necessário a estimação de uma distribuição cumulativa bidimensional dada por,

$$P_{A,T} = \sum_{i=0}^{T} \sum_{j=0}^{T} C_{i,j}, \tag{1.38}$$

$$P_{B,T} = \sum_{i=T+1}^{L} \sum_{j=0}^{T} C_{i,j}, \tag{1.39}$$

$$P_{C,T} = \sum_{i=0}^{T} \sum_{j=T+1}^{L} C_{i,j}, \tag{1.40}$$

$$P_{D,T} = \sum_{i=T+1}^{L} \sum_{j=T+1}^{L} C_{i,j}. \tag{1.41}$$

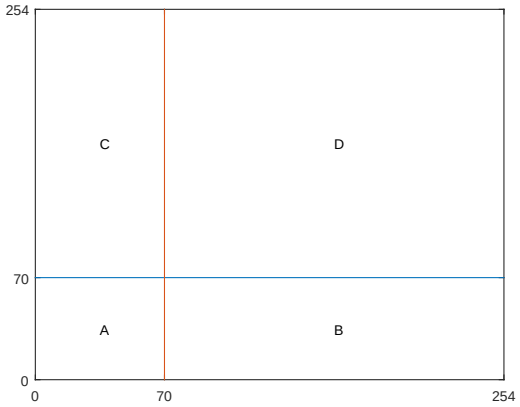


Figura 7 – Matriz de coocorrência.

Tal como ocorre nos métodos anteriores, cada valor da matriz de coocorrência deve ser normalizada com a probabilidade acumulada do quadrante a que pertence. Isso é realizado por meio das equações,

$$p_{i,j}^{A} = C_{i,j}/P_{A,T}, \tag{1.42}$$

$$p_{i,j}^{B} = C_{i,j}/P_{B,T}, \tag{1.43}$$

$$p_{i,j}^{C} = C_{i,j}/P_{C,T}, \tag{1.44}$$

$$p_{i,j}^{D} = C_{i,j}/P_{D,T}. \tag{1.45}$$

As entropias de cada quadrante, por sua vez, são calculadas por meio das equações,

$$H_A(T) = -\frac{1}{2}\sum_{i=0}^{T}\sum_{j=0}^{T} p_{i,j}^{A} \ln p_{i,j}^{A}, \tag{1.46}$$

$$H_B(T) = -\frac{1}{2}\sum_{i=T+1}^{L}\sum_{j=0}^{T} p_{i,j}^{B} \ln p_{i,j}^{B}, \tag{1.47}$$

$$H_C(T) = -\frac{1}{2}\sum_{i=0}^{T}\sum_{j=T+1}^{L} p_{i,j}^{C} \ln p_{i,j}^{C}, \tag{1.48}$$

$$H_D(T) = -\frac{1}{2}\sum_{i=T+1}^{L}\sum_{j=T+1}^{L} p_{i,j}^{D} \ln p_{i,j}^{D}, \tag{1.49}$$

onde $H_A(T)$ e $H_D(T)$ são as entropias de plano de fundo e primeiro plano, respectivamente.

O método de Pal e Pal (1989) possui duas funções objetivo para estimação de limiar. A primeira, descrita pela equação,

$$T_{opt} = \underset{T}{\operatorname{argmax}}[H_A(T) + H_D(T)], \tag{1.50}$$

maximiza a entropia dos quadrantes referentes a classificação do fundo e primeiro plano. $H_B(T)$ e $H_C(T)$ são chamadas de entropias de transição de estados. A maximização dessas entropias intensificam as descontinuidades da imagem. A segunda função do método é

descrita como,

$$T_{opt} = \underset{T}{\text{argmax}}[H_B(T) + H_C(T)]. \tag{1.51}$$

1.11 Considerações finais

Neste capítulo descrevemos técnicas utilizadas em aprendizado probabilístico. Dentre os métodos probabilísticos de aprendizado, foram detalhados dois modelos: a Mistura de Gaussianas e as distorções para métodos não-paramétricos. Esses modelos descrevem o comportamento dos padrões do problema em parâmetros de funções de densidade de probabilidade. Esses parâmetros devem ser estimados por meio de amostras, o método *Expectation-Maximization* foi detalhado e como ele interage com o modelo de misturas.

Dado os padrões de comportamento descritos por meio das funções de densidade de probabilidade, é necessário definir o limiar que separa as classes do problema. Nesta seção descrevemos alguns métodos de limiarização que utilizam diferentes abordagens. Embora este trabalho tenha se concentrado em abordagens de entropia, agrupamento e espaciais, existem outras abordagens de limiarização. Elas utilizam diversas informações da imagem para estimar o limiar de um histograma bimodal. Algumas destas abordagens são a forma do histograma (ROSENFELD; TORRE, 1983; SEZAN, 1990b), os atributos dos objetos (TSAI, 1985; GALLO; SPINELLO, 2000) e a vizinhança dos *pixels* (NIBLACK, 1985; SAUVOLA; PIETIKäINEN, 2000).

2 Sistema de segmentação de veículo proposto e Metodologia utilizada

Este capítulo descreve as propostas deste trabalho: o sistema de segmentação de veículos e a metologia utilizada para a avaliação desse sistema. A metodologia utiliza base de dados propostas neste trabalho e, dessa forma, a metodologia não é detalhada em capítulo isolado.

2.1 Sistema proposto

Esse trabalho propõe uma abordagem para segmentação de veículos por estimação de fundo com limiarização local. O esquema do sistema proposto pode ser observado na Figura 8. Essa figura mostra o sistema dividido em quatro etapas: modelagem de fundo, segmentação, filtragem e realimentação. Há diversos métodos atuando em cada uma destas etapas do sistema. Desse modo, a subdivisão delas permite o detalhamento do processo e objetivo de cada um dos métodos utilizados.

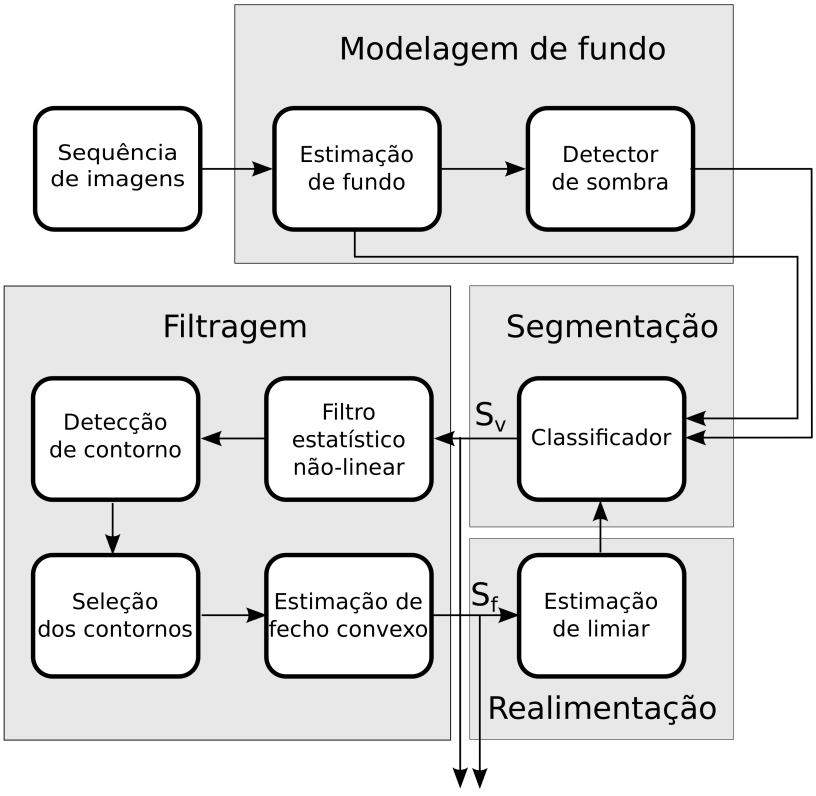


Figura 8 – Esquema do sistema proposto.

Observa-se duas saídas do sistema S_v e S_f, sendo S_v a imagem do processo de

classificação e S_f a imagem resultante do processo de filtragem. Utiliza-se a saída S_v na avaliação dos métodos de classificação. No entanto, a saída S_f é preferível para processos cognitivos. Essa saída já sofreu um processo de filtragem reduzindo custo computacional de módulos de reconhecimento que não são feitos neste trabalho.

O sistema proposto utiliza uma modelagem de duas classes para a classificação dos *pixels*. Essas classes, ou estados, são definidos como BG e FG e representam o plano de fundo e primeiro plano, respectivamente. Eles que são mutualmente exclusivos e formam o universo do problema, de modo que, a classificação pode ser realizada pela estimação de um desses estados. Dentre os dois estados, o plano de fundo é o candidato ideal. Além de possuir o maior número de dados, seu comportamento, embora dinâmico e mutável, pode ser descrito por funções de probabilidade.

De forma sucinta, a modelagem do plano de fundo é composta por dois estágios: o estimador do plano de fundo e um detector de sombras. Enquanto o primeiro classifica os *pixels* de forma ampla observando as intensidade do modelo de cor RGB, o segundo utiliza distorções do padrão para validar as classificações do primeiro estágio e eventualmente corrigir alguns erros.

O modelo estatístico por si só é apenas uma generalização dos dados em parâmetros. É necessário que esses parâmetros sejam estimados para cada alteração da amostra. No caso do modelo de mistura de gaussianas, esses parâmetros são as médias, desvios e pesos das componentes. O histórico do *pixel* é utilizado para calcular esses valores e determinar uma distribuição estatística para BG.

Embora o GMM consiga estimar o padrão de fundo, a precisão desse modelo cai consideravelmente na transição dos padrões das classes. O estágio de detecção de sombra atua nas regiões de transição dos estados BG e FG, onde mudanças de iluminação, como sombras, são frequentes. A detecção de sombra calcula distorções de cromaticidade e brilho dos *pixels* e utiliza esses valores para determinar as sutis mudanças entre os dois estados. Desse modo, esse estágio busca corrigir classificações do primeiro plano reduzindo falsos positivos.

A segmentação possui apenas um único estágio de classificação que separa os dados em duas classes distintas: o plano de fundo e o primeiro plano. Para isso, o classificador utiliza a similaridade dos dados com o modelo e as distorções de cada *pixel* da imagem. Esse processo resulta em uma imagem binária, S_v, com as regiões dinâmicas da cena segmentadas. A classificação é feita por um processo de limiarização, detalhado na Seção 2.1.2.

Os resultados do processo de segmentação frequentemente apresentam erros como ruídos, artefatos e fragmentação de objetos. A etapa de filtragem tem por finalidade corrigir essas falhas da classificação por meio de análises espaciais. Essa etapa é dividida em quatro

estágios onde cada um deles foca na correção de erros específicos. Após as correções desta etapa, os resultados da segmentação são utilizados para aprimorar a segmentação da próxima iteração. A etapa de realimentação manipula a imagem S_f com as correções da segmentação e determina os limiares do classificador da próxima iteração. Esses valores são escolhidos de modo que os erros sejam minimizados.

2.1.1 Modelagem de plano de fundo

Como enfatizado, a segmentação de veículos é modelada em duas classes: BG e FG. A probabilidade a posteriori do estado BG, $p(\vec{x}|BG)$, é estimada através do GMM. Os parâmetros $\widehat{\vec{\mu}}_m$, $\widehat{\vec{\sigma}}_m$ e $\widehat{\pi}_m$ das componentes desse método são calculados pelo método *Expectation-Maximization* utilizando o histórico de valores dos *pixels* de um vídeo para uma coordenada da imagem.

Devido aos estados BG e FG serem mutualmente exclusivos e complementares, $p(BG) + p(FG) = 1$. A decisão bayesiana descrita pela Equação 1.4 é definida como:

$$R = \frac{p(BG|\vec{x})}{p(FG|\vec{x})} = \frac{p(\vec{x}|BG)(1 - p(FG))}{p(\vec{x}|FG)p(FG)}. \tag{2.1}$$

Seja $\frac{p(FG)}{1-p(FG)}$ o coeficiente de primeiro plano c_{pp}, esse coeficiente descreve a proporção de probabilidades para o primeiro plano. Na Equação 2.1 observamos seu inverso c_{pp}^{-1}. O coeficiente c_{pp} é discutido com mais detalhes na Seção 2.1.2 por ser utilizado na realimentação do sistema.

A probabilidade a posteriori de FG, $p(\vec{x}|FG)$, é descrita como uma distribuição constante com o valor p_{thr}. Esse limiar é definido empiricamente visto que ele pode variar para condições diferentes de iluminação. Portanto, a escolha desse limiar não é trivial e depende do problema, de modo que, pequenas variações de p_{thr} influenciam consideravelmente os resultados do sistema. Isso acontece em razão dos objetos de primeiro plano, como carros, serem frequentemente similares ao plano de fundo sob perspectiva de cor.

A estimação da probabilidade a posteriori do plano de fundo não é totalmente precisa na intersecção das classes. Nos trabalhos de Zivkovic (2004) e Kemouche e Aouf (2009), o método de Horprasert, Harwood e Davis (1999) é utilizado para validar os resultados do classificador e identificar alguns erros da classificação. Entre esses erros, as distorções de cromaticidade e brilho, que são calculadas por meio desse método, são utilizadas para detectar mudanças de iluminação que foram classificadas no primeiro plano. A cromaticidade descreve mudança na percepção de cor no modelo, enquanto a distorção de brilho estima as variações de iluminação.

O uso do método não-paramétrico de Horprasert, Harwood e Davis (1999) não tem como enfoque a segmentação de veículos, mas a detecção de realces e as sombras

provocados pela oclusão da luz. Esse problema ocorre com frequência nas vias de tráfego e facilmente resulta na junção de veículos distintos. No entanto, em virtude de veículos de cor escura serem similares à via de tráfego, o método de Horprasert, Harwood e Davis (1999) intensifica falhas de segmentação.

2.1.2 Segmentação

Dada uma imagem I, a segmentação é o processo de classificação dos *pixels* i nas classes BG e FG. Para desempenhar essa função, o classificador analisa a probabilidade do *pixel* assumir o valor de intensidade no modelo e seleciona a classe mais representativa para esse valor. O processo que determina esse valor de separação das classes é chamado de limiarização. A limiarização utiliza pelo sistema proposto é definida pela equação:

$$C(i) = \begin{cases} FG, & \text{se } p(\vec{x}|BG) < p_{thr}c_{pp} \\ BG, & \text{caso contrário} \end{cases}, \tag{2.2}$$

de modo que verifica-se $c_{pp} > 0$. Devido o valor de p_{thr} ser constante, um limiar é criado para substituir a expressão $p_{thr}c_{pp}$. Esse limiar, definido como d_{prob}, é dinâmico e proporcional a c_{pp}.

Como exposto por Zivkovic (2004), a repetição do cálculo da função gaussiana $\mathcal{N}$ gera um custo computacional adicional que pode ser reduzido com medidas de similaridade. O método proposto utiliza a distância Mahalanobis para reduzir o custo e aumentar o desempenho do sistema. A equação de classificação com essa métrica é definida como:

$$C(i) = \begin{cases} FG, & \text{se } d(\vec{x}, \vec{\mu}, \vec{\sigma}) > d_{thr} \\ BG, & \text{caso contrário} \end{cases}, \tag{2.3}$$

onde $\vec{\mu}$ e $\vec{\sigma}$ são os parâmetros da componente mais próxima estimada de $p(\vec{x}|BG)$. O limiar d_{thr}, equivalente ao limiar d_{prob} na Equação 2.2, é dinâmico e determinado pela estimação de crença na classe no *frame*.

É importante notar que as Equações 2.2 e 2.3 não incluem os dados fornecidos pela abordagem não-paramétrica. A classificação envolvendo essas informações é definida como,

$$C(i) = \begin{cases} FG, & \text{se } d(\vec{x}_t, \vec{\mu}, \vec{\sigma}) > d_{thr} \text{ e } (CD_i > \tau_{CD} \text{ ou } \beta_i < \tau_{lo}) \\ BG, & \text{caso contrário} \end{cases}, \tag{2.4}$$

onde CD_i e β_i são as distorções do método de Horprasert, Harwood e Davis (1999) e os limiares τ_{CD} e τ_{lo} são constantes. O limiar d_{thr}, utilizado na Equação 2.4, é discutido com mais detalhes na Seção 2.1.4.

2.1.3 Filtragem

A etapa de filtragem tem como objetivo retificar e realizar as correções necessárias nos resultados do processo de segmentação para que a estimação do limiar d_{thr} seja mais precisa. A segmentação S_v de uma imagem I é submetida aos estágios de filtragem, de modo que os erros de classificação sejam reduzidos. A primeira fase desta etapa, denominada filtragem espacial, tem como função a eliminação dos ruídos gerados pelo sensor. Há duas abordagens populares para essa tarefa: uso conjunto de filtros morfológicos, como erosão e dilatação, e filtros não-lineares, como a mediana (GONZALEZ; WOODS, 2006). Neste trabalho, optou-se pelo filtro de mediana por ter um custo computacional menor.

Após a filtragem espacial, os ruídos segmentados são descartados e as regiões fragmentadas próximas são reconectadas. Normalmente, descritores são utilizados em processos cognitivos para reconhecimento. Entretanto, neste sistema, eles são usados para suavizar as regiões e estimar áreas que podem ter sido ignoradas durante o processo de classificação. Uma abordagem simples e de baixo custo para descrição de formas de objetos é o contorno.

A descrição de contornos é flexível e muitas informações podem ser extraídas dessa descrição, como área, perímetro, convexidade, entre outras (GONZALEZ; WOODS, 2006). O método de Suzuki e Abe (1985) é uma técnica de detecção de contornos robusta e presente na biblioteca de processamento de imagem OpenCV (BRADSKI, 2000). A vantagem dessa técnica, quando comparada a outras técnicas de detecção de contornos, é a organização hierárquica que ela gera dos contornos. Utilizando esses contornos, o sistema pode determinar algumas falhas de segmentação como a ocorrência de cavidades dentro das regiões segmentadas e as deformações na borda dos contornos.

Utilizando e_r, o conjunto dos contornos mais externos detectados após a filtragem espacial, as cavidades internas nas regiões dos objetos são reagrupadas ao objetos. Esta abordagem, no entanto, não detecta cavidades próximas a região de fronteira do contorno dos veículos. Para realizar a correção do contorno, é necessário o uso de fecho convexo. O fecho convexo suaviza a borda do contorno e, desse modo, detecta e agrupa as cavidades na fronteira. Devido a similaridade dos veículos com formas convexas simples, o fecho convexo se torna uma opção eficaz de tratar erros de segmentação.

O método de Sklansky (1982) é utilizado para determinar os fechos convexos hu_n dos contornos e_r. Uma imagem binária S_f é gerada pelo preenchimento dos fechos hu_n estimados pelo método. Essa imagem, denominada neste trabalho como segmentação corrigida, é o resultado do processo de filtragem. A imagem S_f é utilizada para selecionar os limiares do classificador na próxima iteração.

2.1.4 Estimação do limiar

A etapa de realimentação determina os limiares que são utilizados pelo classificador na próxima iteração do sistema. A abordagem de escolha do limiar proposta tem como objetivo a redução dos erros do classificador nas regiões dos veículos. Portanto os limiares estimados não reduzirão os ruídos segmentados fora dessas regiões. Essa abordagem foi escolhida por permitir que a classificação de novos dados possa ocorrer sem interferências.

Em um ambiente dinâmico, a localização dos veículos muda constantemente ao longo de uma sequência de imagens. Para estimar limiares em iterações futuras, é necessário considerar o movimento dos veículos, prever sua trajetória e alterar os coeficientes de primeiro plano da região da localização estimada.

A estimação de movimento de um *pixel* é definida como,

$$I(x_1, y_1, t_1) = I(x_0 + \Delta x, y_0 + \Delta y, t_0 + \Delta t), \tag{2.5}$$

onde x e y são as componentes espaciais e t é a dimensão de tempo. As velocidades em cada dimensão da imagem são definidas como $V_x = \frac{\Delta x}{\Delta t}$ e $V_y = \frac{\Delta y}{\Delta t}$. Supondo que o tempo de execução de uma iteração do sistema é suficientemente baixo, $\lim_{\Delta t \to 0} \frac{\Delta x}{\Delta t}$ e $\lim_{\Delta t \to 0} \frac{\Delta y}{\Delta t}$, o deslocamento dos veículos é pequeno e o rastreamento pode ser descartado assumindo o erro de localização do veículo.

Em variações de tempo pequenas, deduz-se que duas imagens S_v sucessivas são suficientemente similares. Em decorrência disso, a classificação de um *pixel* pode ser auxiliada por seu estado na última iteração, $p(X_t|X_{t-1})$, de modo que $X \in \{BG, FG\}$. Essa estratégia de estimação de estados é conhecida como Modelo de Markov de primeira ordem (MARKOV, 1971). O modelo proposto utiliza a segmentação corrigida S_f para determinar o limiar d_{thr} de cada *pixel*.

O Modelo de Markov utilizado nesse trabalho é descrito por uma matriz de transição constante que segue duas estratégias básicas. A primeira é a redução do limiar d_{thr} em regiões que foram segmentadas ou reparadas no processo da filtragem. Desse modo, o classificador é incentivado a classificar essas regiões como primeiro plano e as correções são reduzidas na próxima iteração. A primeira estratégia é modelada tomando uma distribuição $p(FG|FG) > p(BG|FG)$ na matriz de transição do modelo de Markov.

A segunda estratégia, por sua vez, é estabilizar o limiar d_{thr} nas regiões de plano de fundo. Assim, o classificador tem autonomia para determinar o estado do *pixel* sem interferência da realimentação. Essa estratégia é modelada tomando as distribuições $p(FG|BG) = p(BG|BG)$. É por meio da segunda estratégia que o modelo de fundo consegue classificar novos padrões e atualizar os parâmetros.

Na abordagem proposta, o limiar d_{thr} pode assumir dois valores fixos: l_1 e l_2,

verificando-se $l_2 < l_1$. A seleção desses limiares é realizada por meio da equação,

$$d_{thr} = \begin{cases} l_1, & \text{se } X_{t-1} = BG \\ l_2, & \text{se } X_{t-1} = FG \end{cases}, \tag{2.6}$$

onde X_{t-1} é o estado anterior do *pixel* em S_f. O valor l_1 é chamado de limiar de estabilidade e l_2 é chamado de limiar restritivo.

A Equação 2.6 seleciona o limiar de estabilidade quando não há nenhuma informação sobre a cena ou quando a classificação anterior determinou que não houve mudanças na região. O limiar restritivo é utilizado para aprimorar a segmentação do primeiro plano em regiões específicas da imagem. Ele restringe o processo de segmentação indicando para o classificador as regiões onde se localizam os veículos.

É importante enfatizar que o limiar restritivo l_2 não é utilizado nas regiões de sombras. Portanto, a abordagem proposta não incentiva a segmentação de mudanças de iluminação. Nas regiões de plano de fundo, o método proposto e o método de Zivkovic (2004) funcionam de forma similar. A diferença ocorre nas regiões onde o limiar restritivo é selecionado para o processo de classificação.

2.2 Metodologia

Este capítulo descreve as bases de vídeos utilizadas nos experimentos e detalha as diferenças existentes em cada uma delas. Neste capítulo também são detalhadas todas as métricas de avaliação utilizadas. Esses critérios permitem comparar o sistema proposto a outros métodos existentes na literatura.

2.2.1 Base de dados

Uma das principais dificuldades do problema de segmentação é a escassez de bases de dados. Em consequência disso, uma prática comum dos pesquisadores é a análise dos métodos em bases próprias. Em virtude dessa necessidade, a validação de novos métodos se torna uma etapa dispendiosa.

2.2.2 Base T5i

A base de dado T5i-*Morning* é composta por vídeos gravados utilizando uma câmera Canon T5i. Esse dispositivo foi configurado para fornecer imagens com resolução 640×480 *pixels* e 30 fps (*frames per second*). Entretanto, a resolução das imagens é reduzida para 320×240 *pixels* utilizando a interpolação padrão da biblioteca OpenCV. Essa restrição tem como propósito, aumentar o desempenho do sistema. Um *frame* dos vídeos pode ser observado na Figura 9.

Figura 9 – Imagem retirada da base de dados T5i-*Morning*.

Além da aquisição de bases de dados, é preciso determinar a saída esperada para as imagens da cena. Essa tarefa é executada manualmente por um especialista que determina a saída para todos os *pixels*. Devido ao número excessivo de saídas, essa tarefa se torna inviável. A comparação dos métodos é, portanto, limitada a uma amostra do conjunto de imagens.

A diferença nos resultados dos métodos tradicionais e o método proposto situa-se nas regiões de primeiro plano. A amostragem concentra-se nessas regiões onde as saídas diferem. O conjunto de testes criado nessa fase compõe 107 imagens. Juntas, essas imagens totalizam aproximadamente 10^5 *pixels*. A saída esperada dessas regiões, chamada *ground truth*, é designada por um especialista, manualmente. O resultado desse procedimento é mostrado na Figura 10.

(a) (b)

Figura 10 – (a) Exemplo de região selecionada pela amostragem. (b) Saída esperada para a região selecionada.

2.2.3 Bases NP90

A base T5i-*Morning* é simples e não apresenta muita diversidade de mudanças. No entanto, em uma situação real, as condições do ambiente mudam constantemente. Essa variação por si só não é um obstáculo, pois os métodos descritos neste trabalho são adaptativos. Contudo, as sombras e reflexos afetam o sistema drasticamente. Para uma

análise mais profunda da proposta, é necessário submetê-la ao maior número de mudanças de contexto possíveis. As bases NP90 foram criadas com a finalidade de realizar esse experimento.

O conjunto NP90 é composto por três bases: NP90-*Morning* (200 imagens), NP90-*Midday* (165 imagens) e NP90-*Afternoon* (200 imagens). Essas bases foram feitas nos horários de manhã, meio-dia e tarde. Dessa forma, podemos estimar o comportamento do sistema proposto ao longo de um dia e verificar o comportamento do sistema sob os diferentes efeitos que a luz provoca. É importante enfatizar que foi feita a captura dos vídeos para a base NP90-*Evening*, que tem como foco analisar o comportamento com pouca luminosidade. Entretanto, não foi criado o *ground truth* para essa base específica, consequentemente, não foi possível incluir essa base nos experimentos.

Estas três bases, foram criadas utilizando o mesmo ambiente da base T5i-*Morning* para preservar características como angulação e altura da câmera. A primeira destas bases, a NP90-*Morning*, possui características similares à da base T5i-*Morning*. Essa redundância garante que a única variação para o experimento envolvendo essas três bases é a mudança de características inerentes ao próprio meio, os fatores de variação. Não podendo, nesse caso, relacionar os resultados ao sensor da câmera.

Enquanto a base NP90-*Morning* conta com mudanças graduais de iluminação e sombras e penumbras suaves, a base NP90-*Midday* apresenta intensas mudanças com poucas projeções de sombra. Dessa forma, pode-se analisar a atualização do modelo adaptativo e se ela ocorre de forma satisfatória. A base NP90-*Afternoon*, por sua vez, é a que apresenta a maior complexidade. Nesta base específica, as mudanças de iluminação são intensas e é frequente a ocorrência de sombras e reflexos. Esses reflexos constituem o maior obstáculo desta base. Eles prejudicam a qualidade do vídeo e geram mudanças em grandes regiões da imagem.

2.2.4 Medidas de análise experimental

Criado o conjunto de testes, os vídeos são processados pelos métodos e as saídas são armazenadas. A amostragem realizada para seleção das regiões é aplicada nos resultados dos métodos. A comparação das saídas esperadas e as saídas de cada método geram as matrizes de confusão de cada teste (FAWCETT, 2006). Essa matriz, que é mostrada na Tabela 1, resume-se em determinar as taxas VP (Verdadeiro Positivo), FP (Falso Positivo), FN (Falso Negativo) e VN (Verdadeiro Negativo).

As métricas que permitem a comparação dos métodos são calculadas através da matriz de confusão (FAWCETT, 2006). A primeira, a acurácia, é calculada pela equação,

$$Acurácia = \frac{VP + VN}{FP + FN + VP + VN}. \tag{2.7}$$

Tabela 1 – Matriz de confusão para os estados X_t.

Predito \ **Esperado**	**FG**	**BG**
FG	VP	FP
BG	FN	VN

A acurácia é definida como o acerto geral do sistema. Entretanto, essa taxa não permite uma análise mais detalhada dos resultados. A análise detalhada do sistema é realizada por meio das taxas: sensibilidade, especificidade, valor preditivo positivo (VPP) e valor preditivo negativo (VPN).

A sensibilidade, também chamada de *recall*, mede a capacidade do teste detectar o primeiro plano. Essa taxa, calculada como,

$$Sensibilidade = \frac{VP}{VP + FN}. \tag{2.8}$$

Similar à sensibilidade, a especificidade é calculada como:

$$Especificidade = \frac{VN}{VN + FP}, \tag{2.9}$$

mede a capacidade do estimador em reconhecer o padrão de plano de fundo.

Dado a classificação de um *pixel* nos estados BG e FG, o valor preditivo estima a probabilidade dessa saída estar correta. Tal como a sensibilidade e a especificidade, os valores preditivos analisam as duas saídas do sistema. O valor preditivo positivo, que corresponde a confiabilidade da classificação de primeiro plano, é calculado como:

$$VPP = \frac{VP}{VP + FP}. \tag{2.10}$$

A confiabilidade do modelo de plano de fundo é estimada a partir de,

$$VPN = \frac{VN}{VN + FN}. \tag{2.11}$$

Essa taxa é importante, pois determina a eficácia do modelo de fundo.

Sistemas de classificação, em geral, são avaliados pela sua capacidade de buscar soluções satisfatórias para cada classe do problema. A curva ROC (*Receiver Operating Characteristic*) é uma representação gráfica que analisa as taxas de sensibilidade e especificidade em conjunto. O classificador ótimo é obtido quando na variação do limiar, a sensibilidade e especificidade atingem o acerto máximo. Esse ponto é o mais próximo da coordenada $(0, 1)$.

Essa curva é importante por permitir a análise dos resultados em condições de desbalanceamento nas classes. Nessa situação, o sistema apresenta boas acurácias para determinados limiares que, na verdade, podem apresentar uma tendência de acerto na

classe mais populosa. Em problemas como o abordado neste trabalho, essa métrica é fundamental para comparar os métodos testados.

A curva ROC é uma boa forma de avaliar métodos visualmente. No entanto, quando os métodos são muitos similares a avaliação requer uma medida quantitativa. Uma métrica envolvendo a classificação relativa a cada classe é a F-Measure. Ela é calculada como:

$$F = 2 * \frac{Sensibilidade \cdot Especificidade}{Sensibilidade + Especificidade}. \tag{2.12}$$

2.3 Considerações finais

Neste capítulo detalhamos o funcionamento de cada etapa do sistema proposto e como os métodos utilizados interagem entre si. Foram relatados os objetivos de cada estágio de filtragem e como essas correções são utilizadas para realimentar a classificação e melhorar a precisão da segmentação.

Embora tenhamos utilizado somente o estado anterior no processo Markoviano, é possível utilizar um histórico maior. Contudo, testes preliminares, que não foram explicitados neste trabalho, mostraram que essa abordagem não traz bons resultados sem a estimação de movimento e predição de localização dos veículos. Essa condição foi determinante para este trabalho adotar em seus experimentos um Modelo Markoviano de primeira ordem.

As métricas de avaliação utilizadas são quantitativas, isto é, mensuram os acertos e erros de cada *pixel*. Esses critérios são importantes pois permitem analisar os acertos de cada classe do problema. No entanto, essas métricas não analisam a qualidade das soluções. Falhas de classificação que resultem em cavidades internas são mais fáceis de detectar e corrigir do que fragmentação. Dessa forma, existem outras métricas que podem ser realizadas para comparar a qualidade das soluções.

3 Experimentos e Discussão

Este capítulo descreve detalhadamente os experimentos e resultados. Cada experimento detalhado busca avaliar o sistema proposto sob uma perspectiva diferente. O primeiro considera o comportamento geral do sistema para verificar se ele é uma opção eficaz. O segundo analisa a classificação sob diferentes condições de luminosidade para o mesmo ambiente. O terceiro experimento realiza testes de limiarização adaptativa sob várias abordagens.

3.1 Experimento 1

O método de Zivkovic se concentra na estimação e mantém a classificação com limiares fixos. Este trabalho altera o processo de classificação e preserva a descrição dos dados. O primeiro experimento deste trabalho tem como objetivo analisar se a realimentação pode trazer melhorias para a classificação. Este é o experimento inicial que busca avaliar o comportamento geral da proposta e determinar sua viabilidade em ambientes não-controlados. Em virtude do comportamento ser desconhecido a priori, uma análise de todas as métricas de avaliação se faz necessária. A análise desses dados permite determinar as vantagens e problemas do sistema proposto para que os próximos experimentos enfatizem as situações de falhas.

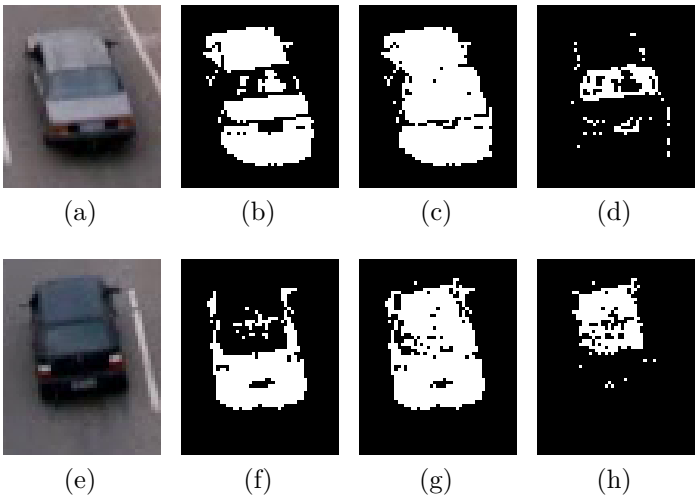
3.1.1 Metodologia do experimento

Como destacado na Seção 2.1.4, a estimação do limiar consiste em selecionar l_1 ou l_2 como limiar d_{thr}. Enquanto l_2 restringe a classificação, l_1 estabiliza-a. Devido a isso, os valores de l_1 e l_2 são limitados à regra $l_2 < l_1$. Esses valores variam nos intervalos $l_1 \in [0{,}8; 7{,}0]$ e $l_2 \in [0{,}6; 6{,}8]$, estes intervalos são amplos o suficiente para observar as mudanças de comportamento e erros de classificação. O limiar c_{thr}, utilizado por Zivkovic, assume a mesma função de l_1, variando no intervalo $c_{thr} \in [0{,}8; 7{,}0]$. Os limiares de distorção τ_{CD} e τ_{lo} assumem os valores padrões indicados por Zivkovic (2004), $\tau_{CD} = 4,0$ e $\tau_{lo} = 0,5$.

A segmentação do método de Zivkovic (2004) e do sistema proposto são comparados ao *ground truth* da base T5i-*Morning*. Essa é a base mais simples e com menor número de regiões, ela é utilizada para estimar o comportamento da proposta em situações triviais do problema. Esse experimento irá avaliar a acurácia, sensibilidade, especificidade, VPP, VPN, e curva ROC das segmentações.

3.1.2 Resultados do experimento

A Figura 11 mostra os resultados do método de Zivkovic (2004) e do sistema proposto para duas regiões selecionadas. Como demonstrado, o método proposto segmenta o primeiro plano com precisão maior. A análise dessa melhoria é realizada por meio da estimação das matrizes de confusão. Essa matriz computa os acertos e erros para cada estado do modelo. Através dessas informações, a classificação é investigada sob várias perspectivas.



(a) (b) (c) (d)

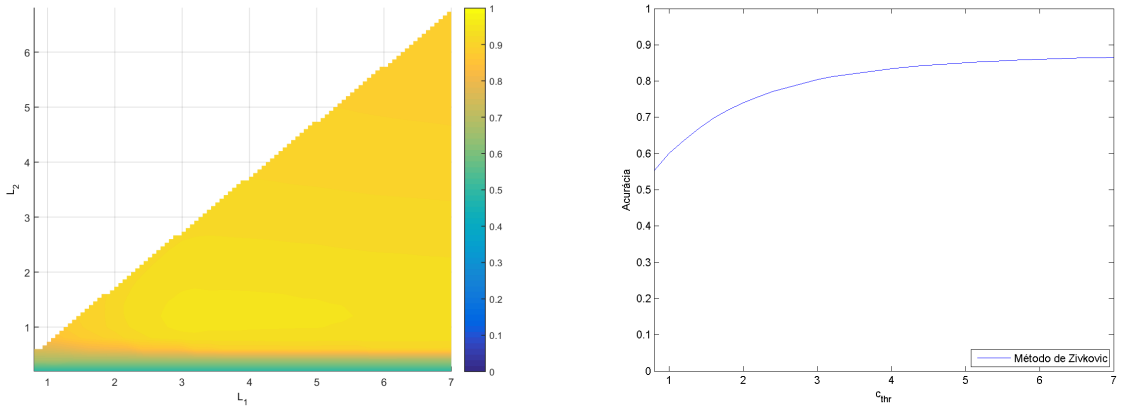
(e) (f) (g) (h)

Figura 11 – Resultados da primeira região: (a) Região da cena, (b) Resultado da classificação pelo método de Zivkovic, (c) Classificação pelo sistema proposto, (d) Diferença entre as classificações. Resultados da segunda região: (e) Região da cena, (f) Resultado da classificação pelo método de Zivkovic, (g) Classificação pelo sistema proposto, (h) Diferença entre as classificações.

A classificação é avaliada por meio das taxas de acurácia, sensibilidade, especificidade, VPP e VPN. Cada uma dessas métricas permite a análise da matriz de confusão sob um aspecto diferente. Em virtude da grande quantidade de testes, os resultados dessas taxas são organizados na forma de gráficos. A Figura 12 apresenta a taxa de acurácia para esses métodos. Enquanto a Figura 12a mostra o comportamento dessa taxa no método proposto, a Figura 12b mostra os resultados da acurácia no método de Zivkovic.

Sob a perspectiva de um limiar l_2 fixo, o comportamento de ambos os métodos é semelhante. Como demonstrado na Figura 12b, a acurácia apresenta um crescimento com o aumento do limiar e converge para a acurácia máxima do método. Sob a perspectiva de l_1, o crescimento da acurácia ocorre de forma brusca. Após uma faixa de valores ótimos, a acurácia retrocede lentamente.

Como pode ser observado na Figura 12a, a faixa de valores ótimos para l_2 é $\Gamma_2 = [0{,}8; 1{,}8]$. Para valores abaixo desse intervalo, o método proposto possui acurácia baixa. A medida que o valor de l_2 se aproxima de l_1, a acurácia declina lentamente até se equiparar à do método de Zivkovic. A estabilização da acurácia ocorre em diferentes pontos



(a) (b)

Figura 12 – (a) Acurácia do método proposto. (b) Acurácia do método de Zivkovic.

para cada método. No método proposto, ela ocorre em $l_1 = 2{,}8$, enquanto no método de Zivkovic ocorre em $c_{thr} = 4{,}0$. O intervalo ótimo para l_1 é descrito como $\Gamma_1 = [2{,}8; 6{,}0]$.

O melhor resultado e a acurácia média, que são mostradas na Tabela 2, evidenciam a melhoria do método proposto. Apesar do pior resultado obtido no sistema proposto ser uma acurácia baixa. Em casos gerais, a abordagem de realimentação utilizada pelo sistema proposto é mostra resultados melhores do que o método de Zivkovic. A dispersão dos resultados é maior no método proposto como pode ser visto na Figura 12a, esses resultados divergentes ocorrem quando l_2 assume valores baixos. Excluindo essa região do espaço de busca, o desvio-padrão dos resultados do sistema proposto converge para valores mais baixos.

Tabela 2 – Dados estatísticos da acurácia.

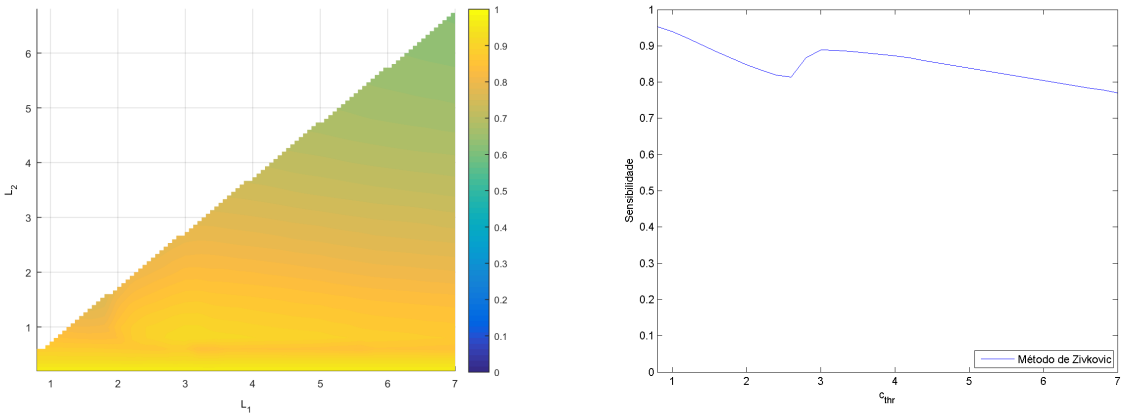
	Máxima	**Mínima**	**Média**	**Desvio-padrão**
Sistema proposto	94,2872%	49,2512%	87,5989%	10,9174
Método de Zivkovic	86,5480%	55,2790%	79,5928%	8,3648

A sensibilidade mostra a porcentagem de dados do primeiro plano que foram corretamente segmentados. Os principais dados dessa taxa são mostrados na Tabela 3. As maiores taxas de sensibilidade encontradas para ambos os métodos são muito próximas. Em virtude da baixa dispersão, a escolha do limiar no método de Zivkovic é imediata. Ao passo que no método proposto é realizada uma busca de valores ótimos para l_2.

Tabela 3 – Dados estatísticos da sensibilidade.

	Máxima	**Mínima**	**Média**	**Desvio-padrão**
Sistema proposto	96,9959%	62,0765%	79,7932%	9,0344
Método de Zivkovic	95,3338%	77,0525%	85,0154%	4,5867

Os gráficos dessa taxa são apresentados para ambos os métodos na Figura 13. Assim como ocorre com a acurácia, l_2 é o parâmetro responsável pela grande dispersão dos resultados. Conforme o valor de l_2 aumenta, a classificação se torna propensa a agrupar os dados no plano de fundo. Como resultado, a taxa de sensibilidade cai significantemente.



(a) (b)

Figura 13 – (a) Sensibilidade do método proposto. (b) Sensibilidade do método de Zivkovic.

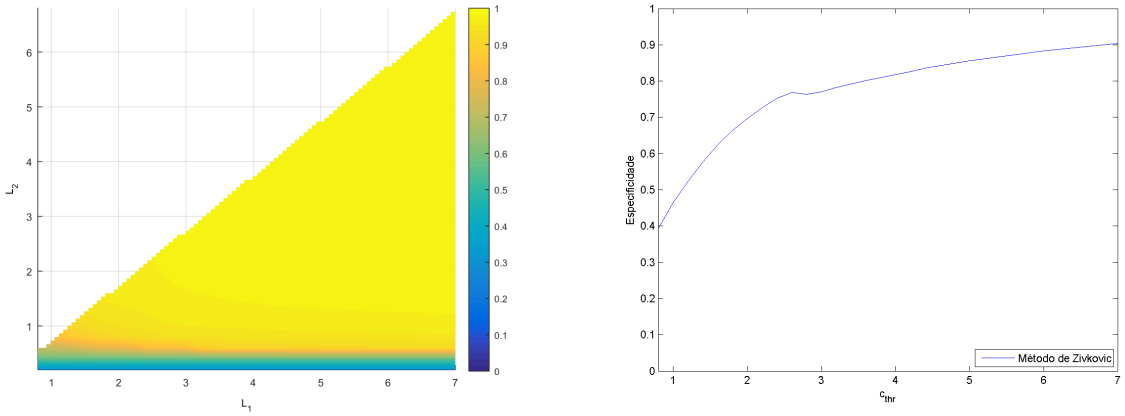
De forma geral, a sensibilidade do método de Zivkovic é melhor. O comportamento dessa taxa se difere em dois aspectos. O primeiro, e mais perceptível, é o declínio mais suave no método de Zivkovic, enquanto no método proposto, a queda dessa taxa é brusca e atinge valores mais baixos. O segundo, entretanto, exige uma atenção maior. Nos intervalos Γ_1 e Γ_2, que foram selecionados por terem as melhores taxas de acurácia, a sensibilidade do método proposto é superior. Nessa abordagem, a sensibilidade dessa região chega à $91,3618\%$, enquanto o método de Zivkovic apresenta $88,9325\%$ na melhor acurácia.

Uma vez que a sensibilidade analisa os resultados da classificação do primeiro plano, a especificidade analisa a taxa de acerto do plano de fundo. Os dados gerais dessa taxa são mostrados na Tabela 4. O método proposto apresenta uma especificidade maior do que a obtida no método de Zivkovic. Consequentemente, essa melhoria contribui para o aumento da acurácia do sistema.

Tabela 4 – Dados estatísticos da especificidade.

	Máxima	**Mínima**	**Média**	**Desvio-padrão**
Sistema proposto	97,9971%	30,1899%	90,7206%	17,2869
Método de Zivkovic	90,3456%	39,2598%	77,4241%	13,1626

Os resultados da especificidade são exibidos na Figura 14. O comportamento dessa taxa apresenta duas diferenças significantes. Como pode ser visto na Figura 14a, a especificidade no sistema proposto possui um crescimento rápido. Ao atingir $l_2 = 1,4$, os resultados se estabilizam em 97% com um aumento ínfimo constante.



(a) (b)

Figura 14 – (a) Especificidade do método proposto. (b) Especificidade do método de Zivkovic.

No método de Zivkovic, entretanto, essa taxa não se estabiliza inicialmente como mostrado na Figura 14b. Em contrapartida ao método proposto, o crescimento é gradual e constante. Após $c_{thr} = 2{,}6$, o sistema converge para a especificidade de 90%. Esses resultados mostram que o método proposto não só melhora a segmentação, mas também aprimora a classificação do plano de fundo.

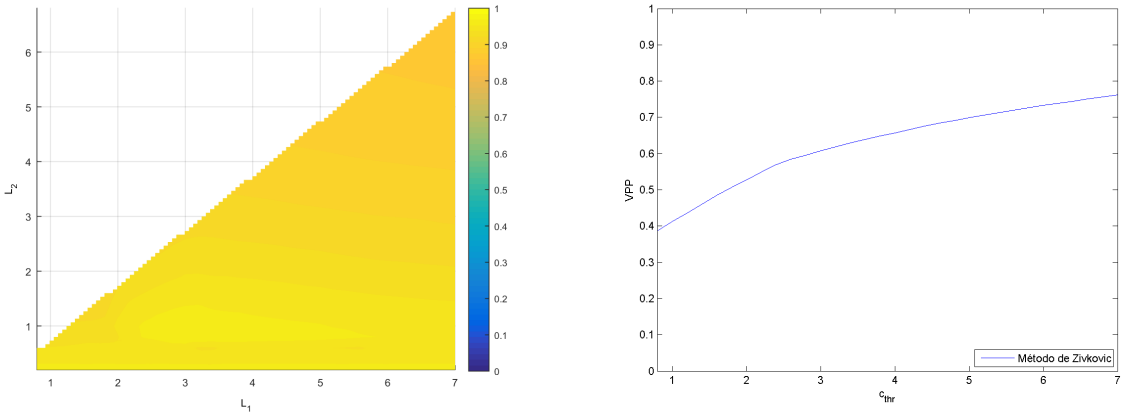
Após a avaliação das taxas de acerto: acurácia, sensibilidade e especificidade, é necessário determinar a precisão do processo de classificação. Os valores preditivos do experimento determinam essa crença nos resultados. A Tabela 5 mostra a precisão da segmentação. Analisando essa tabela, o método proposto possui uma precisão maior que a apresentada pelo método de Zivkovic.

Tabela 5 – Dados estatísticos da taxa VPP.

	Máxima	**Mínima**	**Média**	**Desvio-padrão**
Sistema proposto	93,4094%	35,6995%	85,1371%	16,4082
Método de Zivkovic	76,1445%	38,5640%	62,7483%	10,8287

A Figura 15 apresenta os resultados da taxa VPP visualmente. Como pode ser visto na Figura 15a, assim como ocorre com o limiar c_{thr}, os piores resultados se concentram quando l_2 assume valores baixos. Nessa ocasião, a precisão do método proposto atinge resultados inferiores ao do método de Zivkovic. O crescimento da taxa é excessivamente rápido no método proposto. Após o valor $l_2 = 0{,}8$, o crescimento é reduzido. A precisão da segmentação se estabiliza quando $l_2 \approx 2{,}0$.

Enquanto a taxa VPP determina a confiabilidade do classificador na segmentação do veículo, o valor preditivo negativo estima a precisão do modelo de plano de fundo. A Tabela 6 apresenta de forma geral o comportamento dessa taxa. A abordagem proposta



(a) (b)

Figura 15 – (a) VPP do método proposto. (b) VPP do método de Zivkovic.

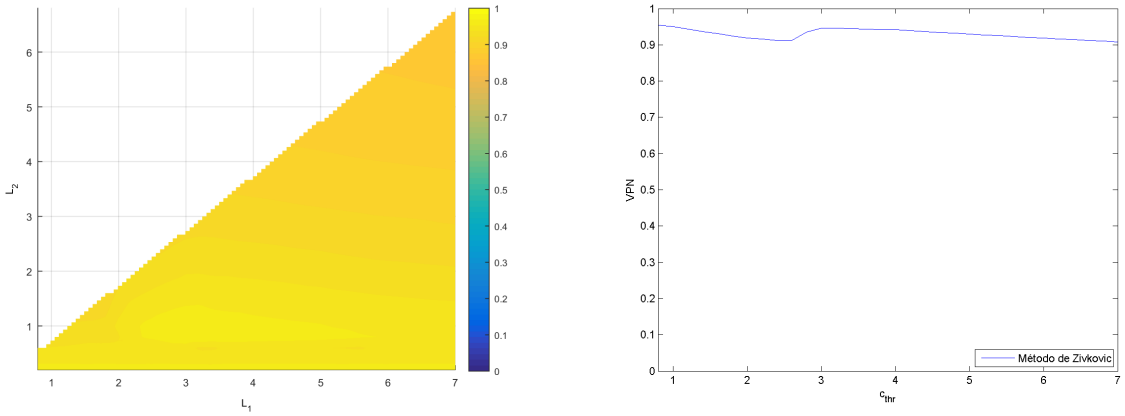
influencia positivamente na precisão da modelagem de plano de fundo. Nessa tabela, o método proposto apresentou um ganho na precisão da classificação do fundo. Entretanto, considerando as médias semelhantes e a dispersão do método de Zivkovic baixa, esse método apresenta uma estabilidade melhor.

Tabela 6 – Dados estatísticos da taxa VPN.

	Máxima	**Mínima**	**Média**	**Desvio-padrão**
Sistema proposto	96,7568%	86,5974%	92,1467%	2,7215
Método de Zivkovic	95,4623%	90,7785%	92,9522%	1,3332

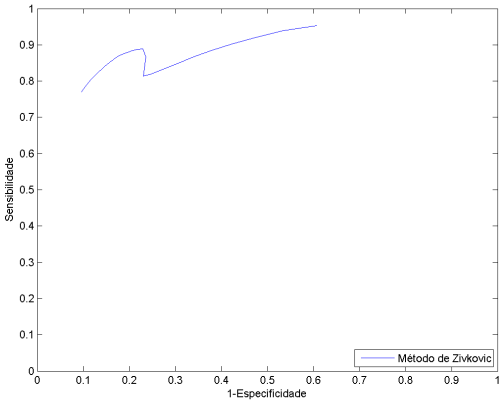
A Figura 16 apresenta todos os resultados da taxa VPN, graficamente. No método de Zivkovic, a Figura 16b mostra que o ápice da taxa está em $c_{thr} = 3$. Após esse valor a precisão entra em declínio devido aos pontos do primeiro plano que são classificados como plano de fundo. No método proposto, como mostrado na Figura 16a, há uma região de ótimos locais em $l_1 \in \Gamma_1$ e $l_2 \in \Gamma_2$. Nesses pontos, o método de Zivkovic apresenta uma precisão para o plano de fundo de 93,0506%, enquanto o sistema proposto apresenta uma precisão de 94,5677%.

Estimados os acertos e a precisão da classificação, a Curva ROC analisa as taxas de sensibilidade e especificidade para encontrar o melhor limiar do experimento. Esse limiar é determinado como o valor no qual essas duas taxas são maximizadas. Ambos os métodos apresentam o mesmo comportamento. Ao passo que a especificidade decai, a sensibilidade cresce até atingir seu ápice. Como pode ser observado na Figura 17, no método de Zivkovic, a melhor classificação do primeiro plano é alcançada com 78% de especificidade. No método proposto, o crescimento da sensibilidade é mais acentuado. No limiar ideal encontrado, $l_1 \approx 3{,}2$ e $l_2 \approx 0{,}8$, a sensibilidade é mais alta. Isso mostra que o método proposto é mais eficaz na classificação do primeiro plano, por conseguinte, na segmentação dos veículos.

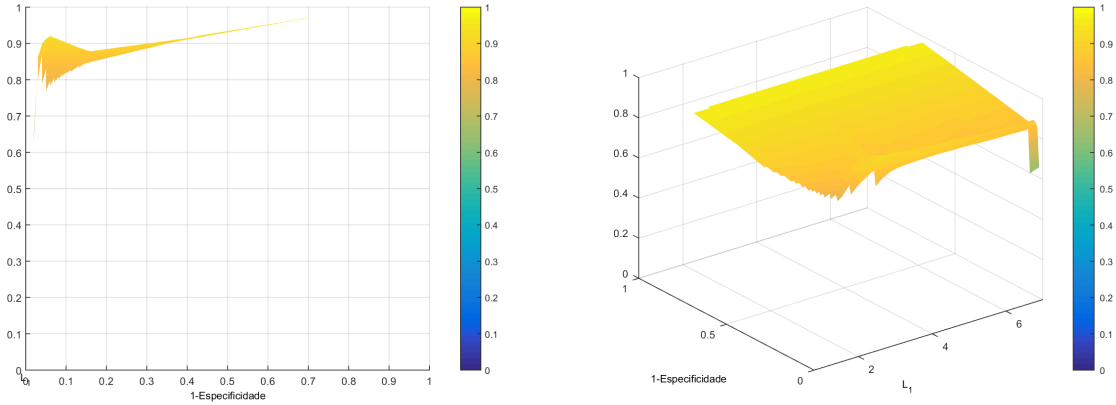


(a) (b)

Figura 16 – (a) VPN do método proposto. (b) VPN do método de Zivkovic.



(a)



(b) (c)

Figura 17 – (a) ROC do método de Zivkovic. (b) ROC do método proposto. (c) ROC do método proposto em perspectiva.

3.1.3 Discussão dos resultados

Dados os resultados das métricas de avaliação, é interessante agrupar os valores de l_1 e l_2 em cinco regiões: R_1, R_2, R_3, R_4 e R_5. A razão desse agrupamento é que os resultados destas regiões apresentam comportamento semelhante. Assim, a divisão facilita a discussão dos resultados. Essa divisão é apresentada pela Figura 18. Na região R_1 se localizam os melhores resultados do método proposto, enquanto em R_5 se localizam os piores resultados. As regiões R_2 e R_3 são regiões de transição e seu comportamento varia em cada métrica de avaliação.

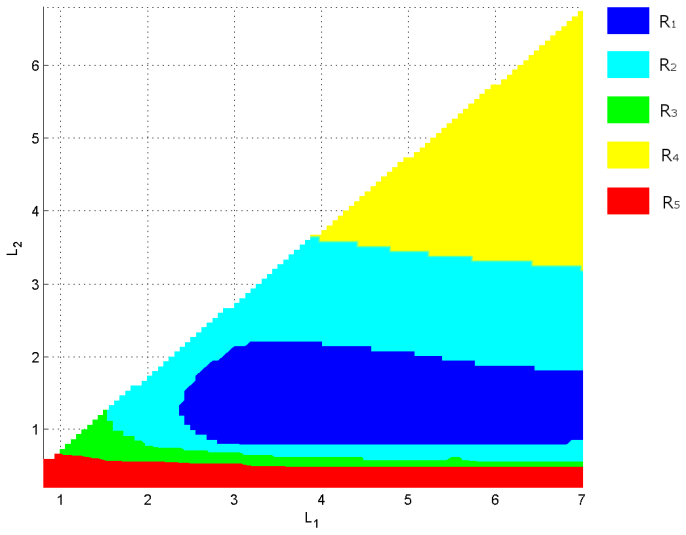


Figura 18 – Divisão dos testes em 4 grupos.

O método de Zivkovic apresenta uma taxa de acurácia sem ótimos locais. O aumento do limiar resulta na convergência da acurácia em aproximadamente 87%. Na abordagem utilizada pelo método proposto, a classificação atinge seu ápice em R_1. Com o aumento de l_1, a abordagem tem um comportamento semelhante ao método de Zivkovic, estabilizando em 92%. A estabilização da taxa de acurácia se inicia em R_2 e permanece em R_4. O rápido crescimento dessa taxa ocorre na região de transição R_3.

Apesar do objetivo ser a maximização da segmentação dos veículos, em virtude do desbalanceamento das classes, os resultados da sensibilidade não influenciam intensamente no acerto geral do sistema. Esse desequilíbrio propicia que os resultados do plano de fundo atuem diretamente na acurácia dos testes. Esse comportamento é visível na região R_5. Nessa região se encontram as maiores taxas de sensibilidade, porém, a especificidade dessa região é baixa. Consequentemente, a acurácia apresenta baixos valores refletindo a taxa de especificidade da região.

Esse erro também é percebido nos valores preditivos. Os resultados do valor preditivo negativo para a região R_5 são excelentes. Entretanto, o valor preditivo positivo dessa

mesma região é baixo. Isso indica que boa parte das regiões de fundo são classificadas como primeiro plano. A Figura 19 mostra o resultado da classificação com limiares baixos.

(a) (b) (c)

Figura 19 – (a) Região selecionada para análise. (b) Resultado do método proposto com limiar da região R_5. (c) Saída esperada para a região selecionada.

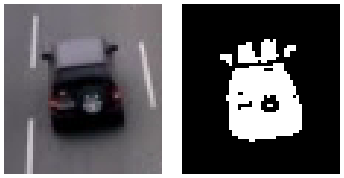
Observando o resultado fora da amostra, a acurácia para limiares na região R_5 é mais baixa do que a identificada no experimento. Como os resultados são referentes a uma amostra focada nos veículos, a influência da taxa de especificidade é reduzida. A Figura 20 mostra que a presença de falhas de segmentação são mais frequentes fora da amostra.

(a) (b)

Figura 20 – (a) Imagem selecionada para análise. (b) Resultado do método proposto com limiar da região R_5.

Na região R_4 ocorre a estabilização da acurácia tal como no método de Zivkovic. Entretanto, isso não significa que as taxas de sensibilidade e especificidade se alteram. Nessa região, a sensibilidade decai gradualmente, enquanto a especificidade se mantém constante. Devido ao desbalanceamento das classes, a queda da acurácia não é brusca. Considerando a taxa VPP dessa região, a precisão da segmentação dos objetos é alta. Assim, as mudanças de iluminação não são classificadas como primeiro plano.

Entretanto, a precisão na classificação do plano de fundo cai consideravelmente. Apesar dos limiares em R_4 apresentarem uma acurácia razoável, a taxa VPN indica que muitas regiões de primeiro plano são classificadas como fundo. Sendo assim, a fragmentação é um problema comum nesses limiares. Essa falha pode ser observada na Figura 21. Em veículos com cores escuras, as diferenças sutis entre as cores da via e do veículo são ignoradas.



(a) (b)

Figura 21 – (a) Região selecionada para análise. (b) Resultado do método proposto com limiar da região R_4.

A região R_3 possui a maior variação nos resultados das métricas. Em virtude do aumento da taxa de especificidade, a acurácia progride nessa região. Entretanto, a especificidade alta não significa uma classificação de fundo perfeita. O valor preditivo dessa região indica que ainda há algumas regiões de veículos que foram incorporadas no plano de fundo. A sensibilidade entra em declínio, entretanto, observa-se um aumento na precisão da segmentação. Desse modo, a segmentação dos veículos melhora em relação a R_5.

A especificidade chega no ponto de máximo em R_2. Nesses limiares, é alcançada a máxima porção de dados do plano de fundo que pode ser classificada pelo método proposto. Essa hipótese é confirmada pela observação da taxa VPP. O valor preditivo positivo dessa região indica que as mudanças de iluminação não estão sendo classificadas como objetos. Porém, a fragmentação ainda é um problema nesses limiares. Em geral, a sensibilidade e o VPN dos limiares em R_2 são iguais ou inferiores às mesmas taxas em R_3.

A especificidade se mantém estável na região R_2 e R_1 e melhor acurácia é atingida na região R_1 com o aprimoramento da classificação do primeiro plano. A abordagem de seleção de limiar é mais efetiva nessas configurações. A segmentação dos veículos aumenta consideravelmente e a fragmentação é reduzida, contudo, não é eliminada. Os resultados das regiões R_2 e R_1 são mostrados na Figura 22. A abordagem proposta resulta em um aumento da sensibilidade e VPP, como consequência, a segmentação dos veículos é melhorada.

(a) (b) (c)

Figura 22 – (a) Região selecionada para análise. (b) Resultado do método proposto com limiar da região R_2. (c) Resultado do método proposto com limiar da região R_1.

No método de Zivkovic, as taxas de sensibilidade e acurácia não convergem para os mesmos valores. A escolha do limiar nesse método é baseado no equilíbrio e na maximização

dos objetivos do problema, neste caso, a sensibilidade. Como os resultados das métricas de sensibilidade e acurácia na abordagem proposta tendem a convergir para a região R_1. A escolha dos valores de l_1 e l_2 é imediata. Sem que haja ponderação das consequências no plano de fundo.

Como descrito no Capítulo 2, a função de l_1 equivale ao do limiar c_{thr} do método de Zivkovic. Essa premissa é corroborada com os resultados da acurácia. Os melhores resultados observados para ambos os métodos foram quando $c_{thr} \approx l_1 \approx 3$. O diferencial da abordagem proposta permitiu o aumento da sensibilidade nas regiões dos veículos. Assim, as cavidades que só seriam segmentadas com limiares menores são detectadas.

3.2 Experimento 2

O experimento descrito na Seção 3.1 permite visualizar o comportamento do sistema sob perspectiva de uma única situação. Essa reação pode se repetir para diferentes circunstâncias do ambiente ou apresentar variações dado os fatores de variação. O experimento descrito nesta seção tem por objetivo analisar a precisão do sistema para essas condições ambientais adversas.

3.2.1 Metodologia do experimento

A análise deste experimento é feita tal como na Seção 3.1.1. Entretanto, foi reduzido o intervalo de escolha dos valores l_1 e l_2 do limiar d_{thr}. Essa redução se justifica devido ao fato de que os resultados apresentados mostram uma convergência com limiares $l_1 > 5$ e $l_2 > 5$. Dessa forma, os limiares são limitados aos intervalos $l_1 \in [0{,}8; 5{,}0]$ e $l_2 \in [0{,}6; 4{,}8]$ para reduzir o tempo gasto no experimento. Os limiares da distorção de cromaticidade τ_{CD} e brilho τ_{lo}, por sua vez, assumem os valores padrões de Zivkovic (2004), $\tau_{CD} = 3{,}0$ e $\tau_{lo} = 0{,}5$.

No experimento anterior, todas as métricas foram utilizadas para comparar o sistema proposto ao método de Zivkovic (2004), embora todas as métricas contribuam para a compreensão dos resultados. Neste e nos próximos experimentos deste trabalho, as medidas de VPP e VPN não serão mais utilizadas, visto que já foi demonstrado que elas são reflexos do desbalanceamento das classes nas imagens. As métricas utilizadas serão acurácia, sensibilidade, especificidade e F-Measure. Embora a curva ROC seja uma boa métrica, ela fornece uma informação mais gráfica. Ela é substituída pela F-Measure nesta análise como uma métrica de avaliação que analisa os acertos ponderados das classes.

3.2.2 Resultados do método de Zivkovic

Zivkovic (2004) utiliza os mesmos valores para os limiares c_{thr} e τ_{CD} que corresponde aos limiares dos atributos M (Distância Mahalanobis) e C (Distorção de cromaticidade) e indica o limiar $\tau_{lo} = 0{,}5$ para o atributo L (Distorção de Luminância). A Figura 23 mostra o resultado do método de Zivkovic (2004) para as três bases NP90. Os resultados desse método para estas bases possuem o mesmo comportamento. A acurácia cresce a medida que o limares c_{thr} e τ_{CD} aumentam e τ_{lo} diminui.

O crescimento da acurácia ocorre a partir de $c_{thr} \geq 2$ e $\tau_{CD} \geq 2$ nas bases NP90-*Morning* e NP90-*Midday*. Entretanto, é perceptível que a distorção de brilho é determinante na classificação dessas bases. Além dos resultados não serem satisfatórios, observa-se que há uma instabilidade dos resultados da acurácia. A acurácia só passa crescer em limiares c_{thr} e τ_{CD} altos.

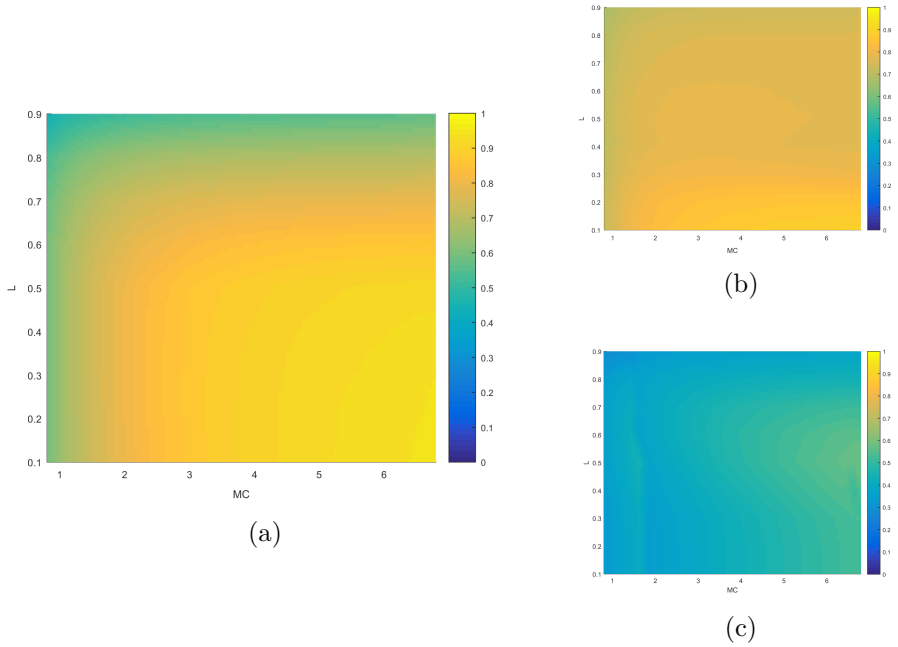
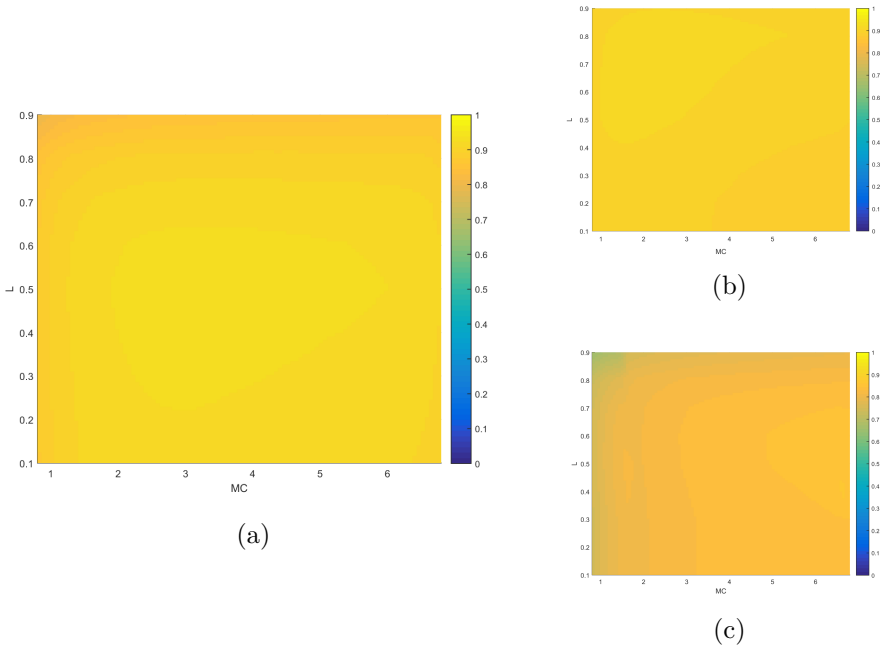


Figura 23 – Acurácia do método de Zivkovic nas bases: (a) NP90-*Morning*, (b) NP90-*Midday*, e (c) NP90-*Afternoon*.

A acurácia por si só não revela o comportamento do sistema de forma detalhada. Para isso é necessário analisar os resultados da sensibilidade e a especificidade do experimento. A Figura 24 mostra os resultados referentes a classificação do primeiro plano. O comportamento dessa métrica específica é diferente em cada uma das bases. Entretanto, essa métrica apresenta um padrão similar. Inicialmente ela cresce até atingir um valor máximo, então entra em declínio. Na base NP90-*Midday*, a classificação do primeiro plano atinge seu ápice rapidamente, enquanto na NP90-*Morning* o crescimento é mais lento. Na base NP90-*Afternoon* o crescimento é bem gradual, não podendo ser observado, na Figura 24, o declínio da taxa.



(a) (b) (c)

Figura 24 – Sensibilidade do método de Zivkovic nas bases: (a) NP90-*Morning*, (b) NP90-*Midday*, e (c) NP90-*Afternoon*.

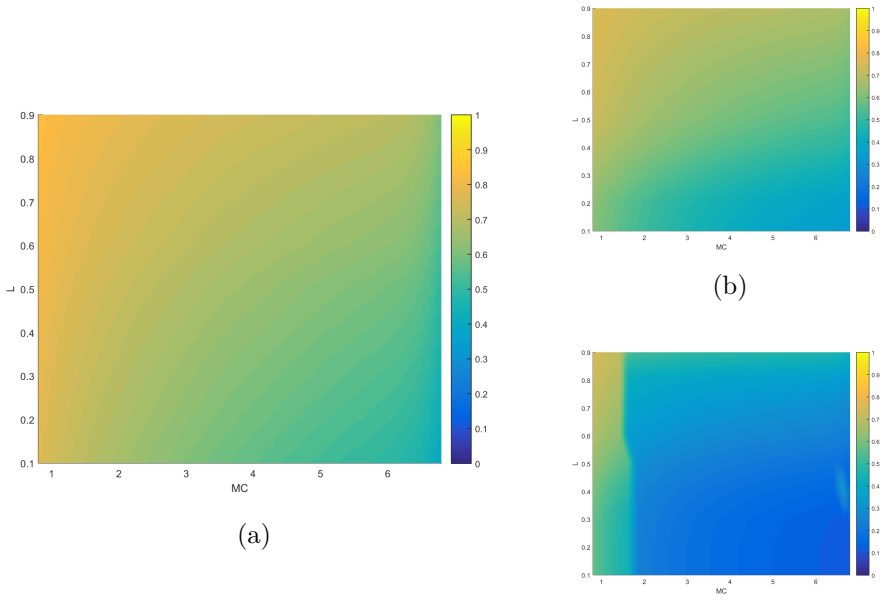
A sensibilidade é uma boa métrica de avaliação porque permite analisar a classificação do primeiro plano e, consequentemente, dos veículos. Embora, em algumas situações, o processo de classificação espere uma proporção inversa entre a sensibilidade e especificidade, na prática, isso pode não ocorrer. A Figura 25 mostra os resultados obtidos para a especificidade. O comportamento da especificidade nestas três bases é o declínio da taxa com o aumento dos limiares c_{thr} e τ_{CD}. Nas bases NP90-*Morning* e NP90-*Midday*, o decaimento da especificidade é gradual, ao passo que, na base NP90-*Afternoon*, é súbito.

Observando as taxas de acurácia, sensibilidade e especificidade, é perceptível que a análise da classificação sob contexto geral, assim como a relativa a uma única classe, não são suficientes pra determinar os melhores limiares para o problema. A métrica F-Measure, mostrada na Figura 26, é utilizada para equilibrar os acertos entre as classes. Esta figura mostra que o método de Zivkovic varia à medida que as condições mudam. A F-Measure assume um comportamento similar à acurácia.

3.2.3 Resultados do sistema proposto sob perspectiva da base

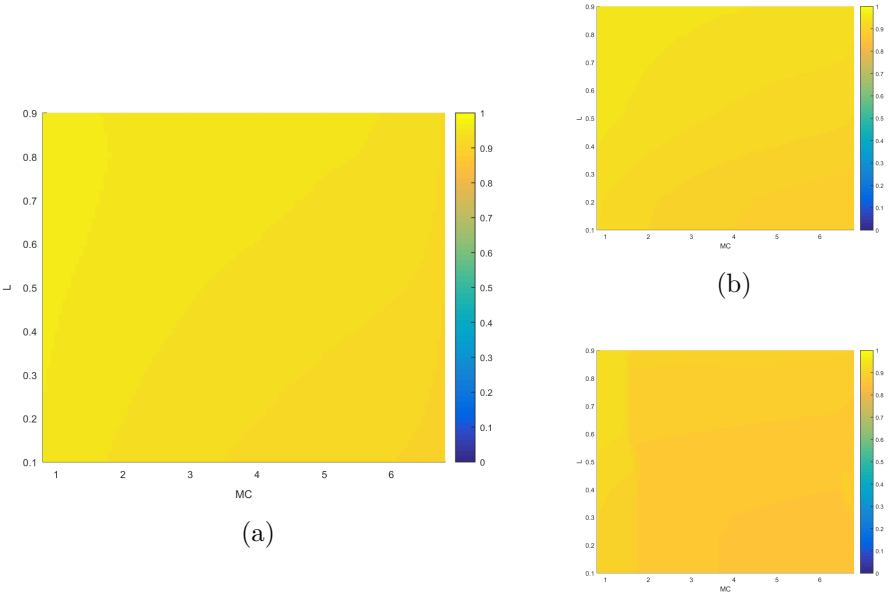
A primeira análise que deve ser feita é sob perspectiva das bases tal como foi realizado com o método de Zivkovic (2004). Esse teste analisa os resultados da alternância do limiar d_{thr}, que é referente ao canal M, entre dois valores pré-definidos para as bases NP90. Neste experimento, foram utilizados os valores $\tau_{CD} = 2,4$ e $\tau_{lo} = 0,5$ para os limiares de distorção de cromaticidade e brilho.

A primeira métrica a ser analisada é a acurácia, mostrada na Figura 27. Em



(a) (b) (c)

Figura 25 – Especificidade do método de Zivkovic nas bases: (a) NP90-*Morning*, (b) NP90-*Midday*, e (c) NP90-*Afternoon*.



(a) (b) (c)

Figura 26 – F-Measure do método de Zivkovic nas bases: (a) NP90-*Morning*, (b) NP90-*Midday*, e (c) NP90-*Afternoon*.

comparação ao método de Zivkovic, o sistema proposto mostra um comportamento estável para todas as bases do experimento. A medida que o limiar l_1 cresce, a acurácia também aumenta. Embora não seja tão perceptível na base NP90-*Morning*, valores mais baixos de l_2 resultam em um aumento de acurácia.

Assim como na acurácia, a sensibilidade também apresenta um padrão comum

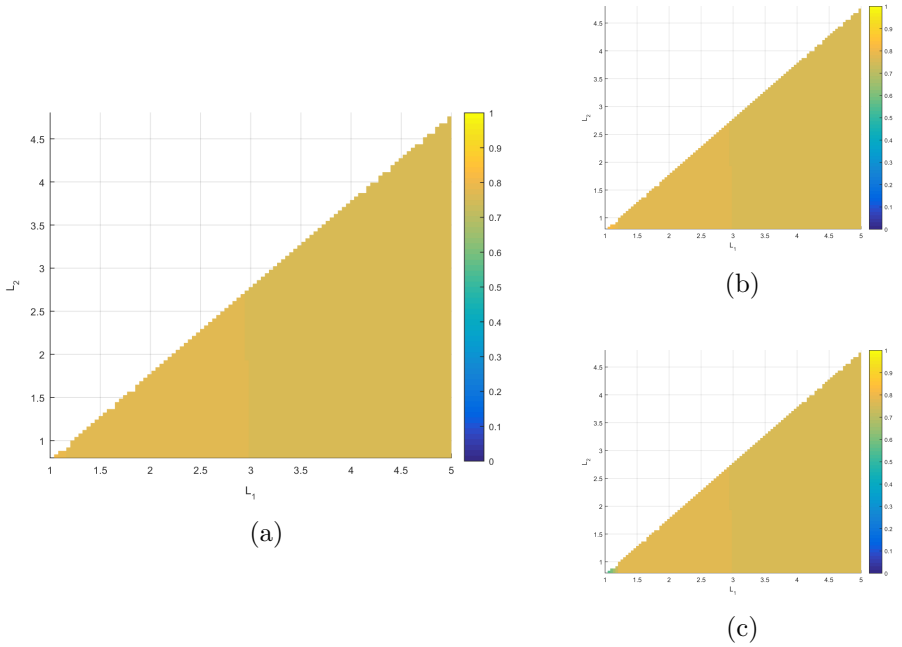


Figura 27 – Acurácia do sistema proposto nas bases: (a) NP90-*Morning*, (b) NP90-*Midday*, e (c) NP90-*Afternoon*.

para as três bases. Essa métrica é mostrada na Figura 28. Embora, possa parecer que exista uma simetria nos resultados, ela não se aplica fora dos limites do intervalo testado devido a abordagem utilizada na realimentação.

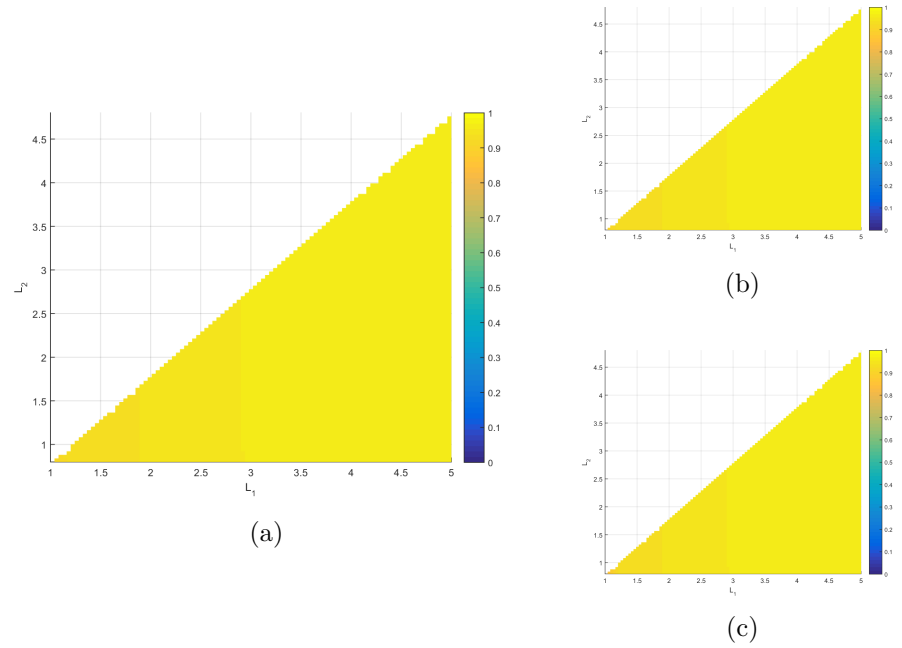
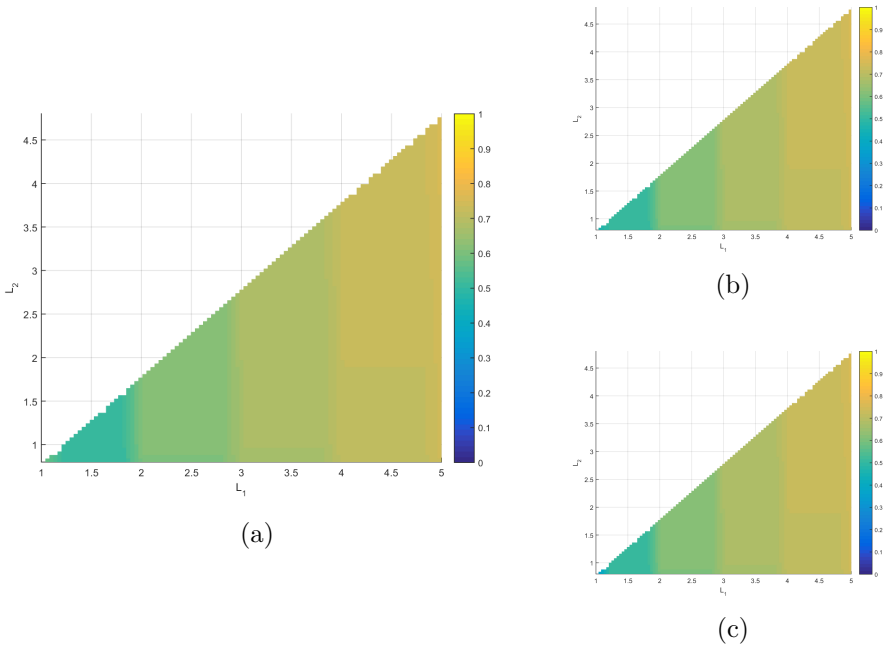


Figura 28 – Sensibilidade do sistema proposto nas bases: (a) NP90-*Morning*, (b) NP90-*Midday*, e (c) NP90-*Afternoon*.

A especificidade segue o mesmo comportamento para as bases, contudo, com algumas variações. Os resultados da especificidade são mostrados na Figura 29. Enquanto

na base NP90-*Morning* os melhores são obtidos com valores l_2 baixos, nas bases NP90-*Midday* e NP90-*Afternoon* são obtidos com valores l_2 mais altos. Observa-se uma redução da especificidade ao longo do dia, no entanto, ela é mais significativa na base NP90-*Afternoon*. Como exposto anteriormente, este experimento utiliza a métrica F-Measure como análise do média de acertos relativos às classes. Os resultados dessa métrica são mostrados na Figura 30. Essa medida de avaliação apresenta um comportamento próximo ao da acurácia e especificidade.



(a) (b) (c)

Figura 29 – Especificidade do sistema proposto nas bases: (a) NP90-*Morning*, (b) NP90-*Midday*, e (c) NP90-*Afternoon*.

3.2.4 Resultados do sistema proposto sob perspectiva do limiar da distorção de cromaticidade

O experimento realizado na Seção 3.2.3 realiza variações apenas no limiar d_{thr}. Entretanto, é necessário observar a influência do limiar τ_{CD} nos resultados do sistema. A primeira métrica a ser analisada sob essa perspectiva é a acurácia. A Figura 31 mostra os resultados dessa métrica com a variação do limiar τ_{CD}. Essa figura mostra que o comportamento é similar sob variação da distorção de cromaticidade. No entanto, o aumento da acurácia é proporcional à τ_{CD}.

Embora a acurácia cresça com o aumento do limiar τ_{CD}, é necessário observar os acertos relativos a cada classe. Essa necessidade se deve ao fato da existência do desequilíbrio entre primeiro plano e plano de fundo. A Figura 32 mostra os resultados da sensibilidade sob variação do limiar de distorção de cromaticidade. O aumento de τ_{CD} resulta em uma queda da sensibilidade. A taxa de especificidade, por sua vez, é mostrada

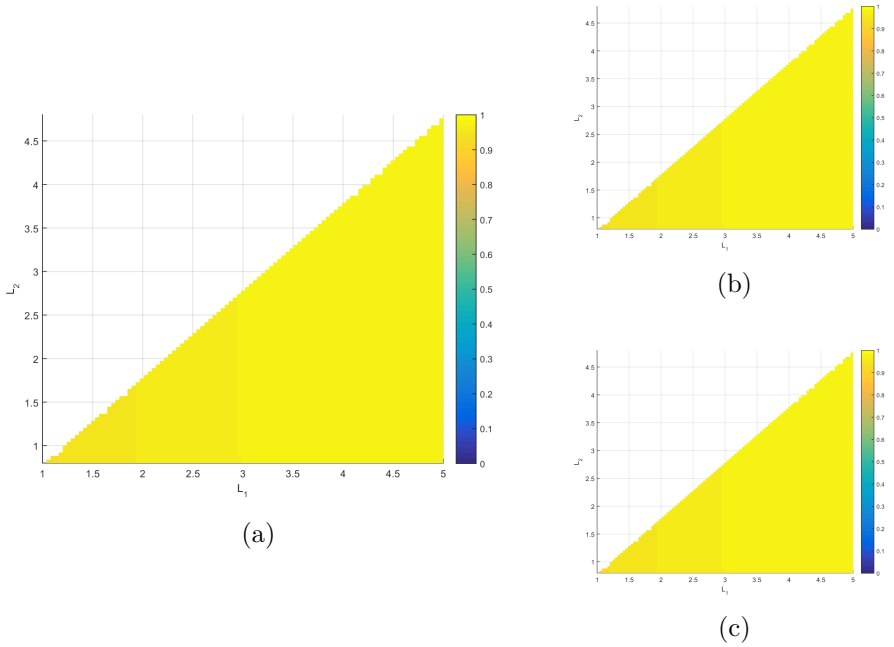


Figura 30 – F-Measure do sistema proposto nas bases: (a) NP90-*Morning*, (b) NP90-*Midday*, e (c) NP90-*Afternoon*.

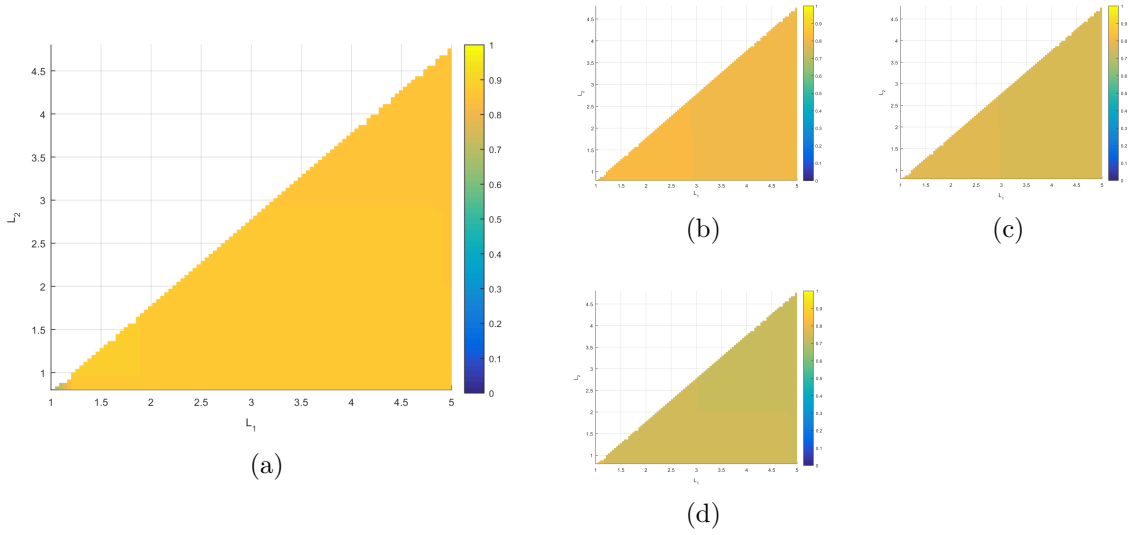


Figura 31 – Acurácia do sistema proposto com na base NP90-*Morning* com limiar: (a) $\tau_{CD} = 1{,}8$, (b) $\tau_{CD} = 2{,}8$, (c) $\tau_{CD} = 3{,}8$, e (d) $\tau_{CD} = 4{,}8$.

na Figura 33. Enquanto a sensibilidade decai com o aumento do limiar τ_{CD}, a especificidade aumenta.

Os resultados para a métrica F-Measure são mostrada na Figura 34. Tal como na análise referente à abordagem de realimentação sob perspectiva do limiar d_{thr} nas bases, a F-Measure assume um comportamento similar ao da acurácia e especificidade. Esse critério de avaliação se mantém mesmo sob variação do limiar τ_{CD}. No entanto, há um crescimento proporcional ao aumento de τ_{CD}.

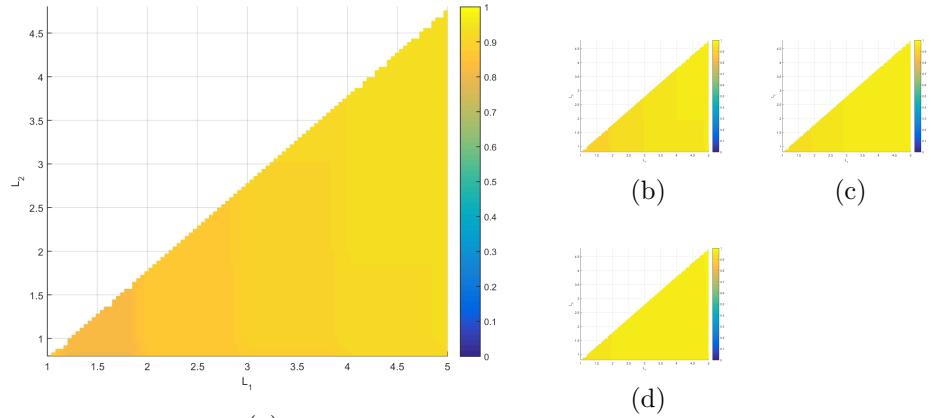


Figura 32 – Sensibilidade do sistema proposto com na base NP90-*Morning* com limiar: (a) $\tau_{CD} = 1{,}8$, (b) $\tau_{CD} = 2{,}8$, (c) $\tau_{CD} = 3{,}8$, e (d) $\tau_{CD} = 4{,}8$.

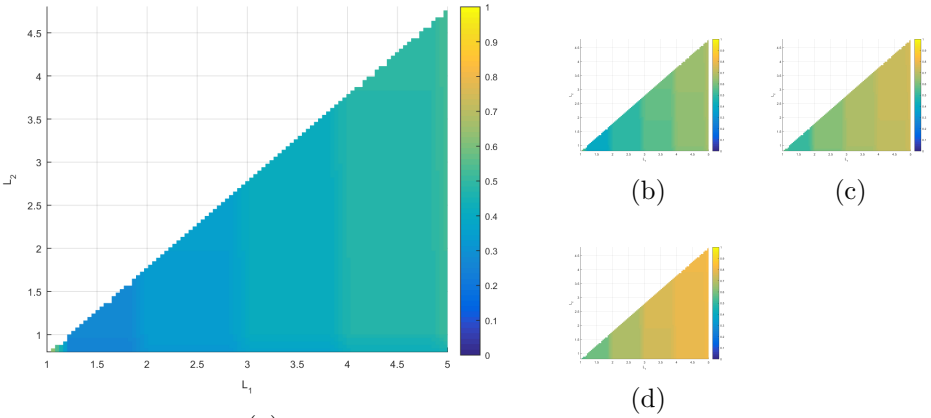


Figura 33 – Especificidade do sistema proposto com na base NP90-*Morning* com limiar: (a) $\tau_{CD} = 1{,}8$, (b) $\tau_{CD} = 2{,}8$, (c) $\tau_{CD} = 3{,}8$, e (d) $\tau_{CD} = 4{,}8$.

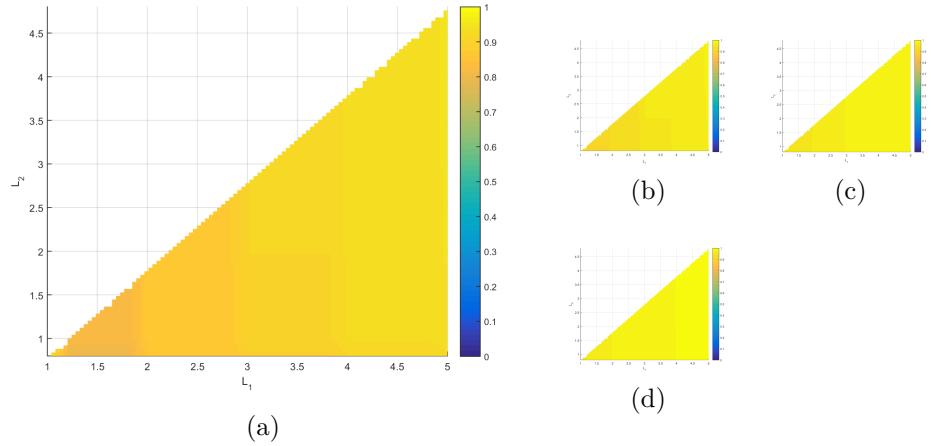


Figura 34 – F-Measure do sistema proposto com na base NP90-*Morning* com limiar: (a) $\tau_{CD} = 1{,}8$, (b) $\tau_{CD} = 2{,}8$, (c) $\tau_{CD} = 3{,}8$, e (d) $\tau_{CD} = 4{,}8$.

3.2.5 Discussão dos resultados

Neste experimento são expostos três testes. O primeiro mostra o comportamento do método de Zivkovic nas bases NP90. Embora esse método apresente uma acurácia satisfatória para o problema de segmentação de veículos, os melhores resultados são obtidos utilizando parâmetros diferentes. Assim, não é possível afirmar a repetibilidade do experimento para outra base e nem estimar a faixa de horários que cada configuração de parâmetros deve ser usada. Sob análise da acurácia, o método de Zivkovic apresenta grandes dificuldades na solução do problema.

A acurácia não é o único critério de análise do método de Zivkovic. Essa métrica não fornece informações sobre quais dificuldades esse método específico apresenta. Deve-se, portanto, analisar as outras métricas para discutir este método como solução do problema abordado. Embora a especificidade apresente resultados baixos, ela exibe uma região ótima comum em todas as bases do experimento. Entretanto, somente um pequeno grupo restrito de limiares dessa região ótima apresenta bons resultados de sensibilidade. Isso torna árdua a tarefa de seleção de parâmetros para o método de Zivkovic (2004), devendo frisar que essa região não é a região ótima da acurácia.

O sistema proposto não apresenta, para vários casos, uma acurácia superior. No entanto, o comportamento é comum a todas as bases do experimento o que torna a escolha do limiar trivial e ótima para as três bases. A acurácia do sistema proposto somente atinge valores superiores ao método de Zivkovic para limiares $\tau_{CD} \geq 3{,}8$. Observando os maiores resultados para as taxas de sensibilidade e especificidade, é perceptível que eles são próximos ao método de Zivkovic. A diferença é a disposição dessas taxas, no método de Zivkovic o aumento da especificidade resulta em uma queda significativa da sensibilidade, enquanto o sistema proposto mantém a taxa de sensibilidade estável para a segmentação dos veículos.

Sob o ponto de vista de reprodutibilidade, o sistema proposto é mais adequado para ambientes não-controlados. Sua previsibilidade facilita a seleção de limiares para o problema da segmentação de veículos. O aumento da sensibilidade não ocorreu como previsto. Isso se deve ao fato do padrão de determinados veículos serem muito próximos ao do plano de fundo e o sistema de descrição de dados de Zivkovic. A abordagem de realimentação utilizada não é suficiente para classificar corretamente esses padrões.

3.3 Experimento 3

Como exposto pelo experimento anterior na Seção 3.2, a abordagem de realimentação utilizada não está apta para classificar corretamente as regiões de intersecção dos padrões de plano de fundo e primeiro plano. Este experimento tem como foco a aplicação de outras abordagens de limiarização na classificação dos *pixels*. Os métodos de limiarização

utilizados neste experimento são baseados em entropia, *clustering* e disposição espacial.

Outro teste do experimento é analisar a classificação individual e conjunta dos *pixels* para cada atributo: distância Mahalanobis (M), distorção de cromaticidade (C), distorção de brilho (L) e a segmentação (MCL) pela Equação 2.4. Dessa forma, pode-se observar a estabilidade de resultados subsequentes para regiões específicas. As bases NP90 são formadas por imagens agrupadas cinco a cinco de veículos específicos ao longo de *frames* consecutivos.

3.3.1 Metodologia do experimento

As métricas de avaliação utilizadas neste experimento limitam-se à acurácia, sensibilidade e especificidade. Embora a F-Measure seja útil para comparar e determinar os acertos relativos às classes, o objetivo deste experimento é verificar se os métodos de limiarização abordados melhoram a classificação do primeiro plano. Essa classificação, determinada pela sensibilidade, deve ser maximizada preservando a taxa de especificidade obtida pelo sistema proposto.

3.3.2 Resultados da limiarização baseada em entropia

Dentre os métodos de limiarização utilizados neste experimento, a primeira categoria a ser analisada é a formada por métodos baseados na entropia. A acurácia dos métodos de Kapur, Sahoo e Wong (1985) e de Yen, Chang e Chang (1995), são mostrados na Figura 35. É perceptível que a componente de distorção de brilho é a que resulta em melhores taxas de acurácia. Esse comportamento é similar em ambos os métodos, no entanto, o método de Yen, Chang e Chang (1995) apresenta acurácias mais baixas utilizando os outros atributos da classificação.

A sensibilidade dos métodos baseados em entropia é alta. Isso pode ser observado na Figura 36 onde são observadas também instabilidades na classificação do primeiro plano na base NP90-*Afternoon*, e parcialmente na base NP90-*Midday*. Dentre essas instabilidades, destaca-se a distância Mahalanobis do modelo GMM. Embora esse atributo tenha resultados baixos para a classificação do primeiro plano. A classificação do plano de fundo é alta utilizando essa distância. Esses resultados podem ser comparados ao da Figura 37. Enquanto a distorção de brilho resulta em altas sensibilidades, esse atributo resulta também em baixos valores de especificidade.

3.3.3 Resultados da limiarização baseada em agrupamento

A próxima categoria deste experimento é a baseada em agrupamento. Esses métodos de limiarização são similares ao da entropia, no entanto, são utilizadas informações de variância das classes. Os métodos utilizados foram o de Otsu (1979) e o de Kittler (SEZAN,

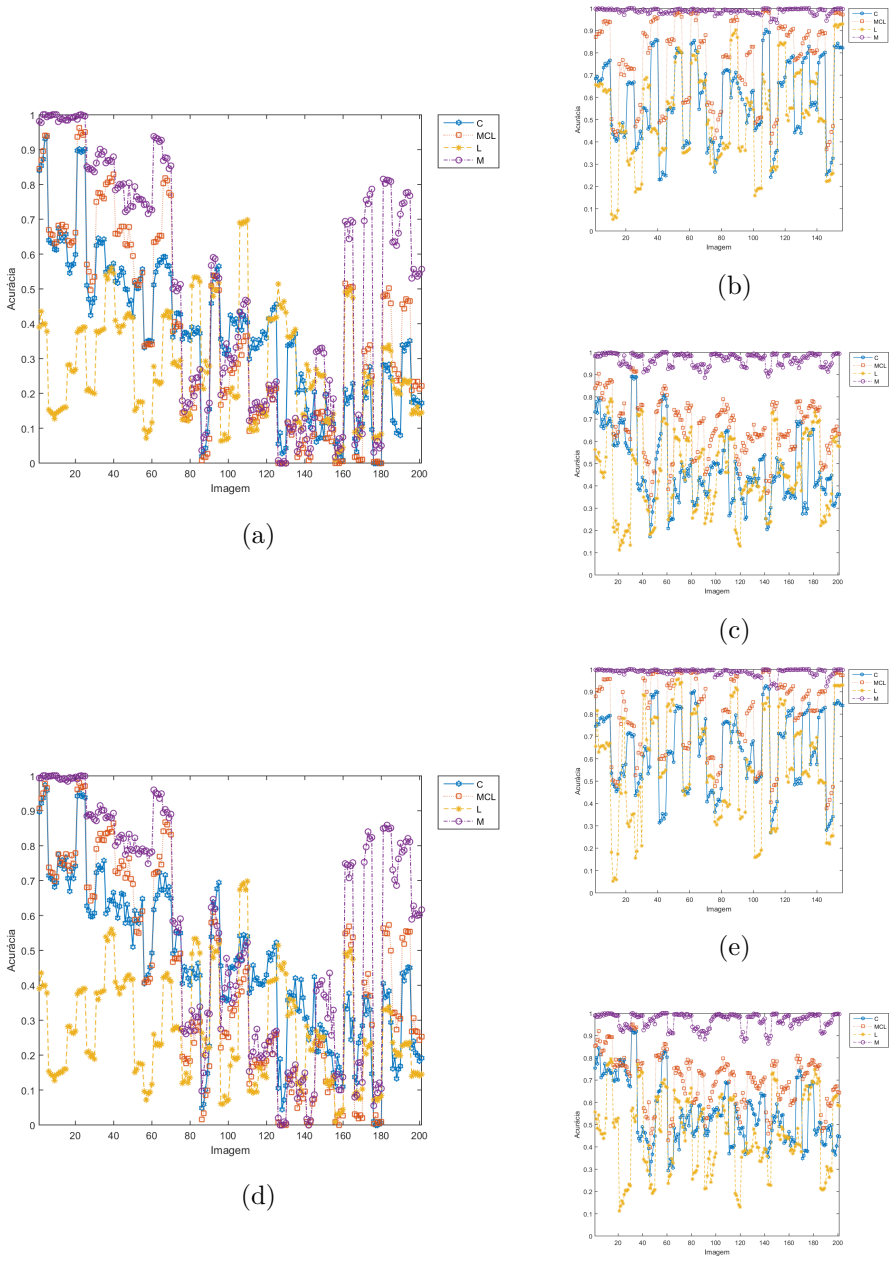


Figura 35 – Acurácia do método de Kapur para as bases: (a) NP90-*Morning*, (b) NP90-*Midday*, e (c) NP90-*Afternoon*. Acurácia do método de Yen para as bases: (d) NP90-*Morning*, (e) NP90-*Midday*, e (f) NP90-*Afternoon*.

1990a). A acurácia destes dois métodos é mostrada na Figura 38. Embora o atributo com melhor acurácia seja a distorção de cromaticidade, observa-se uma instabilidade, e veículos diferentes apresentam resultados diferentes. Dentre estes dois métodos, a limiarização de Otsu (1979) é a que apresenta as melhores acurácias no experimento.

Embora a acurácia da limiarização por agrupamento não apresente resultados maiores do que a limiarização por entropia, a limiarização de Otsu (1979), especificamente, é capaz de maximizar os acertos relativos às classes. Isso pode ser observado na Figura

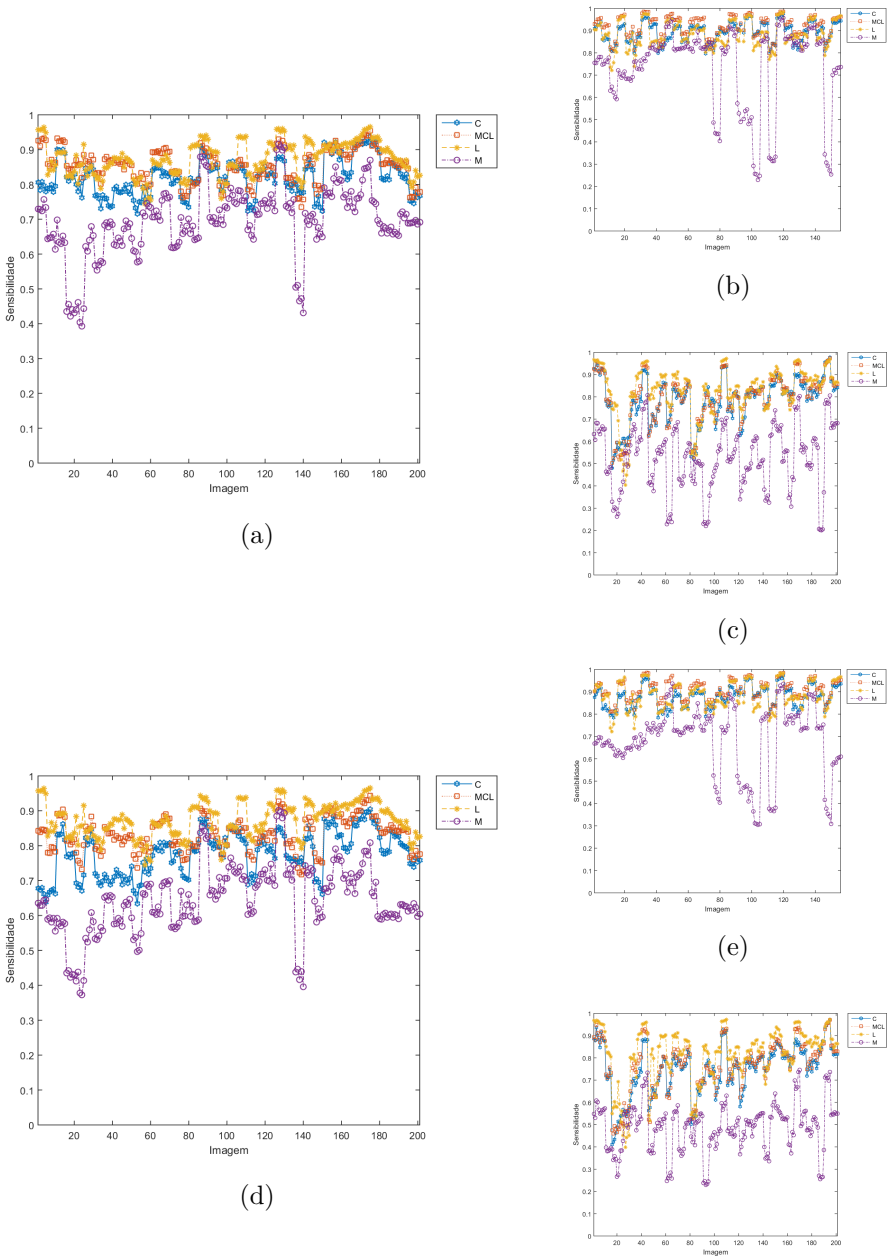


Figura 36 – Sensibilidade do método de Kapur para as bases: (a) NP90-*Morning*, (b) NP90-*Midday*, e (c) NP90-*Afternoon*. Sensibilidade do método de Yen para as bases: (d) NP90-*Morning*, (e) NP90-*Midday*, e (f) NP90-*Afternoon*.

39 e na Figura 40. A especificidade desses métodos é alta e a sensibilidade se mantém estável, com exceção da base NP90-*Afternoon* e em alguns veículos das outras bases do experimento. Assim, o método não é apto a identificar as mudanças de iluminação mais intensas e os reflexos desses veículos.

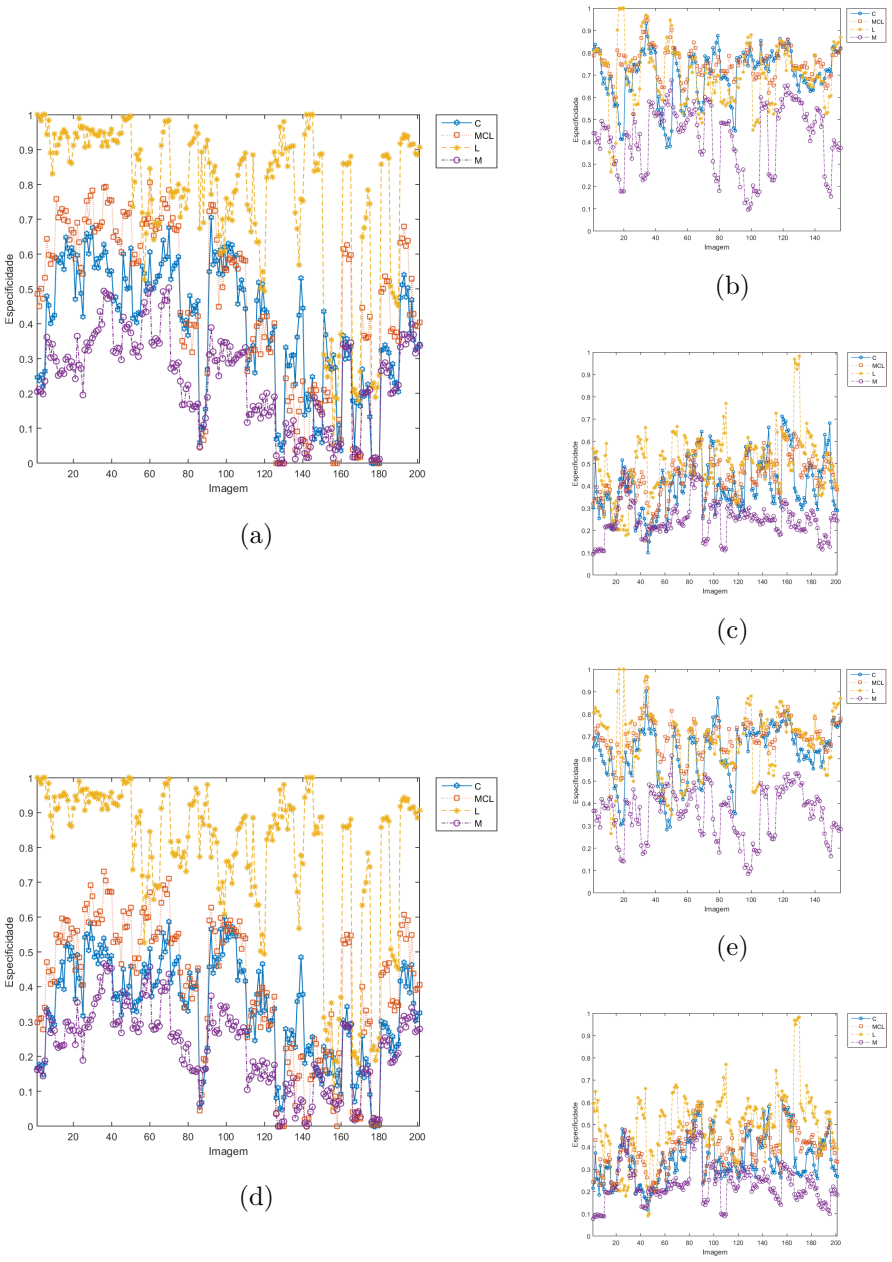


Figura 37 – Especificidade do método de Kapur para as bases: (a) NP90-*Morning*, (b) NP90-*Midday*, e (c) NP90-*Afternoon*. Especificidade do método de Yen para as bases: (d) NP90-*Morning*, (e) NP90-*Midday*, e (f) NP90-*Afternoon*.

3.3.4 Resultados da limiarização por abordagens espaciais

A limiarização da abordagem espacial de Pal e Pal (1989) baseia-se em métodos de entropia. Entretanto, ela analisa a disposição espacial dos valores de intensidade dos *pixels*. A acurácia deste método de limiarização é mostrada na Figura 41. Assim como nos métodos de agrupamento, as acurácias obtidas por este método não foram altas. Uma instabilidade é vista nos resultados independente dos atributos utilizados na limiarização. Essa oscilação da acurácia é mais perceptível na base NP90-*Midday* e ela decai ainda mais

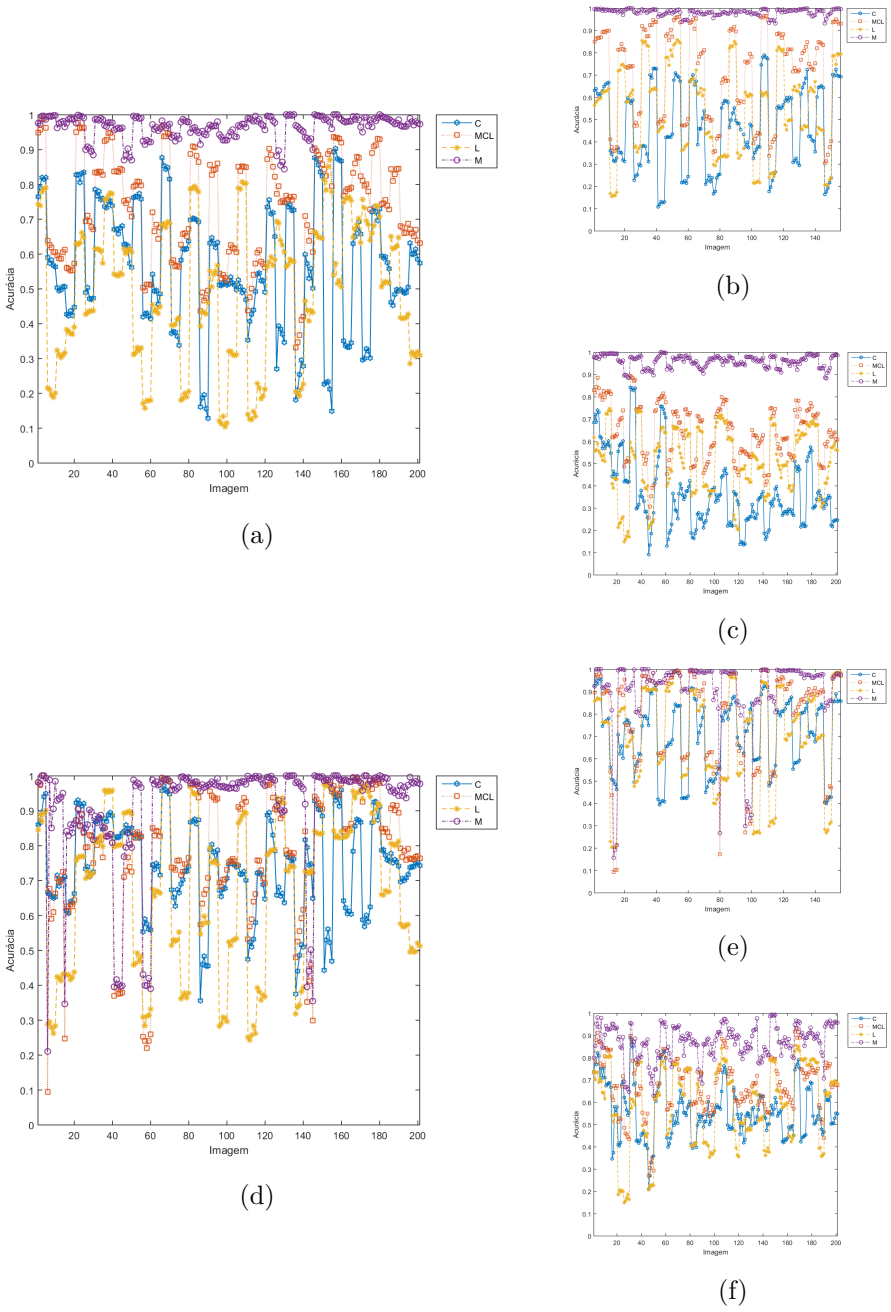


Figura 38 – Acurácia do método de Otsu para as bases: (a) NP90-*Morning*, (b) NP90-*Midday*, e (c) NP90-*Afternoon*. Acurácia do método de Kittler para as bases: (d) NP90-*Morning*, (e) NP90-*Midday*, e (f) NP90-*Afternoon*.

na base NP90-*Afternoon*.

A sensibilidade do método de Pal e Pal (1989) mostra resultados similares para as duas abordagens de limiarização desse método. Os resultados dessa métrica de avaliação são mostrados na Figura 42. Ao passo que a distância Mahalanobis do modelo paramétrico apresenta instabilidades nos resultados da sensibilidade, os atributos de distorções foram relevantes para a classificação do primeiro plano. Essas instabilidades se iniciam na base NP90-*Midday* e permanecem ao longo da base NP90-*Afternoon*.

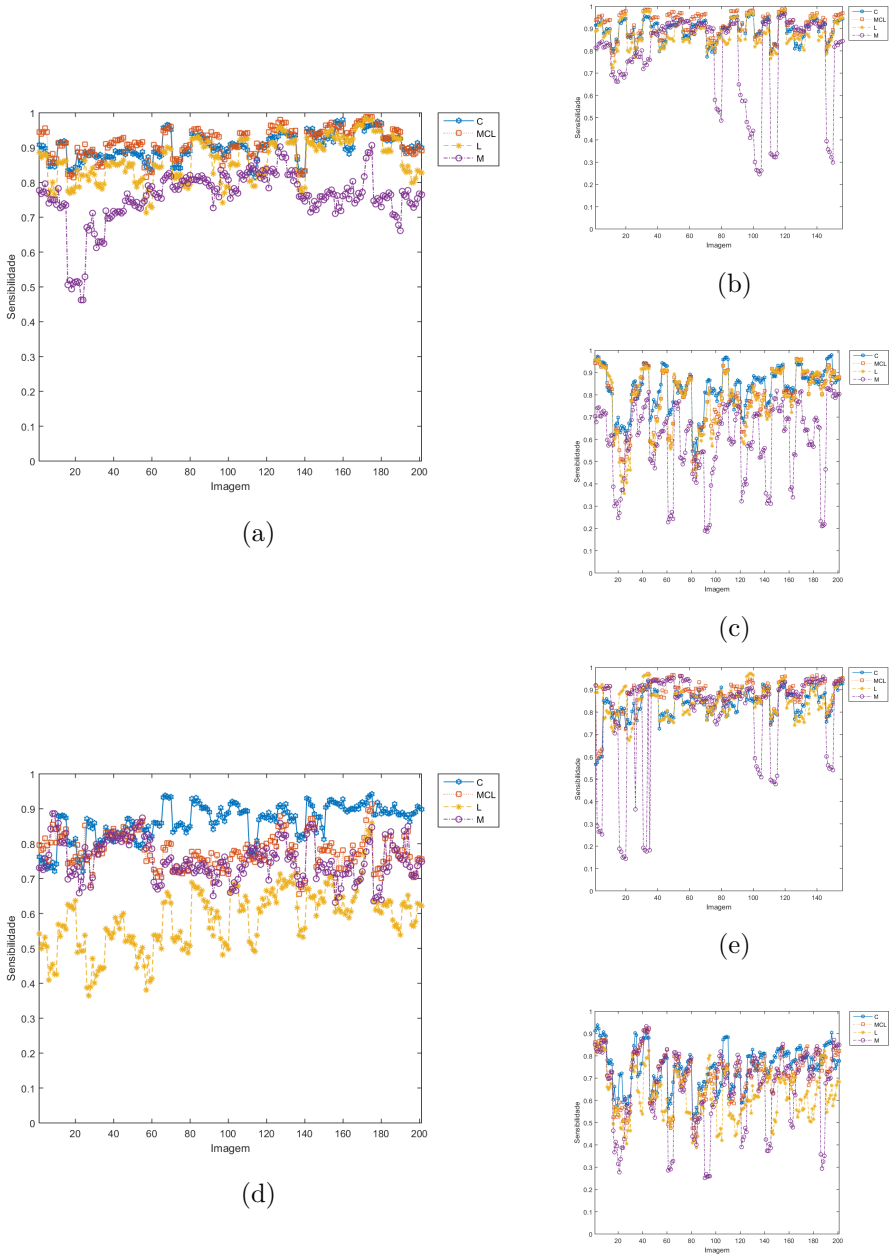


Figura 39 – Sensibilidade do método de Otsu para as bases: (a) NP90-*Morning*, (b) NP90-*Midday*, e (c) NP90-*Afternoon*. Sensibilidade do método de Kittler para as bases: (d) NP90-*Morning*, (e) NP90-*Midday*, e (f) NP90-*Afternoon*.

Embora a classificação do primeiro plano apresente falhas utilizando a distância do modelo GMM, observa-se um comportamento inverso na classificação do plano de fundo. Os resultados da especificidade são mostrados na Figura 43. Enquanto as distorções se mostram instáveis e pouco relevantes, as duas abordagens classificam o plano de fundo com precisão utilizando a distância Mahalanobis.

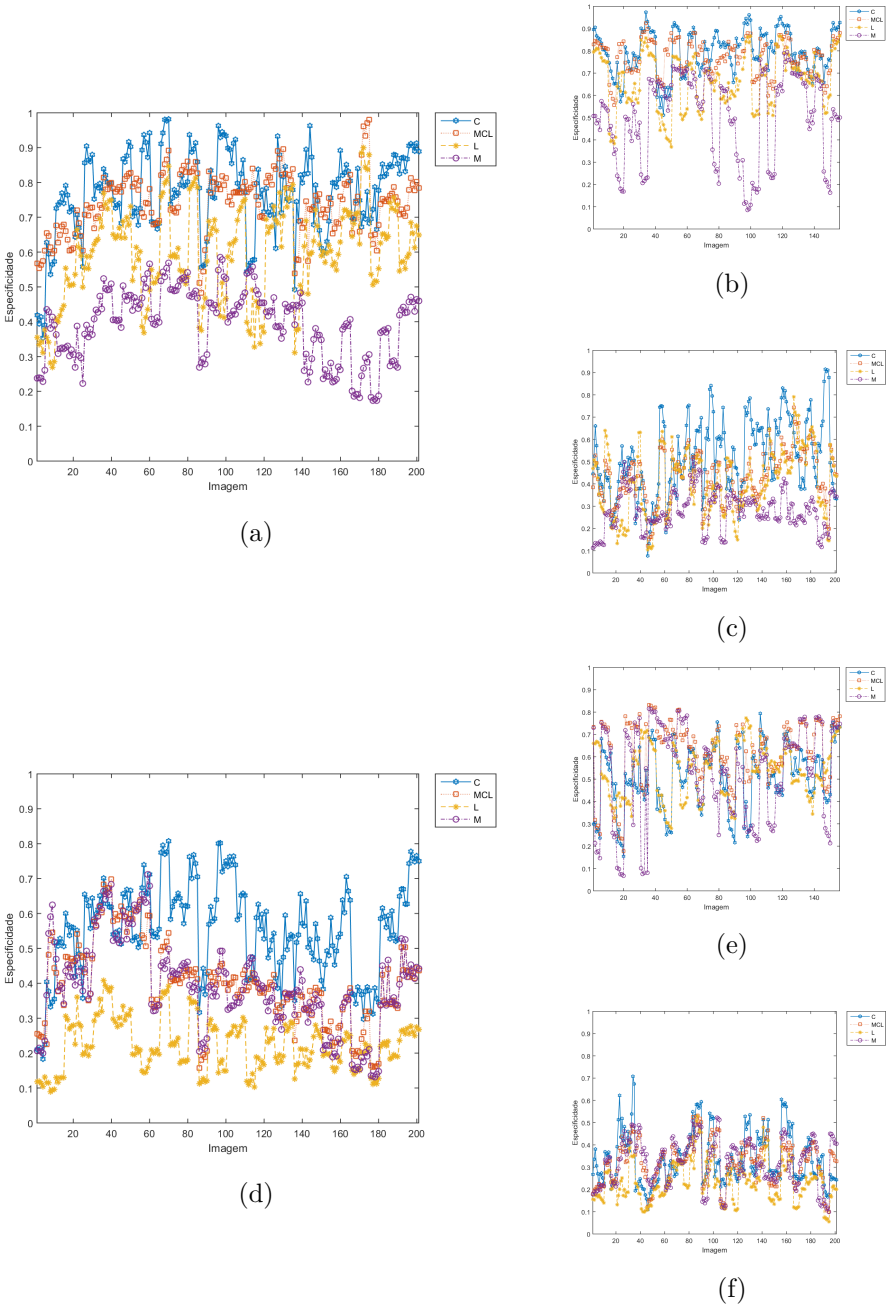


Figura 40 – Especificidade do método de Otsu para as bases: (a) NP90-*Morning*, (b) NP90-*Midday*, e (c) NP90-*Afternoon*. Especificidade do método de Kittler para as bases: (d) NP90-*Morning*, (e) NP90-*Midday*, e (f) NP90-*Afternoon*.

3.3.5 Discussão do experimento

Os métodos de limiarização testados, com exceção dos métodos baseados em agrupamento, não conseguem equilibrar os acertos de cada classe. O método espacial de Pal e Pal, por exemplo, é mais propenso a classificar as distâncias Mahalanobis do modelo GMM como plano de fundo e as distorções como primeiro plano. Os métodos de Kapur, e Yen, embora mostrem resultados mais equilibrados do que a abordagem espacial, apresentam o mesmo comportamento.

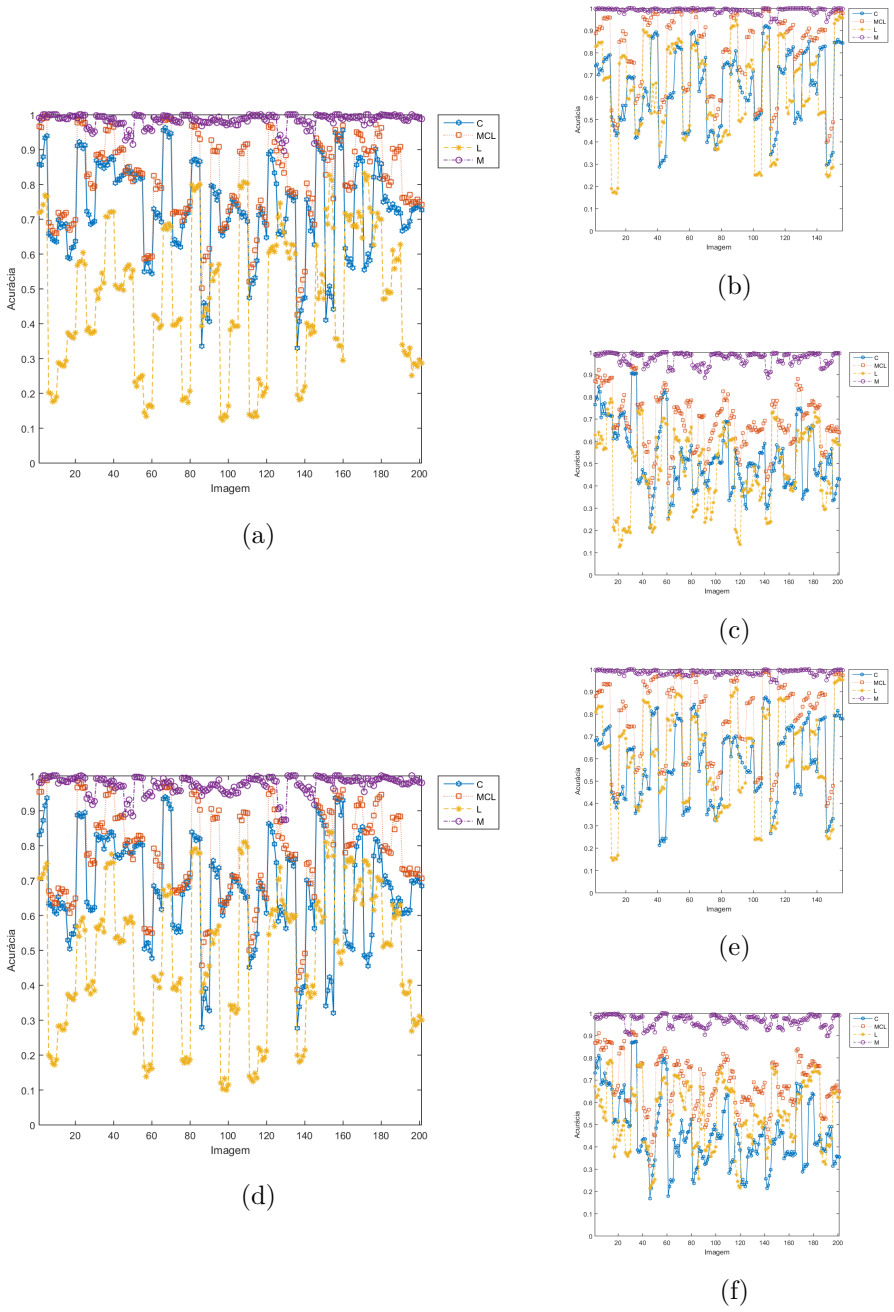


Figura 41 – Acurácia do método de Pal e Pal (Abordagem 1) para as bases: (a) NP90-*Morning*, (b) NP90-*Midday*, e (c) NP90-*Afternoon*. Acurácia do método de Pal e Pal (Abordagem 2) para as bases: (d) NP90-*Morning*, (e) NP90-*Midday*, e (f) NP90-*Afternoon*.

A limiarização por agrupamento apresenta acurácias mais baixas, no entanto, dentre as três abordagem de limiarização analisadas no experimento, é a que melhor mantém em equilíbrio as taxas de sensibilidade e especificidade nas bases. O método de Otsu (1979) é o que apresentou melhor estabilidade e equilíbrio da classificação para a base NP90-*Afternoon*.

Sob a perspectiva dos atributos, a distorção de brilho é uma característica relevante

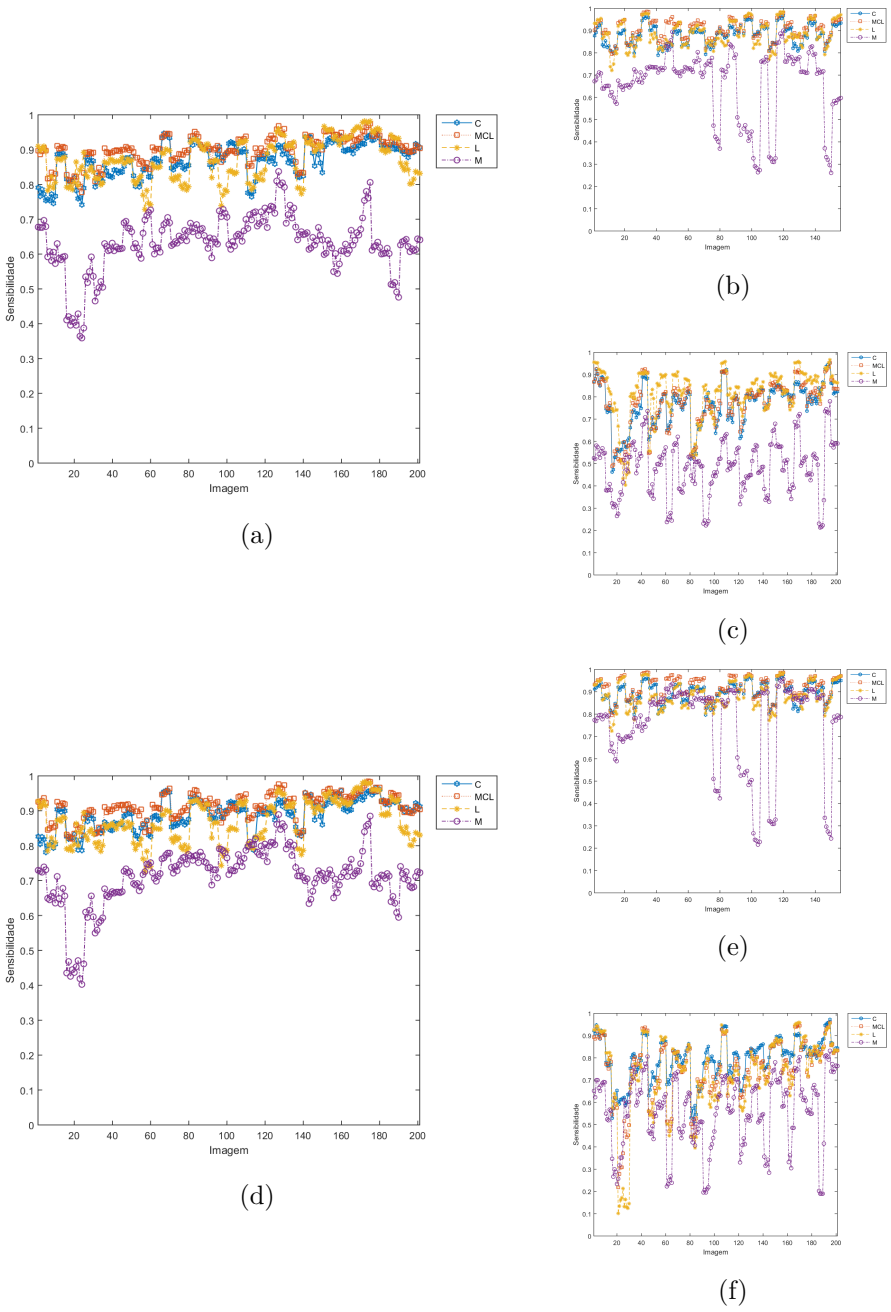


Figura 42 – Sensibilidade do método de Pal e Pal (Abordagem 1) para as bases: (a) NP90-*Morning*, (b) NP90-*Midday*, e (c) NP90-*Afternoon*. Sensibilidade do método de Pal e Pal (Abordagem 2) para as bases: (d) NP90-*Morning*, (e) NP90-*Midday*, e (f) NP90-*Afternoon*.

para a classificação do plano de fundo. No entanto, nas regiões próximas aos veículos, essa característica é determinante para a classificação do primeiro plano. Em virtude das mudanças de iluminação, os valores da distorção de brilho são altos nas proximidades dos veículos. Dessa forma, o uso desse atributo para extração de fundo especificamente nas regiões dos veículos não traz classificações precisas. Considerando a segmentação dos veículos, a deficiência da distorção de brilho se dá nas regiões dos para-brisas. Os baixos valores nessas regiões são responsáveis pela cavidades internas que o sistema proposto

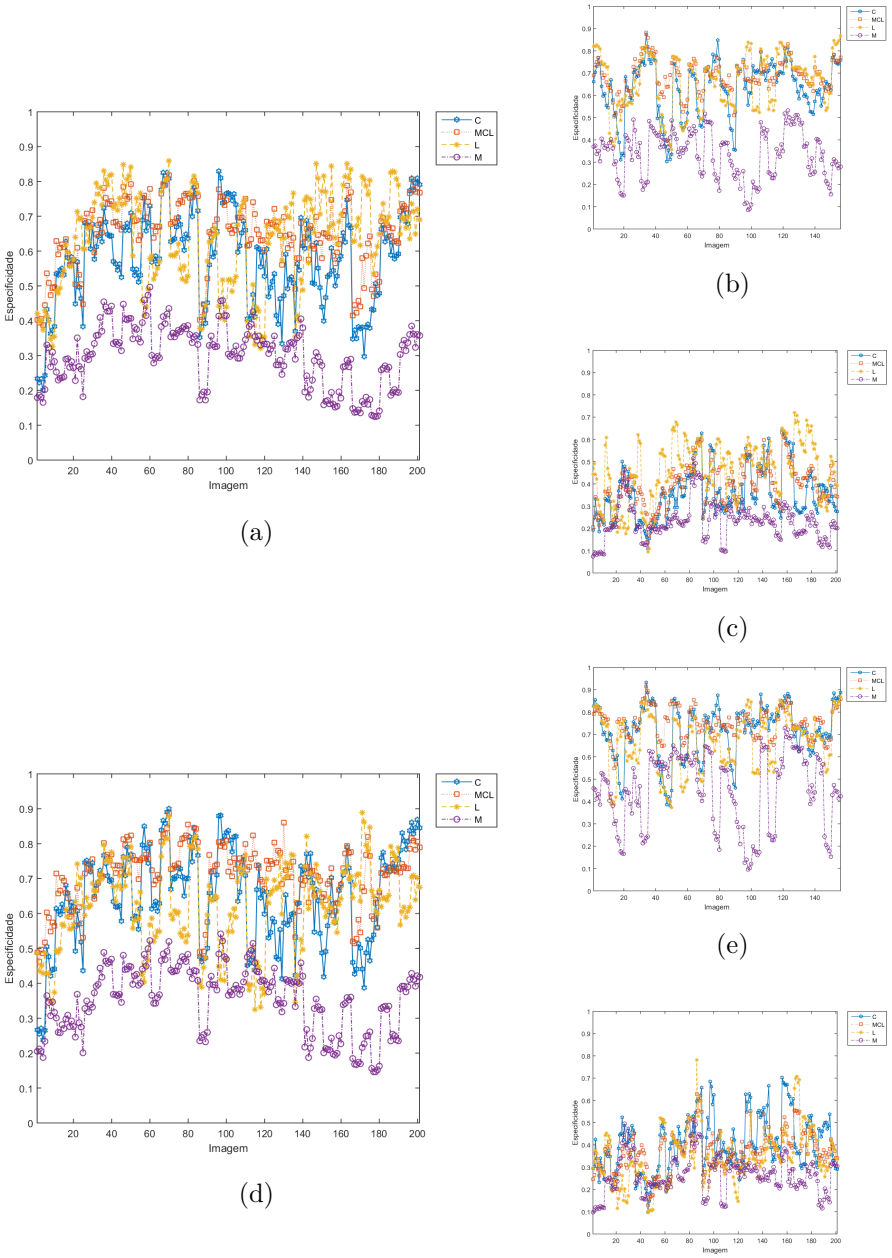


Figura 43 – Especificidade do método de Pal e Pal (Abordagem 1) para as bases: (a) NP90-*Morning*, (b) NP90-*Midday*, e (c) NP90-*Afternoon*. Especificidade do método de Pal e Pal (Abordagem 2) para as bases: (d) NP90-*Morning*, (e) NP90-*Midday*, e (f) NP90-*Afternoon*.

corrige na etapa de filtragem.

A distância Mahalanobis é uma boa representação do modelo probabilístico. No entanto, a sutis diferenças entre as classes não são muito perceptíveis nessa representação. Isso resulta em falhas de segmentação, principalmente nas regiões de sombras e em veículos escuros. No experimento em questão, o desbalanceamento entre as classes resulta na seleção de um limiar que favorece a classificação do plano de fundo. Outra deficiência da distância

Mahalanobis é que ele intensifica ruídos e tende a extrair regiões que não representam veículos.

A distorção da cor assume um comportamento intermediário entre distorção de brilho e a distância Mahalanobis. Assim como a distância Mahalanobis da estimação do GMM, a distorção de cor apresenta muitos ruídos na cena. Na região dos veículos, ela apresenta falhas e descontinuidades. No entanto, as sombras e mudanças de iluminação podem ser identificadas.

3.4 Considerações finais

Este capítulo mostra o comportamento geral para o método de Zivkovic (2004) e sistema proposto em uma base simples. Definido a viabilidade e eficácia da proposta. Os resultados do método de Zivkovic (2004), do sistema proposto e de outros métodos de limiarização são analisados detalhadamente sob perspectiva dos horários do dia em diferentes bases. Esses experimentos permitem verificar a proposta como uma solução real para o problema.

Embora a abordagem de realimentação proposta atue somente em um dos atributos, esse capítulo realiza também uma análise da influência de cada atributo da classificação. Esse estudo é feito utilizando métodos de limiarização de diferentes categorias e abordagens encontradas na literatura. Assim avaliamos se a abordagem de realimentação local nas regiões do veículos melhora a segmentação dos mesmos.

O Experimento 1 teve como resultados as publicações (LIMA; REIS; AIRES, 2015; LIMA; AIRES; REIS, 2015). Trabalhos sobre os outros experimentos já foram submetidos a conferências e periódicos e estão sob revisão dos mesmos.

Conclusões e Continuidade da pesquisa

A abordagem de realimentação proposta se mostrou eficaz para aprimorar a etapa de classificação no problema de segmentação de veículos. O sistema proposto obteve uma acurácia maior e reduziu o problema da fragmentação da segmentação na região do veículo. Entretanto, ainda é necessário a realização de testes qualitativos. Neste capítulo são discutidas algumas das limitações do método proposto e experimentos que podem ser realizados.

Conclusões

O método proposto, que em essência é uma extensão do método de Zivkovic (2004), apresenta bons resultados para o problema abordado. As métricas de avaliação utilizados neste trabalho tem enfoque quantitativo, contudo, existem outras medidas qualitativas que podem ser utilizadas na comparação dos métodos.

A amostragem dos dados para somente algumas regiões é realizada em virtude da quantidade de saídas dos testes. A análise por meio dessa amostragem não traz problemas à comparação entre o método de Zivkovic (2004) e o sistema proposto, pois eles apresentam o mesmo comportamento nas regiões onde não são detectados os veículos. A abordagem de realimentação atua como uma extensão do método de Zivkovic nas regiões dos veículos. No entanto, essa amostragem não permite uma análise completa com os outros métodos de limiarização.

Em problemas de segmentação de veículos, espera-se que o sistema seja apto a funcionar em várias situações climáticas e horários diferentes. Ao longo do dia a percepção de cores muda e os fatores de variação em ambientes não-controlados muitas vezes são imprevisíveis. Isso dificulta a segmentação dos veículos e a análise dos métodos. Apesar da base NP90-*Evening* não ter sido incluída nos experimentos, as bases utilizadas foram suficientes para estimar um comportamento de como cada método atua sob os fatores de variação no decorrer do dia. O comportamento do sistema proposto se manteve sob essas variações o que permite a seleção de um limiar único, o que não ocorreu no método de Zivkovic (2004).

Nos resultados do sistema proposto não foram apresentadas situações de oclusão. Embora a posição da câmera utilizada nos bases dificulte as oclusões, essa situação ocorre com frequência no problema. Sendo assim, é vital verificar se as mudanças de iluminação entre veículos são detectadas. Caso contrário, é necessário incluir uma etapa de tratamento de oclusão na etapa de filtragem. Contudo, o aumento de complexidade do processo de

filtragem deve respeitar a limitação do tempo de processamento.

Os experimentos realizados atestam que a abordagem de realimentação na classificação após a correção da segmentação é eficaz. Essa correção, realizada pela etapa de filtragem no sistema proposto, influencia diretamente o aumento da acurácia do sistema. Testes preliminares indicaram que filtros morfológicos e estatísticos são opções viáveis, porém, há outros métodos na literatura que não foram testados. Esses métodos podem melhorar os resultados dessa etapa que influencia diretamente o resultado do sistema proposto.

É importante observar que o método de Zivkovic (2004) é uma abordagem híbrida que utiliza o método GMM de Stauffer e Grimson (1999) e o modelo não-paramétrico de Horprasert, Harwood e Davis (1999). O limiar dinâmico do sistema proposto, que é estimado na etapa de realimentação, atua somente na classificação do método GMM. Entretanto, os atributos de distorção de cor e brilho da abordagem não-paramétrica também atuam na classificação. Consequentemente, a realimentação pode ser estendida para os limiares desses atributos.

Em virtude da limitação de tempo e escopo, os processos cognitivos não foram incorporados ao sistema e nem foram discutidos nesse trabalho. Porém, esses processos são fundamentais nos ITSs, sendo responsáveis pela função de reconhecimento dos veículos da cena. O desempenho do sistema proposto atua diretamente na acurácia em virtude da abordagem de realimentação utilizada. Desse modo, as execuções dos processos de segmentação e cognitivo devem ser assíncronas e sem interferências mútuas.

Continuidade da pesquisa

O sistema proposto apresentou resultados satisfatórios e se mostrou uma opção viável e melhorada do método de Zivkovic (2004). O sistema proposto pode ser aprimorado analisando outras representações de dados. Estendendo-o aos outros atributos e avaliando novos processos de filtragem. Cada etapa do sistema deve ser exaustivamente testada, novos experimentos podem ser focados na análise da etapa de filtragem e abordagens de estimação do limiar para a classificação.

Os atributos de distância Mahalanobis do modelo, distorção de cromaticidade e distorção de brilho apresentam uma ampla região de incerteza na classificação e, em determinados contextos, não fornecem opções precisas de representação do problema. A mudança do sistema de representação pode ser explorada, assim como os sistemas Fuzzy (COX, 1994) e máquinas de Comitê (HAYKIN, 1998) para explorar a propensão de cada atributo com as classes do problema.

A etapa de realimentação se baseia em uma limiarização analisando características

dos objetos, no caso veículos. Embora esse trabalho tenha incluído experimentações de outras abordagens de limiarização, há outros métodos promissores na literatura como o *Statistical Region Merging* (SRM). Ele analisa o contexto e padrão da região abordagem e pode reduzir a fragmentação do objeto em várias regiões.

Referências

ABUTALEB, A. S. Automatic thresholding of gray-level pictures using two-dimensional entropy. *Computer Vision, Graphics, and Image Processing*, v. 47, n. 1, p. 22–32, 1989. Disponível em: <http://dblp.uni-trier.de/db/journals/cvgip/cvgip47.html#Abutaleb89>. Citado na página 27.

ALVAREZ, S. et al. Monocular target detection on transport infrastructures with dynamic and variable environments. In: *Intelligent Transportation Systems (ITSC), 2012 15th International IEEE Conference on*. Anchorage, AK: IEEE, 2012. p. 61–66. ISSN 2153-0009. Citado na página 6.

Associação Brasileira das Empresas de Sistemas Eletrônicos de Segurança (ABESE). *A cada dia, SP tem dez novas câmeras*. 2012. Disponível em: <http://www.abese.org.br/clipping11-06-2012/>. Acesso em: 19 set. 2015. Citado na página 1.

ATKINS, P.; PAULA, J. D. *Atkins' Physical Chemistry*. Oxford University Press, 2002. ISBN 9780198792857. Disponível em: <https://books.google.com.br/books?id=klOUbwAACAAJ>. Citado na página 23.

BEKMAN, O.; NETO, P. de O. C. *Análise estatística da decisão*. Editora E. Blücher, 1980. ISBN 9788521202141. Disponível em: <https://books.google.com.br/books?id=KwxsPgAACAAJ>. Citado na página 14.

BLACKMAN, S. Multiple hypothesis tracking for multiple target tracking. *Aerospace and Electronic Systems Magazine, IEEE*, v. 19, n. 1, p. 5–18, Jan 2004. ISSN 0885-8985. Citado na página 6.

BRADSKI, G. *Dr. Dobb's Journal of Software Tools*, 2000. Citado 3 vezes nas páginas 5, 7 e 35.

CASELLA, G.; BERGER, R. *Statistical Inference*. Brooks/Cole Publishing Company, 1990. (Duxbury advanced series). ISBN 9780534119584. Disponível em: <https://books.google.com.br/books?id=nA_vAAAAMAAJ>. Citado na página 17.

CHANG, C. I. et al. Survey and comparative analysis of entropy and relative entropy thresholding techniques. *Vision, Image and Signal Processing, IEE Proceedings* -, v. 153, n. 6, p. 837–850, 2006. Disponível em: <http://dx.doi.org/10.1049/ip-vis:20050032>. Citado 2 vezes nas páginas 3 e 22.

COX, E. *The Fuzzy Systems Handbook: A Practitioner's Guide to Building, Using, and Maintaining Fuzzy Systems*. San Diego, CA, USA: Academic Press Professional, Inc., 1994. ISBN 0-12-194270-8. Citado na página 74.

DOWNEY, A. B. *Think Bayes: Bayesian Statistics Made Simple*. Needham, MA, USA: Green Tea Press, 2012. Disponível em: <http://www.greenteapress.com/thinkbayes/>. Citado 3 vezes nas páginas 11, 12 e 13.

FARO, A.; GIORDANO, D.; SPAMPINATO, C. Adaptive background modeling integrated with luminosity sensors and occlusion processing for reliable vehicle detection.

Intelligent Transportation Systems, IEEE Transactions on, v. 12, n. 4, p. 1398–1412, Dec 2011. ISSN 1524-9050. Citado na página 6.

FAWCETT, T. An introduction to {ROC} analysis. *Pattern Recognition Letters*, v. 27, n. 8, p. 861 – 874, 2006. ISSN 0167-8655. {ROC} Analysis in Pattern Recognition. Disponível em: <http://www.sciencedirect.com/science/article/pii/S016786550500303X>. Citado na página 39.

GALLO, G.; SPINELLO, S. Thresholding and fast iso-contour extraction with fuzzy arithmetic. *Pattern Recognition Letters*, v. 21, n. 1, p. 31–44, 2000. Disponível em: <http://dblp.uni-trier.de/db/journals/prl/prl21.html#GalloS00>. Citado na página 29.

GONZALEZ, R. C.; WOODS, R. E. *Digital Image Processing (3rd Edition)*. Upper Saddle River, NJ, USA: Prentice-Hall, Inc., 2006. ISBN 013168728X. Citado 4 vezes nas páginas 2, 22, 23 e 35.

HAYKIN, S. *Neural Networks: A Comprehensive Foundation*. 2nd. ed. Upper Saddle River, NJ, USA: Prentice Hall PTR, 1998. ISBN 0132733501. Citado na página 74.

HORPRASERT, T.; HARWOOD, D.; DAVIS, L. S. A statistical approach for real-time robust background subtraction and shadow detection. In: . [S.l.: s.n.], 1999. p. 1–19. Citado 9 vezes nas páginas 2, 4, 5, 7, 14, 20, 33, 34 e 74.

HOUGH, P. *Method and Means for Recognizing Complex Patterns*. [S.l.]: United States Patent Office, 1962. U.S. Patent 3.069.654. Citado na página 6.

Instituto Brasileiro de Geografia e Estatística (IBGE). *Brasil em números*. Instituto Brasileiro de Geografia e Estatística, 2009. ISSN 1808-1983. Disponível em: <http://biblioteca.ibge.gov.br/visualizacao/monografias/GEBIS%20-%20RJ/brasilnumeros/Brasil_numeros_v17_2009.pdf>. Citado na página 1.

KALMAN, R. E. A new approach to linear filtering and prediction problems. *ASME Journal of Basic Engineering*, 1960. Citado na página 5.

KAPUR, J.; SAHOO, P.; WONG, A. A new method for gray-level picture thresholding using the entropy of the histogram. *Computer Vision, Graphics, and Image Processing*, v. 29, n. 3, p. 273 – 285, 1985. ISSN 0734-189X. Disponível em: <http://www.sciencedirect.com/science/article/pii/0734189X85901252>. Citado 3 vezes nas páginas 24, 25 e 62.

KEMOUCHE, M.; AOUF, N. A gaussian mixture based optical flow modeling for object detection. In: *Crime Detection and Prevention (ICDP 2009), 3rd International Conference on*. London: IET, 2009. p. 1–6. Citado 4 vezes nas páginas 3, 5, 6 e 33.

LEI, M. et al. A video-based real-time vehicle counting system using adaptive background method. In: *Signal Image Technology and Internet Based Systems, 2008. SITIS '08. IEEE International Conference on*. Bali: IEEE, 2008. p. 523–528. Citado na página 5.

LEITHOLD, L. *Cálculo com geometria analítica*. Harbra, 1994. ISBN 9788529400945. Disponível em: <https://books.google.com.br/books?id=6csjAAAACAAJ>. Citado na página 17.

LI, C. H.; LEE, C. K. Minimum cross entropy thresholding. *Pattern Recognition*, v. 26, n. 4, p. 617–625, 1993. Disponível em: <http://dblp.uni-trier.de/db/journals/pr/pr26.html#LiL93>. Citado na página 24.

LI, T. et al. Measurement of the Instantaneous Velocity of a Brownian Particle. *Science*, American Association for the Advancement of Science, v. 328, n. 5986, p. 1673–1675, jun. 2010. ISSN 1095-9203. Disponível em: <http://dx.doi.org/10.1126/science.1189403>. Citado na página 12.

LI, Z.; ZHONG, L. An efficient framework for detecting moving objects and structural lanes in video based surveillance systems. In: *Pervasive Computing (JCPC), 2009 Joint Conferences on*. Tamsui, Taipei: IEEE, 2009. p. 355–358. Citado na página 6.

LIE, W.-N. An efficient threshold-evaluation algorithm for image segmentation based on spatial graylevel co-occurrences. *Signal Process.*, Elsevier North-Holland, Inc., Amsterdam, The Netherlands, The Netherlands, v. 33, n. 1, p. 121–126, jul. 1993. ISSN 0165-1684. Disponível em: <http://dx.doi.org/10.1016/0165-1684(93)90083-M>. Citado na página 27.

LIMA, K. A. B.; AIRES, K. R. T.; REIS, F. W. P. D. Adaptive method for segmentation of vehicles through local threshold in the gaussian mixture model. In: *2015 Brazilian Conference on Intelligent Systems (BRACIS)*. [S.l.: s.n.], 2015. p. 204–209. Citado na página 72.

LIMA, K. A. B.; REIS, F. W. P. dos; AIRES, K. R. T. A. Aprimoramento do modelo de mistura de gaussianas para segmentação de veículos utilizando abordagem de limiarização local. In: *XII Simpósio Brasileiro de Automação Inteligente (SBAI)*. [S.l.: s.n.], 2015. Citado na página 72.

LIPSCHUTZ, S. *Probabilidade*. 4ª edição revisada. ed. McGraw-Hill, 1993. (Coleção Schaum). Disponível em: <https://books.google.com.br/books?id=1YRgrgEACAAJ>. Citado 2 vezes nas páginas 6 e 13.

LUO, J.; ZHU, J. Adaptive gaussian mixture model based on feedback mechanism. In: *Computer Design and Applications (ICCDA), 2010 International Conference on*. Qinhuangdao: IEEE, 2010. v. 2, p. V2–177–V2–181. Citado 2 vezes nas páginas 3 e 6.

MAHALANOBIS, P. C. On the generalised distance in statistics. In: *Proceedings National Institute of Science, India*. [s.n.], 1936. v. 2, n. 1, p. 49–55. Disponível em: <http://ir.isical.ac.in/dspace/handle/1/1268>. Citado na página 21.

MARKOV, A. Extension of the Limit Theorems of Probability Theory to a Sum of Variables Connected in a Chain. In: HOWARD, R. (Ed.). *Dynamic Probabilistic Systems (Volume I: Markov Models)*. New York City: John Wiley & Sons, Inc., 1971. cap. Appendix B, p. 552–577. Citado 2 vezes nas páginas 12 e 36.

NIBLACK, W. *An Introduction to Digital Image Processing*. Birkeroed, Denmark, Denmark: Strandberg Publishing Company, 1985. ISBN 87-872-0055-4. Citado na página 29.

OLIVO, J.-C. Automatic threshold selection using the wavelet transform. *CVGIP: Graphical Model and Image Processing*, v. 56, n. 3, p. 205–218, 1994. Disponível em:

<http://dblp.uni-trier.de/db/journals/cvgip/cvgip56.html#Olivo94>. Citado na página 26.

OTSU, N. A thresholding selection method from gray-level histograms. v. 9, n. 1, p. 62–66, 1979. Citado 4 vezes nas páginas 26, 62, 63 e 69.

PAL, N. R. On minimum cross-entropy thresholding. *Pattern Recognition*, v. 29, n. 4, p. 575–580, 1996. Disponível em: <http://dblp.uni-trier.de/db/journals/pr/pr29.html#Pal96>. Citado na página 24.

PAL, N. R.; PAL, S. K. Entropic thresholding. *Signal processing*, Elsevier, v. 16, n. 2, p. 97–108, 1989. Citado 4 vezes nas páginas 27, 28, 65 e 66.

PEDRINI, H.; SCHWARTZ, W. *Análise de imagens digitais: princípios, algoritmos e aplicações*. THOMSON PIONEIRA, 2008. ISBN 9788522105953. Disponível em: <https://books.google.com.br/books?id=13KAPgAACAAJ>. Citado 3 vezes nas páginas 11, 22 e 26.

POISSON, S. D. *Recherches sur la probabilité des jugements en matière criminelle et en matière civile, précédées des règles générales du calcul des probabilités*. Paris: Bachelier, 1837. Citado 2 vezes nas páginas 12 e 13.

Portal da Saúde. *Acidentes de trânsito matam 7.160 pessoas em SP*. 2011. Disponível em: <http://portalsaude.saude.gov.br/index.php/cidadao/principal/agencia-saude/noticias-anteriores-agencia-saude/867-acidentes-de-transito-matam-7-160-pessoas-em-sp>. Acesso em: 19 set. 2015. Citado na página 1.

Portal da Saúde. *Brasil é o quinto país no mundo em mortes por acidentes no trânsito*. 2015. Disponível em: <http://portalsaude.saude.gov.br/index.php/cidadao/principal/agencia-saude/17821-brasil-e-o-quinto-pais-no-mundo-em-mortes-por-acidentes-no-transito>. Acesso em: 19 set. 2015. Citado na página 1.

PRATI, A. et al. Detecting moving shadows: Formulation, algorithms and evaluation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, v. 25, p. 2003, 2003. Citado 5 vezes nas páginas 2, 4, 5, 14 e 20.

PUN, T. A new method for grey-level picture thresholding using the entropy of the histogram. *Signal Processing*, v. 2, n. 3, p. 223–237, jul. 1980. Disponível em: <http://www.sciencedirect.com/science/article/B6V18-48V265B-BC/1/a3401cd10ffc896591be2eab31c597be>. Citado na página 24.

RASMUSSEN, C. E.; WILLIAMS, C. K. I. *Gaussian Processes for Machine Learning (Adaptive Computation and Machine Learning)*. [S.l.]: The MIT Press, 2005. ISBN 026218253X. Citado na página 12.

ROSENFELD, A.; TORRE, P. de la. Histogram concavity analysis as an aid in threshold selection. *IEEE Transactions on Systems, Man, and Cybernetics*, v. 13, n. 2, p. 231–235, 1983. Disponível em: <http://dblp.uni-trier.de/db/journals/tsmc/tsmc13.html#RosenfeldT83>. Citado na página 29.

RUSSELL, S. J.; NORVIG, P. *Artificial Intelligence: A Modern Approach*. 3rd. ed. [S.l.]: Prentice Hall, 2009. Citado 3 vezes nas páginas 11, 12 e 17.

RYAN, T. *Estatística Moderna para Engenharia*. Elsevier Brasil, 2009. ISBN 9788535250886. Disponível em: <https://books.google.com.br/books?id=gWcVKMMz_zgC>. Citado 3 vezes nas páginas 11, 17 e 18.

SAHOO, P. K.; WILKINS, C.; YEAGER, J. Threshold selection using renyi's entropy. *Pattern Recognition*, v. 30, n. 1, p. 71–84, 1997. Disponível em: <http://dblp.uni-trier.de/db/journals/pr/pr30.html#SahooWY97>. Citado na página 24.

SAUVOLA, J. J.; PIETIKäINEN, M. Adaptive document image binarization. *Pattern Recognition*, v. 33, n. 2, p. 225–236, 2000. Disponível em: <http://dblp.uni-trier.de/db/journals/pr/pr33.html#SauvolaP00>. Citado na página 29.

SEZAN, M. I. A peak detection algorithm and its application to histogram-based image data reduction. *Computer Vision, Graphics, and Image Processing*, v. 49, n. 1, p. 36–51, 1990. Disponível em: <http://dblp.uni-trier.de/db/journals/cvgip/cvgip49.html#Sezan90>. Citado 2 vezes nas páginas 26 e 63.

SEZAN, M. I. A peak detection algorithm and its application to histogram-based image data reduction. *Computer Vision, Graphics, and Image Processing*, v. 49, n. 1, p. 36–51, 1990. Disponível em: <http://dblp.uni-trier.de/db/journals/cvgip/cvgip49.html#Sezan90>. Citado na página 29.

SEZGIN, M.; SANKUR, B. Survey over image thresholding techniques and quantitative performance evaluation. *J. Electronic Imaging*, v. 13, n. 1, p. 146–168, 2004. Disponível em: <http://dblp.uni-trier.de/db/journals/jei/jei13.html#SezginS04>. Citado 5 vezes nas páginas 3, 23, 24, 26 e 27.

SHAO, J. et al. Robust visual surveillance based traffic information analysis and forewarning in urban dynamic scenes. In: *Intelligent Vehicles Symposium (IV), 2010 IEEE*. San Diego, CA: IEEE, 2010. p. 813–818. ISSN 1931-0587. Citado 2 vezes nas páginas 5 e 6.

SKLANSKY, J. Finding the convex hull of a simple polygon. *Pattern Recogn. Lett.*, Elsevier Science Inc., New York, NY, USA, v. 1, n. 2, p. 79–83, dez. 1982. ISSN 0167-8655. Disponível em: <http://dx.doi.org/10.1016/0167-8655(82)90016-2>. Citado na página 35.

SOONG, T. *Modelos probabilisticos em engenharia e ciencias*. LTC, 1986. ISBN 9788521604389. Disponível em: <https://books.google.com.br/books?id=wa87AwEACAAJ>. Citado 2 vezes nas páginas 4 e 13.

STAUFFER, C.; GRIMSON, W. Adaptive background mixture models for real-time tracking. In: *IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 1999.* Fort Collins, CO: IEEE, 1999. v. 2, p. –252 Vol. 2. ISSN 1063-6919. Citado 6 vezes nas páginas 4, 5, 6, 7, 18 e 74.

SUZUKI, S.; ABE, K. Topological structural analysis of digitized binary images by border following. *Computer Vision, Graphics, and Image Processing*, v. 30, n. 1, p. 32–46, 1985. Citado na página 35.

TAO, M. et al. Simpleflow: A non-iterative, sublinear optical flow algorithm. *Comput. Graph. Forum*, The Eurographs Association & John Wiley & Sons, Ltd., Chichester, UK, v. 31, n. 2pt1, p. 345–353, maio 2012. ISSN 0167-7055. Disponível em: <http://dx.doi.org/10.1111/j.1467-8659.2012.03013.x>. Citado na página 5.

TSAI, W.-H. Moment-preserving thresolding: A new approach. *Computer Vision, Graphics, and Image Processing*, v. 29, n. 3, p. 377–393, 1985. Disponível em: <http://dblp.uni-trier.de/db/journals/cvgip/cvgip29.html#Tsai85>. Citado na página 29.

WITHAGEN, P.; SCHUTTE, K.; GROEN, F. Likelihood-based object detection and object tracking using color histograms and em. In: *Image Processing. 2002. Proceedings. 2002 International Conference on*. [S.l.]: IEEE, 2002. v. 1, p. I–589–I–592 vol.1. ISSN 1522-4880. Citado na página 4.

YEN, J.-C.; CHANG, F.-J.; CHANG, S. A new criterion for automatic multilevel thresholding. *IEEE Transactions on Image Processing*, v. 4, n. 3, p. 370–378, 1995. Disponível em: <http://dblp.uni-trier.de/db/journals/tip/tip4.html#YenCC95>. Citado 3 vezes nas páginas 24, 25 e 62.

ZABIH, R.; MILLER, J.; MAI, K. A feature-based algorithm for detecting and classifying production effects. *Multimedia Syst.*, v. 7, n. 2, p. 119–128, 1999. Disponível em: <http://dblp.uni-trier.de/db/journals/mms/mms7.html#ZabihMM99>. Citado na página 6.

ZIVKOVIC, Z. Improved adaptive gaussian mixture model for background subtraction. In: *Pattern Recognition, 2004. ICPR 2004. Proceedings of the 17th International Conference on*. [S.l.]: IEEE, 2004. v. 2, p. 28–31 Vol.2. ISSN 1051-4651. Citado 21 vezes nas páginas 2, 3, 5, 7, 8, 18, 19, 21, 22, 33, 34, 37, 43, 44, 53, 54, 55, 61, 72, 73 e 74.