

Daniel Felipe Silva Santos

## **Contagem automatizada de veículos utilizando segmentação estatística**

Trabalho de conclusão de curso apresentado junto à disciplina Projetos e Implementação de Sistemas II, do curso de Bacharelado em Ciência da Computação, Universidade Estadual Paulista “Júlio de Mesquita Filho”, Faculdade de Ciências, Campus de Bauru, como parte dos requisitos para obtenção do título de Bacharel em Ciência da Computação.

**Orientador:** João Paulo Papa.

**Co-Orientador:** Bárbara Stolte Bezerra.

BAURU

2013

Santos, Daniel Felipe Silva  
Contagem Automatizada de Veículos Utilizando  
Segmentação Estatística / Daniel Felipe Silva  
Santos, 2013  
72 f. : il.

Orientador: João Paulo Papa  
Co-Orientador: Bárbara Stolte Bezerra

Monografia (Graduação)-Universidade Estadual  
Paulista. Faculdade de Ciências, Bauru, 2013

1. Segmentação estatística. 2. Rastreamento.  
3. Reconhecimento. 4. Vetor de estados. 5. Lógica  
de controle. I. Universidade Estadual Paulista.  
Faculdade de Ciências. II. Título.

Daniel Felipe Silva Santos

## **Contagem automatizada de veículos utilizando segmentação estatística**

Trabalho de conclusão de curso apresentado junto à disciplina Projetos e Implementação de Sistemas II, do curso de Bacharelado em Ciência da Computação, Universidade Estadual Paulista “Júlio de Mesquita Filho”, Faculdade de Ciências, Campus de Bauru, como parte dos requisitos para obtenção do título de Bacharel em Ciência da Computação.

**Orientador:** João Paulo Papa.

**Co-Orientador:** Bárbara Stolte Bezerra.

Banca Examinadora:

---

Prof<sup>o</sup>. Dr. João Paulo Papa

UNESP – BAURU

---

Prof<sup>o</sup>. Dr. Aparecido Nilceu Marana

UNESP – BAURU

---

Prof<sup>a</sup>. Dra. Simone das Graças Domingues Prado

UNESP – BAURU

Bauru, 21 de novembro de 2013.

*Dedico este trabalho aos meus pais e minha irmã, que foram  
imprescindíveis para a conclusão desta monografia.*

## **AGRADECIMENTOS**

Quero aqui agradecer primeiramente a Deus, por todas as pessoas boas e extremamente generosas que Ele colocou em minha vida. Dentre estas pessoas primeiramente estão meus pais e minha irmã, que sempre me apoiaram em minhas decisões e me incentivaram a lutar pelos meus ideais.

Agradeço também aos meus grandes amigos que me ajudaram a superar as muitas dificuldades que tive para me adaptar ao curso. Sem a ajuda deles eu provavelmente não conseguiria chegar até aqui.

Quero agradecer também aos meus orientadores, professor João Paulo e professora Bárbara, pois sempre foram muito atenciosos em atender meus pedidos de ajuda.

*Para ser grande, sê inteiro: nada  
Teu exagera ou exclui.  
Sê todo em cada coisa. Põe quanto és  
No mínimo que fazes.  
Assim em cada lago a lua toda  
Brilha, porque alta vive  
(Ricardo Reis)*

## RESUMO

Neste trabalho foi desenvolvido um programa capaz de realizar a contagem automática de veículos em vias. O problema da contagem de veículos está na utilização de técnicas ainda caras para sua realização, técnicas que muitas vezes envolvem contagem manual ou degradação do pavimento. A principal motivação para este trabalho foi a importância que a contagem de veículos representa para a engenharia de tráfego, sendo ela essencial para análise do desempenho das vias, possibilitando mensurar a necessidade de instalação de semáforos, rotatórias, alças de acesso, entre outros meios capazes de garantir um fluxo contínuo e seguro para os veículos. O principal objetivo deste trabalho foi aplicar uma técnica de segmentação estatística recentemente desenvolvida, baseada em um modelo de regressão linear não paramétrica, para solucionar o problema de segmentação do programa contador. O desenvolvimento do programa foi baseado na criação de três grandes módulos, um para a segmentação, outro para o rastreamento e outro para o reconhecimento. Para o desenvolvimento do módulo de segmentação, aplicou-se a técnica estatística combinada com a técnica de segmentação por subtração de fundo, com o intuito de otimizar o processo. O módulo de rastreamento foi desenvolvido com base na utilização de filtros de Kalman e na aplicação de conceitos simples de geometria analítica. Para desenvolver o módulo de reconhecimento, foram utilizados descritores de Fourier e uma rede neural perceptron multicamadas, treinada por retropropagação. Além dos módulos, desenvolveu-se também uma lógica de controle capaz de realizar a interligação entre os módulos, baseada principalmente em uma estrutura de dados denominada estado. A análise de resultados foi aplicada no programa contador e nos seus módulos componentes, sendo que a análise individual serviu como meio para estabelecer os valores dos parâmetros das técnicas utilizadas. O resultado final foi positivo, pois a técnica de segmentação estatística mostrou-se muito útil e o programa gerado foi capaz de contar os veículos pertencentes às três classes desejadas, carro, caminhão e ônibus com uma taxa de acerto em torno de 70% no melhor caso e 50% no pior. Além disso, o programa foi capaz de ser executado a 28 quadros por segundo, sendo um resultado muito promissor para programas de tempo real, que pedem uma taxa de execução em torno de 30 quadros por segundo.

**Palavras-chave:** Segmentação estatística, Rastreamento, Reconhecimento, Vetor de estados, Lógica de controle.

## ABSTRACT

In this work was developed a program capable of performing automatic counting of vehicles on roads. The problem of counting vehicles is using expensive techniques for its realization, techniques which often involve manual counting or degradation of the pavement. The main motivation for this work was the importance that the vehicle counting represents to the Traffic Engineer, being essential to analyze the performance of the roads, allowing to measure the need for installation of traffic lights, roundabouts, access ways, among other means capable of ensuring a continuous flow and safe for vehicles . The main objective of this work was to apply a statistical segmentation technique recently developed, based on a nonparametric linear regression model, to solve the segmentation problem of the program counter. The development program was based on the creation of three major modules, one for the segmentation, another for the tracking and another for the recognition. For the development of the segmentation module, it was applied a statistical technique combined with the segmentation by background difference, in order to optimize the process. The tracking module was developed based on the use of Kalman filters and application of simple concepts of analytical geometry. To develop the recognition module, it was used Fourier descriptors and a neural network multilayer perceptron, trained by backpropagation. Besides the development of the modules, it was also developed a control logic capable of performing the interconnection among the modules, mainly based on a data structure called state. The analysis of the results was applied to the program counter and its component modules, and the individual analysis served as a means to establish the parameter values of techniques used. The find result was positive, since the statistical segmentation technique proved to be very useful and the developed program was able to count the vehicles belonging to the three goal classes, car, truck and bus with a hit rate around 70% in the best case and 50% in the worst. In addition, the program was able to run around 28 frames per second, being a very promising result for real time programs, requesting an execution rate around 30 frames per second.

**Keywords:** Statistical Segmentation, Tracking, Recognition, State Vector, Control Logic.

## Lista de Figuras

|   |    |
|---|----|
| Figura 1 - Reta média gerada por uma Regressão Linear Simples ..... | 23 |
| Figura 2 - Utilização do método de segmentação estatística .....    | 27 |
| Figura 3 - Elementos Estruturantes .....                            | 29 |
| Figura 4 - Erosão e Dilatação .....                                 | 31 |
| Figura 5 - Geração do vetor de características de S .....           | 32 |
| Figura 6 - Aprendizado supervisionado .....                         | 35 |
| Figura 7 - Aprendizado não-supervisionado .....                     | 36 |
| Figura 8 - Neurônio artificial.....                                 | 36 |
| Figura 9 - Perceptron Multicamadas .....                            | 37 |
| Figura 10 - Superfície de erro da Rede Neural.....                  | 39 |
| Figura 11 - Influência do termo momentum .....                      | 39 |
| Figura 12 - Representação esquemática do fluxo de veículos .....    | 42 |
| Figura 13 - Exemplos de objetos aprovados pelo algoritmo NHD .....  | 45 |
| Figura 14 - Segmentação por subtração de fundo .....                | 45 |
| Figura 15 - Fluxograma do processo de segmentação .....             | 47 |
| Figura 16 - Fluxograma de rastreamento .....                        | 50 |
| Figura 17- Primeira etapa de definição .....                        | 51 |
| Figura 18 - Segunda etapa de definição .....                        | 52 |
| Figura 19 - Verificação de coordenada prevista.....                 | 53 |
| Figura 20 - Centro relativo .....                                   | 55 |
| Figura 21 - Fluxograma de reconhecimento .....                      | 56 |
| Figura 22 - Deslocamento de veículo.....                            | 60 |
| Figura 23 - Reconstrução de contornos.....                          | 61 |
| Figura 24 - Gráfico do Espectro de Magnitudes .....                 | 62 |
| Figura 25 - Vídeos de teste.....                                    | 64 |
| Figura 26 - Casos de uso do contador .....                          | 69 |
| Figura 27 - Classes do programa contador.....                       | 70 |
| Figura 28 - Tela Principal do Contador.....                         | 71 |
| Figura 29 - Janela de Parâmetros .....                              | 72 |

## Lista de Tabelas

|  |    |
|--|----|
| Tabela 1 - Dados de estado.....                        | 54 |
| Tabela 2 - Tabela de Classificação.....                | 57 |
| Tabela 3 - Parâmetros da Segmentação Estatística ..... | 59 |
| Tabela 4 – Dados de desempenho do treinamento.....     | 63 |
| Tabela 5- Resultado do Contador.....                   | 65 |

## Sumário

|   |           |
|---|-----------|
| <b>1 - Introdução.....</b>  | <b>12</b> |
| 1.1 - Situação problema.....  | 12        |
| 1.2 - Justificativa .....   | 13        |
| 1.3 - Objetivos .....   | 14        |
| 1.3.1 - Objetivo geral.....   | 14        |
| 1.3.2 - Objetivos específicos .....                                       | 14        |
| 1.4 - Metodologia .....   | 14        |
| <b>2 - Fundamentação teórica .....</b>                                    | <b>16</b> |
| 2.1 - Conceitos estatísticos .....  | 16        |
| 2.1.1 - Medidas de tendência central e de dispersão .....                 | 16        |
| 2.1.2 - Variáveis aleatórias e matrizes aleatórias .....                  | 17        |
| 2.1.3 - Vetor de esperanças e matriz de covariância.....                  | 17        |
| 2.1.4 - Processos estocásticos.....                                       | 18        |
| 2.1.5 - Distribuição normal multivariada .....                            | 19        |
| 2.1.6 - Maximização das estimativas probabilísticas de similaridade ..... | 21        |
| 2.1.7 - Maximização das estimativas probabilísticas a posteriori.....     | 22        |
| 2.1.8 - Regressão linear paramétrica e não paramétrica .....              | 22        |
| 2.2 - Segmentação de vídeo.....   | 24        |
| 2.2.1 - Detecção estatística de mudança.....                              | 25        |
| 2.2.2 - Geração do modelo estatístico da cena.....                        | 27        |
| 2.2.3 - Operações morfológicas em imagens digitais .....                  | 29        |
| 2.2.3.1 - Operação morfológica de erosão .....                            | 30        |
| 2.2.3.1 - Operação morfológica de dilatação .....                         | 30        |
| 2.3 - Identificação e reconhecimento de objetos através de contornos..... | 31        |
| 2.3.1 - Contornos e suas características .....                            | 31        |
| 2.3.2 - Descritores de Fourier.....                                       | 32        |
| 2.4 - Rede neural artificial e métodos de aprendizagem .....              | 34        |
| 2.4.1 - Perceptron .....  | 36        |

|  |           |
|--|-----------|
| 2.4.2 - Rede neural perceptron multicamadas .....              | 37        |
| 2.4.3 - Aprendizado por retropropagação .....                  | 38        |
| 2.5 - Contagem volumétrica de veículos em vias .....           | 41        |
| 2.5.1 - Medidas baseadas na contagem de veículos .....         | 41        |
| 2.5.1.1 - Fluxo ou vazão .....                                 | 41        |
| 2.5.1.2 - Densidade ou concentração .....                      | 42        |
| 2.5.1.3 - Velocidade .....                                     | 42        |
| 2.5.1.4 - Espaçamento .....                                    | 43        |
| 2.5.1.5 - Intervalo .....                                      | 43        |
| 2.5.1.6 - Fluxo de saturação .....                             | 44        |
| 2.6 - Trabalhos correlatos .....                               | 44        |
| <b>3 - Estrutura do programa contador .....</b>                | <b>46</b> |
| 3.1 - Módulo de segmentação .....                              | 46        |
| 3.1.1 - Otimização da segmentação estatística .....            | 48        |
| 3.2 - Módulo de rastreamento .....                             | 49        |
| 3.2.1 - Mecanismo de rastreamento .....                        | 49        |
| 3.2.2 - Atualização de posição .....                           | 51        |
| 3.2.3 - Definição de estado .....                              | 54        |
| 3.3 - Módulo de reconhecimento .....                           | 55        |
| <b>4 - Análise de parâmetros e resultados .....</b>            | <b>58</b> |
| 4.1 - Análise da segmentação .....                             | 58        |
| 4.2 - Análise do rastreamento .....                            | 59        |
| 4.3 - Análise do reconhecimento .....                          | 60        |
| 4.4 - Análise da contagem dos veículos .....                   | 63        |
| <b>5 - Conclusões e sugestões para trabalhos futuros .....</b> | <b>66</b> |
| <b>Referências bibliográficas .....</b>                        | <b>67</b> |
| <b>Apêndices .....</b>   | <b>69</b> |

## **1 - Introdução**

Assim como várias outras tarefas complexas realizadas pelo homem, como ler e interpretar um texto corretamente, discernir entre objetos vistos a longas distâncias, ou em estado de oclusão, o ato de reconhecer e contar objetos em movimento parece simples, por causa principalmente da forma como o cérebro, de maneira muito eficiente, consegue integrar e gerenciar toda a informação captada pelos sensores humanos, tais como olhos e ouvidos, de uma forma muito natural. No entanto, percebe-se que existe ainda uma grande dificuldade e uma grande limitação de recursos para poder resolver esses mesmos problemas de forma artificial. O processo de contagem de veículos faz parte deste imenso conjunto de problemas que dependem não só da utilização de bons sensores para captura de informações, mas também de boas técnicas para lidar com elas. Estas técnicas incluem principalmente a segmentação e reconhecimento dos objetos nas imagens analisadas, que levam também em consideração o rastreamento do veículo que pode estar em movimento ou mesmo parado no momento da análise. Para construir um sistema contador de veículos é necessário recorrer a diversos segmentos de estudo, como visão computacional, estatística, aprendizado de máquina e engenharia de tráfego. Além disto, é preciso integrar, da melhor forma possível, todo este conhecimento adquirido, para que o contador seja eficiente e robusto, capaz de gerar bons resultados nas mais diversas situações em que for utilizado.

### **1.1 - Situação problema**

O problema que aqui se apresenta é a maneira como é realizada a contagem volumétrica de veículos. Existem diversas técnicas para realizar a contagem de fluxo de veículos, só que, no entanto, a maior parte dessas técnicas envolve a instalação de sensores que podem degradar o pavimento, ou então envolve contagem manual e, dessa forma, o processo torna-se extremamente desgastante e caro. Outro problema que surge é com relação à categorização posterior dos dados adquiridos. Isso porque um programa de computador para contagem poderia ser facilmente integrado a outro programa para análise de dados de contagem, enquanto que se o dispositivo de contagem não for integrado, haverá a necessidade de transferir esses dados posteriormente para um computador que possua o software de análise em questão.

## 1.2 - Justificativa

"A contagem volumétrica de veículos em vias é essencial para a área de engenharia de tráfego, pois, a partir dos volumes, é possível estabelecer parâmetros de operação e desempenho dos diversos tipos de vias e seus dispositivos, tais como: semáforos, rotatórias, alças de acesso, interseções rodoviárias entre outros." Trecho retirado de (HCM,2010). Segundo o próprio DNIT (Departamento Nacional de Infraestrutura de Transportes), um sistema de contagem de tráfego seria essencial para o cumprimento das seguintes finalidades:

- Planejar o sistema rodoviário;
- Programar necessidades e prioridades de melhorias no sistema rodoviário;
- Medir a demanda atual de serviços por via rodoviária;
- Estabelecer as tendências de tráfego no futuro;
- Determinar os volumes de viagens de forma a proporcionar justificativa econômica aos investimentos programados;
- Avaliar o fluxo existente de tráfego em relação ao sistema rodoviário atual;
- Estimar os benefícios dos usuários nas rodovias;
- Estabelecer uma classificação do sistema rodoviário;
- Justificar e planejar o policiamento;
- Estabelecer o *veículo de projeto*<sup>1</sup> para fins de *projeto geométrico*<sup>2</sup>;
- Projetar pavimento, obras de arte, seção transversal e outros elementos de rodovia;
- Estudar a localização de postos de pesagem, socorro médico emergencial, etc;
- Analisar a capacidade e estabelecer o Nível de Serviço;
- Realizar análise estatística de acidentes;
- Localizar e projetar instalações para a operação rodoviária.

---

<sup>1</sup> O veículo de projeto é um veículo teórico de certa categoria, que serve como modelo de estudo para os veículos reais pertencentes à mesma categoria.

<sup>2</sup> O projeto geométrico é a fase de estudo das características do traçado geométrico das estradas.

## **1.3 - Objetivos**

### **1.3.1 - Objetivo geral**

O objetivo geral deste trabalho é realizar a contagem de veículos utilizando técnicas aritméticas, geométricas e estatísticas, a fim de, primeiramente segmentar o objeto de interesse, e posteriormente aplicar uma análise de dimensões ao mesmo visando classificá-lo como sendo ou não um veículo.

### **1.3.2 - Objetivos específicos**

Os objetivos específicos deste trabalho são:

1. Estudar e compreender a técnica de segmentação “Statistical Change Detection” proposta em Lanza e Stefano (2011).
2. Realizar a coleta dos vídeos em vias terrestres em diferentes horários do dia.
3. Aplicar e adequar o método “Statistical Change Detection” aos vídeos coletados.
4. Estudar uma forma de contabilizar os objetos segmentados somente se estes forem veículos.
5. Gerar o programa contador para realizar a segmentação e contagem.

## **1.4 - Metodologia**

Na primeira etapa foi estudada a segmentação estatística proposta por (LANZA; STEFANO, 2011), assim como, todo o conteúdo advindo dessa técnica. Alguns dos conteúdos principais seriam processos estocásticos, regressão linear não paramétrica e análise normal multivariada. Na etapa seguinte foram coletados os vídeos em rodovias e vias urbanas em pelo menos dois períodos do dia, das 9h às 12h e das 15h às 18h, utilizando uma câmera Nikon Coolpix P510 fixada a um tripé. O próximo passo foi implementar computacionalmente o algoritmo proposto por (LANZA; STEFANO, 2011) e executá-lo nos vídeos coletados, realizando os ajustes necessários, como redução e análise morfológica das imagens, a fim de obter uma máscara de segmentação precisa em uma velocidade próxima a tempo real.

Para poder realizar a contagem partindo da imagem segmentada, foram estudadas em uma primeira etapa, técnicas de extração e reconhecimento de características de contornos e em uma segunda etapa, essas técnicas estudadas foram aplicadas visando à identificação entre veículo e não veículo em cena, para poder efetuar a contagem. O ambiente de desenvolvimento utilizado foi o Visual Studio 2012, as linguagens utilizadas foram C e C++/CLI e para auxiliar nas operações de processamento de imagens, como redução de ruídos, segmentação e extração de contornos, fora utilizada a biblioteca OpenCV na sua versão 2.4.3. Ao final do trabalho, os resultados gerados pela etapa de segmentação, pela etapa de identificação e, posteriormente, pela contagem propriamente dita, serão expostos por meio de tabelas e gráficos.

## **2 - Fundamentação teórica**

Este tópico abrange todos os conceitos estatísticos fundamentais para compreensão da técnica de segmentação estatística utilizada no programa contador. Além do conceito estatístico, são descritos também os conceitos mais fundamentais do processo de segmentação aplicados neste trabalho, assim como, o conceito de operações morfológicas e o método de segmentação estatística proposto por Lanza e Stefano (2011). Ao final deste tópico, fundamenta-se o método de reconhecimento utilizado e a importância do resultado gerado pelo contador para posterior análise destes dados pela Engenharia de Tráfego.

### **2.1 - Conceitos estatísticos**

Nos tópicos seguintes serão fundamentados os conceitos estatísticos necessários à etapa de segmentação estatística. Nos tópicos, que vão de 2.1.1 à 2.1.3, serão abordados os conceitos estatísticos mais básicos. O tópico 2.1.4 abordará o que são processos estocásticos e suas implicações estatísticas e o tópico 2.1.5 tratará da distribuição normal multivariada. Os tópicos 2.1.6 e 2.1.7 abordarão técnicas de maximização de estimativas probabilísticas. Por fim, o tópico 2.1.8 abordará o assunto de regressões lineares.

#### **2.1.1 - Medidas de tendência central e de dispersão**

Quando se fala em estatística uma das primeiras coisas que logo vem a mente é a noção de probabilidade. No entanto, apesar de ser um conceito muito importante também, a probabilidade isoladamente tratada não consegue explicar muito a respeito das tendências e correlações entre os dados estudados. O uso da estatística só faz sentido então, em um contexto mais amplo, onde outros métodos de análise devem ser aplicados, para se chegar a uma conclusão mais concreta a respeito do comportamento de um objeto de estudo. Dentre estes métodos alternativos, encontram-se as medidas de tendência central e as de variabilidade. No primeiro caso, é feito um estudo do quão distantes do centro da amostra os dados estão, fazendo parte desse estudo a média aritmética, mediana e a moda. Já a forma como os dados variam no interior da amostra, é dada pela variância, covariância e desvio padrão. Maiores explicações em Moore, McCabe e Craig (2009).

### 2.1.2 - Variáveis aleatórias e matrizes aleatórias

Uma amostra estatística é um conjunto de dados coletados a partir do comportamento de um determinado objeto de estudo. Esses valores coletados são mapeados por uma função  $X(\zeta)$  denominada variável aleatória, onde é estabelecida uma relação entre os dados amostrais representados por  $\zeta$  e as ocorrências desses dados mapeadas por  $X$ . Para um melhor esclarecimento do assunto, vide Papoulis e Pillai (2002). Na área de estatística, um objeto de estudo possui várias características a serem analisadas, sendo que cada uma pode estar condicionada a uma variável aleatória particular. Para representar as variáveis aleatórias em um espaço  $n$ -dimensional, utiliza-se uma matriz  $n \times p$  chamada de matriz aleatória, onde  $n$  e  $p$  representam respectivamente cada variável aleatória e a dimensão de suas realizações. Abaixo estão ilustradas essas representações matriciais.

$$X_{(n \times p)} = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1p} \\ x_{21} & x_{22} & \cdots & x_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{np} \end{bmatrix} = \begin{bmatrix} x_1' \\ x_2' \\ \vdots \\ x_n' \end{bmatrix} \quad (1)$$

$$X_{(n \times p)} = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1p} \\ x_{21} & x_{22} & \cdots & x_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{np} \end{bmatrix} = [y_1 \quad y_2 \quad \cdots \quad y_p] \quad (2)$$

### 2.1.3 - Vetor de esperanças e matriz de covariância

A fim de calcular o centro da dispersão controlada por uma determinada variável aleatória discreta ou contínua, utiliza-se a esperança estatística. A partir do resultado encontrado para a esperança, pode-se então estimar a variância entre os dados. Desta forma, para analisar o comportamento de um objeto de estudo, caracterizado estatisticamente por mais de uma variável aleatória, passa a ser necessário reunir os dados do vetor de esperança e da matriz aleatória, e calcular a matriz de covariância. Para formalizar tais conceitos, considera-se o vetor de esperanças como sendo um vetor definido por,

$$E(X) = \begin{bmatrix} E(X_1) \\ E(X_2) \\ \vdots \\ E(X_p) \end{bmatrix} = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_p \end{bmatrix} = \mu \quad (3)$$

Sendo a matriz de covariância dada por:

$$\Sigma = E[(X - \mu)(X - \mu)^T] = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1p} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{p1} & \sigma_{p2} & \cdots & \sigma_{pp} \end{bmatrix} \quad (4)$$

### 2.1.4 - Processos estocásticos

Um processo estocástico é uma coleção de variáveis aleatórias  $(X(t) : t \in T)$  onde  $t$  é um parâmetro que está contido no conjunto indexado por  $T$ . Em geral nós chamamos  $t$  de parâmetro-tempo (ou simplesmente o tempo), e  $T \subseteq R$ . Cada  $X(t)$  possui seus valores contidos em algum conjunto  $S \subseteq R$  chamado de conjunto estado; logo  $X(t)$  é o estado do processo no tempo  $t$ . (STIRZAKER, 2005, p. 45, tradução nossa).

Um processo estocástico trata da análise do comportamento das variáveis aleatórias ao longo de um período  $t$  de observação. É importante que esse tipo de análise seja feita, já que, em muitos casos, uma amostra estatística de dados está condicionada a um intervalo de tempo, e dessa forma os resultados são dependentes desse parâmetro. Para exemplificar, pode-se imaginar um estudo que pretende determinar qual a probabilidade de uma pessoa recém-chegada a uma fila de um caixa ser atendida antes do fechamento dele. Como o tempo de atendimento pode variar de pessoa para pessoa, para determinação da probabilidade de atendimento, uma variável de tempo  $t$  deve ser adicionada a observação da variável aleatória fila.

Ainda segundo Stirzaker (2005), um processo estocástico pode ser de caráter discreto ou contínuo dependendo em qual conjunto  $t$  se enquadra. Sabendo que  $t \in T$ , então caso  $T \subseteq R$ , sendo  $R$  o conjunto dos números reais, tem-se um processo estocástico contínuo.

Caso  $T \subseteq Z$  ou  $T \subseteq Z^+$ , sendo  $Z$  o conjunto dos números inteiros, tem-se um processo estocástico discreto. A fim de sintetizar o que seria um processo estocástico, (PAPOULIS; PILLAI, 2002, v. 4, p.373, tradução nossa) estabelece 4 interpretações para  $X(t)$ , sendo elas:

1. É uma família de funções  $X(t, \zeta)$ . Nesta interpretação,  $t$  e  $\zeta$  são variáveis.
2. É uma função do tempo de uma única variável (ou uma amostra do processo dados). Neste caso,  $t$  é uma variável e  $\zeta$  é fixa.
3. Se  $t$  é fixo e  $\zeta$  é variável, então  $X(t)$  é uma variável aleatória igual ao estado do processo dado em um tempo  $t$ .
4. Se  $t$  e  $\zeta$  são fixos, então  $X(t)$  é um número.

### 2.1.5 - Distribuição normal multivariada

Para compreender melhor o conceito de uma distribuição normal multivariada, é necessário estabelecer um paralelo entre ela e a distribuição normal, que leva em consideração a distribuição normal aplicada apenas a uma única variável aleatória. Uma distribuição normal é uma distribuição simétrica dos dados de uma amostra em torno da própria média, onde a área sobre a curva descrita pela função normal representa a probabilidade de ocorrência de um evento delimitado por este espaço amostral. Essa distribuição é representada por uma função gaussiana parametrizada pela média e pela variância da distribuição. Dessa maneira a função é representada da seguinte forma:

$$f(x | \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-[(x - \mu) / \sigma]^2 / 2} \quad (5)$$

Segundo Johnson e Wichern (2007) esta função é controlada pelo expoente de  $e$ , que é o quadrado da distância de  $x$  até  $\mu$ , distância esta, ponderada por  $\sigma$ . Para melhor explicitar tem-se que:

$$\left(\frac{x - \mu}{\sigma}\right)^2 = (x - \mu)(\sigma^2)(x - \mu) \quad (6)$$

A equação 5 possui algumas características muito importantes, dentre elas a de possuir a média igual ao valor da mediana, ser uma função suave (contínua em todo seu

domínio) e também segundo Moore, McCabe e Craig (2009), conseguir representar de forma bastante precisa experimentos reais, aproximar distribuições aleatórias discretas a contínuas e também servir de aproximação para outros tipos de funções simétricas. A partir então dessa breve conceituação a respeito de uma distribuição normal, uma distribuição normal multivariada pode ser interpretada como uma generalização desse conceito. Essa generalização estende a aplicação da distribuição a uma única variável aleatória  $X$ , a uma análise da distribuição normal aplicada a  $p$  variáveis aleatórias, sendo  $p \geq 2$ .

Partindo então desse ponto,  $\mu$  passa a ser um vetor  $p \times 1$  dimensional dos vetores das médias referentes às  $p$  variáveis aleatórias,  $X$  passa a ser um vetor  $p \times 1$  dimensional das  $p$  variáveis aleatórias observadas e  $\sigma^2$  é substituído pela matriz  $p \times p$  dimensional de variância-covariância  $\Sigma$ . Dessa forma o quadrado da distância monovariada, passa a ser denotado pela fórmula:

$$(X - \mu)^T \Sigma^{-1} (X - \mu) \quad (7)$$

Aplicando as devidas alterações à constante monovariada  $(2\pi)^{-1/2} (\sigma^2)^{-1/2}$ , a fórmula para uma distribuição normal multivariada  $p$  – dimensional de um vetor aleatório  $X^T = [X_1, X_2, \dots, X_p]$  passa a ser descrita por:

$$N_p(X | \mu, \Sigma) = \frac{1}{(2\pi)^{p/2} |\Sigma|^{1/2}} e^{-\frac{(X - \mu)^T \Sigma^{-1} (X - \mu)}{2}} \quad (8)$$

Para maiores esclarecimentos vide Johnson e Wichern (2007).

### 2.1.6 - Maximização das estimativas probabilísticas de similaridade

Uma definição do que seria uma maximização das estimativas probabilísticas de similaridade amostral é dada por Kobayashi, Mark e Turin (p. 528, 2012, tradução nossa), onde é dito que:

Considerando o caso onde nós conhecemos a forma de uma função de distribuição de probabilidade  $f_X(x; \theta)$  (ou função massa de probabilidade  $p_X(x; \theta)$  para o caso discreto), mas o valor do parâmetro  $\theta$  é desconhecido. Nós desejamos estimar  $\theta$  baseados nos dados em  $x = (x_1, x_2, \dots, x_n)$ . Então um procedimento razoável seria encontrar o valor de  $\theta$  mais suscetível a produzir  $x$ .

Ainda segundo Kobayashi, Mark e Tutin (2012), qualquer  $\theta$  que seja um parâmetro maximizador pode ser representado por  $\hat{\theta}$  da seguinte forma,

$$\hat{\theta} = \arg \max_{\theta} L_x(\theta) \quad (9)$$

Partindo da definição acima, uma MEPS (maximização das estimativas probabilísticas de similaridade) pode ser utilizada para encontrar um ou mais pontos de máximo que possam existir em uma função de distribuição de probabilidade. Sendo assim, técnicas que necessitem efetuar separações entre dados, podem se beneficiar desse método. MEPS são comumente associadas a distribuições normais, devido ao caráter contínuo de suas funções de distribuição e sua maior usabilidade em modelos reais.

Para  $n$  variáveis aleatórias em uma distribuição normal multivariada  $N_p(X | \theta)$ , considerando que as variáveis aleatórias são independentes, então o cálculo da MEPS será dado pelo resultado obtido pela probabilidade condicional calculada entre as variáveis aleatórias. O cálculo fica representado por,

$$L_x(\theta) = \prod_{j=1}^n \frac{1}{(2\pi)^{p/2} |\Sigma|^{1/2}} e^{-\frac{(X_j - \mu)^T \Sigma^{-1} (X_j - \mu)}{2}} \quad (10)$$

Para maiores detalhes vide Johnson e Wichern (2007).

### 2.1.7 - Maximização das estimativas probabilísticas a posteriori

Este método de maximização é muito parecido com o método descrito em 1.4 (MEPS), diferenciando-se pela utilização de uma abordagem bayesiana na determinação de um ou mais pontos ótimos na amostra. Enquanto que uma MEPS visa maximizar uma função probabilística de similaridade  $L_x(\theta)$  baseando-se em um parâmetro  $\theta$  constante, uma MEPP trata  $\theta$  como sendo uma variável aleatória  $\Theta$  e que possui uma função própria de distribuição de probabilidade. Sendo assim, o objetivo de uma MEPP é maximizar a probabilidade a posteriori de  $n$  variáveis  $p$ -dimensionais aleatórias  $X_p$  condicionadas a uma probabilidade a priori calculada sobre a distribuição de uma variável aleatória  $p$ -dimensional  $\Theta$ . Para formalizar esse conceito segue a equação abaixo,

$$\hat{\theta}(x) = \arg \max_{\hat{\theta}} \pi(\hat{\theta} | x) \quad (11)$$

sendo a função a posteriori  $\pi(\hat{\theta} | x)$  descrita como,

$$\pi(\hat{\theta} | x) = \frac{L_x(\hat{\theta})\pi(\hat{\theta})}{p_x(x)} \quad (12)$$

Para maiores detalhes vide Kobayashi, Mark e Turin (2012).

### 2.1.8 - Regressão linear paramétrica e não paramétrica

Para muitos dos eventos que ocorrem na natureza não se tem um modelo definido, sendo que apenas se consegue armazenar um conjunto de dados variáveis de acordo com observações de suas ocorrências. Uma das soluções propostas na estatística para este problema, é tratá-lo como uma função linear dos seus dados de ocorrência, realizando para isso um processo denominado regressão linear. Este processo tem por objetivo aproximar os dados distribuídos aleatoriamente a uma reta da melhor forma possível, ou seja, tenta-se encontrar uma reta média que satisfaça a condição de estar a menor distância possível de todos os dados amostrais.

Segundo Moore, McCabe e Craig (p. 561, 2009, tradução nossa),

O modelo estatístico para uma regressão linear simples, também assume que para cada valor de  $x$  os valores observados da variável de resposta  $y$  estão dispostos segundo uma Distribuição Normal com uma média que depende de  $x$ . Em regressões lineares simples, nós assumimos que as médias recaem todas sobre uma reta[...].

Ainda segundo Moorem McCabe e Craig (2009), a equação da reta média  $\mu_y$  fica representada pela equação:

$$\mu_y = \beta_0 + \beta_1 x \quad (13)$$

Para melhor ilustrar as definições acima, tem-se a Figura 1.

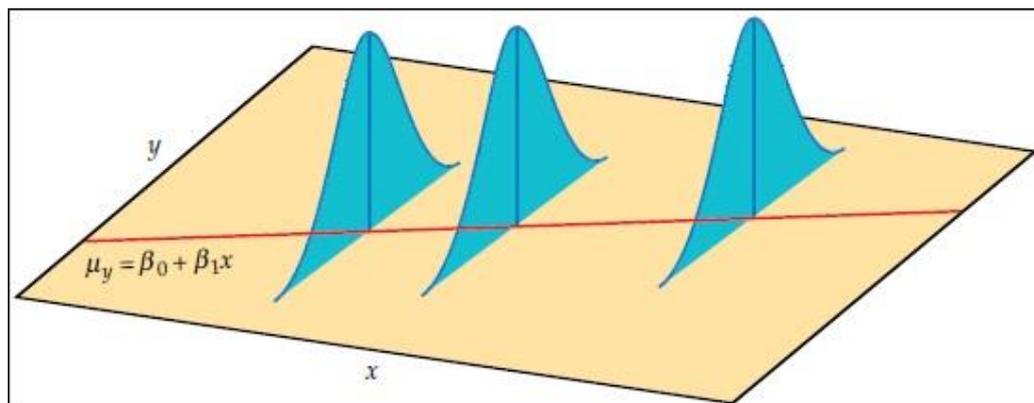


Figura 1- Reta média gerada por uma Regressão Linear Simples  
Fonte: (MOORE; MCCABE; GRAY, v. 6, p. 561, 2009)

Para realizar a regressão linear dois métodos podem ser utilizados, o método paramétrico e o não paramétrico. Em linhas gerais, o método paramétrico busca estimar os coeficientes  $\beta_0$  e  $\beta_1$  (parâmetros) pertencentes à função  $\mu_y$  por meio dos dados estatísticos disponíveis. Apesar da definição de regressão paramétrica estar vinculada a ideia de estimativa de parâmetros, a definição de regressão não paramétrica não vai de encontro à mesma. Em outras palavras, a regressão não paramétrica também busca estimar os coeficientes de uma regressão linear, só que a quantidade de parâmetros a serem estimados pela mesma, não é a principal preocupação do método.

(TAKEZAWA, p. 19, 2006, tradução nossa) diz que:

Regressão paramétrica favorece expressões quanto menor for o número de parâmetros e seleciona uma equação de regressão com um grande número de parâmetros apenas quando o uso de tal equação de regressão for indispensável para representar bem os dados. Por outro lado, regressão não paramétrica, é focada principalmente em atingir o objetivo de extrair tendências benéficas dos dados, sem dar atenção especial à redução de parâmetros.

## **2.2 - Segmentação de vídeo**

A segmentação de imagens, de uma forma geral, pode ser considerada uma etapa de separação do conteúdo de interesse, comumente chamado de imagem de frente, do conteúdo irrelevante, chamado de imagem de fundo. O processo de vídeo segmentação compreende a aplicação de técnicas de segmentação a um conjunto de várias imagens, já que um vídeo nada mais é do que um conjunto de imagens sequencialmente capturadas de um determinado evento, a uma determinada taxa de resolução, que são separadas por um intervalo de tempo mínimo capaz de transmitir uma sensação de continuidade entre elas. De acordo com as principais características de um vídeo, duas necessidades devem ser primordialmente satisfeitas.

A primeira delas é a eficiência, ou seja, o processo de segmentação deve ser suficientemente rápido, e a segunda é que ele deve ser robusto, produzindo bons resultados nas mais diversas situações em que for aplicado. Um dos fatores mais prejudiciais à eficiência de um método de segmentação é a complexidade da própria técnica utilizada, o que pode se tornar ainda mais agravante ao ser levado em conta à necessidade, muitas vezes, de se aplicar pré-processamentos às imagens analisadas. A aplicação de filtros às imagens geralmente é necessária para reduzir ruídos e corrigir alterações bruscas de luminosidade que possam contribuir negativamente para com etapas posteriores que venham a se beneficiar com os resultados da etapa de segmentação, como um processo de reconhecimento, que fará uso da informação contida na região segmentada. Com relação à robustez do método, o problema da segmentação se torna ainda mais evidente e complexo de ser tratado, pois a segmentação é altamente dependente das características de luminosidade dos ambientes de estudo.

Algumas técnicas, recentemente utilizadas, tentam lidar com esses problemas por meio de uma abordagem estatística do ambiente a ser segmentado. Com essas técnicas é

possível estudar alterações espaço temporais que possam ocorrer nos pixels de uma imagem e estabelecer parâmetros de referência para melhor estudar essas mudanças. Entram neste escopo técnicas como Modelos de Mistura em Friedmani, Hastie e Tibshirani (2008) e técnicas de modelos de predições temporais mostradas por Al-Kofahi, Andra, Radke e Roysam (2005) além de Lanza e Stefano (2006). O desafio maior para uma segmentação estatística é elaborar modelos que contemplem as inúmeras possibilidades de variação que um pixel possa sofrer, e dentre elas, conseguir separar as que são provenientes do objeto a ser segmentado e as que fazem parte do fundo de imagem ou qualquer outro ruído não pertencente ao objeto.

### 2.2.1 - Detecção estatística de mudança

Pesquisadores da área de visão computacional, recentemente, estão utilizando da estatística para análise da segmentação de imagens. Vários dos conceitos estatísticos abordados no tópico 1, estão sendo aplicados em técnicas de segmentação, a fim de, explorar as correlações entre as características apresentadas por um determinado objeto em um cenário não determinístico. Dentre essas técnicas, em Lanza e Stefano (2011), o problema da segmentação é abordado como sendo um problema de maximização a posteriori, que possui solução segundo um método de regressão linear não paramétrica. O problema é formalizado pelo autor da seguinte maneira:

$$t(U, f) \underset{U}{\overset{c}{>}} T \quad (14)$$

Segundo (LANZA; STEFANO, p. 1897, 2011, tradução nossa), “[...]  $t(U, f)$  é um teste estatístico que provê uma medida da *distância* entre dois eventos, respectivamente, provenientes apenas de fatores de distúrbios agindo na região analisada, e pelo vetor de características  $f$  notadas.”. No entanto, para compreender melhor a formulação para o problema, deve-se subdividi-lo em etapas. A primeira delas diz respeito a como encontrar o vetor de características  $f$ . O vetor de características é gerado por uma concatenação das tonalidades dos pixels extraídos das vizinhanças quadráticas  $N$ -dimensionais da imagem de fundo e a imagem a ser segmentada, tomados em ordem crescente de valores. Sendo assim, o vetor  $f$  passa a ser um vetor  $2N$ -dimensional de tal forma que,

$$f = (x \parallel y) = (x_1, \dots, x_N, y_1, \dots, y_N) \in F, \text{ onde} \quad (15)$$

$$F = [0, L-1]^{2N} \subset Z^{2N}, \quad (16)$$

é o espaço discreto das características observadas, onde os valores em escala de cinza que os pixels podem assumir limitam-se a  $L$ .

Na segunda etapa, os autores buscam uma forma de modelar o problema da alteração em determinada fração da imagem analisada, com base em  $f$  e em um estudo da maximização das estimativas probabilísticas a posteriori, mencionada no tópico 1.5. Partindo então do fato de que o estado desta fração da imagem, representado por  $c$ , pode assumir os valores:

- 0, caso ocorram apenas alterações por distúrbios na imagem,
- 1, caso realmente uma mudança tenha ocorrido,
- 2, caso ambos os estados ocorram na mesma região analisada, a maximização pode ser aplicada da seguinte forma,

$$c^{MAP} = \arg \max_c p(c | f) = \arg \max_c p(f | c)p(c), \quad (17)$$

onde  $p(c | f)$  é a probabilidade a posteriori de  $c$  que se pretende maximizar,  $p(f | c)$  a função probabilística de similaridade  $f$  (MLE/MEPS, tópico 1.4) e  $p(c)$  a probabilidade a priori calculada sobre o estado  $c$ .

Como última etapa da formalização estatística, os dois estados 1 e 2 que  $c$  pode assumir, são então combinados em um único estado de mudança  $C$  e em contrapartida, o estado de não mudança, passa a ser definido como  $U$ . Feitas essas devidas considerações, os autores resumem a detecção de mudança estatística em uma região de determinada imagem como sendo uma regra binária Bayesiana onde,

$$p(C|f) \underset{U}{\overset{C}{>}} p(U|f) \quad (18)$$

ficando estabelecido que quando a probabilidade a posteriori de ocorrer uma mudança na região da imagem analisada superar a probabilidade a posteriori de não ocorrência, pode-se realizar a segmentação, pois houve uma mudança não proveniente de ruídos ou distúrbios. A Figura 2 ilustra o resultado da segmentação.

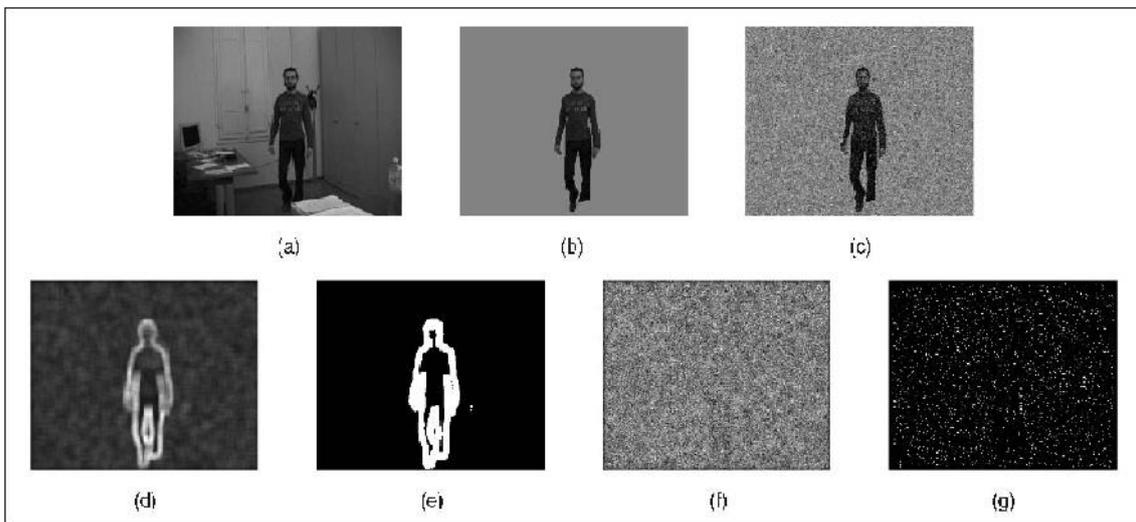


Figura 2 - Utilização do método de segmentação estatística  
 Legenda: As máscaras binárias de mudança (e), (g) foram obtidas limiarizando as imagens correspondentes (d), (f) com taxa de falso positivo de 5%.  
 Fonte: (LANZA; STEFANO, p. 1903, 2011, tradução nossa).

### 2.2.2 - Geração do modelo estatístico da cena

Para realizar a detecção estatística, explicada no tópico imediatamente acima, é necessário que se tenha as medidas de variância para cada pixel na cena, possibilitando o cálculo da regressão linear não paramétrica. Tentando obter um resultado mais robusto para o cálculo dessas variâncias, em Bevilaqua, Lanza e Stefano (2005) é utilizada a função massa de probabilidade gerada de forma não paramétrica por um processo estocástico aplicado a cada pixel individualmente. Formalmente falando, dado um intervalo  $I$  de tempo, onde cada imagem capturada do vídeo possua um total de  $i$  pixels,  $i \in [1; n]$ , é possível construir um processo estocástico segundo a equação abaixo.

$$P_i^I = \{p_i(t) : t \in I\} \quad (19)$$

É importante notar, que assim como dito em 1.2, cada pixel passa a ser tratado como uma variável aleatória, sendo que, as realizações dessa variável compreendem os valores de 0 a 255 correspondentes aos valores assumidos por um pixel pertencente à uma imagem de 8 bits de informação, em tons de cinza.

Sugerindo em seu trabalho que o processo estocástico possui ainda duas fases distintas, Bevilaqua, Lanza e Stefano (2005) propõe a geração de um modelo que combine as características estatísticas do pixel enquanto deterministicamente estacionário, com as características do pixel influenciado pela dinâmica do ruído de câmera, dinâmica esta considerada como única influência relevante na alteração da cor do pixel ao longo do tempo. Sendo assim, tomando  $B_i(t)$  como o processo determinístico e  $N_i(t)$  como o não determinístico, tem-se que:

$$P_i(t) = B_i(t) + N_i(t) \quad (20)$$

Para representar  $B_i(t)$ , os mesmos autores sugerem o cálculo de uma imagem mediana sobre a primeira metade do intervalo de tempo  $I$  do processo estocástico. Já para representar  $N_i(t)$ , sugerem a utilização da imagem mediana como meio de identificação dos pixels estacionários ao longo do tempo. Para os pixels estacionários, determinados pelo cálculo da diferença entre as imagens pertencentes à segunda metade do intervalo de tempo  $I$  e a imagem mediana, criam-se conjuntos de imagens, nas quais o resultado da diferença calculada tenha sido 0.

Ainda para os pixels nos quais a cardinalidade de seus conjuntos de imagens tenham ultrapassado um valor limiar, chamado também de ergodicidade pelos autores, calcula-se um histograma seletivo relativo temporal, baseado na frequência dos valores assumidos pelo pixel ao longo do tempo nas imagens em seu conjunto. Por fim, por meio do modelo de ruído, adotado como sendo uma imagem da média dos valores assumidos pelo pixel ao longo do tempo, constrói-se a função massa de probabilidade. Formalizando, tem-se que para cada  $w = \mu_i^{I^2}$ , ou seja, para cada tom de cinza presente na imagem média, é construída uma  $pdf_w(v)$ , sendo que  $v \in [0;255]$ .

### 2.2.3 - Operações morfológicas em imagens digitais

Como o próprio nome sugere, operações morfológicas estudam as estruturas geométricas presentes em uma imagem, suas formas. Segundo Facon (2011, p. 62): “O princípio básico da morfologia matemática consiste em extrair informações relativas à geometria e à topologia de conjuntos desconhecidos de uma imagem a partir do *elemento estruturante*.”.

O elemento estruturante é um conjunto conhecido e definido pelo usuário a ser aplicado sobre a imagem de interesse. De maneira bastante similar ao processo de convolução de imagens, em que uma matriz, mais conhecida como *kernel*, é transladada por toda a imagem, ocorre o processamento morfológico de imagens. As principais diferenças passam a ser o *kernel*, sendo substituído pelo elemento estruturante, e as operações realizadas pelo mesmo sobre a imagem, que tomam agora por base a álgebra de conjuntos. No quadro abaixo é possível visualizar alguns dos elementos estruturantes mais utilizados.

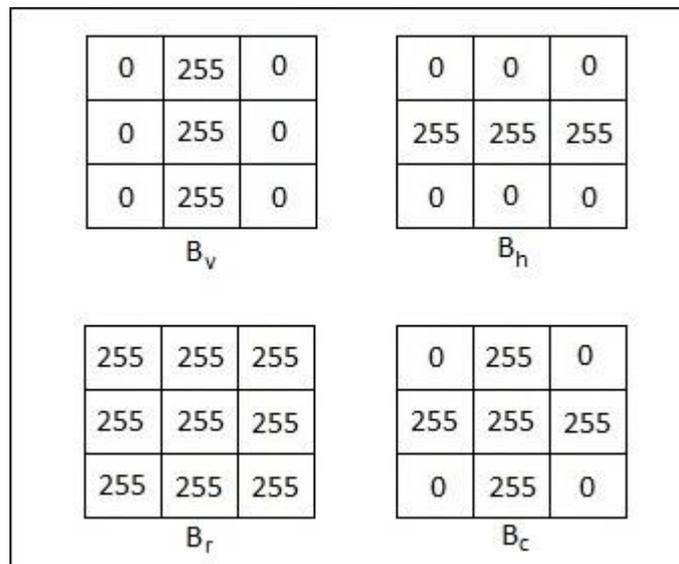


Figura 3 - Elementos Estruturantes

Legenda:  $B_v$ ,  $B_h$ ,  $B_r$  e  $B_c$  representam respectivamente os elementos estruturantes vertical, horizontal, retangular e em cruz.

Fonte: Elaborada pelo autor.

As operações mais básicas são a erosão e a dilatação. A partir delas é possível derivar operações mais complexas como operações de abertura, fechamento e até mesmo extração de contornos. Nos tópicos subsequentes serão formalizadas as operações de erosão e dilatação por serem justamente as mais fundamentais.

### 2.2.3.1 - Operação morfológica de erosão

Segundo Facon (2011 apud SERRA, 1982) a erosão binária  $\varepsilon$  de uma imagem  $f$  por um elemento estruturante  $B$  é definida como,

$$\varepsilon^B(f) = \{x \in \varepsilon : B_x \subset f\} \quad (21)$$

Por meio dessa definição, o autor explica que o processo de erosão binária é realizado posicionando-se o elemento estruturante  $B_x$ , ilustrado na Figura 3, sobre cada píxel da imagem  $f$ . Após o posicionamento, realizam-se comparações entre os pixels adjacentes à  $x$  que se encontram sob  $B$ , e os pixels no elemento estruturante. Dessa forma é possível determinar se o píxel  $x$  será mantido ou removido da imagem erodida resultante, se respectivamente, os valores dos pixels na vizinhança de  $x$  forem diretamente correspondentes aos do elemento estruturante, ou caso um ou mais não correspondam.

### 2.2.3.1 - Operação morfológica de dilatação

Ao contrário da operação morfológica de erosão, a dilatação tem por objetivo adicionar pixels à vizinhança do píxel analisado pelo elemento estruturante. Segundo Facon (2011 apud SERRA, 1982) a erosão binária  $\delta$  de uma imagem  $f$  por um elemento estruturante  $B$  é definida como,

$$\delta^B(f) = \{x \in f : B_x \cap f \neq \emptyset\} \quad (22)$$

A definição acima, de uma forma bastante simplificada, diz que a dilatação será aplicada na vizinhança de um píxel, vizinhança da mesma dimensão do elemento estruturante, sempre que o resultado da intersecção entre o pixel e o elemento estruturante na posição central sobre ele não for vazia. Como neste caso trabalha-se com conjuntos booleanos, imagens binárias, o resultado da dilatação será equivalente ao resultado da aplicação de uma combinação lógica disjuntivas, uma operação *ou* não exclusiva, entre a vizinhança de  $x$  e o elemento estruturante. Na Figura 4 foram ilustradas as operações morfológicas de dilatação e erosão.

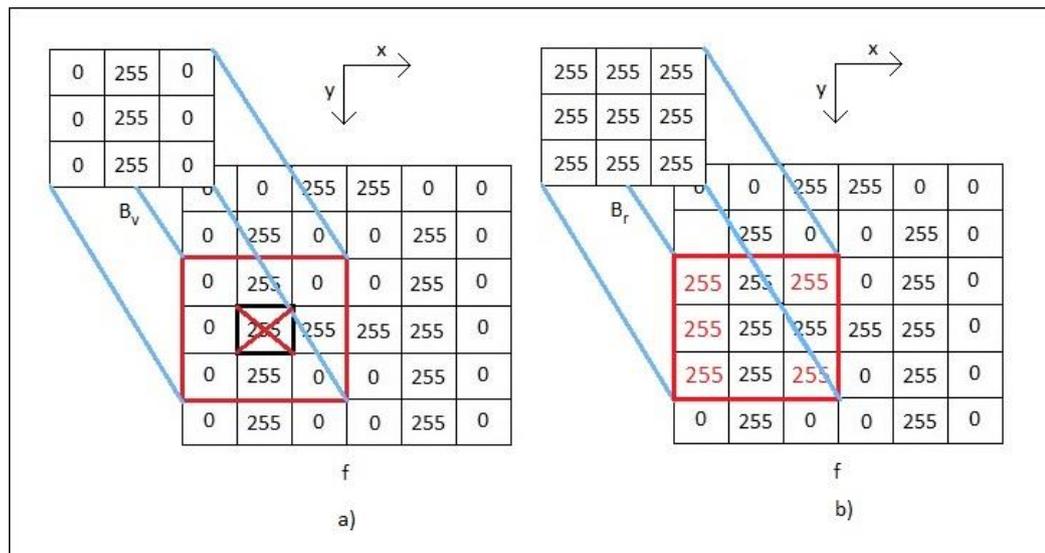


Figura 4 - Erosão e Dilatação

Legenda: Em a) esta representada um erosão feita por um elemento estruturante vertical. Em b) tem-se uma dilatação feita por um elemento estruturante retangular.

Fonte: Elaborada pelo autor.

## 2.3 - Identificação e reconhecimento de objetos através de contornos

Neste tópico serão abordados os principais conceitos a respeito da extração de características relacionadas ao contorno do objeto previamente segmentado e a respeito das técnicas utilizadas para o aprendizado dessas informações para um posterior reconhecimento.

### 2.3.1 - Contornos e suas características

A forma de um objeto está associada principalmente a disposição dos seus pontos delimitadores, que quando conectados formam seu contorno. Por meio da análise de um contorno várias informações podem ser extraídas, tais como altura, largura, perímetro, área, número de vértices, dentre outras, que servirão como características discriminativas para efetuar a identificação do objeto. Segundo Cesar e Costa (2009), a caracterização de um objeto através de sua forma (características do contorno), pode ser feita por meio de um vetor de características de tal forma que cada componente seja gerada por uma transformação  $T_i$  aplicada ao contorno. A Figura 5 ilustra essa caracterização.

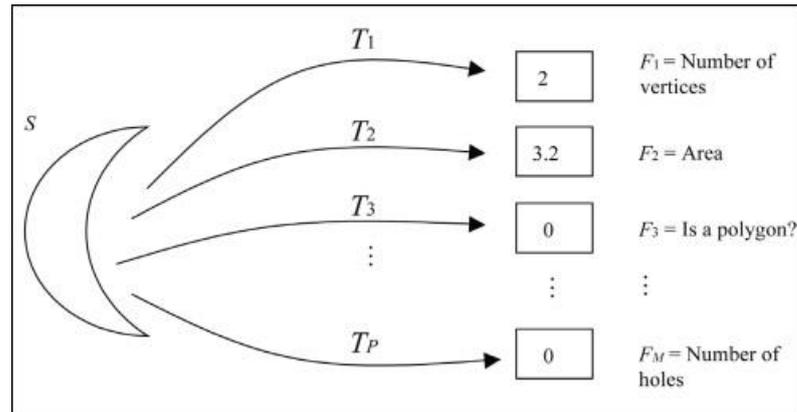


Figura 5 - Geração do vetor de características de S  
 Fonte: (COSTA; CESAR, p. 270, 2009).

### 2.3.2 - Descritores de Fourier

Por terem um papel importante no reconhecimento do objeto, as características, mencionadas anteriormente, devem ser o mais representativas possível e, conseqüentemente, serem também o mais invariáveis possível. A fim de tentar satisfazer principalmente essas condições, são utilizados então os Descritores de Fourier, descritores estes que surgem como resultado da aplicação da Transformada de Fourier aos pontos pertencentes ao contorno do objeto.

A Transformada de Fourier comumente utilizada é a contínua, que é aplicada sobre funções que não possuem descontinuidades em seu domínio. No entanto, em se tratando de uma imagem digital existem descontinuidades e por esse motivo utiliza-se então a Transformada Discreta de Fourier para obtenção dos descritores. Dito em Gonzalez e Woods (2000), a Transformada Discreta de Fourier aplicada a uma função  $f$  é expressa por,

$$F(u) = \frac{1}{N} \sum_{x=0}^{N-1} f(x) e^{-i2\pi ux/N} \quad (23)$$

para  $u = 0, 1, 2, \dots, N - 1$ .

O resultado da transformada será, na maioria das vezes, um valor pertencente ao conjunto dos números complexos, dado que o intuito principal da transformada é realizar somas de funções senóides e cossenóides no domínio dos números complexos. Para evidenciar melhor esse fato, tem-se pela representação de Euler,

$$e^{i\theta} = \cos\theta + i\sin\theta \quad (24)$$

para  $i = \sqrt{-1}$  e  $\theta = [0;360]$

Desta forma então é possível representar  $F$  no domínio dos números complexos de tal forma que,

$$F(u) = |F(u)|e^{\theta(u)} \quad (25)$$

$$\text{e } \theta(u) = \tan^{-1}\left(\frac{I(u)}{R(u)}\right) \quad (26)$$

lembrando que este é outro meio de representar  $F$  em termos de suas coordenadas polares. Para a equação 25,  $|F(u)|$  representa as magnitudes, também chamadas de *espectro de Fourier*, e  $\theta(u)$  representa o *ângulo de fase* da transformada. De posse desta representação, torna-se mais fácil compreender e aplicar as normalizações necessárias aos descritores, algo muito mais difícil de fazer ao se trabalhar com os pontos no domínio espacial. Dentre as normalizações que podem ser aplicadas estão as relacionadas à mudança de escala, translação e rotação.

Alterando um pouco o formato da equação 23, representando agora  $f(x)$  por  $z_p$ , onde  $z_p$  passa a ser o conjunto dos  $p$  pontos do contorno representados no domínio dos números complexos, tem-se as transformadas das mudanças de,

$$\text{translação } F(u) = \frac{1}{N} \sum_{p=0}^{N-1} (z_p + z) e^{-i2\pi up/N} \quad (27)$$

$$\text{escala } F(u) = \frac{1}{N} \sum_{p=0}^{N-1} (cz_p) e^{-i2\pi up/N} \quad (28)$$

para  $u = 0, 1, 2, \dots, N - 1$ .

Na equação 27 é possível observar que a translação acontece segundo a adição de um termo complexo a todos os pontos. Ao desmembrar a equação poder-se-á notar que a influência de  $z$  é a maior possível quando  $p = 0$ . Isso significa que para normalizar os descritores com relação à translação basta desconsiderar o descritor 0. Já na equação seguinte, a mudança de escala acontece segundo a multiplicação da transformada por uma constante  $c$ , o que torna possível a normalização através da escolha de um determinado descritor para ser o divisor de todos os outros. Como o descritor 0 não pode ser escolhido, geralmente escolhe-se o descritor 1 por ser o mais próximo. Por fim, a normalização com relação à rotação pode ser deduzida pela equação 25, onde percebe-se que a variação do ângulo  $\theta$  afeta somente a parte da equação não correspondente às magnitudes. Portanto, o *espectro de Fourier* já possui os valores normalizados para a rotação.

A Transformada Discreta de Fourier ainda demonstra um interessante comportamento, que torna possível, por meio de sua inversa, reconstruir o contorno de um objeto utilizando todos os descritores, para um resultado idêntico ao original, ou o total menos  $n$  descritores, para uma reconstrução aproximada. A equação 29 apresenta a Transformada Inversa Discreta de Fourier,

$$f(x) = \sum_{u=0}^{N-1} F(u) e^{i2\pi ux/N} \quad (29)$$

para  $x = 0, 1, 2, \dots, N - 1$ .

## 2.4 - Rede neural artificial e métodos de aprendizagem

As redes neurais estabeleceram dentro da computação, uma espécie de novo paradigma, isto porque, a necessidade de solucionar problemas cada vez mais complexos, como problemas de classificação e reconhecimento, exigiram a adoção de métodos diferentes dos algoritmos determinísticos até então utilizados. Este novo paradigma, procurou desenvolver algoritmos biologicamente inspirados, sendo que as redes neurais particularmente, foram inspiradas no modo de funcionamento do cérebro humano. Como definido por (BRAGA,LUDEMIR,CARVALHO; p. 1;2000)

RNAs são sistemas paralelos distribuídos compostos por unidades de processamento simples (nodos) que calculam determinadas funções matemáticas (normalmente não-lineares). Tais unidades são dispostas em uma ou mais camadas e interligadas por um grande número de conexões, geralmente unidirecionais. Na maioria dos modelos estas conexões estão associadas a pesos, os quais armazenam o conhecimento representado no modelo e servem para ponderar a entrada recebida por cada neurônio da rede.

Assim como o modelo biológico, o modelo artificial também necessita de um treinamento para estabelecimento do aprendizado. Este treinamento pode ser por supervisão, neste caso chamado de supervisionado, ou sem supervisão, chamado de não supervisionado. O treinamento supervisionado caracteriza-se pela atuação de um tutor no processo, capaz de informar as saídas esperadas e não esperadas para os dados treinados, a fim de, direcionar o aprendizado em função da minimização do erro entre o resultado oferecido pelo tutor e o resultado obtido pela rede.

O treinamento não supervisionado, como o próprio nome já diz, não necessita de um tutor ao longo do aprendizado. Os neurônios na rede neural tendem a se auto-organizarem, criando áreas classificatórias por meio de um processo adaptativo, onde de acordo com o grau de estímulo de determinada região neural diante das características apresentadas, torna-se possível realizar uma identificação. Este processo parece ser bem mais compatível com o processo natural de aprendizagem realizado pelo cérebro humano, no entanto, o modelo supervisionado é mais adotado, dada a maior estabilidade apresentada pelos seus resultados. Nas Figuras 6 e 7 estão ilustrados os dois modelos de aprendizado.

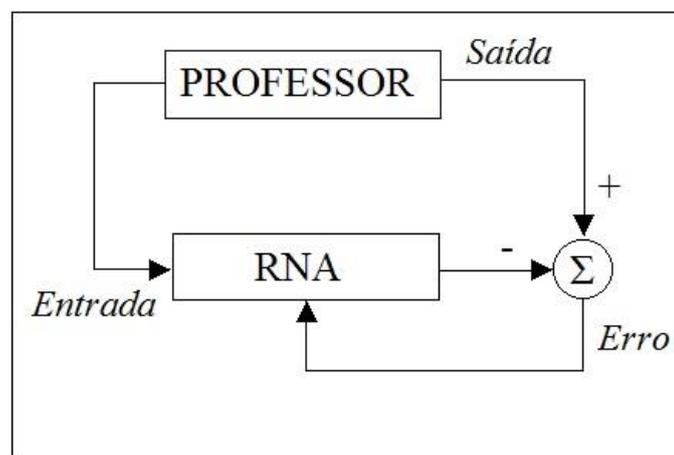


Figura 6 - Aprendizado supervisionado  
Fonte: (BRAGA; CARVALHO; LUDEMIR, p. 17, 2000)



Figura 7 - Aprendizado não-supervisionado  
 Fonte: (BRAGA; LUDEMIR; CARVALHO, p. 19, 2000)

### 2.4.1 - Perceptron

Nos idos de 1943, McCulloch e Pitts estabelecem um paralelo entre o sistema nervoso biológico e um possível sistema artificial. Suas teorias foram pioneiras com relação à criação de um modelo matemático para o funcionamento de um neurônio artificial quase que inteiramente baseado na lógica proposicional. Seguindo essa linha de estudo, Rosembat (1958) propõe um modelo estrutural para o neurônio artificial, o qual denomina de perceptron. Rosembat também estuda a capacidade de um perceptron realizar classificações corretamente após aprender determinadas características.

Segundo Braga, Carvalho e Ludemir (2000), para emular o funcionamento de um neurônio biológico, um perceptron utiliza pesos acoplados as entradas para simular as sinapses, sendo esses pesos positivos ou negativos, o que contribui respectivamente para estimular ou inibir o valor de entrada. O corpo celular é construído por uma função ponderadora dessas entradas juntamente com seus pesos sinápticos e por um valor de limiar, capaz de controlar a ativação do neurônio. O neurônio artificial está ilustrado na Figura 8.

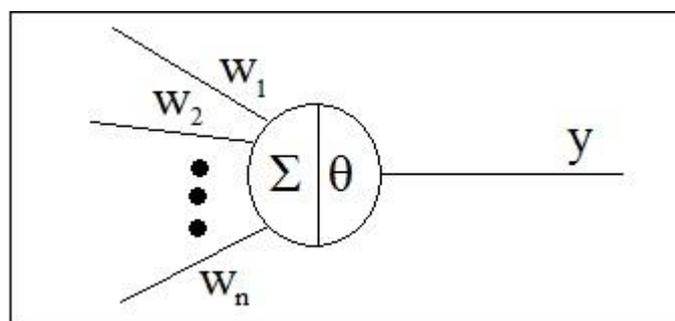


Figura 8 - Neurônio artificial  
 Fonte: (BRAGA; LUDEMIR; CARVALHO, p. 30, 2000)

### 2.4.2 - Rede neural perceptron multicamadas

A arquitetura de uma rede neural perceptron multicamadas baseia-se na utilização de vários perceptrons interconectados, de tal forma que por meio destas conexões seja possível ponderar um conjunto de entradas e obter uma resposta classificatória para um problema de interesse. Essa resposta é apresentada pela última camada da rede, a camada de saída, que possui uma quantidade  $t$  de neurônios que equivalem à quantidade de classes. A camada de entrada é composta por um vetor coluna contendo  $n$  linhas, referentes à quantidade de características a serem treinadas. A rede neural ainda contém  $m$  camadas intermediárias, denominadas de escondidas, sendo cada camada composta por um vetor coluna de tamanho variável, representados na figura 9 por  $p$  e  $q$ . Para melhor compreensão da arquitetura da rede neural ver Figura 9.

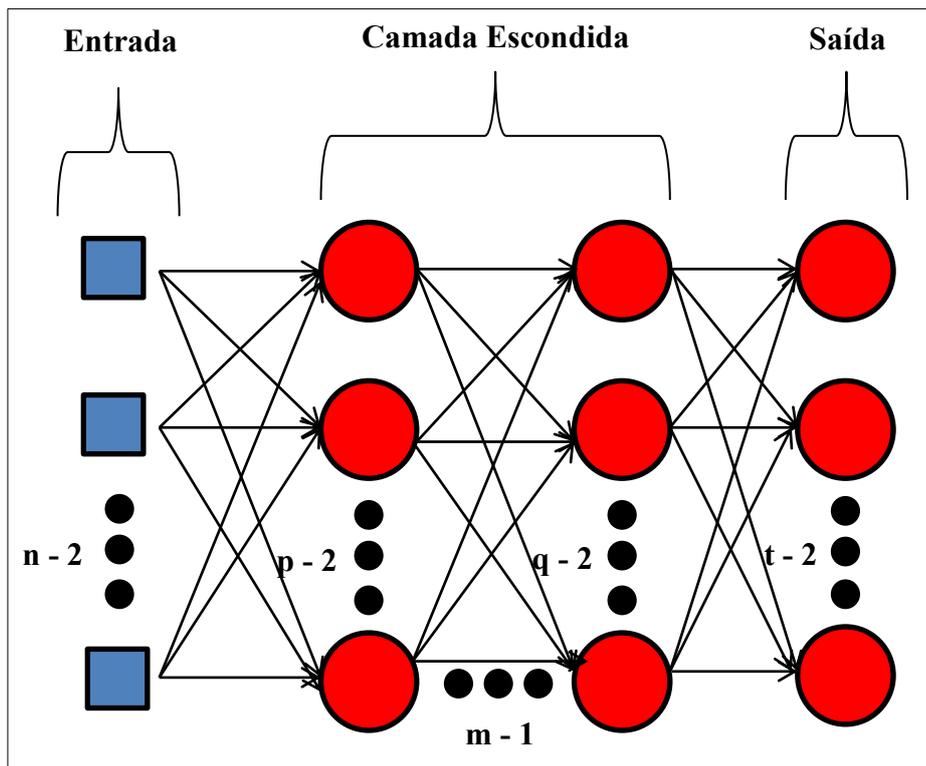


Figura 9 - Perceptron Multicamadas

Legenda: Arquitetura de uma rede neural perceptron multicamadas

Fonte: Elaborado pelo autor

Na Figura 9 é possível observar que o sentido de ativação da rede é unidirecional, sempre em direção à camada de saída. O mecanismo de disparo dos sinais segue sempre um mesmo padrão, onde para cada neurônio calcula-se o produto interno entre as saídas da

camada anterior, que chegam a ele, e os pesos existentes nestas conexões estabelecidas, sendo por fim aplicada uma função de ativação ao resultado obtido. As funções mais comumente usadas são a limiar, logística e tangente hiperbólica. Elas estão representadas abaixo respectivamente pelas funções  $f$ ,  $g$  e  $h$ .

$$f(x) = \begin{cases} 1, & p/x > 0 \\ 0, & p/x \leq 0 \end{cases} \quad (30)$$

$$g(x) = \frac{1}{1 + e^{-x}} \quad (31)$$

$$h(x) = \frac{e^{2x} - 1}{e^{2x} + 1} \quad (32)$$

Para ajudar a entender melhor esse processo, pode-se pensar na existência de uma rede neural com uma camada de entrada  $I$ , apenas uma camada intermediária  $H$  e a camada de saída. O valor de disparo de um único neurônio  $H_j$  da camada intermediária, conectado aos valores de entrada da camada  $I$  por um vetor de pesos  $W_{H_j}$ , fica definido por,

$$\phi(I \bullet W_{H_j}) \quad (33)$$

$$\text{sendo que } I \bullet W_{H_j} = \sum_{i=1}^n w_{i,j} I_i \quad (34)$$

### 2.4.3 - Aprendizado por retropropagação

O aprendizado por retropropagação é utilizado no treinamento supervisionado de uma rede neural perceptron multicamadas. Ele baseia-se na minimização de uma função de erro gerada pelo cálculo da diferença entre a saída obtida e a saída desejada para cada exemplo de treinamento apresentado a rede. A minimização ocorre de acordo com a propagação do erro pelos pesos das conexões estabelecidas entre os neurônios, sendo que, ao final do treinamento, deseja-se obter valores de pesos capazes de ponderar uma dada entrada

gerando para ela um valor de saída o mais próximo possível do valor que a classifique corretamente. Para propagar o erro pela rede utiliza-se uma fração do gradiente da função de erro, denominada taxa de aprendizagem. Também é possível adicionar uma fração do valor do peso anterior, denominada *momentum*, para reduzir as chances do gradiente atingir um mínimo local na função de erro e o aprendizado ficar estagnado. As imagens abaixo ilustram as superfícies de erro sem a influência do momentum e sob a influência dele.

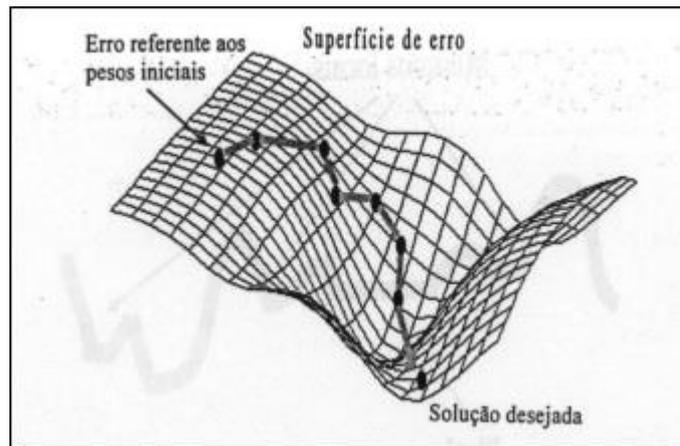


Figura 10 - Superfície de erro da Rede Neural  
Fonte: (BRAGA; LUDEMIR; CARVALHO, p. 62, 2000)

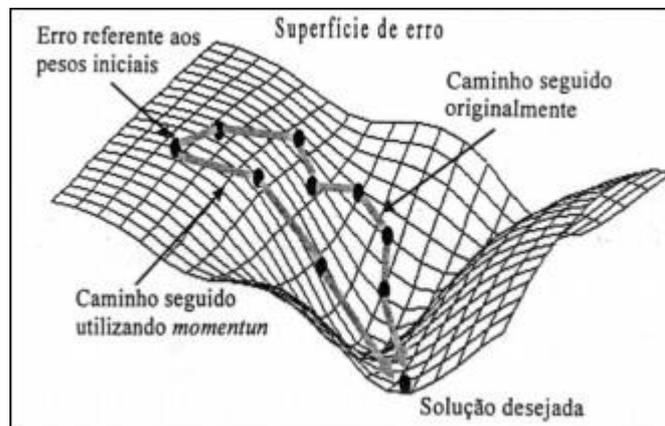


Figura 11 - Influência do termo momentum  
Fonte: (BRAGA; LUDEMIR; CARVALHO, p. 68, 2000)

O cálculo do gradiente é feito por meio da aplicação de derivadas parciais, processo que se inicia pela camada de saída da rede, onde obém-se o valor do erro por meio da equação,

$$E = \frac{1}{2} \sum_{j=1}^k (d_j - y_j)^2 \quad (35)$$

que mede o erro médio quadrático entre o valor esperado  $d$  e o valor de saída limiarizado por uma das funções  $f$ ,  $g$  ou  $h$ , exibidas logo acima, representadas na equação por  $y$ . Sendo assim, o gradiente neste primeiro momento fica representado por:

$$\frac{\partial E}{\partial w_{ji}} = (d_j - y_j) y'(s_j) \quad (36)$$

$$\text{sendo } s_j = \sum_{t=1}^n w_{jt} x_t \quad (37)$$

No segundo momento do cálculo do gradiente, o método de aprendizado passa a considerar os neurônios pertencentes às camadas escondidas da rede. Esta etapa compreende a análise de funções aninhadas, o que transforma a equação 36 em:

$$\frac{\partial E}{\partial w_{ji}} = y'(s_j) \sum_{t=1}^m \frac{\partial E}{\partial s_j} w_{jt} , \quad (38)$$

observando que a principal alteração diz respeito à adição do erro calculado sobre todas as funções que foram sendo aninhadas pelas camadas escondidas anteriores à do neurônio analisado. Tendo sido calculado o erro de aprendizado, a cada novo instante  $t$ , este valor é então utilizado para ajustar os pesos das conexões de acordo com a seguinte equação:

$$w_{ji}(t+1) = w_{ji}(t) + \Delta w_j(t), \quad (39)$$

$$\text{para } \Delta w_j(t) = \varepsilon \frac{\partial E}{\partial w_{ji}}(t) + \alpha \Delta w_j(t-1) \quad (40)$$

## 2.5 - Contagem volumétrica de veículos em vias

A contagem volumétrica visa estimar o volume  $N$  total de veículos que passam num dado ponto ou seção de uma via ou faixa da via durante um determinado período de tempo. Segundo (TMG, p. 3-1, 2001, tradução nossa):

“A medição do volume de tráfego é uma das funções mais básicas do planejamento e gerenciamento de estradas. A contagem do volume de tráfego provê as medições mais comumente empregadas do uso das estradas, sendo essas necessárias para a maioria das análises da engenharia de tráfego.”

Como mencionado anteriormente, a contagem volumétrica serve para auferir resultados fundamentais para a construção e manutenção das estradas de rodagem.

### 2.5.1 - Medidas baseadas na contagem de veículos

Neste tópico serão apresentadas as medições na área de engenharia de tráfego, mensuradas a partir dos dados obtidos pela contagem volumétrica, assim como, uma breve descrição de suas funções. Os dados compreendidos pelos tópicos que vão de 2.5.1.1 à 2.5.1.6 foram retirados de Ferraz e Bezerra (2009).

#### 2.5.1.1 - Fluxo ou vazão

Define-se fluxo ou vazão  $Q$  de uma corrente de tráfego, como sendo uma relação entre o volume de veículos  $N$  que passam num ponto da via (P ou O na Figura 12) e o intervalo de tempo correspondente  $T$  :

$$Q(\text{veic} / h) = \frac{N(\text{veic})}{T(h)} \quad (41)$$

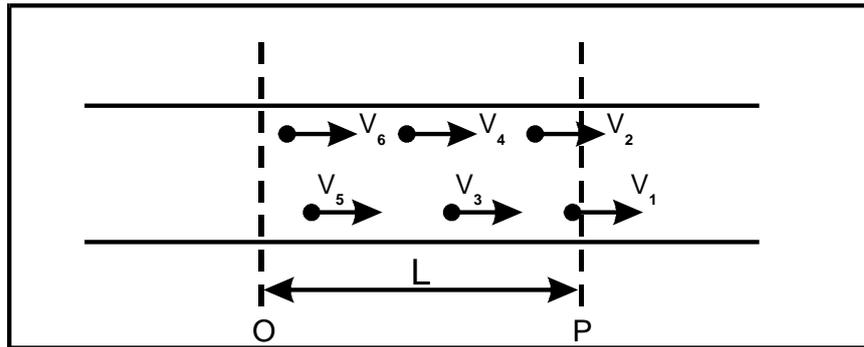


Figura 12 - Representação esquemática do fluxo de veículos  
 Fonte: (FERRAZ; BEZERRA, 2009)

### 2.5.1.2 - Densidade ou concentração

A densidade ou concentração  $K$  é dada pela relação entre a quantidade de veículos num trecho de via  $N$  e a extensão do mesmo  $L$ , conforme ilustrado na Figura 12.

$$K(\text{veic} / \text{km}) = \frac{N(\text{veic})}{L(\text{km})} \quad (42)$$

### 2.5.1.3 - Velocidade

Como os veículos numa mesma corrente de tráfego têm velocidades diferentes, a velocidade do fluxo deve ser caracterizada por um valor médio das velocidades individuais dos veículos. No estudo do tráfego de veículos a velocidade média de viagem é a mais utilizada, por duas razões: é a mais fácil de ser obtida no campo e apresenta uma relação direta com as outras duas variáveis caracterizadoras do tráfego, o fluxo e a densidade. A velocidade média de viagem  $V_v$  é dada pela relação entre a extensão do segmento de via considerado no cálculo ( $L$ : em geral entre 50 e 100m para se obter resultados confiáveis) e o tempo médio que os veículos demoram para percorrê-lo ( $\bar{t}$ ), conforme a seguinte equação:

$$V_v = \frac{L}{\bar{t}}, \text{ sendo} \quad (43)$$

$$\bar{t} = \frac{\sum t_i}{N}, \text{ resulta:} \quad (44)$$

$$V_v = \frac{L}{\frac{\sum t_i}{N}} = \frac{N.L}{\sum t_i} \quad (45)$$

onde:  $t_i$  é o tempo para o veículo  $i$  percorrer o trecho  $L$  e  $N$  é o número de veículos na corrente.

#### 2.5.1.4 - Espaçamento

Denomina-se espaçamento ( $s$ ) a distância entre veículos sucessivos numa corrente de tráfego, medida normalmente de para-choque dianteiro a para-choque dianteiro. O espaçamento médio numa corrente de tráfego é dado pela relação,

$$S = \frac{\sum s_i}{N} \quad (46)$$

sendo o espaçamento médio igual ao inverso da densidade,

$$S(m/veic) = \frac{1}{K(veic/m)} \quad (47)$$

#### 2.5.1.5 - Intervalo

O intervalo ( $h$ ) é definido como o tempo que decorre entre a passagem dos para-choques dianteiros de veículos sucessivos num mesmo ponto da via. O intervalo médio de uma corrente de tráfego é dado pela expressão,

$$H = \frac{\sum h_i}{N} \quad (48)$$

sendo o intervalo médio igual ao inverso do fluxo,

$$H(\text{seg} / \text{veic}) = \frac{1}{Q(\text{veic} / \text{seg})} \quad (49)$$

O espaçamento e o intervalo são considerados como características microscópicas do fluxo de tráfego, pois são relacionadas com pares individuais de veículos e variam dentro da corrente. Quando consideradas de forma agregada (valores médios), essas grandezas dão origem aos parâmetros macroscópicos fluxo e densidade.

### 2.5.1.6 - Fluxo de saturação

De acordo com Spigolon (2010), outra medida que pode ser calculada a partir de dados obtidos pela contagem volumétrica de veículos, é o fluxo de saturação. Essa medida está diretamente relacionada ao fluxo máximo de veículos por uma via que possua semáforos, durante o intervalo de tempo em que a luz esteja verde. Não sendo consideradas interrupções momentâneas na faixa analisada, (SPIGOLON; p. 26; 2010) diz ainda, à repeito de quais estimativas influem no cálculo do fluxo de saturação, que “[...]o fluxo de saturação depende dos seguintes fatores: largura da faixa de tráfego, inclinação longitudinal da via, porcentagem de veículos pesados (ônibus e caminhões) e porcentagem de veículos que fazem conversão à direita e/ou à esquerda.”.

Considerando o valor  $h$  para o intervalo (explicado em 3.1.5) previamente calculado, o fluxo de saturação  $s$  é dado pela seguinte expressão,

$$s = \frac{3.600}{h} \quad (50)$$

## 2.6 - Trabalhos correlatos

Em seu trabalho, Santos (2008) desenvolve um contador classificatório de veículos baseado em segmentação adaptativa, utilizando o algoritmo NHD<sup>3</sup> na etapa de segmentação de imagens e descritores de Fourier juntamente com redes neurais perceptron, para identificar os veículos e contabilizá-los. A Figura 13 ilustra as áreas segmentadas pelo algoritmo NHD.

---

<sup>3</sup> Algoritmo para segmentação adaptativa de imagens em tempo real utilizando sistema de cores Y'C<sub>b</sub>C<sub>r</sub>. Maiores esclarecimentos em Butler et al (2005, p. 2).

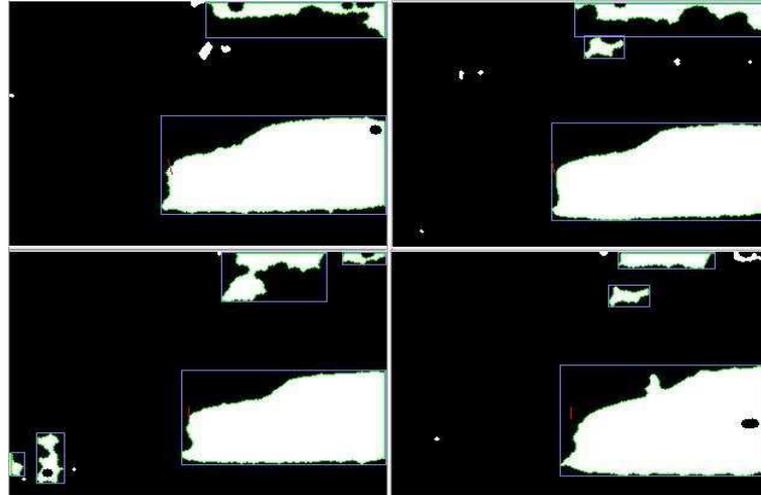


Figura 13 - Exemplos de objetos aprovados pelo algoritmo NHD  
 Fonte: (SANTOS; p. 59; 2008)

Em Feitosa (2012) é desenvolvido um método de contagem volumétrica automatizado utilizando visão computacional e faz-se um estudo da adaptação do mesmo para dispositivos móveis. O estudo utiliza técnicas de subtração de fundo e limiarização de imagem para segmentar os veículos e efetuar a contagem. A segmentação é ilustrada abaixo.

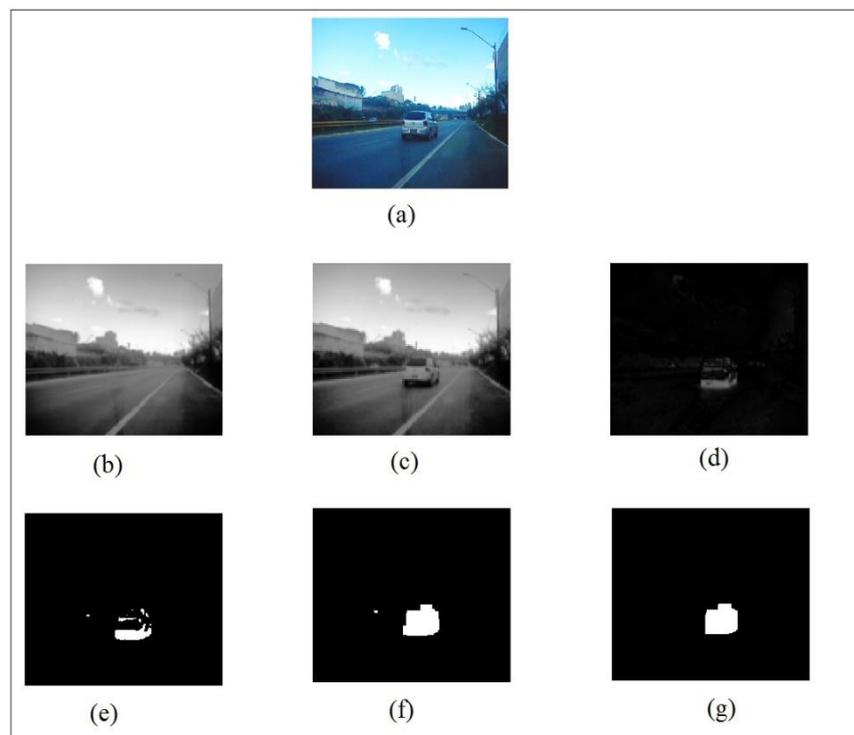


Figura 14 - Segmentação por subtração de fundo  
 Legenda: (a) é o quadro de entrada atual,  
 (b) a (d) são imgback, img fore e imgdiff ,  
 (e) a (g) mostra o processo de separação de objetos do fundo.  
 Fonte: (FEITOSA; p. 88; 2012).

### **3 - Estrutura do programa contador**

Para construção do programa de contagem veicular fora utilizada uma estratégia de divisão por módulos, a fim de facilitar o gerenciamento e a análise dos resultados gerados. A contagem acontece efetivamente após a imagem ter passado por três fases de processamento, segmentação, rastreamento e reconhecimento. Além da modularização do processo foi desenvolvido também um mecanismo de controle, criado para interligar os módulos e gerar o resultado final da contagem. A construção dos módulos será abordada ao longo dos tópicos 3.1 à 3.3, sendo que, o mecanismo de controle será abordado juntamente com o módulo de rastreamento 3.2.

Ao longo do trabalho foram utilizadas algumas palavras na tentativa de simplificar um conceito maior. Dentre elas esta a palavra máscara, que no caso fora utilizada para fazer referência à imagem binária, composta apenas pelos valores 0 e 255, gerada após a aplicação da segmentação. Outra convenção adotada foi à utilização da palavra objeto para fazer referência à região da máscara composta por pixels brancos, que antes do reconhecimento não pode ser chamada de veículo. A palavra quadro fora utilizada como um sinônimo para o termo imagem.

#### **3.1 - Módulo de segmentação**

A segmentação é o ponto de partida para realizar a contagem veicular. Isto porque é preciso, antes de tudo, separar o conteúdo de interesse a ser analisado, ou imagem de frente, do conteúdo irrelevante, dito também imagem de fundo. O processo de segmentação utilizado no programa é exibido no fluxograma da Figura 15.

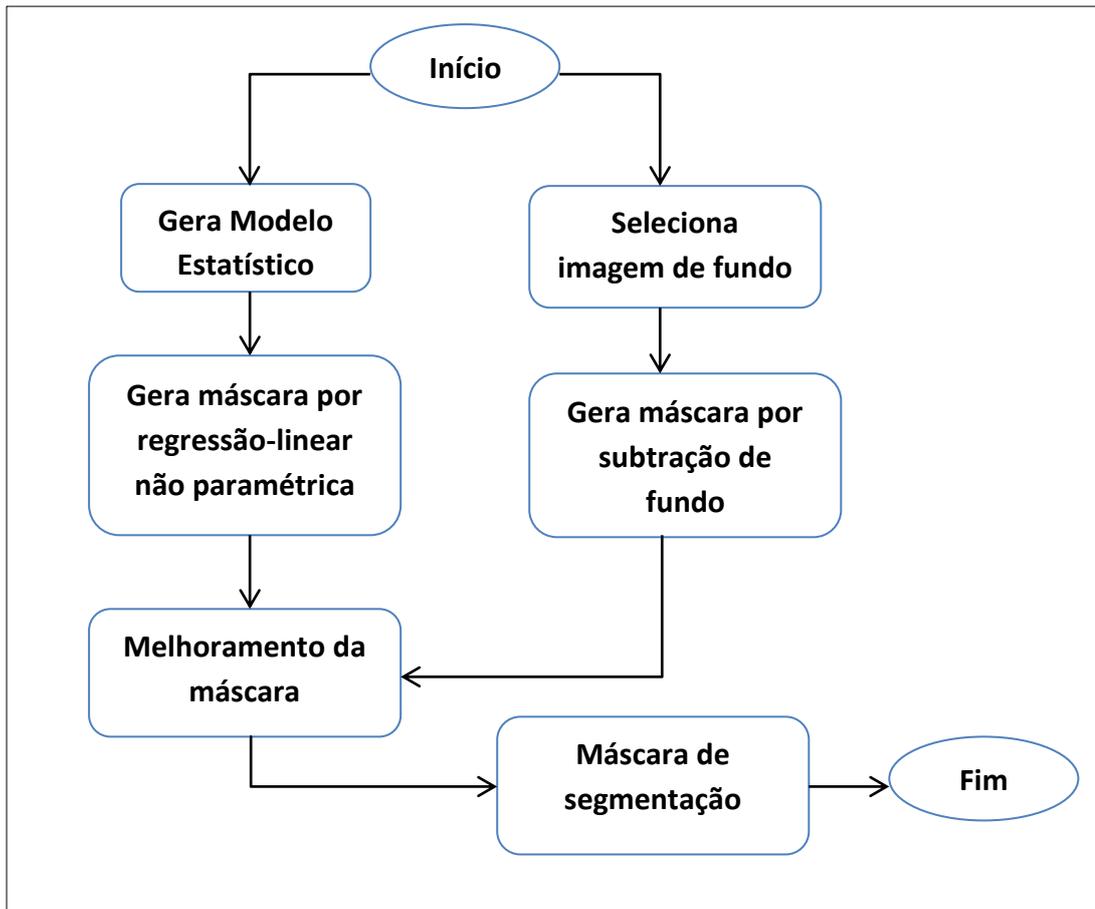


Figura 15 - Fluxograma do processo de segmentação

Legenda: Esquema de funcionamento do Módulo de Segmentação.

Fonte: Elaborado pelo autor

O fluxograma possui duas ramificações partindo do seu estado inicial, apesar de as mesmas não serem executadas de forma concorrente. Estas ramificações servem para evidenciar a existência de dois momentos distintos na etapa de segmentação. Em um primeiro momento, ramificação esquerda do fluxograma da Figura 15, gera-se uma primeira máscara utilizando o método estatístico explicado nos tópicos 2.2.1 e 2.2.2. Em um segundo momento, ramificação direita do fluxograma da Figura 15, gera-se uma segunda máscara por meio de uma segmentação por subtração de fundo. Ao final tem-se como resultado a *Máscara de segmentação*, obtida pela combinação entre as duas máscaras geradas inicialmente. As etapas *Gera Modelo Estatístico* e *Seleciona imagem de fundo* são realizadas apenas no início do programa, enquanto que no decorrer do mesmo, os dados armazenados por estas etapas vão sendo utilizados para segmentar cada novo quadro do vídeo. O principal objetivo da etapa *Gera Modelo Estatístico* é obter o valor das variâncias de cada píxel ao longo do tempo, além

de construir a matriz de permutação que será utilizada no processo de regressão linear não paramétrica do método estatístico de segmentação.

Na etapa *Selecciona imagem de fundo* é selecionado um quadro capaz de representar apenas o fundo da imagem, momento do vídeo sem a existência de veículos. Este quadro é escolhido em um instante  $t \in [1;1000]$ , com base no cálculo de diferenças absolutas entre os 1000 primeiros quadros capturados do vídeo e o valor das médias de pixels restantes nas imagens resultantes das subtrações. Desta forma, por meio do valor da menor média, foi possível estabelecer o quadro base para a segmentação por subtração de fundo. É preciso salientar também, que para evitar falsos positivos, no caso de calcular a diferença entre quadros muito próximos e o resultado do valor da média também ser baixo, utilizou-se uma estratégia de comparação. Esta estratégia envolveu trabalhar com subintervalos de 20 quadros, sendo que para cada quadro pertencente a metade do subintervalo, a diferença fora calculada apenas entre o mesmo e os 5 primeiros e os 5 últimos quadros do subintervalo. Testes mostraram que esta distância de 5 quadros fora suficiente para evitar a maioria dos falsos positivos.

### **3.1.1 - Otimização da segmentação estatística**

A técnica de segmentação estatística por si só pode ser considerada uma técnica robusta, mas não uma técnica rápida o suficiente para ser aplicada sozinha para realizar a video-segmentação. Um dos fatores que mais justificam essa lentidão é a aplicação de uma matriz de permutação à imagem antes de realizar a regressão linear, já que operações matriciais costumam ter uma complexidade quadrática  $O(n^2)$ . A otimização realizada teve por objetivo reduzir ao máximo este atraso, levando também em consideração a necessidade de preservar a qualidade da máscara gerada. Para a diminuição do atraso, fora aplicada uma drástica redução a cada imagem que posteriormente seria segmentada pela técnica estatística. Para preservar, ou melhor dizendo, recuperar a forma original da área segmentada, utilizou-se de uma combinação lógica entre as máscaras geradas estatisticamente e por subtração de fundo. A combinação lógica utilizada fora uma operação de união de conjuntos, sendo que posteriormente, ao resultado obtido, foram aplicadas operações morfológicas para melhorar ainda mais o resultado final.

As operações morfológicas, definidas em 2.2.3, foram aplicadas utilizando um *elemento estruturante*  $B$  do tipo retangular e de dimensões  $5 \times 5$ . Sendo assim, a combinação das máscaras pode ser formalizada da seguinte maneira,

$$\text{sendo } f^- = \varepsilon^B(f), \quad (51)$$

$$\text{tem-se que } f^+ = \delta^B(f^- \cup g), \quad (52)$$

observando que  $f$  e  $g$  representam respectivamente as máscaras obtidas pela segmentação estatística e segmentação por subtração de fundo. Além disso, o sinal das funções serve apenas para lembrar que a erosão tende a reduzir a dimensão do conjunto de píxels, enquanto que a dilatação tende a aumentar.

## 3.2 - Módulo de rastreamento

Realizar o rastreamento do objeto previamente segmentado na imagem foi de fundamental importância para o sistema contador. Por meio do rastreamento foi possível coletar informações referentes à posição e identificação do objeto ao longo do tempo, informações estas, que serviram como parâmetros de tomada de decisão no momento efetivo da contagem. O método de rastreamento desenvolvido neste trabalho teve por objetivo interligar os módulos de segmentação e reconhecimento de uma maneira ágil, evitando ao máximo gargalos de processamento. Sendo assim, optou-se por trabalhar com a informação relacionada à localização central do objeto utilizando filtro de Kalman e conceitos simples de geometria analítica. Nos tópicos seguintes serão descritos com mais detalhes os principais componentes deste módulo, assim como, os meios de controle utilizados.

### 3.2.1 - Mecanismo de rastreamento

O mecanismo de rastreamento foi construído tomando por base a existência de duas situações principais. Uma delas se resume à detecção do objeto na imagem pela primeira vez e a outra se resume à atualização dos estados armazenados para o objeto, estrutura que será explicada mais a frente, no tópico 3.2.3. O fluxograma da Figura 16 ilustra melhor o processo.

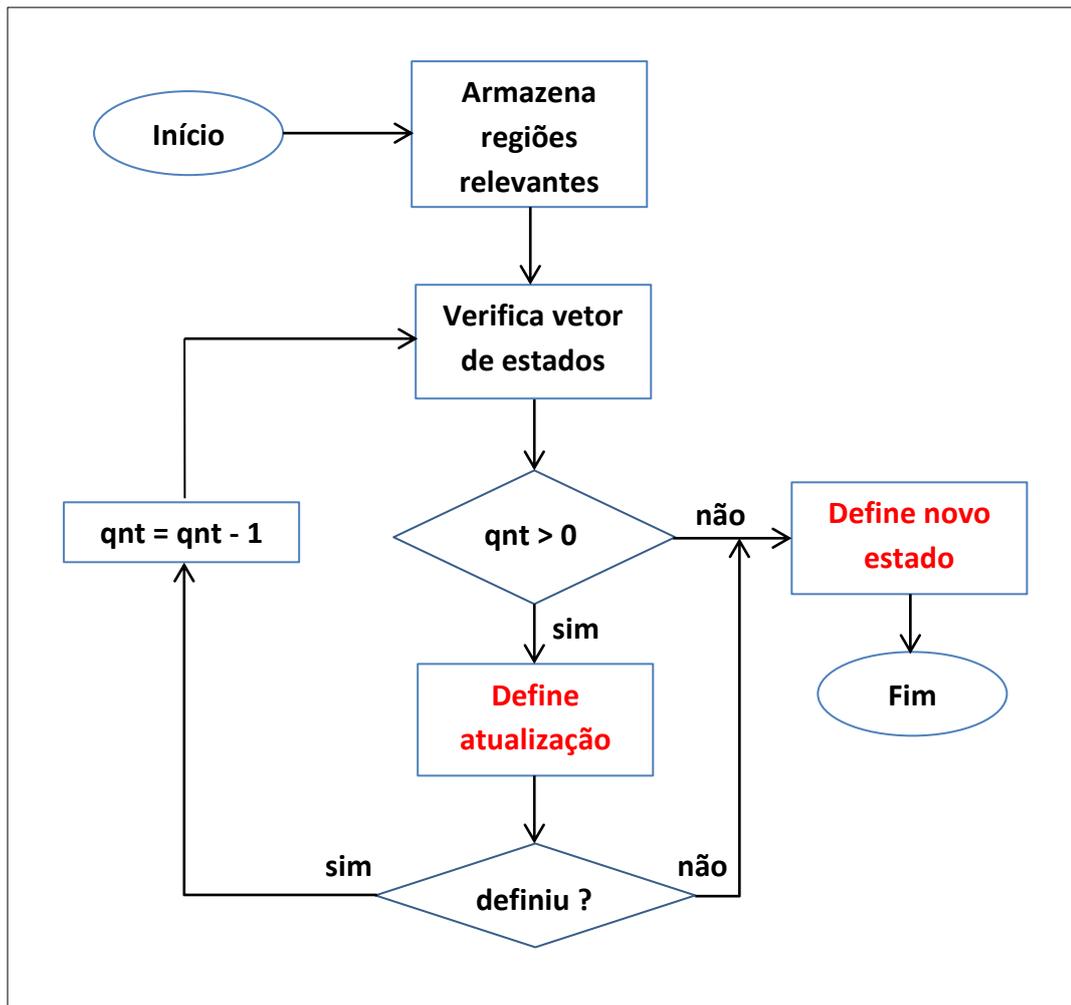


Figura 16 - Fluxograma de rastreamento

Legenda: Fluxo principal do módulo de rastreamento

Fonte: Elaborada pelo autor

Como ilustra o fluxograma da Figura 16, o módulo de rastreamento realiza inicialmente, na etapa *Armazena regiões relevantes*, uma checagem na máscara de segmentação a fim de armazenar os  $n$  conjuntos de objetos encontrados no instante  $t$  de observação, assim como, definir seus estados iniciais. Esses conjuntos serão utilizados na etapa seguinte, *Verifica vetor de estados*, onde é feita uma comparação entre a posição central da região delimitadora do objeto e o resultado previsto para esta posição, gerado pelo filtro de Kalman do objeto. Após a comparação ter sido efetuada por meio do cálculo de uma distância euclidiana entre as coordenadas, a próxima etapa, *Define atualização*, definirá, de acordo com o valor dado pela distância, se o objeto se moveu na imagem, ou então se outro objeto acaba de ser detectado. Esse processo continua enquanto existirem estados no vetor a serem verificados.

Para compreender melhor a função do filtro de Kalman, explicando de uma maneira mais prática e sucinta, ele serve para maximizar uma probabilidade a posteriori, futura, com base em dados de estado fornecidos pelo sistema sobre o qual ele atua e no erro envolvido na medição destes dados, vide tópicos 2.1.4 à 2.1.6 e também Bradsky e Kaehler (2008), para uma formalização maior do conceito. Neste trabalho, o filtro foi utilizado para prever as posições centrais dos objetos no instante  $t$ , sendo esta previsão baseada nos valores de estado armazenados previamente em um instante  $t - 1$ , para serem comparadas com as posições centrais referentes às regiões segmentadas pela máscara também no instante  $t$ . Os tópicos seguintes definem como fora determinada uma atualização de posição em virtude do resultado da comparação realizada, além da estrutura estado criada para armazenar toda a informação relevante da etapa de rastreamento e reconhecimento dos veículos.

### 3.2.2 - Atualização de posição

As Figuras 17 e 18, expandem a etapa *Define atualização*, presente no fluxograma da figura 16.

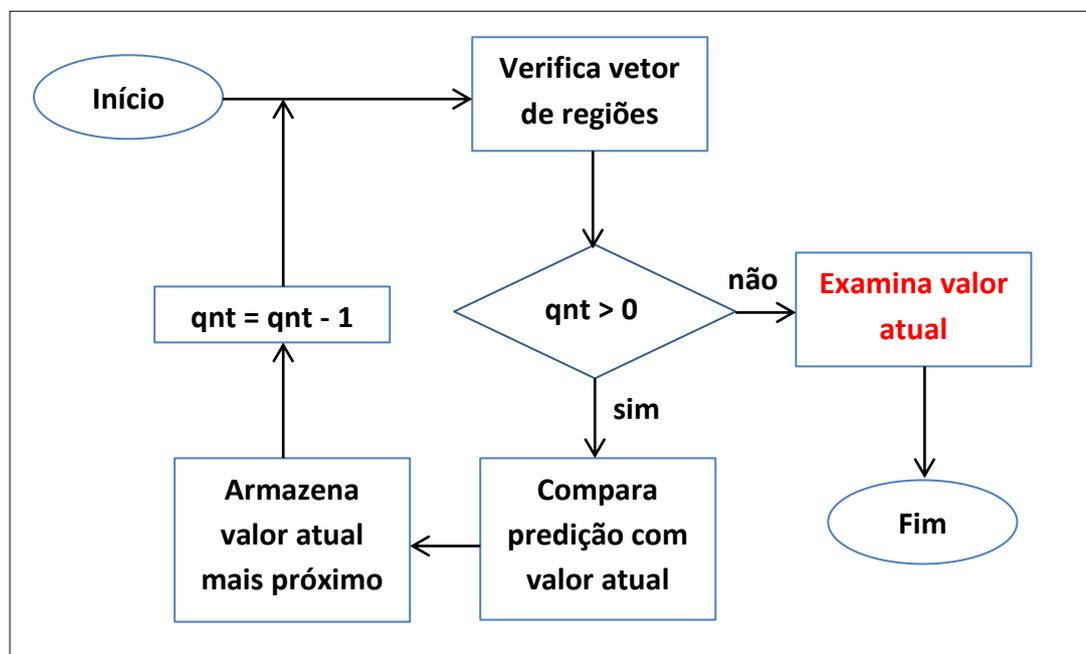


Figura 17- Primeira etapa de definição

Legenda: Fluxo da primeira etapa de definição de atualização

Fonte: Elaborada pelo autor

Nesta primeira fase de atualização de posição é realizado o cálculo das distâncias euclidianas entre as coordenadas dos centros das regiões dos objetos na máscara de

segmentação recém gerada e as coordenadas previstas pelo filtro de Kalman. Esse cálculo permite realizar associações, sendo que cada estado armazenado será atualizado pela nova informação contida na região associada. A associação é feita de acordo com a menor distância possível existente entre as coordenadas comparadas. Expandindo a etapa *Examina valor atual*, presente na Figura 17, obtem-se a segunda metade complementar do processo de atualização de posição, a fase *Examina valor atual*. A figura 18 ilustra esta etapa.

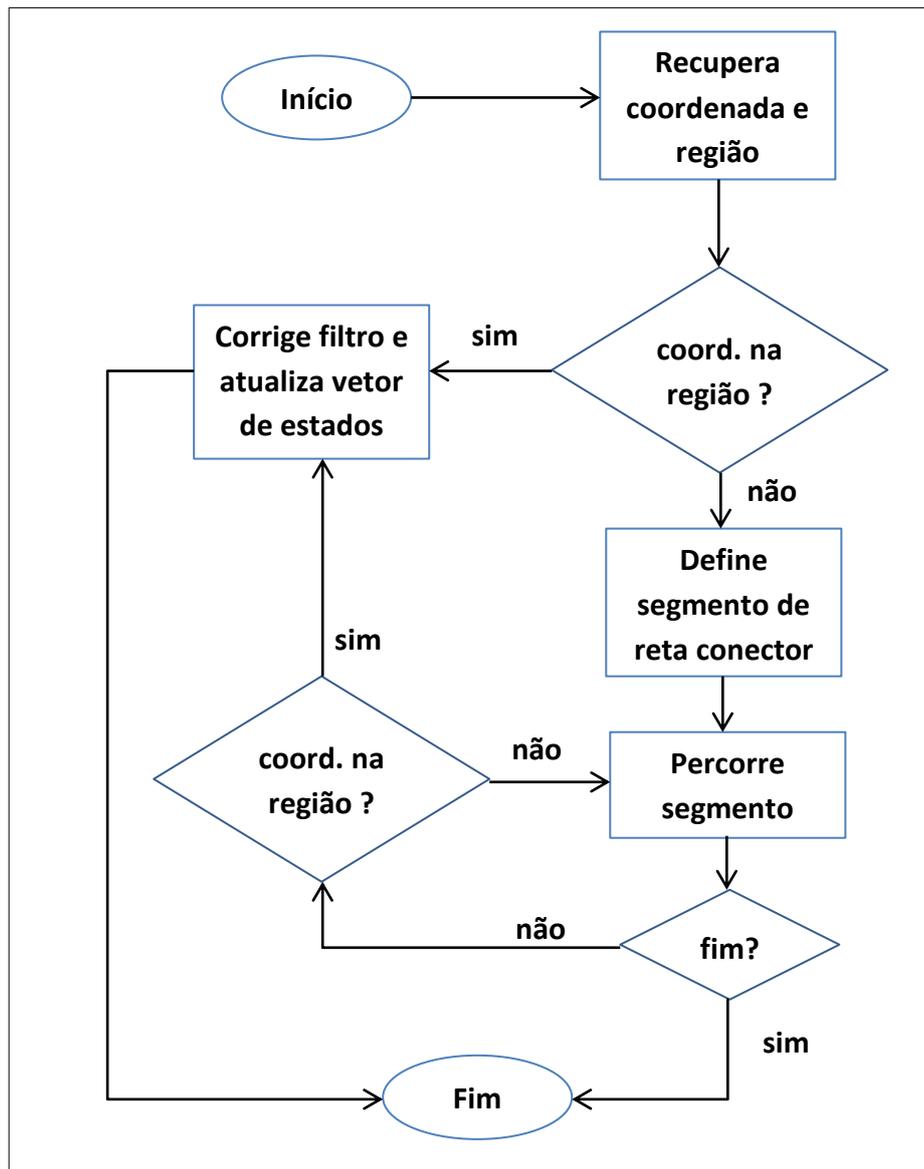


Figura 18 - Segunda etapa de definição

Legenda: Fluxo da segunda etapa de definição de atualização

Fonte: Elaborada pelo autor

Nesta segunda fase é realizada uma análise mais criteriosa da região escolhida para a associação. Primeiramente é verificado se a coordenada central prevista para o objeto, cujo

estado possivelmente será alterado, se encontra no interior da região a ele associada. Ao satisfazer essa condição, concluiu-se imediatamente que o estado do objeto deve ser redefinido pelos valores de estado na região associada. No entanto, em alguns casos, a coordenada prevista se encontra fora da região associada, o que leva ao teste realizado pelas etapas *Define segmento de reta conector* e *Percorre segmento*, presentes no fluxograma da Figura 19. Este teste se baseia na análise das coordenadas dos pontos pertencentes ao segmento de reta que interliga a coordenada prevista ao centro da região associada. Sendo assim, percorre-se o segmento de reta verificando se algum de seus pontos estaria contido em alguma outra região diferente da região escolhida para associação. Caso isto não aconteça, o método também atualiza o estado armazenado para o estado na região associada. A Figura 19 torna mais clara essa ideia.

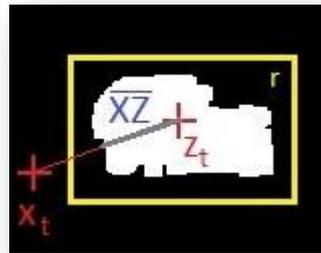


Figura 19 - Verificação de coordenada prevista

Legenda: Checagem da validade da predição utilizando o segmento de reta conectando o valor real  $z_t$  à predição  $x_t$

Fonte: Elaborada pelo autor

Na Figura 19, a parte acinzentada do segmento de reta  $\overline{XZ}$  indica que o estado  $x_t$  será atualizado para o estado  $z_t$  e, portanto, a região  $r$  passará a representar a nova localização do carro, neste caso em específico. Por meio desta mesma figura, é possível ainda observar a aplicação do filtro de Kalman de tal forma que,

$$x_t = \bar{x}_t + K_t(z_t - \bar{z}_t) \quad (53)$$

observando que o sinal de barra indica o valor da variável no momento de medição  $t - 1$ .

### 3.2.3 - Definição de estado

Nos tópicos anteriores, referentes ao módulo de rastreamento do objeto na imagem, muito se falou da utilização de estados, mas uma definição melhor para esta estrutura não fora dada. Para defini-la é preciso, em primeiro lugar, fazer uma distinção importante entre o estado utilizado pelo filtro de Kalman e a estrutura propriamente dita. O filtro utilizou o estado coordenada central do objeto segmentado, enquanto que a estrutura criada serviu para reunir uma quantidade maior de informações. Dentre elas estão as referentes à própria coordenada central do objeto, dados de reconhecimento e também de controle do sistema contador. A Tabela 1 exibe todas as informações consideradas sobre o estado do objeto.

Tabela 1 - Dados de estado

| <b>Dados de Posição</b>       | <b>Dados de Reconhecimento</b>        | <b>Dados de Controle</b> |
|-------------------------------|---------------------------------------|--------------------------|
| cantos da região delimitadora | n <sup>o</sup> carros reconhecidos    | qnt. de vezes rastreado  |
| centro da região delimitadora | n <sup>o</sup> caminhões reconhecidos | qnt. de vezes sem objeto |
| -                             | n <sup>o</sup> ônibus reconhecidos    | -                        |
| -                             | descritores                           | -                        |

Fonte: Elaborada pelo autor

A Tabela 1 apresenta duas colunas com informações ligadas diretamente ao rastreamento, que são as colunas *Dados de Posição* e *Dados de Controle*. Na coluna *Dados de Posição* armazenam-se as coordenadas principais da estrutura retangular, que define a região onde se encontra o objeto, os seus cantos. Outra informação também armazenada nesta coluna é o centro relativo do objeto dentro da região retangular, o estado diretamente ligado ao filtro de Kalman, assunto abordado logo no início deste tópico. A utilização do centro relativo e não do centro da região retangular delimitadora deveu-se ao fato de que em determinadas situações, como a existência de sombra ao redor do veículo, o centro da região delimitadora sofria um deslocamento  $dx$  e  $dy$  com relação às suas coordenadas  $x$  e  $y$  respectivamente. A Figura 20 ilustra melhor essa situação.

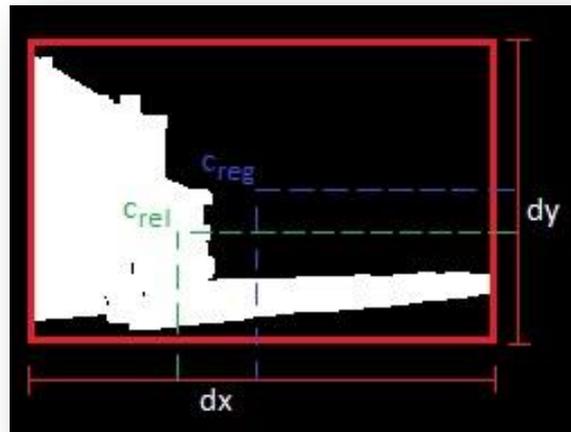


Figura 20 - Centro relativo

Legenda: Deslocamento do centro da região  $C_{reg}$ , em azul

na imagem, para o centro relativo  $C_{rel}$ , em verde na imagem.

Fonte: Elaborada pelo autor

A coluna *Dados de Controle*, por sua vez, serviu como base para consolidação da contagem. Para efetivar uma contagem, analisou-se primeiramente a quantidade de vezes em que um mesmo objeto fora identificado na imagem e posteriormente verificou-se a quantidade de vezes em que seu estado permaneceu inalterado durante um determinado período de tempo, em outras palavras, a quantidade de vezes em que a região delimitadora ficou vazia. Este controle visou otimizar a quantidade de memória utilizada para alocar os estados dos objetos e também reduzir o efeito de falsos positivos na contagem, casos em que áreas segmentadas na imagem não correspondiam à veículos.

### 3.3 - Módulo de reconhecimento

Após terem sido executadas as etapas de segmentação e rastreamento, descritas em 3.1 e 3.2 respectivamente, a última etapa realizada foi a de classificação do objeto. Toda essa etapa fora baseada na utilização de descritores de Fourier, assunto abordado em 2.3.2, e na utilização de redes neurais perceptron multicamadas, descritas em 2.4.2. A Figura 21 ilustra o processo.

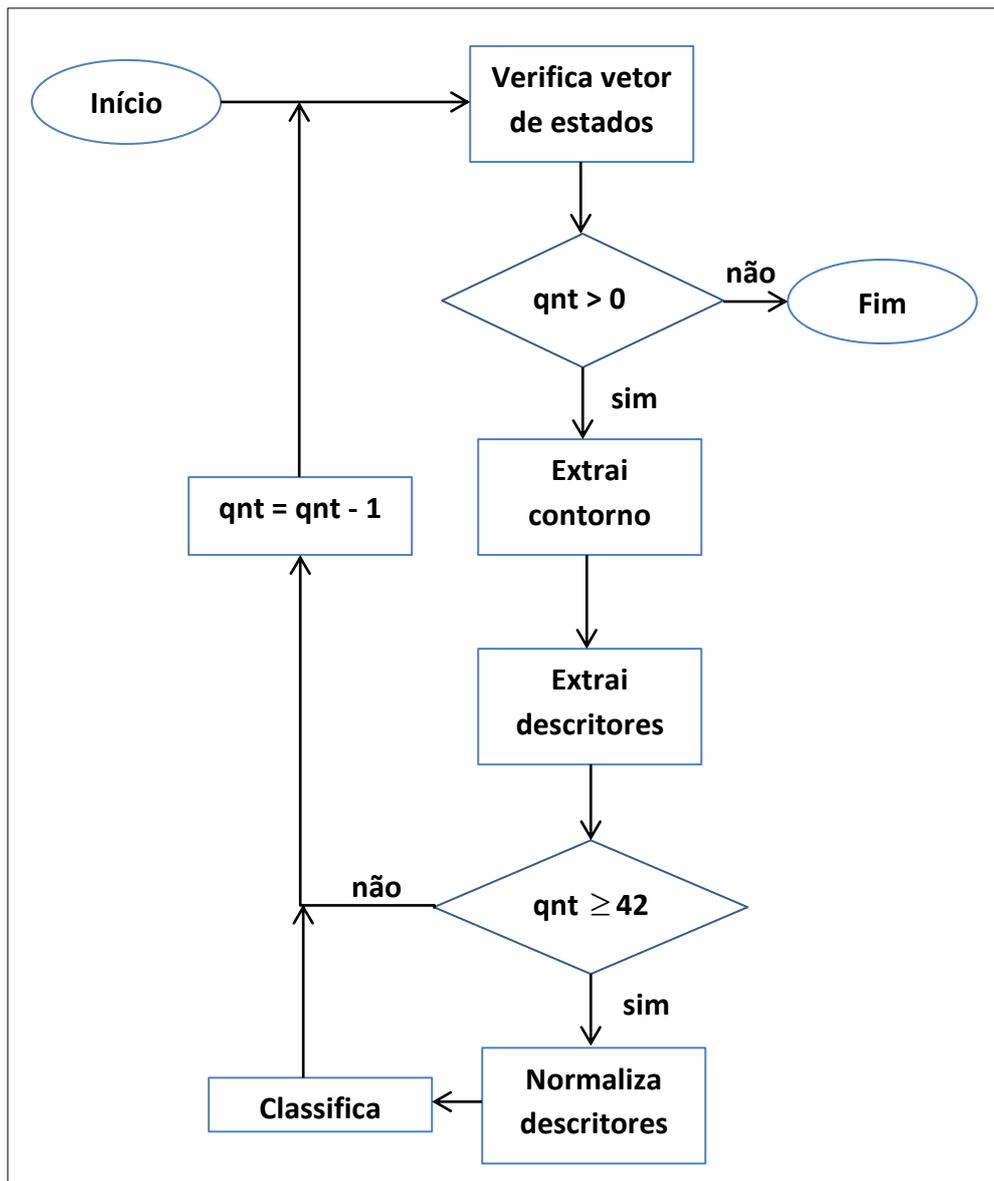


Figura 21 - Fluxograma de reconhecimento

Legenda: Fluxo prindipal do módulo de reconhecimento

Fonte: Elaborada pelo autor

Assim como os módulos anteriormente abordados, este módulo também se baseia na estrutura de estados. Neste caso especificamente, utiliza a informação contida na coluna *Dados de Reconhecimento* da Tabela 1, onde, para cada região armazenada, tem-se como objetivo recuperar apenas o contorno mais externo do objeto em seu interior, pois do mesmo é que serão extraídas as informações de entrada da rede neural, os descritores. O fluxograma da Figura 21 mostra ser necessário que o contorno gere no mínimo 42 descritores para prosseguir com a normalização. O porquê deste valor será explicado mais adiante no tópico 4.3. Após aplicar a normalização, seguindo o fluxograma, tem-se a etapa de classificação, responsável

por identificar o objeto e incrementar a quantidade de vezes da classe identificada na estrutura de estados. Para realizar a classificação, antes de tudo, foram estabelecidos os rótulos identificadores das classes de veículos consideradas para efeito de contagem neste trabalho. A Tabela 2 indica as correspondências realizadas.

Tabela 2 - Tabela de Classificação

| <b>Identificadores</b> | <b>Saídas</b> |   |   |
|------------------------|---------------|---|---|
| Carro                  | 1             | 0 | 0 |
| Caminhão               | 0             | 1 | 0 |
| Ônibus                 | 0             | 0 | 1 |

Fonte: Elaborado pelo autor

## 4 - Análise de parâmetros e resultados

Ao longo de todo este capítulo são discutidos os parâmetros definidos para algumas técnicas utilizadas no programa contador, assim como, suas relações com os resultados obtidos. Devido ao fato de cada módulo possuir suas particularidades, optou-se por abordar os principais resultados obtidos separadamente para cada módulo e ao final o resultado da contagem baseada na execução conjunta de todos eles.

### 4.1 - Análise da segmentação

Além da otimização, explicada em 3.1.1, que fora realizada nesta etapa, mais quatro parâmetros estatísticos, pertencentes às técnicas explicadas em 2.2.1 e 2.2.2, tiveram de ser definidos. O primeiro deles foi o tamanho do intervalo de tempo para realização do processo estocástico. Bevilaqua, Lanza e STEFANO (2005) apenas sugerem que este valor seja o menor possível. Após alguns testes, percebeu-se que um valor de 6 a 10 quadros seria o suficiente para extração das variâncias. Já para o segundo parâmetro, a cardinalidade dos conjuntos de imagens dos pixels estacionários, um valor de 3 foi suficiente para limiarizar o tamanho dos conjuntos.

A definição dos dois últimos parâmetros esteve diretamente relacionada à regressão linear não paramétrica do método proposto por Lanza e Stefano (2011), sendo que, o terceiro parâmetro determinado fora o tamanho da matriz de permutação e o quarto o limiar  $T$  capaz de satisfazer a equação 14. A utilização de uma matriz de permutação de tamanho  $3 \times 3$  ou  $5 \times 5$  não afetou significativamente o tempo de execução do processo de regressão linear não paramétrica. A principal diferença ao utilizar uma das duas opções esteve associada ao fato de que, ao utilizar uma matriz de tamanho  $3 \times 3$  gerou-se uma máscara com mais falhas no interior da região do objeto, enquanto que utilizando uma matriz de tamanho  $5 \times 5$  gerou-se uma segmentação mais homogênea. Mesmo assim, testes mostraram que nem sempre utilizar a matriz de dimensões maiores gerou melhores resultados. Em situações de segmentação de imagens trêmulas, por exemplo, foram gerados vários falsos positivos, muitos dos quais não puderam ser removidos da máscara mesmo aplicando um grande valor de  $T$ .

Por fim, a determinação do valor de limiar  $T$  foi deixada a critério do usuário, que pode modificar este valor enquanto o vídeo é processado, utilizando como referencial a própria máscara estatisticamente gerada. Na Tabela 3 é apresentado um resumo da análise.

Tabela 3 - Parâmetros da Segmentação Estatística

| <b>Imagem Trêmula</b> | <b>Tamanho do Intervalo</b> | <b>Cardinalidade do Conjunto</b> | <b>Matriz de Permutação</b> | <b>Limiar de Mudança</b> |
|-----------------------|-----------------------------|----------------------------------|-----------------------------|--------------------------|
| Sim                   | 6 - 10                      | 3                                | 3 X 3                       | Variável                 |
| Não                   | 6                           | 3                                | 5 X 5                       | Variável                 |

Fonte: Elaborado pelo autor

## 4.2 - Análise do rastreamento

Para fazer o rastreamento utilizou-se como base o filtro de Kalman, que fora introduzido de uma forma mais prática em 3.2.2 e 3.2.3. Ao utilizar o filtro, no entanto, algumas questões tiveram que ser solucionadas, como a determinação da dinâmica do movimento e o erro adicionado ao processo. A equação de movimento que melhor descreveu a trajetória do objeto foi a equação baseada em um MRU (Movimento Retilíneo e Uniforme). Como primeira tentativa fora utilizada uma equação de MUV (Movimento Uniformemente Variado), dado que os veículos em um ambiente real possuem também aceleração, mas os resultados não foram tão bons e essa idéia foi abandonada. Uma das prováveis explicações para esse teste não ter sido favorável, é a falta de um medidor externo que fosse capaz de gerar os dados de aceleração como variáveis de controle.

Para compor as matrizes de covariância relacionadas ao erro adicionado ao processo de predição, foram utilizados os valores dos desvios padrões das coordenadas  $x$  e  $y$  e das velocidades em  $x$  e  $y$ . Após diversos testes, determinou-se que os desvios padrões para as coordenadas de posição  $x$  e  $y$  seriam de 2.0 e que para as velocidades na direção  $x$  e  $y$  seriam de 1.5. Para comprovar a eficácia da utilização desses valores como parâmetros, apresenta-se abaixo uma imagem com as posições de deslocamento observadas e as previstas pelo filtro.

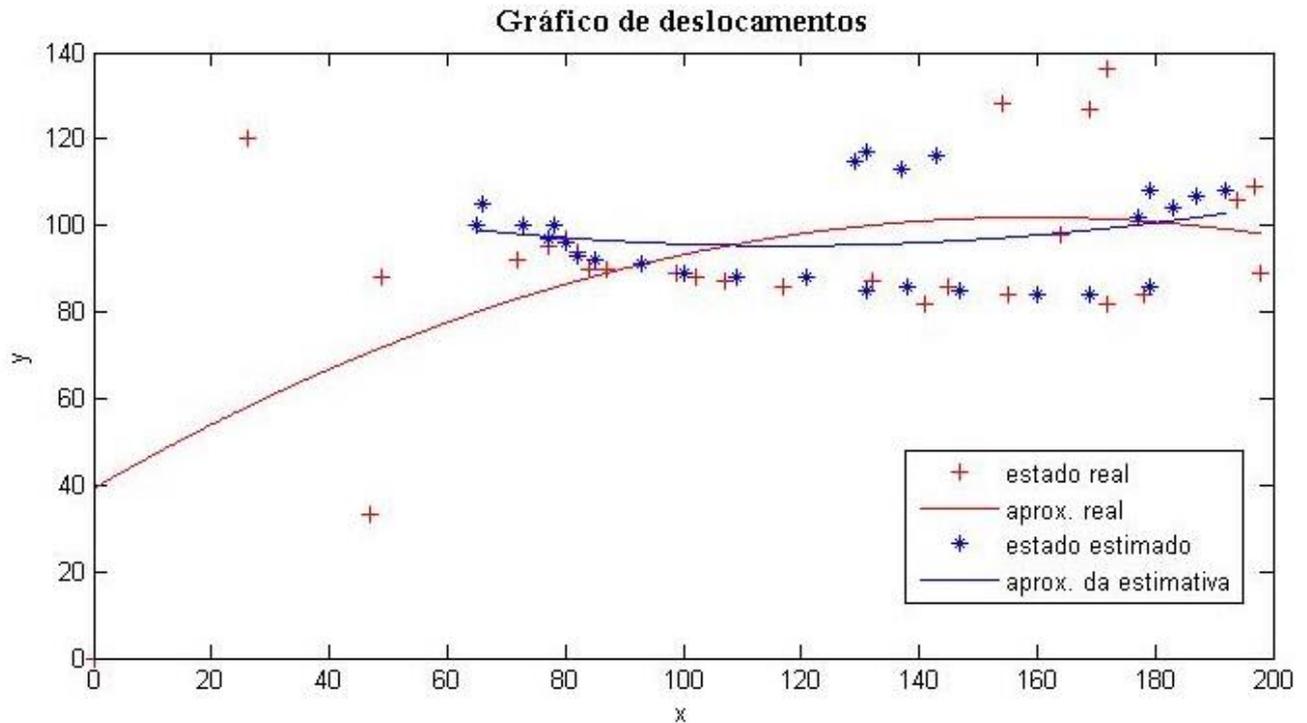


Figura 22 - Deslocamento de veículo

Legenda: Gráfico que mostra o deslocamento real de um veículo, em vermelho, e a previsão deste deslocamento, em azul, gerada pelo filtro de Kalman.

Fonte: Elaborado pelo autor.

### 4.3 - Análise do reconhecimento

Devido ao fato de ter sido utilizada uma rede neural perceptron multicamadas para fazer o reconhecimento dos objetos segmentados, foi necessário realizar o treinamento supervisionado por retropropagação para gerar o arquivo com os pesos das conexões. Por meio do resultado gerado por esse treinamento, foram sendo feitos os ajustes necessários para atingir a maior taxa de acerto possível pela rede neural. Como existiram vários parâmetros que tiveram que ser ajustados, algumas etapas foram seguidas para facilitar a visualização de todos eles. A primeira etapa foi definir a topologia da rede neural, sendo que a camada de entrada foi definida tomando por base a quantidade mínima de descritores capaz de gerar uma reconstrução aceitável dos contornos dos veículos mantendo um grau considerável de separabilidade entre as 3 classes aqui abordadas. Para determinação desta quantidade, utilizou-se de uma análise visual da reconstrução dos contornos de uma amostra de cada veículo e do gráfico das magnitudes dos descritores, dados exibidos nas Figuras 23 e 24. Chegou-se então a uma boa representação e um bom grau de separabilidade utilizando um total de 42 descritores, que após serem normalizados, onde desconsiderou-se os descritores 0

e 1 para efeito de normalização com relação à posição e escala respectivamente, se transformaram em um total de 40 descritores. As imagens abaixo ilustram a reconstrução do contorno dos veículos utilizando 40 descritores e também seus gráficos de magnitudes.

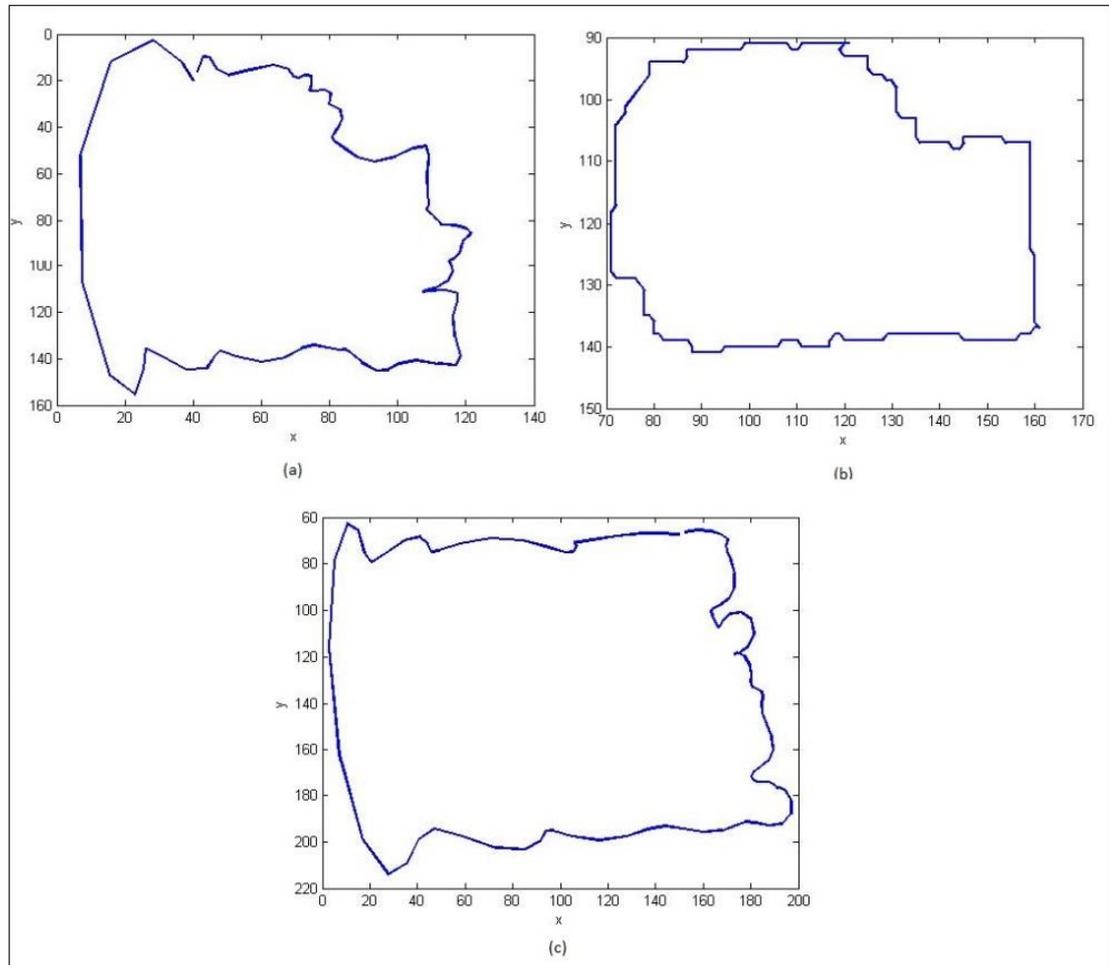


Figura 23 - Reconstrução de contornos

Legenda: Reconstrução dos contornos utilizando 40 descritores. Em (a) tem-se um caminhão, em (b) tem-se um carro e em (c) tem-se um ônibus.

Fonte: Elaborado pelo autor.

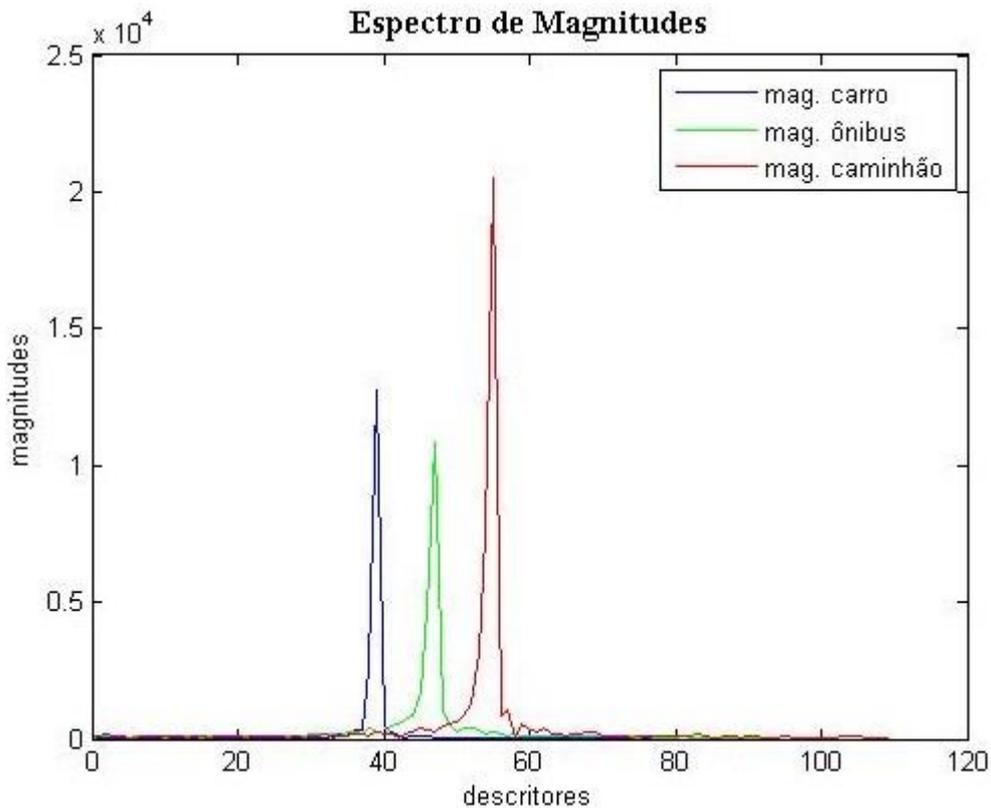


Figura 24 - Gráfico do Espectro de Magnitudes

Legenda: Plotagem das magnitudes dos 3 diferentes tipos de veículos

Fonte: Elaborado pelo autor.

Para construir a parte intermediária da rede, foram utilizadas duas camadas de neurônios, sendo a primeira composta por 40 e a segunda por 3 deles. Sendo que o motivo para utilização da segunda camada escondida foi gerar uma função de erro mais suave, evitando uma maior quantidade de mínimos locais indesejados, vide Figura 11. A camada de saída foi definida pelas classes de veículos existentes, como mostrado no tópico 3.3 na Tabela 2. A segunda etapa foi gerar os dados de treinamento, o que foi feito através da criação de um programa externo capaz de extrair os descritores e normalizá-los. De posse dos dados, como terceira etapa, realizou-se o treinamento e ajuste dos parâmetros da equação de aprendizado (39), no tópico 2.4.3, para atingir a melhor taxa de acerto possível. Utilizando uma quantidade de 35 vetores de entrada para cada classe, configurando um total de 105 amostras de treinamento e 9 amostras para teste, configurando um total de 27 amostras, observando que as amostras de teste não foram utilizadas na fase de treinamento, treinou-se a rede neural. A maior porcentagem de acertos fora atingida a uma taxa de aprendizado de 0.05 com a aplicação de um momentum de 0.02. Para efeitos de comparação utilizou-se também o

algoritmo *Resilient Propagation*, criado por Braun e Riedmiller (1993), na etapa de treinamento da rede, sem alterar seus parâmetros originais. A tabela mostra os resultados da análise.

Tabela 4 – Dados de desempenho do treinamento

| <b>Algoritmo</b> | $\epsilon$ | $\Delta$ | <b>Topologia</b> | <b>Norm.</b> | <b>Acerto (%)</b> |
|------------------|------------|----------|------------------|--------------|-------------------|
| Backprop.        | 0.05       | 0.02     | 40;70-3;3        | $Z_n$        | 55.55             |
| R. prop.         | -          | -        | 40;60-3;3        | $Z_n$        | 51.85             |
| Backprop.        | 0.05       | 0.02     | 40;70-3;3        | $Z_1$        | 44.44             |
| R. prop.         | -          | -        | 40;60-3;3        | $Z_1$        | 25.93             |

Fonte: Elaborado pelo autor

Como mostrado na Tabela 4, o fator que mais influenciou na porcentagem de acerto foi a normalização do espectro com relação à escala. Os melhores resultados foram obtidos utilizando para normalização o descritor  $Z_n$ , sendo este o descritor de maior magnitude escolhido dentre os descritores já normalizados com relação à translação.

#### 4.4 - Análise da contagem dos veículos

Para validar efetivamente o resultado da contagem, três vídeos foram utilizados para teste, sendo que, os vídeos foram escolhidos pelas características bem distintas que apresentaram, proporcionando testar a robustez e tempo de processamento do programa. O vídeo 1 apresentou-se o mais estável quando comparado com os vídeos 2 e 3, sendo esta estabilidade relativa a não existência de imagens trêmulas nas componentes do vídeo e a intensidade luminosa no ambiente pouco variável. No entanto, como característica principal, o vídeo apresentou muita sombra e esse fator foi útil para testar a robustez do Módulo de Rastreamento e do Módulo de Reconhecimento. Os vídeos 2 e 3 apresentaram muitas imagens trêmulas quando comparados com o vídeo 1. No entanto, diferentemente do vídeo 1, os veículos quase não exibiram sombras no seu entorno.

Outras duas características que diferenciaram os vídeos foram a questão da velocidade de deslocamento dos veículos na cena e a proximidade entre eles. Os vídeos 1 e 2 foram coletados de uma rodovia onde a velocidade máxima permitida para carros é de 110Km/h. Já o vídeo 3 foi coletado de uma via urbana, onde a velocidade máxima permitida é de 60Km/h. A influência do fator proximidade na porcentagem de erro de contagem pode

ser melhor observada utilizando o vídeo 3, onde o tráfego de veículos foi mais intenso no período de gravação do mesmo. A Figura 25 ilustra as imagens de fundo, obtidas como explicado no tópico 3.1, para melhor visualizar as diferenças entre os vídeos utilizados para o teste.

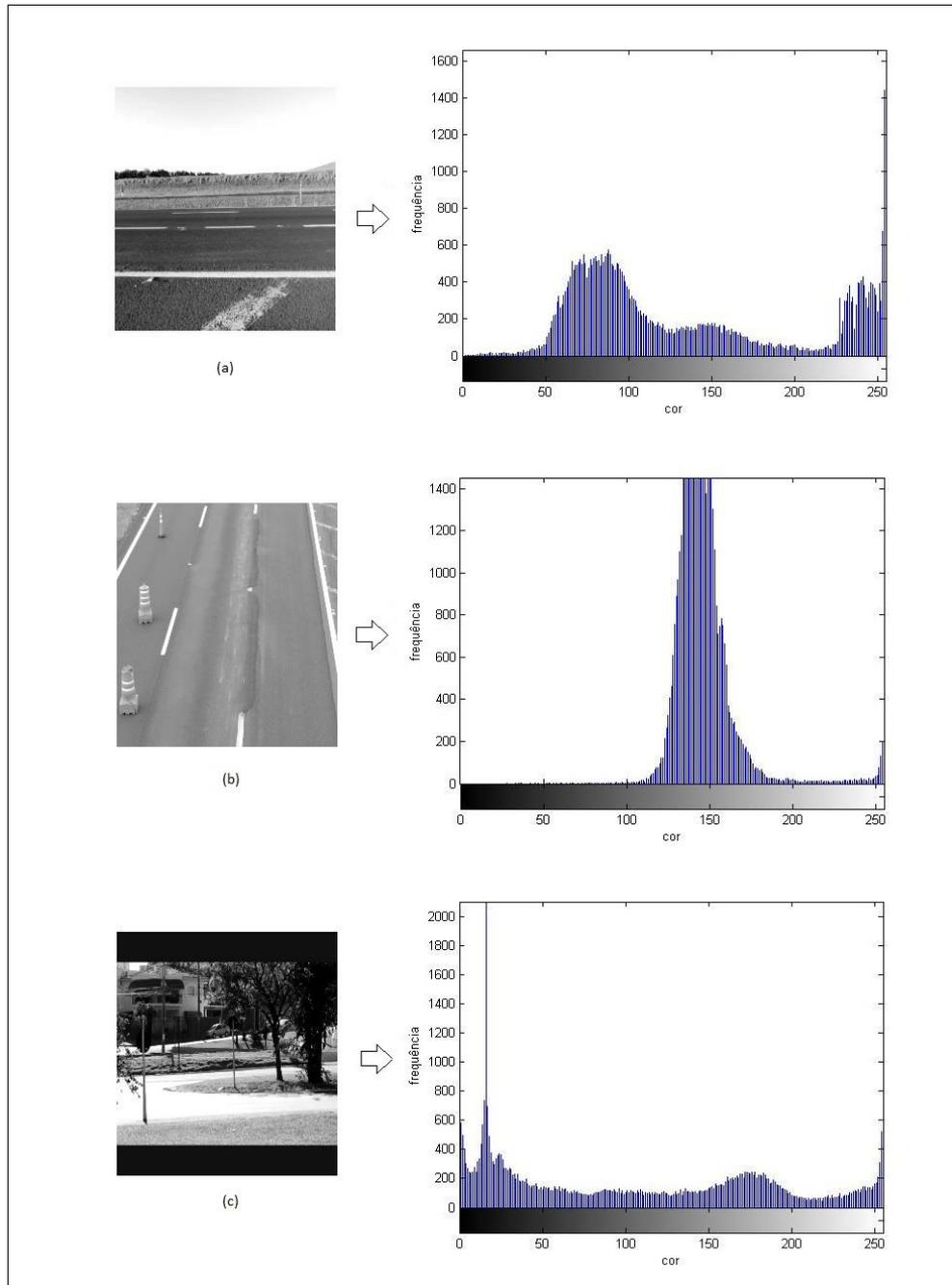


Figura 25 - Vídeos de teste

Legenda: Quadros de fundo acompanhados de seus respectivos histogramas, sendo que (a), (b) e (c) referem-se respectivamente aos vídeos 1, 2 e 3.

Fonte: Elaborado pelo autor.

Na tabela 5 são apresentados os resultados da contagem.

Tabela 5- Resultado do Contador

| <b>Vídeo</b> | <b>Horário de Medição</b> | <b>Dim. Matriz</b> | <b>T</b> | <b><math>\Delta t</math>(ms) modelo estatístico</b> | <b>Qnt. quadros</b> | <b>Fps médio</b> | <b>Acerto(%)</b> |
|--------------|---------------------------|--------------------|----------|---|---------------------|------------------|------------------|
| 1            | 17:00                     | 5 X 5              | 30       | 9080  | 6693                | 28.08            | 71               |
| 2            | 11:00                     | 5 X 5              | 20       | 9220  | 8477                | 28.14            | 72               |
| 3            | 10:00                     | 3 X 3              | 20       | 15351   | 9047                | 27.8             | 50               |

Fonte: Elaborado pelo autor.

A partir dos resultados apresentados pela Tabela 5, é possível ver que contador mostrou-se robusto ao ser aplicado em vídeos com características bem distintas, mantendo uma porcentagem de acerto média em torno de 70%, excetuando-se o terceiro experimento, que fora realizado sobre condições muito extremas. Estas condições estiveram relacionadas principalmente ao vento, que fez com que a câmera se movimentasse muito, mesmo estando fixada a um tripé, e ao ajuste de foco automático, que gerou, eventualmente, uma adição exagerada de ruído a imagem. Com relação à velocidade de execução, o programa mostrou-se veloz, pois mesmo sendo processado pelos módulos de segmentação, rastreamento e reconhecimento, o vídeo conseguiu ser exibido a uma taxa de quadros por segundo próxima a de uma exibição de tempo real, em torno de 28q/s.

## **5 - Conclusões e sugestões para trabalhos futuros**

Ao final deste trabalho foi possível obter um programa capaz de automatizar a tarefa de contagem de veículos como desejado inicialmente. O objetivo relacionado à adequação da técnica de segmentação estatística por regressão linear não paramétrica também fora atingido. Outros objetivos secundários também puderam ser alcançados, dentre eles a geração de uma técnica de rastreamento simples e bastante eficaz e um mecanismo de controle possível de ser adaptado a outros problemas de monitoramento. Mesmo tendo sido concluída a construção do programa contador, alguns pontos carentes de melhorias e outros possíveis de expansão ficam em aberto. Um dos pontos de melhoria estaria ligado à adição de um tratamento de oclusão e de sobreposição para os veículos, que poderia ser adaptado ao módulo de rastreamento. Outro ponto estaria ligado à utilização de outras técnicas mais robustas do que a segmentação por subtração de fundo, para serem combinadas com a segmentação por regressão linear não paramétrica na etapa de otimização da segmentação. Já a expansão do programa estaria relacionada à sua utilização para outros fins diferentes da contagem de veículos. Algumas modificações teriam de ser feitas, principalmente com relação ao Módulo de Reconhecimento, mas uma grande parte do programa poderia ser reaproveitada, como os módulos de segmentação e de rastreamento.

## Referências bibliográficas

- AL-KOFAHI O.; ANDRA S.; RADKE R .J.; ROYSAM B. **Image Change Detection Algorithms: A Systematic Survey**, v. 14, n. 3, p. 294-307, mar. 2005.
- BRAGA, Antônio de P.; CARVALHO, André; LUDEMIR, Teresa B. **Redes Neurais Artificiais: Teoria e Aplicações**. Rio de Janeiro: ETC, 2000.
- BRADSKY,G.; KAEHLER, A. **Learning OpenCV**. 1. ed. Estados Unidos da América: O'Reilly, 2008. 543 p.
- BUTLER, D.; JUNIOR, V. Michael B.; SRIDHARAN, S. **Real-time adaptive background segmentation**. Hindawi, 2005.
- CÉSAR JR, Roberto M.; COSTA, Luciano F. **Shape analysis and classification**. 2. ed. Boca Raton: CRC Press, 2009.
- DEPARTAMENTO NACIONAL DE INFRAESTRUTURA DE TRANSPORTES. Plano Nacional de Contagem de Trânsito. Disponível em: <<http://www.dnit.gov.br/rodovias/operacoes-rodoviaras/postos-de-contagem/plano-nacional-de-contagem-de-transito>>. Acesso em: 27 de mar. de 2013.
- FACON J. A Morfologia Matemática e suas Aplicações em Processamento de Imagens. In: VII WORKSHOP DE VISÃO COMPUTACIONAL - WVC 2011, UFPR, **Minicursos**, Maio 22-25, 2011. p. 64-65.
- FEITOSA, Francisco Coelho C. **Um estudo prático para contagem volumétrica automática de veículos usando Visão Computacional**. Goiânia-GO, 2012. 138p. Dissertação de Mestrado. Instituto de Informática, Universidade Federal de Goiás.
- FERRAZ, A. C. P.; BEZERRA, B. S. **Engenharia de Tráfego - Notas de Aula**. Escola de Engenharia de São Carlos, USP, 2009.
- FRIEDMAN J.; HASTIE T.; TIBSHIRANI R. **The Elements of Statistical Learning: Data Mining, Inference, and Prediction**. 2. ed. Califórnia: Springer-Verlag, 2008. 763 p.
- GONZALEZ, Rafael C.; WOODS, Richard E. **Processamento de imagens digitais**. Tradução de Roberto Marcondes Cesar Junior, Luciano da Fontoura Costa. São Paulo: Edgard Blücher, 2000.
- HIGHWAY CAPACITY MANUAL. Washington, TRB, 2010. Não paginado.
- JOHNSON, Richard A.; WICHERN, Dean W. **Applied Multivariate Statistical Analysis**. 6. ed. New Jersey: Pearson Prentice Hall, 2007.
- KOBAYASHI, Hisashi; MARK, Brian L.; TURIN, William. **Probability, Random Process, and Statistical Analysis**. New York: Cambridge University Press, 2012.

LANZA A.; STEFANO L. D. **Image Statistical Change Detection by the Pool Adjacent Violators Algorithm**, IEEE Transactions on Pattern Analysis and Machine Intelligence, v. 33, n. 9, p. 1894-1910, set. 2011.

LANZA A.; STEFANO L. D. Detecting Changes in Grey Level Sequences by ML Isotonic Regression. In: IEEE INTERNATIONAL CONFERENCE ON VIDEO AND SIGNAL BASED SURVEILLANCE (AVSS'06), Sydney(Austrália), **Proceedings**, Novembro 22-24, 2006. Não paginado.

MCCULLOCH, Warren S.; PITTS, Walter. **A Logical Calculus of the Ideas Immanent in Nervous Activity**, Bulletin of Mathematical Biophysics, v. 5, p. 115-133, 1943.

MOORE, David S.; MCCABE, George P.; CRAIG, Bruce A. **Introduction to the Practice of Statistics**. 6. Ed. New York: W. H. Freeman and Company, 2009.

NGAN, King N.; LI, Hongliang. **Video Segmentation and Its Applications**. Springer Science+Business Media, LLC, 2011.

PAPOULIS, Athanasios; PILLAI S. Unnikrishna. **Probability, Random Variables and Stochastic Process**. 4. ed. New York: McGraw-Hill, 2002.

ROSEMBLAT, Frank. **The Perceptron: A Probabilistic Model for Information Storage and Organization in the Brain**, Psychological Review, v. 65, n. 6, p. 286 – 408, 1958.

SANTOS, Daniel dos. **Sistema Óptico para Identificação de Veículos em Estradas**. Blumenau, 2008. Monografia de Graduação. Universidade Regional de Blumenau Centro de Ciências Exatas e Naturais.

SERRA J. **Image Analysis and Mathematical Morphology**. Academic Press, London, 1982.

SPIGOLON, Luciana M. G. **Semáforo: Grupo Focal Convencional X Grupo Focal com Indicação do Tempo de Verde/Vermelho Restante**. São Carlos, 2010. Dissertação de Mestrado. Universidade de São Paulo, Programa de Pós-graduação em Engenharia de transportes. Área de concentração: Planejamento e operações de Sistema de Transportes.

STIRZAKER, David. **Stochastic Process & Models**. New York: Oxford University Press Inc., 2005).

TAKEZAWA, Kunio; **Introduction to Nonparametric Regression**. New Jersey: John Wiley & Sons, Inc., 2006.

TMG, **Traffic Monitoring Guide**. In: Section 3, Traffic Volume Monitoring. USA: U.S. Department of Transportation, Federal Highway Administration, Office of Highway Policy Information, mai. 2001.

## APÊNDICE A - Diagrama de casos de uso

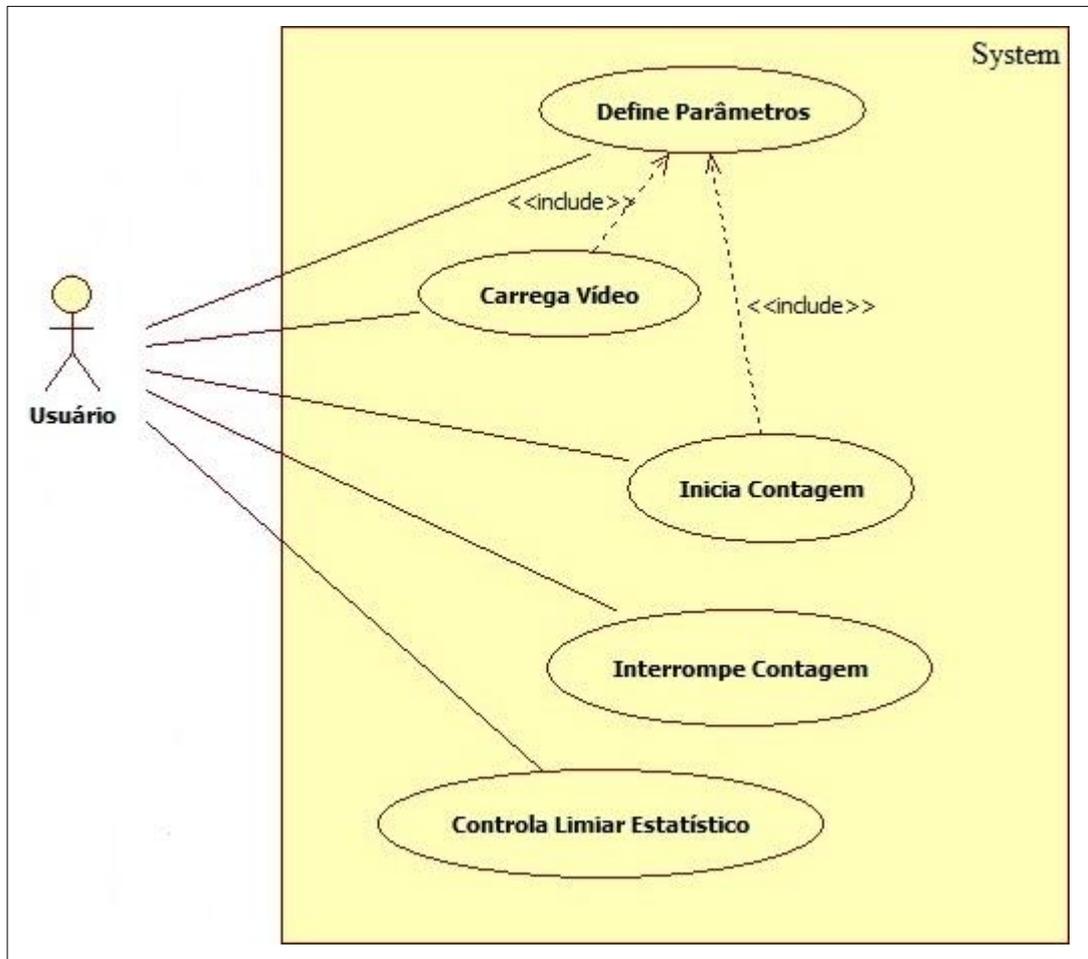


Figura 26 - Casos de uso do contador  
Fonte: Elaborado pelo autor.

## APÊNDICE B - Diagrama de classes

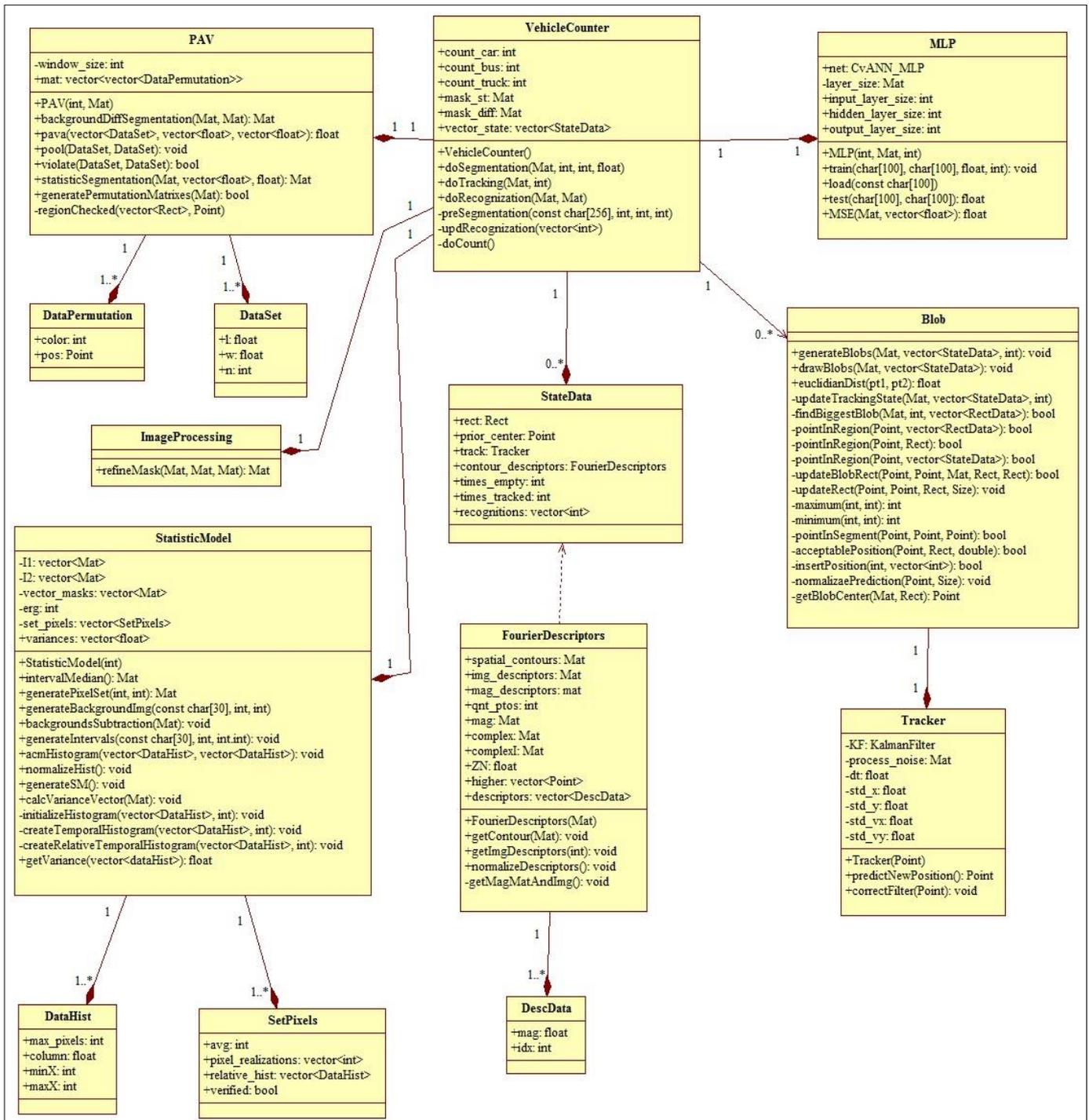


Figura 27 - Classes do programa contador  
Fonte: Elaborado pelo autor.

## APÊNDICE C - Interface do programa contador

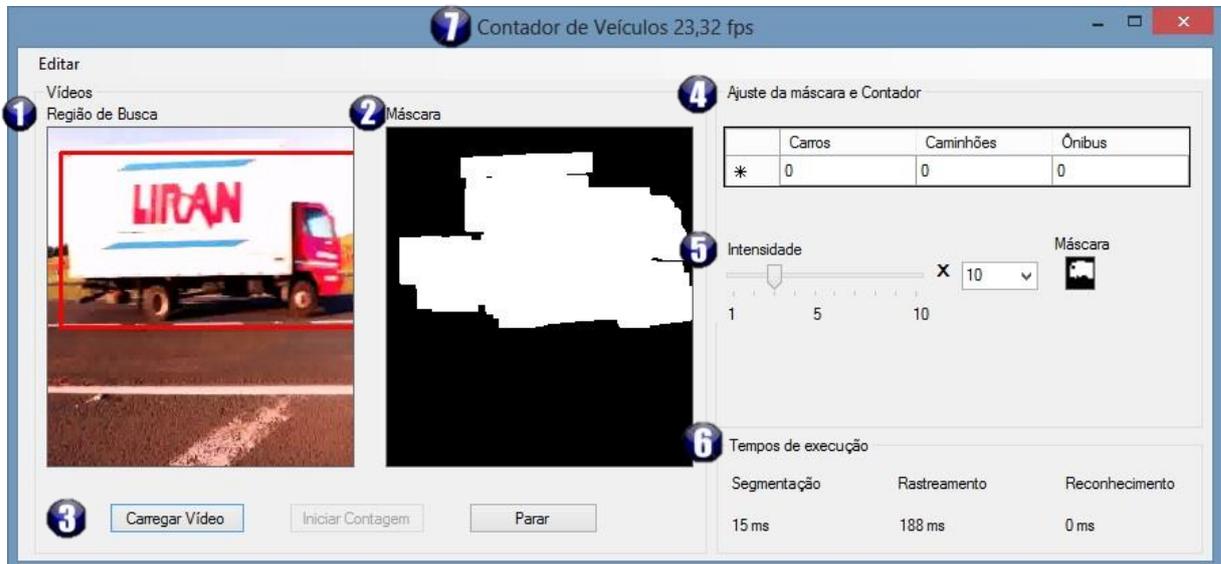


Figura 28 - Tela Principal do Contador  
Fonte: Elaborado pelo autor

- 1- A Região de Busca exibe o vídeo da via analisada contendo os veículos, além da região retangular delimitadora dos mesmos.
- 2- Nesta área é exibida a máscara resultante da combinação entre as máscaras geradas estatisticamente e por subtração de fundo.
- 3- Nesta área estão localizados os botões de controle do vídeo, onde primeiramente carrega-se o mesmo, posteriormente inicia-se o processo de contagem e por fim, é possível interromper o programa a qualquer momento através do botão Parar.
- 4- Área de exibição dos resultados da contagem classificatória.
- 5- Área de controle do limiar de segmentação estatística, controlado pela *track bar*, em função da máscara estatística exibida logo a frente da mesma.
- 6- Área de visualização do tempo de execução de cada módulo separadamente.
- 7- Taxa de exibição do vídeo contabilizando o atraso gerado pelo processamento.

## APÊNDICE D - Janela de definição de parâmetros

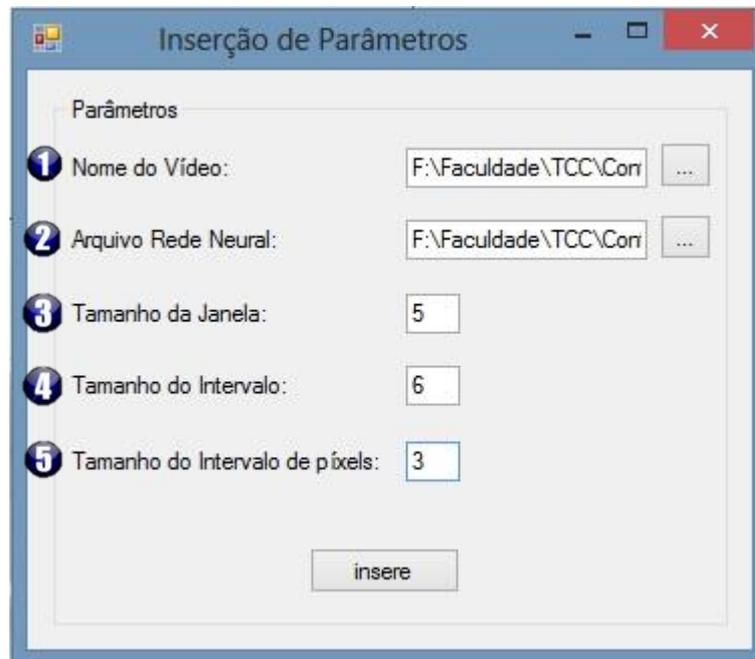


Figura 29 - Janela de Parâmetros

Fonte: Elaborado pelo autor.

- 1- Inserção do caminho até o vídeo.
- 2- Inserção do caminho até o arquivo de treinamento da rede neural.
- 3- Definição do tamanho da matriz quadrada de permutação.
- 4- Define o tamanho do intervalo de quadros para o processo estocástico.
- 5- Define a ergodicidade dos conjuntos para os pixels estacionários.