

Sistema de Medição de Velocidade de Veículos Utilizando Processamento de Imagem

1

Abstract. *Image processing techniques are increasingly present in vehicle traffic monitoring. This paper proposes an initial design of a vehicle speed meter using only a fixed camera and a computer. The background subtraction technique is used to generate a binary image. Morphological operations are performed to increase the success rate in tracking the vehicle and finally the speed calculation is done. The obtained results are compared with values of traditional meters in order to validate the method. The vehicle detection rate is 92,9% for speeds between 45 km/h and 60 km/h, and the average speed error is 7,2% for that same speed range.*

Resumo. *Técnicas de processamento de imagem estão cada vez mais presentes no monitoramento de tráfego de veículos. Este artigo propõe um projeto inicial de um medidor de velocidade de veículos utilizando-se de apenas uma câmera fixa e um computador. Primeiro utiliza-se da técnica de subtração de fundo para gerar uma imagem binária. Operações morfológicas são realizadas para aumentar as chances de sucesso no rastreamento do veículo e por fim é feito o cálculo da velocidade. Os resultados obtidos são comparados com valores de medidores tradicionais a fim de validar o método. A taxa de detecção de veículos ficou em 92,9% para velocidades entre 45 km/h e 60 km/h, e a média do erro percentual da velocidade é 7,2% para essa mesma faixa de velocidade.*

1. Introdução

Com o crescimento constante das pequenas e grandes cidades gerado principalmente pelo desenvolvimento econômico, uma maior parte da população foi capaz de conquistar um veículo próprio, portanto é natural o aumento do fluxo veicular nas estradas e vias dentro das cidades. Em virtude dessa expansão, métodos de controle de tráfego e velocidades são cada vez mais requisitados, possibilitando uma melhor análise das vias em geral. Esse aumento natural do número de automóveis proporciona um aumento significativo nos acidentes causados por alta velocidade. Segundo a Organização Mundial da Saúde (OMS), o número de mortes causadas por acidentes veiculares em 2016 foi de 1,2 milhões, sendo 90% desses acidentes em países de baixa ou média renda apesar de possuírem apenas metade dos veículos do mundo [WHO 2017].

Com esse cenário, surge uma grande necessidade de aprimoramento do controle de tráfego. Para controle de velocidade na vias, o dispositivo mais difundido trata-se de sensores magnéticos, no entanto, esses instrumentos são capazes apenas de realizar a contagem, medição de velocidade e identificar algumas infrações como avanço de semáforo e parada sobre a faixa de pedestre.

O aumento do poder de processamento e o desenvolvimento de algoritmos para processamento de imagens digitais, possibilita uma recente solução para o controle de tráfego através de câmeras de monitoramento. Esse tipo de solução é bem mais simplificado em termos de instalação, o que gera um baixo custo em seu investimento e manutenção se comparado aos medidores de laço indutivo [Alexandre 1997]. Além disso, com esse novo sistema, pode-se realizar todas as funções que um sistema de medição por laço indutivo já faz. Contudo, a utilização de processamento de imagem e visão computacional, ainda abre a possibilidade de aprimoramento dos algoritmos e adição de novas funções ao longo dos anos, sem a necessidade, na maioria dos casos, de troca de equipamentos.

Os radares tradicionais (sensores indutivos) apresentam diversos tipos de dificuldade em sua instalação e manutenção. Problemas como interferências magnéticas, deterioração do pavimento, dificuldade no ajuste de sensibilidade dos laços indutivos, possibilidade de burlar o sistema, entre outros. Porém, muitos desses inconvenientes poderiam ser evitados, por meio de um sistema mais simplificado, utilizando-se de apenas uma câmera de vídeo e aplicando técnicas de processamento de imagem e visão computacional.

2. Metodologia Proposta

Um arquivo de vídeo trata-se de uma sucessão de quadros (imagens) em um determinado período. Geralmente, a sensação natural de movimento é obtida quando tem-se 25 imagens sendo visualizadas em um período de 1 segundo. Será discutido mais a frente que, *Frames per second* (FPS) é um parâmetro importante no cálculo da velocidade. Sabendo disso, é importante ressaltar que as etapas de processamento de imagem serão feitas em cada quadro do vídeo. O método proposto é constituído pela etapas ilustradas na Figura 1. As etapas serão discutidas mais a fundo nas seções seguintes.

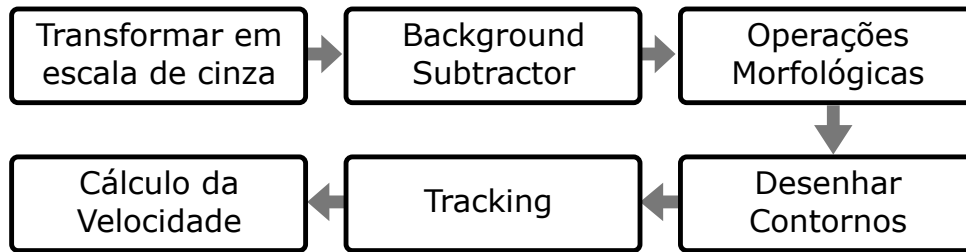


Figura 1. Fluxograma do método proposto

3. Conversão na forma binária

Uma imagem colorida pode ser representadas por diversos modelos tridimensionais, como o RGB (*Red, Green, Blue*), o HSV (*Hue, Saturation, Value*), o CMYK (*cian, magenta, yellow, black*). Devido a dificuldade de operar com esses modelos, escolhe-se trabalhar com imagens binária, onde tem-se apenas as cores branca (valor 255) e preta (valor 0). Isso facilita na construção lógica do algoritmo e evita processamentos desnecessários. Esta etapa é extremamente importante, pois com uma boa imagem binária as chances de sucesso aumentam.

3.1. Background Subtraction

Background subtraction é amplamente usada quando há a necessidade de detectar objetos em movimentos a partir de câmeras estáticas. A técnica de *background subtraction* tem como objetivo comparar duas imagens, *background* e *foreground*. O *background* será tudo aquilo que é estático no vídeo, ou seja, o valor do pixel não se altera, ou altera muito pouco, de um frame para outro. Portanto, *foreground*, será os objetos que se moveram de um frame para outro.



(a) Background.

(b) Foreground.

(c) Resultado do Background Subtraction.

Figura 2. Processo do Background Subtraction.

Para fazer essa análise, é feita a comparação dos valores do pixel da imagem de referência (*background*) e do pixel da imagem subsequente de mesma posição.

$$pixel_value = |(x, y)_1 - (x, y)_{i-1}| \quad (1)$$

Onde:

- $(x, y)_{t-1}$: Valor do pixel na posição (x, y) da imagem de referência. Valores entre 0 e 255 devido a imagem estar em escala de cinza.
- $(x, y)_t$: Valor do pixel na posição (x, y) do frame seguinte. (Valores entre 0 e 255)

A Equação 1 retorna também um valor entre 0 e 255 e esse valor indica a intensidade da mudança do valor do pixel analisado. Sendo que, valores baixos representam poucas mudanças e valores altos representam grandes mudanças. A matriz de pixel formada pelos valores da Equação 1 ainda não é uma imagem binária. Para realizar a conversão de uma imagem em escala de cinza para uma imagem binária é necessário definir um valor de limiar (*threshold*) e aplicar a Equação 2.

$$binary_pixel = \begin{cases} 0 & \text{para } pixel_value < threshold \\ 255 & \text{para } pixel_value > threshold \end{cases} \quad (2)$$

3.2. Morphological Operation

Como pode ser visto na Figura 3, alguns pixels que não fazem parte do objeto em movimento acabam recebendo o valor 255 (branco). Isso se deve a dois motivos: o valor de limiar escolhido e ruídos em imagens digitais. Nos testes do método proposto, identificou-se que ao variar o valor de limiar, pode diminuir a quantidade de pixel ruidosos, porém com o aumento do valor de limiar perde-se detalhes dos objetos em movimentos.

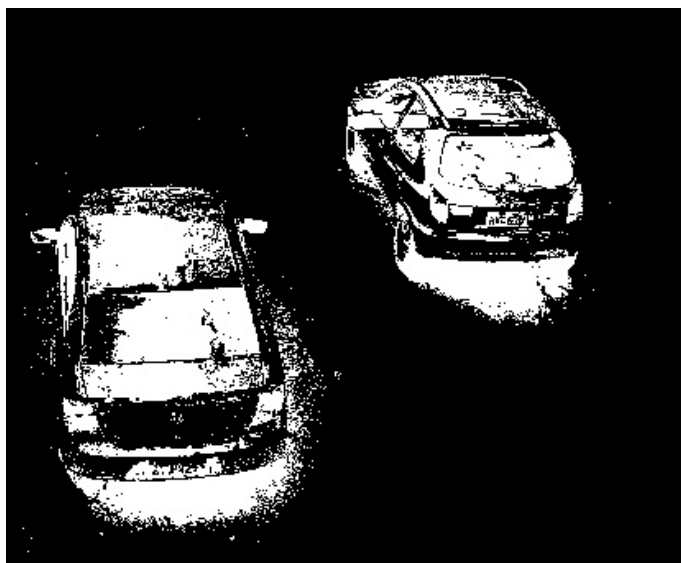


Figura 3. Resultado do Background Subtractor.

Os ruídos em imagens podem ser definidas como efeitos randômicos indesejados, normalmente ocasionado por diversos fatores como tipo de câmeras ou lentes, temperatura, intensidade de luz ou movimentos. O resultado de ruídos é, por exemplo, em uma situação onde dois pixels possuem o mesmo nível de cinza na realidade dificilmente possuirá o mesmo nível em uma imagem que representa esses mesmos pixels [Parker 2010].

A fim de eliminar os pixels ruidosos, utiliza-se de transformações morfológicas. Essa técnica é baseada na morfologia matemática, em que seus princípios surgiram em 1964 a partir das pesquisas feitas pelos franceses Georges Matheron e Jean Serra

[Facon 2011]. Essa operação tem como objetivo extrair informações relativas à geometria e à topologia de conjuntos desconhecidos de uma imagem a partir do elemento estruturante, conhecido também como *kernel* [Facon 2011].

O elemento estruturante trata-se de uma imagem binária de prova, geralmente pequena se comparado com a imagem que irá sofrer a operação morfológica. O *kernel* será posicionado em todos os locais da imagem binária e em cada posição é feita uma verificação para saber se o elemento estruturante está ou não contido na imagem. Na construção do algoritmo, os *kernels* não possuem limitação quanto à sua forma. Podem possuir diferentes formatos de acordo com os valores atribuídos a cada posição da matriz, como pode se visto alguns exemplos de *kernels* na Equação 3. No método proposto adotou-se um *kernel* retangular com a altura possuindo o dobro do valor da largura. Essa escolha se deve ao fato de que no vídeo analisado os veículos passam na direção de vertical e com sentido para cima, por isso o *kernel* deve ter uma altura maior. A largura menor é justificada para não ocorrer erros de detecção, como dois carros de faixas diferentes serem detectados como apenas um.

$$rect_{3 \times 3} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \quad cross_{3 \times 3} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 0 \end{bmatrix} \quad vert_{3 \times 3} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \end{bmatrix} \quad (3)$$

A morfologia matemática possui dois operadores básicos, *erosão* e *dilatação*. A operação de erosão é amplamente usada para fazer a remoção de partículas da imagem binária e a dilatação é utilizada para destacar os objetos na imagem. Os resultados obtidos dessas operações podem ser observadas nas Figuras 4b e 4c.



Figura 4. Resultado das operações morfológicas.

Após o processo de dilatação, é necessário fazer o contornos de cada veículo, pois essa informação será exigida para realizar a próxima etapa, *Convex Hull* [Sklansky 1982]. A função Convex Hull verifica os traços de contornos, identifica os defeitos de convexidade e os corrige. O objetivo dessas duas operações é deixar a aparência do veículo na imagem binária menos pixelizadas e sem lacunas dentro da área do veículo. O resultado obtido é visto na Figura 5.

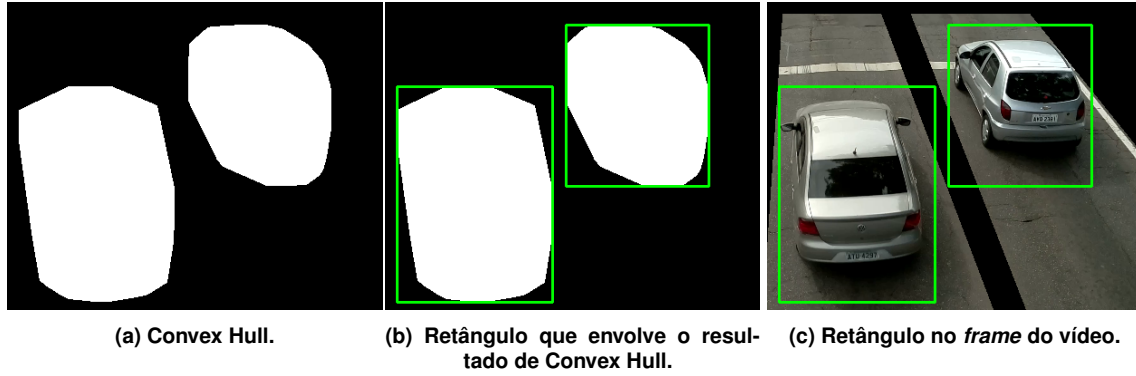


Figura 5. imagem da mask se aplicação da operações morfológicas.

4. Rastreamento do veículo

Com a imagem binária finalizada é possível desenhar um retângulo que envolve todo o veículo. Para o desenho desse retângulo é necessário o ponto de origem (x, y) e a largura (w) e altura (h) . Com os valores de (x, y, w, h) de cada *frame* é possível fazer o rastreamento (*tracking*) dos veículos.

A técnica utilizada no método deste artigo é *Blob Detection*. Ela consiste em traçar a movimentação do objetos no *frame*. Um *blob* é um grupo de pixels identificados como um objeto. O rastreamento funciona da seguinte maneira: com os valores de (x, y, w, h) encontra-se o ponto central do retângulo, esse ponto central então é armazenado e continua-se a análise do próximo *frame*. Com o novo valor do ponto central, traça-se uma linha do ponto anterior ao atual. Esse processo é feito até ter dez pontos centrais consecutivos para um mesmo veículo. O traçado desses dez pontos refletem a distância percorrida pelo veículo nesse intervalo. Logo com essa informação abre a possibilidade de realizar o cálculo da velocidade do veículo.

5. Estimando a Velocidade do Veículo

Com os dez pontos centrais consecutivos obtidos através do rastreamento, pode-se determinar a distância aproximada que o veículo percorreu. Essa distância real é calculada pela Equação 5.

$$\begin{aligned}
 dist_x &= x_i - x_{i-10} \\
 dist_y &= y_i - y_{i-10} \\
 mpp_x &= k \cdot (width \cdot y_i) \\
 mpp_y &= k \cdot y_i
 \end{aligned} \tag{4}$$

As distância percorrida no eixo x ($dist_x$) e no eixo y ($dist_y$) são calculadas pela diferença entre o valor do frame mais atualizado (décimo frame: x_i, y_i) para o frame mais antigo (primeiro frame: x_{i-10}, y_{i-10}). Ainda é necessário estimar o valor do metro por pixel (mpp), para isso deve-se saber a largura em pixel ($width$) do vídeo e ajustá-lo com uma constante (k). Essa constante k é em geral um valor abaixo de 0.1, pois é dependente da altura da câmera do solo e da resolução do vídeo capturado.

$$real_dist = \sqrt{(dist_x)^2 \cdot (mpp_x)^2 + (dist_y)^2 \cdot (mpp_y)^2} \tag{5}$$

Com o valor da distância real estimado é feito o cálculo da velocidade como na Equação 6.

$$speed = real_dist \cdot fps \cdot \frac{250}{3,6} \quad (6)$$

Onde o *fps* (*frame per second*) é a taxa de captura do vídeo e o $(250/3,6)$ é uma conversão do valor máximo de velocidade para km/h. Esse processo de medição da velocidade é realizado até quatro vezes, isso varia de acordo com o número de pontos centrais obtidos na área de medição do vídeo. Então, com esse quatros valores de velocidade é feita uma média e essa velocidade média será o valor final estimado.

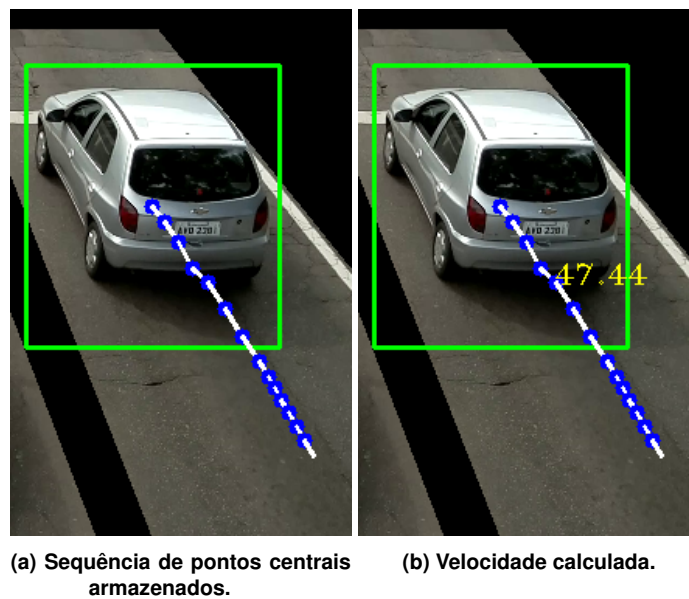


Figura 6. Velocidade calculada a partir dos pontos centrais dos retângulos gerados no processo de detecção.

6. Resultados Experimentais

O método proposto foi aplicado em um vídeo contendo um total de 107 veículos. O resultado obtido foi um erro percentual médio de 9,96% das velocidades calculadas em comparação com os valores obtidos pelos sensores indutivos. Esse erro percentual representa um erro absoluto médio de $\pm 4,21$ km/h nas medições.

Tabela 1. Erros percentuais e absolutos médios obtidos em 87 velocidades calculadas

Erro Absoluto (E)	$E < \pm 3\text{km/h}$	$\pm 3\text{km/h} < E < \pm 5\text{km/h}$	$E > 5 \text{ km/h}$
Erro Absoluto Médio (km/h)	± 1.42	± 4.18	± 8.31
Erro Percentual Médio	3.12%	10.09%	19.94%
% do total de veículos	51%	15%	34%

Com os dados obtidos percebeu-se que o método proposto apresenta melhores resultados na faixa de velocidades entre 45 km/h e 60 km/h. Dentro dessa faixa o erro percentual médio cai para 7,19% e o erro absoluto é de $\pm 3,71$ km/h, próximo do limite de $\pm 3,00$ km/h exigido pelo Instituto Nacional de Metrologia, Qualidade e Tecnologia - INMETRO para instrumentos fixos, estáticos e portáteis.

A taxa de detecção de veículos ficou em 81,3% considerando todos os 107 veículos do vídeo. Esse valor aumenta para 92,9% quando é considerado apenas a faixa entre 45 km/h e 60 km/h, sendo 66 veículos detectados de um total de 71.

7. Vantagens e Desvantagens

Vantagens:

- Baixo custo de implementação e manutenção;
- Sistema simplificado;
- Possibilidade de adição de novas funções sem precisar fazer mudanças no *hardware*.

Desvantagens:

- Oclusões de veículos podem ocorrer dependendo do tamanho do veículo e posição da câmera;
- Mudanças bruscas de iluminação, seja por algum feixe de luz refletido ou limitação da câmera em compensar essa mudança rapidamente, pode diminuir a taxa de detecção de veículos;
- Dias ensolarados podem diminuir a precisão devido as sombras projetadas dos veículos.

8. Conclusão

O método proposto performou satisfatoriamente se comparado com os valores medidos pelos sensores indutivos e suas complexidades de projeto e instalação. Por ser baseado em subtração de fundo, este método se sai melhor quando utilizado em rodovias ou avenidas que possuem fluxo constante. Devido ao ângulo da câmera utilizado não foi possível medir a velocidade de veículos grandes, como ônibus e caminhões.

Ainda assim existem desvantagens quando se trata da utilização de câmeras de monitoramento. Pois muitas, estão sujeitas à vibrações devido ao vento e passagem de veículos pesados. E quando são utilizadas câmeras de vigilância que já estavam previamente instaladas e obviamente a posição não foi escolhida objetivando o uso de visão computacional, aumenta-se a complexidade dos métodos e algoritmos a serem utilizados, para contornar essa questão dos ângulos, iluminações, oclusões, etc.

O uso de processamento de imagem para medição de velocidade tem um grande potencial, pois a estrutura de instalação, na maioria dos casos já está feita, uma vez que os radares tradicionais já utilizam-se de uma câmera para fazer o reconhecimento das placas dos veículos. Além disso, outra vantagem desse sistema simplificado, é possibilidade de aprimoramento nos algoritmos, permitindo não apenas identificar veículos, mas também realizar o reconhecimento dos mesmos a fim de classificá-los em tipos, modelos, marcas, etc. Tudo que possa ser relevante para ajudar na tomada de decisão do controle do fluxo de tráfego de veículos.

Referências

- Alexandre, L. F. B. d. A. (1997). Detecção de movimento e sua aplicação à monitorização de tráfego rodoviário. Master's thesis, Departamento de Engenharia Electrotécnica e de Computadores, Faculdade de Engenharia da Universidade do Porto, Porto, PRT.
- Facon, J. (2011). A morfologia matemática e suas aplicações em processamento de imagens. *Workshop de Visão Computacional*.
- Parker, J. R. (2010). *Algorithms for Image Processing and Computer Vision*. Wiley Publishing, Inc., Indianapolis, USA.
- Sklansky, J. (1982). Finding the convex hull of a simple polygon. *Pattern Recognition Letters*, 1.
- WHO, W. H. O. (2017). Managing speed. Technical report, World Health Organization.