

Detecção de Sonolência ao Volante Utilizando Facial Landmarks e Eye Aspect Ratio

Julia Manayra da Silva Ferreira¹, Ricardo Costa da Silva Marques¹,
Geraldo Braz Junior¹

¹Universidade Federal do Maranhão (UFMA)
Campus do Bacanga - São Luís - MA

{juhmanayra, rcsmarques89, ge.braz}@gmail.com

Abstract. *In Brazil, approximately 28 % of traffic accidents are caused by drivers sleeping at the wheel, this percentage doubles in the United States. Thus, the main objective of this work is to develop a methodology to create a system of detection of sleepy drivers behind the wheel. Therefore, the way to accomplish this task is to analyze the measure taken from the eye region, the Eye Aspect Ratio. This approach takes as a starting point a face detector, then uses the Facial Landmark information of the eye and monitors it frame by frame, to assess whether the driver is sleeping and trigger an alert to wake him. Tests were performed with 18 people and in all cases the system was successful.*

Resumo. *No Brasil, aproximadamente 28% dos acidentes de trânsito são causados por motoristas que dormem ao volante, este percentual chega a dobrar nos Estados Unidos. Dessa forma, o principal objetivo deste trabalho é desenvolver uma metodologia para criação de um sistema de detecção de motoristas sonolentos ao volante. Portanto, a maneira encontrada para realizar esta tarefa é analisar a medida extraída da região dos olhos, o Eye Aspect Ratio. Esta abordagem tem como ponto de partida um detector de faces, em seguida usa a informação Facial Landmark do olho e faz o monitoramento frame a frame, para avaliar se o motorista está dormindo e disparar um alerta para acordá-lo. Foram realizados testes com 18 pessoas e em todos os casos o sistema obteve sucesso.*

1. Introdução

Estatísticas mundiais mostram que entre 26% e 32% dos acidentes de trânsito são provocados por motoristas que dormem na direção. Eles são responsáveis por índices entre 17% e 19% das mortes no local do acidente, de acordo com Marco Túlio de Mello, coordenador do Centro de Estudo Multidisciplinar em Sonolência e Acidentes da Unifesp e chefe da disciplina Medicina e Biologia do Sono.

A estatística nacional para este tipo de acidente chega a marca de mais de 50 mil mortes em acidentes de trânsito a cada ano. Dados da National Highway Traffic Safety Administration (NHTSA) indicam que nos Estados Unidos esse número é o dobro, sendo 1550 fatais, apenas em decorrência de motoristas que dormem ao volante. Pesquisas da Austrália, Inglaterra, Finlândia e de países da Europa apontam o sono ao volante como causa de 10% a 30% dos acidentes.

Números da Polícia Rodoviária Federal também mostram que 28% dos acidentes são causados também por desatenção ao volante, que também é indicada como consequência da perda de sono. Para Simone Prezotti, o sono reduz a capacidade de concentração e de reflexos do motorista, além da sensação de fadiga.

O risco de acidentes aumenta não no horário de maior trânsito, mas naquele em que o ser humano tem um declínio na temperatura corporal, afirma Marco Mello, da Unifesp. Isso ocorre entre 12h30 e 14h e das 22h às 6 da manhã, sendo que o período crítico fica entre 3h30 e 5h30. “Com a temperatura corporal baixa, começamos a liberar melatonina, que induz ao sono.”

Portanto, este trabalho tem como objetivo apresentar uma metodologia para criação de um sistema capaz de evitar acidentes relacionados à sonolência de motoristas ao volante. Para realização dessa tarefa são utilizados conceitos chaves de Visão Computacional e Processamento de Imagens, tais como detecção de face e uso de *Facial Landmarks* [Koestinger et al. 2011] com auxílio das bibliotecas Opencv e dLib [King 2009] para a linguagem Python 3.

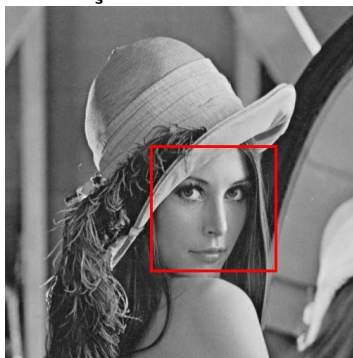
2. Fundamentação Teórica

Esta seção aborda conceitos relacionados ao objeto de estudo e mostra como trabalhos nessa área buscam desenvolver sistemas precisos que possam contribuir com a diminuição de acidentes.

2.1. Detecção de Faces

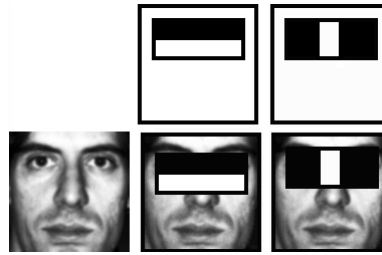
A aplicação mais conhecida em análise facial é o Reconhecimento de Rosto. Para poder identificar uma pessoa em uma imagem, primeiro é preciso detectar em que área da imagem a face está localizada. Portanto, a detecção de rosto a principal tarefa é localizar um rosto em uma imagem e retornar uma região que contém esse rosto (Figura -1). Em 2001, [Viola and Jones 2001] desenvolveram um trabalho intitulado “Detecção de objetos rápidos usando uma cascata de recursos simples”.

Figura 1. Detecção de rosto em uma imagem.



Desde os primeiros dias do OpenCV e até então, uma implementação do detector de rosto Viola e Jones é empregada. Também como contribuição deste estudo é introduzido o conceito de imagens integrais, que permite uma rápida avaliação de características. Estas características são obtidas através de *Haar Features* (Figura-2) para determinar regiões conhecidas como boca, olhos e nariz.

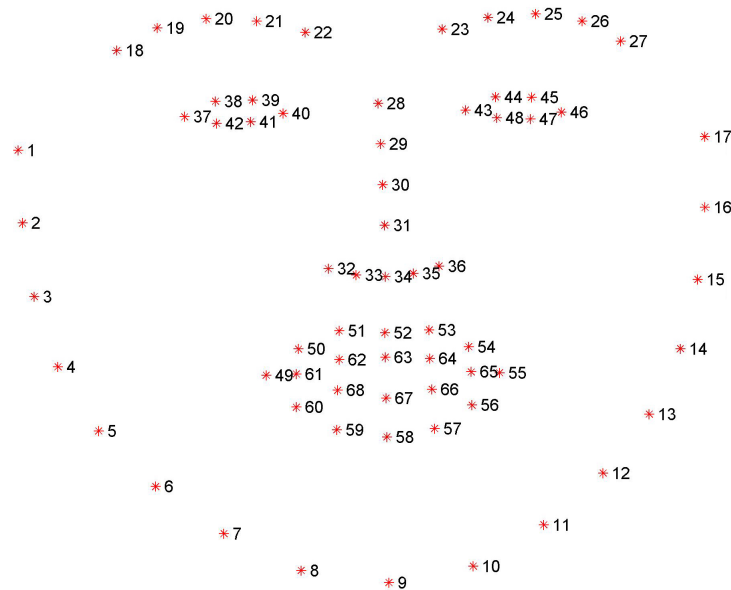
Figura 2. Haar Features.



2.2. Facial Landmarks

Uma vez que se tenha uma região do rosto bem definida, o problema óbvio é ver se é possível encontrar a localização de diferentes características faciais (por exemplo, cantos dos olhos, sobrancelhas e boca, ponta do nariz etc.) com precisão. Essas características são conhecidas na literatura como Facial Landmarks (Figura-3) e podem ser usados em muitos tipos de aplicações.

Figura 3. Facial Landmarks.



O detector de Facial Landmarks utilizado está dentro da biblioteca dLib, ele produz 68 coordenadas (x, y) que mapeiam estruturas específicas da face. Estes 68 pontos são obtidos a partir do treinamento de um modelo de predição de formas chamado iBUG 300-W dataset, e são armazenados em vetor que pode ser usado de forma específica apenas escolhendo o intervalo que a aplicação precisa. Esse vetor está definido da seguinte forma:

- Boca [48, 68].
- Sobrancelha direita [17, 21].
- Sobrancelha esquerda [22, 26].
- Olho direito [36, 41].
- Olho esquerdo [42, 47].
- Nariz [27, 35].
- Mandíbula [0, 16].

2.3. Eye Aspect Ratio

Como mostrado na Seção 2.2, podemos aplicar a detecção e usar Facial Landmark para localizar regiões importantes do rosto, incluindo olhos, sobrancelhas, nariz, orelhas e boca. No contexto deste trabalho é de grande interesse o uso das regiões que correspondem aos olhos esquerdo e direito.

A medida de Eye Aspect Ratio (EAR) é calculada entre a altura e a largura do olho (Figura-5). Onde P_1, \dots, P_6 são os locais das marcas em 2D, representados pelos intervalos [36, 41] e [42, 47] do vetor de Facial Landmark que é calculado pela equação abaixo:

$$EAR = \frac{\|P_2 - P_6\| + \|P_5 - P_5\|}{2 \|P_1 - P_4\|}$$

Nota-se que o EAR é constante quando um olho permanece em um estado e se altera a partir do momento em que o olho pisca. Esse valor pode variar de pessoa para pessoa, sendo mais indicada uma parametrização de qual limiar usar para cada indivíduo. Pois a medida EAR varia de acordo com a proporção que cada olho possui, geralmente esse valor para o olho aberto varia de 0,25 a 0,32. Para valores abaixo deve-se considerar que o motorista está entrando em um estado de sonolência.

3. Metodologia

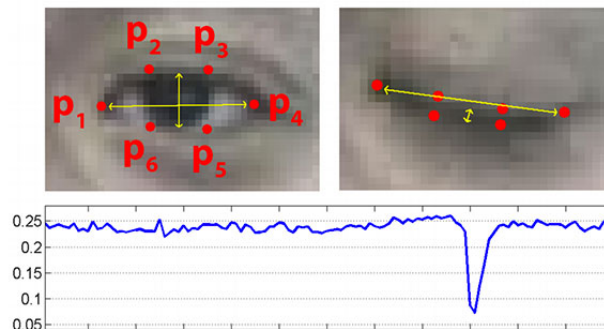
A metodologia está subdividida em 4 etapas principais (Figura - 4), respectivamente: captura dos frames de vídeo, detecção da face, tracking dos olhos, monitoramento da medida EAR. Na etapa de captura dos frames de vídeo, faz-se necessário apenas que a distância entre motorista e câmera seja de 35 a 40 cm, para priorizar uma boa detecção da face.



Na etapa de detecção da face, foi utilizado o modelo treinado da biblioteca dLib para detecção das faces. Este usa uma implementação do algoritmo Histogram of Oriented Gradient [Tsai 2010] para detecção. Assim que a face for detectada, então é aplicada a detecção das Facial Landmarks para que seja feito o tracking dos olhos.

Na etapa final, a medida de EAR é calculada em todos os frames e é disparada uma thread para realizar uma contagem de tempo assim que os olhos forem mantidos fechados. O valor de retorno da EAR será aproximadamente constante quando o olho estiver aberto e diminuirá rapidamente durante um piscar de olhos, porém ao fechar por um instante o sistema deve disparar um alerta. Se o olho estiver fechado, a EAR permanecerá novamente aproximadamente constante (Figura-5), mas será muito menor do

Figura 5. Representação do EAR. Topo esquerda: uma visualização de Landmark do olho quando está aberto. Topo direita: quando o olho está fechado. Parte inferior: proporção de aspecto do olho ao longo do tempo. A queda do valor de EAR indica um piscar de olhos [Soukupová 2016].



que a proporção quando o olho estiver aberto. Isto foi baseado, considerando o estudo de [Soukupová 2016].

Com base nessas informações esta metodologia se adapta para realizar o monitoramento constante dos olhos, fazendo o cálculo do EAR e iniciando um contador para saber a quantidade de frames que o olho se mantém no estado fechado. A ultima tarefa é soar um alarme quando o motorista manter os olhos fechados por um intervalo de 40 a 48 frames consecutivamente.

4. Resultados

A metodologia foi validada apenas em laboratório usando uma webcam aproximadamente a 40 cm de distância do rosto do indivíduo. Foi possível perceber que o sistema se comportou bem e sem falhas, mesmo apresentando algumas limitações já citadas anteriormente.

O resultado (Figura-6) mostrou que o detector de sonolência é capaz de detectar quando uma pessoa permanece com os olhos fechados e soa um alarme para chamar sua atenção. O detector de sonolência ainda é capaz de trabalhar em condições que incluem luz solar direta e iluminação baixa/artificial.

Figura 6. Resultado obtido: com o olho aberto, mostra apenas medição do EAR. Ao fechar os olhos o sistema detecta e alerta.



Uma vez que a análise do fluxo de frames de vídeo estava funcionando corretamente, foram feitos testes com o detector usando 18 pessoas para validar se ele estava

realmente funcionando corretamente. E em 17 dos casos funcionou corretamente inclusive uma pessoa usando óculos (Figura-7).

Figura 7. Caso de acerto com pessoa usando óculos.



Apenas em uma pessoa com a medida de EAR menor que o detector padrão errou (Figura-8).

Figura 8. Caso de erro com medida de EAR para olho aberto fora do limiar configurado.



Para resolver esse problema bastou que fosse modificada a configuração do EAR de acordo com as medidas que cada pessoa possui, ou seja para cada pessoa deve-se calcular um limiar diferente.

5. Conclusões

Este trabalho apresenta uma metodologia para detectar sonolência ao volante utilizando Facial Landmarks mais especificamente nas regiões que correspondem aos olhos. A característica levada em consideração neste trabalho foi a Eye Aspect Ratio, pois ela auxilia no cálculo de abertura dos olhos. Como resultado, foi produzindo um sistema capaz de detectar quando o motorista fecha os olhos com um bom tempo de resposta.

Dessa forma, cabe ressaltar que essa proposta mostrou-se robusta no que diz respeito a sistemas detecção de sonolência ao volante, mesmo com algumas limitações quanto ao posicionamento do rosto diante da camera, como rotação do tronco e alinhamento da face, pois ao virar o rosto o algoritmo de detecção perde a marcação anterior devido a configuração do EAR estar associada a uma imagem 2D, mas a partir do momento que o indivíduo se posiciona de maneira correta o sistema se recompõe rapidamente.

E por fim, embora o estado da arte use uma estratégia apoiada pelo classificador SVM obtenha bons resultados no que concerne a computação de piscadas, foi possível

perceber que ao utilizar a medida Eye Aspect Ratio, calculando-a frame a frame, este trabalho não se mostrou inferior em nenhum momento.

Como trabalho futuro pretende-se relizar um teste de validação em uma situação real, monitorando o percurso de uma pessoa ao volante, e também viabilizar um acessório, pois esta técnica poderia utilizar um esquema de implementação em um Raspberry-Pi com uma camera integrada e acoplado ao veículo.

Referências

- King, D. E. (2009). Dlib-ml: A machine learning toolkit. *Journal of Machine Learning Research*, 10(Jul):1755–1758.
- Koestinger, M., Wohlhart, P., Roth, P. M., and Bischof, H. (2011). Annotated facial landmarks in the wild: A large-scale, real-world database for facial landmark localization. In *Computer Vision Workshops (ICCV Workshops), 2011 IEEE International Conference on*, pages 2144–2151. IEEE.
- Soukupová, T. (2016). Eye blink detection using facial landmarks.
- Tsai, G. (2010). Histogram of oriented gradients. *University of Michigan*.
- Viola, P. and Jones, M. (2001). Rapid object detection using a boosted cascade of simple features. In *Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on*, volume 1, pages I–I. IEEE.