

Detecção automática de sonolência em condutores de veículos utilizando imagens amplas e de baixa resolução

Rafael Alceste Berri¹, Alexandre Gonçalves Silva¹, Rangel Arthur², Elaine Girardi¹

¹Centro de Ciências Tecnológicas-Universidade do Estado de Santa Catarina (CCT-UDESC)
Rua Paulo Malschitzki, s/n – 89219-710 – Joinville – SC – Brasil

²Faculdade de Tecnologia-Universidade Estadual de Campinas (FT-Unicamp)
Limeira – SP – Brasil

rafaelberri@gmail.com, alexandre@joinville.udesc.br
rangel@ft.unicamp.br, elaine.girardi@hotmail.com

Abstract. *In United States around 100,000 accidents are caused by drowsy drivers, resulting in over 1,500 deaths and 71,000 injuries per year. The driver inattention caused half of 126,000 accidents on brazilian federal highways (in 2010). This paper presents methods for face location, analysis of eye state (opened or closed) and drowsiness detection of drivers. In the tests performed, using images of broad surroundings (beyond the face of the driver), and maximum resolutions of 960×480 , the success rate in identifying eye was above 83%. Considering drowsiness as a sequence of frames in which the driver stands with eyes closed, this state was detected in all situations of processed videos.*

Resumo. *Estima-se cerca de 100.000 acidentes causados por motoristas sonolentos nos Estados Unidos, resultando em mais de 1500 mortos e 71.000 feridos por ano. No Brasil, metade dos 126.000 acidentes ocorridos em rodovias federais (em 2010) tem relação com a desatenção do motorista. Este artigo apresenta métodos para a localização da face, análise do estado dos olhos (aberto ou fechado) e detecção de sonolência de motoristas. Nos testes realizados, usando imagens de ambiente amplo (além do rosto do condutor) e de resoluções máximas de 960×480 , a taxa de acerto na identificação dos olhos foi acima de 83%. Considerando sonolência como uma sequência de quadros em que o motorista permanece com os olhos fechados, tal estado foi detectado em todas as situações dos vídeos processados.*

1. Introdução

Nas estradas dos Estados Unidos, a National Highway Traffic Safety Administration (NHTSA) estima cerca de 100.000 acidentes causados por motoristas sonolentos, o que resulta em mais de 1.500 mortos e 71.000 feridos por ano. Isso equivale a 1,6% de todos os acidentes e de 3,6% dos acidentes fatais [NTSB 1999]. Estima-se que aproximadamente um quarto dos acidentes veiculares sejam resultados de desatenção ou distrações na cidade de Nova Iorque [Stutts et al. 2001]. No Brasil, metade dos 126.000 acidentes ocorridos em rodovias federais, em 2010, tem relação com a desatenção do motorista [Salvador 2011].

Dirigir realmente é uma tarefa que precisa de atenção. Estudar formas de detectar distrações é o primeiro passo para minimizar esses números. O objetivo central desse

trabalho é a construção de um sistema para detecção de sonolência utilizando uma câmera de baixa resolução para os padrões atuais (entre 450×350 e 960×480) e um computador (*Car PC*) com suficiente capacidade de processamento para promover respostas em tempo real.

Entre os trabalhos da área, destaca-se a abordagem de Masala e Grosso [Masala and Grosso 2012], baseada em classificadores de redes neurais (sendo necessário, portanto, o treinamento prévio), a partir de imagens de uma única câmara posicionada frontalmente ao condutor. O método mostrou-se mais eficiente na identificação da atenção do motorista e menos eficiente para a fadiga. Outro sistema é apresentado por Lenskiy e Lee [Lenskiy and Lee 2012], os quais efetuam segmentação do condutor por cor da pele e utiliza o SURF (Speeded Up Robust Feature) [Bay et al. 2008] como modelo de extração de características faciais, utilizadas na detecção do estado do olho. O algoritmo proposto por Kumar et. al. [Kumar et al. 2012] utiliza morfologia matemática na segmentação dos olhos do motorista, sendo a detecção do estado de olho, como sendo aberto ou fechado, possível com auxílio da utilização do filtro de Gabor (distribuição de frequência espacial local). O sistema de detecção de sonolência proposto por Albu et. al. [Albu et al. 2008] utiliza-se de modelos (*template-matching*) para detectar quando os olhos do motorista estão fechados. Os modelos de olhos aberto e fechado são adquiridos em uma etapa prévia de calibragem do sistema e devem ser específicos para cada motorista. Há ainda um estudo intrusivo de Wei e Lu [Wei and Lu 2012], no qual a atividade cerebral é analisada no sentido de se estabelecer associação ao piscar dos olhos. Algumas montadoras já disponibilizam sistemas desta natureza, porém de forma proprietária, fechada ou de difícil reprodutibilidade.

Este artigo está organizado em definições preliminares, na Seção 2, para introduzir a terminologia da área e ferramentas computacionais utilizadas. Na Seção 3, são desenvolvidos os algoritmos do sistema. Na Seção 4, por sua vez, são apresentados os experimentos realizados. Por fim, conclusões são estabelecidas na Seção 5.

2. Definições preliminares

Nesta seção, ferramentas matemáticas de suporte ao algoritmo proposto, bem como uma forma usual de detecção de padrões, baseada em classificadores *Haar-like features*, e estruturas do olho, são apresentadas.

2.1. Fundamentos de processamento de imagens

Uma *imagem em níveis de cinza* é $I(x, y)$ como intensidade em $[0, L - 1]$ para cada pixel (x, y) no domínio de coordenadas $x \in [0, H - 1]$ e $y \in [0, W - 1]$, onde H é a altura e W , a largura. A *normalização* de uma imagem $I(x, y)$, com intensidades mínima e máxima iguais a respectivamente I_{min} e I_{max} , supondo $L = 256$ (intensidade representada por um inteiro sem sinal de 8 bits), é dada por $N(x, y) = \frac{1}{255}(I_{max} - I_{min})(I(x, y) - I_{min})$. Além da normalização, para padronização do brilho da imagem, pode ser necessário um filtro para reduzir ruídos do tipo “sal-e-pimenta”, ou seja, com pontos claros ou escuros espalhados. A tradicional forma de resolver este problema é o uso do *filtro mediana* [Gonzalez and Woods 1992], no qual cada pixel é substituído pela mediana dos pixels vizinhos sob uma janela quadrada (ou retangular). Deste modo, valores relativamente baixos ou altos se organizam nos extremos da ordenação, não sendo selecionados.

Outro filtro importante na presente aplicação é o *fechamento morfológico* [Soille 2003], dado por uma dilatação (filtro de máximo), seguido por uma erosão (filtro de mínimo). A ideia é, em uma primeira etapa (dilatação), aumentar ou unir regiões relativamente claras e, na etapa seguinte, retomar as formas da imagem original (porém preservando eventualmente as novas conexões, eliminando pequenas regiões relativamente escuras ou fechando reentrâncias). O *fechamento (preenchimento) de buracos* [Soille 2003], por sua vez, baseia-se no complemento da reconstrução de regiões de fundo (*background*) que tenham interseção com linhas e colunas dos limites do domínio da imagem (moldura formada pela primeira e última linhas e colunas) para que haja preenchimento de todas as regiões totalmente circundadas por regiões mais claras (preenchimento de ilhas escuras).

Por fim, é necessária a formação de uma *imagem binária*, representada por pixels com dois valores possíveis, 0 (falso) ou 1 (verdadeiro), ou seja, $L = 2$, para caracterização das regiões de interesse, obtidas por algum método de segmentação. Um dos mais simples e efetivos é a *limiarização global*, onde $L(x, y) = 1$, se $I(x, y) > T$, e $L(x, y) = 0$ caso contrário. O problema, neste caso, está na escolha de um limiar (*threshold*) T satisfatório.

2.2. Detecção de faces e olhos

Este trabalho utiliza imagens não registradas, ou seja, não focada no condutor, de modo que a câmera posicionada possa visualizar o interior do veículo de forma ampla. Para a detecção de faces de pessoas presentes em uma cena, foram utilizados detectores conhecidos por *Haar-like features* e *adaboost* previamente treinados para a classificação [Viola and Jones 2004]. Uma vez identificada uma face, esta pode ser isolada do restante da imagem. Com esta região de interesse (*ROI*), aplica-se novamente a técnica de Viola e Jones, desta vez treinada para a localização dos olhos, também a partir treinamentos prévios disponibilizados à comunidade. Para este artigo, são utilizados quatro padrões *Haar-like features*: um *Frontal* (`haarcascade_frontalface_alt.xml`¹) e outro para *Perfil* (`haarcascade_profileface.xml`¹) para a localização das faces; *Parojos* (`parojos.xml`²) para a localização dos olhos; e por fim, para a localização de cada olho individualmente, optou-se pelo padrão *Eyeglasses* (`haarcascade_eye_tree_eyeglasses.xml`¹).

Embora os *Haar-like features* sejam uma forma eficiente de identificar objetos, pode ocorrer falha, ou pior, apresentar falsos positivos (podem identificar uma região como contendo um objeto inválido). Esse problema é contornado, na maioria dos casos, pelos procedimentos apresentados na sequência. O método de Viola e Jones, bem como os treinamentos (em ‘xml’) utilizados, são implementações da literatura, no caso, para a biblioteca OpenCV [Bradski 2000]. As contribuições deste trabalho estão na combinação destes resultados e proposta de classificação de um olho em “fechado” ou “aberto” baseada na análise morfológica das imagens.

2.3. Partes do olho

As partes externas do olho humano, evidentemente observadas apenas quando aberto, são formadas pela esclera (ou esclerótica), íris e pupila. Nota-se que a esclera é branca, a íris possui cores variadas (tons de castanho, verde, azul, entre outros) e a pupila é o pequeno círculo central preto que regula a entrada de luz [Kolb et al. 2005], normalmente

¹Disponível em <https://code.ros.org/> (acessado em 30/10/2012).

²Disponível em <ftp://mozart.dis.ulpgc.es/pub/misc/EyeDetectors.zip> (acessado em 30/10/2012).

apresentando, em imagens fotográficas, um reflexo claro associado. Ao verificar o estado de sonolência de uma pessoa, espera-se a exploração destas regiões pelo algoritmo.

3. Detecção de sonolência proposto

O problema da detecção de sonolência do motorista a ser resolvido foi dividido em quatro partes: a detecção da face do motorista, a detecção de seus olhos, a classificação do estado dos olhos (aberto ou fechado) e, por fim, a identificação de sonolência propriamente dita.

3.1. Detecção da face do motorista

Considerando o fato de o motorista ser a pessoa mais próxima à câmera, o algoritmo de detecção de faces deve encontrar aquela que cobrir a maior área (quantidade de *pixels*) na imagem. O motorista, ao dirigir seu veículo, está em constante movimentação, sendo importante a utilização de padrões variados. Neste trabalho, são utilizados: padrão *Frontal* (para motorista olhando à frente) e o padrão *Perfil* (para motorista olhando lateralmente). A região adotada como a face do motorista será a de maior área retornada pelos padrões. Na Figura 1, é exemplificada, em diferentes *frames*, a escolha da maior região. O retângulo, em verde, representa o padrão *Frontal* e, em azul, o padrão *Perfil*.

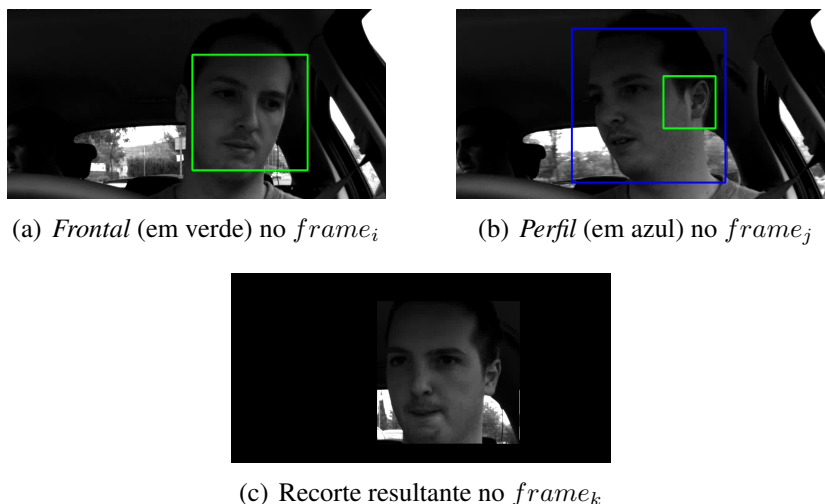


Figura 1. Recorte da face do motorista baseada na análise conjunta de dois padrões de posicionamento de cabeça (*frames* distintos). Vídeo de [Nuevo et al. 2010].

3.2. Detecção dos olhos

A localização dos olhos é importante para abalizar ocorrências de sonolência, sendo aplicada apenas na região de recorte da face do motorista. Novamente, são utilizados dois padrões, o primeiro para a região contendo ambos os olhos e o segundo para cada olho.

3.2.1. Região contendo ambos olhos

Um olho não é localizado diretamente em um primeiro momento. Antes, destaca-se a área de sua provável ocorrência, de modo a reduzir a possibilidade de falsos positivos. Para a localização da região dos olhos, utiliza-se o padrão *Parojos* e adota-se a maior área

encontrada por este padrão. A partir deste momento, é estabelecido um aumento da região retangular em 50% verticalmente (25% para cima e 25% para baixo). Esta folga se dá pelo fato de que, em alguns *frames*, a região dos olhos encontrada é pequena, inviabilizando a identificação posterior de cada olho isoladamente. As sobrancelhas, por exemplo, são importantes nesta detecção e não devem ser removidas. Na Figura 2, a saída desta etapa algoritmo é exemplificada.

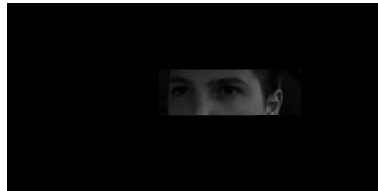


Figura 2. Recorte resultante da detecção da região contendo ambos os olhos.

3.2.2. Região de cada olho

Nem sempre é possível detectar ambos os olhos pois, dependendo do ângulo de movimentação da cabeça (posição de perfil), um deles é ocultado na cena. O padrão utilizado para este algoritmo é o *Eyeglasses* e as duas maiores áreas encontradas na região dos olhos são adotadas como os olhos do motorista. Na Figura 3, tem-se a imagem resultante dos recortes dos dois olhos detectados do motorista. Neste ponto, finaliza-se a utilização de todos os padrões da literatura, combinados coerentemente, neste trabalho, para uma robusta separação dos olhos.

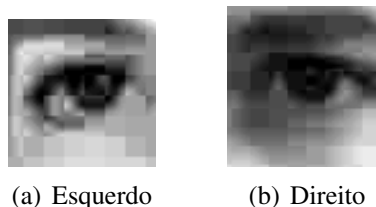


Figura 3. Recortes resultantes da detecção final de cada olho.

3.3. Estado dos olhos

A principal contribuição deste trabalho consiste na classificação do estado dos olhos em “aberto” ou “fechado” para indicação ou não de sonolência. Utiliza-se a ideia de que a pupila é escura, envolvida por uma considerável região mais clara [Kolb et al. 2005]. Por outro lado, o olho, quando fechado, possui uma área predominantemente escura. Portanto, a diferenciação entre olho aberto ou fechado ocorre fundamentalmente pela identificação de uma pequena região escura envolvida por uma região relativamente maior com níveis de cinza também maiores (ilha escura).

O algoritmo de verificação do estado dos olhos, detalhado na Figura 4, foi constituído através de processo empírico e resume-se em: pré-processamento (linhas 2 e 3), segmentação do olho (linhas de 4 a 7) e o teste do olho aberto/fechado (linhas 8 a 14).

```

1: procedure EYESTATUS(imgEye)
2:    $a \leftarrow \text{Normalization}(\text{imgEye})$  ▷ Normalização de 0 a 255
3:    $b \leftarrow \text{Median}(a)$  ▷ Filtro mediana
4:    $c \leftarrow \text{Close}(b, \text{disk}(4))$  ▷ Fechamento morfológico com disco de raio 4
5:    $d \leftarrow \text{CloseHoles}(c)$  ▷ Fechamento (preenchimento) de buracos
6:    $e \leftarrow \text{Subtraction}(d, c)$  ▷ Subtração com saturação em 0
7:    $f \leftarrow \text{Threshold}(e, 20)$  ▷ Limiarização com T=20
8:    $g \leftarrow \text{Threshold}(b \times f, 1)$  ▷ Limiarização com T=1
9:    $h \leftarrow \text{CloseHoles}(g)$  ▷ Fechamento (preenchimento) de buracos
10:  if  $g \neq h$  then
11:    return True ▷ Olho aberto
12:  else
13:    return False ▷ Olho fechado
14:  end if
15: end procedure

```

Figura 4. Algoritmo de verificação do estado do olho.

Inicialmente, aplica-se um filtro de mediana (Figura 5(b)) para homogeneizar (redução de complexidade) a imagem original (Figura 5(a)). A segmentação inicia-se com o fechamento morfológico [Haralick et al. 1987, Gonzalez and Woods 1992] com elemento estruturante disco de raio 4, reduzindo regiões escuras suficientemente pequenas (Figura 5(c)). Em seguida, efetua-se o fechamento de buracos [Gonzalez and Woods 1992, Soille 2003], com um elemento estruturante cruz 3×3 , para preenchimento de ilhas escuras, em especial aquela formada a partir da íris (Figura 5(d)). Na sequência, subtraem-se estas duas últimas imagens no sentido de obter o realce da região dos olhos (Figura 5(e)). Aplica-se então uma limiarização de intensidade 20, de modo a segmentar a região de interesse (*ROI*) do olho (Figura 5(f)).

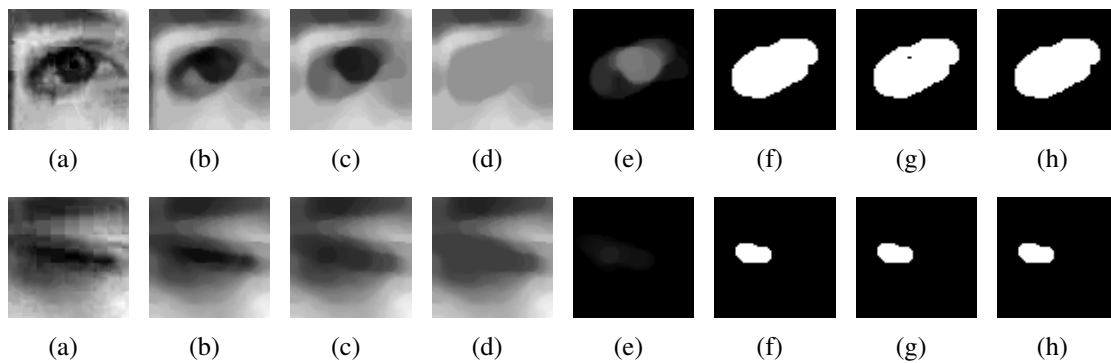


Figura 5. Sequência de operações para identificação do estado do olho “aberto” (linha superior) e “fechado” (linha inferior). (a) imagem do olho normalizada; (b) após mediana; (c) após operação de fechamento; (d) após fechamento de buracos; (e) Subtração entre (d) e (c); (f) Após limiarização de 20 em (e); (g) Binarização $(a) \times (f)$ maior que 0; (h) Fechamento de buracos da imagem (g).

Os *pixels* da imagem (b), sob domínio da *ROI* da imagem binária (f), que tenham intensidades próximas ou iguais a zero, sinalizam a localização da pupila [Kolb et al. 2005]. Somente os *pixels* que possuem intensidade mínima ou não pertencem

ao *ROI*, recebem valor 0 (“falso”) na imagem binarizada resultante (Figura 5(g)). Efetua-se então novo fechamento de buracos [Gonzalez and Woods 1992, Soille 2003] com elemento cruz 3×3 na imagem binária (g), eliminando assim regiões iguais a 0 (“falso”) que estejam completamente envolvidas por regiões iguais a 1 (“verdadeiro”) (Figura 5(h)). Sendo a imagem (g) diferente de (h), declara-se o olho como “aberto”, pois a pupila está visível. Caso contrário, o olho é declarado “fechado”.

3.4. Detectando sonolência do motorista

Com o resultado do algoritmo de identificação do estado dos olhos, é possível detectar a sonolência do motorista. Sabe-se que, para haver um estado de microssono, o fechamento ocular deve se prolongar por cerca de 6 segundos [Heitmann et al. 2001]. Pode-se, portanto, observar o número de quadros em que o motorista permanece com olhos fechados. Sabendo-se o FPS (*frames per second* ou quadros por segundo) do vídeo em questão, é possível calcular a permanência dos olhos fechados. O número de quadros (N) que precisam ser pesquisados, para um tempo desejado (t), em segundos, antes de se afirmar que um motorista está sonolento, é dado pela Equação 1.

$$N = \text{FPS} \times t \quad (1)$$

Um vídeo 30 fps, é necessário, portanto, verificar 180 quadros ($30 \text{ FPS} \times 6$ segundos) para indicar sonolência do motorista. Para promover confirmação, é interessante ainda adotar o *Percent Eye Closure* (PERCLOS) que é o percentual no qual os olhos permanecem fechados em um intervalo de tempo. Usualmente, assume-se a sonolência se houver olhos fechados em 80% ou mais dos quadros [Wierwille et al. 1994, Dinges et al. 1998, Masala and Grosso 2012]. Ao detectar este perigoso estado, o motorista pode ser alertado, por algum tipo de aviso sonoro, a partir da saída do algoritmo.

4. Experimentos

Para a análise dos resultados da solução proposta, optou-se em efetuar testes com três vídeos.

O primeiro é denominado *Fpessoa* e foi gravado por meio do programa *Dell Webcam Center*, utilizando a Webcam Integrada fabricada pela *Quanta Computer Inc.* O vídeo possui 17 segundos de duração, 30 FPS, codec *Windows Media Video 8*, dimensões de 768×480 , 330 *frames* no total e colorido. A câmera focaliza a face frontal de uma pessoa a 55 centímetros de distância. O algoritmo de verificação do estado dos olhos (Figura 4) atribuiu corretamente a situação dos olhos como “aberto” em 527 situações de um total de 652 olhos abertos do vídeo, gerando uma taxa de acerto de 80,83%. Para olhos “fechados”, a taxa de acerto foi de 96,67%, ou seja, identificou corretamente 348 casos em um total de 360 existentes. Em 2 casos não foi possível localizar os olhos, gerando uma taxa de acerto de 99,80% para esta situação. Do segundo 4 até o 10 do vídeo, foram localizados 344 olhos fechados e 16 olhos abertos. Neste período, o motorista foi declarado como sonolento por estar em 4,44% do período de olhos abertos.

O segundo vídeo escolhido é o *RS-DMV07*³ do grupo “RobeSafe Research Group” (RS-DMV) [Nuevo et al. 2010]. O vídeo possui 30 segundos de duração, 30 FPS, codec

³Vídeo “RS-DMV_07_sample.avi” está disponível no site do grupo RS-DMV em <http://www.robeseafe.com/personal/jnuevo/Datasets.html> (acessado em 30/10/2012).

MPEG-4, dimensões de 960×480 , 910 *frames* no total e todos em níveis de cinza. A câmera, na gravação do vídeo, foi colocada do lado esquerdo do carro (visão do motorista) atrás do volante, próxima ao motorista. O algoritmo da Figura 4 atribuiu corretamente a situação dos olhos como “aberto” em 1511 situações de um total de 1535 olhos abertos do vídeo, gerando uma taxa de acerto de 98,44%. Para olhos “fechados”, a taxa de acerto foi de 30,00%, ou seja, identificou corretamente 19 casos em um total de 60 existentes. Neste vídeo, não existem muitos casos de olhos fechados, comprometendo a taxa de acerto. No entanto, a classificação em “sonolência” ou “não sonolência” se mantém funcional. Por apresentar ruídos, tornou-se necessária a utilização da etapa 2 (etapa opcional do filtro de mediana) do algoritmo da Figura 4. Em 138 casos, não foi possível localizar os olhos, gerando uma taxa de acerto de 92,04% para esta situação. Em nenhum momento, há detecção de sonolência do motorista, ocorrendo apenas rápidas piscadas.

O terceiro vídeo é o *DM*⁴. O vídeo possui 63 segundos de duração, 25 FPS, codec MPEG-4, dimensões de 450×360 e todos os seus 1575 *frames* em níveis de cinza. A câmera foi fixada no painel do carro, próximo e frontal ao motorista. Ocorreram 1976 atribuições corretas (pelo algoritmo da Figura 4) para olhos “abertos” de um total de 2583 existentes, gerando uma taxa de acerto de 76,50%. Para olhos “fechados”, a taxa de acerto foi de 74,23%, ou seja, identificou corretamente 193 casos em um total de 260 existentes. Em 180 casos, não foi possível localizar os olhos, gerando uma taxa de acerto de 94,05% para esta situação. A sonolência do motorista não é detectada.

Na Tabela 1, mostram-se os resultados obtidos para os vídeos.

Tabela 1. Resultados obtidos com os vídeos.

Vídeo	<i>Fpessoa</i>		<i>RS-DMV07</i> ³		<i>DM</i> ⁴	
Item analisado	Qtde.	Acerto	Qtde.	Acerto	Qtde.	Acerto
Olhos abertos	652	80,83%	1535	98,44%	2583	76,50%
Olhos fechados	360	96,67%	60	30,00%	260	74,23%
Erro identif. olhos	2	99,80%	138	92,04%	180	94,05%
Total de olhos no vídeo	1014		1733		3023	

5. Conclusão

Neste trabalho, foram utilizados padrões de detecção de faces e olhos da literatura. A proposta deste trabalho se configura no arranjo e combinação de tais padrões para o recorte preciso de cada um dos olhos do motorista (quando não há oclusão). Sua contribuição está especialmente na classificação do estado dos olhos, em “aberto” ou “fechado”, de modo a permitir uma sinalização de sonolência. Para a localização da região de interesse, emprega-se a técnica de Viola e Jones [Viola and Jones 2004] e, para a classificação dos olhos, utilizam-se operações aninhadas fundamentalmente de morfologia matemática. Em relação a outras propostas, o presente trabalho se destaca por ser conceitualmente simples, não necessitando de calibragem prévia ou abordagem intrusiva, como desvantagem, não se adapta a motoristas com olhos cobertos (por exemplo, com óculos escuros) e tende a falhar quando o motorista olha para os lados.

⁴Está disponível em <https://www.youtube.com/watch?v=74Wp75I83gA> (acessado em 30/10/2012).

Para olhos fechados, a taxa de acerto apurada para os vídeos *Fpessoa*, *RS-DMV07* e *DM* foram, respectivamente, de 96,67%, 30,00% e de 74,23%. No caso do *RS-DMV07*, houve pequeno percentual de quadros com olhos fechados, relacionados a piscadas rápidas. Para olhos abertos, obteve-se percentual de acerto de 80,83% para *Fpessoa*, 98,44% para *RS-DMV07* e 76,50% para *DM*.

A detecção dos olhos fechados é essencial na detecção de sonolência. Nos vídeos *RS-DMV07* e *DM* não foi possível detectar sonolência do motorista, pois o mesmo não se manteve de olhos fechados por tempo considerável. Por outro lado, do segundo 4 até o 10 do vídeo *Fpessoa*, foram localizados 344 olhos fechados e 16 olhos abertos. Neste período, o motorista foi declarado como sonolento por estar em 4,44% do período de olhos abertos. O percentual de erro de sonolência utilizado para este *Fpessoa* foi de 8% nos testes, já que a identificação incorreta para olhos fechados é de 3,33%. Este percentual mostra-se seguro, podendo haver até 28 olhos abertos em uma amostra de 360 para o motorista ser declarado como sonolento. Apesar da taxa de acerto, para olhos fechados, do vídeo *RS-DMV07* ser de apenas 30%, não houve comprometimento algum da detecção de sonolência, já que esta necessita do estado do olho fechado em uma sequência relativamente grande de quadros, o que não ocorre neste vídeo. Em ocasiões em que o motorista se encontrava com a face em perfil (ocorrências apenas no vídeo *RS-DMV07*), a detecção dos olhos abertos mostrou-se prejudicada.

Como forma de melhorar a taxa de acerto do algoritmo de detecção dos olhos, pode-se testar ou desenvolver outros arquivos treinados ao reconhecimento de padrões que os localizam (além de *Parojos* para regiões e *Eyeglasses* para localizar os olhos individualmente, utilizados neste artigo). Assim pode-se minimizar a ocorrência de identificação inválida dos olhos. Outra alternativa é acrescentar uma ou mais câmeras em posições diferenciadas, de modo a evitar oclusões dos olhos com a movimentação lateral da cabeça. Como trabalhos futuros, deve-se observar a própria movimentação da cabeça que, em estado de sonolência, tende a se inclinar para baixo e dotar o sistema com visão noturna, utilizando, por exemplo, câmeras sensíveis à luz infravermelha.

Agradecimentos

Os autores agradecem o apoio financeiro da CAPES/DS concedido ao Rafael Alceste Berri do Programa de Pós-Graduação em Computação Aplicada da UDESC, e ao PROBIC/UDESC concedido a Elaine Girardi do curso de Graduação em Engenharia Elétrica.

Referências

- Albu, A. B., Widsten, B., Wang, T., Lan, J., and Mah, J. (2008). A computer vision-based system for real-time detection of sleep onset in fatigued drivers. In *Intelligent vehicles symposium, 2008 IEEE*, pages 25–30. IEEE.
- Bay, H., Ess, A., Tuytelaars, T., and Van Gool, L. (2008). Speeded-up robust features (surf). *Comput. Vis. Image Underst.*, 110(3):346–359.
- Bradski, G. (2000). The OpenCV Library. *Dr. Dobb's Journal of Software Tools*.
- Dinges, D., Mallis, M., Maislin, G., Powell, I., et al. (1998). Evaluation of techniques for ocular measurement as an index of fatigue and the basis for alertness management. Technical report.

- Gonzalez, R. and Woods, R. (1992). Digital imaging processing. *Massachusetts: Addison-Wesley*.
- Haralick, R., Sternberg, S., and Zhuang, X. (1987). Image analysis using mathematical morphology. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, (4):532–550.
- Heitmann, A., Guttkuhn, R., Aguirre, A., Trutschel, U., and Moore-Ede, M. (2001). Technologies for the monitoring and prevention of driver fatigue. In *Proceedings of the First International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design*, pages 81–86.
- Kolb, H., Fernandez, E., and Nelson, R. (2005). Gross anatomy of the eye. In Kolb, H., Fernandez, E., and Nelson, R., editors, *Webvision – The Organization of the Retina and Visual System*, pages 1–10. University of Utah Health Sciences Center.
- Kumar, K., Alkoffash, M., Dange, S., Idarrou, A., Sridevi, N., Sheeba, J., Shah, N., Sharma, S., Elyasi, G., and Saremi, H. (2012). Morphology based facial feature extraction and facial expression recognition for driver vigilance. *International Journal of Computer Applications*, 51(2):17–24.
- Lenskiy, A. and Lee, J. (2012). Driver’s eye blinking detection using novel color and texture segmentation algorithms. *International Journal of Control, Automation and Systems*, 10(2):317–327.
- Masala, G. and Grosso, E. (2012). Detecting driver inattention by rough iconic classification. Technical report.
- NTSB (1999). Evaluation of us department of transportation efforts in the 1990s to address operator fatigue.
- Nuevo, J., Bergasa, L. M., and Jiménez, P. (2010). Rsmat: Robust simultaneous modeling and tracking. *Pattern Recognition Letters*, 31(16):2455 – 2463.
- Salvador, A. (2011). A culpa foi do celular. *Revista Veja*, 21/12/2011.
- Soille, P. (2003). *Morphological image analysis: principles and applications*. Springer-Verlag New York, Inc.
- Stutts, J. et al. (2001). *The role of driver distraction in traffic crashes*. AAA Foundation for Traffic Safety Washington, DC.
- Viola, P. and Jones, M. (2004). Robust real-time face detection. *International journal of computer vision*, 57(2):137–154.
- Wei, Z. and Lu, B. (2012). Online vigilance analysis based on electrooculography. In *Neural Networks (IJCNN), The 2012 International Joint Conference on*, pages 1–7. IEEE.
- Wierwille, W., Wreggit, S., Kirn, C., Ellsworth, L., and Fairbanks, R. (1994). Research on vehicle-based driver status/performance monitoring; development, validation, and refinement of algorithms for detection of driver drowsiness. final report. Technical report.