

Desarrollo de un sistema de detección de pérdidas de atención en la conducción basado en visión computacional

Gonzálo Farré, César Martínez y Enrique M. Albornoz*

¹Instituto de investigación en Señales, Sistemas e Inteligencia Computacional, **sinc**(*i*), FICH-UNL/CONICET, Argentina.

*emalbornoz@sinc.unl.edu.ar

Resumen Los accidentes de tránsito debido a errores humanos son muy habituales a pesar de los avances en seguridad que presentan los vehículos modernos. En este trabajo se presenta el diseño e implementación de un sistema que evalúa las distracciones de los conductores, a través del procesamiento de video obtenido con un dispositivo móvil. El sistema permite evaluar cuando los conductores no prestan la debida atención, que incluye situaciones donde no miran al frente o los parpadeos son excesivos y pueden evidenciar somnolencia. Adicionalmente, se diseñó y registró una base de datos de videos que permite la evaluación de la conducción en diferentes escenarios.

1. Introducción

En las últimas décadas, los avances de la tecnología se han visto plasmados en la producción automotriz y los vehículos se vuelven cada día más sustentables e “inteligentes”. La próxima generación de vehículos tendrán en conjunto: ECU, CPU, GPU, cámaras de alta definición, sensores, pasarelas y dispositivos de almacenamiento, todo conectado a través de redes de alta velocidad. De esta forma, con base en lo percibido se podrán formular sistemas informáticos que tomen decisiones de forma autónoma.

Entre los avances destacados se pueden citar automóviles eléctricos con paneles solares que ayudan a cargar sus baterías¹; monitores del comportamiento de conducción y la detección de choques para aplicaciones telemáticas, de gestión de flotas y seguros²; advertencia de cambio de carril, asistencia de mantenimiento de carril o asistencia de centrado de carril; detección de puntos ciegos; control de cruceo adaptativo (inteligente); seguridad de la ciudad, detección de peatones, advertencia de colisión frontal, frenado automático a baja velocidad³.

¹ <https://sonomotors.com/en/sion/>

² <https://www.businessmotoring.co.uk/bosch-plug-in-telemetry-device-for-fleet-management-just-uses-lighter-socket/>

³ <https://www.extremetech.com/extreme/282494-extremetechs-best-cars-and-car-tech-for-2019>

En este contexto, la responsabilidad de los accidentes está cada vez más influenciada por errores humanos. Entre las principales causas de accidentes está la pérdida de atención al conducir y a la señalización, y que puede ser provocada por múltiples motivos como ser distracciones por la utilización del celular⁴ o del sistema de audio del vehículo, somnolencia, búsqueda de objetos, entre otros [2]. Según el último Informe sobre la situación mundial de la seguridad vial, publicado en 2015 por la Organización Mundial de la Salud ⁵:

Cada año mueren cerca de 1,3 millones de personas en las carreteras del mundo entero, y entre 20 y 50 millones padecen traumatismos no mortales. Los accidentes de tránsito son una de las principales causas de muerte en todos los grupos etarios, y la primera entre personas de entre 15 y 29 años.

Lamentablemente, Argentina tiene uno de los índices más altos de mortalidad producida por accidentes de tránsito: 22 personas mueren por día, casi 8.000 muertos por año, y más de 120.000 heridos por año, sumado a las pérdidas materiales. En comparaciones con países desarrollados, se ve que Argentina no ha disminuido el número de accidentes totales en los últimos 30 años [7]

En la literatura pueden encontrarse diversas aproximaciones a la tarea, la mayoría de las cuales utiliza fusión de información mono o multimodal. En [8] se propone un sistema de evaluación de la atención del conductor utilizando dispositivos con cámaras duales. La fusión de información visual de la cara de la persona con la señal de presión arterial se presenta en [9]. En [4] se combinan datos tomados de sensores del automóvil y diferentes medidas sobre el conductor como punto de mirada, dirección de la cabeza, frecuencia de parpadeo, etc., buscando la combinación más efectiva de características para la clasificación fatigado/no-fatigado. Finalmente, para una revisión reciente del estado del arte, se refiere al lector a [10].

El objetivo de este trabajo es desarrollar un sistema basado en visión computacional para detección de pérdida de atención en la conducción, mediante la evaluación de la dinámica temporal de la mirada del conductor en una única cámara. A continuación se presentan los materiales utilizados y desarrollo de la base de datos, junto a las definiciones principales del sistema y su implementación. En la sección 3 se presentan los resultados y discusiones realizadas sobre el comportamiento del sistema. Finalmente se presentan las conclusiones y los trabajos futuros.

2. Herramientas y desarrollo del sistema

El recurso más fundamental para la evaluación de los desarrollos es la base de datos de videos en situaciones habituales de conducción. Luego de una exploración exhaustiva, se comprobó que los Corpus disponibles no satisfacían las

⁴ <http://www.luchemos.org.ar/es/investigaciones/el-telefono-celular-un-peligro-en-las-manos-de-conductores-y-peatones>

⁵ <http://www.who.int/features/factfiles/roadsafety/es/>

Tabla 1. Distribución de videos del Corpus.

Distracciones	Leves	Moderadas	Peligrosas
12 PM	hombre(24)/mujer(21)	hombre(24)/mujer(21)	hombre(24)/mujer(21)
5 PM	hombre(20)/mujer(30)	hombre(20)/mujer(25)	hombre(19)/mujer(21)
9 PM	hombre(18)/mujer(21)	hombre(18)/mujer(21)	hombre(18)/mujer(21)

necesidades planteadas como requerimientos y por lo tanto se decidió generar un Corpus propio, cuyas definiciones se basaron en el análisis de videos disponibles públicamente⁶.

2.1. Corpus de videos

Luego de varias pruebas, donde se consideró desde no molestar al conductor hasta la mejor posición para la utilización de detección de características del rostro, el protocolo definido para el relevamiento de videos contempla:

- utilizar siempre la misma cámara (Celular Motorola Moto G4 Plus),
- ubicar la cámara siempre en la misma posición, logrando una vista de frente del sujeto,
- utilizar el formato MP4 con resolución 1920×1080 px y 28 FPS,
- realizar sesiones de 1 minuto (aprox.)
- simular una acción de conducción habitual,
- simular 3 distracciones por video, una cada 15[s] (indicada por un tutor),
- utilizar diferentes individuos, en particular fueron: 4 masculinos y 3 femeninos, edades entre 18 y 30 años, estaturas entre 1,70[m] y 1,90[m],
- ubicar al conductor en la forma estándar, es decir, sin situaciones con asientos reclinados o similares,
- relevar los videos a las 12 PM, a las 5 PM y a las 9 PM,
- definir ad-hoc 3 tipos de distracciones:
 - Leves: pérdida de atención de 1 a 3 segundos,
 - Moderadas: pérdida de atención de 3 a 5 segundos,
 - Peligrosas: pérdida de atención mayor a 5 segundos,

Finalmente, se generaron 18 videos (2.25 GB) con la distribución de la Tabla 1, algunos frames pueden verse en la figura 1. Cada uno tiene asociado un archivo de texto con la meta-data asociada: conductor, distracciones, etc.

2.2. Sistema

El sistema está compuesto por una serie de módulo y pueden verse en la figura 2. El primer paso consiste en determinar si la escena es nocturna y si lo es, se

⁶ <http://adas.cvc.uab.es/elektra/enigma-portfolio/cvc11-drivface-dataset/>,
<https://www.kaggle.com/c/state-farm-distracted-driver-detection/data>,
<http://www.robosafe.com/personal/jnuevo/Datasets.html>

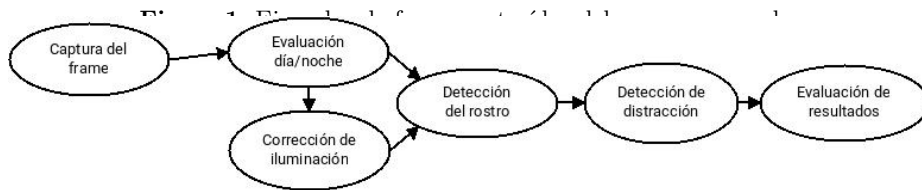


Figura 2. Esquema general del sistema.

aplica una corrección gamma [3]. La evaluación se realiza sobre el histograma RGB de una ROI de la imagen que no contiene al conductor. Para la etapa de *detección de rostro* se exploraron dos alternativas, una basada en LBP (del inglés, Local Binary Pattern) [1] y la otra en clasificadores en cascada basados en características de Haar [11]. Luego de evaluar el comportamiento sobre nuestro Corpus, se decidió utilizar Haar en su implementación de OpenCV [5] ya que se comportó de forma más estable. Sobre el rostro detectado, se obtienen 68 puntos característicos (landmarks) utilizando la librería Dlib [6]. Utilizando 14 de estos puntos, se hace un mapeo $2D \rightarrow 3D$, para poder evaluar la orientación del rostro⁷. Finalmente, es posible obtener los 3 ejes sobre el rostro que indican la orientación de la cabeza (ver figura 3). En la etapa de detección de distracciones, se evalúan los ángulos computados y se comparan con umbrales empíricos. Es importante destacar que las distracciones se evalúan sobre un buffer de resultados que permite evitar errores de cálculo instantáneo.

Adicionalmente a las distracciones se definió un “índice de somnolencia”, basado en la cantidad de frames que los ojos permanecen cerrados por minuto. Para esto, se utilizó una relación de aspecto de los ojos basada en los descriptores de Dlib. Este índice también tiene umbrales empíricos que determinan si es un evento peligroso o no.

Una vez detectado alguno/s de los eventos mencionados, el sistema genera alarmas que alertan la situación. Finalmente, se genera un informe con el análisis del video.

⁷ <http://acodigo.blogspot.com/2017/11/estimar-postura-3d.html>

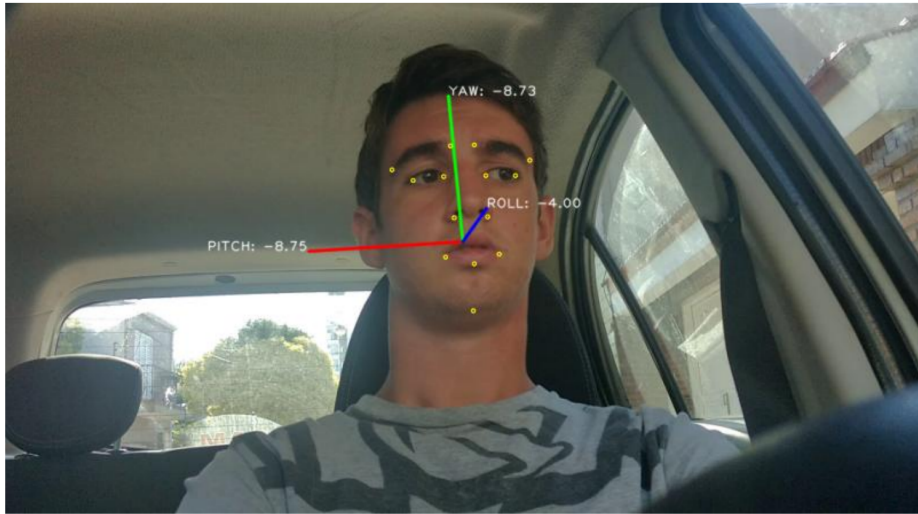


Figura 3. Estimación de postura 3D.

Se realizaron pruebas sobre todos los videos del corpus, y los resultados son muy alentadores. La combinación de la información basada en la orientación de la cabeza en conjunto con el “índice de somnolencia” permite una muy acertada evaluación de la situación. Para poder ver el comportamiento del sistema, se generaron videos que muestran el funcionamiento del sistema⁸.

3. Conclusiones y trabajos futuros

En este trabajo se ha presentado el diseño y desarrollo del prototipo inicial de un sistema que permite evaluar las pérdidas de atención de los conductores, ya sea porque quitan la vista del camino o porque presentan signos de somnolencia. Adicionalmente, se desarrolló una base de datos de videos de conductores de ambos sexos y en distintas condiciones de iluminación natural (a lo largo del día) El sistema funciona de manera offline, los videos son capturados por el dispositivo móvil y su procesamiento se realiza en una PC donde se obtienen los resultados temporales y el informe de las alarmas generadas por el sistema.

La información recolectada y los informes generados junto a los videos permitieron corroborar el buen funcionamiento del sistema, dando lugar a la configuración de distintos tipos de distracción.

Como trabajo futuro se planifica integrar el sistema en una aplicación Android donde se procese el video en tiempo real y se desplieguen las alarmas gráficas y sonoras en el dispositivo móvil o en el sistema multimedia del vehículo mediante bluetooth. Alternativamente, se prevé el desarrollo de un sistema integral de hardware y software que funcione en una mini-PC (como ser una

⁸ https://drive.google.com/drive/folders/105Wt_bxBa6YhwFp8DD1YUNZCd6vk1D1R

Raspberry Pi) y permita la instalación de una cámara dedicada que se integre al ambiente del vehículo.

Agradecimientos

Los autores quieren agradecer el apoyo de la *Universidad Nacional del Litoral* (con CAID-PJ-500 20150100055LI), ANPCyT (con PICT 2015-0977 y PICT 2016-0651) y CONICET.

Referencias

1. Chang-Yeon, J.: Face detection using lbp features. Final Project Report 77 (2008)
2. Dirección General de Tráfico, Ministerio del Interior: Distracciones al volante, <http://publicacionesoficiales.boe.es>, Madrid, España.
3. Gonzalez, R., Woods, R.: Digital image processing: Pearson prentice hall. Upper Saddle River, NJ 1 (2008)
4. Henni, K., Mezghani, N., Gouin-Vallerand, C., Ruer, P., Ouakrim, Y., Vallières, É.: Feature selection for driving fatigue characterization and detection using visual-and signal-based sensors. In: Applied Informatics. vol. 5, p. 7. Springer (2018)
5. Kaehler, A., Bradski, G.: Learning OpenCV 3: computer vision in C++ with the OpenCV library. O'Reilly Media, Inc. (2016)
6. King, D.E.: Dlib-ml: A machine learning toolkit. Journal of Machine Learning Research 10(Jul), 1755–1758 (2009)
7. Luchemos por la Vida: Asociación civil - organización no gubernamental, <http://www.luchemos.org.ar/>
8. Mihai, D., Dumitru, A., Postelnicu, C., Mogan, G.: Video-based evaluation of driver's visual attention using smartphones. In: 2015 6th International Conference on Information, Intelligence, Systems and Applications (IISA). pp. 1–5 (July 2015)
9. Sayeed, A., Sadi, S.A.: Driver drowsiness detection using face monitoring and pressure measurement. Research & Reviews: A Journal of Embedded System & Applications 5(3), 12–18 (2018)
10. Ullah, M.R., Aslam, M., Ullah, M.I., Maria, M.E.A.: Driver's drowsiness detection through computer vision: A review. In: Mexican International Conference on Artificial Intelligence. pp. 272–281. Springer (2017)
11. Viola, P., Jones, M.J.: Robust real-time face detection. International journal of computer vision 57(2), 137–154 (2004)