

ALTERNATIVA TECNOLÓGICA PARA ESTIMAR LA DISTRACCIÓN Y SOMNOLENCIA DEL CONDUCTOR MEDIANTE VISIÓN POR COMPUTADOR E INTELIGENCIA ARTIFICIAL USANDO TOF

*TECHNOLOGICAL ALTERNATIVE TO ESTIMATE DRIVER
DISTRACTION AND SLEEPINESS THROUGH COMPUTER VISION
AND ARTIFICIAL INTELLIGENCE USING TOF*

Fernando Alejandro Guevara Carrillo

Universidad Politécnica Salesiana, Ecuador
fguevarac1@est.ups.edu.ec

William Quitiaquez

Universidad Politécnica Salesiana, Ecuador
wquitiaquez@ups.edu.ec

Oswaldo Sebastián Valencia Bernal

EP Petroecuador, Ecuador
oswaldo.valencia@eppetroecuador.ec

Patricio Quitiaquez

Universidad Politécnica Salesiana, Ecuador
rquitiaquez@ups.edu.ec

Fernando Toapanta-Ramos

Universidad Politécnica Salesiana, Ecuador
ltoapanta@ups.edu.ec

Recepción: 30/octubre/2020

Aceptación: 8/diciembre/2020

Resumen

El documento presenta una alternativa tecnológica que detecta condiciones anormales del conductor a causa de la distracción y/o somnolencia. Utilizando la tecnología de tiempo de vuelo, se determinaron distancias de objetos con la emisión de una señal infrarroja en conjunto con la inteligencia artificial del modelo CANDIDE, además, se mostró el resultado de un algoritmo para seguimiento facial en tiempo real. Para la detección de la distracción fueron considerados tres movimientos faciales, además de un algoritmo de control que relaciona los valores de datos de los vértices, en el rostro del conductor. Finalmente, se obtuvo una interfaz visual

que indicó, mediante alertas visuales y sonoras, cuando el rostro del conductor se encontraba fuera de los parámetros aceptados para la conducción. El sistema es autónomo y no debe ser activado y/o configurado por el conductor, utilizó la visión artificial como método de seguimiento, contribuyendo en la reducción de accidentes de tránsito

Palabras Clave: Detección de objetos, distracción, Candide, Kinect, time-of-flight.

Abstract

The document presents a technological alternative, that detects abnormal driver conditions due to distraction and / or drowsiness. Using time-of-flight technology, objects distances were determined with the emission of an infrared signal, in conjunction with the artificial intelligence of the CANDIDE model, besides, the result of an algorithm for real-time facial tracking was shown. For the detection of distraction, three facial movements were considered, in addition to a control algorithm, that relates the data values of the vertices in the driver's face. Finally, a visual interface was obtained, that indicated by visual and audible alerts, when the driver's face was outside the accepted parameters for driving. The system is autonomous and should not be activated and / or set by the driver, it used artificial vision as a monitoring method, contributing to the reduction of traffic accidents.

Keywords: *Candide, distraction, Kinect, object detection, time-of-flight.*

1. Introducción

La distracción y/o somnolencia al conducir son fenómenos complejos, que implican una reducción en los niveles de alerta y conciencia de parte del conductor. Estas situaciones conllevan accidentes eludibles ante la identificación de situaciones peligrosas y por cuanto a evitar riesgos [Li, 2018]. El cansancio mental y físico provocan el adormecimiento del conductor y representan un factor que contribuye a los accidentes al menos en un 8.1% de ellos; en cambio la distracción contribuye al menos en un 20% de los accidentes en el Ecuador [Agencia Nacional de Tránsito, 2019]. El cansancio y la somnolencia de los conductores durante la conducción son descritas como causas de accidentes de tránsito en la ciudad y las

carreteras. La información está debidamente sustentada por autores como Lyznicki, Doege , Davis y Williams [Lyznicki 1998]. Un conductor fatigado disminuye progresivamente su capacidad de atención y concentración durante el manejo y pierde capacidad de respuesta ante condiciones específicas que exigen reacciones inmediatas [Brown, 1994].

Habitualmente, los accidentes producidos en estas circunstancias tienen alta siniestrabilidad en términos de pasajeros muertos, heridos y pérdidas materiales. Entre el 4 y 30% de los accidentes en carreteras son producidos por el cansancio durante la conducción de vehículos [Mitler, 1997). La Agencia Nacional de Tránsito [Agencia Nacional de Tránsito, 2019] determinó que el factor humano ocasiona el 79.2% de los accidentes de tránsito en el Ecuador, de ellos el 8.1% son ocasionados por circunstancias de somnolencia del conductor y el 20% por distintas distracciones del conductor. Un resumen de las condiciones de los accidentes de tránsito en el Ecuador (enero–septiembre 2019) se visualiza en figura 1 y tabla 1.



Figura 1 Causas de los Accidentes de Tránsito en el Ecuador, enero-septiembre 2019.

Tabla 1 Accidentes de Tránsito en el periodo enero – septiembre 2019.

| CONDICIÓN | CIRCUNSTANCIA | SINIESTROS | PORCENTAJE |
|---|---|--------------|----------------|
| C1 | Conducir en estado de somnolencia o malas condiciones físicas (sueño, cansancio y fatiga) | 203 | 1.11 |
| C2 | Conduce bajo la influencia de alcohol, sustancias estupefacientes o psicotrópicas y/o medicamentos | 1288 | 7.01 |
| C3 | Conducir desatento a las condiciones de tránsito (celular, pantallas de video, comida, maquillaje o cualquier otro elemento distractor) | 3678 | 20.03 |
| C4 | Accidentes por exceso de velocidad | 2923 | 15.92 |
| C5 | Accidentes por culpa de los peatones | 316 | 1.72 |
| C6 | Otras circunstancias de accidentes por factores humanos | 9379 | 51.08 |
| C7 | Accidentes por factores externos | 575 | 3.13 |
| Total de accidentes, enero - septiembre 2019 | | 18362 | 100.00% |

Existen proyectos desarrollados con el seguimiento facial que proporciona la tecnología Time-of-Flight de la cámara 3D Kinect. En las diferentes investigaciones poseen diversos alcances, algunos aplicables a la conducción, pero todos tienen mucha relación al tema del presente documento. Los escáneres tridimensionales tienen el principio de poder mostrar los objetos reales en un software de computador, obteniendo así una forma digital para ser manipulados y modificados a gusto de las personas que tengan acceso a estos escáneres [Musuña, 2014].

Feldstein, Güntner y Bengler [Feldstein, 2015] realizaron un estudio con 15 participantes para examinar los movimientos típicos de la cabeza del conductor mientras se conduce. El estudio reveló en general ligeros movimientos verticales y laterales de la cabeza, así como pequeñas posiciones verticales y laterales diferentes entre los participantes. Se desarrolló un método de seguimiento de cabeza basado en estos resultados empleando conversión 2D a 3D usando la cámara termográfica 3D del sensor Kinect.

Solaz et al. [Solaz, 2016] desarrollaron un sistema capaz de detectar el nivel de somnolencia en función de los movimientos involuntarios del conductor provocados por la respiración capturada por medio de un sistema de cámaras automotrices de tamaño reducido de alta dinámica como sensores de frecuencia respiratoria. Las imágenes capturadas por estos sensores se procesan para obtener el movimiento del pecho y abdomen del conductor.

Darby, Sánchez, Butler y Loram [Darby, 2016] compararon las estimaciones de la postura de la cabeza con Kinect, mediante un registro fundamental en marcadores para establecer la precisión. Cuando las características faciales permanecen visibles, el sensor tiene aplicaciones en la evaluación no invasiva del control de la postura.

Masood et al. [Masood, 2018] presentaron un modelo de aprendizaje automático que utiliza redes neuronales convolucionales para no sólo detectar al conductor distraído, sino también identificar la causa de su distracción, mediante el análisis de las imágenes obtenidas utilizando el módulo de cámara instalada dentro del vehículo. Actualmente, existen dos tipos de sistemas de interacción entre el conductor y el vehículo. El primero es un sistema de información en el vehículo IVIS,

tales como radios de automóviles, dispositivos móviles de manos; que implican típicamente demandas visual-manuales-cognitivas, las mismas que se requieren para la tarea de conducción, por lo tanto, es muy probable que causen interferencias [Wickens, 2002]. El segundo es un sistema avanzado de asistencia al conductor ADAS, los cuales son dispositivos auxiliares electrónicos en los vehículos de motor para apoyo al conductor en determinadas situaciones de manejo [Kutila, 2006].

Los desarrollos tecnológicos anteriormente mencionados guardan relación con la fatiga del operador al volante y se han caracterizado en este trabajo de acuerdo con los siguientes objetivos: prevención de la distracción y/o somnolencia, detección de la distracción y/o somnolencia y prevención de accidentes debidos a la distracción y/o somnolencia.

El tiempo de vuelo (ToF, por sus siglas en inglés) de imágenes se refiere al proceso de medición de la profundidad de una escena, mediante la cuantificación de los cambios que una señal de luz emitida encuentra cuando rebota de objetos en una escena [Chen, 2014]. Con la medición de la profundidad usando la tecnología ToF existen ciertas ventajas: sólo se requiere de una cámara específica, no se requiere de un cálculo manual para la profundidad, adquisición de la geometría de la escena 3D en tiempo real, menor dependencia de iluminación de la escena y casi sin depender de textura de la superficie.

2. Métodos

El software puede realizar el análisis de las señales entrantes para identificar la dirección de una fuente de sonido particular y permitir que el sensor rechace los sonidos no deseados y de los ecos [Miles, 2012]. Kinect tiene la capacidad única de ver en 3D. A diferencia de la mayoría de los otros sistemas de visión por ordenador, el sistema Kinect es capaz de construir un mapa de profundidad de la zona en frente de ella. Este mapa se produce totalmente dentro de la barra de sensor y luego se transmite por el cable USB al host en la misma forma que una imagen de cámara típica transmite su información [Miles, 2012].

El sensor de profundidad utiliza un tipo de radar o transmisor de sonido ultrasónico para medir los objetos que están lejos de la barra de sensores, pero en realidad no

es así. El sensor utiliza una técnica inteligente que consiste en un proyector de infrarrojos y una cámara que puede ver los pequeños puntos que el proyector produce.

Hay dos elementos en la cámara de profundidad del Kinect, como se indica en la figura 2. El primero es el proyector de infrarrojos que proyecta un campo de puntos frente a la escena de la barra de sensores. El segundo elemento es una cámara infrarroja que observa los puntos de la escena. La posición de los puntos en la escena reflejada depende de la posición con la que la cámara refleja la escena [Webb, 2012].



Figura 2 Elementos de la cámara de profundidad del Kinect [Miles, 2012].

CANDIDE es una máscara parametrizada específicamente desarrollada para la codificación de rostros humanos, basados en su modelo y constitución, mostrada en la figura 3. Su bajo número de polígonos (aproximadamente 100) permite la reconstrucción facial rápida con una potencia de cálculo moderada.

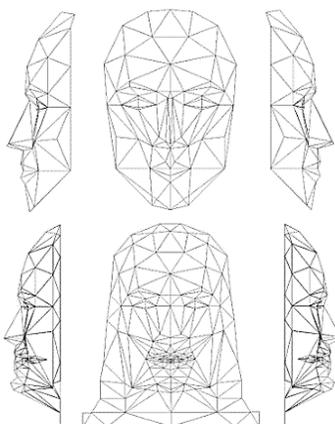


Figura 3 CANDIDE-1 con 79 vértices y 108 triángulos 238 triángulos.

El modelo Candide-3 a diferencia del Modelo Candide-1 ha realizado modificaciones en la estructura de la boca para obtener vértices correspondientes a los puntos característicos externos e internos del labio y se han añadido también vértices entre las esquinas de la boca y el centro de los labios (vértices 79-89, MPEG-4 FFPs en el grupo 2 y 8). Los vértices correspondientes a cada mejilla (vértices 27 y 60) han sido reemplazados por dos vértices correspondientes a los puntos faciales característicos 5.2 y 5.4 (lado derecho) y 5.1 y 5.3 (lado izquierdo), mostrados en la figura 4. Los nuevos vértices tienen los números 90 y 91, respectivamente.

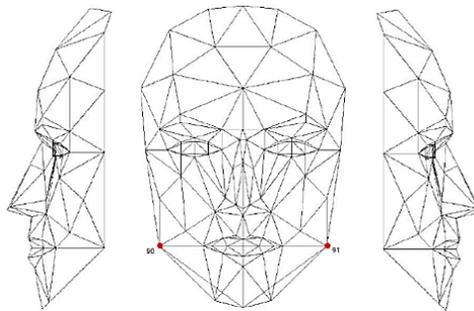


Figura 4 CANDIDE-3 con 113 vértices y 168 triángulos.

Existen tres vértices (92 - 94) en la parte media de la nariz, que corresponden a los nuevos Facial Feature Points (FFPs) agregados: 9.12, 9.13, y 9.14, y dos vértices (111 -112) correspondientes a los nuevos FFPs agregados: 9.4 y 9.5. Se han añadido vértices al contorno de los ojos, haciendo que el diseño sea más realista (vértices 95 - 110). Las unidades de acción definidas en CANDIDE-1 se amplían para incluir nuevos vértices. La animación puede ser realizada por MPEG-4 FAPs, así como por unidades de acción.

CANDIDE es un modelo similar a una malla de alambre con una textura mapeada en su superficie. Las coordenadas de los vértices se almacenan en un archivo *wfm*, y puede ser visto como un vector 3 – *N* dimensional (donde *N* es el número de vértices) que contiene las coordenadas (*x*, *y*, *z*) respectivas de cada vértice. El modelo entonces puede ser reconfigurado según Rydflak (1987) utilizando la ecuación 1.

$$g(\sigma, \alpha) = \bar{g} + S\sigma + A\alpha \quad (1)$$

Donde el vector resultante contiene los nuevos vértices de coordenadas (x, y, z) . Las columnas de S y A son la forma y unidades de animación, respectivamente, por lo tanto, los vectores σ y α contienen la forma y los parámetros de animación. Como también se requiere realizar un movimiento global, se necesita algunos parámetros más para la rotación, el escalamiento, y la traslación. Por lo tanto, se reemplaza la ecuación 1 con la ecuación 2.

$$g = RS(\bar{g} + S\sigma + A\alpha) + t \quad (2)$$

Donde R es la matriz de rotación (ecuación 3), S es la escala y t es el vector de traslación (ecuación 4).

$$R = R(r_x, r_y, r_z) \quad (3)$$

$$t = t(t_x, t_y, t_z) \quad (4)$$

La geometría del vector está por lo tanto parametrizada por el vector de parámetros, como se muestra en la ecuación 5.

$$p = [v, \sigma, \alpha] = (r_x, r_y, r_z, t_x, t_y, t_z, \sigma, \alpha) \quad (5)$$

Donde v es el vector de los parámetros globales de movimiento. El formato de archivo admite tener diferentes escaladas en las tres dimensiones, como se presentan en las ecuaciones 6 y 7.

$$g = RS_3(\bar{g} + F\sigma + A\alpha) + t \quad (6)$$

$$S_3 = S_3(S_x, S_y, S_z) \quad (7)$$

Las herramientas tecnológicas ayudaron a diseñar el algoritmo de control que detectó la somnolencia y/o distracción del conductor, se basaron en la tecnología de tiempo de vuelo para determinar distancias de objetos mediante el envío y recepción de una señal infrarroja, una Interface Natural de Usuario conformada por el sensor Kinect para Windows, que consta de una cámara infrarroja para la detección de objetos en tres dimensiones y el modelo CANDIDE, que entrega un rostro actualizado y parametrizado.

Una vez identificados los parámetros del modelo Candide-3, se implementó un algoritmo de inteligencia artificial, usando la biblioteca SDK de Kinect que detecta

las condiciones anormales de conducción y alerta al conductor de su errónea forma de conducir mediante una alarma.

Dentro del SDK de Windows se encuentra una aplicación básica del seguimiento de rostros, la misma que utiliza las clases (*FaceModel.cs* - *FaceTrack.cs* - *FaceTrackFrame.cs* - *Image.cs* - *Utils.cs*) que forman parte del proyecto *Microsoft.kinect.toolkit.facetracking*, tal y como se muestra en la figura 5. La aplicación forma una malla de triángulos en el rostro permitiendo el reconocimiento de algunas expresiones faciales como apertura y cierre de la boca o movimientos de la cabeza, figura 6. Una vez que se obtuvo la malla el siguiente paso fue obtener los puntos vértices del modelo *CANDIDE-3*, como se visualiza en figura 7.

Para ello se desarrolló un algoritmo dentro del cual se crea una variable, la misma que contuvo una lista de los puntos faciales, cada uno con su respectiva coordenada que entrega la clase *SkeletonFaceTracker*.

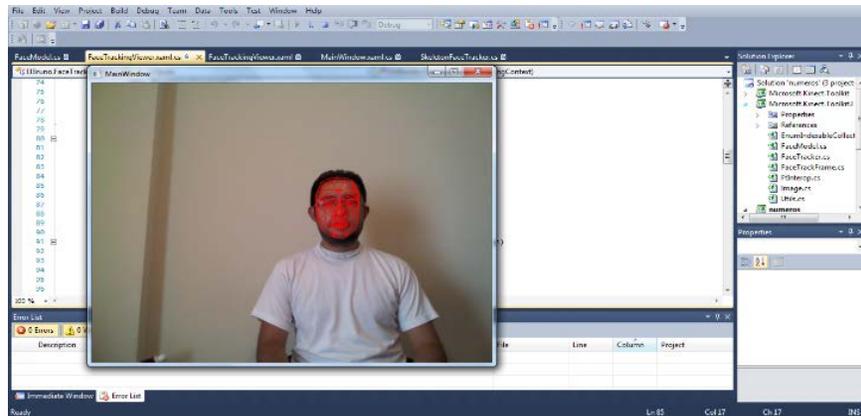


Figura 5 FaceTracking Basic.



Figura 6 FaceTracking Basic Movimiento de la Cabeza.

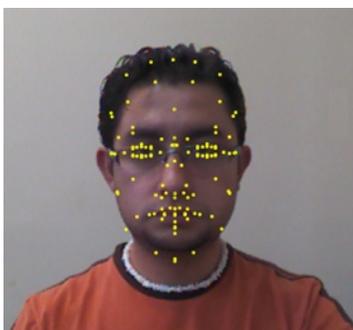


Figura 7 Puntos Faciales CANDIDE-3.

Para la aplicación que se desarrolló no se necesitó todos los puntos del modelo CANDIDE 3, sino un grupo mínimo que permitió conocer la posición actual del conductor. En tabla 2 se describe las diferentes APIs usadas en la presente investigación, las mismas fueron creadas en distintas clases cada una con código de programación que se enlazaron para formar el algoritmo de programación, así como una breve descripción de lo que hizo cada una.

Tabla 2. Descripción de las APIs que conforman el Algoritmo de Programación

| APIs USADAS | DESCRIPCIÓN |
|---|---|
| KinectSensor.Start KinectSensor.Stop | Inicio o parada de la transmisión de datos. |
| ColorImageStream.Enable ColorImageFormat.RgbResolution640 x480Fps30 | Elige el formato color y habilita al sensor para enviar datos a color. |
| DepthImageStream.Enable DepthImageFormat.Resolution320x24 0Fps30 | Elige el formato profundidad y habilita al sensor para enviar datos de profundidad. |
| DepthImageStream.Range | Establece la profundidad del sensor en modo cercano. |
| SkeletonStream.TrackingMode SkeletonTrackingMode.Seated | Habilita el seguimiento del esqueleto en modo sentado. |
| SkeletonStream.Enable | Habilita al sensor a transmitir datos del esqueleto. |
| KinectSensor.AllFramesReady | Agrega un controlador de eventos a todas las transmisiones. El sensor enviará una señal al controlador de eventos cuando cada nuevo cuadro de datos de color, profundidad o esqueleto se encuentre listo. |
| ColorImageFrameReadyEventArgs.O penColorImageFrame | Toma un cuadro de datos de color en un evento controlado. |
| ColorImageFrame.CopyPixelDataTo | Copia un cuadro de datos de color para pre colocarlo en un arreglo, este puede ser escrito a un mapa de bit re escribible opcional. |
| Skeleton.TrackingState SkeletonTrackingState.Tracked SkeletonTrackingState.PositionOnly | Determina si un esqueleto es seguido, y toma su identificación de seguimiento. |
| SkeletonFaceTracker class | Una clase personalizada para el almacenamiento de datos del esqueleto y la elaboración de una rejilla para el rostro posicionada sobre el esqueleto de la cabeza. |
| FaceTrackFrame class | Una clase personalizada que detecta un rostro base en la entrada de datos del Kinect, identifica los puntos característicos y sigue al rostro mientras este se mueve. |

El desarrollo del algoritmo de control ha sido desarrollado en lenguaje C# y consta de tres clases: MainWindows, FaceTrackingViewer y SkeletonTracker.

La clase MainWindows es la principal en la cual se inicializan los parámetros de la cámara para la captura de video.

La clase FaceTrackingViewer es la encargada de realizar el seguimiento del rostro del conductor, esto implica que inicializa las variables de todo el algoritmo, ejecuta el seguimiento del rostro cuando el sensor está listo y da inicio a la presentación de puntos en el rostro.

La clase SkeletonTracker posee el algoritmo destinado al filtro de puntos que se desea que se grafiquen en el rostro, así como la forma en la que se envía la información de estos hacia la clase FaceTrackingViewer, además también se encuentra el algoritmo perteneciente a las alarmas y contadores de los tiempos en que las mismas son accionadas. Dentro de la clase SkeletonTracker se realizó la creación de una variable que hizo de filtro, que contiene una lista de los vértices a ser utilizados, Como se observa en la figura 8.

Los vértices del modelo CANDIDE-3 utilizados fueron: 0 Parte alta del cráneo, 2 Parte media de la frente, 94 Mitad baja en el borde del hueso de la nariz y 10 Parte inferior de la barbilla. El proceso para la generación de alarmas de distracción se presenta en la figura 9.

Al inicializar el sistema, existe un variable que grafica los puntos filtrados en el rostro, luego del cual inicia el contador para determinar si las condiciones de Giro, Balanceo y Cabeceo se encuentran fuera de parámetros normales, activando las alarmas de ser el caso.



Figura 8 Vértices numerados filtrados.

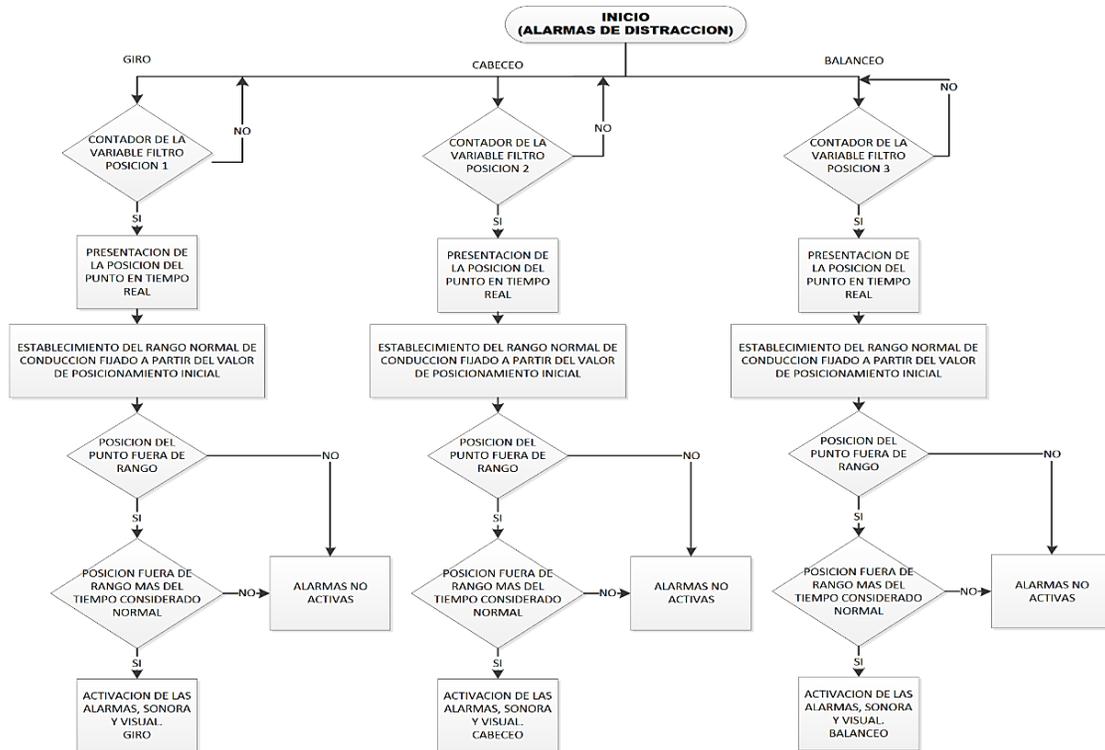


Figura 9 Alarmas para la detección de la distracción [Guevara, 2013].

3. Resultados y Discusión

El algoritmo para la detección de la distracción y/o somnolencia consta de variables para la inicialización de alarmas, contadores para la activación de alarmas y la secuencia para la alarma sonora. Cuando el programa principal se ejecuta se marca en el rostro del conductor los puntos principales. A su vez las posiciones iniciales de dichos puntos se guardan en distintas variables. La posición de cada punto es tomada como referencia para determinar las alarmas. Debido a que no todos los conductores poseen las mismas características físicas a partir de la posición inicial en la que los puntos se marquen en sus rostros se determinarán las alarmas para cada uno.

Las alarmas son accionadas cuando el o los puntos de referencia se desplacen de su posición inicial hacia fuera del rango normal de conducción y haya transcurrido un tiempo considerable de distracción. Para determinar el rango normal de conducción se tomaron las posiciones de los puntos, de varios conductores mientras se encontraban en el simulador y luego en un auto real cada cinco segundos.

Posteriormente se obtuvieron gráficas de la posición de cada punto en función del tiempo. Obteniendo como resultado un rango aceptable de conducción normal. Es complejo determinar un tiempo de distracción adecuado, debido a que este dependerá de la velocidad a la que se encuentra conduciendo el conductor, pero tomando en consideración un viaje a 100 km/h, en dos segundos recorrería una distancia de cincuenta y seis (56) metros, por lo tanto, para poder realizar las pruebas se consideró un tiempo de dos segundos. Además, cada vez que una alarma sea accionada la cámara tomará una foto de la posición del conductor en ese momento.

En la figura 10 se puede observar cómo está establecido el sistema de coordenadas teniendo como origen el centro óptico de la cámara. Las posiciones de los ejes X, Y, Z de la cabeza del conductor se determinaron mediante un sistema de coordenadas utilizando la regla de la mano derecha. Las traslaciones se miden en metros. Es así como la posición de la cabeza del conductor se medirá por tres ángulos: giro, balanceo y cabeceo, tal como se muestra en la figura 11.

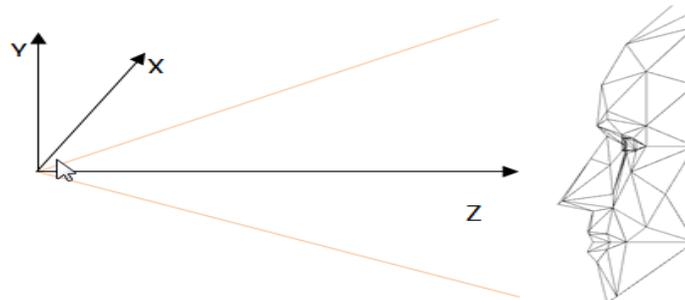


Figura 10 Sistema de coordenadas con origen en el centro óptico del Kinect.

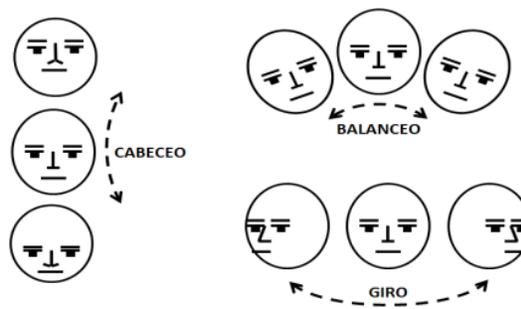


Figura 11 Posicionamiento de la cabeza.

Para el sistema de detección de distracción se creó tres alarmas, la alarma de giro teniendo como vértice referencial al que se encuentra en la mitad baja del borde del hueso de la nariz (94), figura 12. Para la alarma de balanceo el vértice utilizado como referencia es el que está posicionado en lo más alto del cráneo (0), mostrado en la figura 13. Para la alarma de cabeceo se consideró como vértice referencial al que está ubicado en la parte media de la frente (2), tal y como se muestra en la figura 14.



Figura 12 Alarma de giro encendida.



Figura 13 Alarma de balanceo encendida.



Figura 14 Alarma de cabeceo encendida.

El primer síntoma de la somnolencia es el bostezo, mismo que fue tomado como detonante de la alarma en el programa. Para detectar el bostezo del usuario, se tomó en cuenta la unidad de acción que interviene en la abertura de la boca y el tiempo en que la misma permanecía abierta. Para calcular la apertura de la boca, se trabajó directamente con un algoritmo que detecta el movimiento de la mandíbula inferior. Al abrir y cerrar la boca el algoritmo produce diferentes valores: cero cuando la boca se encuentra cerrada y mayor a cero dependiendo de la magnitud de la apertura. Si la magnitud de apertura de la boca es mayor o igual a la magnitud de un bostezo promedio se tendría la primera condición para detonar la alarma. La segunda condición es el tiempo que el usuario permanecería con la boca abierta, ver figura 15. Tomando en consideración un estudio realizado por Galluo, Church y Peregrino [Galluo, 2016] en el cual se indica que el bostezo de un ser humano va desde los tres a los siete segundos, para la programación se considera que deberá superar los tres segundos, dando paso a que la alarma de somnolencia se active.

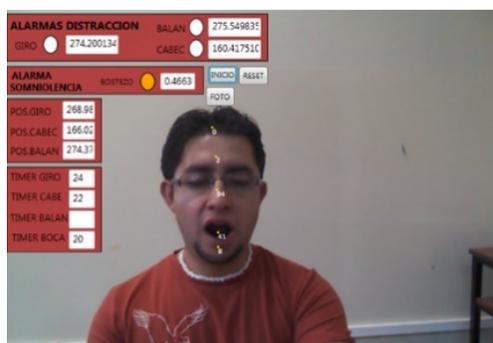


Figura 15 Alarma de somnolencia encendida.

Para la creación de las alarmas se realizaron pruebas de conducción con diferentes sujetos en un tiempo de 225 segundos, llegando a unos valores estándar de los movimientos de cabeceo, balanceo y giro. Estos datos fueron ingresados como valores umbrales en el código de programación. Se puede apreciar que existen valores comunes de tiempo en incremento o decremento en las alarmas de giro y cabeceo, pero con diferente magnitud por su traslación en cada eje. En la figura 16 se presentan los datos obtenidos para los umbrales de alarmas en el algoritmo de detección, de distracción y/o somnolencia.

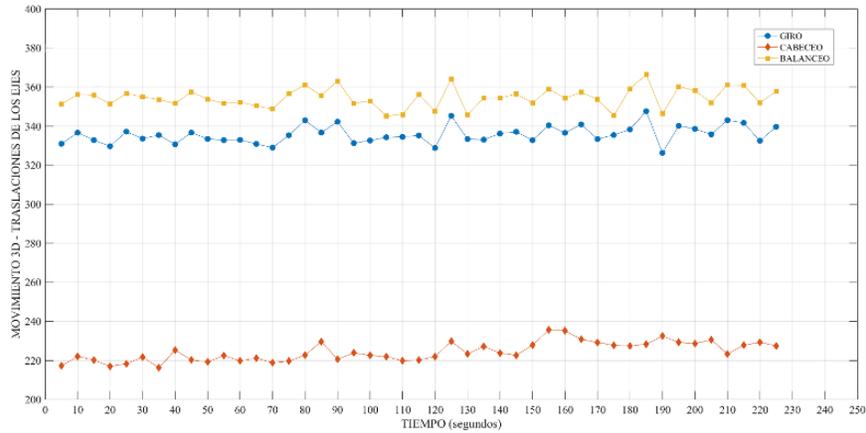


Figura 16 Datos para el ingreso al código de alarmas en el algoritmo.

Una vez que se realizó el código, el siguiente paso será instalar la cámara en un automóvil para la comprobación en una conducción real. Como se indica en las figuras 17 y 18, el algoritmo sirve para una conducción real, alertando las situaciones erróneas del conductor y previniendo accidentes de tránsito.



Figura 17 Alarma de distracción encendida, conducción real.



Figura 18 Alarma de somnolencia encendida, conducción real.

4. Conclusiones

Con el desarrollo de una alternativa tecnológica para detectar la distracción y somnolencia usando la inteligencia artificial de la cámara Kinect que posee tecnología ToF, se obtuvieron los siguientes resultados:

- Al momento de implementar el algoritmo en una conducción real, se modificó el flujo de datos del sensor Kinect debido a la exposición de rayos ultravioletas que cambiaron la recepción del infrarrojo; cambiando el flujo de datos de: 1280 x 960 bits de profundidad con resolución de 12 cuadros por segundo a 640 x 480 bits de profundidad con resolución de 30 cuadros por segundo. Se realizaron pruebas a dos personas, obteniendo resultados de la diferencia máxima entre los valores de las coordenadas y el valor máximo que sirvió como límite en la programación para los siguientes movimientos de distracción: Giro 30.06 - 15.03, Cabeceo 29.53 - 14.76 y Balanceo 28.33 - 14.16. Además, se determinó el valor máximo permisible para la somnolencia mediante la apertura de boca en unidades de acción de 0.2.
- Los límites en grados del campo de visión y giro del ser humano se clasificaron por la condición de sin esfuerzo y con gran esfuerzo, dando como resultado: Giro de la cabeza de 70 – 100 °, giro de los ojos 30 – 30 °, campo de visión aguda 30 – 30 °, zona visible 130 – 160 ° y zona no visible (lo que falta para llegar a los 180 °) 50 – 20 °. Esas condiciones de movimiento del cráneo ayudaron a obtener los rangos normales de conducción para el ingreso al algoritmo.
- La información que se obtiene de los rasgos faciales del conductor, mediante la tecnología ToF a través de una interfaz de visualización, permitirá que la detección de la cabeza del conductor sea indispensable en el futuro. La detección de las alarmas en la posición de la cabeza del conductor también podría usarse para otros fines y aplicaciones. También es posible monitorear la atención del conductor hacia el proceso de conducción y abre nuevas vías para la seguridad de los conductores, al desarrollarse nuevos trabajos de investigación con los datos obtenidos en este modelo.

5. Bibliografía y Referencias

- [1] Agencia Nacional de Tránsito, (09 de 2019). Reporte Nacional de Siniestros de Tránsito Septiembre 2019: <https://www.ant.gob.ec/index.php/descargable/file/6804-siniestros-septiembre-2019>.
- [2] Brown I., (1994). Driver fatigue. *Human Factors*. Human Factors: The Journal of the Human Factors and Ergonomics Society.
- [3] Chen M., Barwolff G., & Schwandt H., (2014). Time-Of-Flight technology applied in pedestrian movement detection. *The Conference on Pedestrian and Evacuation Dynamics*.
- [4] Darby J., Sánchez M., Butler P., & Loram I., (2016). An evaluation of 3D head pose estimation using the Microsoft Kinect v2. *Gait & Posture*, 48, 83-88.
- [5] Feldstein I., Güntner A., & Bengler K., (2015). Infrared-Based In-Vehicle Head Tracking: A Prototype for Tracking Drivers' Head Movements in Real Time. *Procedia Manufacturing*, 3, 829-836.
- [6] Gallup A., Church A., & Pelegrino A., (2016). Yawn duration predicts brain weight and cortical neuron number in mammals. *Biology Letters*.
- [7] Guevara F., & Valencia O., (2013). Análisis del conductor: Estimación de la distracción y somnolencia mediante visión por computador e inteligencia artificial usando tecnología ToF. Sangolquí: Universidad Politécnica del Ejército.
- [8] Kuttila M., (2006). *Methods for Machine Vision Based Driver Monitoring Application*.
- [9] Li Z., Cui L., & Chen J., (2018). Traffic Accident Modelling via Self-Exciting Point Processes. *Reliability Engineering and System Safety*.
- [10] Lyznicki J., Doege T., Davis R., & Williams M., (1998). Sleepiness, driving and motor vehicle crashes. *JAMA*.
- [11] Masood S., Rai A., Aggarwal A., Ahmad D., & Ahmad M., (2018). Detecting Distraction of drivers using Convolutional Neural Network. *Pattern Recognition Letters*.
- [12] Miles R., (2012). *Using Kinect for Windows with XNA, Kinect for Windows SDK*. Edición 1.1.

- [13] Miles R., (2012). Learn Microsoft Kinect API.
- [14] Mitler M., Lipsitz J., Walsh J., & Wylie C., (1997). The sleep of long-haul truck drivers. *N Engl J Med*, 337, 755- 761.
- [15] Musuña J., Zapata B., Oñate L., & Campusano G., (2014). Diseño y construcción de un robot móvil que permita la obtención de una nube de puntos del escaneo de habitaciones utilizando láser y webcams. *Ingenius*, 11, 53-61.
- [16] Rydfalk M., (1987). CANDIDE, A parameterized face. Linköping University, Sweden: Report No. LiTH-ISY-I-866, Dept. of Electrical Engineering.
- [17] Solaz J., Laparra Hernández J., Bande D., Rodríguez N., Veleff S., Gerpe J., & Medina E., (2016). Drowsiness detection based on the analysis of breathing rate obtained from real-time image recognition. *Transportation Research Procedia*, 14, 3867 – 3876.
- [18] Webb J., & Ashley J., (2012). Beginning Kinect Programming with the Microsoft Kinect SDK.
- [19] Wickens C., (2002). Multiple resources and performance prediction. *Theoretical issues in ergonomics science*, 159-177.