

 INSTITUCIÓN UNIVERSITARIA DE ENVIGADO	ARTICULO DE TRABAJO DE GRADO	Código: F-PI-028
		Versión: 01
		Página 1 de 11

DISEÑO DEL SOFTWARE DE RECONOCIMIENTO BIOMÉTRICO PARA LA DETECCIÓN DEL ESTADO DE SOMNOLENCIA EN CONDUCTORES DE AUTOS

HADER FERNEYDER CARDONA SUAREZ

cardona.hader@gmail.com

JORGE HUMBERTO MUÑOZ HINCAPIE

jorhnh@hotmail.com

Resumen: En este trabajo se presenta un diseño de reconocimiento de rostros, para detectar por medio del análisis de ojos, boca y cabeceo la somnolencia de una persona, basado en el procesamiento digital de imágenes; el manejo de imágenes, las formas de análisis y procesamiento por filtros, realces, histogramas y modelos de colores dentro del trabajo.

Se presenta información acerca del software: las necesidades, casos de uso y características que debe contener para que funcione de forma adecuada y con condiciones ideales para un rendimiento alto y las posibilidades de efectividad sean muy cercanas a los objetivos planteados.

Palabras clave: histograma, HSV, Opencv, somnolencia.

Abstract: In this paper, a face recognition design is presented to detect by analysis of eyes, mouth and pitch the drowsiness of a person, based on digital image processing; the image management, the analysis form and filters processing, enhancements, histograms and color models within the work.

It presents information about the software: requirements, use cases and features that must contain for this to work properly and with ideal conditions in order to have high performance and thus, the possibilities of effectiveness are very close to the objectives.

Key words: histograms, HSV, Opencv, somnolence.

 INSTITUCIÓN UNIVERSITARIA DE ENVIGADO	ARTICULO DE TRABAJO DE GRADO	Código: F-PI-028
		Versión: 01
		Página 2 de 11

OBJETIVO GENERAL:

Diseñar un software de reconocimiento biométrico para la detección del estado de somnolencia en conductores de autos.

OBJETIVOS ESPECIFICOS:

-Definir el dispositivo de video, que permita la adquisición de datos en forma digital, para un posterior procesamiento de los mismos.

-Diseñar un sistema de reconocimiento biométrico de rostros utilizando la librería de OpenCV en el lenguaje de programación Python.

-Utilizar técnicas de procesamiento digital de imágenes para determinar estado de somnolencia bajo análisis estadístico.

1. INTRODUCCIÓN

“Los accidentes de tráfico son una de las principales causas de muerte y de lesiones físicas que afectan especialmente a los jóvenes, y que además siguen un patrón creciente en el tiempo.

Se estima que entre el 3% y el 30% de los accidentes de tráfico son atribuibles a la somnolencia, por lo que indudablemente se trata de un tema con una gran repercusión.”(Pertusa, 2010)

A partir de esto, se muestra que la somnolencia es una causa importante de la accidentalidad, ahora en un tiempo en

donde las personas aumentan sus responsabilidades, sufren de estrés y laboran jornadas extensas se convierte en un tema de mucho cuidado a la hora de tomar un volante.

El presente trabajo requirió una investigación inicial en área metropolitana, a nivel nacional y finalmente a nivel mundial, buscando las estadísticas de accidentalidad. Localmente se encontró una situación difícil ya que los accidentes por quedarse dormido se reportaban como otros o de causa no definida (cuando no se determina por parte de los implicados o de testigos el causal), lo que dificultó las estadísticas. En algunos casos nacionales los registros los realizaban las empresas de transporte quienes se encargaban de llevar el seguimiento.

“Según los datos que maneja la ESRS, el sueño está implicado de alguna forma en entre un 20% y un 30% de los accidentes, aunque pocas veces se tiene en cuenta.” (Lucio, 2013). Para este caso la ESRS (la sociedad de investigación del sueño europeo, por sus siglas en inglés) arroja cifras alarmantes para este continente, pero se observa nuevamente que no se le da la importancia que este tema requiere, aunque empresas automovilísticas comienzan a tomar conciencia y desarrollan dispositivos para alertar en casos de adormecimiento.

En cuanto a una salida para este problema, se plantearon varias soluciones desde la ingeniería de sistemas dando como resultado el campo de la biometría, en donde se tratan temas de

reconocimiento de rasgos faciales, pasando por el procesamiento de imágenes, sus definiciones, características y aplicación dentro del trabajo investigativo.

2. ADQUISISION DE LA IMAGEN

Una imagen digital se representa en una matriz $n \times m$ donde cada elemento de la matriz equivale a un pixel, o sea un punto que contiene un color específico dentro de un rango de 0-255 que implica un color de blanco a negro. Además la imagen a color cuenta tres bandas básicas en RGB (rojo, verde, azul).

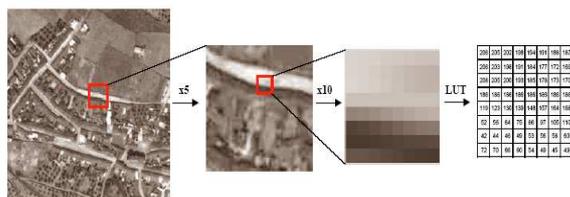


Figura 1. Representación visual y numérica (LINAZA, 2008)

Algunas de las características físicas son el flujo emitido por unidad de ángulo sólido (es el ángulo espacial de un objeto visto desde un punto) y área proyectada (radiancia térmica) y la fracción que se refleja se denomina reflectividad, influye además la radiación solar sobre un objeto que se atraviesa, conocida como transmisividad.

2.1 Tipos de archivos de imágenes.

PCX: formato utilizado por la empresa Zsoft en un programa llamado paintbrush, esto se remonta a la década de los 80, al día de hoy no es común encontrarse un archivo en este formato.

BMP: formato estándar de las imágenes de mapas de bits de Microsoft, generalmente para fondos de pantalla, tienen un problema de baja compresión lo cual los hace más pesados que otros formatos, en los últimos sistemas operativos Microsoft ha migrado a JPG buscando minimizar tamaños en estos.

PSD: formato definido de adobe photoshop, permite realizarle múltiples manipulaciones, otra de las ventajas de este formato es la posibilidad de dividir la imagen por capas o sea trabajar los diferentes elementos de la imagen por separado

EPS: llamado encapsulamiento postscript es admitido en la mayoría de programas de edición, tiene la ventaja de tener gráficos vectoriales como mapas de bits pero una gran desventajas al tratar de imprimir, al igual que los plotter esta imagen debe ser impresa en una impresora postscript.

TIF: formato estrella de Microsoft porque se permite en casi todos las aplicaciones de auto edición y modificación de imágenes, otra ventaja es la compatibilidad con Windows y Mac, tiene una desventaja cuando en un solo tif se unen varias imágenes y comienzan a aumentar de peso, se hace muy difícil navegar por el archivo y puede incluso bloquear el equipo en el cual se abre.

GIF: se usa generalmente para tonos no continuos o cuando se presenta un área grande de un solo tono, ya que se usa con color indexado que cuenta con 256 colores. La gran ventaja que tiene este

 INSTITUCIÓN UNIVERSITARIA DE ENVIGADO	ARTICULO DE TRABAJO DE GRADO	Código: F-PI-028
		Versión: 01
		Página 4 de 11

formato es poder elegir uno o varios colores y colocarlos transparentes y así observar lo que está detrás de estos.

JPG o JPEG: es el formato más utilizado para guardar fotografías y otras imágenes con tono continuo. Tiene la capacidad de guardar toda la información en millones de colores (RGB), sin ser un archivo pesado y sin perder la calidad de la imagen, además la mayoría de aplicaciones y navegadores de internet son compatibles con su formato y su calidad.

PNG: estas son las siglas de Portable Networks Graphics, uno de los pocos formatos con compresión sin pérdidas, esto quiere decir que se puede disminuir en tamaño y peso sin perder su calidad, está especialmente diseñado para el manejo en la red, con capacidad para indexación para colores transparentes, semitransparentes, no se limita a una paleta de 256 colores, al igual que el jpg está diseñado con millones de colores pero no tiene la capacidad de animación como la gif.

2.1 Ruido en imágenes

Un factor importante a tener en cuenta en el momento de la detección de objetos por medio de imágenes es el ruido, puesto que esto puede alterar la calidad de la imagen, confundiendo el software de detección con objetos que no son de interés hasta el punto de no poder determinar con certeza lo que se está buscando.

Este ruido es generado por diversas fuentes entre las cuales está la interferencia electromagnética y eléctrica, defectos en la medición de distancia, fallas en los dispositivos, lo que genera una distorsión considerable en la calidad de la imagen.

2.2 Modelo de colores de imágenes

Los modelos de colores tienen la función específica de facilitar la estandarización de colores de alguna forma, quiere decir que es una representación de coordenadas 3D y un subespacio donde cada color es representado por un punto.

Estos modelos son generalmente orientados al hardware en el caso de los monitores, cámaras, pantallas digitales para dispositivos, programas de diseño con el modelo RGB, el modelo CMY (cian, magenta, amarillo) para los colores específicos en las impresoras a color.

2.2.1 Modelo RGB.

El modelo RGB indica que sus componentes se representan en espectrales primarias es decir, rojo, verde y azul.

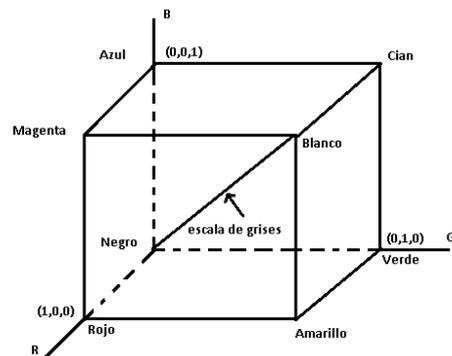


Figura 2. Tetraedro de color RGB

En la figura 2 se observa que los tres colores primarios se encuentran en los tres vértices, y el CMY se sitúan en los otros vértices, el color negro se encuentra en el origen y el blanco está en la diagonal que une el origen y los CMY. La escala de grises está dada por el origen en negro a través de su diagonal hasta el blanco y los colores dentro del tetraedro están definido por vectores, lo que significa que el RGB está en el rango (0,1), asumiendo que todos estos vectores han sido normalizados.

2.2.2 Modelo HSV.

El modelo HSV es un modelo que toma como base el RGB y varía en algunas propiedades del color agregándole (H) tonalidad, (S) saturación y (V) intensidad. Más específicamente el color Hue es el punto de partida, (rojo, naranja, azul, verde, etc), la saturación es la cantidad de blanco en un hue e intensidad o brillo es la luminosidad que un hue emite.

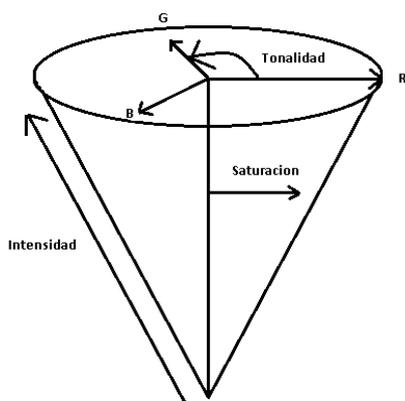


Figura 3. Diagrama del modelo HSV

A diferencia del RGB el modelo HSV puede utilizarse para controlar y medir el

nivel de saturación en una imagen además el brillo que un color pueda emitir para mejorar la calidad en cuanto a lo visual.

2.2.3 Escala de grises.

Es una representación de una imagen en donde cada pixel tiene un valor de luminancia que corresponde a un rango de blanco al negro, en una imagen con profundidad de 8 bits la posibilidad de colores en escala de gris va hasta 256.

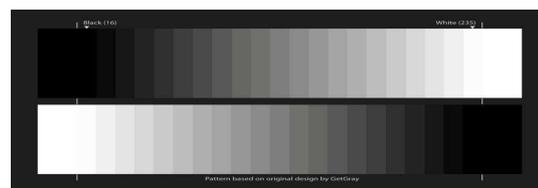


Figura 4. Ejemplo escala de grises (AVACAB-ONLINE, 2013).

3. PREPROCESAMIENTO DE LA IMAGEN

El preprocesamiento es un conjunto de técnicas que se utilizan para modificar de alguna forma las imágenes buscando resaltar o crear alguna característica específica con respecto a la imagen original, para esto es necesario:

Usar técnicas de realce de imagen, reducir el ruido mediante técnicas de filtrado y realce de bordes.

3.1 REALCE DE LA IMAGEN

La imagen digital se define estadísticamente por su histograma, cuando esto se realiza se puede notar el número de píxeles en la imagen en función de sus niveles de intensidad, es decir la frecuencia en al que estos

	ARTICULO DE TRABAJO DE GRADO	Código: F-PI-028
		Versión: 01
		Página 6 de 11

aparecen. Para que su contraste sea el deseado es necesario organizar o modificar la imagen con base a la real, esto significa aumentar o disminuir ciertos picos en donde se concentra la mayor parte de la intensidad, en una escala de grises por ejemplo, los rangos van hasta 256 niveles de gris, lo que significa que el trabajo debe ser correcto en cuanto a la nivelación de dichos colores.

La técnica de realce se basa fundamentalmente en transformar los niveles digitales de la imagen que debe cumplir dos condiciones ideales: que sea reversible el cambio y que la asignación de pixeles se encuentren en el mismo nivel tanto en la imagen original como en la realizada.

3.2 ECUALIZACION DE HISTOGRAMA

La ecualización es una técnica en donde se busca orientar los niveles de gris uniformemente, lo que hace que la imagen de salida quede mejor distribuida en su escala que la imagen de entrada.

Con esto se busca un mejor contraste y un mejor tratamiento del histograma de modo que los pixeles muy claros o muy oscuros estén en concordancia con el resto de la imagen. Además se debe tener en cuenta los colores y su resolución para realizar cambios y poder identificar los objetos, se presentara el siguiente paso del preprocesamiento.

3.3 FILTRADO DE IMAGENES

Continuando con el tratamiento de imágenes se procede a realizarle una convolución para suavizar la imagen y poderla tratar de una forma más pura y más exacta al momento de realizarle la detección.

3.3.1 Convolución

La convolución es un operador matemático denotado por $*$ para funciones cualquiera $f * g$ buscando transformarlas en una tercera, para darle diferentes terminaciones, algunos ejemplos son: enfocar, desenfocar, realzar bordes, detectar bordes, entre otros.

“la teoría aplicada al tratamiento digital de señales es sabido que el filtrado de las mismas se puede realizar bien en el dominio de la frecuencia bien en el dominio espacial. En el dominio de la frecuencia la operación de filtrado se reduce a una simple multiplicación de transformadas mientras que en el dominio espacial resulta ser una operación de convolución” (Pajares & De la Cruz, 2008).

La fórmula que se aplica para el dominio espacial es la siguiente:

$$f(x, y) * g(x, y) = h(x, y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, w)g(x - x, y - w) dx dw$$

Eq.1

Pero en el caso de las imágenes digitales estas se presentan por intervalos, es decir en forma discreta por lo cual es necesario cambiar la integral de dominio espacial a una sumatoria definida así:

$$F'(x, y) = \sum_{i=-g}^g \sum_{j=-h}^h w(i, j) f(x+i, y+j)$$

Eq. 2

	ARTICULO DE TRABAJO DE GRADO	Código: F-PI-028
		Versión: 01
		Página 7 de 11

Dónde:
x: columnas.
y: filas.
w: coeficientes de peso.
-g, g: ventana de filtro.
-h, h: ventana de filtro.

3.3.2 Filtros

Son procesos matemáticos para tratar de normalizar o estandarizar las imágenes por medio de contrastes en escalas de gris, esto quiere decir que se procesan los píxeles para que estos queden de forma similar a sus adyacentes o si se busca resaltar se transforma en un píxel de color aumentado.

En la figura 5 se puede observar un filtro 1/9 de 3X3.

	1	1	1
1/9 x	1	1	1
	1	1	1

Figura 5. kernel para suavizar la imagen

4. SEGMENTACIÓN

La segmentación de imágenes es la división de las imágenes en partes pequeñas en las que se aíslan colores, figuras u otras condiciones que son llamadas regiones de interés.

Estas regiones son llamadas ROI (región of interest), esta región determina una ubicación del objeto de interés y tiene la capacidad de cambiar rápidamente el tamaño de interés para acomodarse a diferentes imágenes que sean ingresadas.

5. EXTRACCIÓN DE CARACTERÍSTICAS

Esto se realiza por medio de fórmulas matemáticas, después de la segmentación y aplicación de ejemplos positivos se procede a detectar el rostro y los ojos de la persona ubicada en frente de la cámara.

5.1 DETECCIÓN DEL ROSTRO

La clave para la detección del rostro es definir de alguna forma el o los objetos a determinar luego de ubicar la región de interés dentro de la imagen, para esto el sistema recurre a xml de reconocimiento para cada parte de interés, este es:

haarcascade_frontalface_alt.xml: este xml permite enfocarse en la cara frontal del usuario.

Adicional a esto el xml contiene aproximadamente 7000 ejemplos positivos, esta definición significa que el programa está diseñado para detectar 7000 posibles registros morfológicos del rostro, algunos ejemplos se muestran en la siguiente imagen.

5.2 DETECCIÓN DE OJOS Y BOCA

Al igual que en el rostro el sistema detecta los ojos y la boca por medio de XML que cuentan con una cantidad entre

5000 y 7000 ejemplos positivos para poder identificar un promedio alto de características en las personas, los nombres de los XML son los siguientes:

haarcascade_mcs_mouth.xml: este xml permite enfocarse en la boca del usuario

haarcascade_eye.xml: este xml permite enfocarse en los ojos del usuario.

5.3 DETECCIÓN DE PARPADEO

Esta detección se realizara aplicando métodos de promedios, desviaciones estándar para determinar el estado del ojo, esto tomando en cuenta las últimas 10 imágenes captadas por la cámara; realizando una cola para que la primera imagen en entrar pueda ser la primera en ser procesada.

6. IDENTIFICACION DE OBJETOS

Para terminar, el software realiza una gráfica de los objetos detectados del rostro y los ojos por medio de cuadrados y diferentes colores para mantener una comunicación con el sistema principal, en caso de no detectar ni el rostro ni los ojos el programa procederá a realizar alertas, un ejemplo de la gráfica se puede ver en la figura 6.

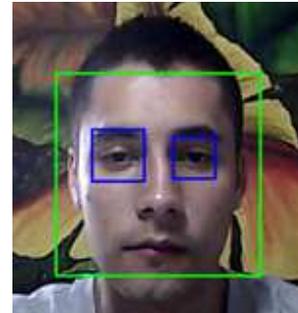


Figura 6. Detección de rostro.

7. CONDICIONES IDEALES DEL SOFTWARE.

Para correcta detección de todas las características faciales, es necesario un entorno ideal, buscando que las pruebas realizadas sean consistentes y puedan ser aplicadas las diferentes técnicas de procesamiento de imágenes digitales para crear la mayor confiabilidad del sistema.

La luz artificial tipo ahorradora tiene una potencia de 60 vatios, debe ser ubicada correctamente para que el rostro pueda verse con todo su color y así encontrar rápidamente los ojos, el contorno facial y el momento del parpadeo, para esto se tomaron las siguientes medidas:

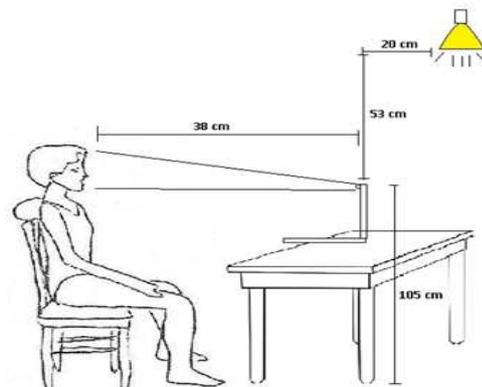


Figura 7. Condiciones ideales

	ARTICULO DE TRABAJO DE GRADO	Código: F-PI-028
		Versión: 01
		Página 9 de 11

CONCLUSIONES Y RECOMENDACIONES

-El mejor instrumento que se utilizó para el análisis y procesamiento de imágenes es la librería OpenCV, debido a que tiene una gran cantidad de funciones optimizadas para el software de biometría que se diseñó.

- Durante las pruebas realizadas al software se encontró una gran variedad de factores físicos que afectan en las personas la efectividad en la determinación del estado de somnolencia, como son el color de los ojos, el color de la piel y accesorios externos.

-Se determinó que el lenguaje de programación Python es una excelente herramienta para desarrollar proyectos de este tipo, porque contiene una fácil sintaxis, documentación y su acceso es sin ningún costo a diferencia de MatLab.

-Fomentar la investigación para los estudiantes de la Institución Universitaria de Envigado, que estos trabajos de investigación sirvan como base para que lo estudiantes se motiven a indagar y a conocer más sobre su campo de estudio.

-Si se implementara en vehículos, ayudaría a disminuir los accidentes debido a que es un mecanismo viable para evitar tantos desastres causados por los

conductores que se quedan dormidos luego de una jornada extensa de trabajo.

- La biometría es un campo amplio que permite a la humanidad solucionar problemas que quizás con otros métodos no se podrían solucionar o serían demasiado complejos y largos para ser aplicados, por ejemplo: reconocimientos faciales y seguridad de autenticación.

-un excelente manejo de las metodologías para desarrollar un software que sea competente en un mercado cambiante y constantemente exigente de mejoras y buenas prácticas.

- Al momento de diseñar un sistema de reconocimiento facial es importante conocer las limitaciones que este tendrá, debido a la diversidad en todas las personas; es recomendable buscar o crear un clasificador que permita disminuir el margen de error por la detección.

- Es importante motivar a los estudiantes a la investigación ya que permite formarse con más conocimientos en los diferentes ámbitos de la ingeniería y los prepara para enfrentar un nuevo proyecto en su ámbito laboral.

- El software de programación libre juega un papel importante en cuanto a la investigación y en la creación de grandes proyectos, esto queda demostrado con el uso de Python que posee características similares al MatLab.

	ARTICULO DE TRABAJO DE GRADO	Código: F-PI-028
		Versión: 01
		Página 10 de 11

- Los semilleros de investigación son indispensables en los trabajos de grado, por esto es importante que se realicen más y cada vez mejores, con su respectiva promoción y haciéndolos atractivos para todo el plantel educativo.

-La biometría es un gran campo investigativo que permite ser tomada desde diferentes aspectos como la identificación, la seguridad, el acceso a lugares, sistemas de votación, entre otros, que pueden ser explotados dentro de la Universidad de Envigado por compañeros de sistemas.

-Realizar un trabajo investigativo es el reto más fuerte en la carrera, por eso es importante un compromiso constante con la investigación y las asesorías que puedan brindarse dentro de la Institución Universitaria de Envigado.

accidentes de tránsito. [En línea]. Recuperado el 10 de octubre de 2013 de: <http://www.hola.com/salud/enciclopedia-salud/2010060645460/viajes/viajar/el-sueno-como-causante-de-accidentes-de-trafico/>.

LINAZA, Carlos. (2008). La imagen digital. [En línea]. Recuperado el 4 de mayo de 2013 de ftp://ftp.unsj.edu.ar/agrimensura/Fotogrametria/Unidad6/Tema_1_Fotogrametria_Digital_La_Imagen_Digital.doc.

AVACOB-ONLINE. (2013). Cómo se debe ver una pantalla: ajuste profesional de imagen [En línea]. Recuperado el 4 de mayo de 2013 de <http://avacablog.avacab-online.com/como-se-debe-ver-una-pantalla-ajuste-profesional-de-imagen/avacab>.

REFERENCIAS

LUCIO G, Cristina. (2013). ¿Cuántos accidentes de tránsito provoca el sueño? . [En línea]. Recuperado el 10 de octubre de 2013. De: <http://www.elmundo.es/elmundosalud/2013/08/05/noticias/1375713107.html>.

Pajares, G & De la Cruz, J. Visión por computador: imágenes y aplicaciones. (2008). México: Alfaomega Grupo Editor.

PERTUSA MARTINEZ, Salvador. (2010) El sueño como causante de

 INSTITUCIÓN UNIVERSITARIA DE ENVIGADO	ARTICULO DEL TRABAJO DE GRADO	Código: F-PI-03
		Versión: 01
		Página 11 de 11