

Detección del centro de la pupila usando la transformada de simetría radial

Autores:

Alex Fernando Collazos González

Rowanlee Einstein 21-S López Valenzuela

Director:

Ruthber Rodríguez Serrezuela

Corporación Universitari del Huila

Facultad de Ingeniería

Ingeniería Mecatrónica

2022

Resumen

El análisis de los movimientos oculares es de suma importancia para el estudio del funcionamiento del sistema vestibular y el diagnóstico de enfermedades neurodegenerativas y neuro-psiquiátricas, entre otras. Aunque la electro-oculografía es la técnica de registro de los movimientos oculares más utilizada, requiere de un diseño electrónico complejo para cumplir las normas de seguridad eléctrica. En cambio, la video-oculografía resulta más factible desde el punto de vista de seguridad eléctrica, costos y generalización. Sin embargo, esta requiere de métodos para detectar con exactitud la posición del centro de la pupila. En esta tesis se implementa un algoritmo, basado en la Transformada de Simetría Radial, para estimar la posición del centro de la pupila. El algoritmo se fundamenta en el cálculo de gradientes para detectar puntos de alta simetría radial y se implementó en MATLAB. El algoritmo se evaluó con imágenes infrarrojas de las bases de datos CASIA. La anotación y la detección del centro de la pupila se llevó a cabo usando la interfaz gráfica "LABELPUP" desarrollada en este trabajo. La evaluación con 400 imágenes demostró resultados competitivos con el estado del arte.

Palabras clave: seguimiento ocular, detección de pupila, video-oculografía, procesamiento digital de imágenes.

Abstract

Eye-movement analysis is crucial for the functional study of the vestibular system and the diagnosis of neurodegenerative and neuro-psychiatric diseases, among others. Although electro-oculography is the most widely used eye movement recording technique, it requires a complex electronic design to comply with electrical safety standards. On the other hand, video-oculography is more feasible from the point of view of electrical safety, costs and generalization. However, it requires processing methods to accurately detect the pupil center position. In this research work, we implement an algorithm based on the radial symmetry transform to estimate the pupil center position. It is based on gradient calculation in order to detect points of high radial symmetry. It was implemented in MATLAB. The algorithm was evaluated with infrared images from CASIA database. Annotation and pupil center detection processes were carried out using the graphical interface "LABELPUP" developed in this work. The evaluation with 400 images showed competitive results with the state of the art.

Keywords: eye tracking, pupil detection, videoculography, digital imaging processing.

TABLA DE CONTENIDOS

INTRODUCCIÓN	1
1 . CONSIDERACIONES GENERALES.....	6
1.1 Fisiología del ojo humano	6
1.2 Seguimiento ocular mediante VOG.....	8
1.2.1 Configuraciones para seguimiento ocular.....	8
1.2.2 Pasos para el seguimiento ocular.....	11
1.2.3 Aplicaciones del seguimiento ocular.....	12
1.3 Detección de pupila.....	12
1.3.1 Operador Integro-diferencial.....	13
1.3.2 Transformada de Hough.....	14
1.3.3 Modelo de contornos activos	16
1.3.4 Ajuste de elipse	17
1.4 Análisis de los movimientos oculares.....	18
2 . DETECCIÓN DEL CENTRO DE LA PUPILA	20
2.1 Bases de datos	20
2.1.1 CASIA.....	21
2.2 Algoritmo de detección implementado	23
2.2.1 Selección de los parámetros	30
2.3 Métodos usados para la anotación de imágenes	34
2.3.1 Ajuste de Elipse.....	34
2.3.2 Ajuste de Círculo	36
3 . RESULTADOS Y DISCUSIÓN	39

3.1	Metodología de evaluación	39
3.2	<i>LABELPUP</i> : Herramienta para anotación y detección del centro de la pupila	40
3.3	Evaluación del algoritmo de detección del centro de la pupila	49
3.3.1	Exactitud	49
3.3.2	Robustez	50
3.3.3	Costo computacional	55
3.4	Comparación de métodos	56
3.5	Discusión general.....	57
CONCLUSIONES Y RECOMENDACIONES		59
REFERENCIAS BIBLIOGRÁFICAS		62
Glosario de términos:		66

INTRODUCCIÓN

Los trastornos neurológicos en su conjunto constituyen una de las principales causas de muerte y discapacidad en el mundo y afectan a todos los grupos de edad. En los últimos años, el aumento de la incidencia de enfermedades neurodegenerativas, como el Parkinson o el Alzheimer, se está convirtiendo rápidamente en uno de los mayores retos de salud pública a nivel mundial. En Colombia se estima que por cada mil personas, 13.1 son diagnosticadas con Alzheimer [1] y 4.4 con Parkinson [2]. Actualmente no existe una cura y su diagnóstico se basa en pruebas invasivas, costosas o cognitivamente exigentes para los pacientes [3]; sin embargo, investigaciones recientes han señalado el estudio de los movimientos oculares como un indicador útil en el diagnóstico temprano y la progresión de los trastornos neurodegenerativos. El análisis de las diferentes características fisiológicas del sistema visual se realiza comúnmente utilizando tecnologías de seguimiento ocular (*eye-tracking*, en inglés), la cual ha sido un área en constante crecimiento. El seguimiento ocular es el proceso de medir el punto donde se fija la mirada. En la actualidad, los sistemas existentes para el seguimiento de la mirada presentan dos problemas principales: i) su precio es excesivo o ii) son demasiado invasivos para su uso rutinario.

Los métodos de seguimiento ocular comúnmente se dividen en tres categorías:

1- Electroculografía (EOG): Es una técnica no invasiva basada en el registro, mediante electrodos, de la diferencia de potencial existente entre la córnea y la retina (potencial corneo-retiniano). El principal inconveniente del uso de la técnica de EOG es que los niveles de la señal obtenida (electro-oculograma) son bajos (50 a 3500 μV) lo que la hace poco inmune a las perturbaciones ocasionadas en el sistema de adquisición y al ruido muscular. En sistemas de registro acoplados en corriente directa, el desapareamiento de los electrodos también puede constituir una fuente de error. Estos factores pueden ocasionar errores de aproximadamente 1° en la estimación de la dirección de la mirada, y por tanto, la dificultad de detectar con precisión los movimientos oculares inferiores a 2° . Los sistemas para seguimiento ocular usando EOG no son baratos por el costo asociado al aislamiento galvánico y a los electrodos.

2- Lentes de contacto: En este método se inserta en el ojo un anillo de silicona que contiene un pequeño enrollado de alambre de cobre. Cuando el sujeto se sitúa en un campo magnético de corriente alterna, la posición del ojo puede determinarse a partir de la amplitud de la corriente de

inducción en el enrollado. El bajo ruido permite lograr altas resoluciones espaciales ($\ll 1^\circ$) y temporales ($\ll 1$ ms) por lo que es considerado como el método más preciso para rastrear la mirada del usuario. Sin embargo, su principal desventaja es que es altamente invasivo e incómodo de usar. El uso de esta tecnología se limita, por lo tanto, a la investigación de laboratorio con modelos animales [4].

3- Videoculografía (VOG): Esta técnica utiliza cámaras para capturar las imágenes de los ojos. La señal de video es digitalizada y posteriormente procesada empleando técnicas de visión artificial para calcular la dirección de la mirada [5]. Esta técnica posee mayor movilidad y adaptabilidad y es preferible para pruebas pediátricas. Los equipos comerciales existentes requieren dispositivos especializados, por ejemplo, cámaras de video de bajo ruido y que transfieren tramas de alta resolución, así como fuentes de luz infrarroja (IR), los cuales elevan considerablemente sus precios.

Este método puede realizar el seguimiento ocular de dos formas: medición de la posición angular del ojo en relación con la cabeza (sistemas de video montados en la cabeza) y medición de la posición del ojo en relación con el entorno (sistemas remotos). Como regla general, los métodos que miden la posición del ojo con respecto a la cabeza son más precisos y están destinados al estudio de la dinámica oculomotora, mientras que los métodos que miden la posición del ojo en relación con el ambiente circundante se utilizan para medir el punto de mirada en una interfaz de usuario [6].

ANTECEDENTES DEL PROBLEMA

Los sistemas de videoculografía son los sistemas de seguimiento más populares hoy en día [4]. Las técnicas basadas en video han permitido reducir la invasividad del sujeto, gracias a la miniaturización de las cámaras, y por consiguiente, del tamaño y peso de estos dispositivos. Aunque resuelve el problema de la invasividad, el mayor inconveniente es el del alto coste de los dispositivos. Los sistemas de seguimiento de la mirada de tipo móvil que se pueden encontrar en el mercado comienzan a partir de los 15000€ (*ASLTM EyeTrac6 Eye Head*), y si además se desea contar con la característica inalámbrica, el precio asciende hasta los 45000€ (*TobiiTM Glasses Eye Tracker*). Claramente, el alto precio de estos dispositivos no es debido al hardware, pues el precio de las cámaras y otros componentes ha bajado exponencialmente en los últimos años. El precio está más asociado al software que se proporciona con estos dispositivos, así como al diseño. Estos programas, aunque muy completos, son raramente modificables para adaptarlos a distintas necesidades, ya que suelen ser de código cerrado y casi la totalidad de ellos corren

sobre *MicroSoft Windows*, con el coste de licencia añadido que ello supone. Todo lo anterior hace que los sistemas de seguimiento de la mirada actuales estén fuera del alcance de muchos de los centros de investigación o empresas, debido al alto coste del producto y a las dificultades actuales de acceso a la financiación.

Para cualquier tipo de sistema de VOG montado en la cabeza o sobre la mesa, los pasos involucrados para el seguimiento ocular son los siguientes: (a) adquisición de imagen/video, (b) localización de la cara, (c) localización del ojo, (d) detección de la posición de la pupila y (e) cálculo de la dirección de la mirada. Los pasos (a) y (b) dependen de la calidad y montaje de las cámaras usadas, que suelen ser infrarrojas porque logran mejor contraste entre el iris y la pupila. Los sistemas de video más usados (y más baratos) son los montados en la cabeza que usan gafas y cascos con cámaras, los cuales prescinden del paso (b). Una vez que el sujeto se ubica en el soporte de barbilla, y se enfoca la cámara en su ojo, el paso (c) es relativamente fácil de implementar y no requiere mucha precisión, por cuanto sólo se desea encuadrar el ojo, de forma tal que la pupila se distinga claramente. El paso (d) es el más complejo e importante porque su exactitud y precisión condiciona al paso (e) y al posterior diagnóstico realizado por el software de estos sistemas, por tanto es el gran reto de estos sistemas y el foco principal de este trabajo.

PLANTEAMIENTO DEL PROBLEMA

Hasta la actualidad han sido propuestos una gran cantidad de métodos y algoritmos con el objetivo de detectar la posición de la pupila. Entre los más usados están los basados en: la Transformada de Hough [7], el ajuste de elipse [8], los operadores integro-diferenciales [9] y la Transformada de Simetría Radial [10]. En una búsqueda realizada en Google académico, estos poseen 8620, 8500, 4175 y 750 citaciones, respectivamente. A pesar de que existen muchos métodos, estos suelen fallar debido a diversas causas como: cambio de la posición ocular respecto a la cámara, captura imperfecta debido a condiciones de iluminación, ruido en los datos adquiridos debido a sensores sucios o en mal estado por condiciones ambientales desfavorables, obstrucción parcial de la pupila por el párpado, pestañas y cabellos así como los destellos de luz, entre otros. Por todo lo anterior el problema de detección de la pupila con exactitud y robustez no se puede considerar como resuelto. La exactitud se refiere a la media del error en la estimación del centro de la pupila. La robustez es la forma en que el método tolera diferentes perturbaciones en la imagen manteniendo una buena exactitud.

Muchos de los métodos citados anteriormente han sido validados con imágenes propias, sin embargo, para la investigación científica resulta valioso el empleo de bases de datos de prestigio internacional debido a que constituyen una referencia para comparar los distintos métodos que se evalúan sobre estas imágenes. Las bases de datos públicas (libres de pago) más utilizadas para el seguimiento ocular son la UBIRIS (VOG de luz visible) y la CASIA (VOG de luz infrarroja). Dichas bases de datos no poseen anotaciones del centro de la pupila por lo que resulta imposible una adecuada evaluación y comparación de los métodos de detección de pupila. Para realizar una marcación (anotación) fiable del centro de la pupila en imágenes con perturbaciones se necesitan, simultáneamente, especialistas de mucha experiencia y herramientas potentes e interactivas que faciliten la labor de estos.

Existen diferentes herramientas software para la anotación manual de imágenes: DataTurks [11], Labelbox [12], AnnoStation [13] y Microsoft VoTT [14], entre otras, las cuales están realizadas en diversas plataformas (JS, HTML, Javascript, CSS, Python, Javascript, MATLAB, PHP, C++, PHP, Drupal, Wordpress, Perl, Adobe Flash y Angular 5, entre otras) y poseen diversos tipos de licencia software (de pago y gratis con diferentes niveles de uso). Sin embargo, todas fueron diseñadas para múltiples aplicaciones de delineación (casas, carros, etc.). Por otro lado, las herramientas especializadas para anotar la posición del centro de la pupila son muy escasas.

De ahí que el problema científico que se desea resolver es:

¿Cómo detectar la posición del centro de la pupila de forma exacta y robusta usando imágenes de VOG infrarroja?

JUSTIFICACIÓN

Como pasos preliminares en el diseño de un sistema que asista al personal médico en el diagnóstico de enfermedades neurodegenerativas a través del estudio de los movimientos oculares y, en aras de avanzar hacia el uso de la VOG, resulta necesario detectar la posición del centro de la pupila y evaluar la exactitud y precisión de dicha detección. El hecho de poder desplegar un sistema de este tipo por instituciones hospitalarias de Colombia representa un sustancial ahorro económico pues la importación del equipo a compañías extranjeras resulta excesivamente costoso.

OBJETIVOS

Objetivo general:

Implementar y evaluar un algoritmo para la detección del centro de la pupila en imágenes de VOG infrarroja.

Objetivos específicos:

- i. Anotar las imágenes de la base de datos CASIA.
- ii. Desarrollar en MATLAB un algoritmo para detección de pupila.
- iii. Evaluar el algoritmo desarrollado con las imágenes de la base de datos CASIA.
- iv. Desarrollar una herramienta para integrar los objetivos específicos (i) y (iii)
- v. Comparar el desempeño del algoritmo implementado con otros existentes.

ALCANCE DEL PROYECTO

El proyecto tiene grandes perspectivas y posibilidades de aplicación en diversas ramas como la medicina, robótica, estudios de mercado y evaluación de choferes, entre otras. En particular, esta tesis se enfoca en la implementación de un algoritmo para la detección de pupila que se pueda incorporar como un módulo a un sistema de VOG.

1 . CONSIDERACIONES GENERALES

En este capítulo se describen algunas características fisiológicas del sistema de la visión haciéndose énfasis en la estructura del ojo humano, específicamente en la pupila. Se resumen los aspectos más importantes sobre el seguimiento ocular usando la técnica de videoculografía.

1.1 Fisiología del ojo humano

El conocimiento completo de la anatomía del ojo es requisito indispensable para la interpretación adecuada de enfermedades con manifestaciones oftálmicas. El ojo humano es un órgano que detecta, filtra y convierte un pequeño espectro de energía electromagnética en información significativa, que puede ser interpretada por el cerebro. Seis músculos, organizados en tres pares, dan al ojo tres grados de libertad. Los diferentes pares de músculos controlan los movimientos horizontales, verticales y de rotación alrededor de la dirección de la vista. Tres partes del ojo son visibles desde el exterior: la esclerótica (blanca y curvada), la pupila (negra), y el iris (de colores) tal como se muestra en la figura 1.1.

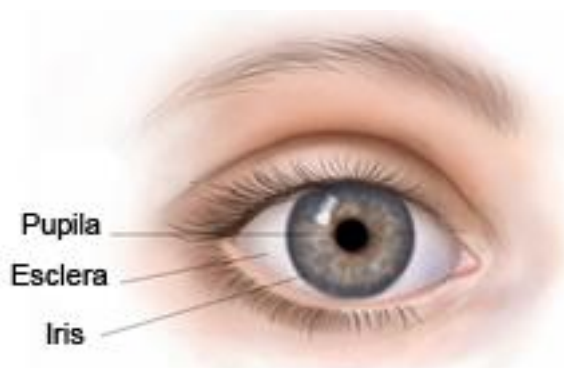


Figura 1.1 Vista frontal del ojo humano. (Fuente:[15])

La forma del ojo, en el adulto normal, es casi esférica y su diámetro anteroposterior mide, en promedio, 24 milímetros. Se encuentra rodeado por cuatro membranas, la córnea, la esclerótica, la coroides y la retina. Esta última es la parte más sensible del ojo, ya que está cubierta por infinidad de células receptoras. La retina, o parte interna del globo ocular está funcionalmente

organizada para que la máxima resolución del detalle del espacio pueda ser obtenido en un punto o área de clara visión llamada fovea.

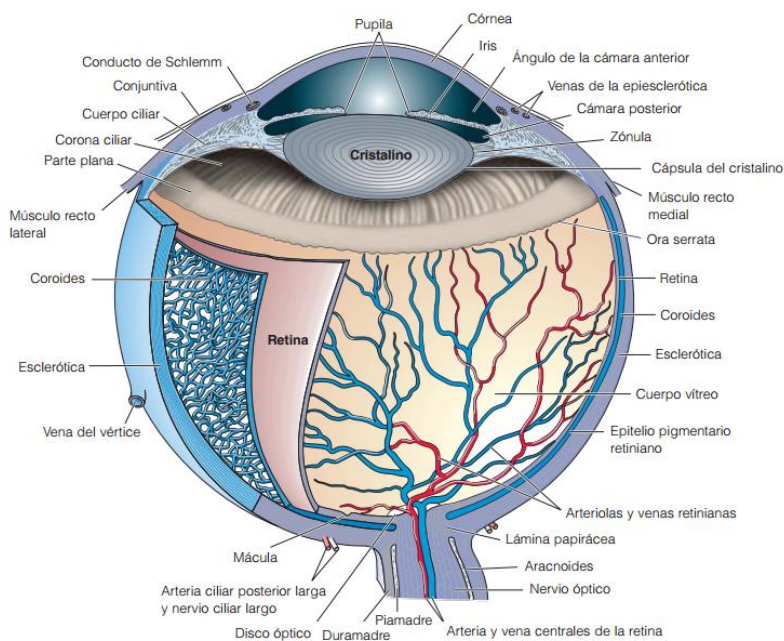


Figura 1.2 Estructuras internas del ojo (Fuente:[16])

La túnica media del ojo, o coroides constituye una membrana altamente pigmentada. Se trata de una red de venas que representa la principal fuente nutricional del ojo. En su porción anterior la coroides se encuentra dividida en dos estructuras: el cuerpo ciliar y el iris. El cuerpo ciliar está formado por fibras musculares radiales y circulares que, junto a otros ligamentos, mantienen al cristalino en su sitio. El iris, se abre o se cierra con el propósito de controlar la cantidad de luz que entra al globo ocular y contiene el pigmento visible del ojo.

La esclerótica está formada por tejido fibroso blanco y duro, excepto su porción anterior denominada córnea, que es de color transparente y se localiza sobre el iris.

La córnea cubre la pupila y el iris con una capa transparente, que refracta la luz antes de que entre en el ojo. La abertura en el centro del iris es la pupila, a través de ella pasa la luz hacia el cristalino y la retina y no es más que un agujero rodeado por el cuerpo muscular del iris.

El diámetro pupilar aumenta a poca luz, y disminuye ante luz brillante e intensa. La pupila de un ojo humano puede comprimirse hasta los 1,5 mm de diámetro, o dilatarse hasta los 8 ó 9 mm y puede reaccionar ante estímulos en 0,2 segundos. El grado de luz que penetra en el globo ocular es proporcional al área de la pupila, por lo cual, considerando las dimensiones previamente

mencionadas, la cantidad de luz que entra al ojo puede cambiar unas 30 veces como resultado de las variaciones pupilares [17].

La constricción y dilatación pupilar dependen principalmente del control que ejerce el sistema nervioso autónomo sobre los músculos del iris [18]. Dicho funcionamiento depende del balance entre el músculo esfínter de la pupila y las fibras del iris, que tienen únicamente inervación autónoma [19].

El tamaño pupilar varía de individuo a individuo e incluso entre los ojos de un mismo sujeto. Su diámetro normalmente se encuentra entre los 3mm y 4mm y tiende a decrecer con la edad. Las respuestas pupilares más comunes, que constituyen variaciones en su diámetro, están relacionadas con cambios en los niveles de luz con los que se estimula el sistema ocular, variaciones de color y movimientos, aunque hay que tener en cuenta que no son las únicas condiciones asociadas a la reactividad pupilar [19].

1.2 Seguimiento ocular mediante VOG.

El seguimiento ocular (en inglés se conoce como *eye tracking*) es el proceso de medir la dirección de la mirada. Existen diversos métodos para medir el movimiento del ojo. En este trabajo se aborda la variante más moderna que es la VOG la cual consiste en utilizar cámaras de vídeo para calcular la dirección de la mirada. Se colocan una o varias cámaras de reducidas dimensiones, enfocando continuamente el ojo o los ojos sobre los que se desea hacer el seguimiento. La señal de video de la cámara es digitalizada y posteriormente, procesada empleando algoritmos de procesamiento digital de imágenes.

1.2.1 Configuraciones para seguimiento ocular

En función de la posición de la cámara se distinguen dos configuraciones:

- Sistema montado en la cabeza: En estos sistemas la cámara se coloca sobre un casco o unas gafas a una corta distancia del ojo, como se muestra en la Figura 1.4.



Figura 1.3 Sistema montado en la cabeza

El sistema es muy preciso (precisiones inferiores a 1° angular) dentro de un rango de medida medio ($\pm 30^\circ$ en horizontal y ± 25 en vertical) y es además intrusivo, ya que requiere que el usuario se coloque un casco o gafa. Es un sistema referenciado a la cabeza, y por lo tanto, los movimientos que detecta son de los ojos con respecto a la cabeza, con lo que no calcula la dirección absoluta de la mirada.

Entre los sistemas de esta configuración más difundidos se encuentra el de *SensoMotoric Instruments* (*SMI Eye Tracking Glasses 2*) (ver Figura 1.4). Esta tecnología captura imágenes con una resolución de 1280x960 píxeles a 24 fps (cuadros por segundo o *frames per second*, en inglés) y con una resolución de 1024x720 píxeles a 30 fps. El precio del paquete mínimo es de 11900 USD.



Figura 1.4: Sistema montado en la cabeza de SensoMotoric Instruments.

- Sistemas remotos: Se coloca la cámara a una cierta distancia del usuario, normalmente abajo o alrededor de la pantalla de la computadora, logrando así una configuración no intrusiva. El usuario tiene la posibilidad de moverse frente a la pantalla, siempre y cuando se mantenga en el campo de visión de la cámara. Para mayor precisión, en algunos casos, se emplea una mentonera (*chinrest*, en inglés) para asegurar que la posición del usuario durante un experimento sea la misma durante la calibración. En la Figura 1.5 se muestra un sistema de seguimiento visual remoto.

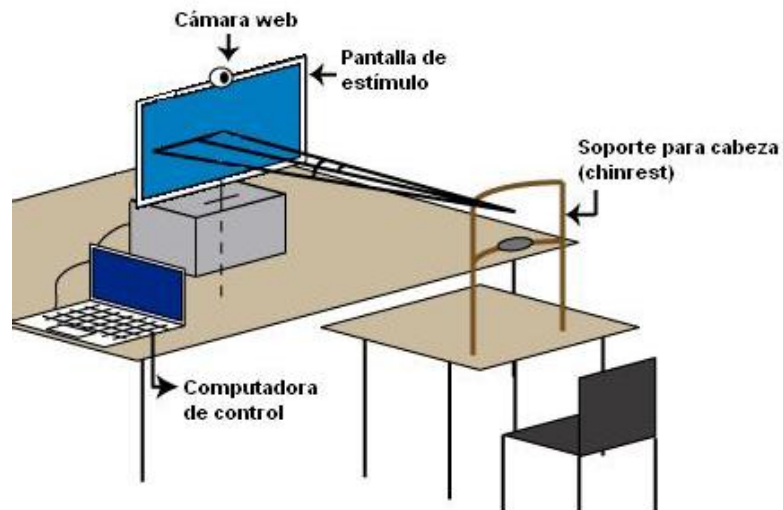


Figura 1.5: Sistema de seguimiento visual remoto.

En este tipo de sistema pueden aparecer problemas debidos principalmente a movimientos de cabeza o del ojo, de modo que éste se sitúe fuera del campo de visión de la cámara. Por tanto, los sistemas montados en la cabeza son los que mejores prestaciones presentan, aun cuando necesitan de un soporte para su fijación.

Entre los sistemas de esta configuración más difundidos se encuentra el de la marca Tobii, con sus diferentes generaciones y modelos comerciales: TX300, X1 y X2, entre otros [20]. El TX300 (ver figura 1.6) captura 300 imágenes por segundo de ambos ojos y es probablemente el mejor seguidor ocular (*eye tracker*) del mundo, pero el precio solamente del hardware es de 45900 USD a los cuales habría que agregar el software, los envíos y la capacitación. Es el sistema más caro de la empresa Tobii.

Tobii dispone también de varios otros sistemas empezando con el X2-30 cuyo precio es 8900 USD. Entre las versiones de software de este sistema están el “*Tobii Studio Professional*” y el “*Tobii Studio Enterprise*” cuyos precios son 8900 USD y 12900 USD, respectivamente.

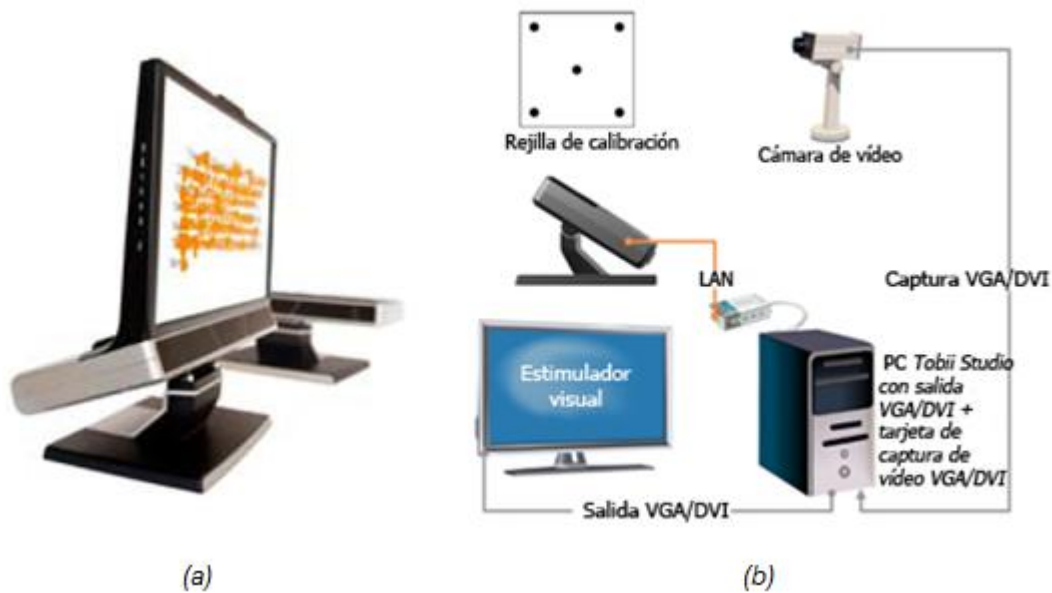


Figura 1.6: Sistema TOBII TX 300: (a) Iluminador y pantalla de seguimiento (b) Sistema completo

1.2.2 Pasos para el seguimiento ocular

Los pasos para el seguimiento ocular son los siguientes:

- Adquisición de imagen/video
- Localización del rostro
- Localización del ojo
- Detección de la pupila
- Cálculo de la dirección de la mirada

Adquisición de imagen/video: Se realiza con una cámara de video convencional, *webcam* o cámara infrarroja (IR) que se suele conectar a la computadora a través del puerto USB. En ocasiones, se requieren varias cámaras para compensar los movimientos de la cabeza.

Localización del rostro: Permite localizar la región de la cara en una imagen usando métodos basados en: color de la piel, inteligencia computacional, rasgos faciales, textura, correlación con un patrón y apariencia, entre otros. La selección de un método depende de las condiciones ambientales, la exactitud deseada, la velocidad de procesamiento y la complejidad, entre otras.

Localización del ojo: Existen dos métodos principales para la detección del ojo: la holística, que localiza los ojos usando representaciones globales, y la abstracta, que extrae y mide rasgos locales discretos empleando técnicas de reconocimiento de patrones. Se pueden citar las siguientes técnicas: modelo de proyección de gris, correlación con un patrón, filtrado de Kalman, propiedades geométricas de los ojos, función de proyección, detección de esquina y combinaciones de los anteriores.

Detección de la pupila: El objetivo es estimar la posición del centro de la pupila.

1.2.3 Aplicaciones del seguimiento ocular

Desde el surgimiento del primer seguidor ocular o *eye tracker* en 1947 en la fuerza aérea de los Estados Unidos para encontrar las mejores posiciones de los controles en la cabina del piloto, hasta nuestros días, se ha reportado un gran número de aplicaciones. Algunas de estas son: interacción humano-computador, videojuegos, rehabilitación, evaluación de choferes, detección de fatiga, ciencia cognitiva, estudios de mercado, estudios de usabilidad, investigaciones médicas, tecleo mediante la mirada, interacción con la mirada y sistemas asistentes en automóviles, entre otras.

Tradicionalmente, las personas discapacitadas que no pueden mover ninguna parte de su cuerpo excepto los ojos usan la interacción por la mirada. Estos sistemas están diseñados para dirigir la computadora de esta forma.

1.3 Detección de pupila

El procedimiento de localización de la pupila es una tarea compleja, puesto que la forma del objeto a segmentar no es necesariamente regular o sus límites no siempre están bien definidos. Generalmente se obtiene primero una zona de interés para facilitar la localización y para que los algoritmos trabajen más rápido. Existen varios métodos para detectar el centro de la pupila, a continuación se explican los más destacados de la literatura.

1.3.1 Operador Integro-diferencial

Daugman [9] hace uso de la técnica de los operadores integro-diferenciales para la localización de los bordes circulares de la pupila y del iris y para identificar las regiones de los arcos de los párpados superiores e inferiores. Matemáticamente está descrito por la expresión 1.1:

$$\max_{(r, x_0, y_0)} \left| G_\sigma(r) * \frac{\partial}{\partial r} \oint_{r, x_0, y_0} \frac{I(x, y)}{2\pi r} ds \right| \quad (1.1)$$

Donde el símbolo $*$ denota convolución, $I(x, y)$ es la intensidad del píxel en las coordenadas (x, y) de la imagen ocular, r es el radio de varias regiones circulares sobre las que se calcula el gradiente. G_σ es una función de suavizado gaussiana con un valor de escala espacial σ y está descrita matemáticamente por la expresión 1.2 :

$$G_\sigma(r) = \left(\frac{1}{2\pi\sigma} \right) e^{-\frac{(r-r_0)}{2\sigma^2}} \quad (1.2)$$

El operador completo se comporta como un detector de bordes circulares que busca iterativamente el borde que maximice el valor del operador integro-diferencial dentro de un círculo con radio r y centro (x_0, y_0) el cuál es la superficie de integración. Este operador se aplica iterativamente con la cantidad de suavizado progresivamente reducida con el fin de lograr la localización exacta. Los párpados están localizados de una manera similar.

La metodología sugerida por Daugman en [9], se ejecuta en dos etapas:

1. La primera etapa tiene como objetivo buscar el borde esclerótico del iris en una zona delimitada por dos conos de 90° cada uno, en oposición a las partes superiores, que cubren prácticamente las zonas laterales derecha e izquierda visibles del iris.
2. La segunda etapa consiste en buscar el contorno de la pupila. Para eso, Daugman usa el mismo operador en una escala más fina y, por lo tanto, un valor más bajo de σ . Los valores de (r, x_0, y_0) varían en los valores previamente localizados para la zona del iris. [21]

Sin embargo, este método falla con frecuencia cuando las imágenes no tienen alto contraste entre las regiones del iris y la esclerótica. Como posibles optimizaciones en [22], se implementan dos operaciones de pre-procesamiento que contribuyen a la mejora de los resultados. Estas son:

- Ecualización del histograma: Se mejora el contraste entre cada región del ojo, contribuyendo así a la correcta aplicación del algoritmo de segmentación.
- Binarización: La aplicación de un umbral en una imagen antes de la ejecución del operador permite optimizar el contraste entre las regiones. Este proceso tiene, sin embargo, una desventaja importante: es muy dependiente en el umbral elegido, y como pueden cambiar

las características de la imagen, los resultados pueden deteriorarse gravemente. Aparte de este hecho, la binarización compromete una de las mayores ventajas de esta metodología: la inexistencia de parámetros definidos por el usuario para la tarea de segmentación.

1.3.2 Transformada de Hough

La transformada de Hough es una herramienta estándar en el análisis de imágenes, que permite el reconocimiento de patrones globales en el espacio de la imagen, reconociendo patrones locales (idealmente un punto) en un espacio de parámetros transformado. Es particularmente útil cuando la información es ruidosa. La idea básica de esta técnica es encontrar curvas tales como rectas, polinomios, círculos y otras, que puedan ser localizadas en un espacio de parámetros adecuado. Aunque la transformada se puede utilizar para dimensiones mayores, se le usa principalmente en dos dimensiones, para encontrar líneas rectas, centros de círculos con radio fijo, parábolas, etc. En específico, la transformada de Hough circular puede emplearse para estimar el radio y las coordenadas del centro de las regiones de la pupila y el iris. El primero que la utilizó para la segmentación de iris fue R. Wildes [23].

Para el proceso de segmentación usando la transformada de Hough primero se obtiene el mapa de bordes de la imagen que es el conjunto de puntos de esta que superan un determinado valor de gradiente. En este punto, sale a relucir una de las mayores dificultades de este método puesto que se requiere la estimación de este umbral de antemano. Luego se cuentan en trayectorias circulares los puntos de la imagen que pertenecen al mapa de bordes. A este proceso se le denomina votación. Las circunferencias más votadas son las que se escogen como bordes del iris. Para realizar estas votaciones se crea el espacio paramétrico de Hough que es una matriz tridimensional cuyas dimensiones son precisamente los parámetros del círculo: las coordenadas del centro (x_0, y_0) y la magnitud de su radio r . Las entradas (x_0, y_0, r) de este espacio paramétrico H se obtienen de la siguiente forma:

$$H(x_0, y_0, r) = \sum_{j=1}^n h(x_j, y_j, x_0, y_0, r) \quad (1.3)$$

dónde:

$$h(x_j, y_j, x_0, y_0, r) = \begin{cases} 1, & \text{si } g(x_j, y_j, x_0, y_0, r) = 0 \\ 0, & \text{de otra forma} \end{cases} \quad (1.4)$$

$$g(x_j, y_j, x_0, y_0, r) = (x_j - x_0)^2 + (y_j - y_0)^2 - r^2 \quad (1.5)$$

Esto significa que para cada punto (x_j, y_j) del mapa de bordes, $g(x_j, y_j, x_0, y_0, r)$ es puesto a cero si la terna paramétrica (x_0, y_0, r) representa un círculo que pase por dicho punto. Por último se elige la terna que maximiza H como el círculo más probable dentro de la imagen. La figura 1.7 ilustra la secuencia de pasos que usa la transformada de Hough para la detección de círculos.



Figura 1.7 Secuencia de operación de la Transformada de Hough para la detección de círculos

El método de segmentación basado en la transformada de Hough es muy útil cuando se trabaja con imágenes con cierto nivel de ruido, pero la carga computacional a la hora de implementarlo es bastante grande. Para resolver esta cuestión se le han hecho muchas mejoras en la implementación del algoritmo donde casi todas se basan en su combinación con otras técnicas de segmentación.

- Liu [24] emplea el detector de bordes de Canny seguido de la transformada de Hough, tratando de simplificar los métodos modelando pupila e iris como dos círculos concéntricos.
- Singh [6] introduce una mejora basada en la ecualización del histograma seguido de un filtro de altas frecuencias.
- Shi [25] binariza la imagen con el objetivo de localizar la pupila y entonces aplica la detección de contornos y posteriormente la transformada de Hough.
- Tian [26] busca los píxeles que tienen un nivel de gris por debajo de un determinado umbral, por medio de estos aproxima el centro de la pupila. Luego aplica la mencionada técnica, o sea, el detector de Canny seguido por la transformada de Hough.
- Sun [27] asume que en una imagen de iris, los valores de gris correspondientes a la región de la pupila son los de menor intensidad, utiliza esto para estimar la pupila y así aplicar Canny y Hough.

El hecho de que el espacio de parámetros sea de tres dimensiones, ya que cada píxel en el espacio de la imagen corresponde a un círculo en el espacio de Hough, hace que la implementación de este método pueda requerir un alto coste computacional. Además, para formas geométricas más complejas (por ejemplo, elipses, que representan una mejor

aproximación a la forma de la pupila), la dimensión del espacio de Hough aumenta con el número de las variables, y el tiempo de cálculo del algoritmo aumenta considerablemente. Por esta razón, dicho método no se implementará en este trabajo.

1.3.3 Modelo de contornos activos

Las técnicas de contornos activos constituyen una herramienta eficaz en múltiples tareas, como por ejemplo, en el análisis de imágenes médicas, donde la baja relación señal/ruido hace insuficientes los resultados obtenidos mediante técnicas clásicas. Son muy robustas frente a la presencia de ruido y otros elementos espurios y permiten segmentar imágenes mucho más complejas.

Un contorno activo (comúnmente llamado *snake*) consiste en una curva elástica aproximada que, colocada sobre una imagen, empieza a deformarse a partir de una forma inicial con el fin de delimitar las regiones de interés. Los contornos activos responden a fuerzas internas y externas que se deforman o mueven a través de una imagen hasta que se establece el equilibrio entre ellas. El contorno contiene una serie de vértices cuyas posiciones se cambian por dos fuerzas opuestas: una interna (que depende de las características deseadas) y una externa (dependiente de la imagen). Cada vértice se mueve entre el tiempo t y $t + 1$ mediante la ecuación:

$$v_i(t + 1) = v_i(t) + F_i(t) + G_i(t) \quad (1.6)$$

donde $F_i(t)$ es la fuerza interna, $G_i(t)$ es la fuerza externa y $v_i(t)$ es la posición del vértice i .

Para ubicar la región de la pupila, las fuerzas internas se calibran para formar un contorno interno que corresponda a un círculo globalmente expandido. Las fuerzas externas se utilizan básicamente para obtener la información de borde de la imagen. Finalmente, se ubica un punto interior a la pupila a partir de una imagen de varianza y luego se crea un contorno activo circular discreto con este punto como su centro. Estas fuerzas localizan a la pupila bajo la condición de equilibrio entre las fuerzas internas y externas [28].

La técnica de segmentación usando los contornos circulares activos ha obtenido buenos resultados en la segmentación de la pupila, pero tiene el inconveniente de una implementación bastante compleja en relación a las otras técnicas mencionadas, por lo cual no fue objetivo de este trabajo.

1.3.4 Ajuste de elipse

La localización del centro de la pupila se determina generalmente mediante la detección de sus bordes y asumiendo que presenta una forma geométrica particular. La aproximación más simple consiste en suponer que es rotacionalmente simétrica [29], sin embargo, la forma elíptica parece ser una mejor opción [30].

El método de ajuste de elipse se basa en detectar el contorno de la pupila y calcular la elipse que mejor se ajusta a él. Primero, se selecciona un valor de umbral para segmentar la región de interés (la pupila) y luego se realiza la operación morfológica de cierre para eliminar el ruido en la imagen que pudiera distorsionar el contorno de la pupila. A continuación se utiliza un detector de bordes para localizar el contorno de la pupila. En muchas ocasiones, las reflexiones debidas al sistema de iluminación pueden aparecer dentro de la región de interés (ver figura 1.8a). Estas no perturbarán el proceso de detección, sin embargo, aquellas cercanas al borde sí pueden interferir en el contorno obtenido. Para eliminar esta interferencia, el tamaño del destello se delimita seleccionando todos los píxeles saturados conectados en la imagen. Si el área así obtenida se cruza con el contorno de la pupila, estos datos se eliminan del conjunto que se ajustará a una elipse. Los datos excluidos son un pequeño porcentaje de los puntos que delimitan la pupila y no afectan la bondad del ajuste. Las coordenadas (x, y) de los píxeles que forman el borde de la pupila se ajustan a la ecuación general de la elipse:

$$Ax^2 + Bxy + Cy^2 + Dx + Ey + F = 0 \quad (1.7)$$

En la figura 1.8 se muestra un ejemplo del empleo de este algoritmo para segmentar a la pupila.

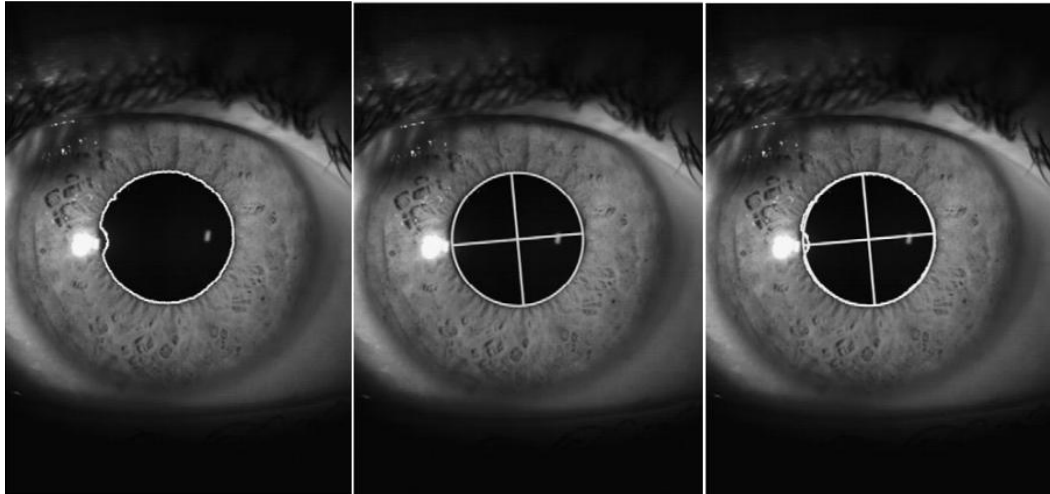


Figura 1.8 Secuencia de operación del ajuste de elipse para la detección la pupila: (a) imagen que representa el contorno de la pupila en presencia de brillo, (b) ajuste de los datos de contorno, (c) imagen que representa el contorno y el ajuste de la pupila.

1.4 Análisis de los movimientos oculares

El objetivo final de la detección de pupila en el contexto de este trabajo es obtener una señal de seguimiento ocular, a la cual se le extraerán parámetros que servirán de apoyo al diagnóstico. En la evaluación de algunas enfermedades del sistema nervioso, por ejemplo, la ataxia, se estimula visualmente al paciente con patrones específicos. El más usado es el estímulo sacádico que consiste en un punto que se mueve de forma impulsiva entre dos posiciones diferentes en una misma dirección horizontal (ver Figura 1.9 a). Los parámetros que se analizan en este movimiento son los siguientes: latencia, velocidad y desviación.

Latencia: Es el tiempo que tarda el ojo en iniciar el movimiento sacádico ante un estímulo (ver Figura 1.9 a).

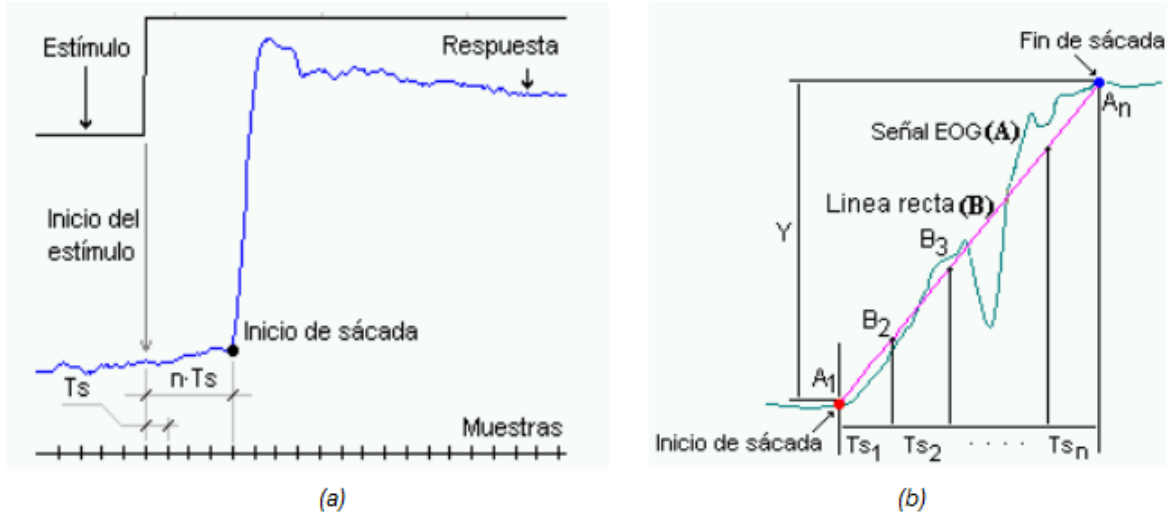


Figura 1.9: Parámetros de diagnóstico en el dominio tiempo para la prueba de seguimiento ocular sacádico. (a) Latencia (demora) entre la señal de estímulo (superior) y la respuesta ocular (señal de seguimiento de la posición de la pupila) (b) Velocidad sacádica lineal.

Velocidad sacádica: Es la rapidez con que cambia la posición de la mirada debido a un movimiento sacádico (ver Figura 1.9 b). Así, la velocidad sacádica lineal se define como:

$$V_{lin} = \frac{A_n - A_1}{(n-1)T_s} \cdot \frac{30^\circ}{V_{cal}} \text{ [}^\circ/\text{s]} \quad (1.8)$$

Dónde: $\overline{V_{cal}}$ es el voltaje de calibración obtenido a un ángulo de 30° , T_s es el período de muestreo (número de imágenes por segundo), A_1 y A_n son las amplitudes (tensiones) de las muestras de inicio y fin, y n es el número de muestras entre el inicio y el fin.

Desviación: Representa, en grados, cuanto se desvió la dirección de la mirada con respecto a la posición de estímulo (Figura 1.10). Si la posición de la mirada sobrepasó la posición del estímulo, ocurre una hipermetría. En el caso contrario, ocurre una hipometría. Este parámetro es muy marcado en sujetos patológicos.

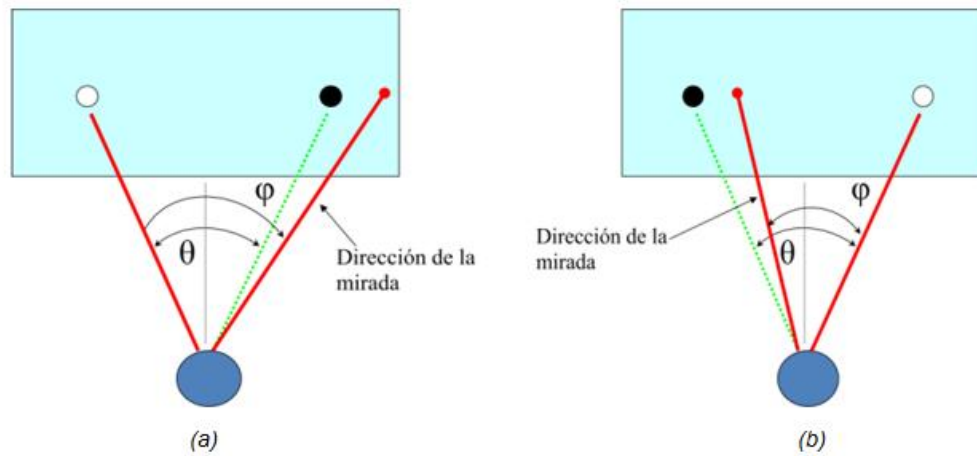


Figura 1.10: Ángulo de desviación (δ) entre los ángulos de dirección de la mirada (ϕ) y de posición del estímulo (θ) para un movimiento sacádico. (a) caso de hipermetría si $\delta = \phi - \theta > 0$, (b) caso de hipometría si $\delta = \phi - \theta < 0$.

2 . DETECCIÓN DEL CENTRO DE LA PUPILA

En este capítulo se describen las imágenes y los algoritmos que soportan la investigación. Inicialmente, se fundamenta la selección de las bases de datos de imágenes utilizadas y se realiza su descripción. Luego, se explica en detalles, el método de detección implementado basado en la transformada de simetría radial (TSR). Finalmente, se describen los algoritmos de Ajuste de Elipse y Ajuste de Círculo, que permiten realizar, de forma interactiva, la anotación semiautomática del centro de la pupila, lo cual es un paso necesario para evaluar la exactitud del algoritmo.

2.1 Bases de datos

En la actualidad, existen diversas bases de datos de imágenes oculares que han sido liberadas para su uso en la investigación universitaria y de institutos especializados, como por ejemplo: CASIA [31], UBIRIS, MMU (Universidad Multimedia, 2004) e ICE (Instituto Nacional de Estándares y Tecnología, 2006), entre otras. Algunas de ellas como la CASIA, MMU e ICE, almacenan imágenes adquiridas con iluminación infrarroja cercana y en otras, como la UBIRIS, las imágenes han sido adquiridas en condiciones de iluminación correspondientes al espectro de luz visible. Los sistemas basados en video utilizan frecuentemente cámaras e iluminación IR, los cuales proporcionan una imagen nítida y de alto contraste entre la pupila y cualquier color de iris que tenga el sujeto. Por el contrario, cuando se ilumina al ojo con luz visible, se hace difícil la detección de un contorno claro del borde de la pupila, ya que se crean multitud de reflejos y el contorno se hace difuso (figura 2.1). Es por esta razón que en este trabajo se ha utilizado la base de datos CASIA para la experimentación.

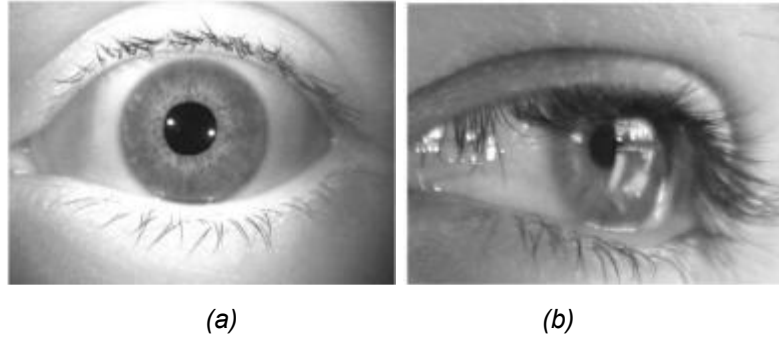


Figura 2.1: Diferencias en la nitidez de la pupila con diferentes fuentes de iluminación: (a) luz infrarroja y (b) luz natural

2.1.1 CASIA

La base de datos de imágenes de CASIA fue desarrollada por el grupo de investigación del Instituto de Automatización de la Academia de Ciencias de China en el 2004. A partir de su fecha de creación, se han producido cuatro versiones: desde la *CASIA - IrisV1* hasta *CASIA-IrisV4*. Estas, a su vez, están divididas en varios subconjuntos, cada uno de ellos con características particulares. A continuación se explicarán solamente aquellos utilizados en este trabajo que son los que más se ajustan al contexto de la aplicación de esta tesis.

CASIA-Iris V3: Esta es la primera base de datos de CASIA que introduce elementos ruidosos. Contiene un total de 22034 imágenes de iris adquiridas en 700 sujetos. La base de datos se divide en 3 subconjuntos bastantes dispares entre sí denominados: *CASIA-Iris-Interval*, *CASIA-Iris-Lamp* y *CASIA-Iris-Twins*. El primer subconjunto usa demasiados iluminadores IR (ocho) lo cual pudiera resultar incómodo para el sujeto y no se corresponde con el sistema VOG actual y en desarrollo. Solamente, de los dos últimos subconjuntos, fueron seleccionadas imágenes de manera aleatoria para probar los algoritmos implementados. El subconjunto *CASIA-Iris-Lamp* fue capturado usando la cámara de mano OKI que se muestra en la figura 2.2 y contiene imágenes influenciadas por las variaciones de iluminación provocadas por una lámpara, que causaron la dilatación de la pupila en las imágenes capturadas (ver figura 2.3). Como se conoce, el iris humano se expande hacia su centro o se contrae contra su límite con la esclerótica según el ojo perciba exceso o falta de iluminación, respectivamente. Este subconjunto contiene 16212 imágenes de iris almacenadas en formato JPG con una resolución de 640×480 píxeles.



Figura 2.2: Sensor OKI Irispass-h



Figura 2.3: Imágenes de muestra de CASIA Iris Lamp con pupila dilatada y contraída

Por otra parte, el subconjunto *CASIA-Iris Twins* contiene 3183 imágenes de 100 pares de gemelos capturadas durante el Festival Anual de Gemelos en Beijing, China. Todas las imágenes se almacenan en formato JPG con una resolución de 640×480 píxeles. En la figura 2.4 se observa una imagen de muestra de este subconjunto. Este subconjunto se caracteriza por presentar un número considerable de imágenes ruidosas que serán usadas para analizar la robustez de los algoritmos implementados.

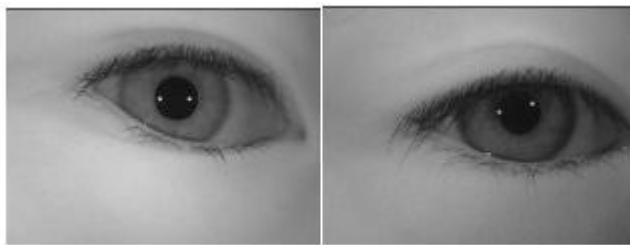


Figura 2.4: Imágenes de muestra de CASIA Iris Twins

CASIA-Iris V4: Es una extensión de *CASIA-Iris V3* con tres subconjuntos más: *CASIA-Iris Distance*, *CASIA-Iris Syn* y *CASIA-Iris Thousand*. La primera corresponde a imágenes adquiridas

a distancias a partir de tres metros para fines de biometría que no se corresponden con la aplicación de interés en este trabajo. La segunda está compuesta por imágenes sintéticas que no corresponden a sujetos reales. Por estas razones ambas fueron excluidas. Por tanto, sólo, el último subconjunto se tomó en consideración para la evaluación del algoritmo implementado.

El subconjunto *CASIA-Iris Thousand*, contiene 20000 imágenes del iris de 1000 sujetos, que se capturaron usando la cámara IKEMB-100 producida por *IrisKing* (ver figura 2.5). La IKEMB-100 es una cámara doble con una amistosa retroalimentación visual. La imagen mostrada en el panel frontal LCD ayuda al usuario a ajustar su postura durante la adquisición de la imagen del iris de alta calidad. Las principales fuentes de variación que presenta este subconjunto son sujetos con espejuelos y reflexiones especulares. En la figura 2.6 se muestra un ejemplo de este subconjunto.

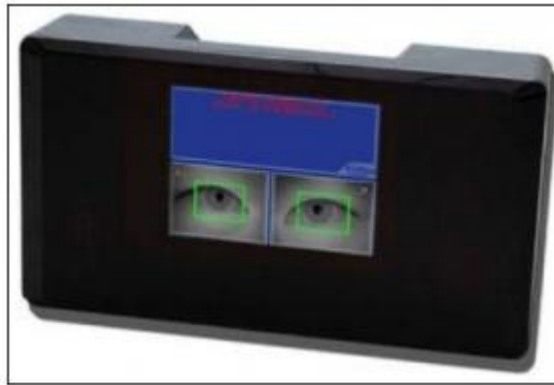


Figura 2.5: Cámara IKEMB-100



Figura 2.6: Subconjunto CASIA-Iris-Thousand adquiridos con la cámara IKEMB-100 producida por *IrisKing*

2.2 Algoritmo de detección implementado

El algoritmo implementado para la detección de la pupila se basa en la Transformada de Simetría Radial [32], la cual se fundamenta en considerar todos los posibles círculos de los que puede

formar parte un píxel del borde como se muestra en la Figura 2.9. Para un píxel del borde, los centroides de una posible familia de círculos se extienden a lo largo de la línea que es ortogonal a la dirección del borde (Figura 2.9a). Cada punto del borde de un círculo vota a lo largo de una línea de posibles radios y estas líneas se interceptan en el centro del círculo, resultando en un pico.

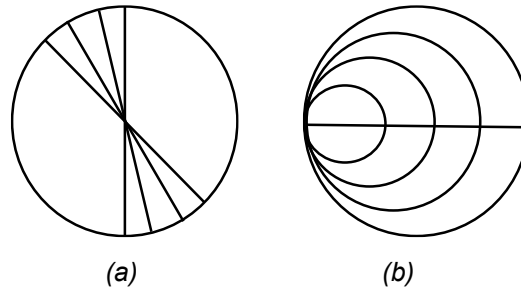


Figura 2.9. Simetría radial: (a) Para cada posible círculo, cada punto del borde vota a lo largo de una línea de posibles radios y estas líneas se interceptan en el centro del círculo, resultando un pico, (b) Búsqueda de todos los posibles círculos de los que puede formar parte un píxel del borde.

El algoritmo mencionado va a disponer a su entrada de una imagen infrarroja del ojo, debido a que esta tiene un mejor contraste entre la región del iris y la pupila, en comparación con una imagen de luz visible. Inicialmente se disminuye la resolución de la imagen original por un factor, para reducir el costo computacional del algoritmo pero manteniendo una calidad de imagen adecuada. Luego, se aplica la TSR, que devuelve un mapa de probabilidades (basado en votaciones) a partir del cual se determina la posición (x, y) del centro de la pupila. Finalmente, las coordenadas se escalan al tamaño de la imagen original multiplicando el valor de estas por el factor por el cual se disminuyó la resolución en el primer paso. A continuación, se explicará en detalles los fundamentos del algoritmo de la TSR.

En la figura 2.10 se muestra los diferentes pasos para obtener la TSR, la cual se calcula en uno o más radios $n \in N$, donde N es el conjunto de radios de las características radialmente simétricas que deben detectarse. El valor de la transformada en el radio n indica la contribución a la simetría radial de los gradientes a una distancia n de cada punto.

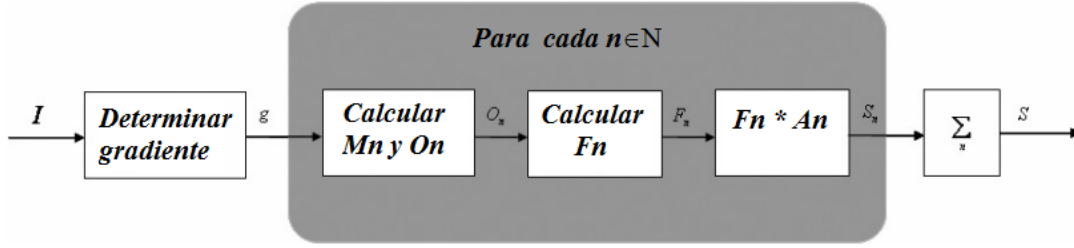


Figura 2.10 Pasos para obtener la Transformada de Simetría Radial.

Primero, se determina el gradiente g de la imagen, que actúa como un detector de bordes. Si su valor es positivo, se analiza la contribución de simetría radial de cada píxel con sus píxeles circundantes. Si no, se pasa a analizar el siguiente píxel de la imagen. Dentro de los vecinos, el valor del gradiente de pares de puntos simétricamente situados sobre el píxel central se usa como evidencia de simetría radial. Para este paso se utilizó la función *gradient* de Matlab que devuelve las componentes x y y del gradiente numérico bidimensional de la matriz que contiene la imagen sub-muestreada (ver figura 2.11).

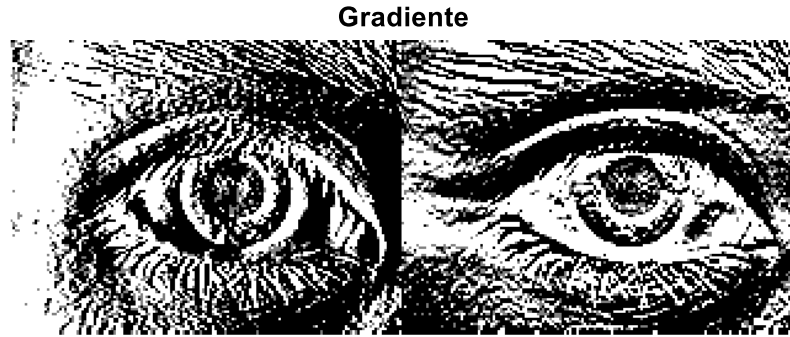


Figura 2.11 Obtención del gradiente de una imagen ocular infrarroja: (a) componente en y (g_y), (b) componente en x (g_x).

Para cada radio n , se calcula una imagen de proyección de orientación O_n y una imagen de proyección de magnitud M_n . Estas imágenes se generan al examinar el gradiente g en cada píxel p a partir del cual se determinan un píxel afectado positivamente p_{+ve} y un píxel afectado negativamente p_{-ve} , como se muestra en la figura 2.12.

El píxel afectado positivamente se define como el píxel donde el vector gradiente está apuntando, a una distancia n de p , y el píxel afectado negativamente se define como el píxel ubicado a la misma distancia n en la dirección opuesta a donde apunta el gradiente.

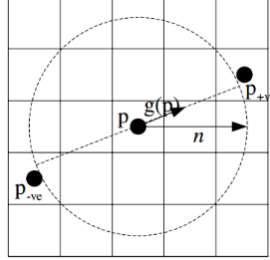


Figura 2.12 Definición de píxeles afectados positivamente p_{+ve} y negativamente p_{-ve} por el elemento de gradiente $g(p)$ para un radio $n = 2$. El círculo con líneas discontinuas muestra todos los píxeles que pueden verse afectados por el gradiente para un radio n .

Las coordenadas de los píxeles afectados positivamente y negativamente están dadas por:

$$p_{+ve}(p) = p + \text{round}\left(\frac{g(p)}{\|g(p)\|} n\right) \quad (2.1)$$

$$p_{-ve}(p) = p - \text{round}\left(\frac{g(p)}{\|g(p)\|} n\right) \quad (2.2)$$

La función "round" redondea cada elemento del vector al entero más cercano. Las imágenes de proyección de orientación y de magnitud se inicializan en cero. Para cada par de píxeles afectados, el píxel correspondiente a p_{+ve} en la imagen de proyección de orientación O_n y en la imagen de proyección de magnitud M_n se incrementa en 1 y $\|g(p)\|$, respectivamente, mientras que el píxel correspondiente a p_{-ve} se reduce en estas mismas cantidades en cada imagen. Lo anterior, se expresa matemáticamente como:

$$O_n(p_{+ve}(p)) = O_n(p_{+ve}(p)) + 1 \quad (2.3)$$

$$O_n(p_{-ve}(p)) = O_n(p_{-ve}(p)) - 1 \quad (2.4)$$

$$M_n(p_{+ve}(p)) = M_n(p_{+ve}(p)) + \|g(p)\| \quad (2.5)$$

$$M_n(p_{-ve}(p)) = M_n(p_{-ve}(p)) - \|g(p)\| \quad (2.6)$$

La transformación se puede ajustar para buscar sólo regiones oscuras o claras de simetría. Para buscar simetría exclusivamente en las regiones oscuras (por las características de la pupila), al determinar M_n y O_n , se deben considerar solamente los píxeles afectados negativamente. Así,

para el caso de nuestra aplicación se ajusta el algoritmo de la TSR para que implemente sólo las ecuaciones 2.4 y 2.6.

En la figura 2.13 se muestran las imágenes M_n y O_n para tres radios diferentes. El radio de la pupila de la imagen que se está procesando es de aproximadamente 48 píxeles en la imagen original lo que corresponde a 12 píxeles en la imagen sub-muestreada. En la figura 2.5 se observa cómo, a medida que el radio de exploración se aproxima al radio real, existen mayor cantidad de votaciones sobre un mismo punto y cómo estas se dispersan a medida que el radio que se analiza se aleja del real.

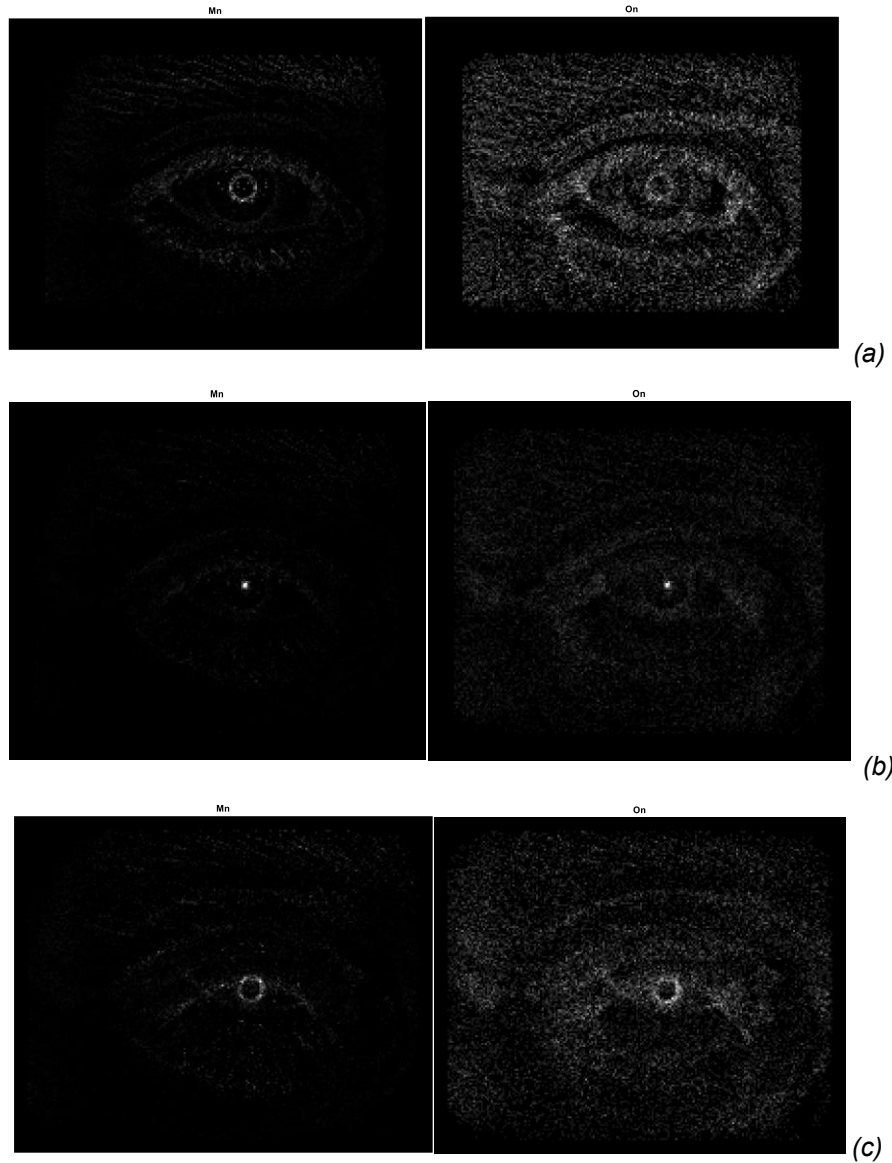


Figura 2.13 Imágenes de proyección de magnitud M_n (a la izquierda) y de orientación O_n (a la derecha) para radio: (a) $n = 6$, (b) $n = 12$ y (c) $n = 18$

La contribución de simetría radial para el radio n se define por la siguiente convolución:

$$S_n = F_n * A_n \quad (2.7)$$

$$F_n(p) = \frac{M_n(p)}{k_n} \left(\frac{|O_n(p)|}{k_n} \right)^\alpha \quad (2.8)$$

Dónde α es el parámetro de rigidez radial, k_n es un factor de escala que normaliza a M_n y O_n a diferentes radios y A_n es una máscara gaussiana bidimensional.

En la figura 2.14 ocurre un proceso análogo al explicado para la figura 2.13. En este caso se muestran las transformaciones que va sufriendo la función F_n a medida que el radio de búsqueda varía. Se observa que para un radio de búsqueda cercano al real, en la imagen se distingue con mayor facilidad un punto brillante que equivale a un máximo.

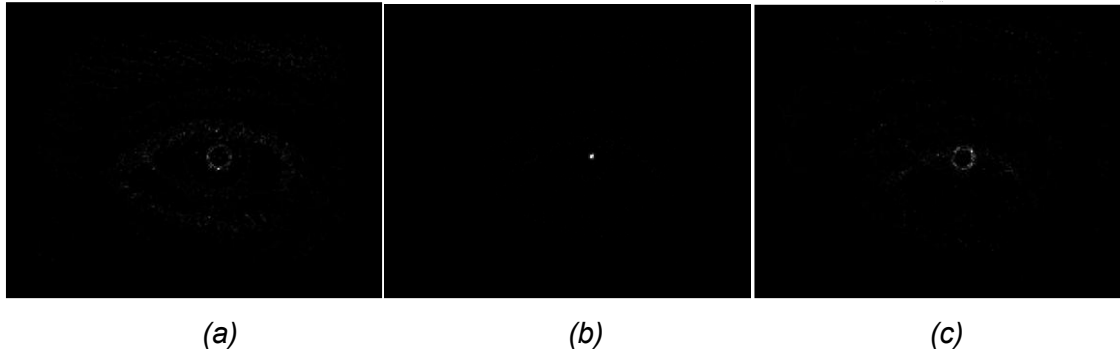


Figura 2.14 Función F_n que integra la información de las imágenes de proyección de magnitud M_n y de orientación O_n para los radios (a) $n = 6$, (b) $n = 12$ y (c) $n = 18$

La transformación completa S se define como el promedio de las contribuciones de simetría en todo el conjunto N de radios considerados, es decir:

$$S = \frac{1}{|N|} \sum_{n \in N} S_n \quad (2.9)$$

En la figura 2.15 se muestra la imagen S resultante de la transformación completa, donde se puede apreciar que el punto de mayor intensidad (más blanco) coincide con el centro de la pupila (figura 2.15b).

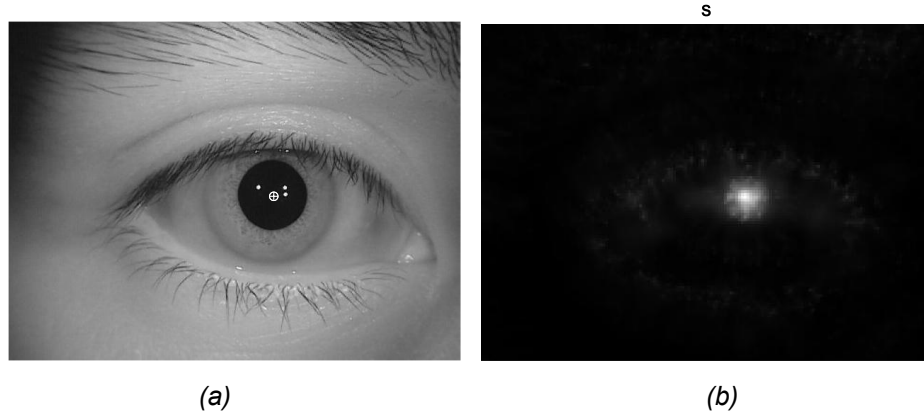


Figura 2.15 El centro de la pupila estimado en (a) coincide con el punto de máxima intensidad de en la imagen (b) obtenida de la transformación S para todos los radios explorados.

El pseudocódigo del algoritmo implementado se muestra a continuación:

Pseudocódigo

1. Sub-muestrear la imagen por un factor de 4
 2. Calcular gradiente de la imagen
 3. Inicializar conjunto de radios de detección ($N_{min} = 5$, $N_{max} = 25$)
 4. Para $n = N_{min}: 1: N_{max}$
 - Calcular coordenadas de $\mathbf{p}_{-ve}$
 - Calcular $O_n(\mathbf{p}_{-ve}(\mathbf{p}))$ y $M_n(\mathbf{p}_{-ve}(\mathbf{p}))$
 - Calcular $F_n(\mathbf{p})$ según ecuación 2.8.
 - Calcular el resultado de la TSR por ecuación 2.7 en el radio de detección n , donde la varianza se escoge $\sigma = 0, 1n$ y el tamaño de la ventana de Gauss es $\text{ceil} \left(\frac{n}{2} \right) \times \text{ceil} \left(\frac{n}{2} \right) *$.
 5. Calcular S por la ecuación 2.9
 6. Buscar en S las coordenadas del valor máximo S_{max} que corresponden a la localización del centro de la pupila.
-

* "ceil" es una función que redondea un número al entero superior

En resumen, para detectar el centro de la pupila de manera efectiva, el algoritmo de la TSR aprovecha al máximo la característica circular simétrica de esta y calcula el píxel afectado negativamente para cada punto en la imagen. Los píxeles afectados de manera negativa apuntan al centro de la pupila debido a que se encuentran ubicados en la dirección negativa del gradiente

que va de un nivel de gris más alto (blanco) a uno más bajo (negro). El resultado final de la TSR se calcula de acuerdo con los dos diagramas de proyección M_n y O_n que, según el radio, se actualizan para guardar las contribuciones de los píxeles afectados negativamente. En el proceso de la TSR, con el cambio de los valores del radio, los puntos del borde de la pupila y del borde del iris se superpondrán cerca del centro de la pupila, donde resultará un valor máximo. La posición del valor máximo de los resultados de la TSR es la ubicación del centro de la pupila. A continuación se detallan algunos criterios para la selección de los parámetros.

2.2.1 Selección de los parámetros

Los parámetros que deben ser considerados para implementar el algoritmo son los siguientes:

- El conjunto de radios $N = \{n_1, n_2, \dots\}$ para los cuales calcular a S_n (contribución para cada radio).
- La máscara gaussiana A_n
- El parámetro de rigidez radial α
- El factor de normalización k_n .

El significado y la selección adecuada de estos parámetros serán abordados en los próximos apartados.

2.2.1.1 Conjunto de radios N

La TSR utiliza un conjunto de radios que deben ser correctamente seleccionados para la optimización y buen rendimiento del algoritmo. En este trabajo fue seleccionado, sobre la base de un estudio previo de las imágenes, el rango de 5 a 25 píxeles los cuales se corresponden a la escala de la imagen sub-muestreada. Este rango contempla todos los posibles radios de la pupila, inclusive cuando esta se haya contraído o dilatado de forma fisiológica. Los radios no necesitan ser continuos, sino que pueden ser alternos, en aras de disminuir el costo computacional, aunque esto podría ocasionar una disminución de la exactitud en la estimación del centro de la pupila. En este trabajo se verificó que, en el rango de radios seleccionado, el incremento en costo computacional para el procesador usado (*Core I5*) no era significativo (fue menor que 0,1s) y se prefirió ganar en exactitud para realizar un mejor diagnóstico.

2.2.1.2 Máscara gaussiana

El propósito de la máscara gaussiana A_n es dispersar la influencia de los píxeles afectados negativamente en función del radio n , lo cual produce un efecto de suavizado entre píxeles vecinos, consiguiendo que las intensidades de los objetos pequeños (por ejemplo, píxeles aislados que no aporten simetría radial) se mezclen con el fondo con el fin de detectar los objetos de mayor tamaño (que aportan simetría radial). Para esto, se utiliza una máscara gaussiana bidimensional, invariante a rotación (su efecto es uniforme en todas las direcciones del gradiente), distribuida a lo largo de las dimensiones x y y , con desviación estándar σ , expresada matemáticamente como:

$$G(x, y) = \left(\frac{1}{2\pi\sigma^2} \right) e^{-\frac{x^2+y^2}{2\sigma^2}} \quad (2.10)$$

Para implementar el método, se seleccionó una desviación estándar que es proporcional al radio, ($\sigma = 0,1n$ [33]), lo cual garantiza que al aumentar el radio n , aumente el área de influencia de la máscara, de forma que en su desplazamiento por la imagen, la convolución de la máscara con la función F_n resalte las nubes de puntos que representen contribuciones elevadas de simetría radial y atenúe los puntos aislados. En la figura 2.16 se muestra este efecto para tres radios diferentes y cómo varían las dimensiones de la máscara gaussiana en función del radio.

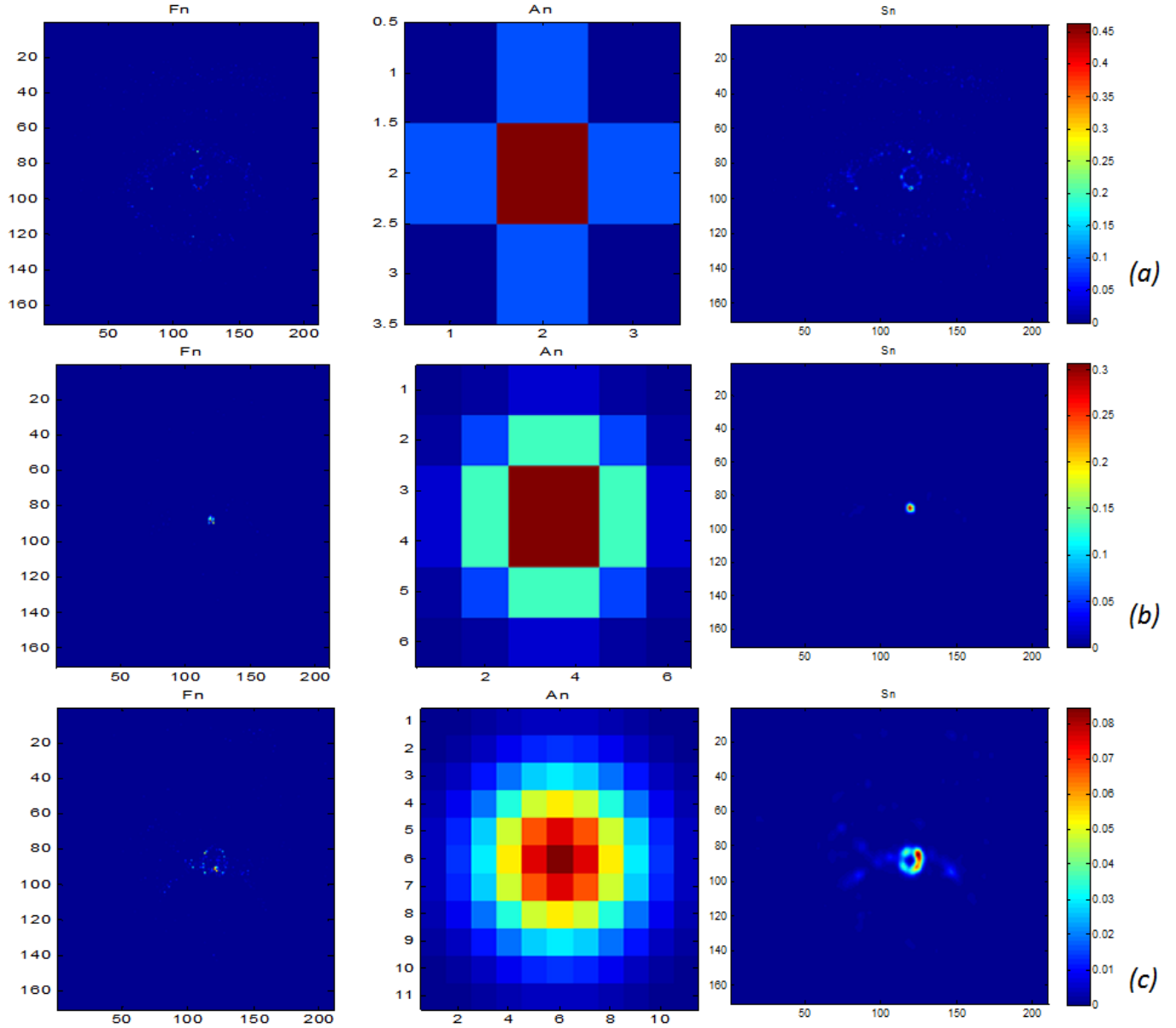


Figura 2.16 De izquierda a derecha: la primera columna representa la función F_n , la segunda a la máscara A_n y la tercera muestra la convolución de las dos anteriores para los radios: (a) $n = 6$, (b) $n = 12$ y (c) $n = 18$

2.2.1.3 Parámetro de rigidez radial

El parámetro de rigidez radial α determina qué tan estrictamente radial debe ser la simetría para que la TSR devuelva un valor de interés alto, es decir, define cuan radial debe ser la simetría del objeto. A mayor α , la TSR elimina rasgos de la imagen que no poseen simetría radial, tales como

cabellos y pestañas, sin embargo el costo computacional aumenta, según se muestra en la expresión (2.8).

En la figura 2.17 se muestra el efecto de tres valores de α para una imagen afectada fuertemente por la presencia de cabellos en la zona del ojo. En este caso, la mejor exactitud se logra para $\alpha = 3$ ya que el algoritmo no tiene en cuenta a los elementos lineales (es decir, sin simetría radial), sin embargo el costo computacional es mayor. En estos casos, para valores de $\alpha = 1$ y $\alpha = 2$, la exactitud disminuye, debido a la alta presencia de elementos lineales.

Dado que en el seguimiento ocular (finalidad del proyecto asociado a este trabajo), al paciente se le orienta, que no debe ocurrir la situación anterior (o sea, no haya presencia de cabello en el ojo), se decidió escoger el valor $\alpha = 2$, que resulta un buen compromiso entre rechazo a otros elementos de simetría no radial (por ejemplo, pestañas), exactitud y costo computacional y es el valor recomendado por varios autores [34], [35].

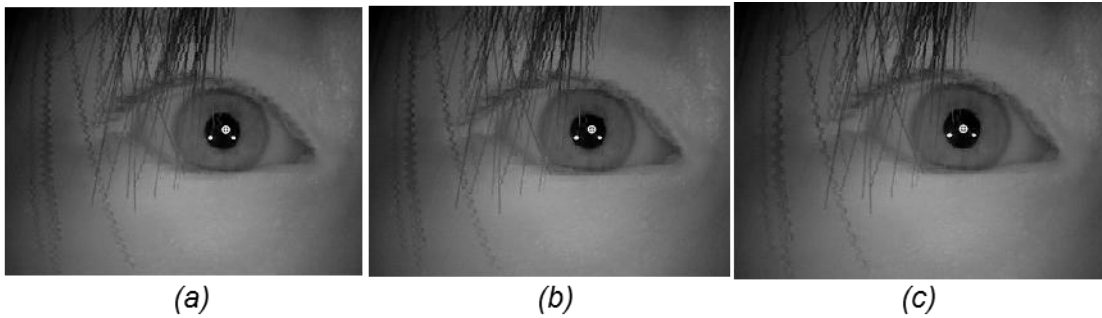


Figura 2.17 Efecto de escoger (a) $\alpha = 1$, (b) $\alpha = 2$, (c) $\alpha = 3$ en una imagen influenciada por el cabello

2.2.1.4 Factor de normalización k_n .

El factor de normalización k_n se usa para poder comparar o combinar en una misma escala, las imágenes de simetría calculadas para diferentes radios. Para normalizar, se divide a O_n y M_n entre sus valores máximos, como se propone en [36]. De esta forma, la ecuación 2.8 se convierte en:

$$F_n(p) = \frac{M_n(p)}{k_{n1}} \left(\frac{|O_n(p)|}{k_{n2}} \right)^\alpha \quad (2.11)$$

donde $k_{n1} = \max(abs(Mn))$ y $k_{n2} = \max(abs(On))$ son los factores de normalización.

Las funciones *abs* y *max* de MATLAB, devuelven el valor absoluto y el máximo de sus argumentos, respectivamente. De esta forma, los elementos de la matriz $F_n(p)$ tendrán valores entre 0 y 1, donde 1 representa la mayor contribución de simetría radial.

2.3 Métodos usados para la anotación de imágenes

Las anotaciones (referencias aceptadas como ideales) son un prerequisite importante para el desarrollo y la evaluación de muchos algoritmos en el área de visión artificial. La anotación manual del centro de la pupila es un proceso laborioso, tedioso y propenso a errores aun cuando la morfología de la pupila y/o el iris se visualice claramente y no se encuentre interferida por diversos elementos como oclusiones, destellos, desenfoques, problemas de contraste iris- pupila, presencia de pestañas y cabellos, entre otros. En los siguientes apartados se explican los fundamentos teóricos de dos métodos implementados para realizar la anotación semi-automática del centro de la pupila en aras de ganar en exactitud y facilitar el trabajo del anotador.

2.3.1 Ajuste de Elipse

Consiste en un método de ajuste de mínimos cuadrados de elipses, el cual confiere ventajas en términos de sensibilidad a las oclusiones y al ruido. El algoritmo de ajuste de la elipse utilizado está basado en el trabajo del Fitzgibbon, Pilu y Fisher [37] que fue concebido, de forma general, para puntos dispersos.

En la Figura 2.18 se muestra el diagrama de flujo correspondiente al algoritmo de ajuste de la elipse, el cual debe recibir como datos de entrada: un vector con las coordenadas x y y de los puntos marcados por el anotador. A continuación, estos datos se normalizan, para situar el centro de la elipse en el origen de coordenadas.

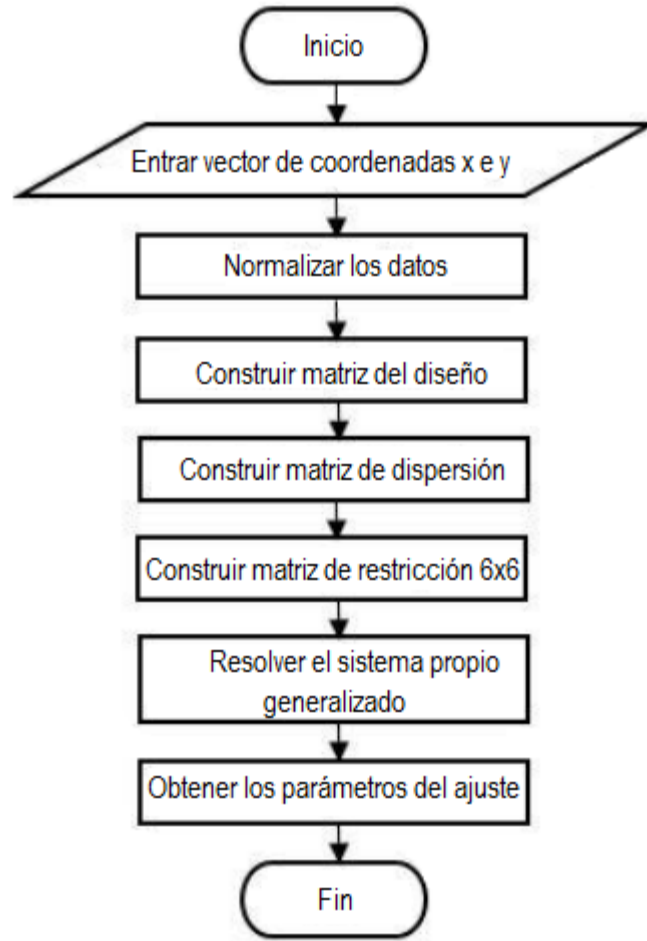


Figura 2.18 Diagrama de flujo del algoritmo de ajuste de elipse

Cualquier cónica, de forma general, puede ser representada por un polinomio de segundo orden como el siguiente:

$$F(a, x) = a \cdot x = ax^2 + bxy + cy^2 + dx + ey + f = 0 \quad (2.12)$$

Donde $a = [a \ b \ c \ d \ e \ f]^T$ y $x = [x^2 \ xy \ y^2 \ x \ y \ 1]^T$;

$F(a, x_i)$ es la llamada “distancia algebraica” de un punto (x, y) a la cónica $F(a, x) = 0$

El ajuste de una cónica general puede ser abordado minimizando la suma de las distancias algebraicas cuadradas de la curva para los N puntos x_i .

$$\mathcal{D}_A(a) = \sum_{i=1}^N F(x_i)^2 \quad (2.13)$$

Para evitar la solución trivial $a = 0$, se aplican restricciones al vector a . La minimización de las distancias puede ser resuelta considerando el sistema de valor propio generalizado:

$$D^T D a = \lambda C a \quad (2.14)$$

Donde:

$D = [x_1 \ x_2 \ \dots \ x_n]^T$ es la matriz de diseño

$D^T D$ es la matriz de dispersión

C es la matriz de restricciones 6×6 .

En el caso específico de la elipse, la restricción es cuadrática de la forma $4ac - b^2 = 1$ y puede expresarse en la forma matricial $a^T C a = 1$ como:

$$a^T \begin{pmatrix} 0 & 0 & 2 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} a = 1 \quad (2.15)$$

Con las ecuaciones y restricciones anteriores se construye y resuelve un sistema propio, el cual posee 6 pares de autovalores y autovectores. Para más detalles sobre la base matemática, véase [37]. Por último, la función de ajuste devuelve los siguientes parámetros: centro de la elipse, radio y orientación. Con estos datos es posible trazar la elipse ajustada a la pupila y determinar el centro de ella.

Este método también se usa para detectar la pupila y será uno de los métodos de comparación usados en este trabajo.

En algunos trabajos la pupila se modela como un círculo y la detección se realiza de acuerdo con este modelo [38], [39]. Sin embargo, el modelo elíptico posee una mayor adaptabilidad, flexibilidad y robustez en imágenes reales.

2.3.2 Ajuste de Círculo

El problema del ajuste de círculo utilizado en el presente trabajo puede formularse de la siguiente manera:

- Dado el conjunto de n puntos $P_i(x_i, y_i)$ para $i = (1, 2, \dots, n)$, encontrar el centro $C(x_c, y_c)$ y el radio r del círculo que pasa más cerca de todos los puntos.

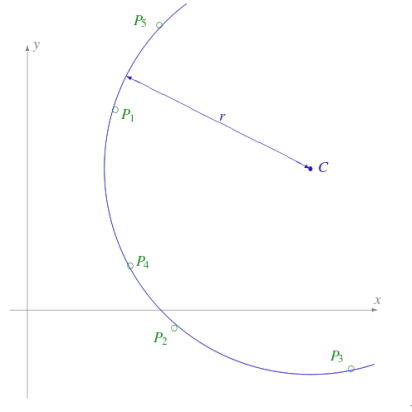


Figura 2.19: Ajuste de círculo con centro C y radio r , dado los puntos P_n .

El método de ajuste de círculo geométrico implementado está basado en dos pasos fundamentales: la selección de un círculo inicial y la minimización de forma iterativa de la distancia entre el círculo inicial y el conjunto de puntos seleccionados por el usuario.

En la inicialización se construye un círculo inicial basado en la propiedad de que, para cualquier triplete dado de puntos no alineados, hay un solo círculo que pasa por los tres puntos, como se observa en la Figura 2.20. El circuncentro triangular está en la intersección de las bisectrices perpendiculares de los lados del triángulo.

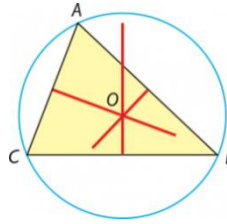


Figura 2.20: Punto circuncentro

Una vez que se tiene la suposición inicial para el centro del círculo, se realiza un proceso iterativo para obtener el mejor ajuste con relación a los puntos establecidos. El uso de un estimador de mínimos cuadrados basado en la distancia euclidiana entre los puntos y el círculo es una opción común para dicho ajuste. Para esto se realiza la minimización de la siguiente expresión:

$$J = \sum_{i=1}^n (d_i - r)^2 \quad (2.16)$$

$$d_i = \sqrt{(x_i - x)^2 + (y_i - y)^2} \quad (2.17)$$

Donde: d_i es la distancia euclidiana del punto $P_i(x_i, y_i)$ al centro del círculo $C(x, y)$ y r es el radio del círculo.

Esta minimización se realiza usando el algoritmo de Levenberg-Marquardt [40] con 50 iteraciones. Este algoritmo es una combinación de los métodos de optimización Gauss–Newton y el gradiente descendente, y es muy utilizado en la resolución de problemas de mínimos cuadrados no lineales. Los resultados de este proceso iterativo son el centro y radio del círculo, ajustados a los puntos seleccionados por el usuario.

3 . RESULTADOS Y DISCUSIÓN

En este capítulo primeramente se explica la metodología de evaluación del algoritmo de la TSR y se analizan los resultados obtenidos. Por último, se comparan estos con los obtenidos por los algoritmos de ajuste de elipse y el operador integro-diferencial y se realiza la discusión de los resultados anteriores.

3.1 Metodología de evaluación

La evaluación del algoritmo de la TSR se realizó con 400 imágenes de la base de datos CASIA descrita en el apartado 2.1.1. De las 400 imágenes de la CASIA fueron seleccionadas aleatoriamente: 200 del subconjunto *CASIA-Iris Thousand*, 100 del *CASIA-Iris Lamp* y 100 del *CASIA-Iris Twins*.

El algoritmo de detección fue aplicado a todas las imágenes. El procedimiento de evaluación abarca los pasos siguientes:

- Análisis de la exactitud

Para el conjunto total de imágenes, se calculó la tasa de aciertos definida como:

$$T_{aciertos} = \frac{P_c}{T_i} \cdot 100\% \quad (3.1)$$

Donde:

P_c es el total de las pupilas detectadas correctamente

T_i es el total de imágenes

Para todas las imágenes evaluadas se anotó el centro geométrico de la pupila con ayuda de la herramienta explicada en el apartado 3.2. Las coordenadas anotadas constituyen el valor de referencia. Se considera una detección de pupila correcta si el error (e) entre el centro estimado y el centro anotado es menor o igual que el 25 % de la distancia euclidiana (d) entre dichos centros, dividida por el radio de la pupila R , lo cual se expresa matemáticamente por la siguiente expresión:

$$e \leq \frac{0,25d}{R} \cdot 100\% \quad (3.2)$$

Esta métrica es más adecuada que considerar a la distancia euclidiana menor de 6 píxeles [33], ya que la pupila se dilata o contrae, lo cual se tiene en cuenta en la métrica relacional propuesta al dividir por el radio de la pupila. Además, es robusta cuando la distancia a la que se captura la imagen del ojo no es la misma. En estos casos, la distancia euclidiana en píxeles equivaldría a diferentes distancias longitudinales, pero la relación entre dicha distancia y el radio de la pupila, da una medida del error de estimación que es más invariante a escala. En este trabajo, el criterio del 25 % fue determinado experimentalmente, a partir de un análisis previo de las 400 imágenes experimentales.

- Análisis de la robustez

La robustez, en este trabajo, se cuantifica analizando el valor de la tasa de aciertos en imágenes con perturbaciones que puedan afectar el desempeño del algoritmo, con respecto a las imágenes claras (sin perturbaciones). Para esto, del conjunto total de 400 imágenes de la CASIA, se subdividieron en: 165 imágenes claras, 85 imágenes influenciadas por el cabello y las pestañas, 50 imágenes influenciadas por el párpado, 42 imágenes con desenfoque y 58 imágenes influenciadas por espejuelos y reflexiones. A cada uno de estos grupos se les calculó la tasa de aciertos a partir de la expresión (3.1).

- Cálculo del costo computacional

El costo computacional se estimó a través del cálculo del tiempo de ejecución promedio en que se procesa la misma imagen por cada algoritmo. Para esto se usó la función “*tíc-toc*” de MATLAB 2020a, que devuelve este tiempo en segundos. Este indicador aporta información valiosa en el contexto de la aplicación final de un sistema de VOG.

3.2 LABELPUP: Herramienta para anotación y detección del centro de la pupila

La herramienta “LABELPUP” fue desarrollada usando la herramienta *GUIDE* de MATLAB 2020a y se obtuvo una versión ejecutable denominada LABELPUP.exe, que permite su uso en computadoras que no tengan instalada la aplicación MATLAB. En la figura 3.1 se muestra la interfaz gráfica de usuario de la herramienta “LABELPUP” en la cual se pueden identificar los siguientes componentes:

- Barra de herramientas: compuesta por 15 botones con distintas prestaciones que facilitan el proceso de marcación.

- Panel principal: área para la visualización de la imagen ocular original o después de aplicarles algunas operaciones.
- Cuadros de texto y botones asociados: se utilizan para definir el nivel empleado para la operación de umbralización y para escoger el algoritmo de detección automática del centro de la pupila que se aplicará a la imagen. En este último caso, el usuario puede elegir entre tres algoritmos diferentes: (1) la Transformada de Simetría Radial, (2) el Operador Integro-diferencial y (3) el Ajuste de Elipse.

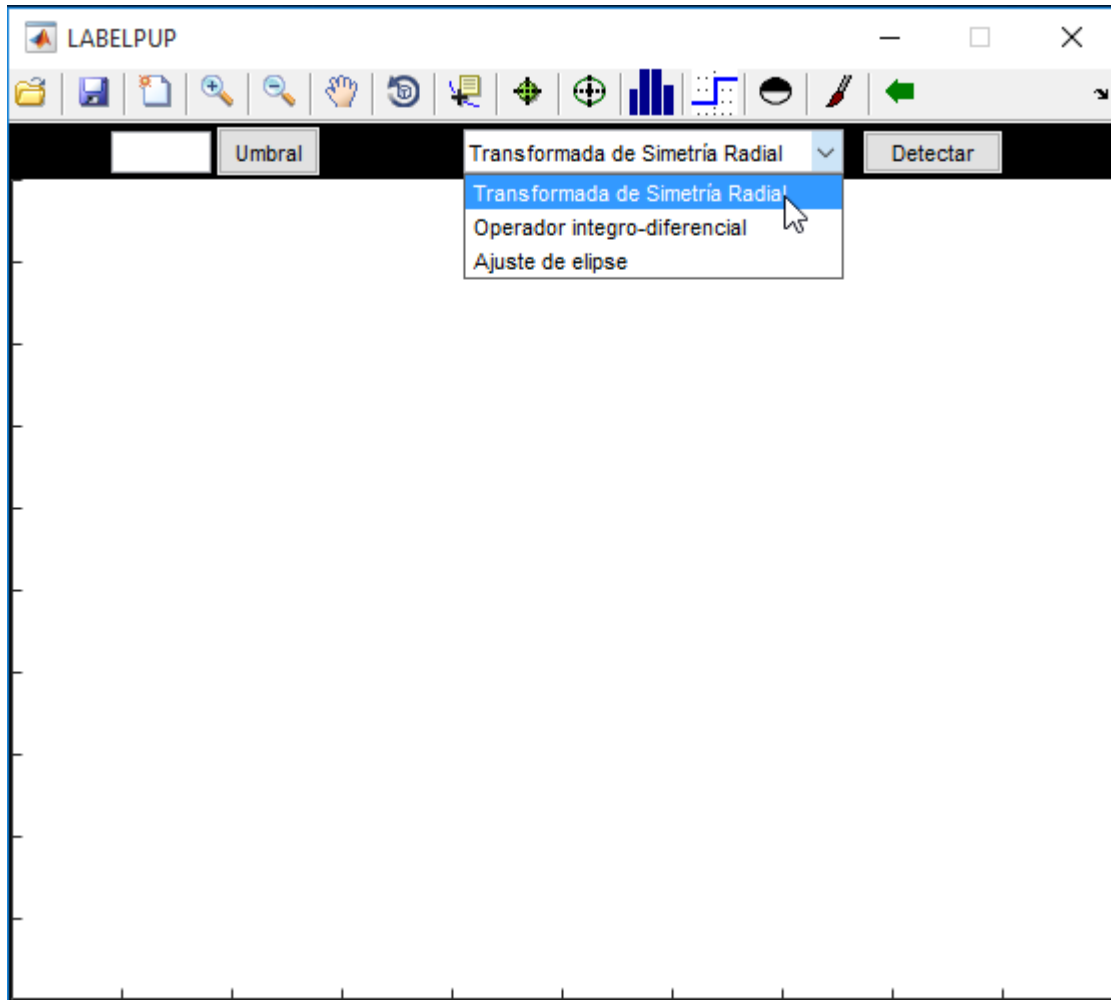


Figura 3.1 Interfaz gráfica de usuario de la herramienta software “LABELPUP”

En la figura 3.2 se muestra la barra de herramientas, cuyos botones se describen a continuación:

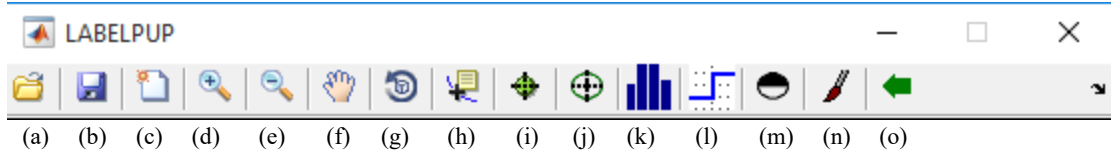


Figura 3.2 Barra de herramientas del software “LABELPUP”

a. Lectura y visualización de la imagen

El botón “Cargar Imagen” (ver figura 3.2 (a)) abre un cuadro de diálogo, mostrado en la figura 3.3, que permite acceder a la ubicación de la imagen ocular (en formato.jpg) para su lectura y visualización en el panel principal.

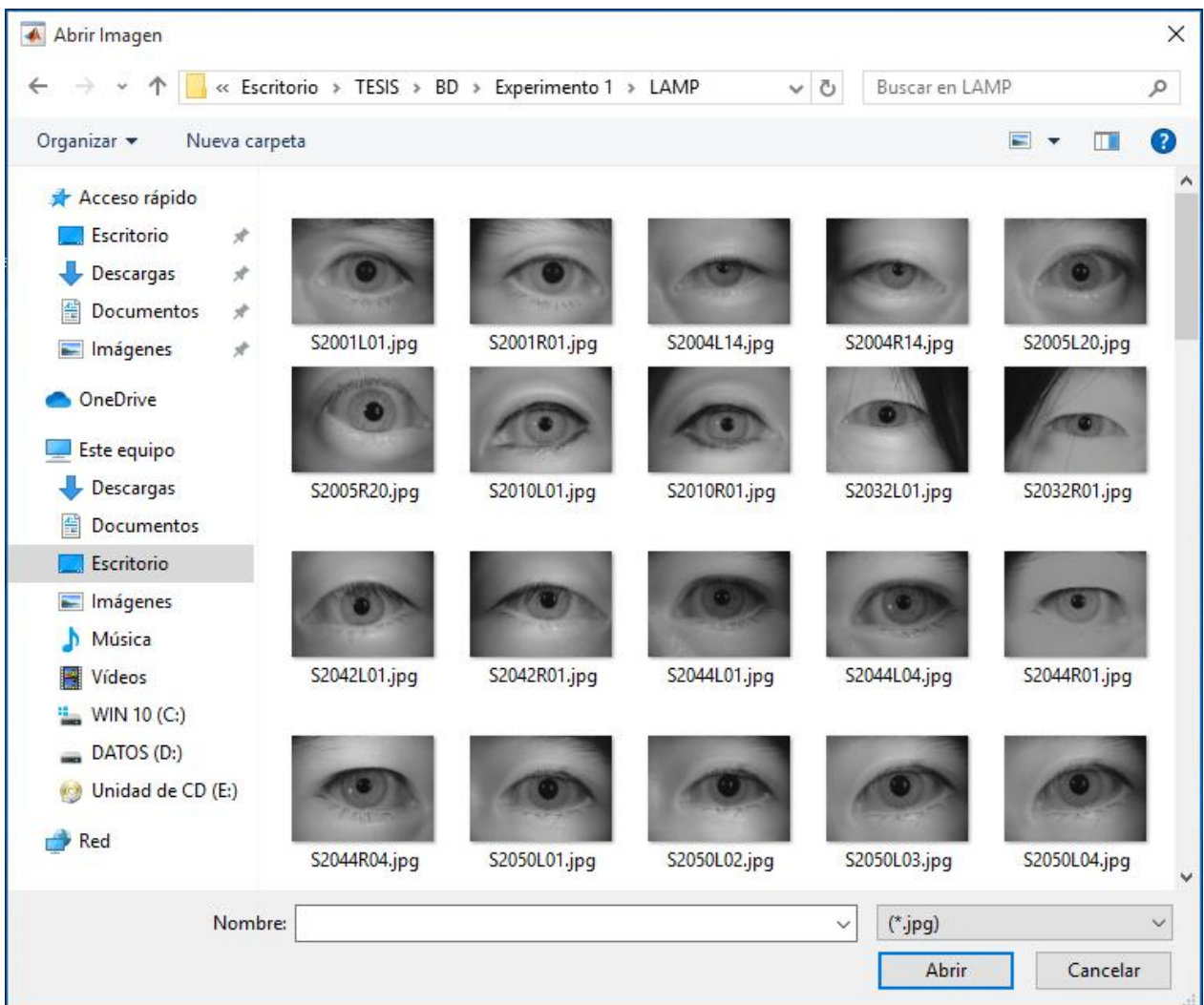


Figura 3.3 Cuadro de diálogo asociado al botón “Cargar imagen”

b. Guardar anotación

El botón “Guardar anotación” (ver figura 3.2 (b)) visualiza en el panel principal un cursor gigante en forma de cruz, mostrado en la figura 3.4, que le permite al usuario estimar visualmente el centro de la pupila a anotar. Una vez que dicho cursor está situado en la posición deseada, se realiza un clic izquierdo del mouse para capturar y salvar en un fichero con formato *.csv, las coordenadas (x, y) de dicha posición. Los archivos CSV (del inglés *comma-separated values*) son un tipo de documento en formato abierto sencillo para representar datos en forma de tabla, en las que las columnas se separan por comas y las filas por saltos de línea. En la figura 3.4 se muestra un ejemplo.

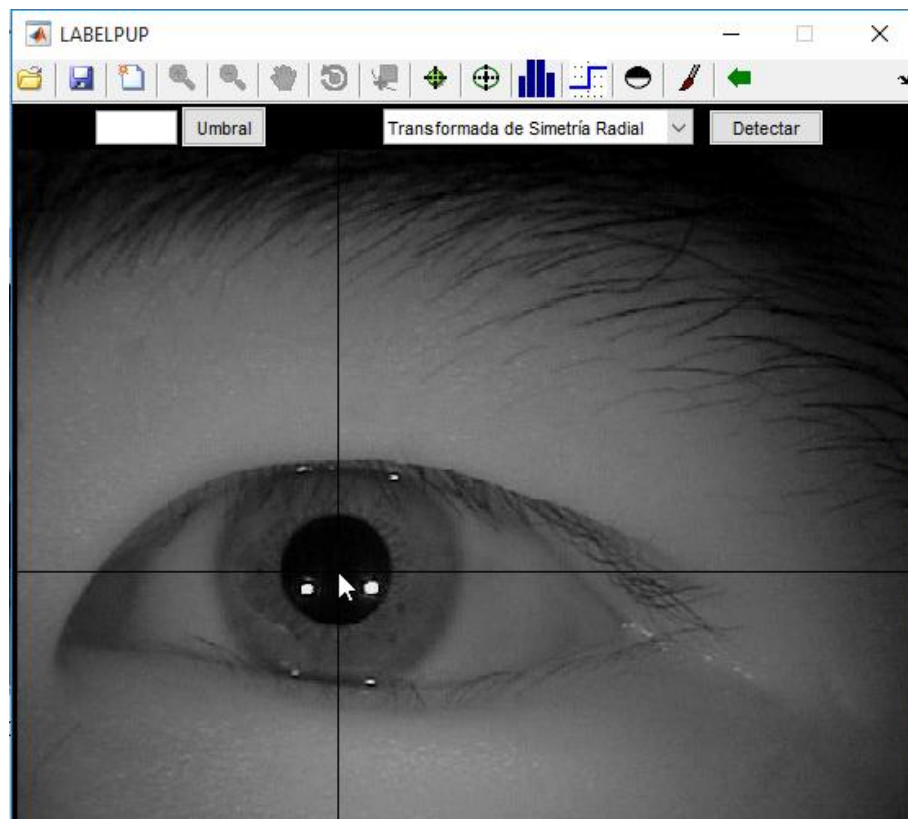


Figura 3.4 Anotación manual de la imagen S5006R00.jpg de CASIA-Iris Thousand

c. Leer anotación

Este botón permite, una vez anotado el centro de la pupila en una imagen dada, leer las coordenadas del archivo *.csv asociado y dibujar un círculo cuyo centro está ubicado en las coordenadas leídas.

d/e. Ampliar/Reducir imagen

Estos botones permiten aumentar y disminuir el área de la pupila para localizar mejor su centro.

f. Desplazamiento en la imagen

Este botón permite al anotador recorrer otras partes de la imagen, posterior al aumento de zoom.

g. Rotar imagen

Permite rotar la imagen en cualquier ángulo. Esto es útil para imágenes adquiridas con cámaras propias.

h. Cursor de estado

Brinda información del estado del píxel donde se coloque el cursor, lo cual incluye, las coordenadas (x,y) y la información del nivel de gris.

i. Ajuste de círculo

Este botón usa el método descrito en la sección 2.3.2. Para esto, el usuario debe marcar cinco puntos (x,y) que se encuentren en el borde de la pupila, los cuales representan la entrada del algoritmo. A partir de aquí, este realiza el ajuste de un círculo que devuelve el centro y el radio del mismo. En la figura 3.5 se ilustra el resultado obtenido para un ojo con oclusión parcial del párpado. Si no existe un círculo que se ajuste a los cinco puntos marcados, entonces el algoritmo no devuelve ningún parámetro. En consecuencia, no se grafica nada. La situación anterior generalmente indica que una elipse se adaptaría mejor al contorno de la pupila. En la figura 3.5 se muestran 2 casos en los cuales el algoritmo de ajuste de círculo facilita la tarea del anotador.

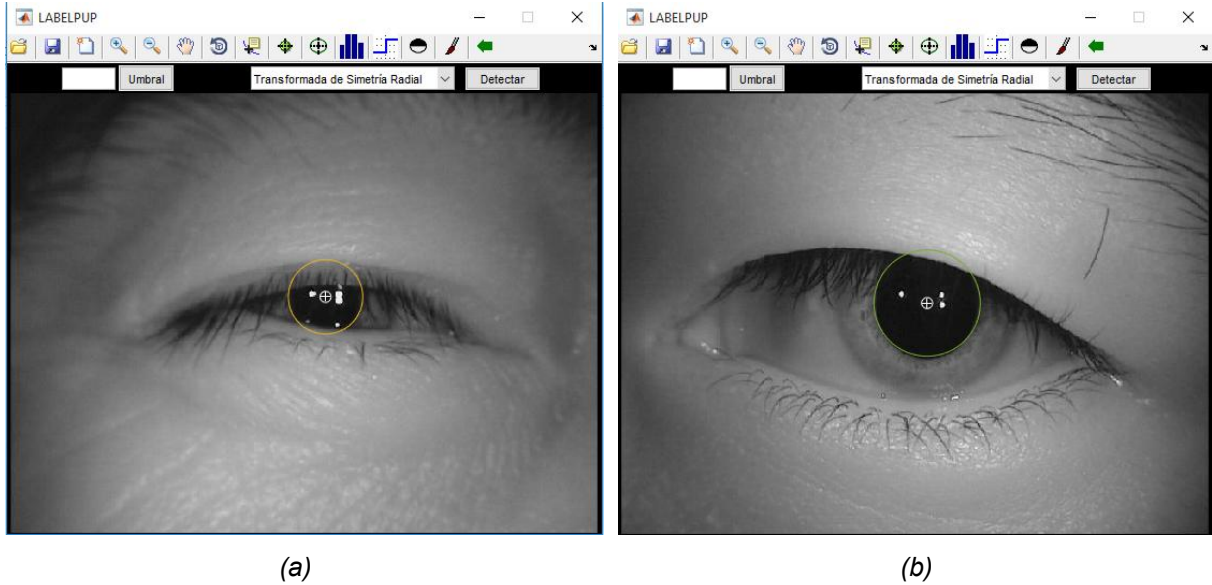


Figura 3.5 Resultados del ajuste de círculo en la imagen: (a) S2244R02.jpg de CASIA-Iris LAMP (b) S2089L01.jpg de CASIA-Iris LAMP. Esto permite predecir la posición del centro de la pupila y facilitar la anotación manual en caso de ojos parcialmente ocluidos.

j. Ajuste de elipse

Este botón proporciona un método de apoyo al usuario en casos donde existe oclusión parcial del párpado o donde, de manera general, no resulta visible la pupila en su totalidad. También puede ser útil en imágenes en donde la pupila se encuentre en perspectiva, como el caso ilustrado en la figura 3.6. En casos como este, se genera una cierta incertidumbre en la posición del centro a la hora de realizar la marcación.

El procedimiento para el empleo de esta funcionalidad se basa en marcar cinco puntos en los lugares donde se pueda visualizar los bordes de la pupila. Estos pares coordenados (x, y) ajustan una elipse siguiendo el algoritmo descrito en el apartado 2.3.1 del capítulo anterior, que devuelve los siguientes parámetros: centro de la elipse, radio y orientación. Con estos datos es posible trazar la elipse ajustada a la pupila y determinar el centro de ella, como se observa en la figura 3.6.

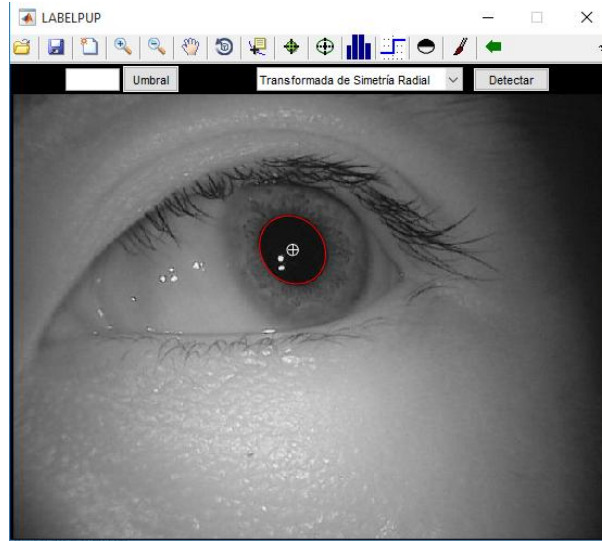


Figura 3.6 Resultados del ajuste de elipse en la imagen S2138L04.jpg de CASIA-Iris LAMP. Esto permite predecir la posición del centro de la pupila y facilitar la anotación manual

k. Ecualización de histograma

Este botón permite mejorar el contraste de la imagen mediante la redistribución de las frecuencias con que aparecen los diferentes niveles de grises en el histograma de una imagen [41]. Para la ecualización del histograma se empleó el método CLAHE [42]. En la figura 3.7 (b) se muestra un ejemplo de la mejora de contraste entre el iris y la pupila luego de aplicar la ecualización de histograma.

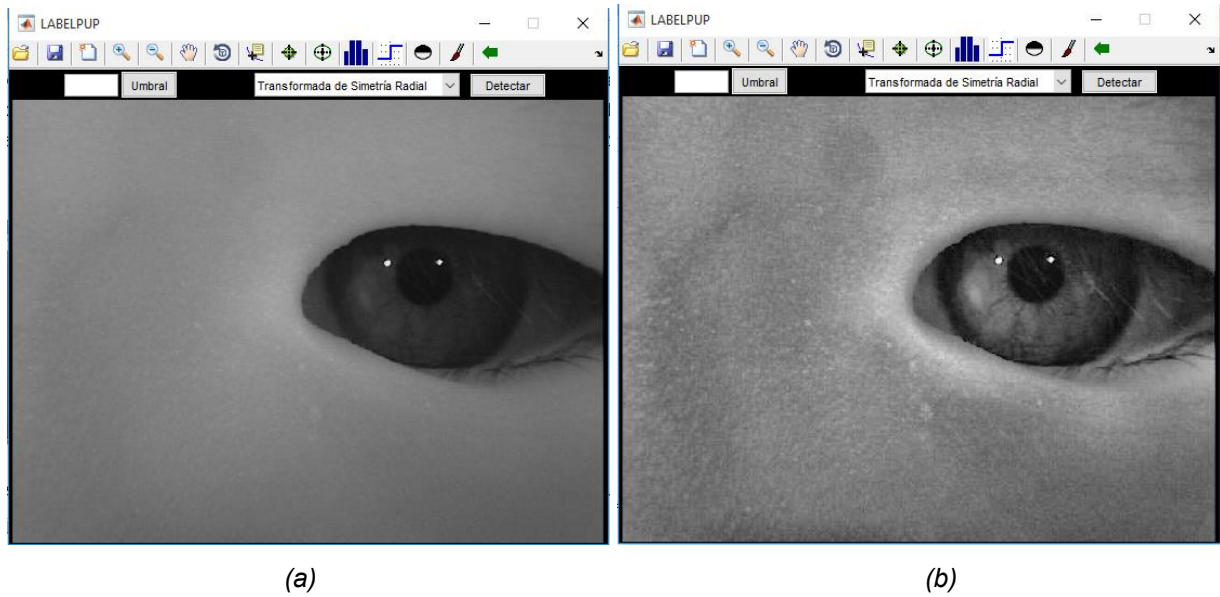


Figura 3.7 Resultados de la ecualización de histograma en la imagen S2138L04.jpg de CASIA-Iris LAMP.

I. Umbralización

La umbralización es uno de los métodos de segmentación más importantes, cuyo objetivo es convertir una imagen en escala de grises a una nueva con sólo dos niveles, de manera que los objetos queden separados del fondo. Una forma de extraer el objeto del fondo es seleccionar un umbral de intensidad T que separe los dos conjuntos. De este modo, cualquier punto (x, y) para el que $f(x, y) > T$, pertenecerá al objeto, en caso contrario, pertenecerá al fondo. Si los valores de gris del objeto y del resto de la imagen difieren claramente, entonces el histograma mostrará una distribución bimodal, con dos máximos distintos, separados por una región vacía. Con lo cual se logrará una separación perfecta entre el objeto y el fondo, al establecer un valor umbral ubicado en esta región del histograma. Este botón realiza la umbralización con un nivel de gris definido por el usuario mediante el botón umbral que se encuentra en el panel principal de la figura 3.1. En la figura 3.8 se muestra el efecto de esta operación.

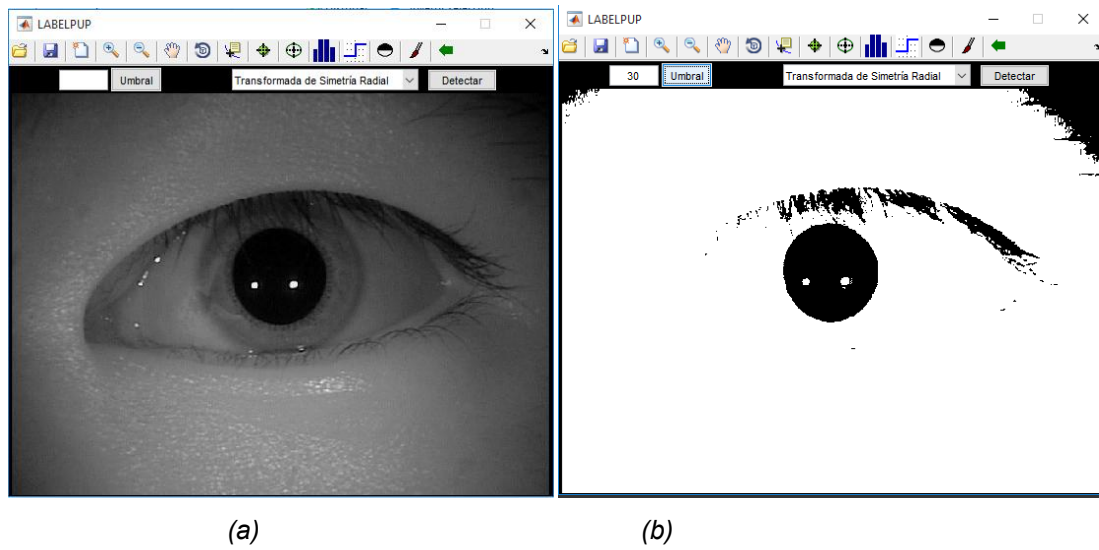


Figura 3.8 Resultados de la umbralización de la imagen S5004R04.jpg de CASIA-Iris Thousand:
(a) imagen original (b) imagen posterior a la umbralización con un nivel de 30

m. Inversión de la imagen

Este botón permite obtener "la inversa" o el "negativo" de la imagen original, es decir, aquellos píxeles cuyos valores era máximos ahora son mínimos y viceversa.

n. Borrar panel principal

Este botón fue concebido para volver al estado inicial sin necesidad de cerrar la herramienta.

o. Restaurar

Este botón retorna a la imagen cargada originalmente para realizar su correspondiente anotación. Es útil en aquellos casos donde las operaciones auxiliares con las que cuenta la herramienta PUPILA no proporcionan una mejoría en la calidad de la imagen que permita realizar la anotación con mayor facilidad.

3.3 Evaluación del algoritmo de detección del centro de la pupila

El algoritmo implementado se evaluó en una computadora con microprocesador Intel Core i5 4670S a 3.1 GHz, 4 GB de memoria RAM DDR3 y sistema operativo *Windows 10 Professional* para arquitectura de 64 bits.

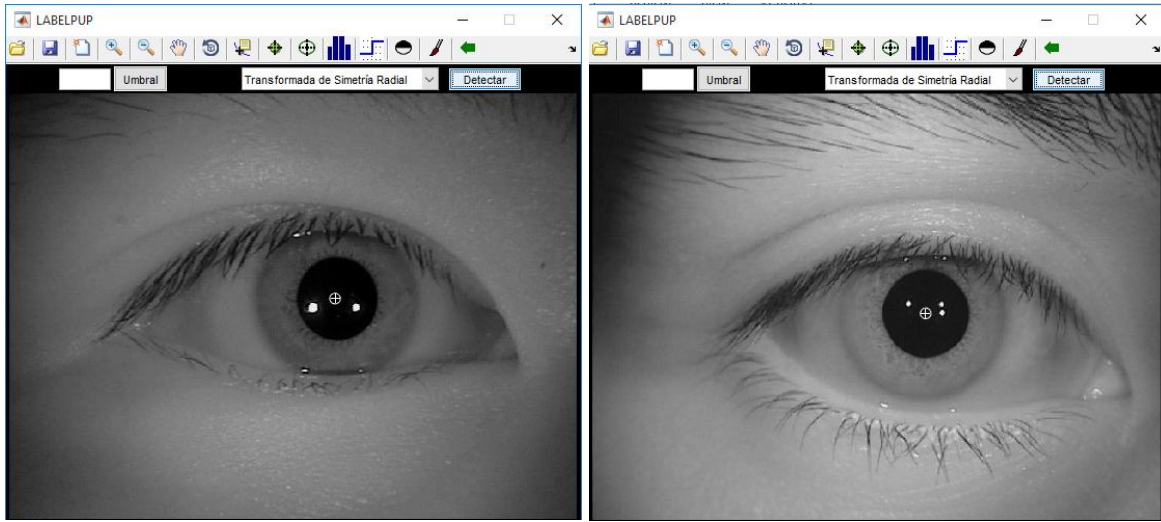
3.3.1 Exactitud

En la Tabla 3.1 se muestran los resultados de la tasa de aciertos para cada subconjunto de la base de datos CASIA. La tasa de aciertos global para todas las imágenes fue de 95 % lo cual representa un buen resultado. Los resultados de la *CASIA-Iris Thousand* y la *CASIA-Iris Lamp* fueron de 98% y 96% respectivamente, mientras que en la *CASIA-Iris Twins* se obtuvo un 91%. Este último resultado se debió a que las 9 imágenes cuya detección fue incorrecta presentaban poco contraste, falta de nitidez, desenfoque y elevada cantidad de ojos semi-abiertos.

Tabla 3.1. Tasa de aciertos para los tres subgrupos de la base de datos CASIA

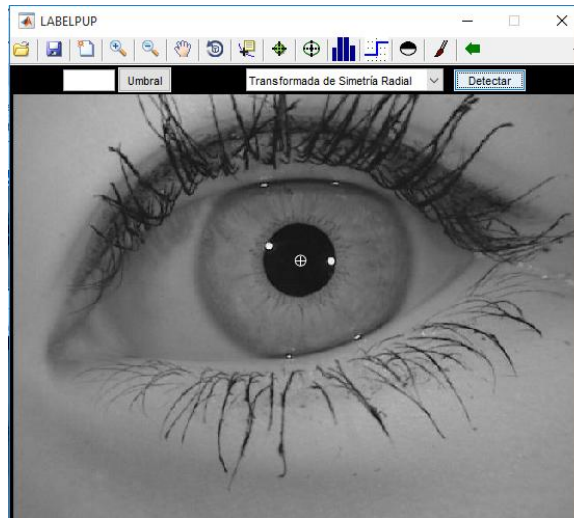
Subconjunto/ Clasificación	<i>CASIA-Iris Thousand</i>	<i>CASIA-Iris Lamp</i>	<i>CASIA-Iris Twins</i>
Detección correcta	196	96	91
Detección incorrecta	4	4	9
Total de imágenes	200	100	100
Tasa de aciertos	98%	96%	91%

En la figura 3.9 se muestra el resultado de la localización de la pupila para imágenes de cada subconjunto.



(a)

(b)



(c)

Figura 3.9 Ejemplos de detección correcta del centro de la pupila utilizando la TSR en las imágenes: (a) S5008L01.jpg de CASIA-Iris Thousand (b) S2001R01.jpg de CASIA-Iris Lamp (c) S3081R01.jpg de CASIA-Iris Twins

3.3.2 Robustez

En esta sección se analiza el funcionamiento del algoritmo en imágenes que se capturan en diversas condiciones. Para demostrar la robustez del algoritmo propuesto, se realizaron cinco experimentos dividiéndose el conjunto de prueba en imágenes: (a) claras, (b) influenciadas por

el cabello y las pestañas, (c) influenciadas por el párpado, (d) influenciadas por espejuelos y reflexiones y (e) con desenfoque.

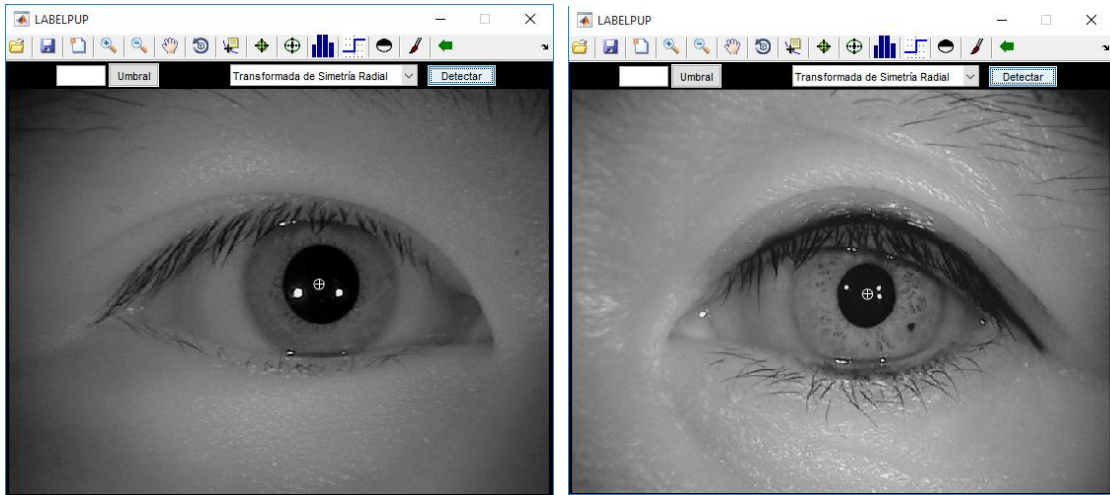
En la tabla 3.2 se presentan los resultados y se observa que el algoritmo basado en la TSR es altamente robusto a reflexiones dentro del área de la pupila y cerca de esta, lo cual representa una gran ventaja ya que es una condición que ocurre a menudo en el entorno. Asimismo, es muy tolerante al uso de espejuelos pues logra una tasa de aciertos elevada. Además, su desempeño es adecuado aún en imágenes que presentan un ligero desenfoque y que se encuentran interferidas por elementos externos tales como cabellos y pestañas.

La tasa de aciertos decreció al procesar imágenes en las que el ojo estaba casi ocluido por el párpado. Sin embargo, los resultados obtenidos aún en estas situaciones desfavorables se consideran aceptables.

Tabla 3.2. Tasa de aciertos para los diferentes experimentos para el análisis de la robustez usando la base de datos CASIA

Experimento/ Clasificación	Correctas	Incorrectas	Total	Tasa de aciertos (%)
Imágenes claras	163	2	165	98,78
Imágenes influenciadas por el pelo y las pestañas	74	11	85	87,05
Imágenes influenciadas por el párpado	42	8	50	84
Imágenes con desenfoque	40	2	42	95,23
Imágenes influenciadas por espejuelos y reflexiones	55	3	58	94,82

En la figura 3.10 se muestra el resultado de la localización de la pupila para imágenes del primer experimento donde se analizaron las imágenes claras. Aquí, de manera general, la región circular de la pupila se encontraba completamente visible y presentaba buen contraste con la zona del iris, de ahí que el método haya tenido un desempeño muy bueno.

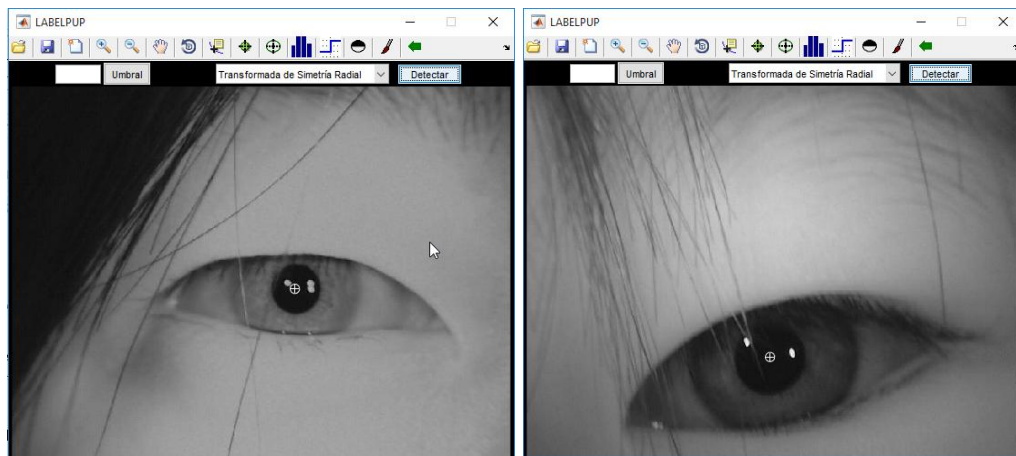


(a)

(b)

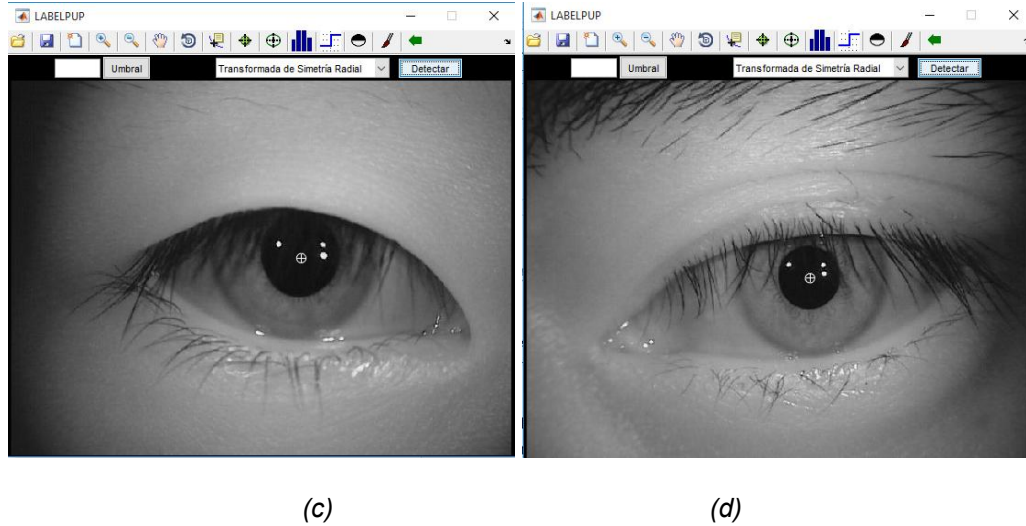
Figura 3.10 Ejemplos de detección correcta del centro de la pupila en dos imágenes del experimento correspondiente a las imágenes claras: (a) S5008L01.jpg y (b) S2247L01.jpg

En la figura 3.11 se evidencia el resultado del algoritmo para las imágenes oculares influenciadas por la presencia de cabellos y pestañas cercanas a la región de interés (pupila). Se comprobó que el algoritmo es tolerante a este tipo de perturbaciones debido a que estas no presentan simetría radial. Los casos de detección incorrecta se debieron fundamentalmente a imágenes oscuras, con poco contraste entre el iris y la pupila. En estos casos el nivel de gris de las pestañas se mezclaba con el de la pupila lo que producía un efecto que destruía la característica circular de esta, que es uno de los rasgos potencialmente empleados por el algoritmo de la TSR.



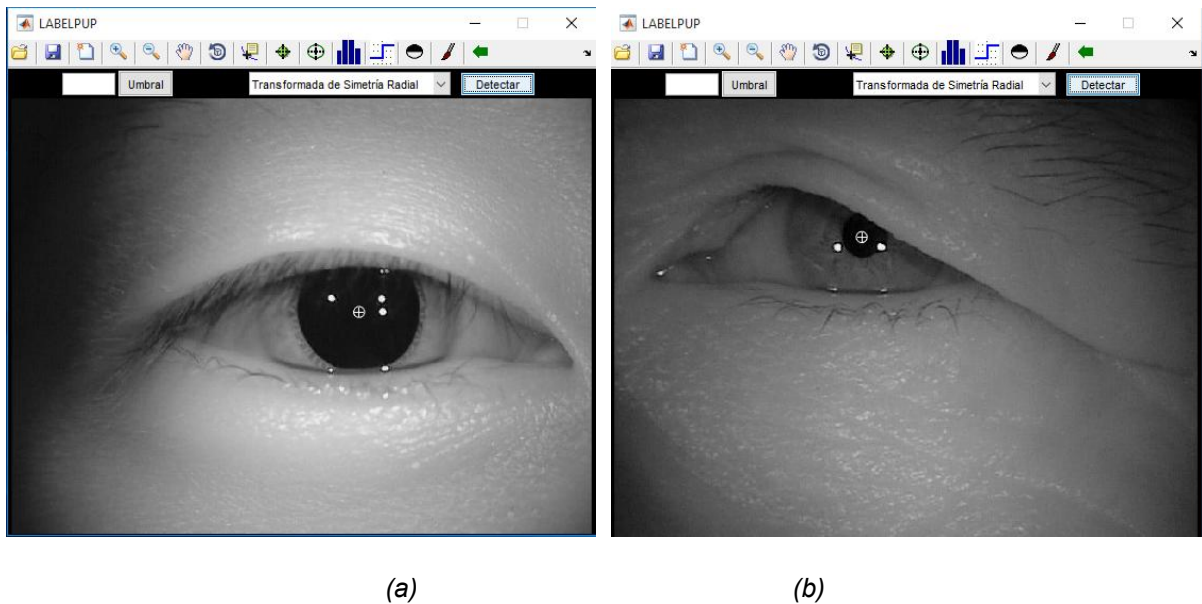
(a)

(b)



*Figura 3.11 Ejemplos de detección en imágenes oculares influenciadas por cabellos y pestañas.
(a) S2032R01.jpg (b) S3191R02.jpg (c) S2164R08.jpg (d) S2155L05.jpg*

En el tercer experimento, correspondiente a la detección de la pupila en imágenes influenciadas por el párpado y el ojo semi-ocluído, el algoritmo mostró la tasa de aciertos más baja. En la Figura 3.12 se muestran los resultados de la detección en estos casos. Se puede observar que en las imágenes en las cuales el área de la pupila se cubrió gradualmente hasta un 10% (figuras 3.12 (a) y (b)), el algoritmo puede funcionar correctamente. Sin embargo, cuando el área de la pupila está cubierta a más del 30% el algoritmo no obtiene resultados aceptables (figura 3.12 (c) y (d)).



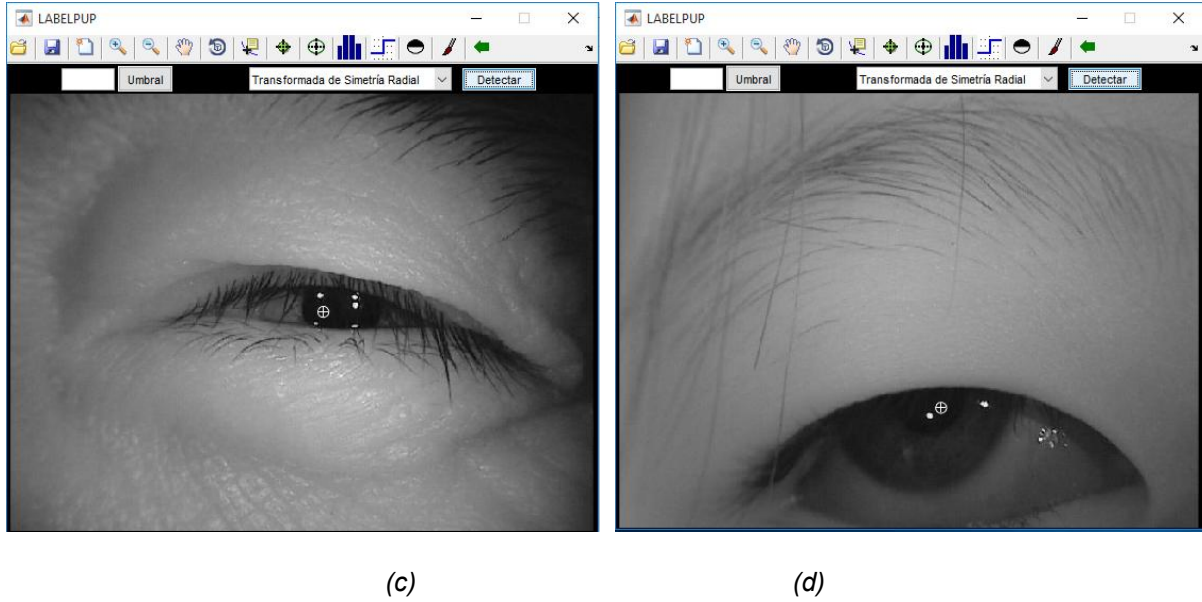


Figura 3.12 Ejemplos de detección en imágenes oculares donde el área de la pupila está cubierta parcialmente. (a) S2051R01.jpg (b) S5355R00.jpg (c) S2244L02.jpg (d) S3191L02.jpg

El algoritmo de la TSR demostró, además, ser muy robusto al localizar correctamente la posición del centro de la pupila en individuos con espejuelos (ver figuras 3.13 (a) y (b)).

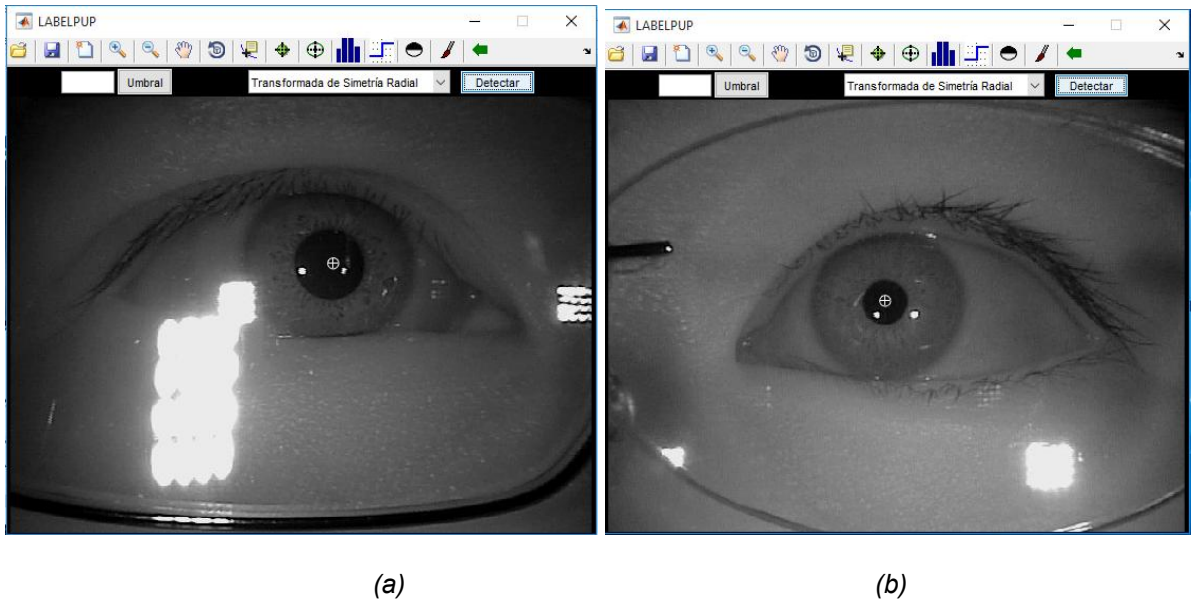


Figura 3.13 Ejemplos de detección en imágenes oculares donde el sujeto porta espejuelos, (a) S5000L03.jpg y (b) S5993R04.jpg pertenecen a la base de datos CASIA

En la figura 3.14 se muestran dos imágenes que presentan un ligero desenfoque sobre las cuales se ha probado el algoritmo. Dado que el algoritmo aprovecha la información de los bordes para

estimar el centro de la pupila, es necesario que la imagen que se le presenta tenga buena nitidez. En la figura 3.14 (a) se aprecia una detección correcta y en la (b) una incorrecta.

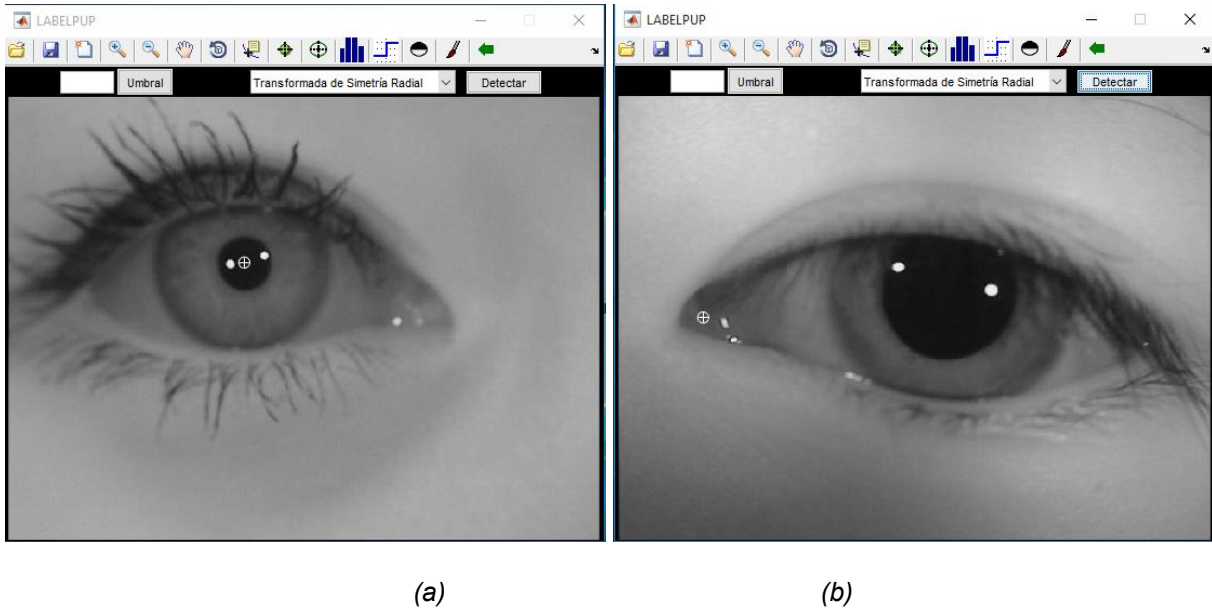


Figura 3.14 Ejemplos de detección en imágenes oculares que presentan un ligero desenfoque: (a) S3081L01.jpg y (b) S3231R01.jpg

3.3.3 Costo computacional

En la figura 3.15 se ilustra gráficamente los tiempos de ejecución medidos al procesar una imagen para los tres algoritmos comparados. Podemos observar que los más eficientes son el basado en: ajuste en elipse, la TSR y los operadores integro-diferenciales, en ese orden. Esta información acerca de la latencia de procesamiento debe ser tomada en cuenta a la hora de la aplicación de un sistema de VOG en la práctica.

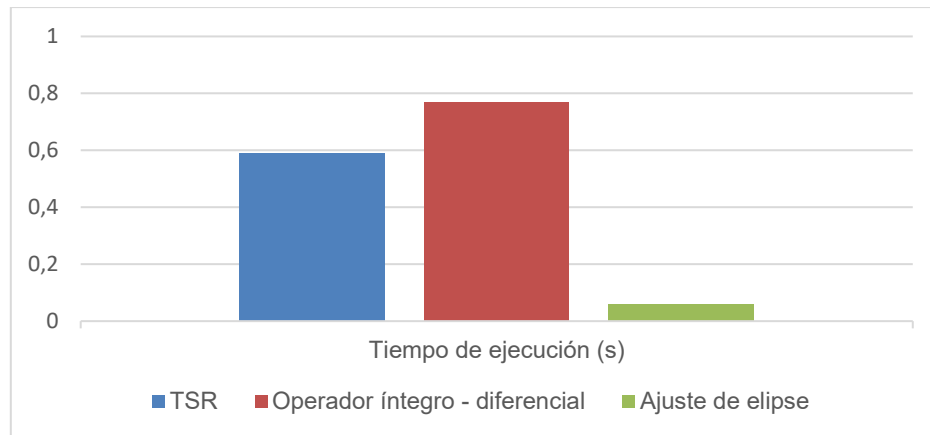


Figura 3.15 Tiempos de ejecución (s) para los tres algoritmos comparados.

3.4 Comparación de métodos

Para cada experimento, los resultados del algoritmo de la TSR fueron comparados los algoritmos de ajuste de elipse y el basado en operadores integro-diferenciales, ampliamente citados en trabajos similares. En la figura 3.16 se muestra un gráfico de barras con los resultados en cada caso. Se observa que, en todos los experimentos, la Transformada de Simetría Radial mostró los mejores resultados. El método de peor desempeño fue el de ajuste de elipse, para la mayoría de los casos. Esto se debe fundamentalmente a la dependencia de umbral de este último.

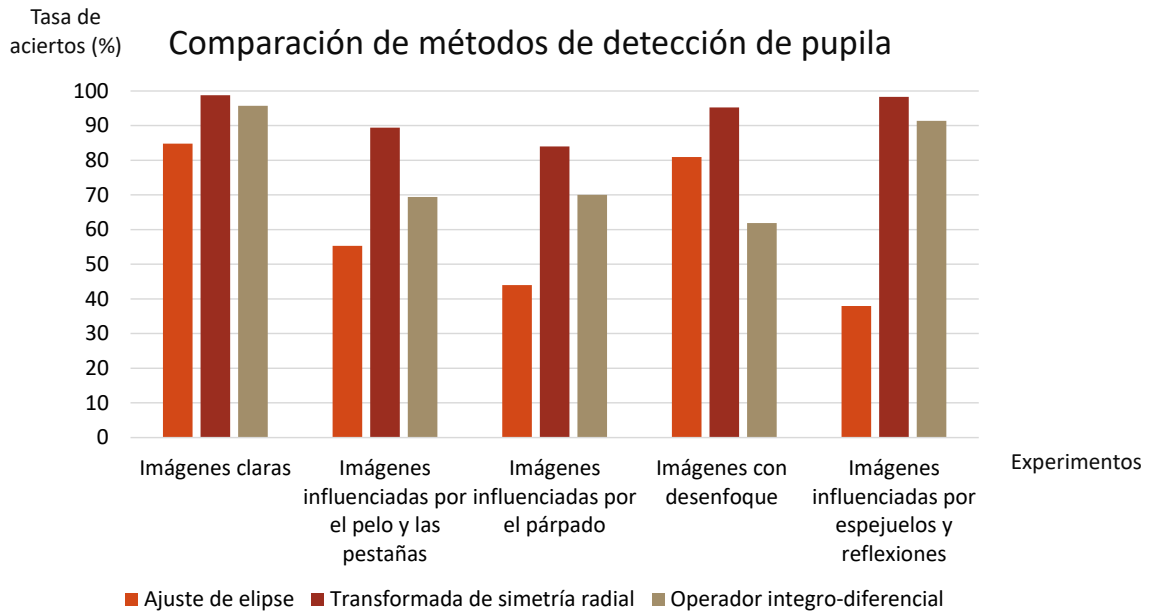


Figura 3.16 Comparación de métodos de detección de pupila con respecto a la tasa de aciertos.

En la figura 3.17 se muestra el resultado de la detección de cada uno de los métodos comparados para una misma imagen ocular (S5033L00.jpg de la base de datos *CASIA-Iris Thousand*) evidenciándose la superioridad del algoritmo de la TSR.

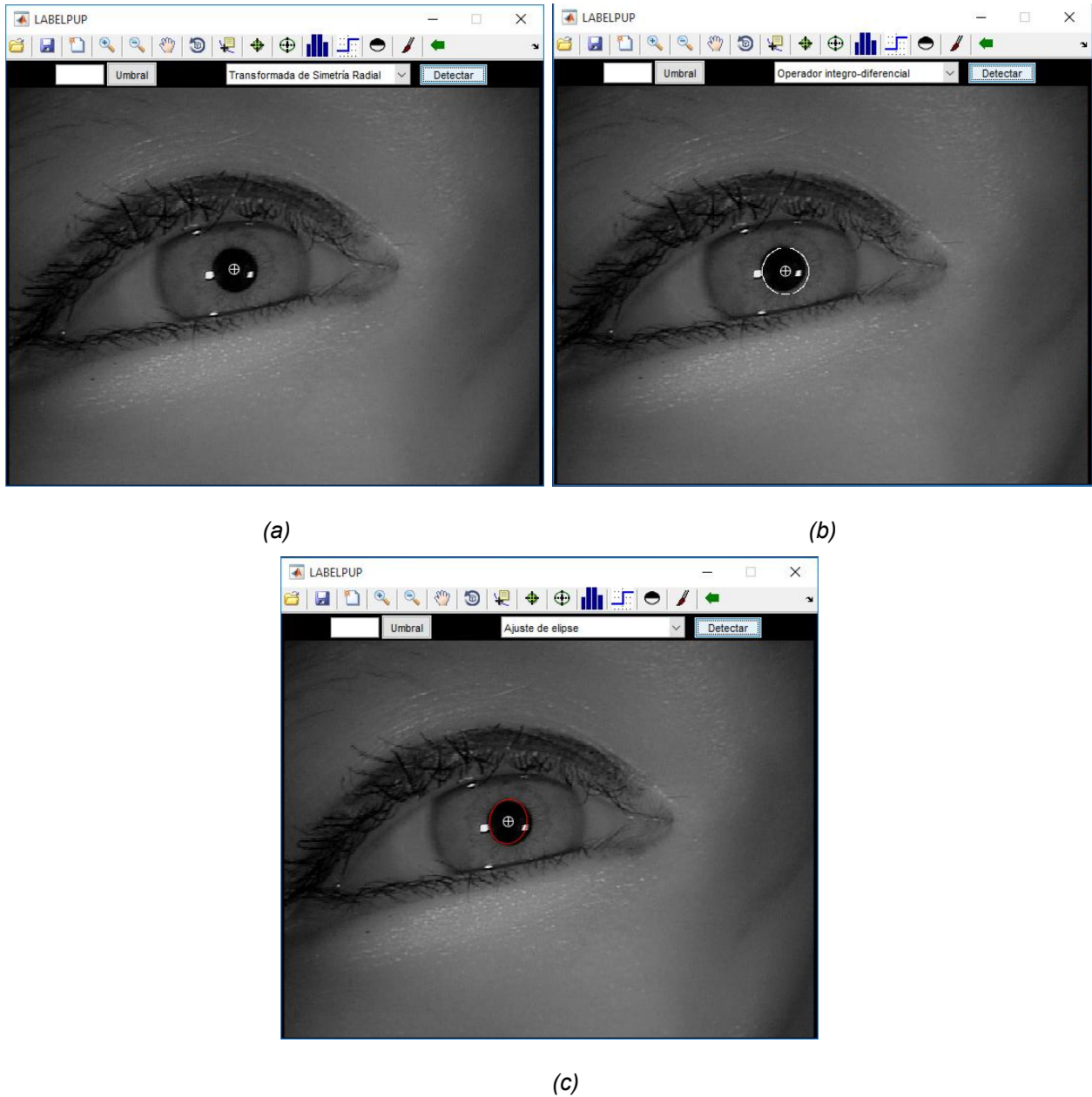


Figura 3.17 Ejemplos de detección realizado por cada uno de los algoritmos comprados en una misma imagen. (a) TSR (b) Operador Integro-diferencial (c) Ajuste de Elipse

3.5 Discusión general

De acuerdo a los análisis cuantitativos y cualitativos realizados, los resultados se consideran satisfactorios. El algoritmo de la TSR demostró ser robusto en imágenes donde aparecen espejos, gran densidad de pestañas o destellos de luz. Debido al hecho de que las pestañas y el marco de los espejos no tienen simetría radial, no importa cómo cambie el radio de

detección, sus resultados no producirán superposición y por lo tanto, no afectarán la detección del centro de la pupila. De manera similar, debido al hecho de que los destellos luminosos poseen un valor alto de nivel de gris (cercano a 255, blanco), la dirección negativa del gradiente en los puntos del borde del destello luminoso se desvía del centro del círculo, y no representará una contribución a los resultados de la TSR. Por lo tanto, el punto de luz tampoco afectará la detección del centro de la pupila. A través de los resultados experimentales se identificó que el algoritmo de la TSR presenta dificultad para procesar imágenes altamente influenciadas por el párpado. Este resultado era esperado teniendo en cuenta que esta oclusión en el área de la pupila destruye su carácter circular, lo cual también afecta al resto de los métodos analizados. En estos casos, se deberían investigar otros métodos de predicción alternativos.

La comparación del desempeño de los métodos mostró que, en la totalidad de experimentos realizados, la TSR supera los algoritmos basados en ajuste de elipse y en operadores integro-diferenciales. La explicación para este resultado no es más que la robustez que supone el uso de la característica radial.

En resumen, los resultados mostraron que este método representa una alternativa robusta, capaz de hacer frente a la mayoría de los artefactos e inconvenientes que pueden ser encontrados en imágenes de la vida real. La integración de varios algoritmos de pre-procesamiento y procesamiento en la herramienta resulta ventajosa debido a que aporta practicidad durante el trabajo con las imágenes. El trabajo realizado representa una aproximación preliminar a un sistema de VOG, especialmente en lo relacionado a los pasos de procesamiento de imagen. Su relevancia radica en su aplicación potencial y con un rápido despliegue en salas de hospitales, consultas y centros de investigación.

CONCLUSIONES Y RECOMENDACIONES

Conclusiones

- En este trabajo de investigación se realizó una revisión de los métodos más relevantes que han sido propuestos para la detección del centro de la pupila en imágenes oculares. De acuerdo a los resultados y aspectos teóricos encontrados en la literatura y los experimentos realizados en este trabajo se concluye que la Transformada de Simetría Radial es un método altamente competitivo para la determinación del centro de la pupila.
- Se logró implementar exitosamente un script en MATLAB para la detección del centro de la pupila basado en la Transformada de Simetría Radial, siendo este el resultado más relevante del trabajo. De manera similar, se implementaron otros algoritmos para la detección basados en los operadores integro-diferenciales y ajuste elipse.
- Se logró implementar en Matlab la herramienta software “LABELPUP” para realizar anotaciones y llevar a término la detección automática. La misma incluye atractivas funcionalidades para realizar la marcación manual y semiautomática facilitando su uso, exactitud y confort del especialista como usuario final. En cuanto a la detección automática, integra las tres implementaciones algorítmicas basadas en la Transformada de Simetría Radial, Operadores Integro-diferenciales y Ajuste Elipse. Esta herramienta abre nuevas posibilidades para el desarrollo de nuevos algoritmos basados en métodos de aprendizaje automático supervisado y ayuda a incrementar el número de anotaciones que puedan ser usadas por otros especialistas en futuros experimentos.
- La evaluación de la exactitud del algoritmo desarrollado con las imágenes de la CASIA mostró un alto índice de desempeño, siendo de 95 %. Además, se evidenció la robustez del algoritmo ante perturbaciones tales como influencia de espejuelos, reflexiones en la pupila y/o el iris, desenfoque de la cámara de adquisición, presencia de pelo y pestañas. De la comparación realizada del desempeño del algoritmo de la TSR con los otros dos implementados se concluye que, en la totalidad de experimentos realizados, la TSR supera los algoritmos basados en ajuste de elipse y en los operadores integro-diferenciales. Esto demostró que la robustez se sustenta en el uso de la característica

radial de la pupila. El algoritmo de la TSR ocupó el segundo lugar en términos de costo computacional, sólo superado por el del ajuste elipse que a su vez fue el que peor desempeño mostró en términos de exactitud y robustez.

Recomendaciones

Se recomienda aumentar las prestaciones de la herramienta desarrollada con otros algoritmos de procesamiento. Además, la herramienta podría beneficiarse del cálculo de otros indicadores. Por ejemplo, un medidor de diámetro pupilar, el cual brinda información del estado psicofísico del sujeto y es un indicador del nivel de vigilia o anestesia, así como de la presencia de narcóticos.

REFERENCIAS BIBLIOGRÁFICAS

- [1] M. F. Lara Díaz, J. C. Beltán Rojas, S. R. Rodríguez Montoya, and S. M. Araque Jaramillo, “Análisis de la percepción de eventos estáticos y dinámicos en personas con enfermedad de Alzheimer,” *Univ. Psychol.*, vol. 15, no. 5, 2016, doi: 10.11144/Javeriana.upsy15-5.apee.
- [2] L. M. Ariza-Serrano, J. Guerrero-Vega, P. Ortiz, and C. L. Moreno-Lopez, “Caracterización de pacientes con enfermedad de Parkinson en un centro de referencia de la ciudad de Bogotá , Colombia Characterization of patients with Parkinson ’ s disease in a reference center of Bogotá , Colombia,” *Acta Neurológica Colomb.*, vol. 32, no. 3, pp. 203–208, 2016.
- [3] Albert, M. *et al.*, “The diagnosis of mild cognitive impairment due to Alzheimer’s disease: recommendations from the National Institute on Aging- Alzheimer’s Association workgroups on diagnostic guidelines for Alzheimer’s disease,” *Alzheimers Dement*, vol. 7, no. 3, pp. 270–279, 2011, doi: 10.1016/j.jalz.2011.03.008.
- [4] J. S. Agustin, *Off-the-Shelf Gaze Interaction*. IT University of Copenhagen, 2010.
- [5] E. Abdulin, L. Friedman, and O. Komogortsev, “Custom Video-Oculography Device and Its Application to Fourth Purkinje Image Detection during Saccades,” no. April, 2019.
- [6] H. Singh and J. Singh, “Human Eye Tracking and Related Issues: A Review,” *Int. J. Sci. Res. Publ.*, vol. 2, no. 1, pp. 2250–3153, 2012, [Online]. Available: www.ijsrp.org.
- [7] R. G. Bozomitu, A. Păsărică, V. Cehan, C. Rotariu, and C. Barabaşa, “Pupil centre coordinates detection using the circular Hough transform technique,” in *38th International Spring Seminar on Electronics Technology*, 2015, pp. 462–465, doi: 10.1109/ISSE.2015.7248041.
- [8] A. B. Roig, M. Morales, J. Espinosa, J. Perez, D. Mas, and C. Illueca, “Pupil detection and tracking for analysis of fixational eye micromovements,” *Optik*, vol. 123, no. 1. pp. 11–15, 2012, doi: 10.1016/j.ijleo.2010.10.049.
- [9] J. Daugman, “How Iris Recognition Works,” *IEEE Trans. CIRCUITS Syst. VIDEO*

- Technol.*, vol. 14, no. 1, pp. 715–739, 2004, doi: 10.1016/B978-0-12-374457-9.00025-1.
- [10] L. Bai, L. Shen, and Y. Wang, “A Novel Eye Location Algorithm based on Radial Symmetry Transform,” in *18th International Conference on Pattern Recognition (ICPR '06)*, 2006, vol. 3, pp. 511–514, doi: 10.1109/ICPR.2006.136.
- [11] “Dataturks – Online tool to build Image Bounding Box, NER, NLP and other ML datasets.” <https://dataturks.com/> (accessed Nov. 17, 2020).
- [12] “Labelbox: The best way to create and manage training data.” <https://www.labelbox.com/> (accessed Nov. 17, 2020).
- [13] “HELLA Aglaia Mobile Vision GmbH.” <https://hella-aglaia.com/products/annostation-labeling-system> (accessed Nov. 17, 2020).
- [14] “microsoft/VoTT: Visual Object Tagging Tool: An electron app for building end to end Object Detection Models from Images and Videos.” <https://github.com/Microsoft/VoTT/>. (accessed Nov. 17, 2020).
- [15] T. Winslow, “Intraocular Melanoma Of The Iris.” <http://www.harvard-wm.org/the-body-of-human-eyes/intraocular-melanoma-of-theiris/> (accessed Dec. 10, 2020).
- [16] D. Vaughan, T. Asbury, P. Riordan-Eva, and J. M. Jane, *Oftalmología general*. 1997.
- [17] H. Hermann and J. F. Cier, *Tratado de fisiología*, Decimoprim. Elsevier, 2012.
- [18] J. L. Andreassi, *Psychophysiology: Human Behavior and Physiological Response*, 5th ed., vol. 94, no. 2. 2007.
- [19] F. E. León-Sarmiento, D. G. Prada, and C. Gutiérrez, “Pupila, pupilometría y pupilografía,” *Acta Neurológica Colomb.*, vol. 24, pp. 188–197, 2008.
- [20] Tobii Pro, “Eye tracking software and devices - Products by Tobii Pro.” <https://www.tobii.com/product-listing/> (accessed Dec. 19, 2020).
- [21] I. Khanfir Kallel, D. Sellami Masmoudi, and N. Derbel, “Fast pupil location for better iris detection,” 2009, doi: 10.1109/SSD.2009.4956750.
- [22] H. Proença and L. A. Alexandre, “Iris segmentation methodology for non-cooperative recognition,” *IEE Proc. - Vision, Image, Signal Process*, vol. 153, no. 2, p. 199, 2006.

- [23] R. P. Wildes, "Iris recognition: An emerging biometric technology," *Proc. IEEE*, vol. 85, no. 9, pp. 1348–1363, 1997.
- [24] Y. Liu, S. Yuan, X. Zhu, and Q. Cui, "A practical iris acquisition system and a fast edges locating algorithm in iris recognition," *Proc. 20th IEEE Instrum. Technol. Conf.*, vol. 1, pp. 166–168, 2003.
- [25] X. He and P. Shi, "A Novel Iris Segmentation Method for Hand-Held Capture Device." Springer Berlin Heidelberg, pp. 479–485, 2005.
- [26] Q. C. Tian, Q. Pan, Y. M. Cheng, and Q. X. Gao, "Fast algorithm and application of hough transform in iris segmentation," *Proc. 2004 Int. Conf. Mach. Learn. Cybern.*, vol. 7, no. August, pp. 3977–3980, 2004, doi: 10.1109/icmlc.2004.1384533.
- [27] C. Sun, C. Zhou, and Y. Liang, "Study and Improvement of Iris Location Algorithm." Springer, Berlin, Heidelberg, pp. 436–442, 2005.
- [28] N. Ritter, R. Owens, J. Cooper, and P. P. Van Saarloos, "Location of the pupil-iris border in slit-lamp images of the cornea," in *Proceedings 10th International Conference on Image Analysis and Processing*, 1999, pp. 740–745.
- [29] D. R. Iskander, S. Mioschek, M. Trunk, and W. Werth, "Detecting eyes in digital images," *Proc. - 7th Int. Symp. Signal Process. Its Appl.*, vol. 2, pp. 21–24, 2003.
- [30] D. Zhu, S. T. Moore, and T. Raphan, "Robust pupil center detection using a curvature algorithm," *Comput. Methods Programs Biomed.*, vol. 59, no. 3, pp. 145–157, 1999.
- [31] "CASIA Iris Image Database, Center of Biometrics and Security Research – Institute of Automation, Chinese Academy of Sciences (CASIA)." <http://www.cbsr.ia.ac.cn/english/IrisDatabase.asp>.
- [32] G. Loy and A. Zelinsky, "Fast radial symmetry for detecting points of interest," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 25, no. 8, pp. 959–973, 2003, doi: 10.1109/TPAMI.2003.1217601.
- [33] W. Wei, Y. Zhou, J. Zhou, and Z. Zhang, "Pupil Center Location Based on Skin Color Segmentation and Radial Symmetry Transform," 2015.

- [34] W. Zhang, B. Li, X. Ye, Z. Zhuang, and K. Wang, "A Robust Iris Localization Algorithm Via Radial Symmetry for Nonideal Capturing Condition," *Int. J. Pattern Recognit. Artif. Intell.*, vol. 23, no. 1, pp. 45–57, 2009.
- [35] B. Yan, J. Li, S. Liu, and H. Yuan, "A robust algorithm for pupil center detection," in *2011 6th IEEE Conference on Industrial Electronics and Applications, ICIEA 2011*, 2011, pp. 413–417, doi: 10.1109/ICIEA.2011.5975619.
- [36] G. Loy and A. Zelinsky, "A Fast Radial Symmetry Operator for Detecting Points of Interest in Images."
- [37] A. Fitzgibbon, M. Pilu, and R. B. Fisher, "Direct least square fitting of ellipses," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 21, no. 5, pp. 476–480, 1999, doi: 10.1109/34.765658.
- [38] X. Yuan and P. Shi, "A Robust Coarse-to-Fine Method for Pupil Localization in Non-ideal Eye Images," 2007.
- [39] M. Soltany, S. Zadeh, and H. Pourreza, "Fast and accurate pupil positioning algorithm using circular Hough transform and gray projection," *Int. Conf. ...*, vol. 5, pp. 556–561, 2011, [Online]. Available: <http://ipcsit.net/vol5/102-ICCCM2011-C094.pdf>.
- [40] J. J. Moré, "The Levenberg-Marquardt algorithm: Implementation and theory." Springer, Berlin, Heidelberg, pp. 105–116, 1978.
- [41] R. C. Gonzalez and R. E. Woods, *Digital image processing*, Third Edit. New Jersey, USA: Pearson Prentice-Hall, 2016.
- [42] R. P. Singh and M. Dixit, "Histogram Equalization: A Strong Technique for Image Enhancement," *Int. J. Signal Process. Image Process. Pattern Recognit*, vol. 8, no. 8, pp. 345–352.

GLOSARIO DE TÉRMINOS:

EOG: Electroculografía

fps: cuadros por segundo (*frames per second*)

IR: infrarrojo, infrarroja

Píxel: Elemento de imagen (*Picture- Element*).

TSR: Transformada de Simetría Radial

VOG: Video-oculografía