



**DEPARTAMENTO DE CIENCIAS DE LA ENERGÍA Y
MECÁNICA**

CARRERA DE INGENIERÍA MECATRÓNICA

**TRABAJO DE TITULACIÓN PREVIO A LA OBTENCIÓN DEL
TÍTULO DE INGENIERO MECATRÓNICO**

**TEMA: DISEÑO Y CONSTRUCCIÓN DE UN DISPOSITIVO
LECTOR DE TEXTOS PARA PERSONAS CON DEFICIENCIA
VISUAL**

GRANDA PAZ, JUAN PABLO

ORTEGA GUTIÉRREZ, ALEX ALBERTO

DIRECTORA: ING. TOBAR, JOHANNA

SANGOLQUÍ

2017



DEPARTAMENTO DE CIENCIAS DE LA ENERGÍA Y MECÁNICA

CARRERA DE INGENIERÍA MECATRÓNICA

CERTIFICACIÓN

Certifico que el trabajo de titulación, ***“DISEÑO Y CONSTRUCCIÓN DE UN DISPOSITIVO LECTOR DE TEXTOS PARA PERSONAS CON DEFICIENCIA VISUAL”*** realizado por los señores ***GRANDA PAZ JUAN PABLO*** y ***ORTEGA GUTIÉRREZ ALEX ALBERTO***, ha sido revisado en su totalidad y analizado por el software anti-plagio, el mismo cumple con los requisitos teóricos, científicos, técnicos, metodológicos y legales establecidos por la Universidad de Fuerzas Armadas ESPE, por lo tanto me permito acreditarlo y autorizar a los señores ***GRANDA PAZ JUAN PABLO*** y ***ORTEGA GUTIÉRREZ ALEX ALBERTO*** para que lo sustente públicamente.

Sangolquí, 7 de marzo del 2017

ING. JOHANNA BELEN TOBAR QUEVEDO
DIRECTORA



DEPARTAMENTO DE CIENCIAS DE LA ENERGÍA Y MECÁNICA

CARRERA DE INGENIERÍA MECATRÓNICA

AUTORÍA DE RESPONSABILIDAD

Los señores, **GRANDA PAZ JUAN PABLO** con CI 1723038533 y **ORTEGA GUTIÉRREZ ALEX ALBERTO** con CI 0401492244, declaramos que éste trabajo de titulación: "**DISEÑO Y CONSTRUCCIÓN DE UN DISPOSITIVO LECTOR DE TEXTOS PARA PERSONAS CON DEFICIENCIA VISUAL**" ha sido desarrollado considerando los métodos de investigación existentes, así como también se han respetado los derechos intelectuales de terceros considerándose en las citas bibliográficas.

Consecuentemente declaramos que este trabajo es de nuestra autoría, en virtud de ello nos declaramos responsables del contenido, veracidad y alcance de la investigación mencionada.

Sangolquí, 7 de marzo del 2017

JUAN PABLO GRANDA PAZ
CI 1723038533

ALEX ALBERTO ORTEGA GUTIERREZ
CI 0401492244



DEPARTAMENTO DE CIENCIAS DE LA ENERGÍA Y MECÁNICA

CARRERA DE INGENIERÍA MECATRÓNICA

AUTORIZACIÓN

Los señores, **GRANDA PAZ JUAN PABLO** con CI 1723038533 y **ORTEGA GUTIÉRREZ ALEX ALBERTO** con CI 0401492244, autorizamos a la Universidad de las Fuerzas Armadas - ESPE publicar en la Biblioteca Virtual de la institución el presente trabajo de titulación "**DISEÑO Y CONSTRUCCIÓN DE UN DISPOSITIVO LECTOR DE TEXTOS PARA PERSONAS CON DEFICIENCIA VISUAL**" cuyo contenido, ideas y criterios son de nuestra autoría y responsabilidad.

Sangolquí, 7 de marzo del 2017

JUAN PABLO GRANDA PAZ
CI 1723038533

ALEX ALBERTO ORTEGA GUTIERREZ
CI 0401492244

DEDICATORIA

A mis padres por ser mi apoyo incondicional en todo momento.

A Daysi por ser mi luz y compañera de vida.

A Camila por inspirarme a ser su mejor ejemplo.

Juan Pablo

A mis padres por confiar en mis capacidades, porque gracias a su esfuerzo he logrado convertirme en profesional.

A mis amigos por transformarse en mi familia, por compartir conmigo esta etapa de la vida.

Alex

AGRADECIMIENTO

A la Ing. Johanna Tobar y al Ing. David Loza por invertir su tiempo en nuestro aprendizaje, por ser portavoces de cada maestro que aportó en nuestra formación académica con sus conocimientos y experiencia.

Al Dr. José Benavides y al Sgos. Edison Yamberla por su orientación y apoyo durante la realización de este proyecto.

A quien lee este apartado, por permitir a nuestra investigación incurrir dentro de su repertorio de información mental.

Juan Pablo y Alex

ÍNDICE

CARATULA	i
CERTIFICACIÓN	ii
AUTORÍA DE RESPONSABILIDAD	iii
AUTORIZACIÓN	iv
DEDICATORIA	v
AGRADECIMIENTO	vi
ÍNDICE	vii
ÍNDICE DE TABLAS	ix
ÍNDICE DE FIGURAS	x
RESUMEN	xii
ABSTRACT	xiii
CAPÍTULO 1	1
INTRODUCCIÓN	1
1.1 Planteamiento del problema	1
1.2 Antecedentes	2
1.3 Justificación e importancia	4
1.4 Objetivos	5
1.4.1 <i>Objetivo General</i>	5
1.4.2 <i>Objetivos Específicos</i>	5
1.5 Alcance	6
1.6 Estructura del documento	6
CAPÍTULO 2	7
MARCO TEÓRICO	7
2.1 Concepto de deficiencia	7
2.1.1 <i>Deficiencia visual</i>	8
2.1.2 <i>Tipos de deficiencia visual</i>	11
2.1.3 <i>Personas con deficiencia visual en Ecuador</i>	12
2.2 Visión artificial para el procesamiento de imágenes	14
2.2.1 <i>Cámaras para visión artificial</i>	14
2.2.2 <i>Librerías para el procesamiento de imágenes</i>	16
2.2.3 <i>Tratamiento de imágenes</i>	17
2.3 Reconocimiento óptico de caracteres	23

	viii
2.3.1 <i>Alternativas del OCR</i>	24
2.3.2 <i>Algoritmos para el desarrollo de un OCR</i>	25
2.3.3 <i>Aplicaciones del OCR</i>	26
2.4 <i>Conversión TTS</i>	27
2.4.1 <i>Alternativas de TTS</i>	27
2.4.2 <i>Algoritmo de un TTS</i>	28
CAPÍTULO 3	30
DISEÑO E IMPLEMENTACIÓN	30
3.1 <i>Identificación de necesidades</i>	30
3.2 <i>Definición de las especificaciones</i>	33
3.3 <i>Definición de la arquitectura y diseño de los módulos</i>	33
3.3.1 <i>Módulo cámara</i>	34
3.3.2 <i>Soporte</i>	35
3.3.3 <i>Placa electrónica</i>	41
3.3.4 <i>Iluminación</i>	44
3.3.5 <i>Control de iluminación</i>	48
3.3.6 <i>Retroalimentación</i>	49
3.3.7 <i>Controlador</i>	53
3.3.8 <i>Proceso genérico</i>	54
3.3.9 <i>Concepto final</i>	59
CAPÍTULO 4	61
IMPLEMENTACIÓN Y PRUEBAS	61
4.1 <i>Implementación</i>	61
4.2 <i>Pruebas</i>	64
4.2.1 <i>Dimensionales</i>	64
4.2.2 <i>Ensamble y Conexiones</i>	64
4.2.3 <i>Funcionales</i>	67
4.2.4 <i>Computacionales</i>	75
CAPÍTULO 5	78
CONCLUSIONES Y RECOMENDACIONES	78
5.1 <i>Conclusiones</i>	78
5.2 <i>Recomendaciones</i>	79
BIBLIOGRAFÍA	81
ANEXOS	90

ÍNDICE DE TABLAS

Tabla 1 Tipos de deficiencia visual según la agudeza visual	11
Tabla 2 Identificación de necesidades	31
Tabla 3 Qué´s y Cómo´s para matriz QFD	32
Tabla 4 Importancia relativa de las necesidades	32
Tabla 5 Características Microsoft LifeCam HD 3000	35
Tabla 6 Criterios a evaluar para el case.	36
Tabla 7 Matriz comparativa entre criterios del case.....	37
Tabla 8 Matriz criterio A Vs alternativas diseño del case.....	37
Tabla 9 Matriz criterio B Vs alternativas diseño del case.....	37
Tabla 10 Matriz criterio C Vs alternativas diseño del case.....	38
Tabla 11 Matriz criterio D Vs alternativas diseño del case.....	38
Tabla 12 Matriz resultante diseño del case.....	38
Tabla 13 Propiedades mecánicas del ABS y PLA.	40
Tabla 14 Características Placa Electrónica	41
Tabla 15 Características LED SMD 3030.	44
Tabla 16 Variación de la foto-resistencia	48
Tabla 17 Características del motor vibrador de disco	49
Tabla 18 Elementos Concepto final	59
Tabla 19 Resultados prueba dimensional	64
Tabla 20 Resultados prueba de conexiones.....	66
Tabla 21 Matriz para la prueba de adquisición de imágenes.....	67
Tabla 22 Matriz para la prueba del control de iluminación.....	69
Tabla 23 Matriz para la prueba del control de iluminación.....	70
Tabla 24 Matriz para la prueba del control de retroalimentación	71
Tabla 25 Indicadores de las señales de retroalimentación	72
Tabla 26 Matriz para la prueba de reconocimiento de caracteres	73
Tabla 27 Matriz para la prueba de tiempo de locución	74
Tabla 28 Tiempo de locución de líneas de texto.....	75
Tabla 29 Indicadores de las señales de retroalimentación	76
Tabla 30 Recursos del ordenador.....	76

ÍNDICE DE FIGURAS

Figura 1 Dispositivos locales para personas con discapacidades	2
Figura 2 Dispositivos extranjeros para personas con deficiencia visual	3
Figura 3 Jóvenes con deficiencia visual.....	8
Figura 4 Test de Snellen para probar agudeza visual.....	9
Figura 5 Campo visual normal	10
Figura 6 Porcentaje de personas según su nivel de discapacidad	13
Figura 7 Porcentaje de personas según su tipo de discapacidad.....	13
Figura 8 Tipos de cámaras para visión artificial	14
Figura 9 Búsqueda de contornos.....	19
Figura 10 Imagen y su histograma.....	19
Figura 11 Imagen normalizada	20
Figura 12 Diferentes tipos de umbralización.....	21
Figura 13 Diferentes tipos de umbralización.....	22
Figura 14 Transformación morfológica de cierre.....	22
Figura 15 Salida de un texto de caracteres latinos sin tratar	23
Figura 16 Esquema de trabajo de un OCR.....	26
Figura 17 Esquema de reconocimiento de caracteres.....	26
Figura 18 Diagrama de bloques de un TTS	28
Figura 19 Arquitectura del dispositivo	34
Figura 20 Opciones de diseño 3D.....	36
Figura 21 Posición ergonómica uso del computador	39
Figura 22 Estudio de caída de Von Mises	40
Figura 23 Gráfica seleccionar sección de la pista.....	42
Figura 24 Gráfica seleccionar sección de la pista.....	42
Figura 25 Diseño PCB	43
Figura 26 Circuito de corte y saturación	45
Figura 27 Rama hacia el colector del transistor.....	45
Figura 28 Rama hacia la base del transistor.....	47
Figura 29 Divisor de tensión	49
Figura 30 Rama hacia el colector del transistor.....	51
Figura 31 Diagrama de flujo del controlador	54

Figura 32 Proceso genérico del dispositivo	54
Figura 33 Máscaras	56
Figura 34 Resultado de la aplicación de filtros.....	56
Figura 35 Base de datos de dedos en contraste con un papel	57
Figura 36 Región de interés.....	57
Figura 37 Línea de texto	58
Figura 38 Conversión de imagen a cadena de caracteres.....	58
Figura 39 Vista explosionada del concepto final	60
Figura 40 Piezas impresas en 3D	61
Figura 41 Piezas impresas en 3D	61
Figura 42 Elaboración PCB	62
Figura 43 Montaje de elementos al case	62
Figura 44 Montaje elementos al case frontal	63
Figura 45 Dispositivo ensamblado	63
Figura 46 Acople del anillo y el case posterior	65
Figura 47 Vistas del acople entre case frontal y posterior	65
Figura 48 Acople de los vibradores al anillo	66
Figura 49 Captura de texto en el fotograma	68
Figura 50 Captura de imagen con lente wide angle.....	68
Figura 51 Ciclos de trabajo para las secuencias de vibración	71
Figura 52 Reconocimiento erróneo de caracteres	73
Figura 53 Uso del CPU de la aplicación	77
Figura 54 Uso de memoria RAM de la aplicación	77

RESUMEN

El proyecto trata sobre el diseño y construcción de un dispositivo lector de textos para personas con deficiencia visual. Se creó un dispositivo en forma de anillo que se adapta al dedo índice del usuario, está provisto de una cámara web matricial que captura imágenes de las cuales se obtiene texto y se lo reproduce en forma de audio. El dispositivo tiene un peso de 100 gramos, y dimensiones de 6,5x4,5x6 centímetros, lee textos en español con fuente de texto arial de tamaño mínimo 14 puntos e interlineado mínimo de 1,5 espacios. La arquitectura modular de tipo ranura permite alojar los elementos funcionales del dispositivo en un soporte de material plástico y flexible diseñado en un software CAD bajo normas que proporcionan ergonomía y comodidad al usuario. El algoritmo de control fue desarrollado en el entorno libre de programación Python utilizando la librería para visión artificial OpenCV, Tesseract como motor del reconocimiento óptico de caracteres y Espeak como TTS ligado a Mbrola para la reproducción de audio. Los resultados obtenidos en las pruebas de funcionamiento confirman la eficiencia del dispositivo bajo ciertas condiciones de uso.

PALABRAS CLAVE:

- **DEFICIENCIA VISUAL**
- **RECONOCIMIENTO ÓPTICO DE CARACTERES**
- **REDES NEURONALES**

ABSTRACT

The project is about the design and construction of a text-reading device for visually impaired people. A ring-shaped device that fits the index finger of the user was created, it is provided by a matrix camera that captures images from which text is extracted and reproduced as audio. The device has a weight of 100 grams, and dimensions of 6,5x4,5x6 centimeters, it reads texts in Spanish with arial text font of 14 points minimum size high and 1.5 spaces minimum line. The slot-type modular architecture allows to insert the functional elements of the device in a support of plastic and flexible material designed in CAD software under standards that provide ergonomics and comfort to the user. The control algorithm was developed in the Python free programming environment using the OpenCV artificial vision library, Tesseract as optical character recognition engine and Espeak as TTS linked to Mbrola for audio reproduction. The results obtained in the performance tests confirm the efficiency of the device under certain conditions of use.

KEYWORDS:

- **VISUAL IMPAIRMENT**
- **OPTICAL CHARACTER RECOGNITION**
- **NEURONAL NETWORKS**

CAPÍTULO 1

INTRODUCCIÓN

El capítulo contiene el planteamiento del problema, los antecedentes, los objetivos, la justificación e importancia y el alcance del dispositivo lector de textos. Se realiza una descripción de dispositivos similares desarrollados para personas con discapacidades.

1.1 Planteamiento del problema

Las personas con deficiencia visual tienen limitación para acceder a la información escrita. Los audiolibros y el sistema en braille son alternativas para solventar los obstáculos, pero poseen la desventaja de que deben ser previamente elaborados para un texto en específico.

El proyecto tiene como finalidad aportar en la vida cotidiana de las personas con deficiencia visual. La meta final consiste en la obtención de un dispositivo en forma de anillo que se adapte al dedo índice de la persona, el mismo que estará provisto de una cámara capaz de captar texto impreso, procesarlo y reproducirlo en forma de audio mediante un sintetizador de palabras.

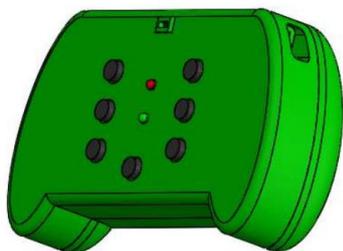
El usuario deberá pasar el dedo que porta el anillo por el texto deseado, el algoritmo rastreará el movimiento del dedo, identificará las palabras y procesará la información, al mismo tiempo emitirá señales vibratorias advirtiendo al usuario cuando se aparta de la línea de texto, al inicio o al final de cada una de éstas.

Para su consecución, el enfoque principal se realizará en la ciencia de la visión artificial la cual ayudará a desarrollar la base teórica y algorítmica mediante la que se extraerá y analizará la información de una secuencia de imágenes.

1.2 Antecedentes

En los últimos años la tecnología ha buscado crear dispositivos que permitan a los discapacitados tener una vida independiente y disfrutar de actividades cotidianas sin limitantes.

A nivel nacional se han desarrollado algunas herramientas y dispositivos para mejorar la calidad de vida de personas con discapacidades. Por ejemplo Garzón Jeréz (2015) diseña un dispositivo para personas sordociegos que permite aumentar el nivel de comunicación sin necesidad de conocer el sistema braille. Mientras que, Loayza Jaramillo (2016) presenta un sistema electrónico de detección de obstáculos que proporciona información al usuario mediante servomotores (ver figura 1).



a) Dispositivo para comunicación Braille para sordo ciegos



b) Dispositivo para detección de obstáculos para no videntes

Figura 1 Dispositivos locales para personas con discapacidades

Fuente: (Garzón Jeréz, 2015), (Loayza Jaramillo, 2016)

Entre los dispositivos y aplicaciones que se han desarrollado en el extranjero se encuentra el lector de pantalla (ver figura 2.a) que se opera desde el teclado de la computadora y el programa verbalizará la información visible en el monitor, de modo que una voz sintética se escuchará por los altavoces de la computadora (Alsina, 2012).

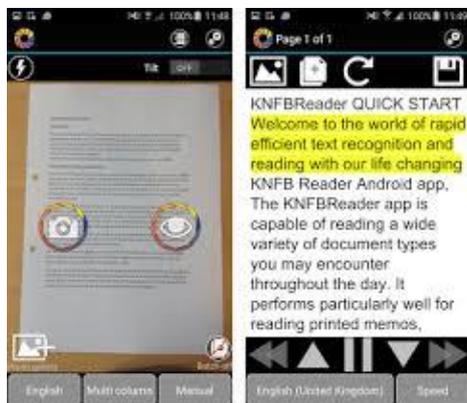
Al 2013 se desarrolla Eytalk (ver figura 2.b), unas gafas que permiten a los usuarios ciegos escuchar lo que está escrito, por ejemplo, en productos de un supermercado, libros o vallas publicitarias (Sénderov, 2013).



a) Lector de pantalla



b) Eytalk



c) KNF Reader



d) FingerReader

Figura 2 Dispositivos extranjeros de lectura para personas con deficiencia visual

Fuente: (Cetpro, 2012), (Chw Net, 2013), (Apkmonk, 2014), (Shilkrot, 2014)

De las aplicaciones desarrolladas para smartphones se encuentra KNFB Reader (ver figura 2.c), con una foto del texto impreso, la aplicación lee en voz alta. Se puede configurar para leer distintos tipos de documento y es capaz de orientar al usuario mediante comandos de voz y vibraciones para que sea capaz de colocar la cámara en la posición correcta al capturar el documento (Barbuzano, 2015).

Investigadores del MIT Media Laboratory desarrollan FingerReader (ver figura 2.d), un anillo que utiliza un algoritmo creado especialmente para

reconocer las palabras, que pasan a un programa que las lee en voz alta (Zahumenszky, 2015).

1.3 Justificación e importancia

Los dispositivos existentes en el extranjero, que representan una ayuda en la lectura de textos a partir de una imagen, están realizados para idiomas diferentes al español y no corresponden al léxico y dialecto manejado en el país. El desarrollo del proyecto pretende ser un aporte a las personas con deficiencia visual, con el desarrollo de un dispositivo cómodo, pequeño y fácil de usar que permita leer textos. Se pretende mejorar sustancialmente la calidad de vida de los usuarios del dispositivo, a la vez que ayudará a que los invidentes tengan igualdad de oportunidades y sean más independientes. Hasta ahora, ningún otro dispositivo ha reunido la movilidad y el acceso rápido a textos.

En la Constitución de la República del Ecuador, en su sección sexta personas con discapacidad, artículo 47, numeral 7, el estado reconoce a las personas con discapacidad varios derechos y proclama que se garantizará su educación dentro de la educación regular y que los establecimientos educativos cumplirán normas de accesibilidad para personas con discapacidad (Asamblea Nacional, 2008).

El trabajo busca contribuir a futuras investigaciones en torno al desarrollo de tecnología que ayude a personas no videntes, planteando el uso y aplicación de hardware y software de libre acceso, lo que facilita la continuidad de la investigación y el desarrollo de mejores dispositivos.

1.4 Objetivos

1.4.1 Objetivo General

Diseñar, construir e implementar un dispositivo lector de textos para personas con deficiencia visual moderada.

1.4.2 Objetivos Específicos

- Realizar el modelado 3D del dispositivo y seleccionar los materiales adecuados para la estructura y el anillo, de manera que sirva de soporte y protección de los componentes internos y, otorgue ergonomía y comodidad al usuario.
- Desarrollar un algoritmo capaz de tratar imágenes adquiridas con una cámara web aplicando técnicas de visión artificial por computador.
- Implementar un sistema de iluminación y señales de retroalimentación controladas para conseguir imágenes legibles y guiar al usuario durante el proceso de lectura.
- Entrenar un OCR con un diccionario básico de palabras en español que permita el reconocimiento del texto presente en las imágenes.
- Integrar un conversor TTS que sintetice el texto reconocido por el algoritmo de control permitiendo la emisión de audio.

1.5 Alcance

El dispositivo lector de textos pretende brindar un apoyo en la lectura de las personas con deficiencia visual moderada, dotándolas de un dispositivo a modo de anillo para que su lectura no se limite a los textos en braille o audiolibros.

El texto para la lectura debe estar impreso en un formato de hoja A4 con márgenes normales de 2,5 centímetros superior e inferior y de 3 centímetros izquierdo y derecho. El tipo de fuente debe ser arial de 14 puntos o más e interlineado de 1,5 espacios. El lugar para realizar la lectura debe contar con luz natural o artificial para obtener buena funcionalidad del dispositivo.

1.6 Estructura del documento

Capítulo 1, se plantea las generalidades del proyecto. También se realiza un análisis de los dispositivos similares desarrollados de forma nacional e internacional.

Capítulo 2, se describe las discapacidades en el Ecuador, en específico, se aborda la deficiencia visual. También se realiza la base teórica de visión artificial para el OCR y conversión texto a audio.

Capítulo 3, se desarrolla la selección y dimensionamiento de los elementos del dispositivo. Se llevan a cabo las fases de diseño para llegar al concepto final.

Capítulo 4, se detalla el proceso de implementación y las pruebas de funcionamiento del dispositivo.

Capítulo 5, se plasman las conclusiones y recomendaciones del proceso de diseño, construcción y pruebas del proyecto.

CAPÍTULO 2

MARCO TEÓRICO

El capítulo refleja una descripción acerca de las deficiencias, en específico tratándose la deficiencia visual y su panorama en el Ecuador. También se construye la base teórica de visión artificial para el procesamiento de imágenes, reconocimiento óptico de caracteres y conversión de texto a audio.

2.1 Concepto de deficiencia

La Organización Mundial de la Salud (OMS) (citada en Zaragoza Baquero, 2012) sostiene que la deficiencia es “toda pérdida o anomalía de una estructura o función psicológica, fisiológica o anatómica. Puede ser temporal o permanente y en principio solo afecta al órgano” (pg. 35).

Diferentes autores definen que la deficiencia no podría considerarse como una enfermedad, sino más bien se puede identificar como la reducida capacidad que presenta un individuo para realizar una determinada acción de modo normal, considerándose como normal, la forma en que el resto de personas de su medio llevan a cabo dicha acción.

Según (Martínez Cuervo, 2011) las deficiencias se pueden clasificar de la siguiente manera:

- Deficiencias intelectuales
- Deficiencias del lenguaje
- Deficiencias de la audición
- Deficiencias de la visión
- Deficiencias músculo - esqueléticas

2.1.1 Deficiencia visual

Para Latorre, Bisseto, & Teruel (2012) la deficiencia visual abarca “una referencia a toda diversidad de situaciones viso perceptivas, que siguen un continuo, un gradiente, que avanza desde trastornos leves hasta la deficiencia total o ceguera” (pág. 16).

Se puede definir también a la deficiencia visual como la carencia “en la estructura o funcionamiento de los órganos visuales, cualquiera que sea la naturaleza o extensión de la misma que causa una limitación, que aún con la mejor corrección, interfiere en el desenvolvimiento de la persona” (Fernández I. , 2014, pág. 44). En la Figura 3 se muestra dos jóvenes tocando acordeón y güiro, de una forma muy normal y capaz a pesar de poseer deficiencia visual.



Figura 3 Jóvenes con deficiencia visual.

Fuente: (Fundación Las Dos Orillas, 2015)

Al ser complejo el sistema de visión humano, los grados de deficiencia visual pueden ser varios, para limitarlos se debe tomar en cuenta los parámetros de agudeza visual y campo visual.

a) Agudeza visual

Es la “habilidad para discriminar o diferenciar detalles entre dos estímulos visuales distintos a una determinada distancia, y para percibir la figura y forma de los objetos” (Fernández, Junquera, & Ribes, 2013, pág. 100).

Se puede determinar la agudeza visual mediante el test de Snellen (ver figura 4), un ojo normal tiene un valor de uno y para eso según (Núñez, 2011) se debe cumplir con las siguientes condiciones:

- El estado de refracción ocular debe ser de emetropía. Si existiera algún defecto de refracción (ametropía) deberá estar corregido por cualquier método posible.
- Las estructuras oculares que son atravesadas por la luz deben mantener la transparencia.
- La mácula (retina central) y la vía óptica, así como el área 17 del córtex, tienen que estar en condiciones de normalidad.

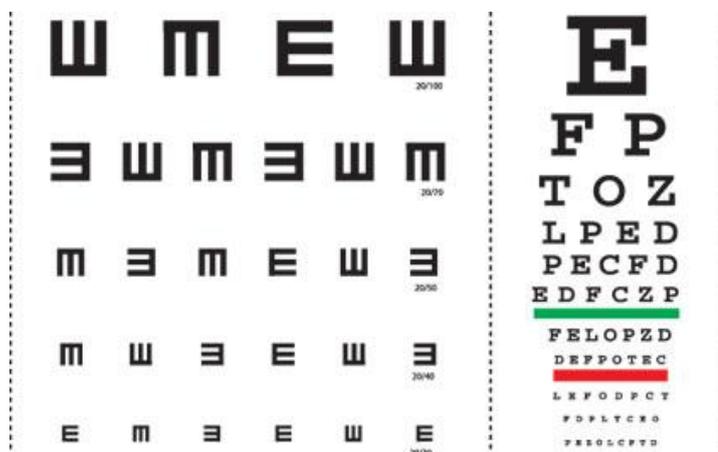


Figura 4 Test de Snellen para probar agudeza visual

Fuente: (CRC_online, 2016)

b) Campo Visual

Es el “espacio que rodea a la persona, el grado de mayor excentricidad que puede abarcar el ojo humano en cualquier dirección” (Fernández, Junquera, & Ribes, 2013, pág. 100).

En la figura 5 se presenta el campo de visión normal de una persona (Núñez, 2011), que posee los siguientes límites:

- Parte externa o temporal: 100°
- Parte interna o nasal: 60°
- Parte superior: 50°
- Parte inferior: 70°

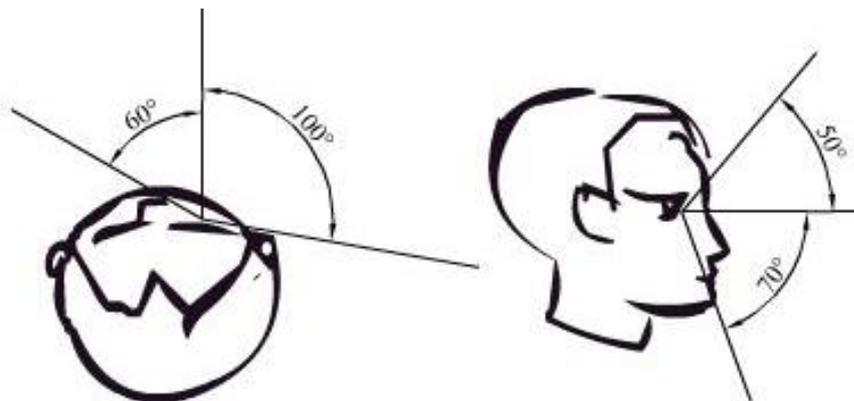


Figura 5 Campo visual normal

Fuente: (Reyes, 2016)

Para tener un campo visual dentro de los parámetros mencionados se debe poseer ciertas características según (Núñez, 2011):

- La transparencia de córnea, cristalino y vítreo.
- La integridad de la retina tanto central (macular) como periférica.
- El ojo en óptimas condiciones refractivas, para valorar el campo visual central.

2.1.2 Tipos de deficiencia visual

Existen varios términos que engloban el daño en el sistema de visión, como lo son: la ceguera, baja visión o deficiencia visual, la diferencia en todos estos casos es la afectación de la vista que puede ser ligera, moderada, grave, profunda o total.

Se puede clasificar las deficiencias visuales según el nivel de pérdida visual, según la agudeza visual y según el órgano afectado. Por el fin que persigue el proyecto de titulación se analizará los tipos de deficiencia visual según la agudeza visual que presenta el individuo.

En la tabla 1 se refleja el tipo de deficiencia visual hasta llegar a la ceguera total, en función del valor de agudeza visual definido con anterioridad.

Tabla 1

Tipos de deficiencia visual según la agudeza visual

Agudeza Visual	Estado
De 1 a 0,8	Visión normal
De 0,8 a 0,3	Deficiencia ligera
De 0,3 a 0,12	Deficiencia moderada
De 0,12 a 0,05	Deficiencia grave
De 0,05 a 0,02	Deficiencia profunda
Menor a 0,02	Ceguera

Fuente: (Castejón & Navas, 2013)

La sociedad suele pensar y encajonar a todas las personas con deficiencia visual en un solo matiz, cuando existen algunos trastornos visuales, la similitud de estos estados se la determina por dos factores: el grado de pérdida visual y el momento que se presentó la deficiencia.

De esta manera Castejón y Navas (2013) mencionan que “no es lo mismo percibir solo sombras que tener reducido el campo visual, no es lo mismo perder la vista gradualmente que bruscamente, no es igual ser invidente congénito que perder la vista en el transcurso de la vida” (pág. 260). El desarrollo y las posibilidades de aprendizaje son diferentes para cada individuo con deficiencia visual, dependiendo del grado de la misma y el momento en el que presentó la deficiencia

La ceguera es un estado en el cual las personas “no tienen visión en lo absoluto o no tienen la capacidad para distinguir entre la luz y la oscuridad, es aquella afectación que impide que la visión sea útil para adquirir algún conocimiento, pudiendo ser de nacimiento o adquirida” (Arteaga Jiménez & Elizalde Ruiz, 2011).

2.1.3 Personas con deficiencia visual en Ecuador

En Ecuador hay 418001 personas con discapacidad registradas (Conadis, 2016). Los tipos de discapacidad que considera el CONADIS (ver Figura 6) son:

- Auditiva
- Física
- Intelectual
- Lenguaje
- Psicosocial
- Visual

Por el grado de discapacidad se verifica en la figura 6 que la de tipo grave domina al resto con un porcentaje de 35,25%, siguiendo la moderada con 27,68%.

De las persona con discapacidad registradas por el Consejo Nacional para la Igualdad de Discapacidades (CONADIS) el 11,80% presentan discapacidad visual, como lo muestra la figura 7.

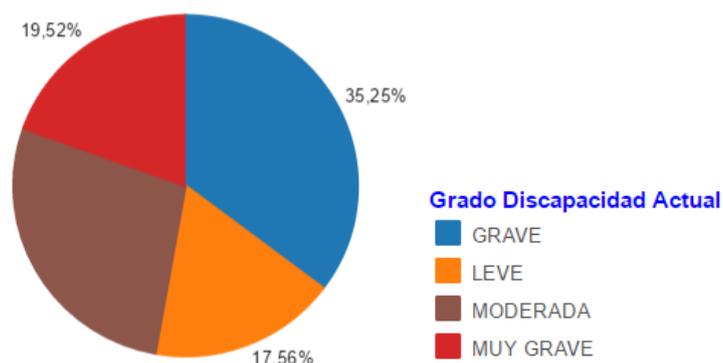


Figura 6 Porcentaje de personas según su nivel de discapacidad

Fuente: (Conadis, 2016)

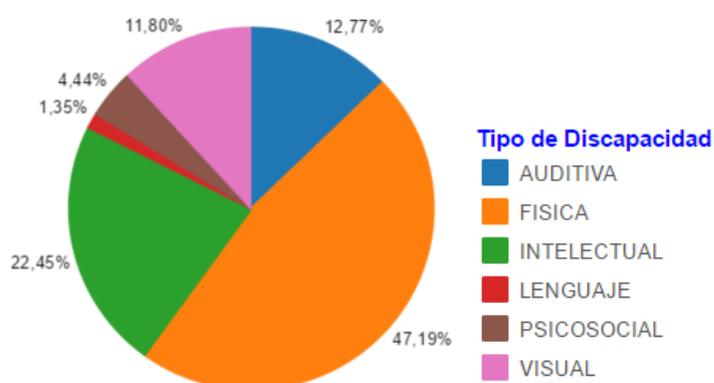


Figura 7 Porcentaje de personas según su tipo de discapacidad

Fuente: (Conadis, 2016)

La Asamblea según su página web “aprobó el Tratado de Marrakech para facilitar el acceso a las obras publicadas a personas ciegas, con discapacidad visual o con otras dificultades para acceder al texto impreso” (Asamblea Nacional, 2016). El acuerdo suscrito por la Asamblea, es un tratado internacional que busca medidas sobre la propiedad intelectual y derechos de autor del conocimiento impreso, para facilitar el libre acceso, modificación y distribución del conocimiento impreso en el sistema braille.

El avance tecnológico en el país es escaso dentro del tema de discapacidades, por esto es que con la aplicación de visión artificial y redes

neuronales se puede desarrollar un dispositivo que ayude en el proceso de lectura a las personas con discapacidad visual.

2.2 Visión artificial para el procesamiento de imágenes

La visión artificial es una herramienta para “el procesamiento, el análisis y la interpretación de imágenes, hace posible que un ordenador procese imágenes o fotografías bidimensionales” (Vicomtech, 2017). Se pretende conseguir a través de la visión artificial es obtener una imagen binaria y limpia de ruido para que en la siguiente fase se realice el reconocimiento óptico de caracteres.

2.2.1 Cámaras para visión artificial

“En los sistemas de visión artificial, la función de las cámaras de visión es capturar la imagen proyectada en el sensor, vía las ópticas, para poder transferirla a un sistema electrónico” (Infaimon, 2017).

En el mercado existen algunos tipos de cámaras para realizar sistemas de visión artificial. Se presentan las siguientes opciones como las más representativas (ver figura 8).

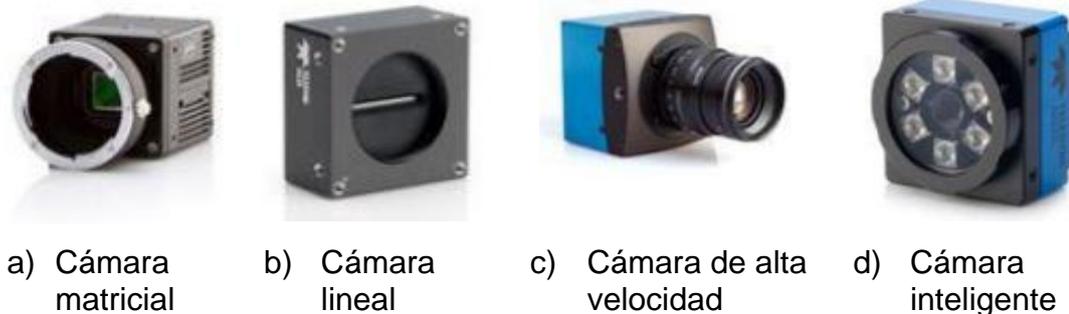


Figura 8 Tipos de cámaras para visión artificial

Fuente: (Infaimon, 2017)

a) Cámaras matriciales

“El sensor cubre un área que está formada por una matriz de píxeles. Los sensores de las cámaras modernas son todos de tecnología CMOS formados por miles de diodos fotosensibles posicionados de forma muy precisa en la matriz” (Centro Integrado Politécnico ETI, 2016).

b) Cámaras lineales

Se basa en la “construcción de una imagen línea a línea empleando un sensor de estructura lineal, esta imagen se forma mediante el movimiento del objeto con respecto a la cámara o bien mediante el movimiento de la cámara con respecto al objeto” (Infaimon, 2017).

c) Cámaras de alta velocidad

“Presentan sistemas de grabación de muy alta velocidad capaces de capturar desde 1.000 a 1.000.000 imágenes por segundo” (Infaimon, 2017).

d) Cámaras inteligentes

“Incorporan los elementos tradicionales de las cámaras convencionales pero además se acompañan de un procesador, memoria y sistema de comunicaciones con el exterior, componen en sí mismas un sistema completo de visión artificial” (Infaimon, 2017).

2.2.2 Librerías para el procesamiento de imágenes

a) Librería Open CV

OpenCV es “una librería diseñada para la eficiencia computacional y con un fuerte enfoque en aplicaciones en tiempo real con procesamiento multi-núcleo construida para proporcionar una infraestructura común para aplicaciones de visión por ordenador” (Cosenza, 2016).

b) Librería JavaVis

JavaVis es “una librería de uso libre para el desarrollo de algoritmos y métodos de visión artificial diseñada en Java en la cual se integran clases y técnicas básicas como definir el tipo de datos de la imagen, manejar lectura y escritura de ficheros o usar una interfaz gráfica; métodos de los cuales se puede hacer uso para desarrollar aplicaciones más complejas” (Cazorla & Viejo, 2015).

c) Image Processing Toolbox

Image Processing Toolbox es “un toolbox proporcionado por el software matemático Matlab que incluye un conjunto de funciones que permiten el desarrollo de aplicaciones con algoritmos para el proceso y análisis de imágenes las cuales son tratadas como matrices de datos” (García, 2009).

2.2.3 Tratamiento de imágenes

Para el tratamiento de imágenes se emplea una serie de técnicas y filtros, entre las herramientas de este tipo más útiles se describen las siguientes:

a) Técnicas

- Segmentación

La segmentación de una imagen implica la detección de los contornos o regiones de la imagen, basándose en la información de intensidad para poder reconocer cada uno de los caracteres. Se descompone el texto en diferentes entidades lógicas, suficientemente invariables, independientes del escritor, y significativas para su reconocimiento (Imbaquingo Esparza, 2017).

Con la segmentación se localizan las zonas de interés caracterizadas por atributos comunes tales como dimensión, superficie, densidad, inclinación, longitud del trazo, etc. Uno de los métodos usados consiste en la determinación de los agrupamientos a partir del histograma, de tal forma que permitan una clasificación o moralización de los píxeles en regiones homogéneas (Grupo de Investigación CATAM, 2013).

Otros métodos para la detección de bordes son:

- Detector de bordes de Sobel.
- Operador de Kirsch.
- Detector de bordes de Marr-Hildreth.
- Detector de bordes de Canny.

- Adelgazamiento

Adelgazar o simplificar la forma de cada componente conexo de la imagen consiste en ir borrando sucesivamente los puntos de los contornos de forma que se conserve su tipología (Imbaquingo Esparza, 2017).

Las condiciones que determinan si un punto se puede borrar se encuentran relacionadas con dos conceptos:

Punto simple es un “pixel negro P del borde de la imagen se considera simple si el conjunto de los vecinos en negro de P tienen exactamente una componente conexas que es adyacente a P” y punto final es “si tiene exactamente un vecino negro; un punto final no es más que un punto extremo de la imagen”. Es decir, un punto del borde de cada componente se puede eliminar si es simple y no es final. (Grupo de Investigación CATAM, 2013).

Para ello se debe seguir un esquema de barridos sucesivos en paralelo, es decir, señalar los píxeles borrables para eliminarlos todos a la vez de forma que la imagen no se deforme y continúe teniendo las mismas proporciones que la original.

- Búsqueda de contornos

En ocasiones es necesario desechar todos aquellos objetos que no llegan a un tamaño determinado (ruido). Esto se obtiene binarizando la imagen de manera que se pueda aplicar una función para buscar sus contornos. Un contorno es una curva que une todos los puntos continuos de igual color o intensidad.

Para tener mejores resultados, se binariza la imagen de manera que el objeto a ser encontrado sea blanco dentro de un fondo puramente negro (ver figura 9).

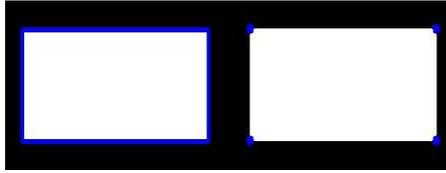


Figura 9 Búsqueda de contornos

Fuente: (OpenCV, 2017)

Para eliminar el ruido se calcula el área de todos los objetos detectados y se almacena en otra máscara únicamente aquéllos cuya área supera un valor determinado. Para ello se utiliza un condicional estándar que se obtiene a modo de prueba y error.

- Histograma de una imagen

El histograma de una imagen es un gráfico que muestra la distribución de la intensidad de color en una imagen. El rango de tono de color en el eje x y el número de píxeles correspondientes en el eje y. En la figura 10 se muestra una imagen y el histograma para la imagen en escala de grises. La región izquierda del histograma muestra la cantidad de píxeles más oscuros en la imagen y la región derecha la cantidad de píxeles más claros.

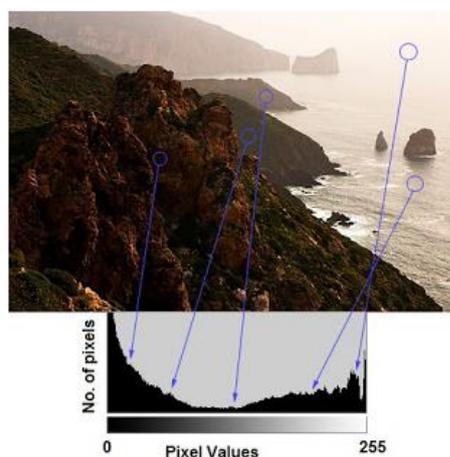


Figura 10 Imagen y su histograma

Fuente: (Cambridge in Colour, 2017)

b) Filtros

- Normalización RGB

La luz juega un papel crucial en el tratamiento de imágenes. Por ejemplo, si existe una fuente unidireccional de luz, ésta crea sombra y diferentes tonos de colores. Algunas partes están resaltadas y algunas sombreadas. Si se opta por detectar un color usando los valores RGB aproximados solo se lo detectará parcialmente. Para reducir los efectos de la luz, la normalización del espacio de color es de gran utilidad (ver figura 11).

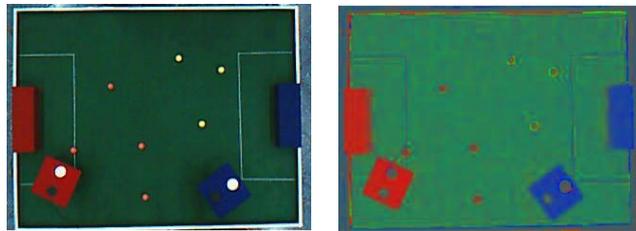


Figura 11 Imagen normalizada

Fuente: (Sinha, 2017)

- Espacio de color Gray Scale

El espacio de color Gray Scale (escala de grises), define un modo de codificar los colores de una imagen restringiéndola a contener únicamente colores entre negro, blanco y matices de gris.

- Umbralización

La umbralización constituye uno de los métodos de segmentación más simples que existe, si el valor de píxel es mayor que un valor umbral, se le asigna un valor (puede ser blanco), de lo contrario se le asigna otro valor (puede ser negro). Este método permite conseguir binarizar directamente una imagen de modo que se puedan diferenciar los objetos de la imagen del fondo (ver figura 12).

La umbralización directa no siempre es buena cuando la imagen está sometida a diferentes condiciones de luz en diferentes áreas.

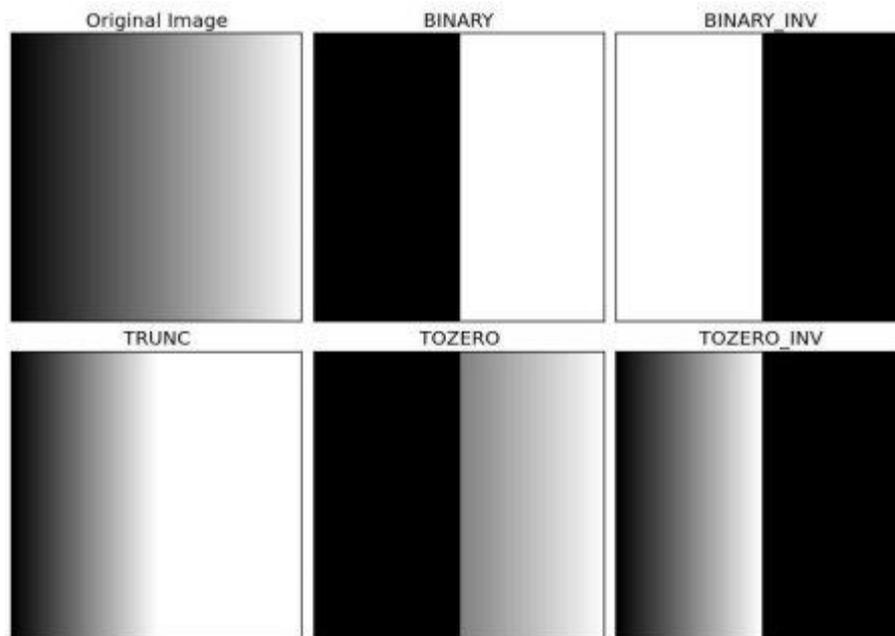


Figura 12 Diferentes tipos de umbralización

Fuente: (OpenCV, 2017)

- Filtro gaussiano

Los filtros de umbralización adaptativa calculan el valor de umbral para una pequeña región de la imagen, así se obtienen diferentes umbrales para diferentes regiones de la misma imagen obteniendo mejores resultados para imágenes con iluminación variable. El método gaussiano calcula el valor de umbral a partir de la suma ponderada de valores de intensidad de píxeles vecinos.

- Binarización Otsu

La imagen que se obtiene después de aplicar un filtro gaussiano es más nítida, sin embargo requiere todavía de eliminación de ruido. Si se considera a la región de interés como una imagen bimodal (el histograma de la imagen tiene dos picos) se puede tomar un valor en el centro de esos dos

picos como valor umbral. Así es como binarización Otsu calcula este valor automáticamente (ver figura 13).

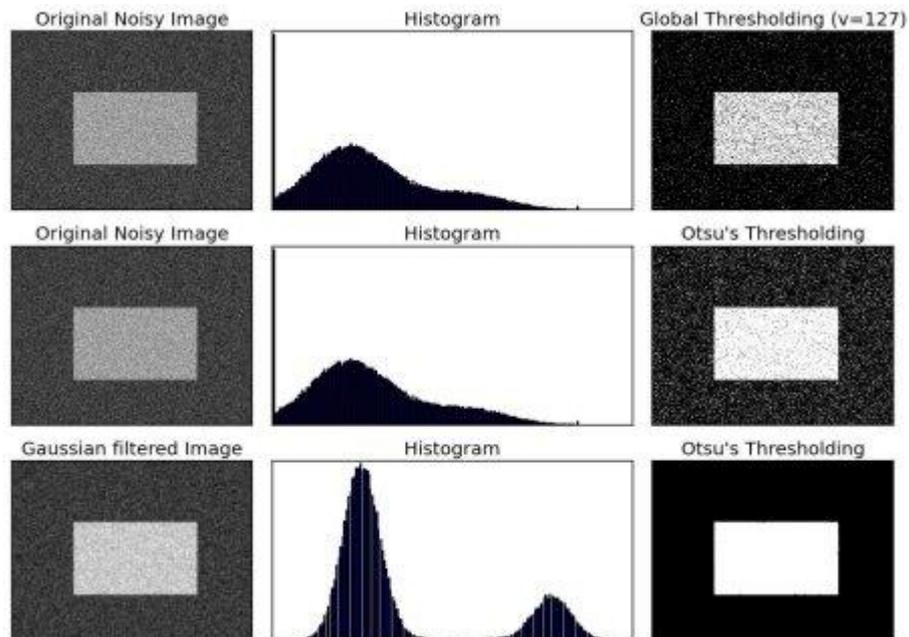


Figura 13 Diferentes tipos de umbralización

Fuente: (OpenCV, 2017)

- Morfología

Una transformación morfológica es una operación simple basada en la forma de la imagen. Necesita dos entradas, una es la imagen original, la segunda se llama elemento estructurante, quien decide la naturaleza de la operación. Una operación morfológica de cierre aplica una operación de dilatación seguida por una de erosión (ver figura 14).



Figura 14 Transformación morfológica de cierre

Fuente: (OpenCV, 2017)

2.3 Reconocimiento óptico de caracteres

El reconocimiento óptico de caracteres (OCR por sus siglas en inglés), es el proceso enfocado en la digitalización de textos, los cuales se identifican automáticamente a partir de imágenes patrón de símbolos o caracteres que pertenecen a un determinado alfabeto (ver figura 15), para luego almacenar sus propiedades de identificación en forma de datos (Hidalgo & Sánchez, 2015).

<p>The (quick) [brown] {fox} jumps! Over the \$43,456.78 <lazy> #90 dog & duck/goose, as 12.5% of E-mail from aspammer@website.com is spam. Der „schnelle“ braune Fuchs springt über den faulen Hund. Le renard brun «rapide» saute par-dessus le chien paresseux. La volpe marrone rapida salta sopra il cane pigro. El zorro marrón rápido salta sobre el perro perezoso. A raposa marrom rápida salta sobre o cão preguiçoso.</p>	<pre>OCR output: The (quick) [brown] {fox} jumps! Over the \$43,456.78 <lazy> #90 dog & duck/goose, as 12.5% of E-mail from aspammer@website.com is spam. Der „schnelle“ braune Fuchs springt über den faulen Hund. Le renard brun «rapide» saute par-dessus le chien paresseux. La volpe marrone rapida salta sopra il cane pigro. El zorro marrón rápido salta sobre el perro perezoso. A raposa marrom rápida salta sobre o cão preguiçoso.</pre>
-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Figura 15 Salida de un texto de caracteres latinos sin tratar

Fuente: (Hidalgo & Sánchez, 2015)

La comparación de patrones es la etapa en la que se establece el buen funcionamiento del OCR, se comparan los caracteres obtenidos anteriormente con unos teóricos (patrones) almacenados en una base de datos. Para el (Grupo de Investigación CATAM, 2013) existen varios métodos para realizar la comparación:

- Método de proyección: se obtienen proyecciones verticales y horizontales del carácter y se comparan con el alfabeto hasta encontrar la máxima coincidencia.
- Métodos estructurales: esqueletizan los caracteres extraídos de la imagen, detectando contornos interiores y exteriores reconociendo una serie de puntos singulares que permiten realizar una vectorización.

- Métodos geométricos o estadísticos: se extrae un conjunto de medidas que constituyen la representación de un vector de un espacio de representación.
- Métodos neuro-miméticos: basados en la utilización de redes neuronales, disponen de una gran base de datos de aprendizaje con la cual se obtienen buenos resultados.
- Métodos Markovianos: se hacen tablas de búsqueda de secuencia de señales de caracteres pero variables en el tiempo.
- Métodos de Zadeh: se basan en la lógica borrosa con la capacidad de realizar OCR con un tipo de datos impreciso.
- Métodos basados en IA: aplican sistemas de decisión en base a reglas utilizando algoritmos de búsqueda de árboles.
- Métodos mixtos: un sistema combina un árbol de decisión con una red neuronal.

2.3.1 Alternativas del OCR

a) Tesseract OCR

Es un OCR originalmente desarrollado desde 1984 a 1994 por Hewlett Packard (HP) bajo licencia de pago, posteriormente en 1995 fue enviado a la University of Nevada Las Vegas (UNLV) donde fue comparado frente otros reconocedores comerciales y se demostró que realmente valía. En el 2005 HP y la UNLV liberan su código fuente, siendo desarrollado y distribuido al presente por Google, es considerado uno de los motores OCR libres con mayor precisión disponible (Sandip, y otros, 2012).

b) Open Cognitive OCR

“Es un programa gratuito de reconocimiento óptico de caracteres. Es capaz de reconocer texto en 23 idiomas y guardar el resultado en archivos” (Makeuseof, 2017).

2.3.2 Algoritmos para el desarrollo de un OCR

El algoritmo de OCR de la figura 16 supone que su entrada es una imagen binaria con regiones poligonales de texto opcionalmente definidas. El procesamiento sigue un enfoque paso a paso tradicional. El primer paso es un análisis de los componentes conectados en el que se almacenan sus contornos. Los contornos simplemente se anidan en grupos a los cuales más tarde se organiza en líneas de texto, y las líneas y regiones se analizan en busca de texto distribuido de forma proporcional. Posteriormente las líneas de texto se dividen en palabras según el espaciado entre caracteres. El reconocimiento entonces procede como un proceso de dos pasos. En el primer paso, se intenta reconocer cada palabra. Cada palabra satisfactoria pasa a un clasificador como dato de entrenamiento. Dado que el clasificador puede haber aprendido algo útil demasiado tarde en el segundo paso, las palabras que no fueron reconocidas lo suficientemente bien se reconocen de nuevo (Smith, 2007).

En la figura 17 se verifica la estructura de otro algoritmo para realizar el reconocimiento óptico de caracteres.

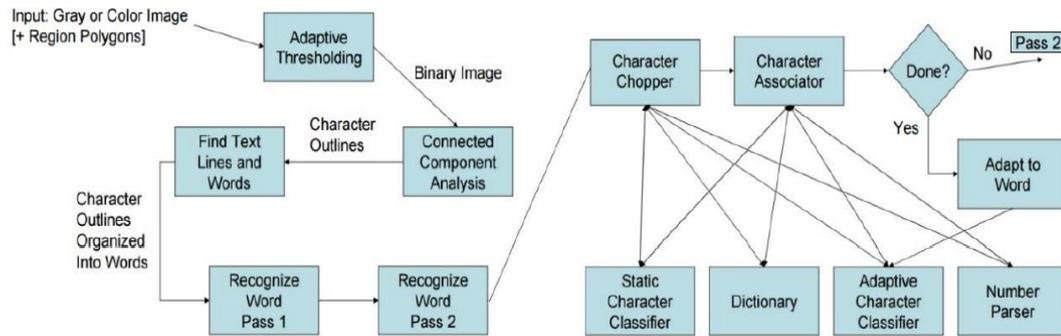


Figura 16 Esquema de trabajo de un OCR

Fuente: (Smith, 2007)

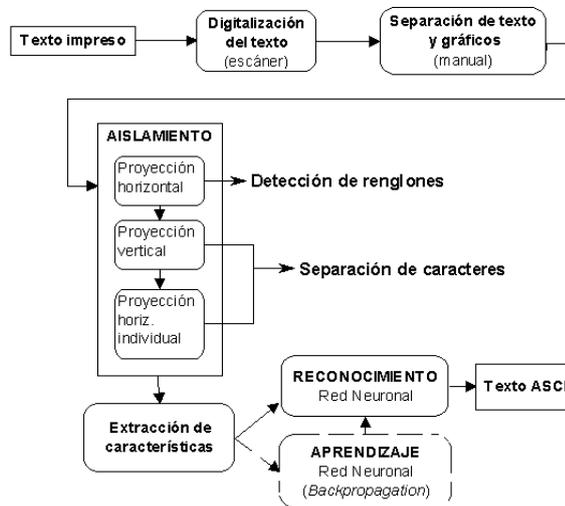


Figura 17 Esquema de reconocimiento de caracteres

Fuente: (Bejarano & Sánchez, 2015)

2.3.3 Aplicaciones del OCR

Algunas de las más relevantes aplicaciones que utilizan algoritmos de reconocimiento de caracteres OCR son:

- Reconocimiento de texto manuscrito: la segmentación de texto continuo es un procedimiento complejo por lo que es necesario utilizar algoritmos robustos que usen una segmentación previa (Kultura 2.0, 2017).

- Reconocimiento de matrículas: un radar debe ser capaz de localizar la matrícula de un vehículo en condiciones variables. Se buscan texturas similares a la de una matrícula, se aísla el área rectangular y se aplica un proceso de clasificación de píxeles obteniendo una cadena de caracteres que se ajusta a un modelo conocido (Herrera Martínez, 2013).
- Procesado automático de documentos: el tratamiento de documentos se traduce en la exploración óptica de los mismos con los metadatos obtenidos, la determinación de su tipo entre diversas plantillas previamente registradas, la localización y el reconocimiento de los campos informativos, validación y catalogación automática (López Rincón, 2015).

2.4 Conversión TTS

La conversión TTS (del inglés Text To Speech – Texto a voz) es la generación por medios automáticos de una voz artificial o sintética que genera el sonido producido por una persona al leer un texto cualquiera en voz alta o una voz artificial (Agüero, 2012).

2.4.1 Alternativas de TTS

a) Espeak

Para la pronunciación del texto se utiliza un sintetizador compacto de voz de código abierto. Espeak puede traducir texto en códigos de fonemas con información de tono y longitud, por lo que se puede adaptar como interfaz para los dífonos de Mbrola (eSpeak, 2017).

b) Festival voices

Festival es un sintetizador de voz disponible libremente basado en concatenación de dífonos y selección de unidades, está disponible para español, inglés británico y americano y galés (Norris, 2017).

c) M-brola

Es un cúmulo de “sintetizadores de voz para el mayor número posible de idiomas y proporcionarlos gratuitamente para aplicaciones no comerciales” (TCTS Lab, 2017), se basa en la “concatenación de dífonos, toma como entrada una lista de fonemas junto con la información prosódica y produce una salida de audio a 16 bits” (Correa , Rueda, & Arguello, 2014).

2.4.2 Algoritmo de un TTS

Consiste de tres fases principales como se observa en la figura 18 donde se esquematiza el proceso desde la entrada de texto hasta la salida de voz.

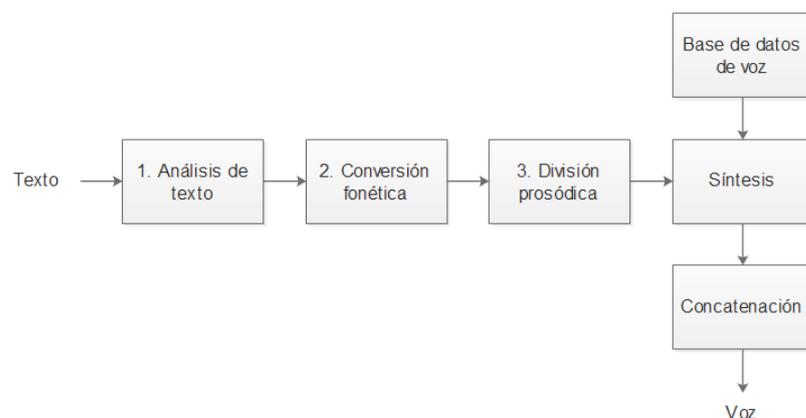


Figura 18 Diagrama de bloques de un TTS

Resumen

Se ha realizado un análisis de las cifras concernientes a las discapacidades en el Ecuador. También, una revisión de las diferentes opciones en cuanto a filtros y técnicas de visión artificial, algoritmos OCR y software para la conversión de texto a audio.

CAPÍTULO 3

DISEÑO E IMPLEMENTACIÓN

El capítulo da a conocer el desarrollo necesario para la selección y dimensionamiento de los elementos del dispositivo lector de textos. Se realiza la identificación de necesidades, definición de especificaciones y arquitectura, análisis de alternativas de diseño, selección de conceptos, definición de diseño y concepto final para realizar la implementación de sus partes.

Se aplica el diseño concurrente, propio de la mecatrónica, donde se pueden desarrollar varias etapas al mismo tiempo. Cada etapa es parte de un proceso interactivo, se puede regresar a fases anteriores para realizar cambios y con los mismos seguir avanzando dentro de los procesos.

3.1 Identificación de necesidades

Se plantea identificar las necesidades que se busca satisfacer con el dispositivo. Es esencial en el desarrollo del proyecto saber los requerimientos de los usuarios potenciales.

Ulrich & Eppinger (2013), analiza un método para cerciorarse que el producto se oriente hacia los requerimientos del usuario y se identifique las necesidades realizando los pasos siguientes:

1. Recopilar datos sin procesar de los clientes.
2. Interpretar los datos sin procesar en términos de las necesidades de clientes.
3. Organizar las necesidades en jerarquías.
4. Establecer la importancia relativa de las necesidades.
5. Reflexionar en los resultados.

En el paso uno y dos se debe tomar contacto con el usuario. En la Tabla 2 se visualizan los enunciados del cliente y las necesidades que se interpretan.

Tabla 2

Identificación de necesidades

Enunciado cliente	Necesidad interpretada
El dispositivo debe capturar imágenes	El dispositivo captura imágenes
El dispositivo debe reconocer texto	El dispositivo detecta texto en español
El dispositivo debe convertir el texto a audio	El dispositivo convierte el texto a audio en español
El dispositivo debe reproducir el audio	El dispositivo pronuncia el audio obtenido
El dispositivo debe ser pequeño y ligero	El dispositivo es de diseño compacto y optimizado
El dispositivo debe ser fácil de usar	El dispositivo es plug & play
El dispositivo debe soportar caídas	El dispositivo es resistente a impactos
El costo de elaboración debe ser lo más bajo posible sin disminuir su calidad	El dispositivo es de buena calidad y bajo costo
El dispositivo debe tener comunicación con el usuario	El dispositivo posee señales de retroalimentación

Para apoyar el proceso de desarrollo de los pasos siguientes (3, 4 y 5) es necesario realizar una matriz QFD. A partir de las necesidades del cliente (Qué's) se procede a plasmar en la tabla 3 las alternativas de solución (Cómo's). En la tabla 4 se muestra la importancia relativa de las necesidades y en el Anexo A se verifica la matriz QFD completa.

Tabla 3**Qué's y Cómo's para matriz QFD**

Qué's	Cómo's
Adquisición de imágenes	Programación, iluminación, selección elementos electrónicos
Detección de texto	Visión artificial, OCR, programación
Reconocimiento de texto	Programación, OCR
Conversión TTS	Normalización del texto, sintetizador de voz
Diseño compacto y optimizado	Selección de material, diseño 3D
Plug & play	Programación
Resistente	Selección de material, diseño mecánico
Calidad y bajo costo	Selección de material, diseño ergonómico
Bajo Costo	Selección de material, diseño 3D
Señales de retroalimentación	Programación, actuadores

Tabla 4**Importancia relativa de las necesidades**

Necesidad	Importancia Relativa
Adquisición de imágenes	12,8 %
Detección de texto	12,8 %
Reconocimiento de texto	12,8 %
Conversión TTS	12,8 %
Diseño compacto y optimizado	7,7 %
Plug & play	7,7 %
Resistente	10,3 %
Calidad y bajo costo	7,7 %
Bajo Costo	7,7 %
Señales de retroalimentación	7,7 %

3.2 Definición de las especificaciones

A partir de las necesidades del usuario y la importancia relativa, es posible definir los requisitos del dispositivo lector de textos.

- Medidas máximas: 7x7x5 cm.
- Peso máximo: 200g.
- Idioma de lectura: Español.
- Fuente de lectura recomendada: arial 14 e interlineado de 1,5.
- Diseño estético y ergonómico.
- Procesador del ordenador: Intel Core i5, 4 Gb de RAM.
- Puertos USB mínimos del ordenador: 2.
- Salida de audio del ordenador: Jack 3,5 mm.
- Resolución mínima de la cámara: 1080x720 pixeles
- Ángulo de visión mínimo de la cámara: 60°
- Distancia focal de la cámara: 35mm
- Salidas digitales mínimas del controlador: 3
- Entradas analógicas mínimas del controlador: 1

3.3 Definición de la arquitectura y diseño de los módulos

Al conocer las especificaciones del dispositivo, se precisa definir la arquitectura modular del dispositivo. Tomando en cuenta los elementos funcionales se ha generado la distribución mostrada en la figura 19.

En la arquitectura del dispositivo se observa al case como elemento de mayor tamaño, ya que cumple con la función de soportar y proteger a los módulos: cámara, placa electrónica, motores vibradores, leds y LDR. La iluminación está provista por los leds y se realiza su control por medio de la LDR. Se realiza la adquisición de imágenes por la cámara. Los motores vibradores envían señales de retroalimentación al usuario. El controlador

comunica los actuadores y LDR con el ordenador, también los controla. El ordenador procesa las imágenes, reconoce el texto y lo convierte en audio.

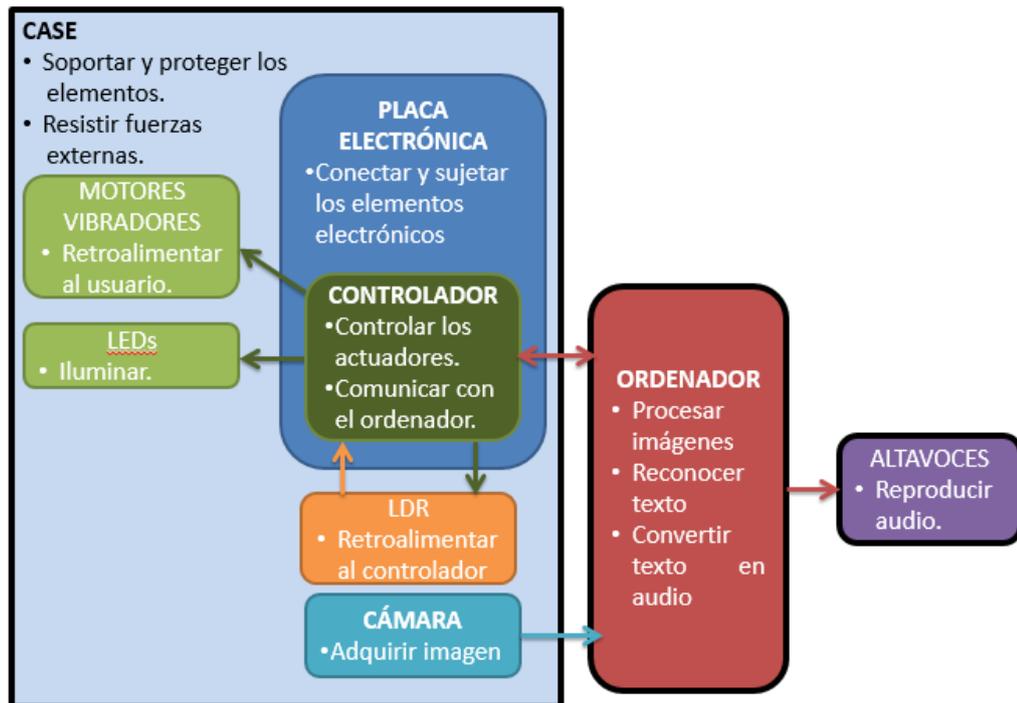


Figura 19 Arquitectura del dispositivo

A continuación se realiza un análisis acerca de las alternativas de solución, selección y definición para cada módulo especificado en la arquitectura, así como para los procesos que se ejecutan en el ordenador y controlador.

3.3.1 Módulo cámara

La captura de imágenes es una función principal para el dispositivo lector de textos. Las opciones de cámaras para visión artificial fueron analizadas en el capítulo 2. Para la adquisición de imágenes se requiere de una cámara que cumpla con las especificaciones definidas anteriormente.

Se selecciona la cámara Microsoft LifeCam HD 3000 del tipo matricial, en la tabla 5 se muestra las características técnicas.

Tabla 5**Características Microsoft LifeCam HD 3000**

Característica	Valor
Sensor	CMOS
Resolución	1280x720
Definición	720p HD
Frames por segundo	30
Ángulo de visión	68,5°
Distancia focal	35 mm
Enfoque	Manual

Fuente: (Microsoft, 2017)

3.3.2 Soporte

La estructura del case para el dispositivo se la realiza tomando en cuenta su función como soporte para la cámara, placa electrónica, leds, fotorresistencia y motores vibradores. El diseño debe facilitar la ubicación de los componentes, sabiendo que el dispositivo debe ser versátil y de tamaño reducido para su fácil operación.

La norma técnica UNE-EN ISO 9241-9:2015 provee los requerimientos ergonómicos para dispositivos de uso en contacto con el usuario, de la norma se han tomado los criterios de diseño que se aplican al dispositivo (ISO, 2015).

- Proporcionar una retroalimentación adecuada al usuario.
- Ser eficiente, confortable, estable y seguro mientras está en funcionamiento.
- Estar diseñado de forma que:
 - Sea fácil aprender a utilizarlo.
 - Minimice la carga física y mental del usuario.

- No limite la capacidad física o mental del usuario.
- Se tenga en cuenta la posición relativa en relación con otros componentes de la tarea.
- No debe presentar perfiles o partes agudas.
- Permitir que el usuario descanse sin que éste se active.
- Estar contruidos con materiales poco conductores del calor.

En base a las funciones que debe cumplir el soporte y los criterios de ergonomía se realizan tres opciones de diseño (ver figura 20).

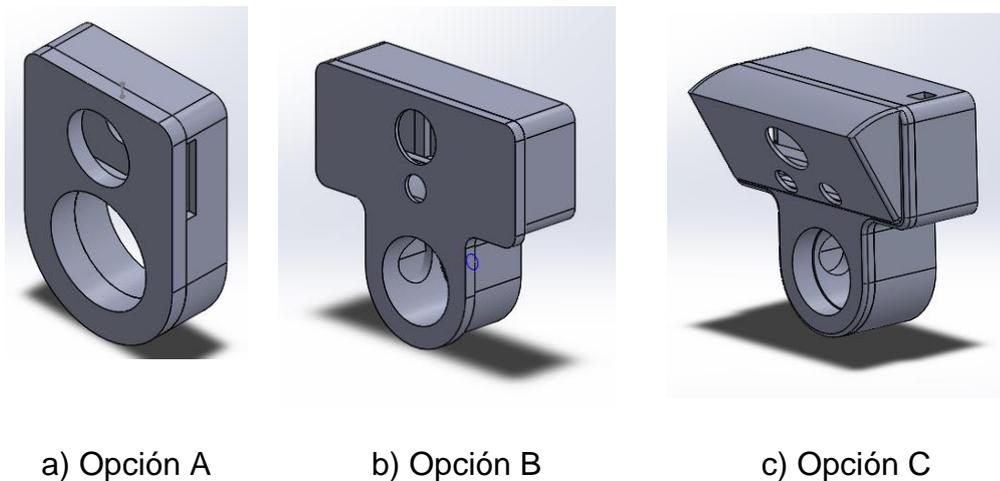


Figura 20 Opciones de diseño 3D

Para realizar la selección se establecen criterios (ver tabla 6), se comparan los criterios (ver tabla 7), se pondera cada opción en función de cada criterio (ver tablas 8, 9, 10, 11 y 12) y se selecciona la mejor puntuada.

Tabla 6

Criterios a evaluar para el case

Criterio	Letra
Espacio interno	A
Cobertura área de lectura	B
Ergonomía	C
Tamaño	D

Tabla 7**Matriz comparativa entre criterios del case**

Criterios	A	B	C	D	$\sum +1$	Peso
A		1	0,5	0,5	3	0,30
B	0		0,5	0,5	2	0,20
C	0,5	0,5		0,5	2,5	0,25
D	0,5	0,5	0,5		2,5	0,25
Total					10	1,00

Tabla 8**Matriz criterio A Vs alternativas diseño del case**

	A	B	C	$\sum +1$	Peso
Opción A		0	0	1	0,17
Opción B	1		0,5	2,5	0,42
Opción C	1	0,5		2,5	0,42
Total				6	1,00

Tabla 9**Matriz criterio B Vs alternativas diseño del case**

	A	B	C	$\sum +1$	Peso
Opción A		0,5	0	1,5	0,25
Opción B	0,5		0	1,5	0,25
Opción C	1	1		3	0,50
Total				6	1,00

Tabla 10**Matriz criterio C Vs alternativas diseño del case**

	A	B	C	$\sum +1$	Peso
Opción A		0	0	1	0,17
Opción B	1		0,5	2,5	0,42
Opción C	1	0,5		2,5	0,42
Total				6	1,00

Tabla 11**Matriz criterio D Vs alternativas diseño del case**

	A	B	C	$\sum +1$	Peso
Opción A		1	1	3	0,50
Opción B	0		0,5	1,5	0,25
Opción C	0	0,5		1,5	0,25
Total				6	1,00

Tabla 12**Matriz resultante diseño del case**

Criterio	A	B	C	D	\sum	Selección
Opción A	0,05	0,05	0,04	0,13	0,27	3
Opción B	0,13	0,05	0,10	0,06	0,34	2
Opción C	0,13	0,10	0,10	0,06	0,39	1
Total					1	

Por lo tanto la opción que más se aproxima a los criterios de diseño 3D es la opción C.

Se plantea un estudio de caída, que permita saber la tensión máxima que puede sufrir el dispositivo en función del material seleccionado sin sufrir ninguna deformación.

Para determinar la altura de caída, se toma como referencia la posición ergonómica de una persona utilizando un computador, ya que para el uso del dispositivo se requiere apoyarlo sobre el texto de lectura ubicado sobre una mesa.

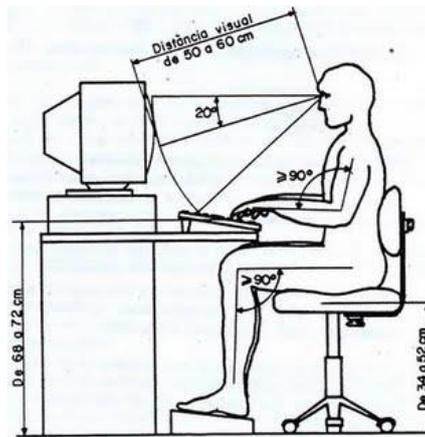


Figura 21 Posición ergonómica uso del computador

Fuente: (Restrepo, 2013)

Según la figura 19 la altura para la caída será 72 cm y se calcula la velocidad de impacto contra el suelo.

$$V_f^2 = V_o^2 + 2gh$$

$$V_f = \sqrt{V_o^2 + 2gh}$$

$$V_f = \sqrt{0^2 + 2 \times 9,8 \frac{m}{s^2} \times 0,72 m}$$

$$V_f = 3,75 \frac{m}{s}$$

En la tabla 13 se visualiza las características del ABS y PLA, se selecciona al ABS para realizar el análisis de plasticidad de Von Mises.

Tabla 13

Propiedades mecánicas del ABS y PLA.

Propiedad	ABS	PLA
Gravedad elástica	1,05 /cm ³	1,25 /cm ³
Límite elástico	38,8 MPa	32 MPa
Resistencia a la tensión	50 MPa	38 MPa
Módulo de elasticidad	2,1 GPa	3,63 GPa
Coefficiente de Poisson	0,394	0,45
Elongación	10 al 30 %	6 %

Fuente: (Groover, 2012)

El estudio de impacto realizado muestra como esfuerzo máximo el valor de 44,19 GPa siendo el límite elástico 3,38 MPa. Se verifica que la geometría y material seleccionado brinda al dispositivo resistencia a caídas (ver figura 4).

Se debe considerar que el material seleccionado es para la realización de prototipos y no se lo puede atribuir como material de selección final. Se hará una modificación para que la porción del dispositivo que está en contacto con el usuario sea de material flexible para que mejore su sujeción y comodidad.

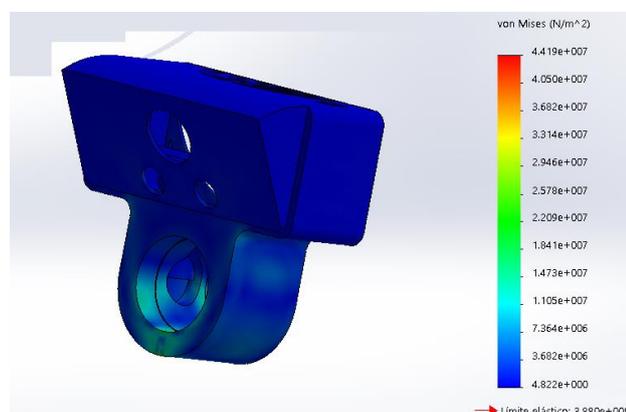


Figura 22 Estudio de caída de Von Mises

3.3.3 Placa electrónica

Se requiere del diseño y fabricación de una placa electrónica para interconectar los componentes y circuitos.

La Asociación de Industrias de Conexiones Electrónicas (IPC) estandariza el diseño y elaboración de placas electrónicas por medio de la norma ANSI/IPC-2221. Se extrae de la norma el cálculo para el ancho pista y reglas de diseño a tomar en cuenta para el diseño de la PCB (IPC, 2012).

Para realizar el cálculo del ancho de pista se requiere la temperatura de funcionamiento, temperatura ambiente, corriente máxima y el material del conductor (fibra de vidrio), se verifica éstas características en la tabla 14.

Tabla 14

Características Placa Electrónica

Característica	Valor	Unidad
Temperatura de funcionamiento máxima	130	°C
Temperatura ambiente	30	°C
Material – FR4	1	oz/ft ²
Corriente máxima	0,5	A

A partir de la siguiente ecuación se calcula la temperatura de trabajo.

$$T \text{ trabajo} = T \text{ operación} - T \text{ ambiente}$$

$$T \text{ trabajo} = 130 - 30$$

$$T \text{ trabajo} = 100 \text{ °C}$$

Para saber la sección de la pista se ingresa con la temperatura de trabajo 100°C y la corriente máxima 0,5A a la figura 5. Se obtiene el valor de 10 milis para la sección. Con el dato de 10 milis y el del material de la placa se puede determinar en la Figura 6 el grosor de la pista.

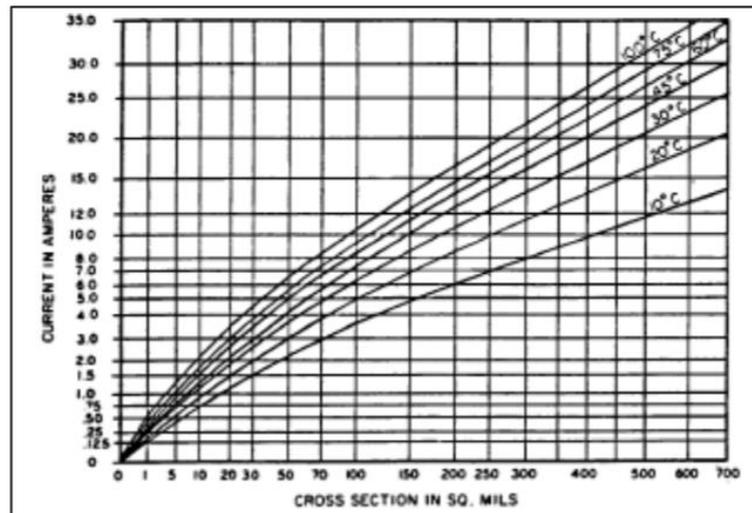


Figura 23 Gráfica seleccionar sección de la pista

Fuente: (IPC, 2012)

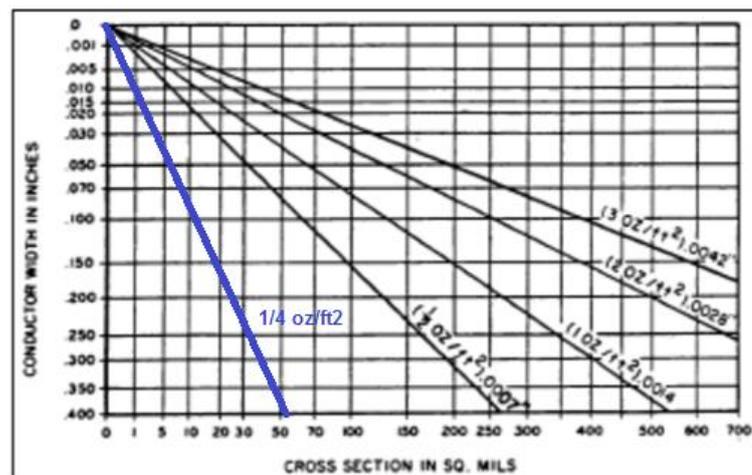


Figura 24 Gráfica seleccionar sección de la pista

Fuente: (IPC, 2012)

Según los datos calculados, en la figura 24 se determina el ancho de la pista, que sería de 0,1" transformando a milímetros se tiene 2,5mm.

Las reglas extraídas de la norma IPC – 2221 que se aplican para el diseño de la PCB son las siguientes (IPC, 2012).

- Se debe evitar ángulos agudos.
- En pistas paralelas la distancia de las mismas ha de ser constante.
- La pista debe poseer unión radial con el pad.
- No se pueden unir de forma directa dos o más pads.
- Solo pueden intersecarse hasta cuatro pistas en un mismo pad.
- Mientras menos pistas y más cortas sean es mejor.
- El diámetro de los pads debe ser mínimo el doble de las pistas.
- De haber tierras de prueba y no deben estar a menos de 3mm del borde.
- No se debe conectar clavijas de control directamente a tierra.
- Los conectores deben ser semejantes en su geometría y estar ubicados de forma uniforme.
- En función de la corriente y temperatura se diseña ancho de las pistas.
- El espacio entre conductores internos debe ser de 0,05mm si el voltaje entre éstos va de 0 a 15 V.
- Debe presentar disipación de calor por convección, radiación o conducción.
- De ser necesario se puede incluir un puente de punto a punto, será considerado parte del montaje, más no de la PCB.
- El espesor mínimo de la pista para ¼ oz/ft debe ser de 6µm.

Con el dimensionamiento realizado y tomando en cuenta las reglas de la norma IPC 2221, se diseña el PCB en un software especializado (ver figura 25).

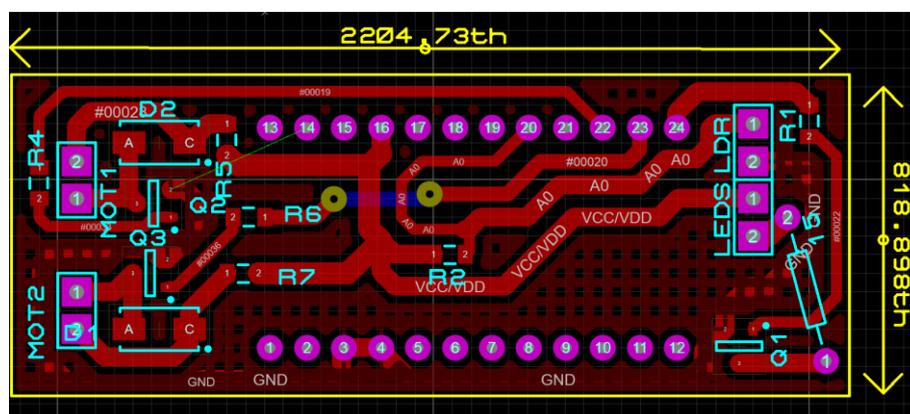


Figura 25 Diseño PCB

3.3.4 Iluminación

Se requiere de iluminación para mejorar las capturas tomadas por la cámara, ya que la presencia de luz influye notablemente en el proceso de obtención de texto de la imagen.

La luz es un factor crítico para que se realice el proceso de adquisición de imagen. Se debe añadir una fuente de luz y se opta por los LEDs del tipo SMD 3030 por su tamaño reducido, bajo consumo y alta potencia (ver tabla 15).

Tabla 15

Características LED SMD 3030

Tipo	Tamaño [mm]	Corriente [mA]	Voltaje [V]	Potencia [mW]	Intensidad lumínica [lm]
3030	3x3	70	3	210	25

Fuente: (Wayjun Technology, 2017)

Para controlar la luminosidad de los leds se diseña un circuito con transistores 2N3904 en corte y saturación, controlados por PWM desde la placa Arduino.

El circuito base para la realización de los cálculos de resistencias de base y colector es el que se muestra en la figura 26.

La corriente nominal del LED SMD 3030 es de 70mA y al utilizarse dos leds conectados en paralelo, es necesario que circule en el colector una corriente de 140mA.

$$I_C = 140mA$$

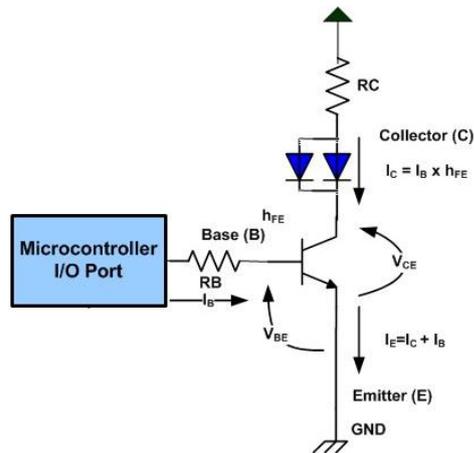


Figura 26 Circuito de corte y saturación

Fuente: (Hermosa Donate, 2012)

A continuación se calcula la resistencia del colector R_c para obtener la corriente de colector I_c , en el circuito de la figura 27 se aplica la diferencia de potencial en la rama.

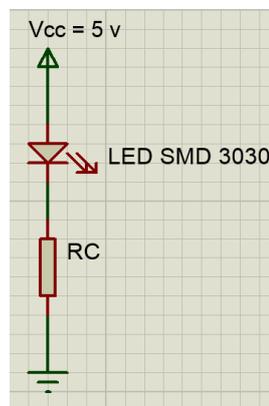


Figura 27 Rama hacia el colector del transistor

$$V_{cc} = V_{led} + V_{Rc}$$

$$V_{cc} = 3\text{ V} + I_c \times R_c$$

$$V_{cc} = 3\text{ V} + 140\text{ mA} \times R_c$$

$$R_c = \frac{5\text{ V} - 3\text{ V}}{140\text{ mA}}$$

$$R_c = 14,29\ \Omega$$

Se aproxima el valor obtenido a una resistencia comercial.

$$R_c \approx 15 \Omega$$

Se dimensiona la potencia de la resistencia de la siguiente manera.

$$P = I \times V \quad \mathbf{3.1}$$

$$P = I_c \times (V_{cc} - V_{led})$$

$$P = 140mA \times (5V - 3V)$$

$$P = 0,28 w$$

Se duplica la potencia calculada para tener un margen de seguridad en el componente, siendo esta de:

$$P = 0,28w \times 2$$

$$P = 0,56 w$$

Por lo que la resistencia del colector dimensionada tiene un valor de 15Ω a una potencia de 0,5 watt.

Para calcular la corriente de la base I_b se toma en cuenta la ganancia de corriente del transistor donde se relaciona la corriente de base I_b con la corriente de colector I_c y la ganancia $hFE=100$, que es tomada de la hoja técnica del transistor 2N3904 (On Semiconductor, 2017).

$$I_c = hFE \times I_b$$

$$I_b = \frac{I_c}{hFE} \quad \mathbf{3.2}$$

$$I_b = \frac{140 mA}{100}$$

$$I_b = 1,4 mA$$

Con la corriente de base calculada resta calcular la resistencia de base R_b , tomando en cuenta el dato de la hoja técnica del transistor donde el voltaje entre la base y el emisor es $V_{be} = 0,7 \text{ V}$ (On Semiconductor, 2017) y aplicando el voltaje total en la rama se obtiene R_b (ver figura 28).

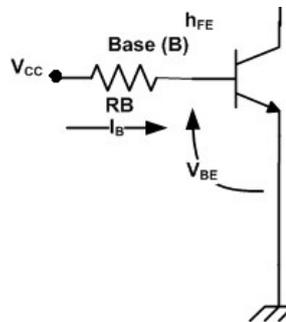


Figura 28 Rama hacia la base del transistor

$$V_{cc} = V_{BE} + V_{RB} \quad 3.3$$

$$V_{cc} = 0,7 \text{ V} + I_b \times R_b$$

$$V_{cc} = 0,7 \text{ V} + 1,4 \text{ mA} \times R_b$$

$$R_b = \frac{5 \text{ V} - 0,7 \text{ V}}{1,4 \text{ mA}}$$

$$R_b = 3,7 \text{ k}\Omega$$

Se aproxima el valor obtenido a una resistencia comercial.

$$R_b \approx 3,9 \text{ k}\Omega$$

Se dimensiona la potencia de la resistencia partir de la Ecuación 3.1, de la siguiente manera.

$$P = I_b \times (V_{cc} - V_{BE})$$

$$P = 1,4 \text{ mA} \times (5 \text{ V} - 0,7 \text{ V})$$

$$P = 6 \text{ mw}$$

Se duplica la potencia calculada para tener un margen de seguridad en el componente, siendo esta de

$$P = 6mw \times 2$$

$$P = 12 mw$$

Por lo que la resistencia de la base dimensionada tiene un valor de 3,9k Ω a una potencia de 0,25 watt de forma comercial.

3.3.5 Control de iluminación

La luz emitida por los leds debe variar dependiendo de la luz natural del lugar en donde se vaya a usar el dispositivo. Se puede realizar un divisor de tensión con una resistencia y una LDR, para controlar la intensidad de corriente enviada a los leds.

Para que exista una variación significativa en el valor analógico obtenido del divisor de tensión, la resistencia debe tener un valor promedio en el rango de máximo y mínimo de variación de la LDR. Se asume un valor de 10k Ω para la resistencia que conforma el divisor de tensión (ver tabla 16).

Tabla 16

Variación de la foto-resistencia

Hora	Valor mínimo [Ω]	Valor máximo [Ω]
10:00	2.35k	5.15k
12:00	1.79k	4.12k
16:00	0.39k	3.55k
18:00	1.14k	12.9k

En la Figura 29 se aprecia el divisor de tensión a utilizarse para el control de la luminosidad de los LED's.

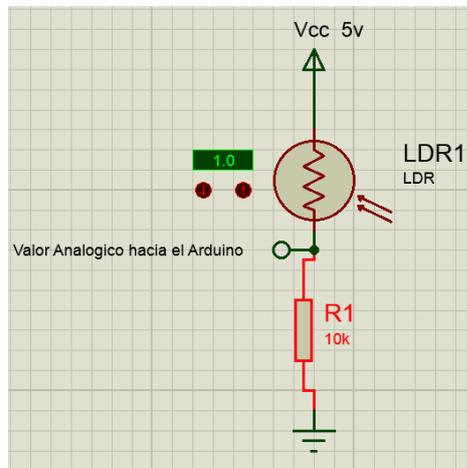


Figura 29 Divisor de tensión

3.3.6 Retroalimentación

Para que el usuario tenga una retroalimentación sensorial, se plantea el uso de motores vibradores, con el objetivo de que la persona pueda recibir señales de alerta del dispositivo.

En la tabla 17 se analizan las características técnicas de un motor vibrador de disco, es la opción seleccionada principalmente por su tamaño reducido, RPM y bajo consumo.

Tabla 17

Características del motor vibrador de disco

Tipo	Tamaño [mm]	Peso [g]	Voltaje [V]	Velocidad [RPM]	Corriente [mA]
Vibrador de disco	$\Phi=10$ espesor=2	0,8	2,5 a 3,5	14500	60

Fuente: (Agelectronica, 2017)

Para el control de los motores vibradores de disco, se utiliza el mismo circuito de la figura 26, pero cada vibrador tendrá su propio circuito y se realiza el cálculo de la resistencia de base y de colector de la misma manera.

Para comprobar la corriente nominal del motor vibrador de disco, se alimenta el mismo con una tensión de 3V, con lo que se obtiene una corriente de 66mA.

$$V = 3V \rightarrow I = 66mA$$

Se calcula la potencia del vibrador con la Ecuación 3.1.

$$P = 66mA \times 3V$$

$$P = 198mW$$

Con este dato se puede calcular la corriente necesaria en el colector para que el vibrador consuma la potencia nominal, con una alimentación de 5V provista por el Arduino, de la Ecuación 3.1 se despeja la corriente para el cálculo.

$$I = \frac{P}{V}$$

$$I = \frac{198mW}{5V}$$

$$I = 39,6mA \approx 40mA$$

$$I_C = 40mA$$

A continuación se calcula la resistencia del colector Rc para obtener la corriente de colector Ic, en el circuito de la figura 30 se aplica la diferencia de potencial en la rama.



Figura 30 Rama hacia el colector del transistor

$$V_{cc} = V_{motor} + V_{Rc}$$

$$V_{cc} = 3\text{ V} + I_c \times R_c$$

$$V_{cc} = 3\text{ V} + 40\text{ mA} \times R_c$$

$$R_c = \frac{5\text{ V} - 3\text{ V}}{40\text{ mA}}$$

$$R_c = 50\ \Omega$$

Se aproxima este valor a una resistencia comercial.

$$R_c \approx 47\ \Omega$$

Se dimensiona la potencia de la resistencia, a partir de la Ecuación 3.1 de la siguiente manera.

$$P = I_c \times (V_{cc} - V_{motor})$$

$$P = 40\text{ mA} \times (5\text{ V} - 3\text{ V})$$

$$P = 80\text{ mw}$$

Se duplica la potencia calculada para tener un margen de seguridad en el componente.

$$P = 80 \text{ mw} \times 2$$

$$P = 0,16 \text{ w}$$

Por lo que la resistencia del colector dimensionada tiene un valor de 47Ω a una potencia de 0,25 watt.

Para calcular la corriente de la base I_b se toma en cuenta la ganancia de corriente del transistor donde se relaciona la corriente de base I_b con la corriente de colector I_c , la ganancia $hFE=100$ se lo toma de la hoja técnica del transistor 2N3904 (On Semiconductor, 2017).

Para calcular la corriente de la base I_b se toma en cuenta la Ecuación 3.2 desarrollada en el apartado anterior se toma el mismo valor de ganancia ya que se utilizará el mismo transistor.

$$I_b = \frac{40 \text{ mA}}{100}$$

$$I_b = 0,4 \text{ mA}$$

Con la corriente de base I_b resta calcular la resistencia de base R_b , a partir de la Ecuación 3.3 se realiza el mismo cálculo con el mismo voltaje de base emisor.

$$V_{cc} = 0,7 \text{ V} + I_b \times R_b$$

$$V_{cc} = 0,7 \text{ V} + 0,4 \text{ mA} \times R_b$$

$$R_b = \frac{5 \text{ V} - 0,7 \text{ V}}{0,4 \text{ mA}}$$

$$R_b = 10,7 \text{ k}\Omega$$

Se aproxima este valor a una resistencia comercial.

$$R_b \approx 10 \text{ k}\Omega$$

Se dimensiona la potencia de la resistencia a partir de la Ecuación 3.1, de la siguiente manera.

$$P = I_b \times (V_{cc} - V_{motor})$$

$$P = 40\text{mA} \times (5\text{V} - 3\text{V})$$

$$P = 80 \text{ mw}$$

Se duplica la potencia calculada para tener un margen de seguridad en el componente, siendo esta de

$$P = 80 \text{ mw} \times 2$$

$$P = 0,16 \text{ w}$$

Por lo que la resistencia del colector dimensionada tiene un valor de $10\text{k}\Omega$ a una potencia de 0,25 watt.

3.3.7 Controlador

Se programa el microcontrolador de la placa Arduino Pro Micro (Sparkfun, 2017) para configurar las terminales de los motores y leds en modo escritura, la lectura analógica del fotorresistor para el control de los leds por ancho de pulso y, la velocidad de transmisión en el puerto serie de manera que pueda enlazarse con el algoritmo de control principal. Cuando el puerto serie recibe datos a través del buffer de entrada se validan los códigos asignados para las diferentes secuencias de activación de los motores. En la Figura 31 se resume el diagrama de flujo del controlador.

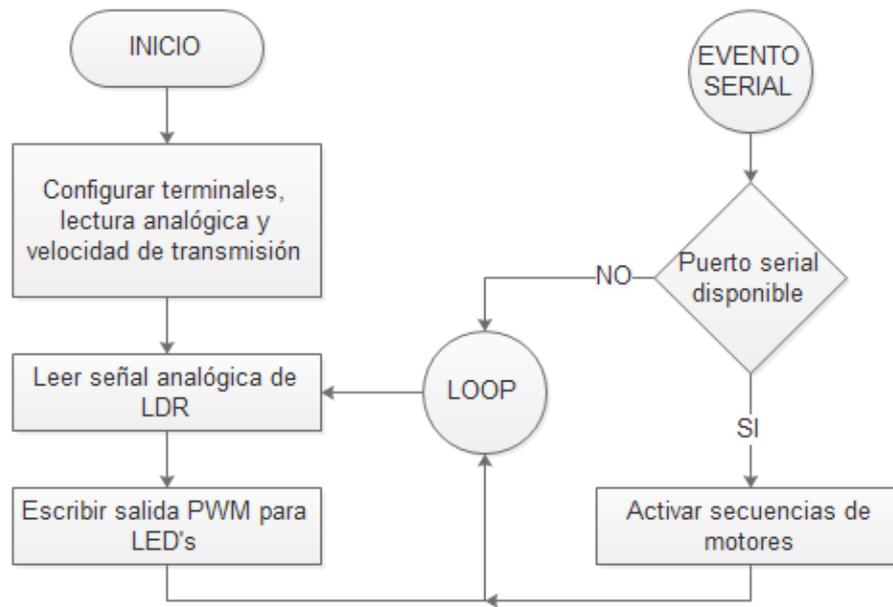


Figura 31 Diagrama de flujo del controlador

3.3.8 Proceso genérico

En la figura 32 se muestra el proceso genérico para la generación de audio a partir de una imagen. A continuación se explica cada bloque funcional y los subprocesos que intervienen en el diseño del algoritmo de control.

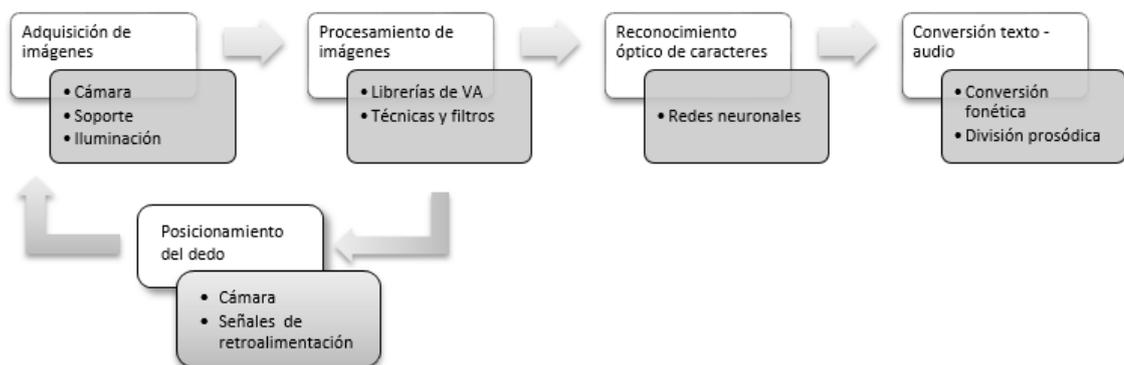


Figura 32 Proceso genérico del dispositivo

a) Adquisición de imágenes

El ancho del fotograma así como el ángulo de visión de la cámara son características determinantes en la adquisición de la imagen ya que determinan la cantidad de texto contenida dentro de un fotograma.

Las imágenes adquiridas son imágenes digitales básicas compuestas por tres elementos espectrales que corresponden a cada uno de los canales del espacio de color RGB. Todos los elementos son representados en una matriz de posiciones (x, y) y toman un valor numérico en un rango entre $[0, 255]$ debido a la representación de 8 bits de cada par coordenado. Cada uno de estos elementos corresponde a un pixel.

La definición HD de la cámara web otorga una resolución aproximada de 1 megapíxel que proporciona un buen detalle de imagen considerando que su propósito es capturar texto impreso.

b) Procesamiento de imágenes

El procesamiento de imágenes se divide en partes secuenciales, en cada una de ellas se utilizan técnicas y metodologías proporcionadas principalmente por la librería de visión artificial multiplataforma OpenCV.

- Detección de dedo índice

El paso inicial trata de acertar si la cámara enfoca un dedo en contraste con el papel. La yema del dedo índice portador del dispositivo trabaja como cursor guía sobre el texto que se pretende leer.

Detectar la piel en la imagen adquirida no es posible por las diversas gamas de colores que se presentan (ver figura 33.a). La normalización elimina las regiones resaltadas, las sombras y hace que la piel sea más fácil

de detectar como se muestra en la figura 33.b. Para extraer el tono de piel se convierte la imagen normalizada en HSV (ver figura 33.c), se determina un umbral para su gama de color y se crea una máscara con pixeles dentro de un rango definido (ver figura 33.d). Otra máscara binarizada permite la eliminación de ruido (ver figura 33.e).



Figura 33 Máscaras a) Imagen adquirida b) Imagen normalizada c) Imagen en espacio de color HSV d) Máscara con umbral de tono de piel e) Máscara sin ruido

Finalmente una transformación morfológica de cierre resulta útil para definir el contorno externo mayor logrando obtener como resultado una imagen totalmente limpia (ver figura 34.a) de la cual se obtiene el pixel blanco superior que actúa como guía y limitante de la región de interés (ver figura 34.b).

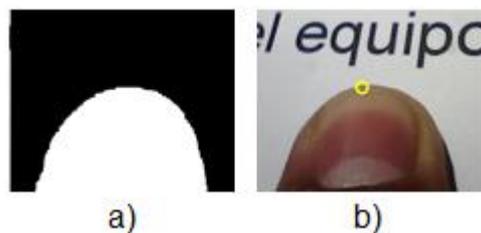


Figura 34 Resultado de la aplicación de filtros a) Máscara con transformación morfológica b) Frame original

Para determinar si la imagen capturada es válida, el histograma de la última máscara se compara con una base de datos de imágenes típicas pre almacenadas de dedos sobre un papel (ver figura 35).



Figura 35 Base de datos de dedos en contraste con un papel

- Región de interés

La región de interés toma como referencia el pixel guía para extraer una porción horizontal de imagen dentro de la cual se desarrolla la búsqueda de texto (ver figura 36.a). Se utiliza un filtro gaussiano y se aplica el método de binarización Otsu para obtener una imagen nítida como se muestra en la figura 36.b). Los contornos que se ajustan al tamaño promedio de un carácter facilitan la proyección de una línea guía como se grafica en la figura 36.c) y se agrupan en palabras (ver figura 36.d).



Figura 36 Región de interés a) Frame original b) Región de interés filtrada c) Línea guía en región interés d) Palabras en región de interés

c) Posicionamiento del dedo

El posicionamiento del dedo determina enviar señales de retroalimentación conforme a las siguientes circunstancias:

- Si la cámara no detecta la presencia del dedo de acuerdo con la comparación con la base de datos.
- Si el valor de la pendiente de la línea guía es mayor a 5° o menor a -5° se indica girar el dedo a la derecha o a la izquierda respectivamente.
- Si en la región de interés no existe al menos una palabra para enviar a procesar.
- Si la ubicación actual del dedo indica el inicio o el fin de una línea de texto en base a la primera y última palabra de la región de interés (ver figura 37).



Figura 37 Línea de texto a) Inicio b) Fin

d) Reconocimiento óptico de caracteres

Pytesseract es un módulo para el reconocimiento óptico de caracteres, toma como entrada una imagen binaria y devuelve una cadena de caracteres (ver figura 38) llamando externamente al ejecutable de Tesseract OCR.

```
Leer es tan simple → ==== RESTART: tesseract2.py  
>>> text: Leer es tan simple
```

Figura 38 Conversión de imagen a cadena de caracteres

Un subproceso que se ejecuta en paralelo al bucle principal de la aplicación gráfica permite validar las cadenas de caracteres obtenidas cada 3 segundos. El proceso de obtención de texto también es un subproceso que se ejecuta cada 0,5 segundos, tiempo aproximado que tarda el algoritmo en

convertir la imagen en cadena de caracteres. En cada ciclo se intenta eliminar las palabras o caracteres que se repiten y almacenar una única vez cada palabra obteniendo una cadena de texto concatenada desde el inicio hasta el final de la línea.

e) Conversión de texto a audio

Pyttsx es un módulo que soporta motores de texto a voz comunes como Espeak, la aplicación habilitada para TTS que permite escuchar la salida de audio resultante de la lectura del archivo temporal con caracteres no repetidos.

3.3.9. Concepto final

Se puede verificar el concepto final con todos los elementos seleccionados y previamente dimensionados en el despiece de la figura 39, y en la tabla 18 la descripción.

Tabla 18

Elementos Concepto final

N°	Parte
1	Lente wide angle
2	Soporte lente
3	Case frontal en ABS
4	Cámara web
5	Alojamiento vibradores en Ninjaflex
6	Placa electrónica
7	Case posterior en ABS
8	LDR
9	Motores vibradores de disco
10	Leds SMD 3030

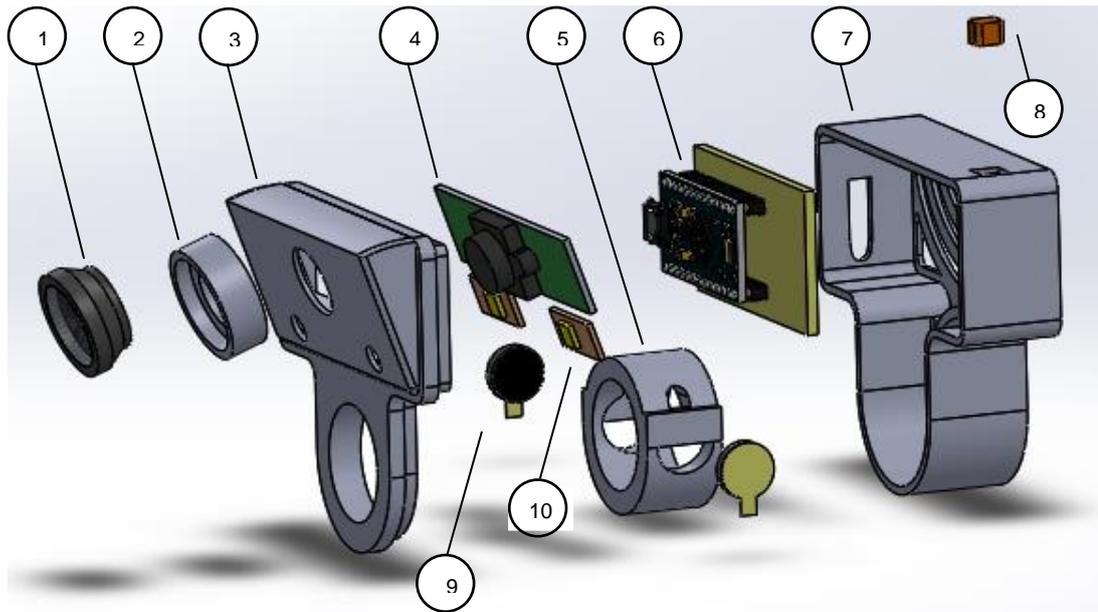


Figura 39 Vista explosionada del concepto final

Resumen

El capítulo 3 contempla el diseño del dispositivo lector de textos, que contiene 8 módulos que son: cámara, soporte, tarjeta electrónica, iluminación, control iluminación, señales de retroalimentación, controlador y proceso genérico de control.

CAPÍTULO 4

IMPLEMENTACIÓN Y PRUEBAS

El capítulo aborda la implementación del dispositivo de forma general, las pruebas aplicadas y los resultados obtenidos.

4.1 Implementación

En la figura 40 se presenta un esquema de los pasos a seguir para la implementación del proyecto.

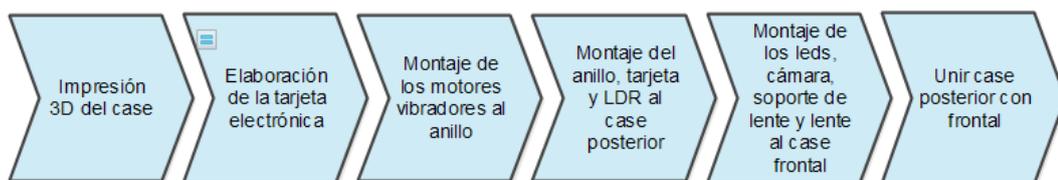


Figura 40 Piezas impresas en 3D

Se realizó una impresión 3D del case frontal, case posterior y soporte del lente en plástico ABS, el anillo para acoplar los motores vibradores se imprimió en material flexible TPE. En la figura 41 se muestran las piezas.

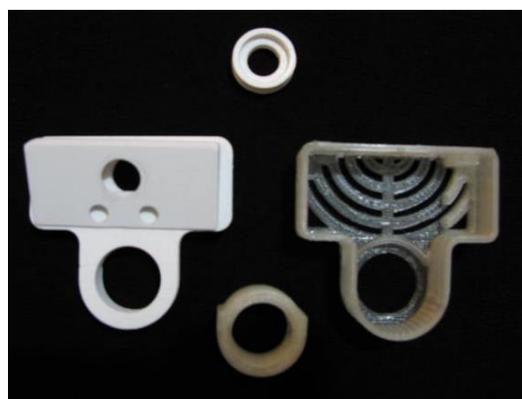


Figura 41 Piezas impresas en 3D

En la figura 42 se muestra la placa con los componentes soldados.

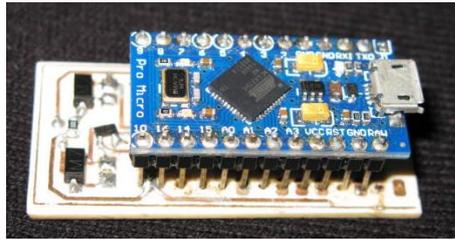
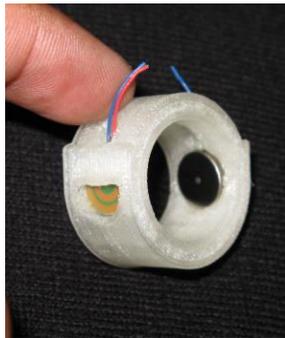


Figura 42 Elaboración PCB

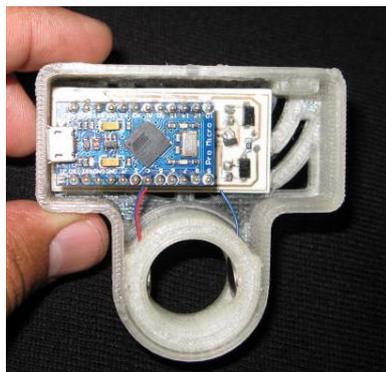
En la figura 43.a se realiza el montaje de los motores vibradores al anillo, a continuación en la figura 43.b se coloca el anillo sobre el case posterior, en la figura 43.c se fija la placa en el case posterior, finalmente se inserta la LDR en la parte superior del case posterior en la figura 43.d.



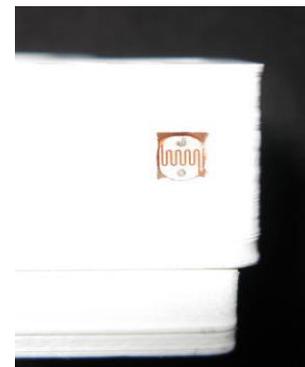
a) Vibradores al anillo flexible



b) Anillo al case posterior



c) Placa al case posterior



d) LDR a la parte superior del case posterior

Figura 43 Montaje de elementos al case

En la figura 44.a se realiza el montaje de los leds, se muestra en la figura 44.b como se posiciona la cámara, el soporte del lente se ubica en la figura 44.c y el lente como tal en la figura 44.d.



a) Led's al case frontal



b) Cámara al case frontal



c) Led's al case frontal



d) Cámara al case frontal

Figura 44 Montaje elementos al case frontal

En la figura 45 se muestra el dispositivo ensamblado por completo.



Figura 45 Dispositivo ensamblado

4.2 Pruebas

4.2.1. Dimensionales

Se realiza las pruebas dimensionales de las partes previo al ensamble del dispositivo, para corroborar si cada una de ellas cumple con las medidas planteadas en el diseño. En la tabla 19 se muestran los resultados de la prueba.

Tabla 19

Resultados prueba dimensional

Elemento	Plano	Dimensión evaluada	Resultado
Case frontal	DLT-Soporte01	Geometría	Aprobada
Case posterior	DLT-Soporte02	Geometría	Aprobada
Acople vibradores	DLT-Acople01	Geometría	Aprobada
Acople lente	DLT-Soporte01	Geometría	Aprobada

4.2.2. Ensamble y Conexiones

Posterior al proceso de implementación se obtiene un dispositivo ensamblado. La prueba que se va a realizar es para verificar que las piezas y elementos encajen para su correcta interacción.

En las figuras 46 y 47 se comprueba que el case frontal, el case posterior y el anillo para los vibradores se sujetan y acoplan.

La cámara, los leds, la fotorresistencia y la tarjeta electrónica de igual forma se ajustan a las partes respectivas del case (ver figura 43 y 44).



Figura 46 Acople del anillo y el case posterior



Figura 47 Vistas del acople entre case frontal y posterior

En la figura 48 se aprecia como los motores vibradores se posicionan en el anillo, diseñado para que exista una superficie de contacto con el dedo del usuario y a la vez puedan transmitir las señales vibratorias.

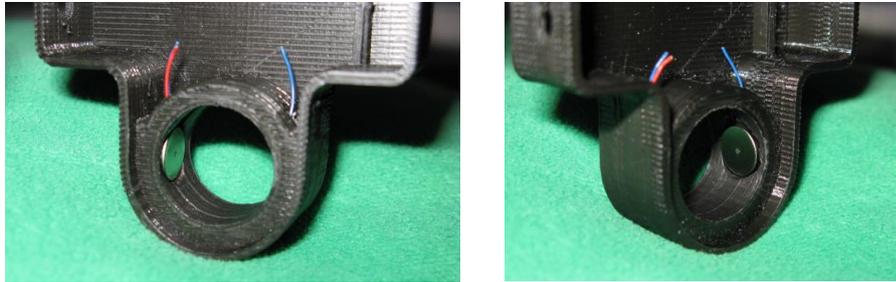


Figura 48 Acople de los vibradores al anillo

Se verifica la comunicación serial entre el controlador y el ordenador enviando las señales para que se produzca la activación de los actuadores del dispositivo. En la tabla 20 se detalla los resultados de la prueba, de esta manera se validan las conexiones internas del dispositivo hacia la tarjeta electrónica que fue previamente testada.

Tabla 20

Resultados de la prueba de conexiones

Elemento	Activación mediante comunicación serial	
	Correcto	Incorrecto
Leds	X	
Vibrador izquierda	X	
Vibrador derecha	X	
LDR	X	
Cámara	X	

4.2.3. Funcionales

Se puso a prueba cada bloque funcional del dispositivo, para validar su correcta operación se utilizó una matriz que establece los componentes y los parámetros que intervienen en la ejecución de la prueba.

a) Adquisición de imagen

En la tabla 21 se despliega la matriz para la prueba de adquisición de imágenes.

Tabla 21

Matríz para la prueba de adquisición de imágenes

Identificación	Prueba de adquisición de imágenes
Tipo	Funcional
Objetivo	Verificar el ángulo de visión de la cámara y la cantidad de texto comprendida en un fotograma
Componentes sometidos	Ordenador, cámara, lente wide angle, soporte
Parámetros que intervienen	Ángulo de visión, distancia focal, resolución de la cámara
Descripción	<ol style="list-style-type: none"> 1. Colocar el anillo en el dedo índice del usuario 2. Ejecutar la aplicación para acceder a la webcam desde el ordenador 3. Ubicar el dedo en posición de lectura 4. Contabilizar el número de palabras dentro del fotograma 5. Montar un lente wide angle sobre la cámara y volver a contabilizar el número de palabras por fotograma 6. Comparar resultados

Para la realización de la prueba se extrajo un párrafo de una noticia en línea, se adecuó éste a los márgenes establecidos para el formato A4 con fuente de texto Arial 14 e interlineado de 1,5 puntos. Cada línea de texto contiene entre 8 y 13 palabras. Se determinó que dentro del fotograma se logran encerrar entre tres y cuatro palabras aproximadamente (ver figura 49).

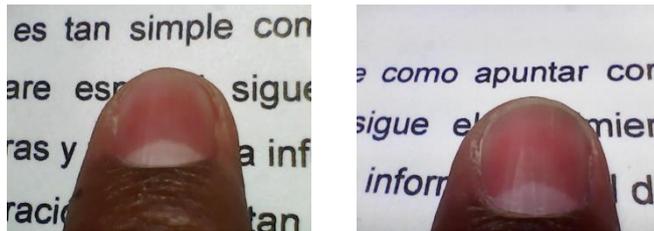


Figura 49 Captura de texto en el fotograma

Aunque la definición HD de la cámara web permite capturar imágenes con resolución máxima de 1280x720 pixeles se decidió montar un lente wide angle (Arturo Goga, 2017) para concentrar más palabras dentro del fotograma. El lente wide angle tiene como ventaja abarcar mayor porción de escena desde distancias inferiores conservando proporciones y evitando la introducción exagerada del efecto de viñeteado (ver figura 50).

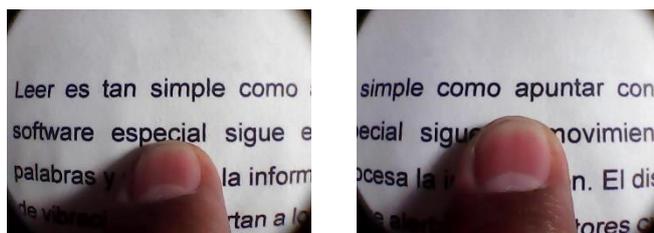


Figura 50 Captura de imagen con lente wide angle

El número de palabras dentro del fotograma asciende hasta seis, de acuerdo al algoritmo de programación una mayor cantidad de palabras procesadas por frame permite obtener porcentajes de similitud entre cadenas más altos y por consecuencia emitir audio más extenso reduciendo

las pausas por lo cual se decidió fijar de manera permanente el lente wide angle a la cámara.

b) Control de iluminación

En la tabla 22 se despliega la matriz para la prueba de control de iluminación.

Tabla 22

Matriz para la prueba del control de iluminación

Identificación	Prueba de control de iluminación
Tipo	Funcional
Objetivo	Comprobar el comportamiento del sistema de iluminación frente a diferentes ambientes
Componentes sometidos	Ordenador, controlador, led's, ldr
Parámetros que intervienen	Luz ambiental, PWM de led's
Descripción	<ol style="list-style-type: none"> 1. Colocar el anillo en el dedo índice del usuario 2. Ejecutar la aplicación para acceder a la webcam desde el ordenador 3. Encender el sistema de iluminación 4. Ubicar el dedo en posición de lectura 5. Situar el anillo en un ambiente con luminosidad baja 6. Medir los valores de entrada del sensor ldr y de salida de la señal PWM 7. Cuantificar la cantidad de lúmenes emitida por los led's 8. Repetir el paso 5, 6 y 7 en ambientes con luminosidad media y alta 9. Comparar resultados

Para la realización de la prueba se registró el funcionamiento del sistema de iluminación en ambientes cerrados con diferente intensidad de luz. Como se visualiza en la tabla 23, se comprobó que la señal de entrada del sensor Idr corresponde proporcionalmente a la variación de tensión de salida mediante PWM para los led's.

Tabla 23

Matriz para la prueba del control de iluminación

Luminosidad del ambiente	Señal de entrada [V]	Ciclo de trabajo [%]	Luz emitida [lúmenes]
Baja	4,75	95	48
Media	2,75	55	28
Alta	0,75	15	8

Se prevé que el uso del dispositivo sea en un ambiente cerrado con luminosidad media, por ello se configuró la variación del ciclo de trabajo entre el 50 y 80% de modo que se emitan de 25 a 40 lúmenes.

c) Control de retroalimentación

En la tabla 24 se despliega la matriz para la prueba de control de retroalimentación.

Se utilizó el método experimental para determinar tres diferentes patrones de vibración, la figura 51 muestra el ciclo de trabajo de cada secuencia. Cada llamada de ejecución se emite durante 600ms, tiempo suficiente para percibir la señal de retroalimentación.

Tabla 24

Matriz para la prueba del control de retroalimentación

Identificación	Prueba de control de retroalimentación
Tipo	Funcional
Objetivo	Evidenciar el comportamiento del sistema de retroalimentación con el uso del dispositivo
Componentes sometidos	Ordenador, cámara, controlador, motores, soporte
Parámetros que intervienen	Posición del dedo índice, frecuencia de vibración de los motores
Descripción	<ol style="list-style-type: none"> 1. Colocar el anillo en el dedo índice del usuario 2. Ejecutar el programa de cumplimiento de secuencias 3. Diferenciar tres diferentes secuencias de vibración variando la frecuencia de operación 4. Determinar los tiempos de funcionamiento de las secuencias de vibración 5. Comparar resultados

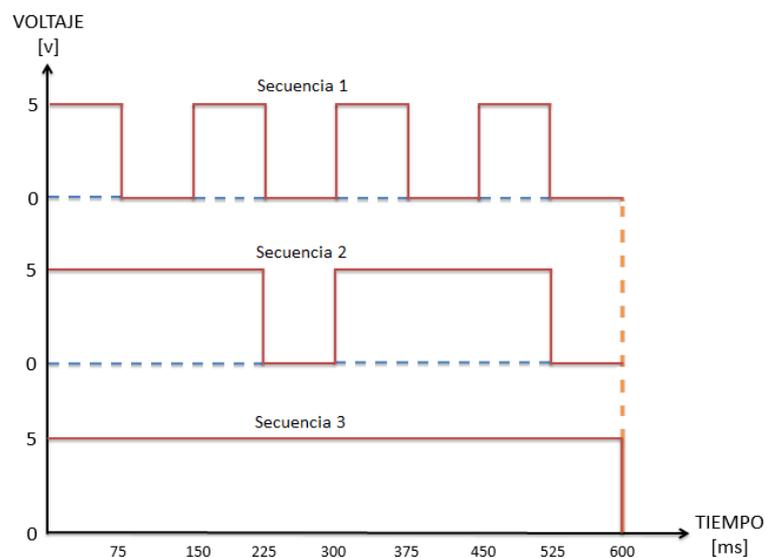


Figura 51 Ciclos de trabajo para las secuencias de vibración

La frecuencia de funcionamiento de cada secuencia permite diferenciar entre las señales de retroalimentación que se indican en la tabla 25. Para evitar la confusión del usuario se decidió usar señales auditivas que indican el inicio y el final de una línea de texto.

Tabla 25

Indicadores de las señales de retroalimentación

Indicador	Señal
Inclinación con pendiente positiva	Secuencia 1
Inclinación con pendiente negativa	Secuencia 2
Fotograma no válido	Secuencia 3
Fotograma borroso	Secuencia 3
No hay presencia de dedo índice	Secuencia 3
Inicio de línea	1 beep de 1s a 2000Hz
Fin de línea	3 beep de 0,3s a 1000Hz

d) Reconocimiento de caracteres

En la tabla 23 se despliega la matriz para la prueba de reconocimiento de caracteres.

El reconocimiento de caracteres genera inconvenientes con las letras 'i', 'j', 'l', 'o' y 'v', los números '1' y '0', los signos de admiración '¡' y '!' y los caracteres especiales '\', '|', '/'. La figura 52 muestra el ejemplo del reconocimiento erróneo de la letra 'l' como el carácter especial de barra invertida '\'.

Tabla 26

Matriz para la prueba de reconocimiento de caracteres

Identificación	Prueba de reconocimiento de caracteres
Tipo	Funcional
Objetivo	Determinar el porcentaje de acierto en el reconocimiento de caracteres
Componentes sometidos	Ordenador, dispositivo
Parámetros que intervienen	Iluminación, posición e inclinación del dedo índice
Descripción	<ol style="list-style-type: none"> 1. Colocar el anillo en el dedo índice del usuario 2. Ejecutar el programa principal 3. Ubicar el dedo en posición de lectura 4. Reconocer letras, palabras y frases que incluyan signos de puntuación y caracteres especiales 5. Proceder al reconocimiento de palabras 6. Continuar con el reconocimiento de frases 7. Evaluar resultados

del dispositivo lector → `==== RESTART: C:\tesseract2.py`
`>>> text: del dispositivo \lector`

los lectores cuanc → `==== RESTART: C:\tesseract2.py`
`>>> text: los léctores cuanc`

v simple → `==== RESTART: C:\tesseract2.py`
`>>> text: v simple`

Figura 52 Reconocimiento erróneo de caracteres

Los errores de este tipo generalmente ocurren por factores como la inclinación del dedo en el momento en el que se captura el fotograma, el almacenamiento de regiones de interés se almacena con caracteres

cortados y la presencia de ruido aún después de filtros aplicados. Hay que mencionar que Tesseract OCR presenta un margen de error de 2,5% en el reconocimiento de caracteres y 5% en el reconocimiento de palabras (Smith, 2007).

e) Tiempo de locución

En la tabla 27 se despliega la matriz para la prueba del tiempo de locución.

Tabla 27

Matriz para la prueba de tiempo de locución

Identificación	Prueba de tiempo de locución
Tipo	Funcional
Objetivo	Determinar el tiempo que tarda el dispositivo en convertir en audio una línea de texto completa
Componentes sometidos	Ordenador, dispositivo
Parámetros que intervienen	Tiempo de procesamiento
Descripción	<ol style="list-style-type: none"> 1. Colocar el anillo en el dedo índice del usuario 2. Ejecutar el programa principal 3. Ubicar el dedo en posición de lectura 4. Pasar el dedo a través de una línea de texto 5. Determinar el tiempo que tarda el dispositivo en procesar la línea de inicio a fin 6. Comparar resultados

Se realizó la prueba con un párrafo de texto extraído de la web. El párrafo contiene 9 líneas de texto en total. La tabla 28 detalla el tiempo, el

número de palabras reconocidas por línea y el tiempo que tomó su procesamiento.

Tabla 28

Tiempo de locución de líneas de texto

Línea	# Palabras procesadas	# Palabras reconocidas	Tiempo de lectura [s]
1	6	6	27
2	12	10	47
3	10	9	43
4	11	9	41
5	10	8	41
6	9	7	47
7	8	8	43
8	10	8	46
9	4	4	21

Se logró determinar que para que se efectuó una buena lectura, la velocidad a la que debe moverse el dedo a través del texto no debe superar los 10 milímetros por segundo.

Se establece entonces que el porcentaje promedio de lectura de una línea de texto es 44 segundos lo que corresponde a 14 palabras por minuto. Un valor muy alejado de la tasa media de expresión de una persona que oscila entre 80 y 150 palabras por minuto (Presentación Revista, 2017).

4.2.4. Computacionales

En la tabla 29 se despliega la matriz para la prueba de requerimiento computacional.

Para la realización de la prueba se lanzó la aplicación que contiene el algoritmo principal sin otros programas en primer plano. Se determinó que los recursos consumidos durante la ejecución del algoritmo se encuentran entre valores usuales como se muestra en la tabla 30.

Tabla 29

Indicadores de las señales de retroalimentación

Identificación	Prueba de requerimiento computacional
Tipo	Computacional
Objetivo	Diagnosticar los recursos utilizados durante la ejecución de la aplicación en el ordenador
Componentes sometidos	Ordenador, dispositivo
Parámetros que intervienen	Porcentaje de uso del microprocesador, uso de memoria física
Descripción	<ol style="list-style-type: none"> 1. Colocar el anillo en el dedo índice del usuario 2. Ejecutar el programa principal 3. Ubicar el dedo en posición de lectura 4. Utilizar el monitor de recursos del ordenador para evaluar el comportamiento de la aplicación 5. Relejar resultados

Las características que posee el ordenador satisfacen el coste computacional del dispositivo, los requerimientos del aplicativo no llegan a superar el 60% de uso del CPU ni el 80% de la memoria RAM; lo cual advierte que el microprocesador del ordenador no se sobrecarga con picos de trabajo.

Tabla 30

Recursos del ordenador

Recurso	Detalle
Velocidad del procesador	2.2 Ghz
Memoria RAM	4 Gb
Uso del CPU	16%
Uso de memoria física	581 Mb

En la figura 53 se visualiza el porcentaje de uso del CPU, valor que indica cuanto de la capacidad del procesador está en uso durante la ejecución del programa.

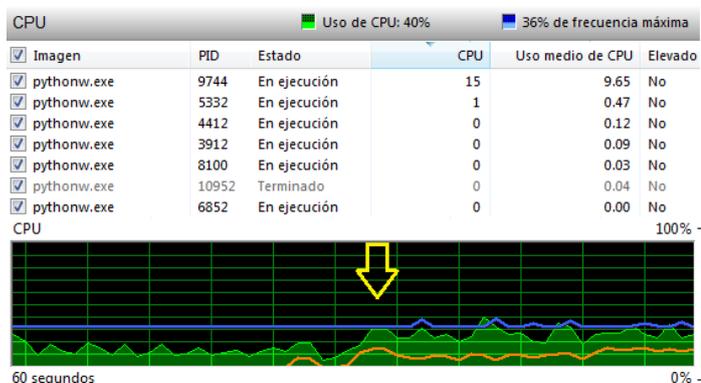


Figura 53 Uso del CPU de la aplicación

La figura 54 muestra el incremento de uso de memoria RAM al momento de iniciar el proceso de la aplicación.

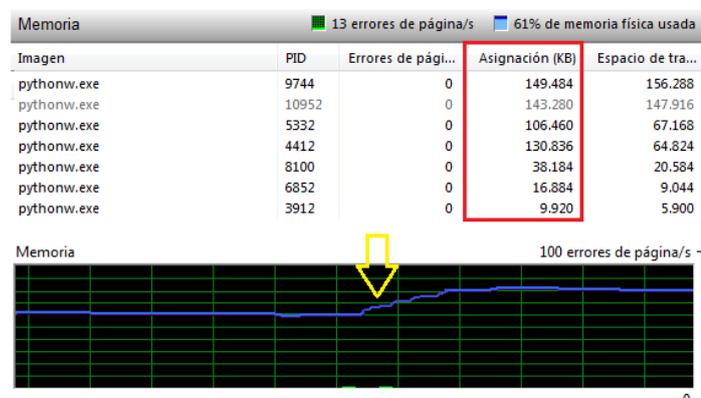


Figura 54 Uso de memoria RAM de la aplicación

Resumen

En el capítulo 4 se describieron los pasos para realizar la implementación del dispositivo lector de textos. Se realizaron pruebas dimensionales, de ensamble, conexiones, funcionales, computacionales y se mostraron los resultados obtenidos.

CAPÍTULO 5

CONCLUSIONES Y RECOMENDACIONES

A continuación se presentan las conclusiones y recomendaciones del trabajo realizado.

5.1 Conclusiones

- Se ha diseñado e implementado un dispositivo para personas con deficiencia visual moderada, capaz de leer textos impresos en letra Arial de mínimo 14 puntos de altura e interlineado de 1,5. La adquisición de la imagen se logra utilizando una cámara Microsoft LifeCam HD 3000, el procesamiento se lo realiza aplicando técnicas y filtros proporcionados por la librería de visión artificial OpenCV. Se seleccionó a Tesseract OCR para el reconocimiento de caracteres y a Espeak como conversor TTS para sintetizar y reproducir el audio generado.
- Mediante un software de diseño CAD se realizó el modelado 3D del case del dispositivo. Se imprimió la parte frontal y posterior en material ABS y el anillo que aloja los motores vibradores en material flexible por tener contacto con el dedo del usuario. Al implementar todas las piezas confirmamos el tamaño compacto del dispositivo que con medidas de 6,5 cm de ancho, 6 cm de alto y 4 cm de espesor soporta todos los elementos internos.
- Se entrenó al motor Tesseract OCR con un diccionario de 3000 palabras para el reconocimiento de caracteres en español. Se realizaron modificaciones al código fuente de la librería Pytesseract

disponible solo para el idioma inglés para enlazar el motor entrenado con el entorno de programación.

- Se diseñó un sistema de iluminación en el dispositivo para compensar la cantidad de luz durante la adquisición de imágenes, se verificó que la luz percibida por la cámara para el tratamiento de imágenes es un factor crítico por lo que se requiere de un promedio de 32 lúmenes para conseguir un ambiente adecuado donde el algoritmo pueda trabajar sin interferencias.
- El conversor TTS Espeak actuó como conversor fonético. Se necesitó del motor Mbrola para ejecutar el sintetizador, se seleccionó la voz más comprensible en idioma español de la base de datos del proyecto colaborativo Mbrola.

5.2 Recomendaciones

- Se recomienda utilizar técnicas avanzadas de programación como Super Vector Machine para el reconocimiento de caracteres que permita detectar texto de cualquier fuente.
- Se sugiere realizar pruebas con materiales distintos a los habituales para el prototipado 3D, de manera que se pueda encontrar uno que brinde características de ergonomía y comodidad al usuario y no disperse las señales vibratorias a través de todo el dispositivo.
- La aplicación desarrollada no requiere de un coste computacional elevado para un ordenador sin embargo, se recomienda reducir el tiempo de ejecución y el espacio de memoria utilizado eliminando redundancias, parametrizando procesos, reordenando la estructura del código, etc. de modo que los recursos necesarios sean mínimos

y se pueda ejecutar el procesamiento en un minicomputador como la Raspberry Pi convirtiendo al dispositivo portátil.

- El uso de threads para la ejecución en paralelo de los procesos de reconocimiento y validación de texto logra suponer que las tareas se ejecutan al mismo tiempo sin embargo la programación multiproceso puede resultar una alternativa más eficiente permitiendo incluso reducir el tiempo de procesamiento.
- El procesamiento de imagen para la detección del dedo índice como cursor guía a través del texto podría omitirse implementando en el dispositivo un led láser que marque un punto fijo sobre el fotograma y mantenga constante el área de captura de la región de interés.
- Indagar en el mercado acerca de cámaras para visión artificial con mejores tecnologías, una cámara con enfoque automático y ángulo de visión superior a los 100° permitiría abarcar más texto dentro de un mismo fotograma disminuyendo el tiempo de lectura por línea.
- Se sugiere la implementación de lentes wide angle o eyefish en la cámara para aumentar el porcentaje de captura de texto en el fotograma, su uso incurre la formación de fotogramas semiesféricos con efecto viñeteado que deben corregirse con procedimientos matemáticos como la transformación lineal.

BIBLIOGRAFÍA

- Asamblea Nacional. (21 de febrero de 2017). Obtenido de <http://www.asambleanacional.gob.ec/es/blogs/comision-de-soberania-integracion-relaciones-internacionales-y-6>
- Adichip. (5 de enero de 2017). *Adichip*. Obtenido de <http://www.adichip.com.ar/es/interfaces-humano-maquina/1661-vibrating-mini-motor-disc.html>
- Agelectronica. (15 de febrero de 2017). *Agspecinfo*. Obtenido de <http://www.agspecinfo.com/pdfs/A/ADA-1201.PDF>
- Agüero, P. D. (2012). *Síntesis de voz aplicada a la traducción voz a voz*. Barcelona: Universidad Politécnica de Cataluña (UPC).
- Alsina, A. (10 de Septiembre de 2012). *Tiflolibros*. Recuperado el 11 de Febrero de 2016, de Tiflolibros: <http://www.tiflolibros.com.ar/contenido/lectura.htm>
- Amaya Amaya, J. (2011). *Sistemas de Información Gerenciales*. Bogotá: Ecoe Ediciones.
- Apkmonk. (12 de Junio de 2014). Obtenido de <http://www.apkmonk.com/app/com.sensotec.knfbreader/>
- Arada Vera, Á. (2014). *Instalación y parametrización del software*. Málaga: IC Editorial.
- Arteaga Jiménez, G., & Elizalde Ruiz, C. (2011). *La creación de un taller de danza para niños ciegos*. Puebla: Universidad de las Américas de Puebla.
- Arturo Goga. (23 de febrero de 2017). *LENTE GRAN ANGULAR*. Obtenido de <https://www.arturogoga.com/as-es-como-se-vende-un-lente-gran-angular-que-es-un-wide-angle-lens/>
- Asamblea Nacional. (2008). *Constitucion Política Del Ecuador*. Quito.

- Asamblea Nacional. (7 de abril de 2016). *Blog de la comisión de soberanía, integración, relaciones internacionales y seguridad integral* . Obtenido de <http://www.asambleanacional.gob.ec/es/blogs/comision-de-soberania-integracion-relaciones-internacionales-y-6>
- Barbuzano, J. (15 de Abril de 2015). *Openmind*. Recuperado el 11 de Febrero de 2016, de Openmind: <https://www.bbvaopenmind.com/tecnologia-para-invidentes-mas-alla-del-braille/>
- Bejarano, I., & Sánchez, R. (2015). *Reconocimiento de caracteres mediante imágenes en contadores de gas en entornos reales*. Madrid.
- Braben, D. (1 de Abril de 2016). *About: The Raspberry Pi Foundation*. Obtenido de <https://www.raspberrypi.org/>
- Cambridge in Colour. (14 de Febrero de 2017). *Camera histograms: tones & contrast*. Obtenido de <http://www.cambridgeincolour.com/tutorials/histograms1.htm>
- Castejón, J. L., & Navas, L. (2013). *Aprendizaje, desarrollo y disfunciones: implicaciones para la enseñanza en la Educación Secundaria*. Alicante: Editorial Club Universitario ECU.
- Cazorla, M., & Viejo, D. (2015). JavaVis: An integrated computer vision library for teaching computer vision. *Computer Applications in Engineering Education*, 258-267.
- Centro Integrado Politécnico ETI. (2016). *Etitudela*. Recuperado el 18 de febrero de 2017, de <http://www.etitudela.com/celula/downloads/visionartificial.pdf>
- Cetpro. (31 de agosto de 2012). *Computacion del Cetpro*. Obtenido de Computacion del Cetpro: http://computaciondelcetproalcidessalomon.blogspot.com/2012_08_01_archive.html

- Chw Net. (25 de abril de 2013). *Chile Hardware*. Obtenido de <http://www.chw.net/2013/04/eye-talk-anteojos-que-leen-para-los-ciegos/>
- Conadis. (23 de octubre de 2016). *Estadísticas de personas con discapacidad*. Obtenido de <http://www.consejodiscapacidades.gob.ec/estadistica/index.html>
- Correa , C., Rueda, H., & Arguello, H. (2014). *Síntesis de Voz por Concatenación de Difonemas para el Español de Colombia*. Bucaramanga,: Universidad Industrial de Santander .
- Cosenza, M. (1 de Abril de 2016). *Acerca: Open Source Computer Vision*. Obtenido de www.opencv.org/about
- CRC_online. (diciembre de 2016). *Centros de Reconocimiento de Conductores*.
- Cuartielles, D. (1 de Abril de 2016). *Arduino*. Obtenido de <https://www.arduino.cc/en/Main/ArduinoBoardMega2560>
- Deambrogio, V. (2013). *Impresión 3D al descubierto*. Alicante: Minidocs.
- Dedutel. (10 de 8 de 2016). *Desarrollo en educación tecnológica*. Obtenido de <http://www.dedutel.com/pdf/STRATASYS.pdf>
- eSpeak. (4 de enero de 2017). *eSpeak text to speech*. Obtenido de <http://espeak.sourceforge.net/>
- Fernández, C., Junquera, C., & Ribes, M. (2013). *Cuidadores de la Comunidad de Navarra*. Sevilla: MAD S.L.
- Fernández, I. (2014). *Discapacidad visual*. Guayaquil: Dirección Provincial de Educació del Guayas.
- Fundación Las Dos Orillas. (2 de diciembre de 2015). *Las 2 orillas* . Obtenido de <http://www.las2orillas.co/vallenato-el-sonido-ilumina-la-vida-de-80-ninos-ciegos-de-colegio-de-bogota/>
- Gallego, J. (2012). *PCPI - Montaje de componentes informáticos*. Barcelona: Editex.

- García Jiménez, J. L., Gómez Morales, T., Navarro, J., & Agueda Casado, E. (2012). *Automoción: estructura del vehículo*. Madrid: Paraninfo.
- García, F. (30 de 11 de 2009). Reconocimiento de objetos en una cocina con una webcam. Madrid, Madrid, España.
- Garzón Jeréz, C. M. (2015). *DISEÑO E IMPLEMENTACIÓN DE UN DISPOSITIVO CON SISTEMA DE COMUNICACIÓN BRAILLE E INTERFAZ ANDROID PARA LA INTERCOMUNICACIÓN DE PERSONAS SORDOCIEGAS*. Latacunga: Universidad de las Fuerzas Armadas ESPE.
- Genius. (2015). *Genius online store* . Obtenido de Genius online store : <http://us.geniusnet.com/scans>
- Groover, M. (2012). *Fundamentos de Manufactura Moderna*. México: Pearson Education.
- Grupo de Investigación CATAM. (23 de septiembre de 2013). *Procesamiento de Imágenes Digitales, Universidad de Sevilla*. Obtenido de <http://alojamientos.us.es/gtocoma/pid/tema4.pdf>
- Guerra Pérez, F., & Vallejo Martín, H. (28 de enero de 2017). *Acido poliláctico*. Obtenido de <http://www.eis.uva.es/~macromol/curso08-09/pla/Pag%20web/acido%20polilactico.html>
- Hermosa Donate, A. (2012). *Electrónica aplicada*. Barcelona: Marcombo S.A.
- Herrera Martínez, M. (2013). *Reconocimiento de marca y modelo de vehículos en entornos*. Puebla: Inaoe. Obtenido de <https://inaoe.repositorioinstitucional.mx/jspui/bitstream/1009/263/1/HerreraMA.pdf>
- Hidalgo, I., & Sánchez, R. (Junio de 2015). Reconocimiento de caracteres mediante imágenes en contadores de gas en entornos reales. Madrid, España.
- Hilera, J., Romero, J., & Gutiérrez, J. (1996). Sistema de reconocimiento de caracteres (OCR) con redes neuronales. *I Jornadas sobre Inteligencia*

Artificial, Control y Sistemas Expertos, Universidad de Alcalá, (págs. 171-183). Madrid.

Imbaquingo Esparza, D. E. (22 de febrero de 2017). *Repositorio UTN*. Obtenido de

<https://www.google.com.ec/url?sa=t&rct=j&q=&esrc=s&source=web&cd=3&ved=0ahUKEwjv0MD0uaPSAhVFfiYKHe1OADIQFggmMAI&url=http%3A%2F%2Frepositorio.utn.edu.ec%2Fbitstream%2F123456789%2F559%2F2%2FCAP%25C3%258DTULO%2520II%2520ALGORITMOS.doc&usg=AFQjCNF2Zc6ILZNXpfl>

Impresoras3D.com. (10 de agosto de 2016). *Impresoras3d.com*. Obtenido de Impresoras3d.com:

<https://impresoras3d.com/blogs/noticias/102883975-tipos-de-impresoras-3d>

Infaimon. (2017). *Cámaras de visión industrial*. Recuperado el 18 de febrero de 2017, de <http://www.infaimon.com/es/camaras-industria>

Intel. (3 de febrero de 2017). *Intel*. Obtenido de <http://www.intel.es/content/www/es/es/nuc/overview.html>

IPC. (2012). *Generic Standard on Printed Board Design*. Illinois.

ISO. (9 de septiembre de 2015). Norma Técnica UNE-EN ISO 9241-9: 2015. Madrid.

Kalpakjian, S., & Schmid, S. (2012). *Manufactura Ingeniería y Tecnología*. México: Pearson Educación.

Kultura 2.0. (21 de febrero de 2017). *OCR: tecnología para el reconocimiento óptico de caracteres en una imagen*. Obtenido de http://www.kultura.ejgv.euskadi.eus/r46-19123/es/contenidos/informacion/kultura2_0_prestakuntza/es_k20_form/adjuntos/pildora-OCR-2.pdf

Latorre, Á., Bisseto, D., & Teruel, J. (2012). *Trastornos y dificultades del desarrollo*. Valencia: PUV S.L.

- Loayza Jaramillo, M. C. (2016). *Investigación de un sistema electrónico con detección de obstáculos para personas no videntes*. Latacunga: Universidad de las Fuerzas Armadas ESPE.
- López Rincón, C. (2015). *Diseño e implementación de un sistema automático de análisis documental*. Madrid: Dpto. de Lógica y Filosofía de la Ciencia Facultad de Filosofía.
- Luxlite. (5 de enero de 2017). *Luxlite*. Obtenido de http://www.ecoluxlite.com/home/index.php?option=com_content&view=article&id=231:i-que-son-los-leds-smd&Itemid=426&lang=en
- MakerShop. (3 de febrero de 2017). *MakerShop*. Obtenido de <https://makershopbcn.com/abs-vs-pla-que-diferencia-existe-entre-estos-dos-filamentos-para-impresora-3d>
- Makeuseof. (21 de febrero de 2017). *The 3 Best Free OCR Tools to Convert Your Files Back Into Editable Documents*. Obtenido de <http://www.makeuseof.com/tag/3-free-ocr-tools-convert-files-editable-documents/>
- Martínez Cuervo, F. (2011). *Auxiliares Sanitarios*. Sevilla: MAD S.L.
- Maturana, J. (20 de Febrero de 2014). Obtenido de Xataka: <http://www.xataka.com/perifericos/estas-son-las-tecnologias-de-impresion-3d-que-hay-sobre-la-mesa-y-lo-que-puedes-esperar-de-ellas>
- Maximenko, E. (28 de diciembre de 2016). *Apuntes y ejercicios de matemáticas IPN*. Obtenido de <http://esfm.egormaximenko.com/matlab.html>
- Microsoft. (5 de enero de 2017). *Accesorios Microsoft*. Obtenido de <https://www.microsoft.com/accessories/es-es/products/webcams/lifecam-hd-3000/t3h-00002>
- Microsoft. (15 de Febrero de 2017). *Microsoft Developer Network*. Obtenido de [https://msdn.microsoft.com/en-us/library/ms720410\(v=vs.85\).aspx](https://msdn.microsoft.com/en-us/library/ms720410(v=vs.85).aspx)

Norris, C. (4 de enero de 2017). *Ubuntu - ES*. Obtenido de <http://www.ubuntu-es.org/user/7871#.WG2b3FPhB0w>

Núñez, M. A. (2011). La deficiencia visual. *III Congreso “La Atención a la Diversidad en el Sistema Educativo”* (pág. 3). Salamanca: Universidad de Salamanca. Instituto Universitario de Integración en la Comunidad (INICO).

On Semiconductor. (17 de enero de 2017). *brown.edu*. Obtenido de [http://www.brown.edu/Departments/Engineering/Courses/En162/Data Sheets/2N3903-D.pdf](http://www.brown.edu/Departments/Engineering/Courses/En162/Data%20Sheets/2N3903-D.pdf)

OpenCV. (21 de Febrero de 2017). *Open Source Computer Vision 3.2.0*. Obtenido de http://docs.opencv.org/3.2.0/d4/d73/tutorial_py_contours_begin.html

OpenCV. (21 de Febrero de 2017). *Open Source Computer Vision 3.2.0*. Obtenido de http://docs.opencv.org/3.2.0/d7/d4d/tutorial_py_thresholding.html

OpenCV. (21 de Febrero de 2017). *OpenCV 3.0.0-dev documentation*. Obtenido de http://docs.opencv.org/3.0-beta/doc/py_tutorials/py_imgproc/py_morphological_ops/py_morphological_ops.html

Pardal , J., & Pena, M. (2014). *Arduino y node.js*. Pontevedra: Inforede SL.

Patagoniatec. (5 de enero de 2017). *Patagoniatec*. Obtenido de <http://saber.patagoniatec.com/arduino-pro-micro-atmega32u4-arduino-argentina-ptec-leonardo/>

Presentación Revista. (24 de Febrero de 2017). *UNCOMO*. Obtenido de <https://salud.uncomo.com/articulo/como-contar-las-palabras-por-minuto-en-un-discurso-4272.html>

QueCamaraFLex. (5 de enero de 2017). *Qué cámara?* Obtenido de <http://quecamarareflex.com/como-funciona-el-sensor-de-una-camara-digital/>

Restrepo, J. (2013). *Computadoras para todos*. Bogotá: Archi SA.

- Reyes, J. M. (17 de agosto de 2016). *La nube artística*. Obtenido de http://www.lanubear artistica.es/dibujo_artistico_2/unidad5/DA2_U5_T1_Contenidos_v01/11_visin_binocular_y_tridimensionalidad.html
- Roy, R. (1 de Abril de 2016). *ODROID*. Obtenido de http://www.hardkernel.com/main/products/prdt_info.php?g_code=G145457216438
- Sánchez, C., & Consuegra, V. (2010). *Reconocimiento Óptico de Caracteres (OCR)*. Madrid.
- Sandip, R., Amitava, K., Mrinmoy, M., Subhajt, M., Satwika, M., & Subhadip, B. (2012). *Recognition of Handwritten Roman NUMerals Using Tesseract Open Source OCR Engine*. Kolkata: Techno India College of Technology .
- Sénderov, S. (14 de Abril de 2013). *Actualidad RT*. Recuperado el 11 de Febrero de 2016, de Actualidad RT: <https://actualidad.rt.com/ciencias/view/91758-gafas-ciegos-eyetalk>
- Shilkrot, R. (13 de marzo de 2014). *Fluid Interfaces*. Obtenido de <http://fluid.media.mit.edu/projects/fingerreader>
- Silicon. (8 de agosto de 2016). *Impresión 3D: Qué materiales usar*. Obtenido de <http://www.silicon.es/impresion-3d-que-materiales-usar-y-donde-comprarlos-50135>
- Sinha, U. (21 de Febrero de 2017). *AI Shack*. Obtenido de <http://aishack.in/tutorials/normalized-rgb/>
- Smith, R. (2007). *An Overview of the Tesseract OCR Engine*. Obtenido de <https://research.google.com/pubs/pub33418.html>
- Smith, R. (23 de Septiembre de 2007). *An Overview of the Tesseract OCR Engine*.
- Sparkfun. (23 de febrero de 2017). *Pro Micro - 5V/16MHz*. Obtenido de <https://www.sparkfun.com/products/12640>

- Stallman, R. (20 de Marzo de 2016). *El sistema operativo GNU*. Obtenido de <http://www.gnu.org/gnu/linux-and-gnu.es.html>
- TCTS Lab. (4 de enero de 2017). *Théorie des Circuits et Traitement du Signal*. Obtenido de <http://tcts.fpms.ac.be/synthesis/mbrola.html>
- Thompson, M. (1 de Abril de 2016). *Raspbian*. Obtenido de <https://www.raspbian.org/>
- UDOO Org. (4 de enero de 2017). *Udoo Shop*. Obtenido de <http://shop.udoo.org/other/home.html>
- Ulrich, K. T., & Eppinger, S. D. (2013). *Diseño y desarrollo de productos*. México: McGRAW-HILL/INTERAMERICANA EDITORES, S.A. DE C.V.
- Van Rossum, G. (1 de Abril de 2016). *Acerca: Python*. Obtenido de Python: www.python.org/about
- Velasco Imedio, I. (2013). *Procesos grupales y educativos en el tiempo libre infantil y juvenil*. Málaga: IC Editorial.
- Vicomtech. (21 de febrero de 2017). *Visual interaction & communication technologies*. Obtenido de <http://www.vicomtech.org/t1/e7/procesamiento-de-imagenes>
- Wayjun Technology. (16 de enero de 2017). *Wayjun Technology*. Obtenido de <http://www.wayjun.com/>
- Wikipedia. (6 de Abril de 2016). *Wikipedia, la enciclopedia libre*. Obtenido de https://es.wikipedia.org/wiki/Visi%C3%B3n_artificial#/media/File:CVoverview2-ES.jpg
- Zahumenszky, C. (14 de Marzo de 2015). *Gizmodo*. Recuperado el 11 de Febrero de 2016, de Gizmodo: <http://es.gizmodo.com/este-anillo-del-mit-permite-a-personas-ciegas-leer-text-1691463893>
- Zambrano, B. (2013). *Técnicas y campos de la inteligencia artificial*. Panamá: Universidad Latina de Panamá.
- Zaragoza Baquero, I. (2012). *Discapacidad y deficiencia*. Andaluz: Lulu S.A.:

ANEXOS