

# Informe Proyecto Final - Lector Automatizado

Bryan Fallas L. / Michael Fallas L. / Carlos Coronel R. / Sergio Alfaro E. / Mayron Jiménez

**Resumen—Implementación de un sistema de lectura automatizada capaz de capturar imágenes de textos y realizar el proceso necesario para su reproducción en audio del mismo; así como el cambio automático de páginas de un libro.**

## I. DESCRIPCIÓN DEL PROYECTO

El presente documento describe el proyecto desarrollado para el curso de Microcontroladores, el cual consiste en un lector de libros automatizado que lleva a cabo el proceso necesario para la reproducción en audio a partir de una imagen capturada de la página en lectura. Además, este sistema es capaz de emplear las acciones necesarias para el paso automático de hojas para continuar la lectura del libro. El proceso de lectura del libro está basado en el uso de una Raspberry Pi, la cual es la encargada de llevar a cabo las diferentes funciones para esto: software que permita la manipulación de la cámara para la captura de imágenes, el reconocimiento óptico de caracteres (OCR) con el fin de convertir la imagen en texto y, finalmente, el convertidor de texto a voz (TTS) que permita la reproducción del texto en audio. Se emplea el lenguaje de programación Python para esta tarjeta. A su vez, la Raspberry Pi estará en comunicación con un Arduino por medio de los puertos USB, el cual es el encargado de llevar a cabo todo el proceso de control de los diferentes tipos de motores para el cambio de páginas del libro. Este último es programado utilizando el lenguaje de programación C++.

## II. OBJETIVOS A ALCANZAR

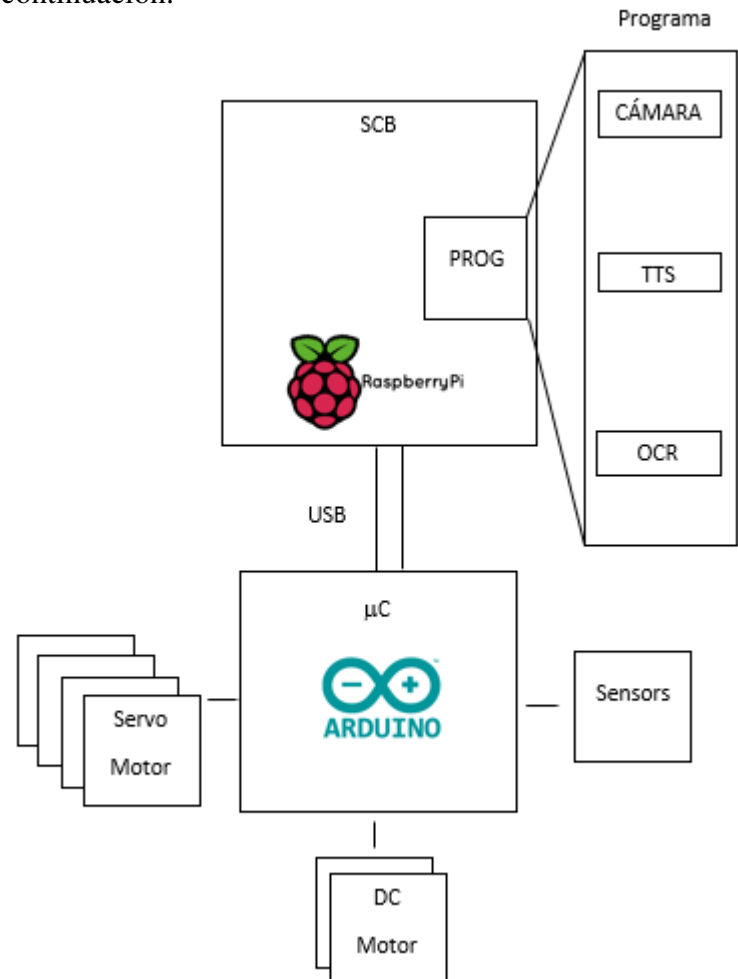
Para lograr el correcto funcionamiento del sistema se plantean las siguientes metas a alcanzar:

- Lograr la captura de una imagen por medio de una cámara conectada a la Raspberry Pi, así como el control de captura de estos por medio de esta tarjeta.
- Implementar la programación necesaria para el reconocimiento óptico de caracteres (OCR) de una imagen capturada de manera eficiente.

- Alcanzar la reproducción eficaz en audio de un texto (TTS).
- Diseñar e implementar un sistema que permita el paso eficiente de hojas de un libro.
- Lograr la comunicación serial de la Raspberry Pi con un Arduino.

## III. DIAGRAMA DEL PROYECTO

El diagrama del sistema implementado se muestra a continuación:



## IV. LISTADO DE COMPONENTES REQUERIDOS

Para implementar el sistema se requiere contar con:

- Tarjeta Raspberry Pi (Modelo B preferiblemente).
- Módulo de cámara para Raspberry Pi.
- Tarjeta de memoria micro SD 8GB (o mayor).
- Fuente de alimentación para la Raspberry Pi.
- Parlantes o audifonos.
- Tarjeta Arduino Mega 2560 (o equivalente).
- Módulo controlador de motores compatible con Arduino.
- Motor DC.
- Motor servo.
- Cable USB.
- Cables para conexión.

## V. DIVISIÓN DEL TRABAJO

Para la eficiente ejecución del proyecto, se hará una división en tres subgrupos de trabajo de los cuales, cada uno tendrá sus diferentes tareas y objetivos a cumplir. El primer subgrupo estará compuesto por: Bryan Fallas y Michael Fallas. Este grupo estará a cargo de:

- Control y uso de la cámara.
- Reconocimiento óptico de caracteres (OCR).

El segundo subgrupo estará compuesto únicamente por Carlos Coronel, el cual la tarea a cargo es:

- Implementar el TTS (Text to Speech) en el sistema.

Y, finalmente, el tercer subgrupo estará integrado por: Sergio Alfaro y Mayron Jiménez; en donde la tarea asignada es:

- Control de los diferentes motores y sensores del sistema para el paso automático de las hojas del libro o texto.

## VI. LIBRERÍAS UTILIZADAS

Para llevar a cabo el proyecto, se utilizaron las siguientes librerías:

- Python-Picamera.
- Tesseract OCR.
- eSpeak.
- Serial-Arduino
- AFmotor.h
- Servo.h

## VII. RESULTADOS OBTENIDOS

El comportamiento del sistema en general es bastante positivo. Con respecto a la cámara, se tuvieron que realizar algunos ajustes en el módulo del

mismo con el fin de ajustar el lente para el enfoque automático de ella. Con esto se logran capturar imágenes con la calidad deseada. Sin embargo, esta no permite la buena resolución con un tamaño de fuente inferior a 12 (esto con pruebas realizadas con fuentes de tipo Arial). Con respecto al OCR, la librería utilizada realiza la correcta conversión a palabras, números y símbolos; sin embargo, el sistema tiene problemas para el reconocimiento de caracteres propios del idioma español, tales como tildes en las letras, así como la letra 'ñ'. Otro aspecto a señalar del OCR, es el fallo al intentar la conversión a textos cuando las imágenes capturadas tienen cierto grado de inclinación. Por otra parte, la librería utilizada para el TTS, eSpeak, presenta errores al momento de reproducir textos grandes (superior a los 30 caracteres reproducidos); por lo que se opta por realizar diferentes ajustes a la librería, entre el cual destaca la subdivisión de la reproducción del texto en pequeños extractos; cuyos símbolos de subdivisión fueron definidos previamente. Con esto se logra la lectura del texto con la fluidez deseada. Con respecto a la parte del Hardware y la comunicación serial entre las tarjetas Arduino y Raspberry Pi, se presentan diversos inconvenientes; dado que no se completa el paso efectivo de las páginas; así como la correcta comunicación para la activación de los motores.

## VIII. ASPECTOS A MEJORAR

Con el fin de optimizar el sistema, entre los aspectos que pueden ser mejorados se encuentran:

- Mayor velocidad en la conversión de los caracteres a texto; o bien, implementar un sistema de multi-tarea para optimizar el tiempo que tarda el sistema.
- Mejorar la resolución de las imágenes con el fin de capturar imágenes de libros con letra más pequeña y lograr mediante ella una correcta conversión a texto de la misma.
- Implementar una voz menos robótica en el TTS para mejorar el entendimiento de lo que se vaya a reproducir.
- Mejoras en la parte del hardware y control de los motores a emplear para el paso de hojas.

## IX. CONCEPTOS APRENDIDOS

- La librería eSpeak no tiene la capacidad de reproducción de textos largos; por lo que para

su correcto funcionamiento hay que dividir el texto completo en extractos más pequeños y reproducir cada uno de estos por separado.

- Es posible ajustar el enfoque y resolución del módulo de la cámara de la Raspberry Pi al girar el lente.
- La librería tesseract no es apropiada para la conversión de textos que contienen características propias del idioma español; tales como palabras tildadas o la letra 'ñ'.
- Es posible ajustar los parámetros de la cámara mediante la librería Python-Picamera, tales como brillo, sombras, saturación, entre otros.
- El lenguaje de programación Python es un lenguaje práctico y sencillo de utilizar para programar.