

CAPÍTULO

7

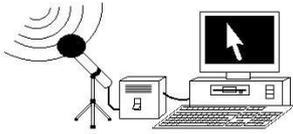
CONCLUSIONES

7.1 AGRADECIMIENTOS

7.2 BIBLIOGRAFÍA

Las pruebas realizadas con los prototipos diseñados para crear un sistema autónomo para accionamiento por voz de un ratón de ordenador proporcionaron unos resultados satisfactorios que ponen de relevancia la bondad de las soluciones obtenidas. La memoria del proyecto incluye la definición de los dos prototipos. Al primero de ellos (prototipo principal) se le han dedicado los cinco primeros capítulos, por constituir de los dos el sistema más elaborado al utilizar una red neuronal como núcleo básico del sistema de reconocimiento. Toda la carga de investigación y de desarrollo se ha aplicado al prototipo principal, y a modo de sistema de pruebas, durante el desarrollo de la implementación para comprobar el buen funcionamiento del sistema diseñado, se realizó un segundo prototipo donde el sistema de reconocimiento no era una red neuronal. En este diseño auxiliar de pruebas el reconocimiento se realizaba comparando los parámetros obtenidos de la señal de audio captada con los parámetros patrones de las cinco vocales, no estando estos parámetros patrones asignados, sino que se adquieren al inicio del sistema solicitando al usuario, a través del display de leds, la pronunciación de las cinco vocales. Este segundo prototipo no constituye la solución final, pues como se ha mencionado, ha sido un resultado residual en el proceso de investigación y desarrollo, pero se ha incluido la descripción del segundo prototipo en el capítulo sexto de la memoria del proyecto por constituir en sí mismo un diseño de suficiente entidad, fiabilidad y operabilidad.

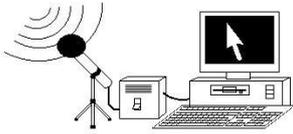
Debido a las restricciones en memoria y a la capacidad física del dispositivo FPGA para implementar un circuito de mayor complejidad la red neuronal utilizada sólo dispone de 32 neuronas. Estas condiciones no han favorecido el proceso de aprendizaje de la red neuronal para sintonizar con las características vocálicas. En esta situación, los pesos sinápticos de las neuronas se han obtenido con el algoritmo de aprendizaje de la red a partir de la exposición a múltiples pronunciaciones de un sólo usuario (yo, en este caso), de modo que el sistema mantiene cierta dependencia con el usuario. Un dispositivo FPGA de mayor capacidad que el utilizado permitiría una red neuronal de más neuronas, capaz de disociar características vocálicas de diversas personas y hacer el diseño independiente del usuario. Por otro lado, el segundo prototipo es completamente independiente del usuario gracias a que los patrones vocálicos son captados del mismo usuario que lo utiliza durante la inicialización del dispositivo.



En las pruebas finales de los prototipos y gracias también a la utilidad “Click-N-Type” de Lake Software que, como ya se mencionó en el capítulo quinto, permite emular el teclado desde el ratón, se realizaron algunas de las operaciones más comunes que se realizan en un ordenador, sólo con la pronunciación de las cinco vocales (lectura de correos electrónicos, escritura y envío de correos electrónicos, escritura de cartas, navegación por la red). Queda así patente la utilidad que el dispositivo ofrece a personas con imposibilidad de movimientos. La utilización del dispositivo amplía las posibilidades de comunicación de los minusválidos con incapacidad de movimientos al permitir la escritura de documentos y el uso de todas las posibilidades que ofrece un ordenador personal. Abre una ventana al mundo a personas con incapacidad, al constituir una herramienta que posibilita navegar por Internet sólo con el uso de la voz.

En este sentido, merece la pena la dedicación y el trabajo desarrollado durante la investigación y el desarrollo del proyecto, a pesar de los problemas encontrados en el camino. No ha sido sencilla la fase de investigación para encontrar en primer lugar una parametrización adecuada, y posteriormente, una técnica funcional de reconocimiento. Muchas pruebas las realicé en la plataforma MATLAB® hasta encontrar un camino viable, incluso llegué a tener en un estado muy avanzado algoritmos de reconocimiento utilizando una red neuronal de funciones de base radial (RBF). Durante la fase de desarrollo el problema principal ha sido la imposibilidad de implementar los diversos diseños elaborados en la FPGA, pues se desbordaban las capacidades internas del dispositivo. No soy un estudiante de la especialidad de electrónica, sino de control de procesos, por lo que el trabajo ha sido tedioso y lento hasta que se han depurado mis técnicas de diseño. Han sido necesarias múltiples versiones y varias modificaciones desde el principio para cambiar completamente la estructura del circuito. Incluso en la fase de implementación han surgido problemas que han detenido el avance en el proyecto al ser incapaz en un principio de poner en funcionamiento el CODEC utilizado (el resultado era siempre 0). Tras revisar todos los puntos de la placa de expansión tuve que realizar algunas soldaduras para que el circuito cumpliera las especificaciones del propuesto en el manual de referencia del CODEC (ejemplo: en los amplificadores operacionales la tierra no estaba fijada). En ese momento el CODEC ya mostraba dos valores (0 y -1). Fijé todas las señales de control del CODEC a un valor '1' lógico y observé que una de ellas no alcanzaba el umbral necesario para que el CODEC tomase el valor '1' (1.2 V y el umbral es de 2 V). Modifiqué el pin de la FPGA asignado a esa señal por el pin que establecía la tensión más alta en la señal, 1.6 V en este caso, pero parece ser suficiente para que el dispositivo funcione, aunque no en todas las ocasiones. Finalmente, noté que si realizo una descarga al ordenador de algunos valores almacenados en la memoria FLASH de la tarjeta XSA aumenta la probabilidad de que el dispositivo funcione.

Al constituir el circuito un sistema de reconocimiento de vocales se podría aplicar el mecanismo de reconocimiento empleado a otros ámbitos donde fuera útil el control por voz de un dispositivo. Sólo habría que especificar un protocolo que generara los comandos necesarios a partir de la pronunciación de las distintas vocales. Incluso podría estudiarse un sistema de reconocimiento de personas, basado en la similitud de los coeficientes cepstrales asociados a una determinada vocal. Quizás pueda utilizarse el material que he elaborado en uno de esos dos procedimientos.



7.1 AGRADECIMIENTOS

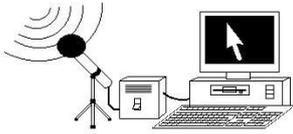
A D. Miguel Ángel Aguirre Echanove, por la ayuda aportada en la realización del proyecto.

Al Departamento de Electrónica de la Escuela Superior de Ingenieros de Sevilla, por proporcionar el material necesario en la realización de los prototipos.

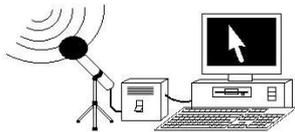
A mis padres, Luis y María, a mis hermanos, Luis Miguel y Eva María, y a Conchi, por su apoyo incondicional y comprensión.

7.2 BIBLIOGRAFÍA

- "SPOKEN. Language processing". Dr Raj Reddy. Editorial Prentice Hall.
- "Tratamiento de señales en tiempo discreto". Alan V. Oppenheim, Ronald W. Schaffer. Editorial Prentice Hall.
- "Tratamiento digital de señales. Principios, algoritmos y aplicaciones". John G. Proakis, Dimitris G. Manolakis. Editorial Prentice Hall.
- "Speech and audio signal processing. Processing and perception of speech and music". Ben Gold, Nelson Morgan. Editorial John Wiley & sons, INC.
- "Speech Processing. A dynamic and optimization-oriented approach". Li Deng, Douglas O'Shaughnessy. Editorial Board.
- "Reconocimiento de voz y fonética acústica". Jesús Bernal Bermúdez, Jesús Bobadilla Sancho, Pedro Gómez Vilda. Editorial Ra-Ma.
- "Redes neuronales y sistemas borrosos". B. Martín del Brío, A. Sanz Molina. Editorial Ra-Ma.
- "Neural Networks. A comprehensive foundation". Simon Haykin. Editorial Macmillan.
- "Redes neuronales artificiales. Fundamentos, modelos y aplicaciones". José R. Hilera, Víctor J. Martínez. Editorial Ra-Ma.



- "VHDL. Lenguaje estándar de diseño electrónico". Lluís Terés, Yago Torroja, Serafin Olcoz, Eugenio Villar. Editorial Mac Graw Hill.
- "VHDL, lenguaje para síntesis y modelado de circuitos". Fernando Pardo, José A. Boluda. Editorial Ra-Ma.
- Revista "Electrónica práctica", número 9. Editorial Prensa Técnica.
- http://www.gte.us.es/~aguirre/Web_plugs/CODEC.htm
- <http://www.computer-engineering.org/ps2protocol/>
- <http://www.computer-engineering.org/ps2mouse/>
- <http://www.atmel.com/>
- <http://www.texas-instruments.com/>
- <http://www.xilinx.com/>
- <http://www.xess.com/>
- <http://www.lakefolks.org/cnt>



ÍNDICE

CAPÍTULO 1. INTRODUCCIÓN. Pág. 1

CAPÍTULO 2. PARAMETRIZACIÓN. Pág. 8

2.1 INTRODUCCIÓN. Pág. 8

2.2 EL CONVERTOR ANALÓGICO-DIGITAL. Pág. 10

2.3 TRANSFORMADA RÁPIDA DE FOURIER. Pág. 16

2.3.1 Cálculo directo de la DFT.

2.3.2 Metodología de “divide y vencerás” para el cálculo de la DFT.

2.3.3 Algoritmo para la FFT base-2.

2.3.4 Utilización particular del algoritmo FFT en base-2 para el cálculo de la Transformada de Fourier de 512 muestras de la señal de audio.

2.3.5 Simulación MATLAB®.

2.4 MÓDULO DE LA TRANSFORMADA RÁPIDA DE FOURIER. Pág. 47

2.4.1 Simulación MATLAB®.

2.5 ESCALADO MEL. Pág. 50

2.5.1 Simulación MATLAB®.

2.6 LOGARITMO EN BASE DOS DE LOS COEFICIENTES MEL. Pág. 57

2.6.1 Simulación MATLAB®.

2.7 COEFICIENTES CEPSTRALES. Pág. 61

2.7.1 Simulación MATLAB®.

CAPÍTULO 3. RECONOCIMIENTO. Pág. 67

3.1. RECONOCIMIENTO PARAMÉTRICO Pág. 67

3.2 RECONOCIMIENTO MEDIANTE UNA RED NEURONAL. Pág. 67

3.2.1 Estructura de un sistema neuronal artificial

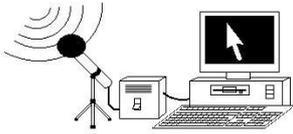
3.2.2 Modelo estándar de una neurona artificial.

3.2.3 Arquitectura de redes neuronales.

3.2.4 Modos de operación: Recuerdo y Aprendizaje.

3.2.4.1 Fase de aprendizaje. Convergencia.

3.2.4.2 Fase de recuerdo o ejecución.



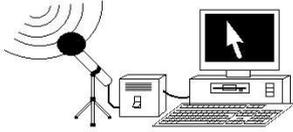
- 3.3 MAPAS AUTOORGANIZADOS. Pág. 73
 - 3.3.1 Modelo de mapas autoorganizados (Kohonen, 1982).
 - 3.3.2 Algoritmo de aprendizaje.
 - 3.3.2.1 Consideraciones prácticas: ritmo de aprendizaje y función de vecindad.
- 3.4 SIMULACIÓN MATLAB®. Pág. 78
 - 3.4.1 Aprendizaje de la red neuronal en MATLAB®.
 - 3.4.2 Recuerdo por la red neuronal en MATLAB®.
 - 3.4.3 Mapeado final.

CAPÍTULO 4. EMULACIÓN RATÓN PS/2. Pág. 107

- 4.1 INTRODUCCIÓN Pág. 107
- 4.2 PROTOCOLO VOCÁLICO. Pág. 108
- 4.3 PROTOCOLO SERIE PS/2. Pág. 109
 - 4.3.1 Introducción.
 - 4.3.2 El conector.
 - 4.3.3 Descripción general.
 - 4.3.4 Comunicación dispositivo-host.
 - 4.3.5 Comunicación host a dispositivo.
 - 4.3.6 Entradas, resolución y escalado.
 - 4.3.7 Paquete de datos de movimiento.
 - 4.3.8 Modos de operación.
 - 4.3.9 Modo reset.
 - 4.3.10 Modo corriente.
 - 4.3.11 Modo remoto.
 - 4.3.12 Modo página.
 - 4.3.13 Extensiones de Intellimouse.
 - 4.3.14 Juego de comandos.
 - 4.3.15 Inicialización.
 - 4.3.15.1 Encendido/inicialización ratón PS/2 estándar.
 - 4.3.15.2 Encendido/inicialización ratón PS/2 intellimouse (scrolling).
 - 4.3.15.3 Encendido/inicialización ratón PS/2 intellimouse (scrolling y 5 botones).
 - 4.3.16 Características imprescindibles.

CAPÍTULO 5. DISEÑO E IMPLEMENTACIÓN. Pág. 135

- 5.1 INTRODUCCIÓN. Pág. 135
- 5.2 UNIDAD FÍSICA DEL SISTEMA: FPGA. Pág. 135



5.3 LA TARJETA DE DESARROLLO XSA. Pág. 136

5.4 PLACA DE EXPANSIÓN PARA LA TARJETA XSA-50. Pág. 139

5.5 VHDL. Pág. 140

5.6 UNIDADES DE DISEÑO. Pág. 141

5.7 FICHERO UCF. Pág. 430

5.8 RECURSOS DEL SISTEMA. Pág. 434

5.9 PRUEBA FINAL. Pág. 436

CAPÍTULO 6. SEGUNDA SOLUCIÓN. Pág. 439

6.1 INTRODUCCIÓN. Pág. 439

6.2 PARAMETRIZACIÓN. Pág. 439

6.3 RECONOCIMIENTO. Pág. 440

6.4 EMULACIÓN RATÓN PS/2. Pág. 441

6.5 DISEÑO E IMPLEMENTACIÓN. Pág. 441

6.5.1 Hardware utilizado.

6.5.2 Fichero UCF.

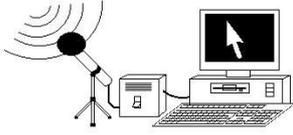
6.5.3 Recursos del sistema.

6.5.4 Prueba final.

CAPÍTULO 7. CONCLUSIONES. Pág. 486

7.1 AGRADECIMIENTOS. Pág. 488

7.2 BIBLIOGRAFÍA Pág. 488



PROYECTO FIN DE CARRERA
Ingeniería de Telecomunicaciones
**SISTEMA AUTÓNOMO PARA ACCIONAMIENTO
POR VOZ DE UN RATÓN DE ORDENADOR**

Dedicado a:

Mis padres, mis hermanos
y Conchi