

CONSTRUCCIÓN DE LA PROTESIS TACTIL “PAUPER TANGO”

Antonio García Rozo*, Roberto Perry C**, Javier Villegas***

RESUMEN

En este artículo se presentan los aspectos mas relevantes del desarrollo de una prótesis táctil para individuos con deficiencias auditivas (Pauper Tango), inspirada completamente en el uso de información suprasegmental construida de manera concreta sobre señales de habla en el idioma español.

1. INTRODUCCIÓN

Las personas con deficiencias auditivas pueden lograr comunicarse a través del lenguaje de señas o lectura de labios pero ninguno de estos métodos permite transmitir información importante pero propia del lenguaje hablado (p.e. entonación); el trabajo que se ha desarrollado es la construcción de una prótesis que transmita información relevante en el idioma español a través de vibraciones en la piel y cuya ubicación permita al usuario su utilización diaria de forma inadvertida para las personas que lo rodean. Esta opción vibro táctil se ha considerado como una alternativa interesante para emitir información sonora en contraposición con las costosos implantes cocleares

2. INFORMACIÓN RELEVANTE

La selección que se ha hecho de la información que puede resultar importante está basada en el papel de los suprasegmentales [1] en el lenguaje español, este tipo de información permanece oculta para los individuos con deficiencias auditivas pero resulta ser de gran importancia para la comunicación a nivel de ora-

ciones. El sentido completo de una frase u oración puede cambiar. La figura 1 muestra el oscilograma y la extracción de la frecuencia fundamental para la misma frase dicha primero como afirmación y luego como pregunta. En la actualidad las prótesis vibro táctiles comerciales existentes (TACTAID) enfocan sus esfuerzos a la transducción de fonemas, este tipo de información puede ya inferirse a través de la lectura de labios. En este proyecto se ha hecho un análisis lingüístico cuidadoso y la información que se ha seleccionado es la frecuencia fundamental y la intensidad.

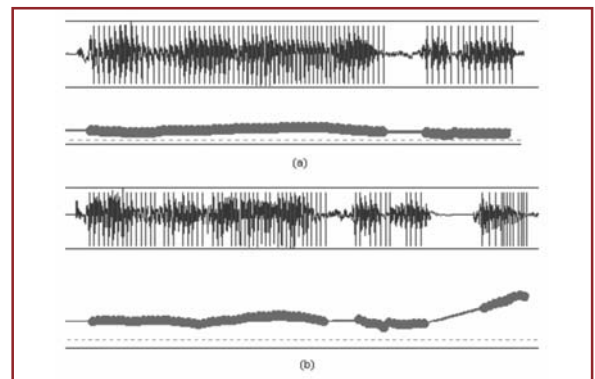


Figura 1. Extracción de la frecuencia fundamental para (a) Una afirmación (b) Una pregunta

* Profesor Titular Uniandes angarcia@uniandes.edu.co

** Profesor Universidad Nacional de Colombia rperry@uniandes.edu.co

*** Investigador Uniandes ja-ville@uniandes.edu.co - Centro de Microelectrónica Universidad de los Andes

3. DESCRIPCIÓN DEL SISTEMA

El dispositivo inicialmente recibe una señal de voz a través de un micrófono y es necesario que la extracción de la información se realice en tiempo real, para esto se utilizó el DSP56002 ([3],[4]) un procesador digital de señales de Motorola de 24 bits que fue manipulado a través de la tarjeta de desarrollo DSP56002EVM [5]. Esta plataforma cuenta además con un conversor análogo digital y digital análogo sigma delta de 16 bits estéreo y con flexibilidad en la frecuencia de muestreo utilizada [2]. Una vez que la información se ha procesado adecuadamente se envía a dos vibradores independientes que están colocados en las muñecas del usuario.

4. EXTRACCIÓN DE LA INFORMACIÓN

El sistema consiste en dos líneas de procesamiento prácticamente independientes, uno de ellas destinada a detectar la frecuencia fundamental y la otra a extraer y procesar el valor de la energía de la señal; todos los cálculos fueron realizados en tiempo real sobre bloques de datos no solapados de 512 muestras usando una frecuencia de muestreo de 16K Hz y 16 bits de cuantización.

Para la detección de la frecuencia fundamental son necesarias varias etapas (Fig. 2 parte inferior):

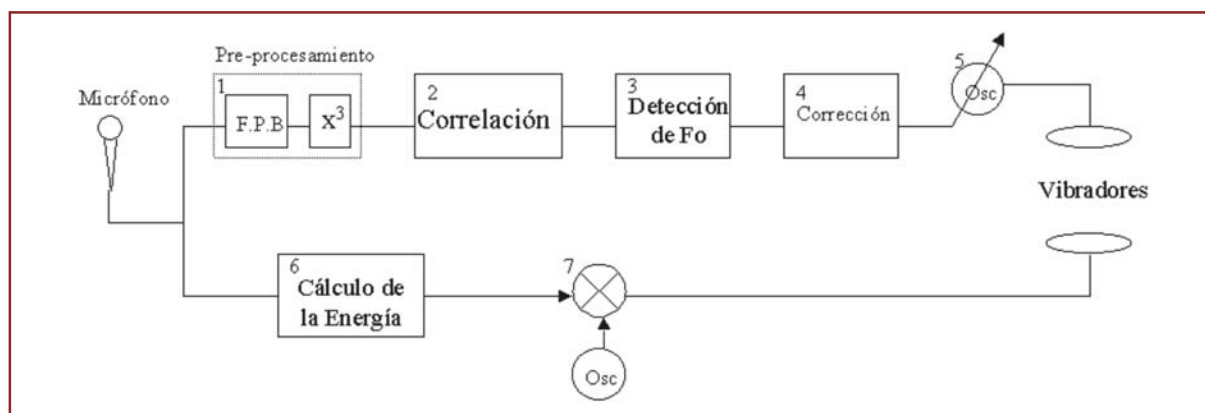


Figura 2. Diagrama en bloques del sistema

Pre-procesamiento.

El trabajo previo que se hace sobre la señal consiste en un filtrado pasa bajos con un filtro digital de orden 3 con frecuencia de corte cercana a 2k Hz para realzar la señal en la banda de interés, seguido de una elevación a una potencia impar (3) que permite resaltar los picos de la señal de voz conservando el signo.[6]

Correlación

Históricamente para encontrar las periodicidades de la señal de voz que representan la frecuencia fundamental se han utilizado distintas alternativas ([6],[8] y [9]), un grupo grande de estas alternativas lo conforman los métodos basados en la autocorrelación y este fue el

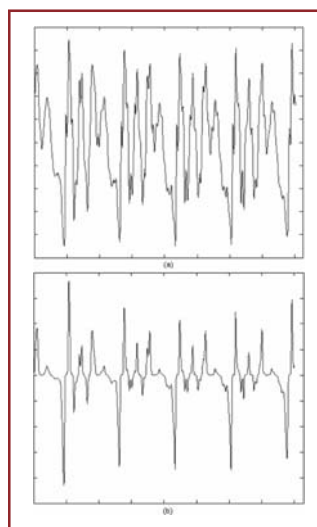


Figura 3. (a) Señal original (b) Señal filtrada y elevada al cubo.

camino seguido en el presente trabajo. Por la cantidad de puntos utilizados en cada bloque (512) el algoritmo de la autocorrelación es computacionalmente más eficiente si se realiza en el dominio de la frecuencia. Para calcular la autocorrelación de la señal de entrada se utiliza primero una FFT en base dos con diezmado en tiempo de 512 puntos[7]; el resultado se multiplica por el conjugado para obtener la densidad espectral de potencia y finalmente sobre esta secuencia se calcula la FFT inversa. El cálculo de la frecuencia fundamental se hace midiendo la distancia entre dos picos máximos de la función de autocorrelación[8].

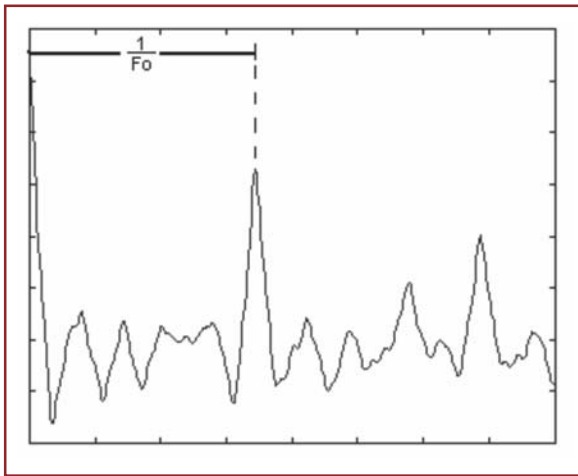


Figura 4. Autocorrelación de la señal de la figura 3b donde se ve el pico que indica la frecuencia fundamental.

Corrección

En los algoritmos de detección de frecuencia fundamental es común encontrar errores por duplicación de la frecuencia detectada (pitch doubling). Una de las soluciones tradicionales es pasar una o dos veces los resultados a través de un filtro mediana [9]; sin embargo la limitación de trabajar en tiempo real hace imposible considerar esta alternativa y en su lugar se lleva a cabo una etapa de validación donde el sistema se asegura de que no existan cambios repentinos entre el valor anterior y el último medido.

Oscilador de frecuencia variable.

Una onda senoidal de amplitud constante y con

frecuencia igual a la detectada es sintetizada utilizando un oscilador digital con la técnica de doble integrador. Esta estrategia es conveniente porque la necesidad de tener frecuencia variable hace poco práctico el empleo de tablas y la cuantización de los coeficientes impide hacer uso de otras estructuras para sistemas de segundo orden.

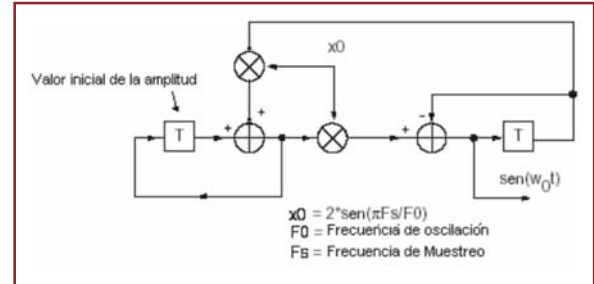


Figura 5. Oscilador digital con la técnica de doble integración.

Para procesar y extraer el valor de la energía se utilizan los siguientes bloques(Fig. 2. Parte Inferior):

Calculo de la Energía.

En esta etapa se halla la suma del cuadrado de bloques de cada 512 muestras.

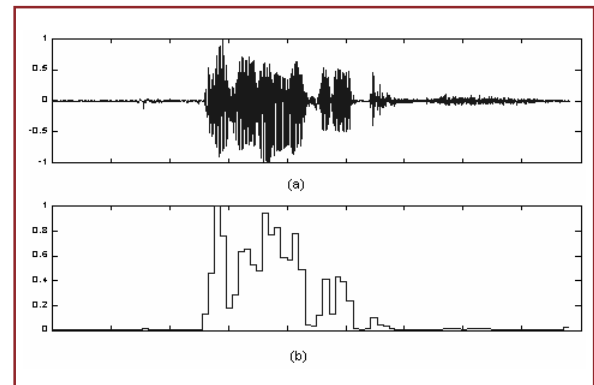


Figura 6. (a) Señal original (b) Señal de energía en bloques de 512 muestras

Modulación

Una señal que consiste en la suma de tres ondas senoidales de 125 Hz, 250 Hz y 500 Hz es modulada en amplitud de acuerdo al valor de la energía anteriormente calculado.

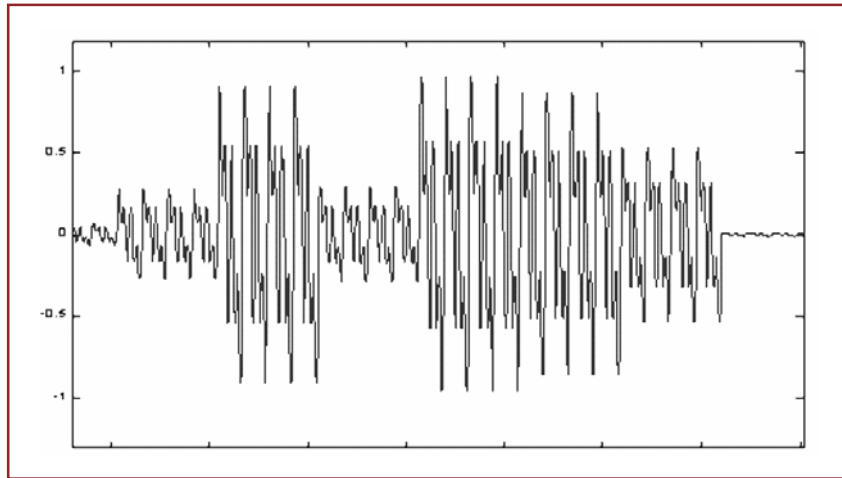


Figura 7. Suma de señales senoidales modulando una parte de la señal de la figura 6 b.

5. CONCLUSIONES

La extracción de la frecuencia fundamental es una labor complicada por las peculiaridades de la señal de voz y el extenso rango dinámico que se debe cubrir. La exigencia adicional de tener que realizarla en tiempo real hace de esta una tarea mucho más complicada y vale la pena explorar alternativas distintas a la correlación [9].

Para la etapa de corrección no es conveniente el uso de filtro mediana (muy usado en aplicaciones que no requieren el resultado en tiempo real) ya que el retardo es notorio incluso si el filtro es de ventana de longitud 3.

Las modificaciones a las que deba ser sujeto el sistema dependen directamente de los resultados que sean obtenidos en terapia.

6. REFERENCIAS

- [1] Beckman, Mary, “Notas sobre los suprasegmentales” The Ohio State University, Columbus, 1992.
- [2] CRYSTAL Semiconductor Corporation. *CS4215 16-bit multimedia audio codec*. 1993.
- [3] MOTOROLA, Digital Signal Processing Division, *DSP56000 Digital Signal Processor Family Manual*.
- [4] MOTOROLA, Digital Signal Processing Division, *DSP56002 24 bits Digital Signal Processor User Manual*.
- [5] MOTOROLA, Digital Signal Processing Division, *DSP56002EVM Quick Start Guide*.
- [6] Witten, H., *Principles of computer speech*, Academic Press, 1982.
- [7] MOTOROLA, Digital Signal Processing Division, *Implementation on Fast Fourier Transform on Motorola's Digital Signal Processors*.
- [8] Dutoit, T., *An introduction to text to speech synthesis*, Kuwer A.P. 1997.
- [9]. Gold, B, Morgan N, *Speech and audio signal processing*, Wiley & Sons, New York, 2000.