

RESTAURACIÓN DE VOZ DETERIORADA POR RUIDO DE PULSOS

Jorge Brito Collantes; Melissa Cortez Vega

Ing. Patricia Chávez

Ing. Juan Avilés

Facultad de Ingeniería en Electricidad y Computación

Escuela Superior Politécnica del Litoral

Campus Gustavo Galindo. Km. 30.5 vía Perimetral,

Apartado 09-01-5863 Guayaquil-Ecuador

jbrito@fiec.espol.edu.ec; mcortez@fiec.espol.edu.ec;

pchavez@fiec.espol.edu.ec; javiles@fiec.espol.edu.ec;

RESUMEN

Las señales de audio son susceptibles al ruido de pulsos en muchos casos ya sea la interferencia electromagnética en los sistemas móviles, mala sincronización en grabaciones digitales o incluso el ruido de rayado en un disco LP viejo. Algunos programas para la cancelación de ruido se han desarrollado, pero son caros y usan algoritmos complejos. Este paper presenta un método adaptativo para la supresión de ruido de pulsos en señales de voz específicamente. Los pulsos son eliminados por la aplicación del Median Filter solo en las regiones contaminadas de la señal, evitando pérdidas de frecuencias altas en las partes no afectadas.

ABSTRACT

Audio signals are susceptible to pulse noise in many cases such as electromagnetic interference in mobile systems, poor synchronization in digital recording or even scratch noise of an old LP record to name a few. A number of software packages for noise cancellation have been developed but they are expensive and use complex algorithms. This paper presents an adaptive method for suppressing pulse noise specifically from voice signals. Pulses are removed by selectively applying median filtering to contaminated signal regions only, avoiding high frequency losses in the undamaged sections.

1. INTRODUCCIÓN

El ruido en el procesamiento de señales es cualquier señal indeseada, que se introduce en los datos a través de cualquier sistema eléctrico utilizado para almacenamiento, transmisión y/o procesamiento.

La reducción del ruido es el proceso de eliminación de un porcentaje considerable del ruido de la señal.

Todos los dispositivos de grabación digital o analógica tienen rasgos que los hacen susceptibles al ruido; cuando se utiliza la tecnología de grabación analógica,

grabaciones de sonido presentan un tipo de ruido conocido como siseo de cinta; además las señales de audio en un rango amplio de aplicaciones están sujetas a ruidos de pulso, ya sea por el rayado de un disco LP, mala sincronización en grabaciones digitales, o los efectos de propagación de los sistemas móviles [1]; la deterioración de la señal de audio puede ser severa.

2. METODOLOGÍA Y MARCO TEÓRICO

La supresión del ruido de pulsos se realizará a través de un Median Filter (MF).

El MF es un filtro no lineal que utiliza una técnica para la suavización de la señal y supresión del ruido de pulsos. En el caso de una dimensión se compone de una ventana deslizante de un número de elementos impares a través de la señal, reemplazando la muestra del centro por la mediana de las muestras en la ventana [2].

Se quiere presentar la implementación de un método adaptativo fácil y eficiente basado en una predicción lineal y filtrado [3] para la cancelación de un ruido de pulso y obtener una considerable mejoría en la calidad de la señal.

El MF es deshabilitado salvo en las regiones donde los pulsos han sido detectados.

$$y_n = \begin{cases} m_n & \text{if } g_n > 0 \\ x_n & \text{otherwise} \end{cases}$$

Donde:

$$m_n = \text{med}\{x_{n-k}, \dots, x_n, \dots, x_{n+k}\}$$

Es el valor de mediana a través de una ventana de longitud $N=2k+1$ muestras x_n y y_n son la entrada y secuencia de salida respectivamente, g_n es una señal de conmutación que se proporciona en el momento correcto para que el MF se active cuando existan pulsos. Es conocido que los MFs suprimen los pulsos en anchos menores a:

$$k = \frac{N-1}{2} \quad [4].$$

Para una detección de pulso más sólida de la señal de entrada, se obtiene la segunda derivada discreta para que estos pulsos del ruido de la señal aumenten en amplitud.

$$D^2\{x_n\} = x_{n-1} - 2x_n + x_{n+1} \quad [5].$$

3. IMPLEMENTACIÓN

El software utilizado es **Simulink** (herramienta de MATLAB), desarrollado por The MathWorks, es una plataforma comercial para modelar, simular y analizar sistemas dinámicos multidominio el cual posee un juego de herramientas incorporado de gran alcance en el tratamiento de señales.

Su interfaz gráfica es una herramienta de diagramas de bloques personalizables y un conjunto de bibliotecas de bloques. Que ofrece una estrecha integración con el resto del entorno de MATLAB.

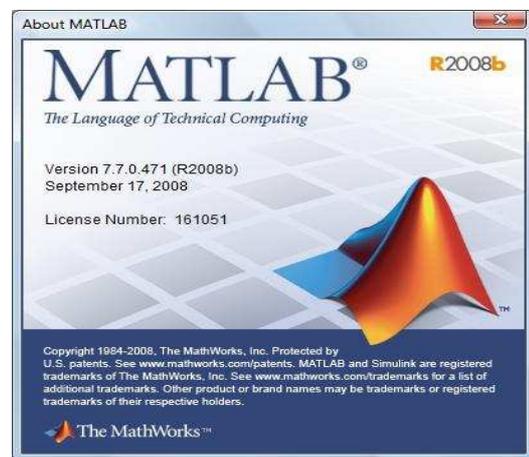


Fig. 1. Versión de MATLAB usada

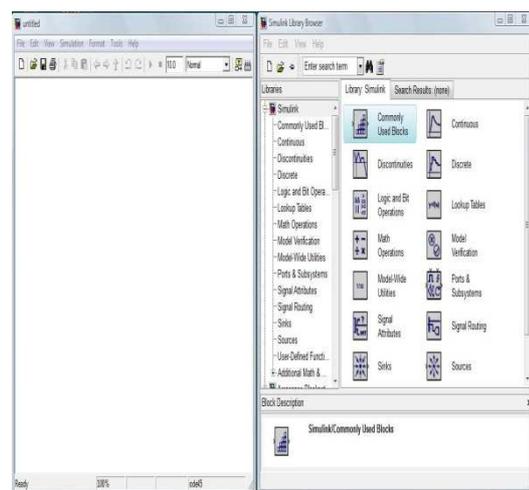


Fig. 2. Interfaz gráfica de Simulink

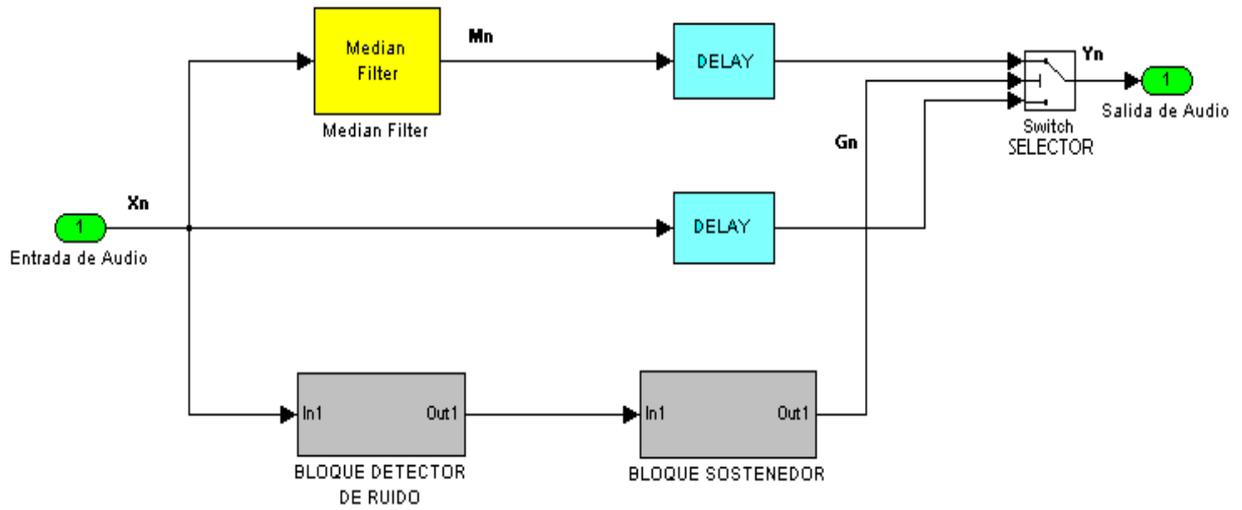


Fig. 3. Diagrama de Bloques general

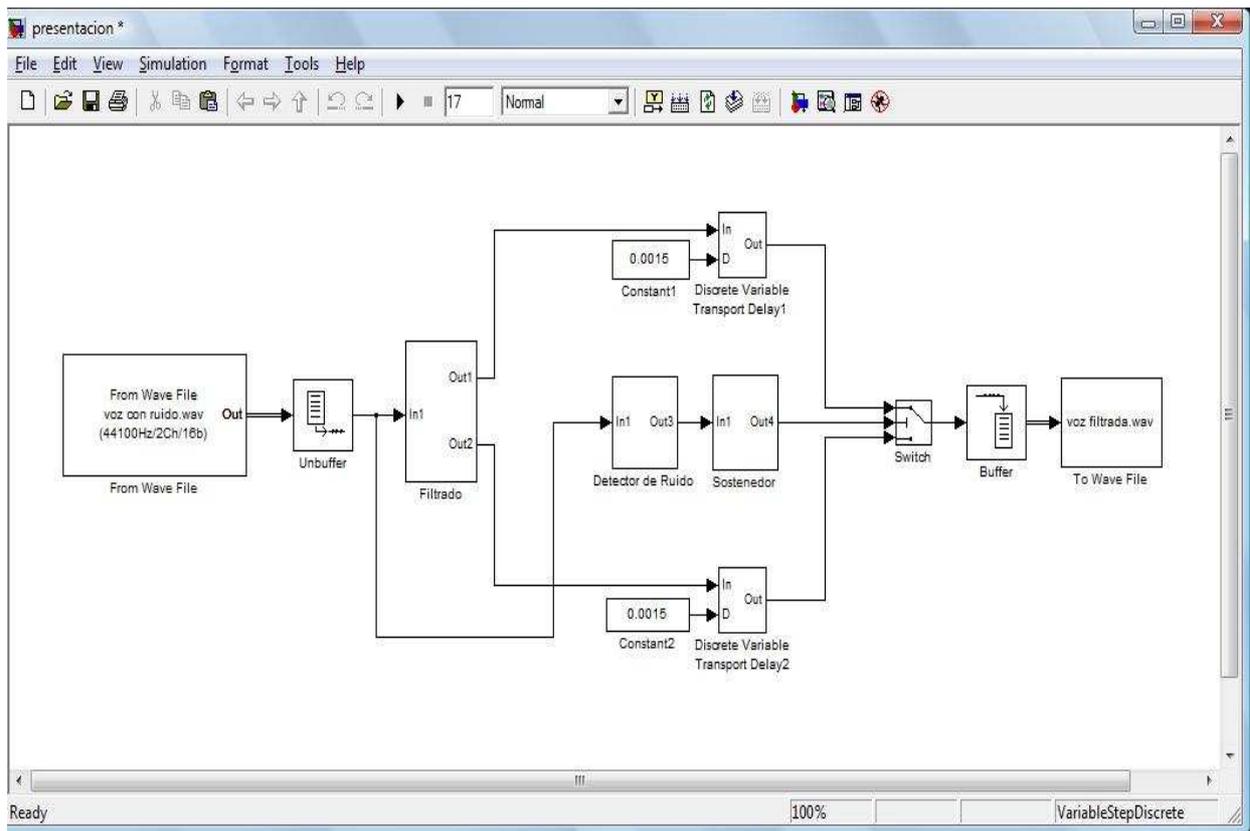


Fig. 4. Esquema de implementación

Adquisición de archivo de audio:

El archivo de sonido usado es una señal de audio (voz) con una duración de 15 segundos con una frecuencia de muestra de datos de 44100Hz.

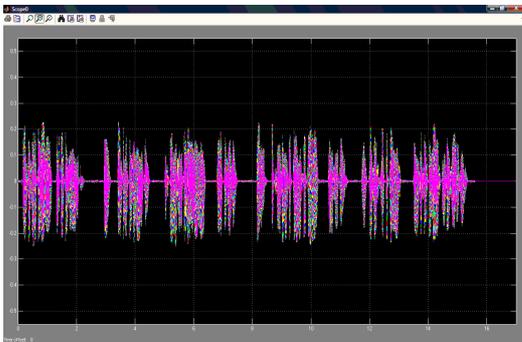


Fig. 5. Señal de audio (voz)

El ruido que afectará a la señal de audio, es un ruido de pulsos (clicks y pops) con una frecuencia de muestra de datos de 44100Hz y está conformado por 11 pulsos que se repiten periódicamente.

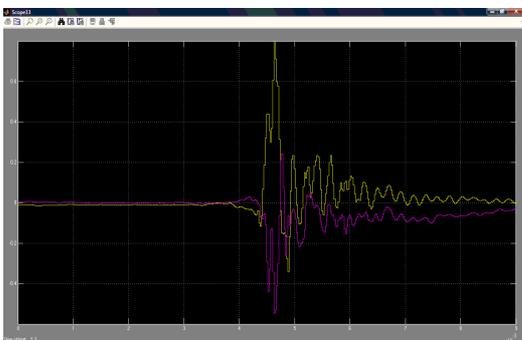


Fig. 6. Ruido de pulso



Fig. 7. Señal de ruido (pulsos)

La señal de entrada al sistema es un archivo de audio real formado por la combinación de voz más un ruido de pulsos que provoca una deterioración en “voz con ruido.wav”; que tiene formato Windows PCM estéreo de 2 canales con un tamaño de 2.62 MB, velocidad de transmisión de 1411 Kbps a una frecuencia de muestra de datos de 44100Hz.

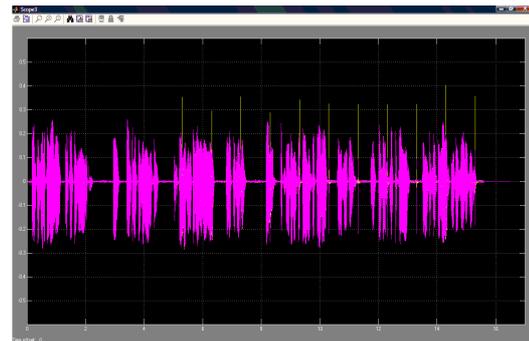


Fig. 8. Señal de entrada al sistema

Acondicionamiento de la señal:

Los bloques encargados de adaptar el archivo de audio a la interface de Simulink son “From Wave File” y “Unbuffer” los cuales terminan por proporcionar dos señales basadas en muestras de 1x256 que representan los 2 canales estéreo.

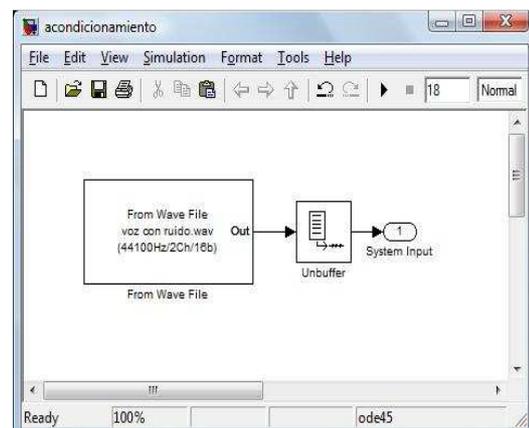


Fig. 9. Acondicionamiento de la señal

Bloque de Filtrado:

Este bloque tiene la finalidad de obtener 2 señales, una filtrada (Out1) y otra sin filtrar (Out2) estas serán usadas posteriormente en el bloque selector.

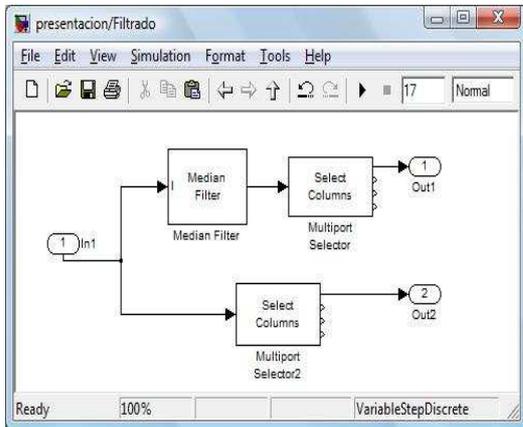


Fig. 10. Bloque de Filtrado

Bloque Detector de Ruido:

La finalidad de este bloque es la obtención de una señal de pulsos de amplitud uno (Out3) para la identificación o ubicación del ruido. Esta señal será usada posteriormente en el bloque selector.

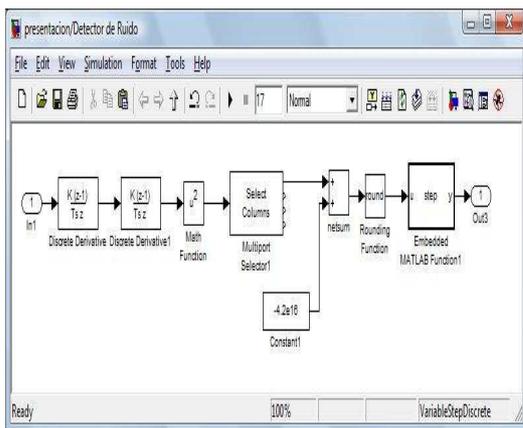


Fig. 11. Bloque Detector de Ruido

Bloque Sostenedor:

Este bloque actúa como un bloque auxiliar, su función es recibir la señal de pulsos generada en el Bloque Detector de Ruido

como entrada, y aumentar el ancho de los mismos para la eficiencia en la conmutación del bloque selector.

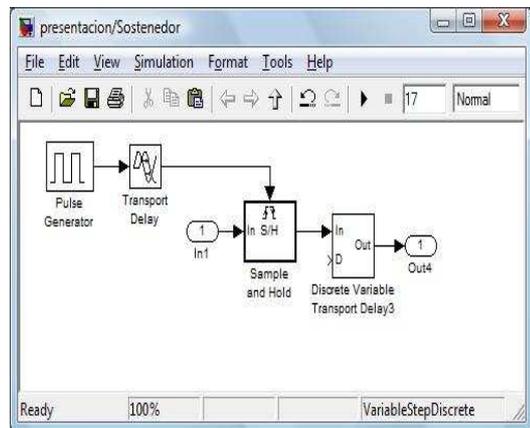


Fig. 12. Bloque Sostenedor

Bloque Selector:

Este bloque está conformado por un "Switch" que realiza la conmutación entre la señal de salida Out1 y Out2 del Bloque de Filtrado, a través de una señal de control (Out4) proporcionada por el Bloque Sostenedor. Estas son las entradas In1, In2 e In3 respectivamente.

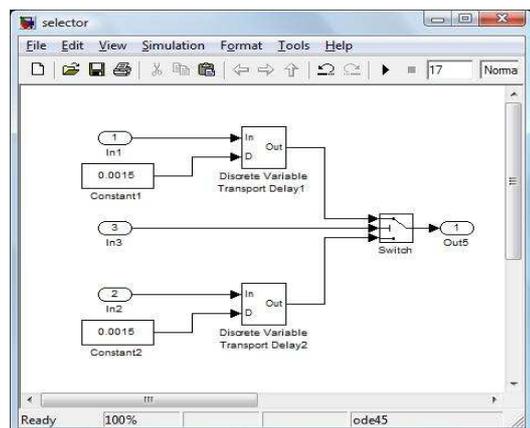


Fig. 13. Bloque Selector

Conversión a un archivo de audio:

Finalmente la salida del switch entra a los bloques "Buffer" y "To Wave File" crean el archivo de audio "voz filtrada.wav"; que

tiene formato Windows PCM mono con un tamaño de 665 KB, velocidad de transmisión de 320 Kbps a una frecuencia de muestra de datos de 20000Hz.

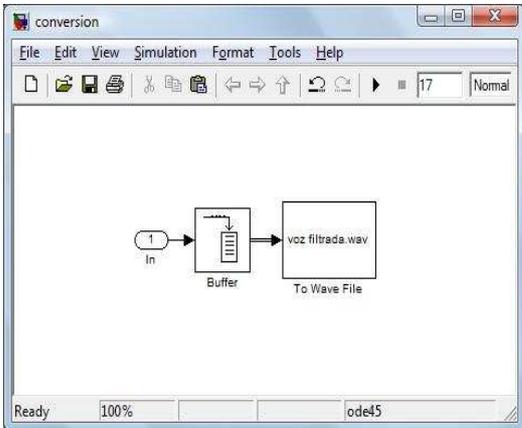


Fig. 14. Conversión a un archivo de audio

4. ANÁLISIS DE RESULTADOS

La señal obtenida como resultado muestra la significativa disminución del ruido como podemos observar a continuación.

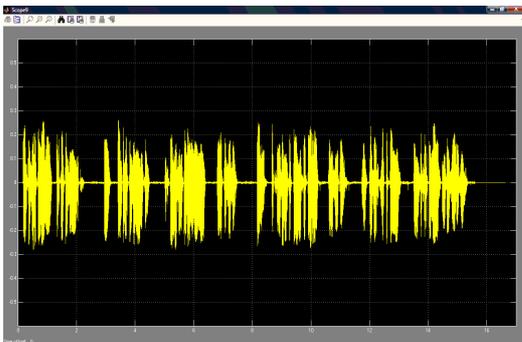


Fig. 15. Señal obtenida a la salida del sistema

Un factor importante para la eficiencia del sistema es el valor de ancho de pulsos a usar en el Bloque Sostenedor.

En la primera prueba el valor del ancho de pulso de 0.003 segundos, en este caso la totalidad del ruido no era filtrado.

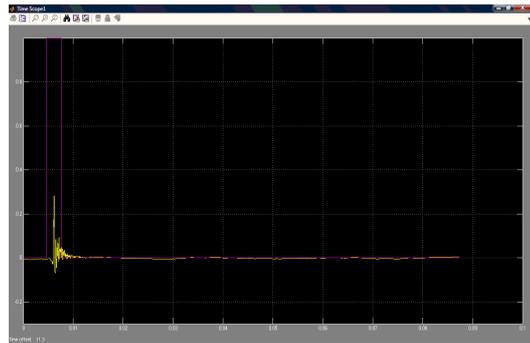
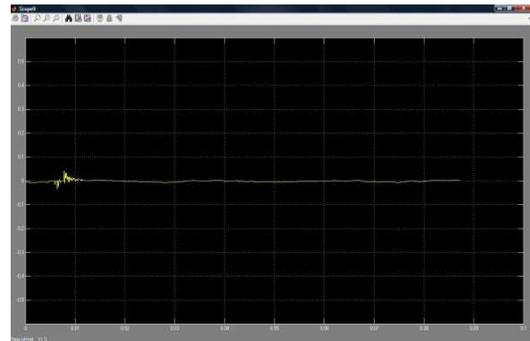


Fig. 16. Prueba con un ancho de pulso de 0.003



Con un valor de 0.009 segundos baja la eficiencia del sistema ya que hay pérdidas en amplitud de la señal original.

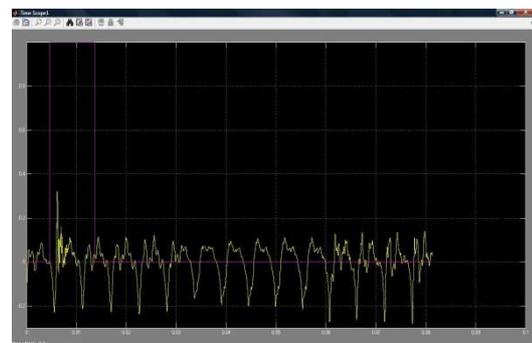
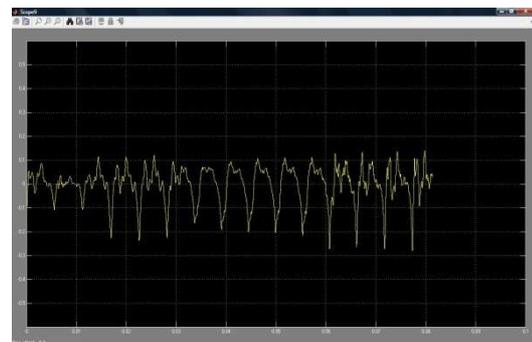


Fig. 17. Prueba con un ancho de pulso de 0.009

Finalmente usando un ancho de pulsos de 0.00625 segundos; dicho pulso cubre toda la componente del ruido que deseamos eliminar, sin tomar componentes adicionales de la señal original.



Fig. 18. Prueba con un ancho de pulso de 0.00625

Es importante tener en cuenta que para poder seleccionar el valor del *ancho de pulso* correcto para la eficiencia del sistema, en las pruebas que se realizaron además de usar un criterio analítico y visual también se usó un criterio auditivo, es decir, se compararon los archivos “voz con ruido.wav” y “voz filtrada.wav” por cada valor de prueba tomado del ancho de pulso, hasta determinar el más eficiente en este caso (0.00625 segundos).

Dentro de la implementación del sistema, un aspecto notorio es el cambio de calidad del archivo de audio. El archivo “voz con ruido.wav” es de dos canales con calidad estéreo; en contraste a “voz filtrada.wav” que es de un canal con calidad mono.

Para efectos de esta aplicación, el cambio de calidad en el archivo de audio no es muy relevante, por no ser muy perceptible ante un criterio auditivo.

5. CONCLUSIONES

El proyecto presentado en Simulink, herramienta fácil de manejar, tiene un tiempo de procesamiento relativamente bajo y gran eficiencia. Por lo cual es una posible solución para las aplicaciones menores que no podrían tolerar el tiempo de procesamiento de otros métodos más sofisticados.

Basados en los resultados favorables de las pruebas, descritos en los “Análisis de Resultados”, se determinó que el filtro de mediana, no solo es funcional para el procesamiento de imágenes, sino que también lo es para aplicaciones de señales de audio afectadas por ruido de pulsos.

La transmisión de la información es clara a pesar de que existe una pérdida en la calidad del audio. Esta pérdida es producto del cambio de formato entre el archivo de audio de entrada y de salida del sistema (*estéreo a mono*), sin embargo es poco perceptible ya que se trata de un archivo exclusivamente de voz.

Un factor determinante para la graduación de este filtro adaptativo es el ancho de pulso de la señal de conmutación; ya que los resultados experimentales muestran que con un ancho de pulso de 0.00625 segundos se obtiene el máximo rendimiento del sistema.

6. REFERENCIAS

1. Rabiner L., Shoter R.; **Digital Processing of Speech Signals**. NJ. Prentice Hall. 1978.
2. Wikipedia, the Free Encyclopedia; **Median Filter**. http://en.wikipedia.org/wiki/Median_filter. 2008.
3. Lasparis T., Lane J.; **Adaptive Scratch Noise Filtering**. IEEE Transactions on Consumer Electronics, vol.39 N.04. 1993.
4. Hodgson R., Bailey D., Naylor M., Ng A., McNeill S.; **Properties, Implementations and Applications of Rank Filters, Image and Vision Computing**. Butterworth and Co. Ltda. vol.3 N.1. 1985.
5. Yli-Harja O., Shmulevich I.; **Efficient Representation and Implementation of Recursive Median Filters and their Cascade Compositions**. 1999.