

ANÁLISIS DE LATENCIA DE CONTROLADORES DE AUDIO PARA APLICACIONES DE METROLOGÍA ACÚSTICA

**José Y. Simes^a, Sebastián P. Ferreyra^a, Ana M. Moreno^a, Juan F. López^a,
Juan I. Morales^b, Julian Rametta^a, Gabriel A. Cravero^a y Hugo C. Longoni^a**

^a*Centro de Investigación y Transferencia en Acústica (CINTRA), Unidad Asociada Conicet, Facultad Regional Córdoba, Universidad Tecnológica Nacional. Maestro M. López esq. Cruz Roja Argentina, Ciudad Universitaria, Córdoba, Argentina. X5016ZAA.
acustica@frc.utn.edu.ar, <http://www.investigacion.frc.utn.edu.ar/cintra>*

^b*Instituto de Investigaciones en Ingeniería Eléctrica (IIIE) Alfredo Desages (UNS - CONICET), Depto. Ing. Eléctrica y Computadoras, Universidad Nacional del Sur. Avda. Alem 1253, Bahía Blanca, Argentina. B8000CPB
iiie@uns.edu.ar, <http://www.iiie-conicet.gob.ar>*

Palabras Clave: latencia, controlador de audio, adquisición/reproducción de señales acústicas, mediciones acústicas.

Resumen. La adquisición y reproducción en alta definición de señales acústicas implica cumplimentar diversos requisitos en función de las características de las señales y del sistema interviniente. La fidelidad del proceso de adquisición/reproducción de sonido, la latencia, la inmunidad al ruido y la diafonía son algunos de los factores que definen la calidad de un sistema de adquisición/reproducción de señales acústicas. Por lo general, en aplicaciones de metrología acústica, estos sistemas son de elevado costo y uso privativo. En este trabajo se presenta la comparación de la medición de latencia de dos controladores de audio que operan un sistema para aplicaciones de metrología acústica de bajo costo, evidenciando diferencias significativas entre ambos para dos condiciones de ensayo. Dicho sistema está conformado por computador personal, una tarjeta de adquisición/reproducción de sonido estándar y programas de cómputo numérico. El trabajo se presenta en el marco del proyecto de diseño y desarrollo de un sistema automático de medición de dispositivos acústicos, el cual permite caracterizar dispositivos acústicos activos y pasivos.

1 INTRODUCCIÓN

Una medición tiene por objetivo determinar el valor numérico de una magnitud. La cantidad a medir, denominada *mesurando*, no siempre está completamente definida. Para que la definición sea exhaustiva deberían especificarse todas las condiciones que podrían influir en el valor del mesurando. Para asegurar la calidad de una medición el operador aplica un procedimiento de medición que implementa un método de medición con el fin de obtener un resultado en forma directa o indirecta a través de un modelo de medición (Miyara, 2013). Este procedimiento describe con precisión los pasos a seguir, abarcando la selección del instrumental y de las muestras, el relevamiento de variables ambientales o de otra naturaleza que pudieran influir en el resultado, la especificación de las condiciones de medición, las recomendaciones y precauciones, las fórmulas y cálculos a realizar y la determinación de la incertidumbre de medición, entre otros. Por lo general, para la realización de diversas mediciones acústicas se siguen las directrices que establecen normas de alcance nacional o internacional.

El objeto de las mediciones acústicas es extraer información representativa de una señal variable en el tiempo y por lo general correspondiente a un fenómeno irreplicable. En algunos casos, el propio sensor (transductor), el instrumento o los accesorios utilizados y hasta la presencia del operador pueden alterar significativamente la señal a medir.

Actualmente, la adquisición/reproducción en alta definición de datos en aplicaciones de metrología acústica, se realiza mediante tarjetas de adquisición/reproducción de sonido (TARS) en combinación con transductores de diseño específico. Las TARS se caracterizan principalmente por su baja latencia, alta fidelidad, bajo ruido, operación en modo full dúplex y/o multicanal, baja diafonía, entre otros. Por lo general, su costo es elevado y utilizan programas de tipo privativo. Sin embargo, es posible lograr resultados plausibles al optimizar el controlador -programa que opera una TARS- de un sistema de medición basado en un computador personal estándar, una TARS de bajo costo, transductores y otros dispositivos utilizados en metrología acústica.

En este trabajo se presenta un análisis comparativo de la medición de latencia de dos controladores de audio operando en idénticas condiciones un sistema para aplicaciones de metrología acústica de bajo costo. Este último está conformado por una TARS estándar, la cual se comunica con un computador personal y transductores electroacústicos (micrófono y altavoz). El sistema de medición utiliza un programa de cómputo numérico como MATLAB ú OCTAVE. Ambos, permiten desarrollar funciones avanzadas de adquisición/reproducción de sonido, análisis y procesamiento digital de señales. Sin embargo, la compilación del lenguaje de alto nivel que utilizan estos programas se desarrolla en serie, imposibilitando la comunicación en modo full dúplex con la TARS, ya que una tarea se ejecuta finalizada la anterior. Para salvar este inconveniente, el sistema implementa la librería especializada *Psychtoolbox-3* (PTB-3) (Kleiner, 2007), la cual es un conjunto de funciones de distribución libre y código abierto para procesamiento de señales audiovisuales, que permite la operación de la TARS en modo full dúplex en programas de cómputo numérico.

2 CONTROLADOR DE AUDIO

Un controlador de audio es un algoritmo de bajo nivel que maneja las conexiones de datos entre un hardware seleccionado y el sistema operativo que gestiona el procesamiento

de datos. Por lo general, cada sistema operativo cuenta con diversos controladores de audio, aunque no todos son aptos para realizar mediciones acústicas. Los sistemas basados en UNIX utilizan una *arquitectura de sonido avanzada de linux* (ASLA) (Phillips, 2005) eficiente que permite manejar la configuración de una TARS. Las principales características configurables son la operación en modo full-dúplex, la mezcla de múltiples canales en hardware, el soporte multiprocesador, seguridad en cuanto a virus informático y la utilización de hardware con síntesis de interfaz digital de instrumentos musicales.

En este trabajo se utilizan controladores de audio disponibles para sistemas operativos Windows Vista, 7 y 10, desarrollados por los fabricantes de las TARS, sin embargo estos sistemas operativos permiten ejecutar otros controladores disponibles, como los *codecs* de alta definición de audio de Realtek Company Ltda., que posibilitan la configuración de las TARS conectadas al computador personal, pero sin poder modificar su codificación interna, ni controlar parámetros que pueden ser de utilidad en tareas de metrología. Los fabricantes de microprocesadores como Intel o AMD también ofrecen sus propios controladores de audio, con similares limitaciones. Para mejorar las prestaciones e introducir baja latencia de adquisición/reproducción, los sistemas operativos Windows utilizan la librería *DirectSound* desarrollada por Microsoft para TARS.

Debido a las limitaciones mencionadas anteriormente, y con el fin de realizar aplicaciones metrológicas, en este trabajo se utilizó el programa MATLAB combinado con librerías específicas y el protocolo de controlador de TARS con entradas y salidas de audio directo (ASIO: Audio Stream Input/Output) desarrollado por Steinberg Media Technologies Ltda., este último establece una interfaz de baja latencia entre el programa de cómputo numérico y la TARS, atravesando distintas capas que establece el sistema operativo con la TARS externa, reduciendo tiempos de retraso entre cada nivel asociado. Sin embargo, tiene la limitación de que su distribución no es compatible con todos los sistemas de procesamiento digital de audio. Es por ello que se incorpora la alternativa del controlador de audio ASIO4all -desarrollado por Michael Tippach-, que ofrece baja latencia para aplicaciones que no fueron diseñadas para tal fin.

2.1 Adquisición de Datos

MATLAB permite incorporar el conjunto de funciones *DAQ Toolbox* para conectarse y realizar configuraciones sobre distintas tarjetas de adquisición de datos. Mediante este paquete de funciones se puede acceder a las características específicas de las TARS, sincronizar los datos requeridos desde múltiples dispositivos, controlar entradas/salidas analógico/digitales, contadores/relojes, entre otros.

Sin embargo, la compilación de las funciones mencionadas se ejecuta en serie. Es decir, el tipo de estructura secuencial del programa no permite realizar dos tareas simultáneamente, razón por la cual no sería posible adquirir una señal al mismo tiempo que se reproduce otra. Por este motivo, la necesidad de implementar una comunicación full-dúplex agrega a su vez un mayor nivel de complejidad al sistema. La versión 2016 del programa MATLAB viene provista de una solución a esta limitación. Cabe mencionar que existen otras librerías de audio alternativas para tal fin. Algunos ejemplo de estas librerías son: *Windows Audio*, *ASLA*, *Simple DirectMedia Layer*, *FFmpeg* y *GStremear*, entre otras (Pelli, 1997), en este trabajo se aplicó la última.

GStremear (2017) es un framework multimedia, de código abierto y distribución libre. Diseñado para trabajar en múltiples sistemas operativos, escrito en lenguaje de

programación C, permite la adquisición/reproducción de sonido en forma directa y simultánea, otorgando las cualidades requeridas anteriormente. GStremear se basa en un sistema conocido como *pipeline*. Estos sistemas son un conjunto de elementos de datos procesados en serie, donde la salida de uno es la entrada del siguiente (Quinn, 2004). De esta forma, los elementos de una pipeline pueden ser ejecutados en paralelo, con el defecto de que se aumenta la memoria de *buffer*.

GStremear es implementado en funciones especializadas de audio y video, que forman parte del paquete PTB-3 mencionado anteriormente. El mismo ofrece un nuevo controlador de audio *PsychPortAudio* que puede ser utilizado específicamente en tareas de psicoacústica (Brainard, 1997). El controlador permite la comunicación full-dúplex entre múltiples canales, pudiendo el usuario modificar su código para aplicaciones específicas.

2.2 Tarjeta de Adquisición/Reproducción de Sonido

Una TARS es un dispositivo periférico que permite la adquisición y reproducción de audio manipulado por el controlador de audio. La principal característica de este hardware es la utilización de convertidores Analógico-Digital (ADC) y Digital-Analógico (DAC) para dicho proceso. Un sistema de adquisición/reproducción de sonido se conforma de tres elementos principales: transductores, la TARS y un programa para procesamiento de señales. La TARS actúa como interfaz entre el computador personal y los transductores. Actualmente, en metrología acústica existen diversos modelos que varían en precio y calidad, algunos ejemplos son las TARS de Brüel & Kjaer, National Instruments, Measurement Computing que brindan excelentes niveles de relación señal-ruido (SNR: signal to noise ratio), como también baja distorsión armónica total (THD: Total Harmonic Distortion) y baja latencia (< 50 ms). Sin embargo, su costo es elevado. En este trabajo, se utiliza una TARS de bajo costo, cuantización de 24 bits y frecuencia de muestreo de 44100 Hz. Los cual, establecen óptimas condiciones de SNR para la realización de mediciones acústicas. Esta última, estará principalmente definida por el ruido intrínseco de los transductores utilizados en el sistema de medición. En la Figura 1, se presenta el diagrama de bloque de un sistema genérico de adquisición/reproducción de sonido como el descrito anteriormente.

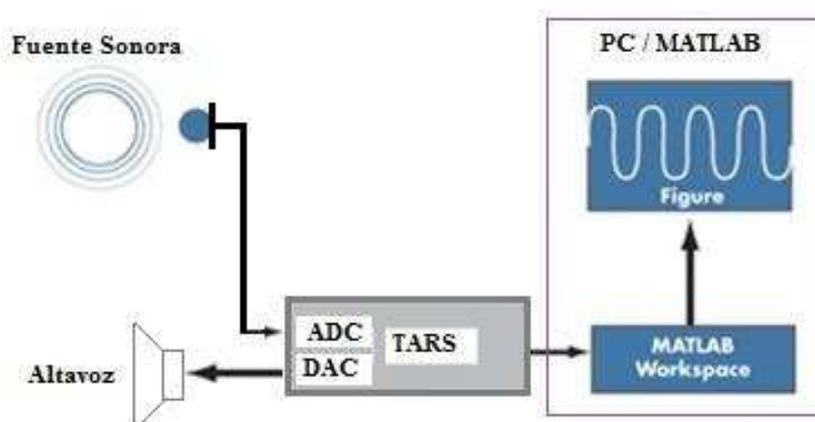


Figura 1: Diagrama de bloques de un sistema de adquisición/reproducción de sonido utilizando MATLAB.

3 CONFIGURACION PARA MEDICIONES ACÚSTICAS

El sistema de medición desarrollado utilizó la librería PTB-3, un programa en MATLAB 2013, una TARS (PRESONUS, AudioBox 44VSL) y controladores ASIO para disminuir la latencia del sistema. Utilizando PTB-3 se configuró el controlador de audio, número de canales, frecuencia de muestreo, la TARS externa, modo full-dúplex y el tamaño del buffer. Esta configuración fue implementada en el software del sistema de medición automático de dispositivos acústicos (SAMSoft), desarrollado por el CINTRA (Ferreyra et al, 2016). El mismo, genera diferentes señales de excitación para realizar diversas mediciones acústicas (Moreno et al, 2016). Un requerimiento específico del SAMSoft es asegurar que al emitir una señal de excitación, la adquisición en simultáneo haya comenzado previamente. Es por este motivo, que se agregan muestras de amplitud nula que pueden ser consideradas de latencia en la reproducción. A través del controlador *PsychPortAudio* se reproducen y concatenan las señales con información disponible para luego ser procesadas o sincronizadas. En el caso de mediciones de directividad de transductores electroacústicos (Ferreyra et al, 2010), la cantidad de muestras que debe tomarse en cada dirección establece un proceso prolongado. El registro que caracteriza el comportamiento temporal del sistema es a su vez lento y requiere de importantes recursos del sistema. Estos procesos multiplicados por cada dirección que se desea analizar demandan que se reduzca el tiempo en las demás funciones del sistema, por lo que es conveniente reducir la latencia y número de muestras nulas que se utilizan para realizar las mediciones.

4 MEDICIONES Y RESULTADOS

Con el objetivo de cuantificar la latencia existente entre la señal emitida y registrada se realizaron mediciones modificando únicamente el controlador de audio en el sistema. Para tal fin, se realizaron dos tipos de ensayos: uno donde el canal de comunicación es eléctrico, y otro electroacústico (interviniendo dos transductores electroacústicos y el medio de propagación). Los componentes principales del sistema de medición se detallan en la Tabla 1. El sistema operó en modo full-dúplex, utilizando 2 canales de la TARS. Durante los ensayos se configuró la TARS con una frecuencia de muestreo de 44100 Hz, y 16 bits de cuantización para la digitalización de la señal. Se utilizó la librería *PsychToolbox-3* (en ambos casos) y dos controladores de audio: ASIO Audio Box (provista por el fabricante de la TARS) y ASIO4all.

Micrófono electrostático electret	Behringer, ECM 8000
Caja acústica activa	Mackie, MKII MR8
Tarjeta de adquisición y reproducción de sonido (TARS)	Presonus, AudioBox 44 VSL
Computador personal	GigaByteBrix, GB-Bace-3150. Intel Celeron N3150, DDR3L 4GB.
Cableado balanceado	AWG, 1 m.
Sistema Operativo	Windows 10
Programa de cómputo numérico	MATLAB 2013a
Cámara acústica	CINTRA, Cámara A

Tabla 1: Componentes del sistema de medición.

4.1 Ensayo de latencia: canal eléctrico.

Este ensayo determina la latencia del sistema considerando exclusivamente los retardos de tiempo que introduce la TARS, el controlador de audio y el cableado. En la Figura 2, se presenta el diagrama de bloques del sistema de medición implementado. La señal emitida a través del DAC de la TARS se comunica por un conductor eléctrico de 1 m de longitud, transportando la señal diferencial hasta la entrada de la TARS para ser adquirida por su ADC. Como señal de excitación se utilizó una señal sinusoidal de 1 kHz de frecuencia de amplitud constante y 2 s de duración sintetizada con MATLAB.

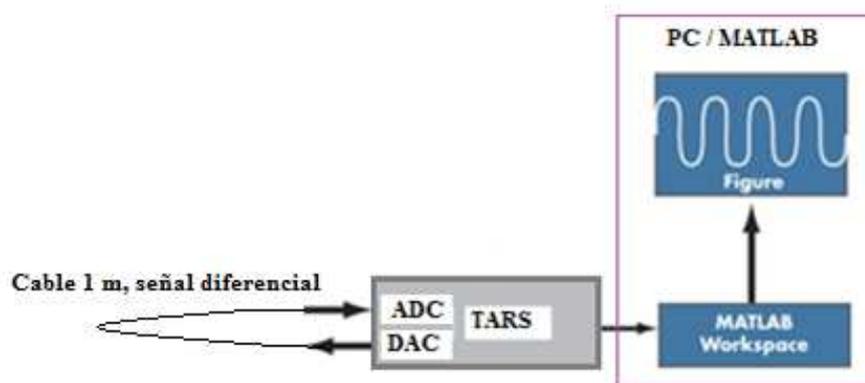


Figura 2: Diagrama de bloques del sistema de medición de latencia para canal de comunicación eléctrico.

Se efectuaron un total de 10 repeticiones para cada controlador de audio, con el fin de realizar un análisis estadístico determinando el valor medio y el desvío estándar asociado. En la figura 3 se visualiza la señal de excitación mencionada.

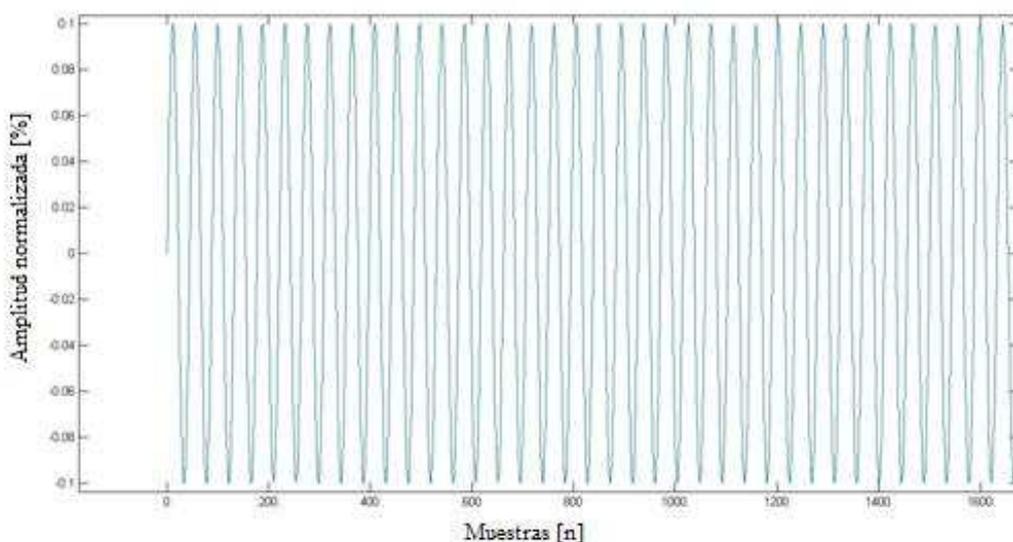


Figura 3: Señal de excitación sinusoidal de 1 kHz de frecuencia, amplitud constante y 2 s de duración.

En la Figura 4, se ilustra el registro de captura de datos de la primera repetición utilizando el controlador de audio ASIO Audio Box. En la Figura 5 se muestra el detalle del inicio de la señal registrada, su fase, y el retardo correspondiente a la muestra número 2390, para un aumento del 10 % de la amplitud de la señal.

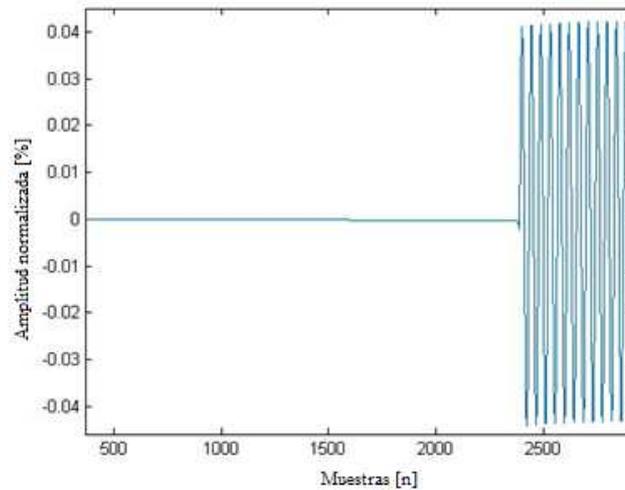


Figura 4: Registro de la señal adquirida, canal eléctrico, controlador de audio ASIO Audio Box.

En la Figura 6, se observa el registro de captura de datos de la primera repetición utilizando el controlador de audio ASIO4all, donde se puede apreciar la latencia del sistema. En la Figura 7, se muestra el detalle del comienzo de la señal adquirida, su fase y el retardo correspondiente a la muestra número 1422, para un aumento del 10 % de la amplitud de la señal.

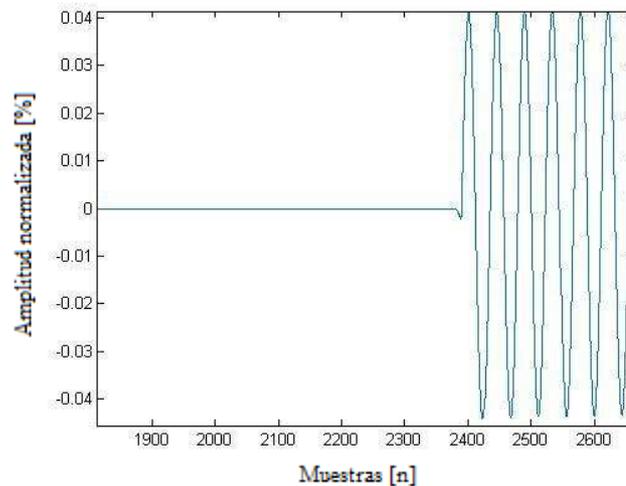


Figura 5: Detalle del inicio de la señal adquirida, fase y retardo introducido, canal eléctrico, ASIO Audio Box.

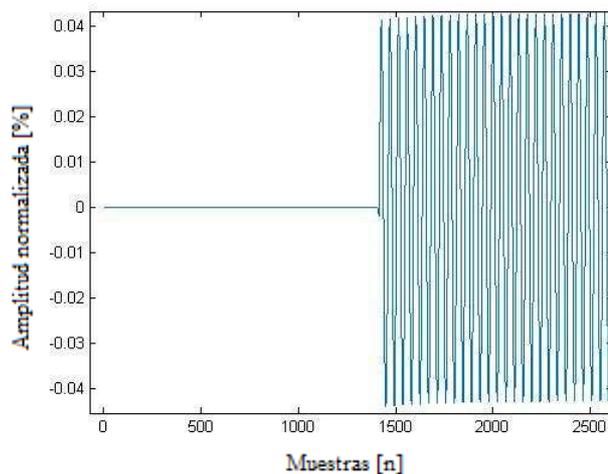


Figura 6: Registro de la señal adquirida, canal eléctrico, controlador de audio ASIO4all.

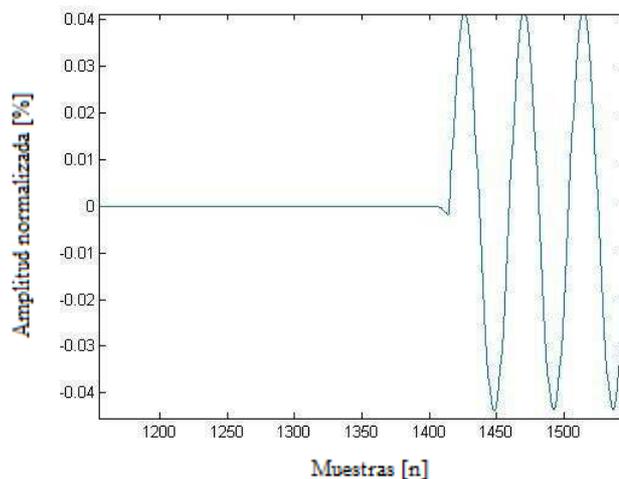


Figura 7: Detalle del inicio de la señal adquirida, fase y retardo introducido, canal eléctrico, ASIO4all.

En la Tabla 2, se presentan los resultados de las mediciones que evidencian, para iguales condiciones de ensayo, que el controlador de audio ASIO4all introduce menor latencia que el ASIO Audio Box. Sin embargo, el análisis estadístico demuestra que el desvío estándar es 131 veces mayor para el controlador ASIO4all comparado con ASIO Audio Box, lo cual puede condicionar su aplicación para ciertas aplicaciones en metrología acústica.

Controlador de audio	ASIO Audio Box	ASIO4all
Señal de excitación	Tono senoidal, 1 kHz	Tono senoidal, 1 kHz
Frecuencia muestreo	44100 Hz.	44100 Hz.
Canal de información	Eléctrico	Eléctrico
Latencia: valor medio y desvío estándar.	54,190 ms \pm 0,016 ms	30,002 ms \pm 2,100 ms

Tabla 2: Resultados de medición de latencia, canal eléctrico.

4.2 Ensayo de latencia: canal electroacústico.

Este ensayo determina la latencia del sistema considerando los retardos de tiempo que introducen los transductores electroacústicos y el medio de propagación, caracterizando la latencia del sistema en condiciones típicas para aplicaciones de metrología acústica. Las mediciones se realizaron en una cámara acústica con un nivel de ruido de fondo de 17,2 dBA y un tiempo de reverberación para frecuencias medias de 0,140 s.

El procedimiento aplicado fue similar al anterior, en la Figura 8 se observa el diagrama de bloques del sistema de medición implementado, donde la señal de excitación se aplica en la entrada de una caja acústica activa, la cual emite la señal acústica que es censada por un micrófono de medición, ambos diafragmas se encuentran enfrentados con una distancia de separación de 10,0 cm. Posteriormente, la señal de salida del micrófono es adquirida por la TARS.

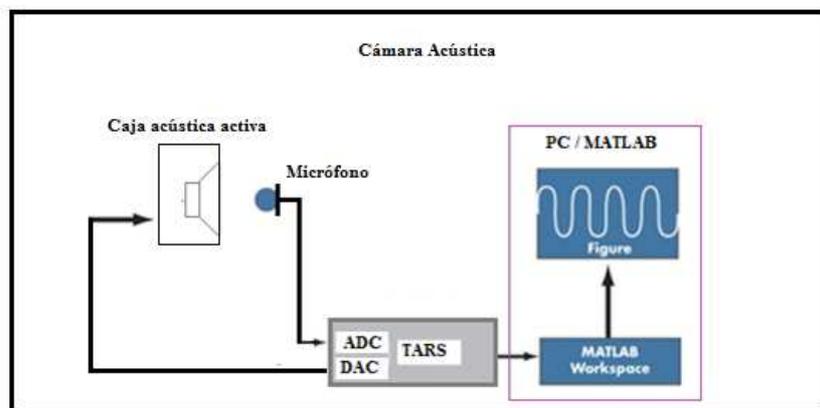


Figura 8: Diagrama de bloques del sistema de medición de latencia canal de comunicación electroacústico.

Nuevamente el sistema opera en modo full-dúplex utilizando 2 canales de la TARS. La señal de excitación es la misma que en el caso anterior, ver Figura 3. En la Figura 9, se observa el registro de captura de datos de la primera repetición utilizando el controlador de audio ASIO Audio Box.

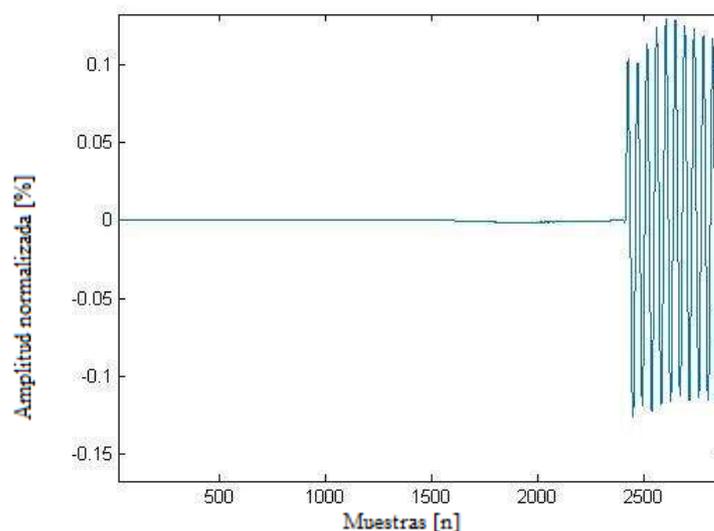


Figura 9: Registro de la señal adquirida, canal electroacústico, controlador de audio ASIO Audio Box.

En la Figura 10 se muestra el detalle el inicio de la señal adquirida, su fase y el retardo correspondiente a la muestra número 2425, para un aumento del 10 % de la amplitud de la señal.

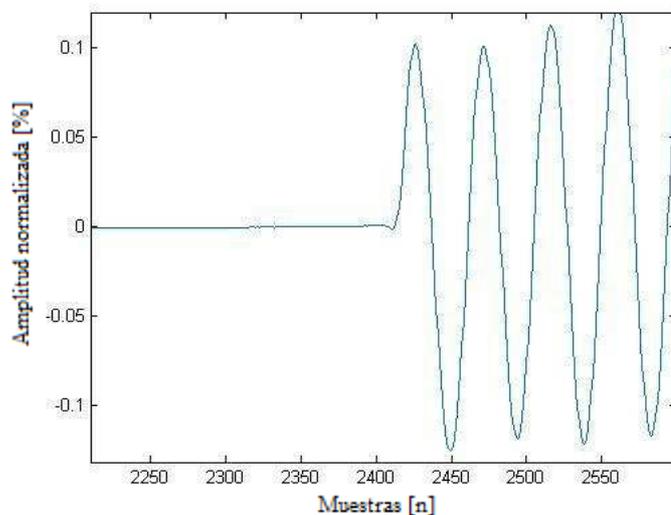


Figura 10: Detalle inicio señal adquirida, fase y retardo introducido, canal electroacústico, ASIO Audio Box.

En la Figura 11, se observa el registro de captura de datos de la primera repetición utilizando el controlador de audio ASIO4all para el canal electroacústico. En la Figura 12 se muestra el detalle el inicio de la señal adquirida, su fase y el retardo correspondiente a la muestra número 1271, para un aumento del 10 % de la amplitud de la señal.

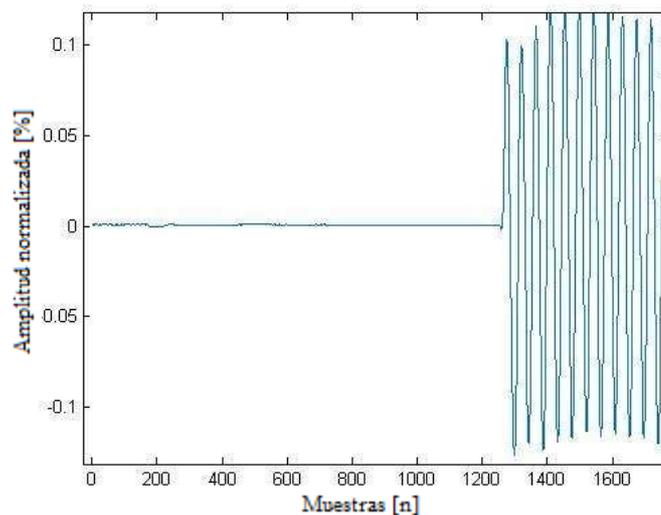


Figura 11: Registro de la señal adquirida, canal electroacústico, controlador de audio ASIO4all.

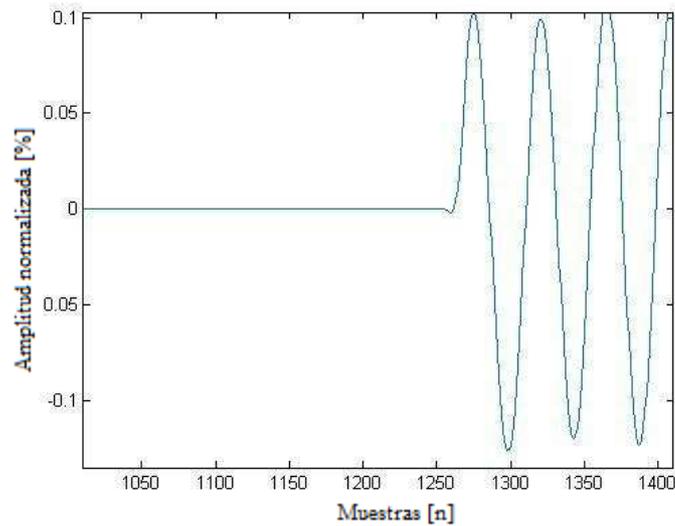


Figura 12: Detalle inicio señal adquirida, fase y retardo introducido, canal electroacústico, ASIO4all.

En la Tabla 3 se presentan los resultados evidenciando que, para iguales condiciones de ensayo, el controlador de audio ASIO4all introduce menor latencia que ASIO Audio Box, con una diferencia de -24,940 ms. Sin embargo, el análisis estadístico demuestra que el desvío estándar es 111 veces mayor para el controlador de audio ASIO4all comparado con ASIO Audio Box, lo cual vuelve a ser consistente con los resultados obtenidos en el ensayo del canal eléctrico.

Controlador de audio	ASIO Audio Box	ASIO4all
Señal de excitación	Tono senoidal, 1 kHz	Tono senoidal, 1 kHz
Frecuencia muestreo	44100 Hz.	44100 Hz.
Canal de información	Electroacústico	Electroacústico
Latencia: valor medio y desvío estándar.	54,680 ms \pm 0,015 ms	29,740 ms \pm 1,671 ms

Tabla 3: Resultados de medición de latencia, canal electroacústico.

5 CONCLUSIONES

El trabajo realizado evidencia que un sistema de medición de bajo costo basado en un computador personal, una TARS estándar, en combinación con programas de cómputo numérico y librerías específicas para adquisición y procesamiento digital de señales puede ser utilizado en aplicaciones de metrología acústica donde se necesita confiabilidad, estabilidad, precisión, repetibilidad y baja latencia como características intrínsecas del sistema. En particular, los ensayos realizados para determinar la latencia del sistema utilizando dos controladores de audio (ASIO Audio Box y ASIO4all) en idénticas condiciones, evidencian la existencia de significativas variaciones entre ambos, tanto para un canal de comunicaciones eléctrico como para uno electroacústico. Además, el análisis estadístico realizado concluye que el controlador de audio ASIO 4all, que introduce menor latencia, presenta un desvío estándar significativamente mayor en comparación con el controlador de audio ASIO Audio Box. Finalmente, el estudio realizado comprueba que este tipo de sistema puede operar de manera confiable y robusta en modo full-dúplex utilizando la librería PTB-3 en combinación con un programa de cómputo numérico.

6 REFERENCIAS

- Brainard D.H., The Psychophysics Toolbox, *Spatial Vision*, 10:443-446, 1997.
- Ferreya S.P., Barcia C. Garcia G., Sistema automático de medición de directividad de transductores electroacústicos (SAMDir). Actas del II Congreso Internacional de Acústica, UNTreF, Buenos Aires Argentina, Septiembre 8-10, 2010. In CD-ROM.
- Ferreya S.P., Moreno A.M., Morales J.I., Tommasini C.F., Budde L., Novillo D.A., Cravero G.A., Longoni C.H., López J.F., Ramos O.A., SAMSoft: Acoustic Device Automatic Measurement System Software. Proceedings of the XXII International Congress on Acoustics, Buenos Aires, Argentina, Septiembre 5-9. ISBN 978- 987-24713-6-1, 2016.
- GStremear, Proyecto GNU de código abierto [en línea]. Disponible en: www.freedesktop.org [Consulta: 10 de abril de 2017].
- Kleiner M., Brainard D. and Pelli D., Abstract Talk: What's new in Psychtoolbox-3?, 30th European Conference on Visual Perception, Arezzo, Italy, Perception, 36: 1-16 (ECP Abstract Supplement), 2007.
- Miyara F., *Mediciones acústicas basadas en software*, Asociación de Acústicos Argentinos (AdAA), Argentina, ISBN 978-987-24713-2-3, 2013.
- Moreno A.M., Ferreya S.P., Morales J.I., López J.F., Simes J.Y., Gelerstein Moreyra S.I., Cravero G.A., Tommasini F. C., Longoni H.C., Gilberto L.G., Proceso computacional para la obtención de descriptores de la directividad de transductores electroacústicos. *Mecánica Computacional*, Vol 34, n 7, pp 393-409. Asociación Argentina de Mecánica Computacional. ISSN: 1666-6070, 2016.
- Pelli, D.G., The VideoToolbox software for visual psychophysics transforming numbers into movies, *Spatial Vision*, 10:437-442, 1997.
- Phillips, D., "A User's Guide to ALSA", *Linux Journal*, 2005.
- Quinn, Michael J., *Parallel Programming in C with MPI and openMP*. Dubuque, Iowa: McGraw-Hill Professional. ISBN 0072822562, 2004.