

SISTEMA DE DETECCIÓN DE MENTIRAS IMPLEMENTADO EN MATLAB

Pablo A. Mora*, Ruben D. Hernández**, Robinson Jiménez**, Jairo Cortes*

*Maestría de Ingeniería Mecatrónica, Universidad Militar Nueva Granada

**Tecnología en Electrónica y Comunicaciones, Universidad Militar Nueva Granada

Grupo de investigación Gi-Ittec Carrera 11 N° 101-80 Bogotá, Colombia

u2900211, ruben.hernandez, robinson.jimenez, jairo.cortes{@unimilitar.edu.co}

Resumen—En este artículo se presenta la implementación de un modelo matemática basado en la transformada rápida de Fourier, la cual es aplicada en el procesamiento de señales de audio con el fin de detectar cuando una persona esta diciendo una mentira; el programa se ha desarrollado en el software de Matlab® con una interfaz gráfica en GUIDE la cual se limita a la captura de señales de voz simples como lo son Si y No y las compara con las afirmaciones verdaderas previamente guardadas; biológicamente cuando se dice una mentira las cuerdas bucales producen una vibración extra conocida como estrés en la voz, lo que permite que el programa diferencie la señal frente a una referencia cargada. Una vez capturada la señal de entrada, se hace el procesamiento e identificación de los parámetros de la voz y los compara con la información previamente almacenados permitiendo determinar si la respuesta es verdadera o falsa.

Index Terms—Detector de mentiras, transformada rápida de Fourier, procesamiento de señales de audio.

I. INTRODUCCIÓN

Cada día nos encontramos más, con infinidad de aplicaciones de los sistemas de cómputo, con capacidad de interactuar con los usuarios mediante el reconocimiento y síntesis de voz. Estos van desde aplicaciones simples en el reconocimiento de comandos (palabras) aislados, hasta el reconocimiento de frases para ejecutar acciones a manos libres: teléfonos celulares, control por voz de instrumental de ayuda en la cirugía, acceso a servicios de compra por teléfonos, llenado de solicitudes, reservaciones de pasaje, entre otros; la búsqueda en Internet es una de las muestras más impactantes de estas aplicaciones [1]. Ahora bien una de las aplicaciones menos comunes pero muy útiles es el reconocimiento del estrés de voz para detectar respuestas falsas en una entrevista, este detector de mentiras funciona a través de un sistema que captura las modulaciones de la voz del evaluado; por su recepción de audio se establece las frecuencias o micro vibraciones que identifica en tiempo real y las transforma en una gráfica. Cada frecuencia es asociada en muestras de audio.

De esta manera, es posible determinar los efectos de los cambios fisiológicos impuestos en la voz humana por el sistema nervioso autónomo, ante la tensión que genera el estímulo que dirige el examinador; aunque estas variaciones no son perceptibles al oído humano, el instrumento las percibe y convierte en gráficas que el evaluador se encargará de evaluar [2]. Dado que el estudio se realiza con la voz del individuo, este puede ser realizado mediante entrevistas personales o

telefónicas o incluso ser grabadas para su posterior análisis. A diferencia del polígrafo, el VSA (Voice Stress Analysis) se da en condiciones normales y no requiere un método invasivo para obtener las respuestas de lo que se quiere establecer. En el desarrollo propuesto se presenta una interfaz gráfica que permite realizar una calibración de la voz del usuario a ser evaluado por medio de preguntas básicas, las cuales deberá contestar con verdad y mentira de forma que el programa distinga y caracterice las señales de respuesta, lo que generara un cuadro señales básicas para comparar las futuras respuestas a partir de un Si y No verdaderos, y así mismo un Si y No falsos. El usuario tiene la opción de grabar su respuesta durante dos segundos a partir de que la señal de inicio es silenciada.

El presente artículo se estructura como sigue: la sección II presenta el procesamiento de audio, la sección III el procedimiento matemático, la sección IV el análisis de resultados y finalmente la sección V las conclusiones obtenidas.

II. PROCESAMIENTO DE AUDIO

Para realizar la captura del audio a procesar se utiliza un equipo de grabación profesional de música, de la marca LINE6®, el cual cuenta con una resolución de captura de 32 bits a 48000Hz, lo que permite observar en detalle la señal de audio. La grabación se realiza bajo condiciones ideales para evitar sonidos de altas y bajas frecuencias que posiblemente alterarían la señal, y a pesar de implementar filtros se pierden datos que pueden ser útiles para la comparación de resultados. Luego de realizar la grabación se procede a cargar el audio matricialmente en Matlab, para poder empezar su tratamiento. Básicamente los datos que se obtienen son la frecuencia de muestreo y el valor de la amplitud de la señal aplicando los siguientes parámetros:

A. Normalización

La normalización consiste en aplicar una cantidad constante de ganancia a una grabación de audio para llevar la amplitud de pico promedio a un nivel objetivo (la norma). Debido a que se aplica la misma cantidad de ganancia a todo el rango, la relación señal-ruido generalmente no cambia. La normalización se diferencia de la compresión de rango dinámico, entre los diferentes niveles de ganancia a una grabación para que la amplitud se encuentre dentro de un rango mínimo

y máximo. La normalización es una de las funciones más comunes proporcionadas por una estación de trabajo de audio digital. [3]

En la gráfica de la figura 1, se puede observar la señal original de color azul, que corresponde a la grabación sin ningún tipo de procesamiento, la gráfica color rojo presenta la señal resultado de la normalización, que básicamente amplifica la señal en estudio y deja el ruido con una amplitud pequeña, lo que permite reconocer donde se encuentra el área a trabajar y el resto recortarlo y/u omitirlo.

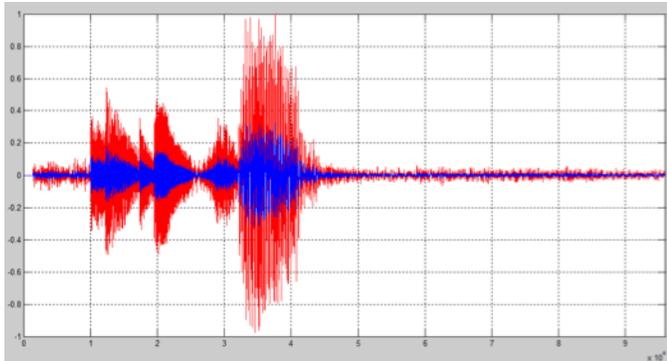


Figura 1. Normalización de la señal

B. Transformada rápida de Fourier

Las series de Fourier son útiles para el estudio de señales periódicas pero, desafortunadamente, este tipo de señales no son tan frecuentes en la práctica como las no-periódicas. Esta situación requiere el desarrollo de una teoría matemática más profunda como se presenta en [4]. La palabra “transformada” indica que se aplica una herramienta que tiene la capacidad para cambiar un tipo determinado de problema en otro. De hecho, la transformada de Fourier es útil para simplificar el estudio de la solución de cierto tipo de ecuaciones diferenciales, convirtiendo el problema de la solución de una ecuación diferencial en un problema de solución de ecuaciones algebraicas. La motivación para dicho estudio está en el hecho de que la transformada de Fourier posee buenas propiedades algebraicas cuando se aplica a las derivadas sucesivas de una señal, o al trasladar la señal, entre otras [4-5]. En la ecuación 1 se presenta la descripción matemática de la transformada rápida de Fourier en tiempo continuo de una función.

$$F\{f\} : \xi \mapsto \hat{f}(\xi) := \int_{-\infty}^{\infty} f(x) e^{-2\pi i \xi x} dx \quad (1)$$

Donde $f(x)$ corresponde a la función a transformar y las variables x y ξ corresponden al cambio tiempo frecuencia.

En la gráfica de la figura 2 se presenta la transformada rápida de Fourier aplicada a la respuesta del usuario.

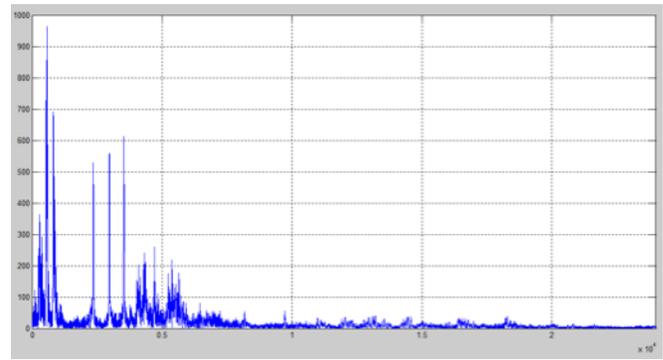


Figura 2. Respuesta de la transformada rápida de Fourier

III. PROCEDIMIENTO MATEMÁTICO

A. Media de una señal

En matemáticas y estadística, la media aritmética (también llamada promedio o simplemente media) de un conjunto finito de números es el valor característico de una serie de datos cuantitativos objeto de estudio que parte del principio de la esperanza matemática o valor esperado, se obtiene a partir de la suma de todos sus valores dividida entre el número de sumandos. Cuando el conjunto es una muestra aleatoria recibe el nombre de media muestra siendo uno de los principales estadísticos muestrales [6]. En la ecuación 2, se presenta la ecuación de la media, que en este caso es aplicada a la matriz de los valores de amplitud de la señal de voz, para n muestras con diversos valores a_i .

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n a_i = \frac{a_1 + a_2 + \dots + a_n}{n} \quad (2)$$

En el cuadro 1 se presenta el resultado de la media de cada una de las señales de audio capturadas [6].

Cuadro I
RESULTADOS DE LA MEDIA DE LAS SEÑALES DE AUDIO CAPTURADAS

Señal de Audio	Media
SI-verdadero	7.6246
NO-verdadero	8.2312
Si-Falso	8.1327
NO-Falso	7.9285
Voz de entrada	11.3885

B. Error entre el audio deseado y el obtenido

Luego de obtener el valor absoluto de la media, se decide obtener el error entre la señal de entrada y los valores previamente cargados con el fin de determinar las señales que tienen más parecido. Las señales que menos error tengan entre si son presentadas como respuesta. En la ecuación 3 se presenta el error estimado para cada señal, en donde n corresponde al número del error calculado y m es la media.

$$error(n) = m(|valor\ deseado|) - m(|valor\ obtenido|) \quad (3)$$

IV. ANÁLISIS DE RESULTADOS

Luego de realizar el cálculo del error se obtiene (cuadro 2) que para este caso la señal con menos error es la diferencia del Si-Verdadero con la respuesta del usuario, por ende esta es la respuesta que se presenta en la interfaz.

Cuadro II
RESULTADO DEL CÁLCULO DEL ERROR

Señal de Audio	Error
Error 1	7.0017
Error 2	9.9418
Error 3	8.1532
Error 4	10.5444

En la figura 3 se presenta la interfaz gráfica de programa de detección de mentiras, en la parte izquierda se encuentran los botones de calibración de las pregunta iniciales y en la parte derecha se graba, reproduce e identifica la respuesta del usuario.

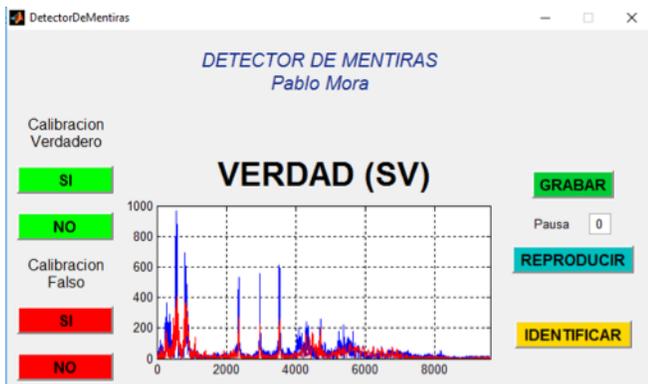


Figura 3. Interfaz gráfica del programa desarrollada en GUI-Matlab®

En la figura 4 se tiene en azul la señal de respuesta del usuario, y en rojo la señal correspondiente a la calibración de un Si verdadero, el error entre la media de estas dos señales es del 7 % aproximadamente, con el solo hecho de observar las gráficas es de notar el parecido entre sí, por ende se selecciona esta respuesta como la correcta.

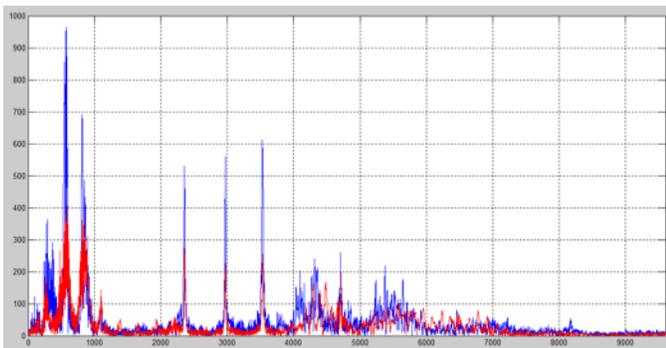


Figura 4. Comparación del SI verdadero con respuesta

En la figura 5 se tiene en azul la señal de respuesta del usuario, y en rojo la señal correspondiente a la calibración de un No verdadero, el error entre la media de estas dos señales es

del 10 % aproximadamente, con el solo hecho de observar las gráficas es de notar que existen diferencias en las señales como por ejemplo en el valor de 1500 del eje x, correspondiente a la frecuencia de muestreo, como también en el valor 4200 del eje x, por esta razón no se considera como coincidencia.

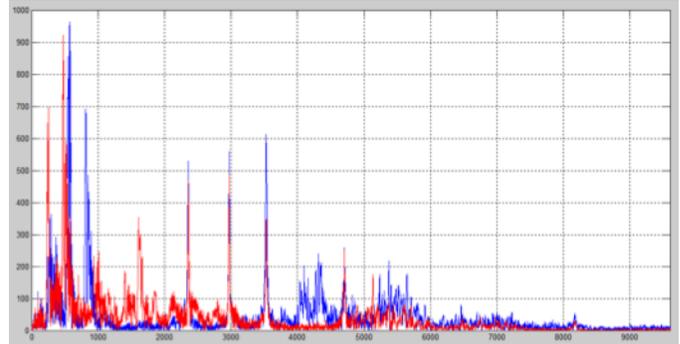


Figura 5. Comparación del NO verdadero con respuesta

En la figura 6 se tiene en azul la señal de respuesta del usuario, y en rojo la señal correspondiente a la calibración de un No falso, el error entre la media de estas dos señales es del 8 % aproximadamente, las diferencias gráficas entre las señales son pocas y difícilmente notorias, sin embargo el método número de la estimación de error dice que existe mayor coincidencia entre el Si verdadero que entre el No falso, este criterio de elección puede generar errores.

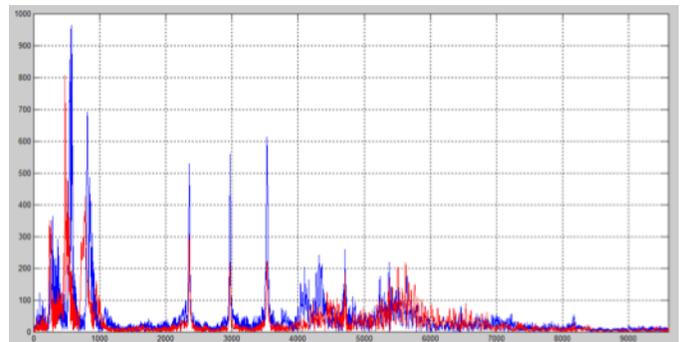


Figura 6. Comparación del SI falso con respuesta

En la figura 7 se tiene en azul la señal de respuesta del usuario, y en rojo la señal correspondiente a la calibración de un No falso, el error entre la media de estas dos señales es del 11 % aproximadamente, con el solo hecho de observar las gráficas es de notar que existen diferencias en las señales como por ejemplo en el valor de 200 del eje x, correspondiente a la frecuencia de muestreo, como también en los valores 1800 y 2000 del eje x, por esta razón no se considera como coincidencia.

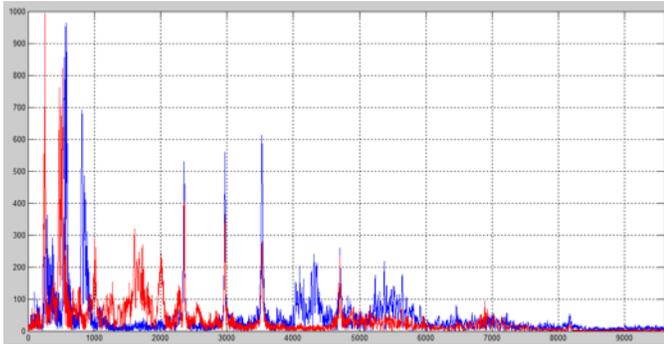


Figura 7. Comparación del NO falso con respuesta

V. CONCLUSIONES

Se presenta un programa que usa la frecuencia natural de la voz para realizar un detector de mentiras a partir de la voz del usuario aplicando conceptos básicos de procesamiento de señales como la transformada rápida de Fourier, normalización, promedios o medias, etc. Por otro lado la obtención de la media y posteriormente del error puede causar problemas ya que la diferencia entre algunos de los errores es muy poca y el criterio de selección juega un papel fundamental para dar una respuesta. Con el fin de obtener una respuesta más acertada se pueden realizar filtros de tipo Butterworth, pasa bajos y pasa altos a la señal de audio antes de aplicarle la transformada rápida de Fourier, lo que eliminaría datos irrelevantes considerados ruido, sin embargo existen métodos que revisan la frecuencia fundamental de la voz, lo que permite observar el nivel de estrés de la voz a un mayor detalle.

REFERENCIAS

- [1] BARROBÉS, Helena Duxans; RUIZ, Marta. "Procesamiento de audio". UOC, 2012.
- [2] MASIP, Jaume; GARRIDO, Eugenio; HERRERO, Carmen. "La detección de la mentira mediante la medida de la tensión en la voz: Una revisión crítica. Estudios de Psicología", 2004, vol. 25, no 1, p. 13-30.
- [3] CHVATAL, L., et al. "Apparatus and method for providing corrected normalized signals corresponding to the composition of a material". U.S. Patent No 3,733,474, 15 Mayo 1973.
- [4] ATHANASIOS PAPOULIS. "Sistemas digitales y analógicos, transformadas de Fourier, estimación espectral". Marcombo, 1985.
- [5] WELCH, Peter D. "The use of fast Fourier transform for the estimation of power spectra: A method based on time averaging over short, modied periodograms." IEEE Transactions on audio and electroacoustics, 1967, vol. 15, no 2, p. 70-73. [6] CANAVOS, George C.; MEDAL, Edmundo Gerardo Urbina. "Probabilidad y estadística". McGraw Hill, 1987.