

# Software para Contabilizar Automáticamente la Publicidad Transmitida en Radio o Televisión, Mediante el Reconocimiento de Audio.

Juan Orrala Moreno, Ing. Jaime E. Velarde, Escuela Politécnica Nacional (EPN), Quito - Ecuador

**Resumen** - Este trabajo, denominado Identificador de Spots<sup>[1]</sup> (SpotID), detecta y contabiliza el audio contenido en un anuncio publicitario que se ha difundido en radio o televisión; mediante la comparación de archivos de audio con formato WAV, que previamente son almacenados. Para “cuantificar el grado de similitud entre las señales de audio”<sup>[2]</sup> se usa la Correlación Cruzada, este proceso se realiza en Matlab en el que incluye una interfaz gráfica al usuario. Durante el desarrollo se realizó decenas de pruebas alcanzando un grado de asertividad del 100%. El prototipo optimiza el tiempo, al realizar el mismo trabajo hecho en forma manual, pero en menor tiempo.

**Índices** – Comparar audios, Correlación Cruzada, Detectar, Publicidad, wav.

## I. INTRODUCCIÓN

El monitoreo de la publicidad tiene el propósito de obtener información, lo más precisa posible, que permita controlar los recursos que se invierten para hacer más efectiva la promoción de un producto, de una idea, etc. El Identificador de Spots (SpotID) es parte de este sistema, al escanear, clasificar y presentar un informe del monitoreo de un determinado comercial transmitido en Radio o TV.

Utilizando los avances tecnológicos, las empresas que realizan el monitoreo de publicidad tienen la expectativa de lograr dos de los objetivos, que podría parecer un solo: contar la publicidad y hacerlo en el menor tiempo posible. El prototipo desarrollado, cumple con esas dos expectativas, analizando exclusivamente en las señales de audio las formas de onda y su fase, dejando de lado el procesamiento de frecuencias, que demanda mayor consumo de recursos computacionales que implica un mayor tiempo de análisis.

## II. DESARROLLO DEL TRABAJO

### A. Diseño del SpotID

Se usa el paquete de software Matlab y su herramienta GUIDE para desarrollar la parte de programación y la interfaz gráfica, respectivamente. SpotID ha sido diseñado para trabajar con archivos grabados desde el capturador de señales ZOGIS Real Angel 400U Dongle USB 2.0 PRO. Este dispositivo tiene con un CD para instalar la

aplicación VivaTV, que permite la visualización, captura y reproducción del video o audio obtenido desde el sintonizador de estaciones emisoras, que tiene incorporado.

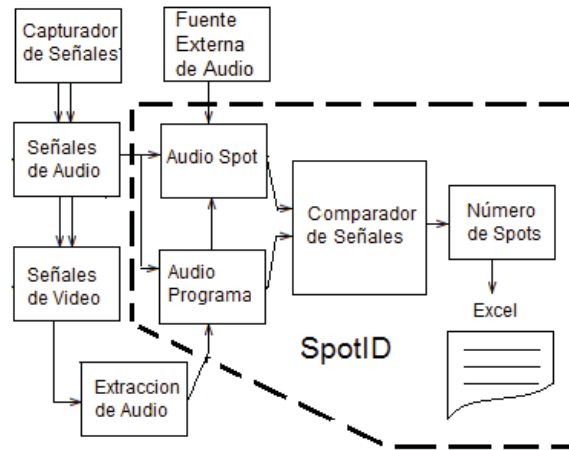


Figura 1 Diagrama de Bloques del Sistema SpotID

En la Figura 1 se muestra el diagrama de bloques de la implementación del SpotID, donde se observan los dos archivos de audio a comparar, el Audio Spot (AS) y el Audio Programa (AP). Todo el desarrollo del prototipo gira en torno a estas dos clases de archivos de audio.

### B. Captura de las Señales de Televisión

Con el software de VivaTV, se realiza la captura o grabación de la programación de cualquier canal de televisión con los formatos de la Tabla I y Tabla II.

TABLA I

FORMATO DE GRABACIÓN DE VIDEOS		
Formato Video	Resolución	Velocidad de bits
MPG	720x480p	2224kbps

TABLA II

FORMATO DE GRABACIÓN DE AUDIO DEL VIDEO		
Formato Audio	Velocidad de bits	Frecuencia de muestreo
MP3	224kbps	48khz

VivaTV tiene una nomenclatura establecida para asignar el nombre a los archivos. En la Figura 2 se muestra el detalle de esa nomenclatura.

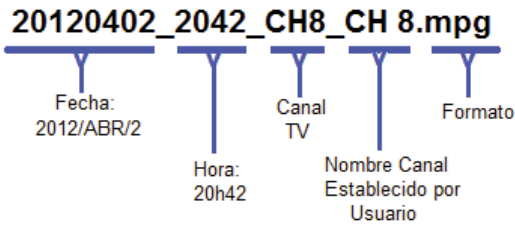


Figura 2 Nomenclatura de Archivo Programación

A partir de esta estructura, el programa diseñado puede procesar la siguiente información:

- Fecha de Grabación
- Hora de Grabación
- Duración de la Grabación.
- Canal y Nombre de Estación
- Franja de Audiencia (A, AA, AAA)

La captura del video se debe limitar a tres horas continuas de grabación, con el fin de que no se vea afectado el rendimiento del computador por el hecho de manejar archivos demasiado grandes. Aproximadamente un archivo de 3 horas de duración requiere de 3 Gb de memoria.



Figura 3 Ejemplo de Archivos Programación.

Cabe recalcar, que en esta etapa del proceso se debe ajustar la antena del sintonizador en busca de una buena calidad de la señal, que está a criterio de la percepción visual y auditiva del operador; se considera técnicamente como “buena” a la relación señal a ruido de 53dB<sup>[3]</sup>.

Sin embargo, la utilización de la Correlación Cruzada permite comparar archivos de audio con deficiente calidad sonora y obtener resultados satisfactorios, como se observa más adelante en el apartado *Comparación de los Archivos de Audio*.

C. Obtención de los Archivos de Audio

Una vez obtenido la captura de video se procede a extraer y acondicionar las señales de audio para obtener los archivos AP y AS.

Tomando en cuenta que Matlab trabaja directamente con archivos de audio con formato WAV y que el audio capturado, según la Tabla II, se encuentra en formato MP3, hay que realizar una conversión de formatos al momento de extraer la señal de audio. Lo que se consigue usando un software externo y que en el presente trabajo se hace mediante el software DVDVideoSoft Free Studio<sup>[4]</sup>.

Por lo tanto, el AP es el resultado de extraer toda la señal de audio que se encuentra dentro de la captura del video, esa extracción se ejecuta con los siguientes parámetros: formato WAV de 8 KHz para la frecuencia de muestreo, con calidad monofónica. Este valor en la frecuencia de muestreo, facilita la manipulación con Matlab en lo que respecta al tamaño del archivo y al tiempo de procesamiento.

En cambio, el AS es una pequeña sección o parte del AP, como se muestra en la Figura 4, la misma que se obtiene con cualquier software utilizado en la edición de audio y con formato de iguales características al AP.

El AS también puede ser obtenido desde otras fuentes externas, siempre y cuando cumpla con el formato establecido.

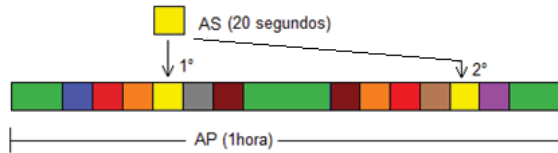


Figura 4 Representación de 2 apariciones de AS en todo el AP

Una vez que se tiene estas dos clases de archivos, se cargan al SpotID para determinar el número de veces que se encuentra el AS dentro del AP y en que instantes de tiempo ocurren.

D. Comparación de los Archivos de Audio

Para determinar que las señales contenidas en dos archivos audio son iguales, se utiliza la función *xcorr*<sup>[5]</sup> implementada en Matlab, función que aplica la Correlación Cruzada entre dos señales y que corresponde a: “la medida de la similitud entre dos señales”, frecuentemente es usada para encontrar características relevantes en una señal desconocida por medio de la comparación con otra que sí se conoce. Es función del tiempo relativo entre las señales y que tiene aplicaciones en el reconocimiento de patrones y en criptoanálisis<sup>[6]</sup>.

El ejemplo gráfico de la Figura 5, muestra dos señales senoidales iguales en amplitud pero desfasadas en el tiempo; el valor máximo de la correlación P0 y P1 será periódico debido a que en más de una vez se encontrará la misma señal en fase. Lo que significa tener más de una vez repetida.

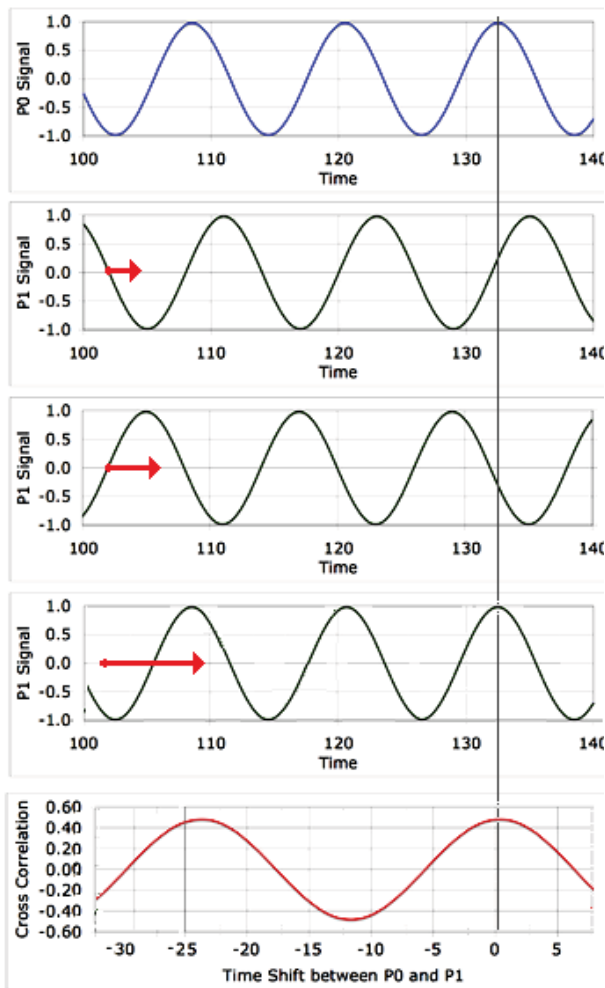


Figura 5 Correlación Cruzada de dos Señales Senoidales<sup>[7]</sup>

El cálculo de la correlación cruzada entre las dos señales P0 y P1, permitirá decidir si P1 se “asemeja” a la señal P0 y en qué posición temporal se encuentra.

La señal P1 es desplazada en el tiempo retrasando o adelantándola. En el caso de la Figura 5, el punto máximo de correlación será cuando coincidan ambas señales en amplitud y fase y será el momento en que P1 será muy semejante a la señal P0. Además la posición del pico nos dará el desplazamiento de la señal en el tiempo.

El valor máximo de la correlación cruzada es independiente de algún tipo de acondicionamiento a la señal por ejemplo normalización, filtrado, atenuación o amplificación.

El ejemplo de la Figura 6 muestra las variaciones de amplitud pero semejanzas en “secuencias” y fase. Esta señal, aún puede reflejar un valor alto de correlación a pesar de los acondicionamientos o distorsiones recibidas. En este trabajo se procesan señales de audio, y cabe resaltar que, mediante pruebas existe correlación cruzada nula, si el audio que ha recibido acondicionamientos o distorsiones deja de ser inteligible al oído humano.

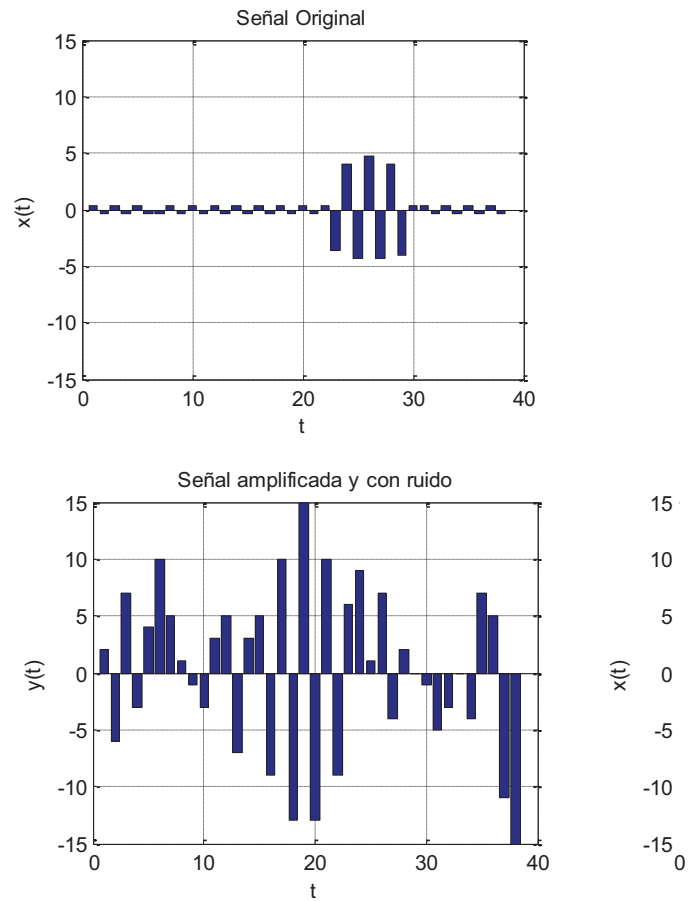


Figura 6 Señal  $x(t)$  Amplificada con Ruido

El proceso de completo de este trabajo sigue la secuencia del diagrama de flujo de la Figura 7, que muestra el ingreso de las dos señales mediante los archivos de audio: AP y AS, luego pasan por la etapa de cotejamiento el par de archivos. El resultado se muestra en pantalla y al mismo se almacena y con la posibilidad de exportar los resultados a Excel, quedando SpotID listo para realizar un nuevo cotejamiento o salir del programa.

Para el cotejamiento de AP y AS dentro del diagrama del SpotID es la función *Semejanza.m*, que corresponde el algoritmo mostrado en la Figura 8. Donde, debido a la necesidad de acelerar el procesamiento se fragmenta al archivo AP, este archivo puede tener una longitud entre 1 a 24 horas, pero dentro de la función se procesan secciones del archivo o fragmentos.

Cada fragmento del AP tiene una longitud aproximada a 2 minutos. Esta fragmentación del AP se realiza en una etapa previa a la correlación.

El número de fragmentos  $i$  es dependiente de la longitud total del AP, el valor  $i$  se genera automáticamente en una variable, que luego es utilizada dentro de la función *Semejanza.m*.

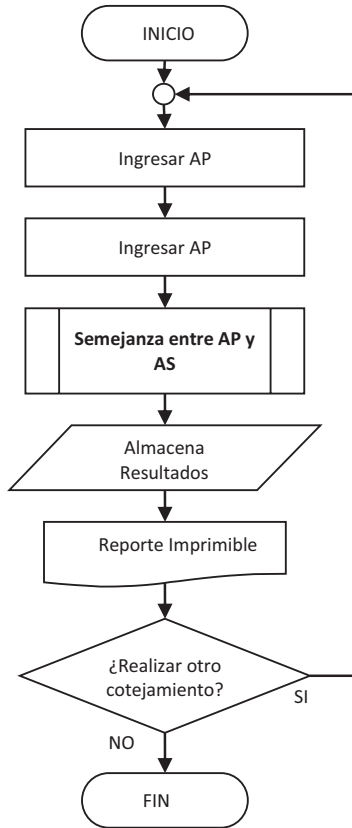


Figura 7 Diagrama de Flujo de SpotID

Luego, se analiza las características de la forma de onda de cada fragmento  $AP[i]$  conjuntamente con las del AS. En cada iteración se genera un grado de semejanza, ese valor es un número  $n$  entero, donde  $n > 0$ , el cual se almacena en un vector  $n[i]$ . Este vector pasa por un discriminador con la intención de encontrar algún valor  $n[i]$  que cumpla con:  $n[i] \geq NMS$ ; donde NMS representa a un *Nivel Mínimo de Semejanza*.

El Nivel Mínimo de Semejanza asegura, que, el grado de similitud entre un  $AP[i]$  y AS sea muy probable. Este umbral numérico se ha determinado mediante pruebas, está por encima de los 1500.

Un ejemplo de la correlación de los audios de la Figura 9; donde AS es de 21 segundos y AP es de 2 minutos y 10 segundos, tiene como resultado los valores de correlación mostrados en la Figura 10, la cual muestra que el AS tiene un *Grado de Semejanza Máxima (GSM)* de 440, aproximadamente en los 124 segundos de AP.

El GSM es el valor  $n[i]$  máximo absoluto que se alcanza en el vector  $n$ . Lo que indicaría, para este caso, que al estar  $n[i]_{max}$  por debajo del NMS, estas señales tienen un grado similitud pobre.

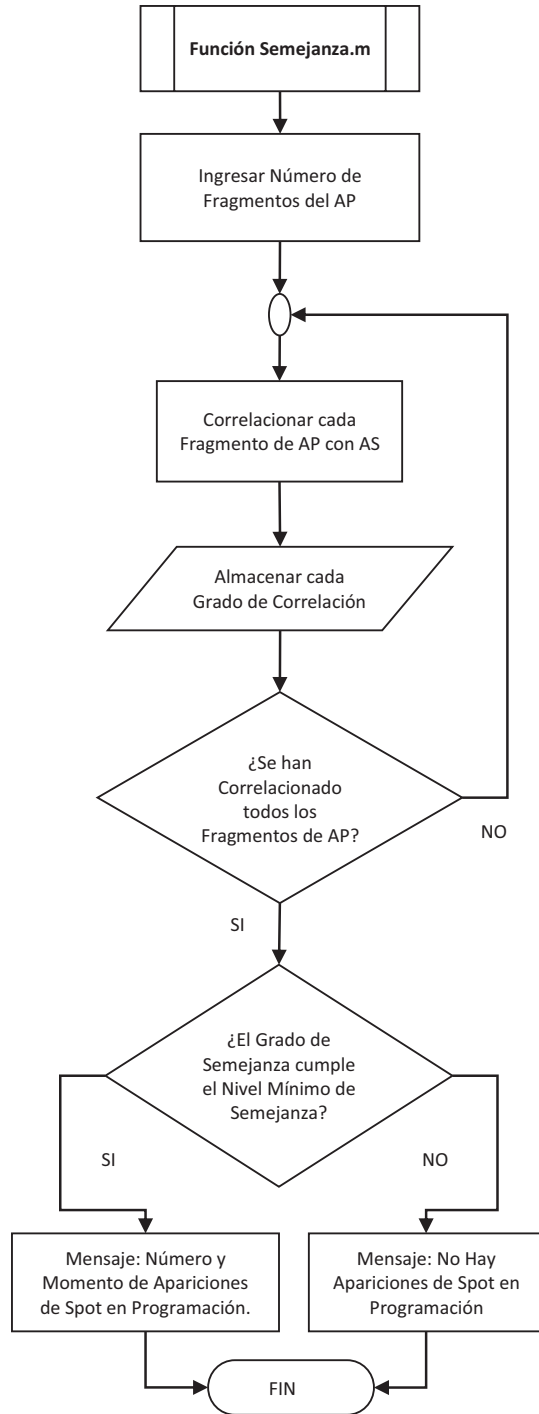


Figura 8 Diagrama de Flujo de subrutina para cotejar dos archivos y determinar su semejanza.

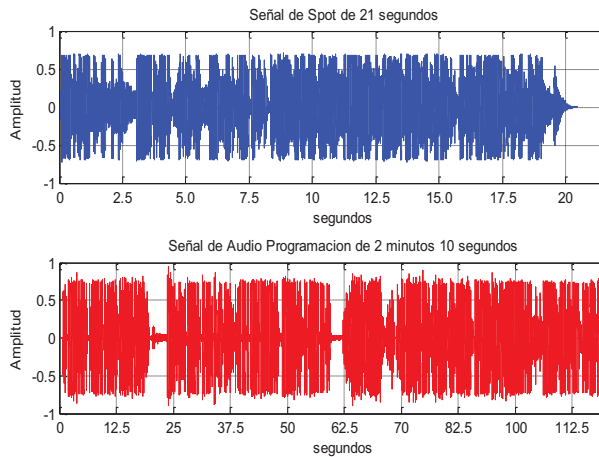


Figura 9 Representación de Señales Audios en Amplitud vs Tiempo

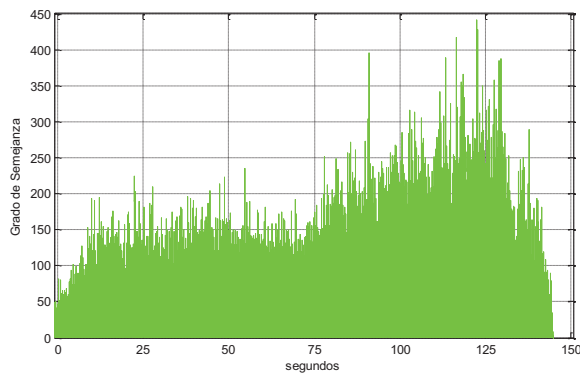


Figura 10 Representación de Similitud en Amplitud y Tiempo

Tomando otro ejemplo, si se realizara el cotejamiento de dos señales iguales, es decir, realizar la correlación de la misma señal; esto se denomina como una Auto Correlación o Auto Semejanza. La Figura 11 muestra el resultado de la auto-semejanza y permite aclarar el panorama del Nivel Mínimo de Semejanza.

Como se han cotejado dos señales idénticas, esto se refleja que el Grado de Semejanza se encuentra alrededor de 12000.

Al cotejar otras dos señales idénticas se tiene el resultado mostrado en la Figura 12. En esta ocasión el grado de correlación máxima llega hasta casi a los 20000.

La variación del valor máximo de la auto semejanza entre una y otra, dependerá exclusivamente, de las características de onda del audio auto correlacionado.

Así, en la correlación de los Audios Programación y Spot, como la Figura 10, lo que se toma en cuenta, es que, el Grado de Semejanza sobrepase el Valor Mínimo requerido para sentenciar si existe o no una similitud entre dos señales, que como se dijo antes está por encima de los 1500.

Se realizaron 100 auto-correlaciones y todas ellas mostraban valores mayores a los 10000. Se considera

como señales no similares a las que están por debajo de los 500.

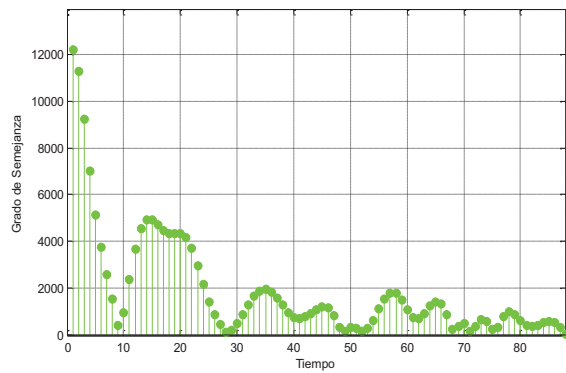


Figura 11 Primera Representación de Auto Correlación

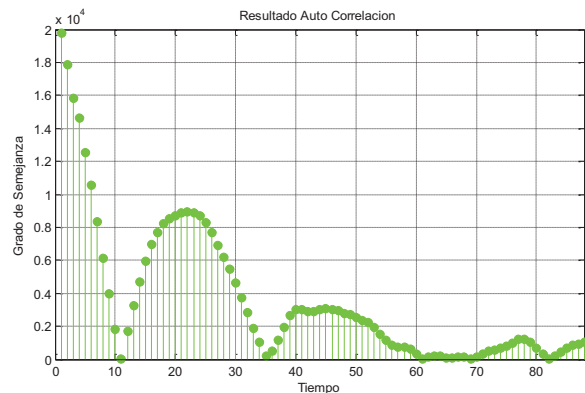


Figura 12 Segunda Representación de Auto Correlación

### E. Presentación de Informe

Los resultados del cotejamiento o correlación son presentados, mediante una interfaz gráfica, mediante un formato que permitan ser entendidos por el operador, que posea conocimientos básicos del sistema. Además de mostrar estos resultados en pantalla, existe la posibilidad de obtener una hoja imprimible.

Los datos generados en el procesamiento son almacenados en un archivo denominado *reporte.xls*. Posteriormente se puede obtener la opción de imprimir un reporte.

El reporte imprimible se visualiza por medio del programa Microsoft Excel 2007. Matlab permite la interacción con Microsoft Excel 2007, gracias a funciones como: *xlsread* y *xlswrite*. Estas funciones permiten abrir y editar, respectivamente, archivos con extensión xls y.xlsx.

## III. RESULTADOS

Se realizó el cotejamiento de 3 AP diferentes, frente a 30 AS. Para conseguir los objetivos propuestos en este trabajo, se analiza la detección positiva, detección negativa, falsos positivos y los tiempos de respuesta de SpotID.

A. DETECCION POSITIVA

De antemano se tiene conocimiento que solo 18 AS deberían aparecer dentro de los AP. Luego de 90 correlaciones, los resultados muestran una eficacia del 100%. Lo que representa que SpotID reconoció todas las apariciones que se deseaba encontrar y no halló algún Falso Positivo (spot erróneo). También se debe considerar que exista la probabilidad que, al terminar el escaneo, no se revelen coincidencias cuando en realidad sí existen, esto se ha declara como Detección Negativa.

Considerando lo anterior, se debería tener al menos 72 falsos positivos que pudieran haberse detectado, y tal como muestra la TABLA III no se reporta ni uno.

TABLA III  
RESULTADO DE COTEJAMIENTO

Cotejamiento	Detección Positiva	Detección Negativa	Falso Positivo
A	9/9	0/9	0/21
B	5/5	0/5	0/25
C	4/4	0/4	0/26
	<b>18/18</b>	<b>0/4</b>	<b>0/72</b>
Porcentaje Total Hallado	<b>100%</b>	<b>0%</b>	<b>0%</b>

B. TIEMPOS DE RESPUESTA

Los tiempos de cada cotejamiento se tabulan para luego promediarlos, con el fin de obtener una tendencia del tiempo que emplea SpotID en analizar una propaganda.

TABLA IV  
TIEMPOS DE RESPUESTAS POR COTEJAMIENTO

AS \ AP	Tiempo de Cotejamiento (segundos)		
	A	B	C
Asepsia.wav	32.14	32.76	30.80
Atunlsa.wav	32.41	32.87	32.70
Avon.wav	32.15	31.94	31.62
Axe.wav	32.21	32.57	31.20
BanPich.wav	32.01	32.46	31.72
BigCola.wav	50.5	51.16	55.22
Cicatricure.wav	32.37	32.79	33.77
Claro.wav	34.1	32.39	31.41
Colineal.wav	32.34	33.24	31.58
CoopNac.wav	32.62	32.9	32.19
CruzAzul.wav	32.31	32.78	26.26
Chaide.wav	32.6	33.51	32.39
Diners.wav	32.53	33.46	33.03
Experta.wav	33.17	33.62	32.78

AS \ AP	Tiempo de Cotejamiento (segundos)		
	A	B	C
Favorita.wav	32.86	33.62	31.84
Girasol.wav	32.49	33.13	32.27
HInds.wav	33.11	33.93	31.87
Lotto.wav	32.8	33.11	31.55
Marcimex.wav	32.31	32.72	33.98
Mensajea.wav	32.15	32.96	31.00
Movistar.wav	50.94	51.07	52.37
NatGard.wav	32.28	32.85	31.51
OpGMO.wav	33.02	33.73	31.44
Pañalin.wav	32.09	32.24	32.43
Pepsjumb.wav	32.81	33.68	32.12
Pozogem.wav	32.14	32.66	31.26
RespMej.wav	32.38	32.88	31.48
RevCiu.wav	48.81	49.08	49.43
Scot.wav	32.19	32.65	30.95
Sensodyne.wav	32.61	32.81	31.19

En la Figura 13 muestra los tiempos promediados de los cotejamientos A, B y C. Así, de manera grafica puede deducirse que en promedio, el tiempo que SpotID invierte en buscar un anuncio publicitario AS mayor a 20 segundos de duración en una programación AP de 60 minutos, es de 32 segundos.

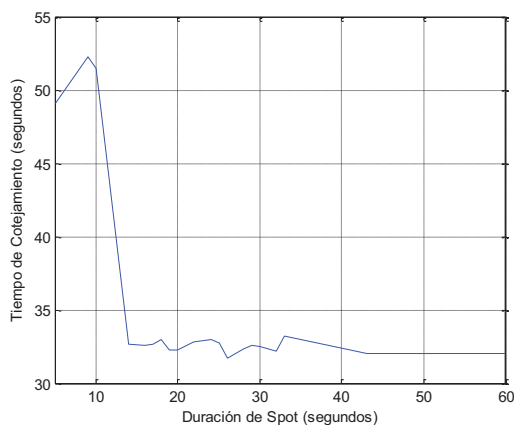


Figura 13 Tendencia de Cotejamiento

También en la Figura 13 se puede notar que mientras el archivo de spot sea menor a los 15 segundos, SpotID tarda más en procesarlos. Esto se debe exclusivamente a la manera de trabajar de la función *xcorr* de Matlab, sin que esto sea causa de falsos positivos.

Sumando el tiempo que conlleva cada proceso del sistema, se puede concluir que: Desde el momento de capturar una hora de programa hasta obtener una respuesta, se emplean alrededor menos de 65 minutos, según la TABLA V

TABLA V

COMPARACIÓN DE TIEMPOS DE BÚSQUEDA

Tiempo Empleado en Hallar un Spot (minutos)		
Actividad	Manualmente	SpotID
Captura	60	60
Extracción Audios	0	2.5
Búsqueda y Presentación de Resultados	30	0.7
<b>Total</b>	90	63.2

#### IV. CONCLUSIONES

Basado en los resultados de las pruebas, se puede concluir que el objetivo de reconocer publicidad se cumple. Se destaca el grado de asertividad de un 100%, reforzando el hecho con las decenas de pruebas realizadas durante el desarrollo y que no han sido transcritas en este trabajo escrito.

Además, el prototipo cumple con el segundo objetivo propuesto, que es la optimización del tiempo; pues, realiza el mismo trabajo de un ser humano pero en menor tiempo.

Aunque, durante las pruebas de SpotID, se manifiesta que el cerebro humano logra rebasar a la máquina, puesto que en menos de diez segundos, el hombre puede detectar y reconocer varias imágenes de comerciales al mismo momento en que las está viendo, aunque esta sea, la primera vez que lo vea. El desafío para el cerebro empieza cuando tiene que revisar videos de más de dos horas de duración.

La calidad de grabación del audio tiene un efecto mínimo. Según pruebas durante el diseño, se obtienen los mismos resultados si se procesan archivos con mayor velocidad de bits a 8kbps (audios con mayor calidad) lo que también aumenta el tiempo de procesamiento y el mayor consumo de memoria. Cabe recalcar que SpotID sólo trabaja con registros de 8kbps monofónicos.

El mismo efecto se tendría si la calidad de la grabación se la registra en modo estéreo; sólo aumentaría el tiempo de procesamiento y mayor consumo de memoria de almacenamiento.

El resultado de la correlación cruzada son valores negativos y positivos, en este proyecto la correlación cruzada es presentada como valor absoluto, sin ver afectación alguna en el cotejamiento y los resultados.

La correlación cruzada realiza desplazamiento del tiempo hacia adelante y atrás del archivo AS, solo es necesario para los tiempos positivos debido a que la muestra del spot únicamente se lo hace en función de la programación en adelante.

El tiempo de captura de los AP debe ser menor a 12 horas para facilitar el manejo y procesamiento de estos archivos. Aunque si el poder de procesamiento del computador lo permite, SpotID está en la capacidad de

manejar el tamaño máximo de un archivo WAV que es alrededor de 4.3GB.

#### V. REFERENCIAS

- [1] Spot homepage on RAE [Online]: <http://lema.rae.es/drae/?val=spot>
- [2] John G. Proarkis, D. G. Tratamiento Digital de Señales. Prentice Hall. Madrid: 1998.
- [3] Transmisiones Radiofónicas y de Televisión. Valor Único de Relación Señal/Ruido para todos los Sistemas de Televisión. Recomendación UIT-T J.62
- [4] (2010) DVDvideoSoft website. [Online]. Disponible: <http://www.dvdvideosoftware.com/>
- [5] xcorr página de inicio de MathWorks [Online]: <http://www.mathworks.es/help/toolbox/signal/ref/xcorr.html>
- [6] Bertran, Eduard. Señales y Sistemas de tiempo discreto. Ediciones UPC. Catalunya. 2003.
- [7] Tomado y adaptado de definición de Multilateración en : <http://en.wikipedia.org/wiki/Multilateration>

#### VI. BIOGRAFÍAS

##### Juan Alfredo Orrala Moreno



Nació en Santo Domingo de los Tsáchilas, en 1979. Sus estudios secundarios los realizó en el Colegio Nacional Técnico “Dr. Luis Céleri Avilés”, de la ciudad de La Libertad, provincia de Santa Elena. Ingeniero en Electrónica y Telecomunicaciones de la

Escuela Politécnica Nacional 2012. Ejerce su profesión en el área de Infraestructura de Red del Departamento de Proyectos UPC a Nivel Nacional. Sus áreas de interés involucran el desarrollo de software y aplicaciones iOS y Android.

##### Ing. Jaime Edison Velarde Guevara



Graduado en Electrónica y Telecomunicaciones - 1980 EPN. Profesor del DETRI desde 1977. Participante por la EPN en la Red ALFA: “Creación de una biblioteca de Objetos de Aprendizaje, orientados a la formación en las TICs”.

Administrador del primer Campus Virtual de la EPN para la capacitación de los docentes con FATLA “Fundación para la aplicación de la tecnología en Latinoamérica”. Participante en el Proyecto de Investigación: SIMON-VE “Sistema Integral de Monitoreo de Vehículos”. [jaime.velarde@epn.edu.ec](mailto:jaime.velarde@epn.edu.ec)