



DEPARTAMENTO DE CIENCIAS FORENSES
ORGANISMO DE INVESTIGACION JUDICIAL (OIJ)
PODER JUDICIAL, COSTA RICA
**Comparación biométrica de voces mediante
software Batvox**

**PROCEDIMIENTO DE
OPERACIÓN NORMADO
ESPECIFICO**
P-DCF-ECT-ISF-17

VERSION: 05

Rige desde: 22/12/2022

PAGINA: 1 de 8

| | |
|--|--|
| Elaborado o modificado por: Lic. Laura Guerra Vargas Técnico en Video Forense | Revisado por Líder Técnico: Lic. José Alfonso Rodríguez Arias Líder Técnico de Sección/Unidad de Acústica Forense |
| Visto Bueno Encargado de Calidad: Lic. José Alfonso Rodríguez Arias Encargado de Calidad de la Sección de Imagen y Sonido Forense | Aprobado por: Lic. Marco Herrera Charraun Jefatura, Sección Imagen y Sonido Forense |

CONTROL DE CAMBIOS A LA DOCUMENTACIÓN

| Versión | Fecha de Aprobación | Fecha de Revisión | Descripción del Cambio | SCD | Solicitado por |
|----------------|----------------------------|--------------------------|--|------------------|-----------------------|
| 01 | 29/02/2016 | 06/05/2020 | Versión Inicial del Procedimiento | - | RVF |
| 02 | 06/05/2020 | 28/09/2021 | Actualización y cambios en redacción | 028-SCD-ISF-2020 | MHC |
| 03 | 28/09/2021 | 16/09/2022 | Agregar verificación de software y mención de población de referencia utilizada. Cambio de imágenes. | 002-SCD-ISF-2021 | MHC |
| 04 | 16/09/2022 | 22/12/2022 | Inclusión acerca de indicación de resultado e interpretación en el dictamen | 004-SCD-ISF-2022 | MHC |
| 05 | 22/12/2022 | - | Agregar indicación de validación, cambios en redacción y eliminación de imágenes | 005-SCD-ISF-SCD | MHC |

**ESTE PROCEDIMIENTO ES UN DOCUMENTO CONFIDENCIAL
PARA USO INTERNO DEL DEPARTAMENTO DE CIENCIAS FORENSES
SE PROHÍBE CUALQUIER REPRODUCCIÓN QUE NO SEA PARA ESTE FIN**

La versión oficial digital es la que se mantiene en la ubicación que la Unidad de Gestión de Calidad defina. La versión oficial impresa es la que se encuentra en la Unidad de Gestión de Calidad. Cualquier otro documento impreso o digital será considerado como copia no controlada .

| | | |
|--|------------------|----------------|
| DEPARTAMENTO DE CIENCIAS FORENSES | VERSIÓN 05 | PAGINA: 2 de 8 |
| Comparación biométrica de voces mediante software Batvox | P-DCF-ECT-ISF-17 | |

1 Objetivo:

Estandarizar la metodología para realizar análisis de comparación biométrica de voz, emitiendo resultados mediante cálculos de LR.

2 Alcance:

Este procedimiento se refiere a la comparación de dos o más grabaciones de audio en las que intervengan voces, con la finalidad de clasificar una señal de voz de origen desconocido, como perteneciente a una señal entre un conjunto de posibles locutores, por medio de análisis biométrico de la voz, utilizando el software forense Batvox.

El uso de este procedimiento es de acatamiento obligatorio para todos los miembros de la Unidad de Acústica Forense, que realicen comparaciones, utilizando el software biométrico Batvox, de locutores en dos o más grabaciones donde intervengan locutores que se requieran someter a comparaciones.

[Esta metodología se encuentra validada en el Informe de validación 001-ISF-VAL-2022.](#)

3 Referencias:

- Agnitio Voice ID. (2015). Manual de Usuario. Batvox Basic 4.1. V1.1.
- Andrzej Drygajlo, Michael Jessen, Stefan Gfroerer, Isolde Wagner, Jos Vermeulen and Tuija Niemi. Methodological Guidelines for Best Practice in Forensic Semiautomatic and Automatic Speaker Recognition, including Guidance on the Conduct of Proficiency Testing and Collaborative Exercises. European Network of Forensic Science Institutes (ENFSI). 2015.
- Barros, José Luis y Moser, Michael. (2009). Ingeniería Acústica. España: Springer.
- European Network of Forensic Science Institutes (ENFSI). (2015). Guideline for Evaluative Reporting in Forensic Science.
- Evett IW, Jackson G, Lambert JA and McCrossan. "The impact of the principles of evidence interpretation on the structure and content of statements". Science & Justice, 40 (4) 233-239, 2000.
- Evett IW. Towards a uniform framework for reporting opinions in forensic science casework. Science & Justice; 38(3) 198-202. 1998.
- Procedimiento "Gestión de indicios de imagen y sonido mediante QuickDME". Sección Imagen y Sonido Forense. Departamento de Ciencias Forenses. 2018.

4 Equipos y Materiales:

4.1 Equipos

- Audífonos profesionales de orejera completa.
- Computadora de trabajo.
- Software Adobe Audition.

| | | |
|--|------------------|----------------|
| DEPARTAMENTO DE CIENCIAS FORENSES | VERSIÓN 05 | PAGINA: 3 de 8 |
| Comparación biométrica de voces mediante software Batvox | P-DCF-ECT-ISF-17 | |

- Software Batvox 4.1.2.

4.2 Materiales

- Formulario "Mantenimiento Preventivo y/o Verificación de Software".
- Formulario de Trabajo "Análisis Preliminar de Muestras de Voz y Comparación Biométrica".
- Formulario de Trabajo "Espacio Adicional (Anexo de todos los formularios periciales)".

5 Reactivos y Materiales de Referencia:

N/A.

6 Condiciones Ambientales:

N/A.

7 Procedimiento:

Nota 1: Realice verificación del funcionamiento del software Batvox al menos una vez al año con muestras de voz indubitadas, deje registro de esto en el formulario de "Mantenimiento Preventivo y/o Verificación de Software", indicando cuáles muestras se utilizaron y el resultado del mismo, anote además si el resultado es el esperado, en caso de no serlo verifique nuevamente. Si el resultado sigue siendo distinto al esperado pese a realizar varias veces la verificación coloque un rótulo de "Equipo en mal estado" hasta que se logre determinar por medio del soporte técnico especializado el motivo del resultado distinto y corregirlo posteriormente.

Nota 2: Los archivos de audio que se sometan a este procedimiento de comparación, deben cumplir con los parámetros establecidos en el Procedimiento Análisis Preliminar de Muestras de Voz para Comparación de voces y estar consignado por medio del formulario de trabajo "Análisis Preliminar de Muestras de Voz y Comparación Biométrica".

7.1 Prepare la muestra en el programa Adobe Audition o similar para ser ingresada al Batvox, con las siguientes características digitales:

-Formato WAV con codificación lineal PCM sin compresión.

-Frecuencia de muestreo 8 KHz.

-Resolución 16 bits.

-Canales: mono.

-Se deben eliminar vocalizaciones que no son del habla, como por ejemplo: risas, tos, carraspeos, clics y respiración, o estilos de habla inusuales, susurros, gritos, entre otros. De igual forma se deben eliminar pausas.

-En caso de que un archivo de audio presente más de una voz, se debe proceder a separar las mismas, de manera que se genere un archivo de audio para cada uno de los hablantes.

7.2 Deje registro de todo lo indicado en este punto en el formulario de trabajo "Análisis Preliminar de Muestras de Voz y Comparación Biométrica".

| | | |
|--|------------------|----------------|
| DEPARTAMENTO DE CIENCIAS FORENSES | VERSIÓN 05 | PAGINA: 4 de 8 |
| Comparación biométrica de voces mediante software Batvox | P-DCF-ECT-ISF-17 | |

7.3 Accese al software forense de comparación de voces Batvox y cree un nuevo caso.

7.4 Identifique el caso en el sistema con el número único y/o número de orden de trabajo, si tiene algún comentario que agregar realícelo en esta misma ventana.

7.5 Cargue todas las muestras de audio que se van a utilizar durante la comparación (Audios de entrenamiento, de test, o impostores).

7.6 Seleccione las características de grabación (canal, idioma, tipo de habla y si es analógico) dentro de las cajas de diálogo que el programa solicita, marcando las que correspondan al audio o audios que va a cargar. Repita esta acción si hay diferencias entre los audios que va a ingresar, por ejemplo si entre los audios dubitados e indubitados no coincide el tipo de habla, el canal u otro.

7.7 Entrene la o las muestras que serán utilizadas como audio de entrenamiento o modelo.

7.8 Realice una verificación de la idoneidad de los audios ingresados. Debe posicionarse con el cursor sobre los mismos uno por uno, y en la parte inferior de la ventana se despliega la información de cada uno. Debe cumplir con un mínimo de 30 segundos para el audio modelo y 20 segundos para los audios de test, y tener un mínimo de 10 dB SNR.

7.9 Cree una nueva sesión, identificándola con el número único y/o número de orden de trabajo. Si tiene comentarios agréguelos en el espacio destinado para este fin.

7.10 Dentro de la nueva sesión genere un nuevo cálculo LR, identifique el mismo con las siglas LR más el número único y/o número de orden de trabajo, por ejemplo "LR 2019-05879-ISF" o "LR 12-00000X-00XX-PE".

7.11 Sobre el nuevo cálculo LR, despliegue las opciones y asocie los audios que serán utilizados en el análisis a la pestaña que corresponda: audios de test, audios impostores o modelos.

7.12 Una vez asociados los audios, seleccione la población de referencia a cada uno de los audios que se utilizarán como de entrenamiento o modelo.

7.13 Debe especificar una por una las características que correspondan al modelo y marcar la opción de "Población Óptima".

7.14 Evalúe la calidad de la población de referencia elegida para cada modelo, mediante la opción "Lanzar Calidad Población".

7.15 Verifique la información del resultado que el programa despliega donde indica si la población representa correctamente al modelo. En caso de que el mensaje indicado sea "La población seleccionada no parece representar correctamente las características del modelo de sospechoso", proceda a revisar las características del modelo del sospechoso y la población de referencia que se utilizó, para que las características de ambos coincidan. En caso de que siga dando esa alerta proceda a ingresar una población de audios impostores como otra forma de normalización. Si el error persiste debe volver a iniciar el caso hasta que el mensaje sea: "La población seleccionada parece representar correctamente las características del modelo sospechoso".

| | | |
|--|------------------|----------------|
| DEPARTAMENTO DE CIENCIAS FORENSES | VERSIÓN 05 | PAGINA: 5 de 8 |
| Comparación biométrica de voces mediante software Batvox | P-DCF-ECT-ISF-17 | |

Nota 3: Dicha población de referencia no se considera una colección de referencia debido a que estos audios son utilizados para realizar los procesos de normalización recomendados para llevar a cabo tareas de reconocimiento del software, no para ser utilizados como audios de test o modelo, por lo tanto no se requiere una identificación única para los audios de esa población ya que no se van a realizar identificaciones con la misma.

7.16 Ejecute el cálculo LR mediante la opción "Lanzar Cálculo LR".

7.17 Observe la tabla de resultados que genera el programa, los indicados en color rojo son los que están por debajo del umbral 1 o con LR inferior a 1, los verdes indican LR superior a 1.

7.18 Guarde los datos de los resultados, el log de entrenamiento y las imágenes de los gráficos de curvas de intervariabilidad e intravariabilidad.

7.19 Convierta el reporte de resultados y el log de entrenamiento a formato PDF y agréguelos al legajo digital del SADCF.

7.20 Incluya las imágenes de los gráficos de curvas de intervariabilidad e intravariabilidad dentro del formulario "Espacio Adicional", posteriormente incorpore el formulario al legajo digital del caso en el SADCF.

7.21 Deje registro de cada uno de los resultados en el formulario de trabajo "Análisis Preliminar de Muestras de Voz y Comparación Biométrica", incorpore el mismo al legajo digital del caso en el SADCF.

8 Criterios de Aceptación o Rechazo de Resultados:

8.1 Los criterios de aceptación o rechazo se definen en el procedimiento de "Análisis Preliminar de Muestras de Voz y Comparación Biométrica".

9 Cálculos y evaluación de la incertidumbre:

N/A.

10 Reporte de Análisis y Resultados:

Nota 4: El LR es una relación de dos probabilidades, que la voz del sospechoso sea la del dubitado o que la voz del sospechoso pertenezca a un tercero cuya grabación se hubiera realizado en condiciones acústicamente similares, es decir que no corresponda a la del dubitado.

Si el resultado del LR es igual a 1, la voz del sospechoso pertenece al dubitado, con una baja probabilidad. Cuanto más aumente el número del LR, mayor será la probabilidad de que la voz del sospechoso pertenece al dubitado. Cuando el resultado es superior a 1, la línea de evidencia, queda dentro de la curva de intravariabilidad (color azul). Indica que es más probable que el audio modelo sea la misma persona respecto al audio cuestionado, y muestra cómo puntúa el sospechoso contra sí mismo.

Si el LR es menor a 1, y cuanto más bajo sea, aumenta la probabilidad de que los audios en cuestión no corresponden a la misma persona. Cuando el resultado del LR es inferior a 1, la

| | | |
|--|------------------|----------------|
| DEPARTAMENTO DE CIENCIAS FORENSES | VERSIÓN 05 | PAGINA: 6 de 8 |
| Comparación biométrica de voces mediante software Batvox | P-DCF-ECT-ISF-17 | |

línea de evidencia, queda dentro de la curva de intervariabilidad (color rojo). Indica que es más probable que el audio modelo no sea la misma persona respecto al audio cuestionado.

10.1 En el dictamen pericial indique como resultado el valor numérico que da el software, el mismo puede ser mostrado mediante la imagen de los resultados brindada por el software.

10.2 En el Dictamen Pericial, indique como interpretación si existe o no correspondencia entre las voces comparadas, explicando lo que significa en términos de probabilidad el valor de LR otorgado por el software y brinde un apoyo a una hipótesis según corresponda, siguiendo como referencia la escala verbal para expresar LR adaptada de las directrices de ENFSI para informes forenses, según lo indicado en los cuadros 1 y 2.

Por ejemplo: "Según el resultado del análisis y el LR obtenido, la comparación entre la voz contenida en la grabación del archivo X y la voz contenida en la grabación del archivo Y es LR 0,00047 (sin redondear 0,000470655546406432), lo que significa que la probabilidad de que ambas grabaciones pertenezcan a la misma persona es 0,00047.

Tomando como referencia el LR obtenido y la escala verbal para expresarlo, el perito que firma este dictamen brinda un apoyo fuerte a la hipótesis de que la voz en el archivo "X" no corresponde con la voz en el archivo "Y".

Cuadro 1

Cuando el valor del LR es superior a 1, se utiliza la escala verbal del siguiente cuadro:

| Soporte a hipótesis | Escala verbal | LR |
|---|----------------------------|------------------------|
| Voz dubitada corresponde con la del sospechoso | Apoyo limitado | $1 < LR \leq 10$ |
| | Apoyo moderado | $10 < LR \leq 100$ |
| | Apoyo moderadamente fuerte | $100 < LR \leq 1000$ |
| | Apoyo fuerte | $1000 < LR \leq 10000$ |
| | Apoyo muy fuerte | $LR > 10000$ |

Cuadro 2

Si el valor del LR es inferior a 1, se utiliza la escala verbal del siguiente cuadro:

| Soporte a hipótesis | Escala verbal | LR |
|--|----------------------------|--------------------------|
| Voz dubitada no corresponde con la del sospechoso | Apoyo limitado | $0.1 \leq LR < 1$ |
| | Apoyo moderado | $0.01 \leq LR < 0.1$ |
| | Apoyo moderadamente fuerte | $0.001 \leq LR < 0.01$ |
| | Apoyo fuerte | $0.0001 \leq LR < 0.001$ |
| | Apoyo muy fuerte | $LR < 0.0001$ |

| | | |
|--|------------------|----------------|
| DEPARTAMENTO DE CIENCIAS FORENSES | VERSIÓN 05 | PAGINA: 7 de 8 |
| Comparación biométrica de voces mediante software Batvox | P-DCF-ECT-ISF-17 | |

10.3 Indique en el Dictamen Pericial la población de referencia utilizada y sus características en el análisis de comparación de voces.

11 Medidas de Seguridad y Salud Ocupacional:

11.1 Nunca coloque los audífonos sobre los oídos antes de que el audio esté sonando, si el sonido inicia de golpe muy fuerte puede lastimarse los oídos.

12 Simbología:

dB: decibeles.

ENFSI: European Network of Forensic Science Institutes.

Kbps: Kilobit por segundo.

KHz: KiloHertz.

Hz: Hertz.

N/A: No aplica.

PDF: Portable Document Format.

PNG: Portable Network Graphics.

SADCF: Sistema Automatizado Departamento de Ciencias Forenses.

13 Terminología:

Audios de entrenamiento: Los audios de entrenamiento son aquellos audios con los que se generará un modelo del sospechoso que se utilizará en un cálculo biométrico.

Audios de impostor: Impostores son aquellos audios que no pertenecen al sospechoso y que han sido grabados bajo las mismas condiciones que los audios de test. Al igual que las poblaciones de referencia, los impostores se utilizan para llevar a cabo procesos de normalización para realizar tareas de reconocimiento fiables.

Audio de test: Audio que contiene voz cuya identidad es desconocida y que se utiliza para compararse con modelos de sospechoso con el objetivo de encontrar alguna coincidencia que permita identificar.

Batvox: Programa biométrico forense de comparación de voces.

Código hash: Algoritmo que consigue crear a partir de información registrada, un código alfanumérico de longitud normalmente fija que representa el contenido y asegura que no se ha modificado una transmisión o duplicación

Compresión: Proceso por el cual se reduce la tasa de bits de una señal digital de audio buscando como fin la reducción de su peso en disco, a medida que se comprime el audio se reducen estos valores, con lo cual se pierde fidelidad con respecto al sonido físico.

Curva de intervariabilidad: representa el rango de puntuaciones del audio de test cuando se enfrenta a una población de referencia (modelos que no son la misma persona que el audio de test).

| | | |
|--|------------------|----------------|
| DEPARTAMENTO DE CIENCIAS FORENSES | VERSIÓN 05 | PAGINA: 8 de 8 |
| Comparación biométrica de voces mediante software Batvox | P-DCF-ECT-ISF-17 | |

Curva de intravariabilidad: estimará como suele puntuar el sospechoso contra sí mismo.

Decibeles (dB): unidad de medida de la presión sonora.

Entrenamiento de modelo: Es el proceso biométrico en el que se extraen las características de la voz de una persona y se genera a partir de toda esta información un modelo biométrico.

Escala verbal para expresar LR: equivalente verbal utilizado en análisis científicos cuya unidad de medida es el likelihood ratio (LR).

Espectrograma: representación gráfica de la energía del contenido frecuencial de la señal de una de una grabación de sonido y sus variaciones en el transcurso del tiempo que ésta dura.

Formato de grabación: propiedades digitales de la forma en la que se almacenan los datos de una grabación de audio y cómo puede utilizarse en un sistema de administración de archivos.

Formato wav: formato de audio digital el cual normalmente se utiliza sin compresión de datos.

Frecuencia de muestreo: Es la cantidad de muestras de audio que se toman en un segundo. Esta se mide en Hertz (Hz). La cantidad de muestras tomadas definen la calidad de la muestra y su correspondencia respecto al original.

LR: Una relación de verosimilitud, o "likelihood ratios", es un valor probabilístico obtenido por aproximación Bayesiana. Un LR es una relación de probabilidades: por un lado la probabilidad de que la voz desconocida o audio de test pertenezca al sospechoso y por otro estará la probabilidad de que la voz pertenezca al resto del mundo (es decir, que no pertenezca al sospechoso).

Normalización: Proceso que se aplica en los cálculos de reconocimiento para corregir los desalineamientos que se producen en las puntuaciones debido a la naturaleza de los audios.

Habla neta: tiempo real de duración de un relato leído, espontáneo o dialogado sin silencios ni interrupciones de un mismo locutor.

Hertz (Hz): unidad de medida de frecuencia del Sistema Internacional de Unidades.

PCM: Se refiere a un método para representar una forma de onda de audio analógica con una serie de valores de muestra digital cuantizados.

Población de referencia: Una población de referencia es un grupo de modelos utilizados para realizar los procesos de normalización recomendados para realizar tareas de reconocimiento fiables. Cada modelo de sospechoso implicado en una tarea de reconocimiento tendrá asociada una población de referencia y dichos modelos de referencia deberán tener exactamente las mismas características acústicas que las del modelo de sospechoso.

14 Anexos

N/A.