

La identificación forense de hablantes

Guía de buenas prácticas y el servicio FORENSIA de cotejo de voces

•Miguel Martínez Soler, Jorge Gurlekian, Pedro Univaso, Humberto Torres, Diego Evin y German Stalker

Programa Ciencia y Justicia – CONICET y BlackVox

Identificación forense de hablantes

- Es decidir si la voz de una persona no conocida (evidencia, dubitada, “offender” en inglés) corresponde o no a la voz de una persona que si es conocida (sospechoso).



*“Conseguime
la guita”*



Identificación Forense

Marco de referencia: Bayes

Existen 2 Hipótesis

H1: E proviene del hablante sospechoso

H0: E proviene de cualquier otro
hablante

$$LR = \frac{p(E|H_1)}{p(E|H_0)}$$

LR > 0 aceptar H₁

LR < 0 aceptar H₀

LR es la Relación de Verosimilitudes

¿Como se interpreta la pericia?

LR	LLR	Interpretación
0,0001	-4	Falta de certeza muy fuerte
0,001	-3	Falta de certeza fuerte
0,01	-2	Falta de certeza moderada
0,1	-1	Falta de certeza limitada
1	0	Indefinición
10	1	Certeza limitada
100	2	Certeza moderada
1000	3	Certeza fuerte
10000	4	Certeza muy fuerte

Objetivos

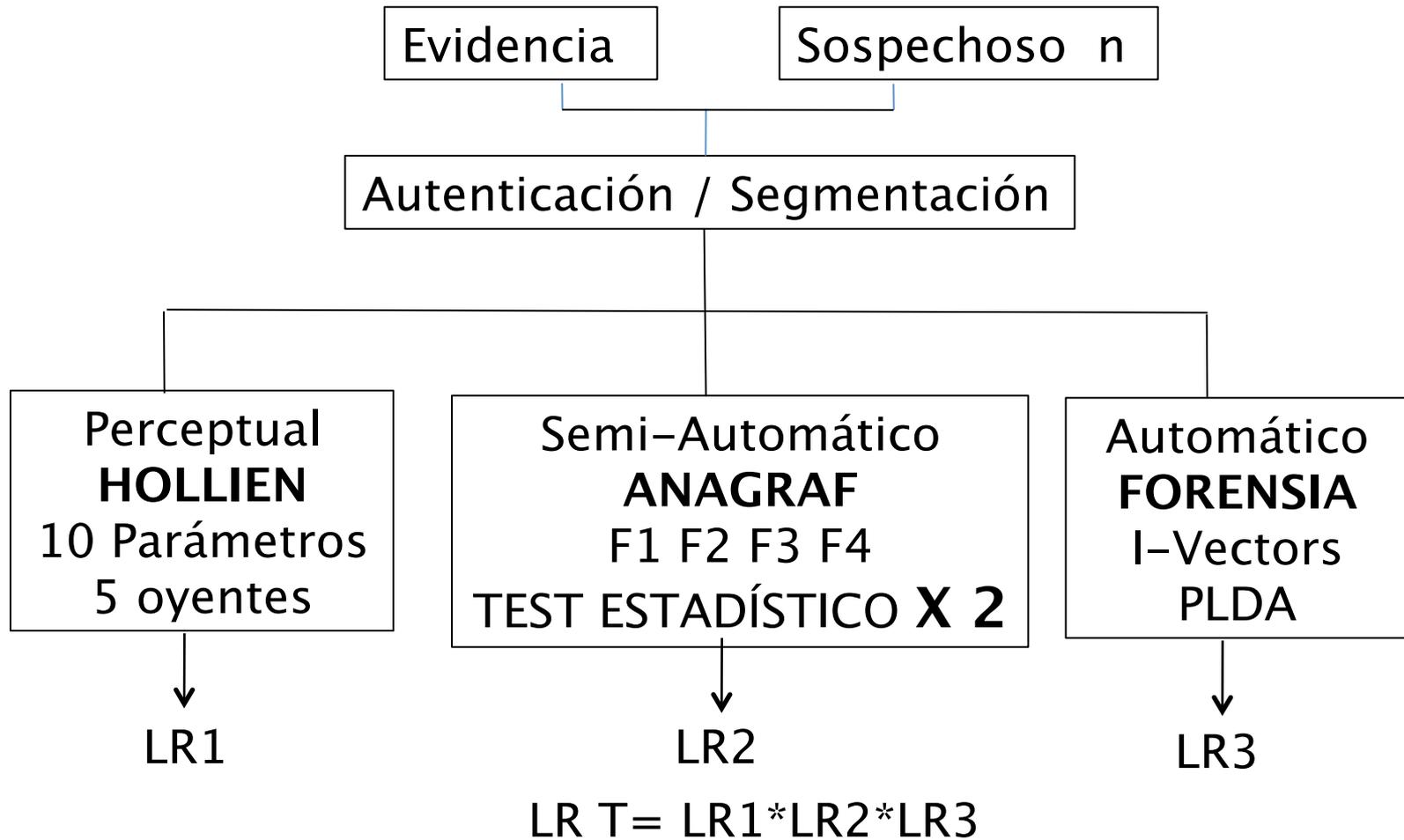
- Sistematizar capacidades, tecnologías y asesorías que se proveen al Poder Judicial
- Fortalecer vínculos con la Justicia
- Ofrecer servicios forenses y capacitación.

Guía de buenas prácticas: etapas en las pericias de voz

1. La recepción de las grabaciones *dubitadas*.
2. La preparación del oficio judicial solicitando la pericia a un perito independiente o institución.
3. La citación del sospechoso o indubitado.
4. Tareas del Perito
 - a. La verificación de la autenticidad y calidad .
 - b. La creación de las grabaciones *indubitadas*.
 - c. El análisis y cotejo de voces.
 - d. La preparación del informe pericial.

El servicio FORENSIA de cotejo de voces

Sistemas de Identificación Forense



FORENSIA: Sistema de Identificación Forense de hablantes

- **Realizado por:** BlackVox–CONICET para realizar el cotejo de voces
- **Entrada:** la grabaciones de la evidencia y del sospechoso.
- **Salida:** LLR medido en relación a la base de datos universal y las bases de referencia.
- **Base datos universal:** con grabaciones de las distintas regiones de la Argentina y grabaciones de bases de datos extranjeras.
- **Motor de reconocimiento:** basado en l-vectors

Especificaciones técnicas

Pre-procesamiento y Extracción de rasgos

- Técnica de supresión de ruidos y filtrado lineal .
- Ventana de 25ms cada 10 ms
- Banco 24 filtros de 40 a 4000Hz. 12 Coeficientes MFCC, Energía y derivadas.
- VAD basado en Energía para etiquetar habla y no habla
- Normalización de cada segmento con el valor medio y varianza del cepstrum.

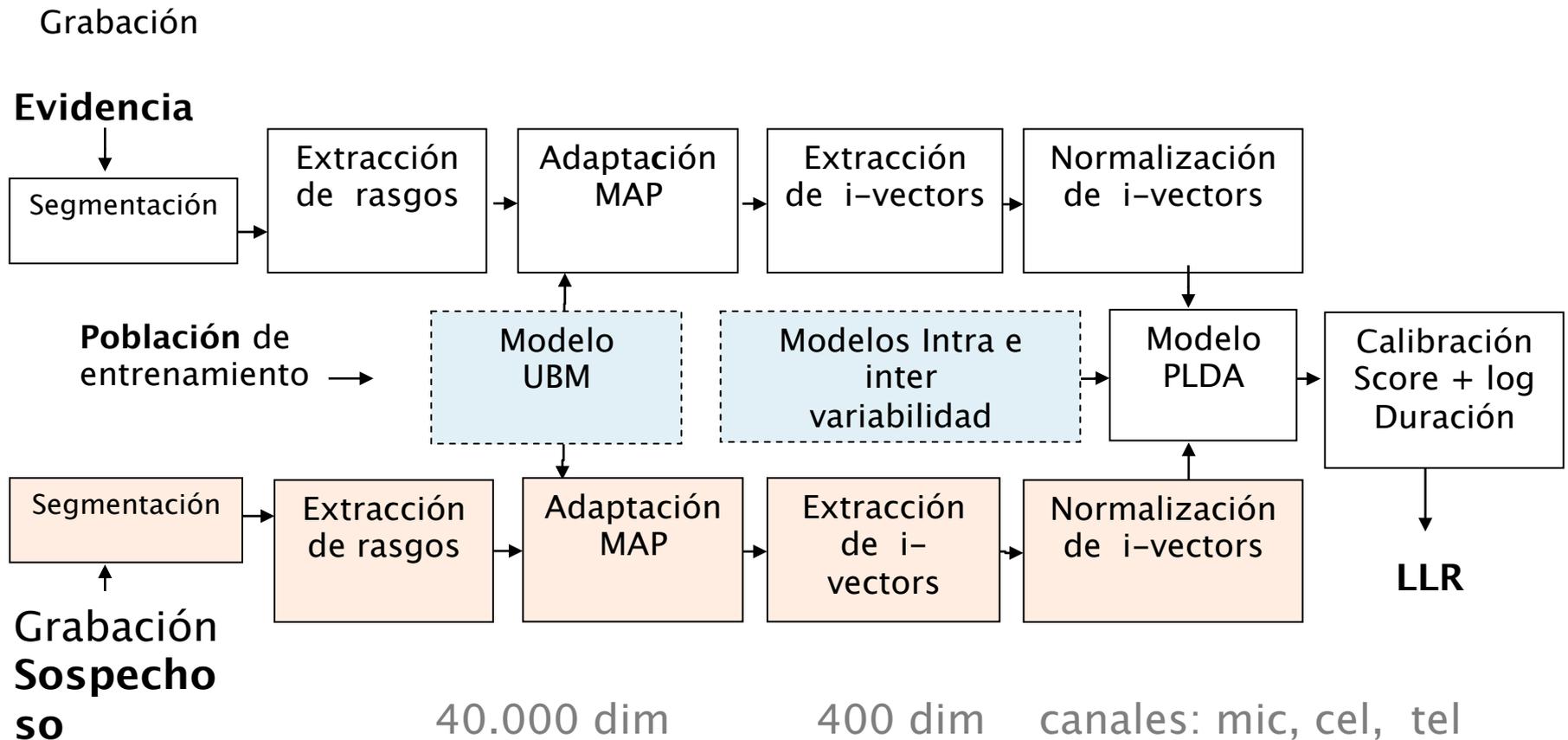
Sistema

- Modelo UBM de 1024 mezclas gaussianas de 40.000 componentes
- Modelo de variabilidad total entrenado con un conjunto de desarrollo para producir I-vectores de dimensión 400.
- Normalización de ruido de los I-vectores seguido de un LDA sin reducción de la dimensionalidad.
- Modelo PLDA entrenado con el conjunto de desarrollo conteniendo 300 dimensiones para la voz y 100 dimensiones para el canal.

Calibración

- Combinación lineal del puntaje de salida y el logaritmo de la duración de los segmentos a evaluar.

Sistema Automático FORENSIA



<http://www.BlackVOX.com.ar>

Principios de Daubert

- El método utilizado debe ser conocido en su totalidad. Se debe tener acceso público.
- Debe darse información sobre los márgenes de error del método.
- Deben indicarse las bases de datos utilizadas en los métodos automáticos y semi automáticos: Universal UBM, de referencia, del sospechoso y de calibración.
- El método debe estar validado con una base de datos de la institución que valida.

Validación del Sistema FORENSIA

- **Marca y modelo del sistema a validar:** FORENSIA2.0 BlackVOX CONICET
- **Metodología:** basada en la guía metodológica propuesta por la Red Europea de Institutos de Ciencias Forenses.
- **Base de datos de evaluación:**
 - **SITW 2016 de SRI International:** 648 pares de emisiones empleando canales de telefonía–micrófono (cellphone–lapel) y 110 pares de emisiones empleando canales micrófono–micrófono (lapel–lapel).
 - **Morrison 2016** conformada por 1250 pares de emisiones empleando canales de telefonía–telefonía.
- Las grabaciones han sido previamente editadas manualmente eliminando segmentos de habla superpuestas, ruidos y otras anomalías que se incluyen en las grabaciones originales (perturbaciones técnicas locales y sonidos no usuales). No se realizó ningún proceso de mejoramiento del audio (e.g. Filtrado).
- Morrison, G. S., & Enzinger, E. (2016). Multi-laboratory evaluation of forensic voice comparison systems under conditions reflecting those of a real forensic case (forensic_eval_01)–Introduction. *Speech Communication*, 85, 119–126.

Métricas de evaluación

- Dato numérico que describe la performance del sistema de identificación en términos de precisión, confiabilidad, poder de discriminación y calibración.
- **Cllr**: la función de costo logarítmica (Cllr) mide la efectiva cantidad de información que el sistema de identificación le brinda al usuario, en forma independiente de la aplicación empleada.
- **Cllr min**: es el valor mínimo que puede obtenerse del sistema de identificación.
- **Cllr cal**: es la pérdida de calibración del sistema de identificación. $Cllr = Cllr\ min + Cllr\ cal$
- **EER%**: la tasa de igual error (EER%) es la tasa de error en la que la probabilidad de falsas alarmas es igual a la probabilidad de casos perdidos.
- **PME H0**: probabilidad de evidencia engañosa en favor de la hipótesis H1. Es la probabilidad de todos los LLR que son menores de 0, sabiendo que la hipótesis H0 es verdadera.
- **PME H1**: probabilidad de evidencia engañosa en favor de la hipótesis H0. Es la probabilidad de todos los LLR que son mayores de 0, sabiendo que la hipótesis H1 es verdadera.
- **95% CI**: intervalo de credibilidad de que en el 95% de los casos una medición de LLR se encuentra acotada por dicho intervalo. Corresponde a la confiabilidad de las mediciones del sistema (1).
- (1) Morrison, G. S. (2011). Measuring the validity and reliability of forensic likelihood-ratio systems. *Science & Justice*, 51(3), 91–98.

Validación Forensia: Micrófono – Micrófono

- El criterio de validación presenta condiciones relacionadas con las métricas de evaluación que deben cumplirse como una condición necesaria para que el sistema de identificación sea considerado válido.

Métricas de evaluación	Resultado	Criterio de validación	Cumplimiento del criterio de validación	Tipo de parámetro
Cllr	0.18	0.65	Si	Precisión
Cllr Min	0.17	0.46	Si	Discriminación
Cll cal	0.01	0.05	Si	Calibración
EER%	6%	13%	Si	Discriminación
PME H1	0.06	0.30	Si	Precisión
PMEH0	0.06	0.15	Si	Precisión
95% CI	1.38	1.50	Si	Confiabilidad

Validación Forense: Teléfono – Teléfono

- El criterio de validación presenta condiciones relacionadas con las métricas de evaluación que deben cumplirse como una condición necesaria para que el sistema de identificación sea considerado válido.

Métricas de evaluación	Resultado	Criterio de validación	Cumplimiento del criterio de validación	Tipo de parámetro
Cllr	0.46	0.65	Si	Precisión
Cllr Min	0.46	0.46	Si	Discriminación
Cll cal	0.00	0.05	Si	Calibración
EER%	8%	13%	Si	Discriminación
PME H1	0.08	0.30	Si	Precisión
PMEH0	0.08	0.15	Si	Precisión
95% CI	1.38	1.50	Si	Confiabilidad

Conclusiones

Validación

- El sistema de identificación de hablantes Forensia cumple los criterios de validación de precisión, poder de discriminación y calibración obtenidos con muestras certificadas de otro laboratorio, siendo por lo tanto válido para su uso en el ámbito forense.

Verificación

- El sistema Forensia ha participado del Concurso NIST 2012, donde se ha verificado su calidad en la identificación de los 20 casos más difíciles en condiciones de “mismatch” de datos. (4to puesto).
- El sistema Forensia ha participado en el desafío SITW 2016 convocado por SRI, mejorando su desempeño con el agregado de modelos de duración.

Homologación

- No se han definido aun criterios universales de homologación para la identificación forense de hablantes.

Contactos

- **Black Vox**
- **www.blackvox.com.ar**

- **Programa Ciencia y Justicia**
cienciayjusticia@conicet.gov.ar

Referencias

Drygajlo, M. Jessen, S. Gfroerer, I. Wagner, J. Vermeulen & T. Niemi, “Methodological Guidelines for Best Practice in Forensic Semiautomatic and Automatic Speaker Recognition”, in European Network of Forensic Science Institutes (ENFSI), 2015.