

UNIVERSIDAD DE COSTA RICA
VICERRECTORÍA DE INVESTIGACIÓN
Unidad de Gestión de Proyectos

FORMULARIO PARA LA PRESENTACIÓN DEL INFORME FINAL

Comentario Inicial

Este proyecto ha generado muchos frutos, si bien es cierto que no se desarrolló al 100% de la idea original. En particular, se logró involucrar una gran cantidad de estudiantes (alrededor de 20), lo cual convierte a este proyecto en un ejemplo pedagógico/investigativo que podría reproducirse en otros contextos, pensamos (nuestro país no tiene en este momento las condiciones para copiar la situación de otros países con respecto a estudios de posgrado; por ejemplo, acá la mayoría de los estudiantes trabajan tiempo completo después de obtener su título de pregrado).

Adicionalmente, la ideas de este proyecto ayudaron a concebir y ejecutar otras líneas investigativas (proyecto de pronunciación de diptongos en San José con Annette Calvo de lingüística, y proyecto de acústica de canto de ranas, con Eduardo Boza y otras personas de biología; ambos proyectos se hayan en ejecución).

El objetivo general se cumplió, pues se tiene articulado un sistema de identificación de voz. Además, se tiene redactado un borrador avanzado de un artículo sobre identificación de voz y sentimiento (enojo), el cual será sometido en los próximos días a la consideración del Journal of the Acoustical Society of America.

El proyecto en un momento dado se volvió demasiado ambicioso. La parte telefónica, por ejemplo, resultó ser incompatible con el carácter más bien pedagógico del proyecto. Una continuación del proyecto debería tomar en cuenta este aspecto.

Los principales RESULTADOS del proyecto se pueden resumir de la siguiente manera:

Primero. Se comprobó que los cuatro vectores propuestos (Long-term spectra analysis, F0 analysis, Formant analysis, Time-energy distribution) funcionan de manera aceptable para realizar identificación de locutor, incluso cuando se emplean individualmente. Entre esta comprobación básica y el desarrollo de un producto listo para ser usado en una corte, hay mucha distancia claro está.

Segundo. Se implementó la técnica para el caso particular de Long-term spectra analysis (LTS). Los números finales son los siguientes:

correlación entre distintos hablantes = 0.890 = μ_1 (definición de μ_1)
desv estándar de la muestra = 0.045 = σ_1
desv estándar del promedio = 0.010

correlación de un hablante consigo mismo = 0.955 = μ_2
desv estándar de la muestra = 0.024 = σ_2
desv estándar del promedio = 0.005

Estos datos son suficientes para implementar el método de manera aplicada. El likelihood ratio (LR) o “cociente de verosimilitud” viene dado por:

$$LR(x) = (\sigma_1/\sigma_2) \exp[-(x-\mu_2)^2/2\sigma_2^2] \exp[-(x-\mu_1)^2/2\sigma_1^2] .$$

En esta fórmula, x es el valor de la correlación entre la muestra de referencia y la muestra indagada. El LR es la cantidad que el perito forense entrega al sistema jurídico. Un LR de 2, por ejemplo, significa que cualquier probabilidad previa que el sistema jurídico estuviera dispuesto a darle a la inocencia de un sospechoso, ahora se reduce por un factor de 2.

Los demás vectores no han sido articulados de esta manera todavía.

Tercero. Como se mencionó arriba, se tiene un borrador de un artículo sobre identificación de voz mediante LTS y estado de enojo del hablante.

1. Información general

Código:	805-B2-175		
Nombre del proyecto, actividad o programa:	Articulación de un sistema de identificación de locutor con fines forenses		
Programa de adscripción:	N/A		
Unidad base:	Centro de Investigaciones Geofísicas	Otras unidades colaboradoras:	Esc. Física
Período vigencia original:	1 enero 2012 hasta 31 diciembre 2013		
Fecha de presentación de informe:	29 febrero 2016		
Investigador(a) principal (Nombre, carga, período):	Manuel Ortega Rodríguez, ¼ tiempo, 1 enero 2012 - 31 diciembre 2015		
Otros(as) investigadores(as) participantes (Nombre, carga, período):	Hugo Solís Sánchez, sin carga, mismo período		
Período de ampliación:	1 enero 2014 - 31 diciembre 2015		

2. Resumen ejecutivo del proyecto, programa o actividad (mínimo 200, máximo 500 palabras)

La investigación tiene que ver con la identificación automatizada de locutor (Voice ID) con fines forenses (lo cual convierte a este proyecto en un trabajo de fuerte aplicación en temas de sociedad). El objetivo general se cumplió, pues se tiene articulado un sistema de identificación de voz. Además, se tiene redactado un borrador avanzado de un artículo sobre identificación de voz y sentimiento (enojo), el cual será sometido en los próximos días a la consideración del Journal of the Acoustical Society of America. El proyecto en un momento dado se volvió demasiado ambicioso. La parte telefónica,

Los principales resultados del proyecto se pueden resumir de la siguiente manera:

Primero. Se comprobó que los cuatro vectores propuestos (Long-term spectra analysis, F0 analysis, Formant analysis, Time-energy distribution) funcionan de manera aceptable para realizar identificación de locutor, incluso cuando se emplean individualmente. Entre esta comprobación básica y el desarrollo de un producto listo para ser usado en una corte, hay mucha distancia claro está.

Segundo. Se implementó la técnica para el caso particular de Long-term spectra analysis (LTS). Los demás vectores no fueron articulados de esta manera todavía.

Tercero. Como se mencionó arriba, se tiene un borrador de un artículo sobre identificación de voz mediante LTS y estado de enojo del hablante.

3. Desarrollo y ejecución (qué se hizo y cómo).

Objetivos específicos y metas	Actividades desarrolladas	Dificultades, formas de resolverlas y pendientes
Siguiendo la metodología descrita en la formulación del proyecto, crear un prototipo que logre efectuar una identificación de voz en condiciones normales. Meta: Que el prototipo sea totalmente funcional.	1 enero a 30 junio 2012, estudio de transformadas de Fourier de señales acústicas en general	No hubo mayores dificultades. No quedaron aspectos pendientes.
	1 julio a 31 diciembre 2012, establecimiento de una métrica para poder así determinar la "distancia" entre dos señales acústicas	1 enero a 30 junio 2013, desarrollo de análisis estadístico empleando bases de datos

	1 julio a 31 diciembre 2013, validación del método con pruebas finales.	
Continuando con la metodología propuesta, y siguiendo la lógica del proyecto, crear un sistema que logre efectuar una identificación de voz en situaciones reales (por ejemplo, incluyendo canales telefónicos).	1 enero a 30 junio 2014, repetición de los procedimientos de los primeros dos años pero aumentando la estadística (el número de pruebas) en un factor de 10	El proyecto en un momento dado se volvió demasiado ambicioso. La parte telefónica, por ejemplo, resultó ser incompatible con el carácter más bien pedagógico del proyecto. Una continuación del proyecto debería tomar en cuenta este aspecto. El desarrollo del proyecto se varió de la manera indicada al comienzo de este informe.
META	1 julio a 31 diciembre 2014, implementación del método de cociente de verosimilitud para los datos	
Que el algoritmo sea totalmente funcional en situaciones realistas (es decir, en grabaciones que uno se encuentra en el mundo real, que son de menor calidad que las de laboratorio).	1 enero a 30 junio 2015, implementación del método de distractores para los datos	
	1 julio a 31 diciembre 2015, articulación de todos los métodos en un producto final	

*Agregar las celdas que sean necesarias

COMENTARIO CON RESPECTO AL OBJETIVO ORIGINAL

El objetivo original del proyecto es (citado textualmente de la propuesta original):

OBJETIVO GENERAL: Articulación de un sistema de identificación de locutor con fines forenses, que sienta fuertemente las bases para una implementación funcional del mismo a mediano plazo.

(con correspondiente **OBJETIVO ESPECÍFICO:** Siguiendo la metodología descrita, crear un prototipo que logre efectuar una identificación de voz en condiciones normales.)

Tomando en cuenta lo explicado en la primera sección (“Comentario Inicial”), y en vista de este objetivo, se puede decir que sí se cumplió el objetivo original, si bien es cierto no utilizando todos los vectores sino solamente el primero de ellos (Long-Term Spectra Analysis). Me explico:

La fórmula

$$LR(x) = (\sigma_1/\sigma_2) \exp[-(x-\mu_2)^2/2\sigma_2^2] \exp[-(x-\mu_1)^2/2\sigma_1^2]$$

da el “likelihood ratio” LR, que es la cantidad que el físico forense entrega al sistema jurídico. En esta fórmula, x es el valor de la correlación entre las transformadas de Fourier (promediadas en la duración total de la muestra) de la muestra de referencia y la de la muestra indagada.

Lo único que uno podría pulir y seguir trabajando es la parte de toma de datos de la población general (μ_1 , σ_1), así como poner a prueba el método en condiciones rigurosas (doblemente ciegas, con unos 30 casos aleatorios de la población general). Digamos que a criterio de este servidor, se hizo un 80% del objetivo original.

4. Resultados y alcances (utilice la cantidad de palabras necesarias).

Los resultados se encuentran expuestos en detalle en la primera sección (“Comentario Inicial”) de este informe.

Impacto, difusión y divulgación realizada al interior y exterior de la Universidad:

Los resultados han sido presentados en mini congresos internos a la Universidad.

5. Producción académica (adjuntar anexos a criterio de Consejo Científico o Comisión de Investigación).

Referencias de los manuscritos artículos en prensa o publicados:

N/A

Títulos de las ponencias y participación en eventos, lugar y fechas:

Otros productos:

6. Trabajos de graduación y participación estudiantil.

Asistentes	Labores ejecutadas
Diana Valverde, Bryan Matamoros, Ariadna Venegas, Jonatán Alvarado, Camilo Agüero, Adrián González, Gerardo Salazar, Andrés Chavarría, Armando Gómez, Ricardo Gamboa, Gustavo Ramírez, Lizdenia Arce, Oscar Calvo	Los estudiantes formaron parte activa del proceso de investigación. Dos de ellos figuran como co-autoras del artículo (en proceso) sobre enojo.

*Agregar las celdas que sean necesarias

Título de trabajos finales de graduación involucradas en el proyecto	Nivel (Licenciatura, Maestría, Doctorado, Diplomado, Especialidad)	Nombre de los/las estudiantes
--	--	-------------------------------

N/A

*Agregar las celdas que sean necesarias

7. Observaciones finales.

8. Informe financiero (adjuntar informe respectivo aportado por jefatura administrativa).

9. Aspectos éticos (adjuntar la “Aplicación para revisión continua o para cerrar el estudio” del CEC). N/A

10. Autorización para incorporar el informe final ejecutivo en los repositorios de la UCR.

(x) SI () NO Autorizo a la Vicerrectoría de Investigación para incluir en los repositorios institucionales el anterior informe.

Fecha

Firma

Manuel Ortega Rodríguez

Investigador(a) principal

Versión 26-08-2013.