

Sistema de visualización del habla (VISHA)

*Grupo de Ayuda a Discapacitados
Departamento de Ingeniería Electrónica
E.T.S.I. de Telecomunicación
Ciudad Universitaria, Madrid*

Introducción

En el mercado existen numerosos sistemas basados en proceso de voz, con aplicación en disciplinas muy diversas. Algunos de estos sistemas pueden ser:

- Sistemas para evaluar las pérdidas auditivas, basados en la realización de audiometrías.
- Sistemas para el estudio de la señal de voz, basados en la grabación de la misma, para posteriormente visualizar los parámetros más representativos de la misma.
- Sistemas para la rehabilitación del lenguaje, basados en la representación gráfica, y en tiempo real, de los parámetros de la señal de voz.

Además se pueden encontrar sistemas para la manipulación frecuencial de la señal de voz que hacen un aprovechamiento óptimo de los restos auditivos en personas hipoacúsicas; sistemas de suministro de información oral, como son los sintetizadores de voz; sistemas de reconocimiento de habla, etc.

Estos sistemas, que suelen ser elementos autónomos o bien periféricos conectables a un ordenador de propósito general, se caracterizan por necesitar de hardware propio para realizar cada una de las aplicaciones, así como por ser elementos poco versátiles, donde la mayoría de ellos no admiten ni modificaciones ni ampliaciones soportadas por el mismo hardware.

En el Departamento de Ingeniería Electrónica de la Universidad Politécnica de Madrid se pensó en desarrollar un potente sistema configurable por software, que sobre un único hardware, de bajo coste, permitiese realizar todas las aplicaciones descritas anteriormente, en muchos casos con prestaciones similares a las de los sistemas convencionales, disponibles actualmente en el mercado. Al día de hoy este objetivo se puede considerar como alcanzado. Se dispone de una serie de aplicaciones ya en uso por parte de especialistas de forma satisfactoria.

A continuación se describen cuáles son de entre todas las aplicaciones desarrolladas las que pueden resultar más interesantes desde el punto de vista que nos ocupa.

Realización de Audiometrías (PcAud)

Las pruebas audiológicas, destinadas a proporcionar la respuesta en frecuencia del oído pueden ser de dos tipos: subjetivas y objetivas.

Entre las pruebas subjetivas (que son las más utilizadas aunque requieren la colaboración del paciente) se encuentran las audiometrías tonales. Se realizan excitando el oído con tonos puros, que van variando su energía, y comprobando dónde se encuentra el umbral de audición en cada uno de ellos.

Nuestro sistema permite realizar audiometrías tonales guardando la información en un fichero, con el que se puede seguir la evolución de un paciente mediante la comparación de audiometrías realizadas en distin-

tas fechas. La información de estas audiometrías es suministrada de forma gráfica tanto por el monitor del ordenador como por la impresora.

Con objeto de proporcionar un audiómetro lo más completo posible, se ha incorporado también al sistema la posibilidad de realizar audiometrías verbales. Se pueden llevar a cabo tanto curvas de inteligibilidad como tests de rasgos distintivos y mediante estas pruebas determinar parámetros importantes para los audiólogos como son el umbral de recepción verbal, la discriminación máxima del paciente, los fonemas que pueden causarle problemas, etc.

Todas estas pruebas audiométricas pueden realizarse tanto por conducción aérea como por conducción ósea. Esto ofrece nuevas posibilidades como, por ejemplo, el cálculo del Rinne audiométrico y la determinación de la situación de la anomalía del oído en estudio (si se halla en el oído externo o es un problema más grave de otra parte del aparato auditivo).

Además en el desarrollo del sistema se ha puesto un cuidado muy especial en aquellas pruebas destinadas a los niños y que denominamos Audiometrías lúdicas. Son pruebas tonales pero con una presentación y mecanismos de ejecución adecuados a las características del paciente destinatario, y que tienen como objetivo captar su atención y conseguir su colaboración durante el tiempo necesario como para obtener resultados completos y adecuados.

El programa presenta un interfaz gráfico muy sencillo de manejar incluso para usuarios inexpertos. Los distintos pasos se dan siguiendo un sistema de ventanas que van marcando el proceso que se emplea para llegar a cualquiera de las funcionalidades del sistema.

Realización para el estudio de la señal de voz (PcVox)

El estudio de la señal de voz es necesario en diversas actividades, puramente ingenieriles como codificación, transmisión, síntesis o reconocimiento, o de tipo psicológico, lingüístico y pedagógico.

Básicamente, este sistema permite generar un fichero con la señal de voz que un locutor emite delante de un micrófono y sobre este fichero realizar una gran variedad de procesos, que van desde calcular y visualizar los parámetros más representativos de la voz, manipular dichos ficheros mediante filtrado, hasta la audición de toda o parte de una frase grabada.

Es fácilmente manejable y su uso requiere un entrenamiento mínimo. Para la selección de funciones se ha elegido un entorno de menús, manejados desde el teclado o ratón, mientras que para la visualización de los distintos parámetros utilizamos un sistema de ventanas, fácilmente creables y almacenables por el usuario. De cualquiera de las imágenes que aparecen en pantalla puede obtenerse copia impresa.

Los ficheros que utiliza nuestro sistema tienen formato ILS, estándar ampliamente utilizado, que permite el intercambio de información con otros sistemas que utilicen el mismo formato. Pueden generarse por micrófono, magnetófono o por encadenamiento de segmentos de voz procedentes de otros ficheros. El usuario selecciona la frecuencia de muestreo y la de corte de los filtros.

Los parámetros que el sistema representa, pudiendo seleccionar cualquier escala tanto en el eje de tiempos como en el del valor del parámetro representado, son:

- Forma de onda.
- Energía.
- Entonación.
- Cruces por cero.
- Espectogramas, con diferentes anchos de los filtros paso banda y escalas de frecuencia y tiempo.

El sistema dispone de salida de audio con capacidad para excitar un altavoz, auriculares o vibrador óseo. Se puede escuchar todo o parte de un fichero, por ejemplo, sólo algunos fonemas.

Sobre un fichero de voz es posible realizar distintas funciones de procesado y edición de la misma, como es la inserción de segmentos de voz o la edición desde el punto de vista fonético.

Sistema para rehabilitación del lenguaje (Isoton)

Permite entrenar tres parámetros del habla: intensidad, sonoridad y tono fundamental. Tiene dos modos de funcionamiento: imitación de patrones y control de videojuegos por voz.

En el primer modelo, se representa sobre la misma pantalla la pronunciación del profesor y del alumno. El profesor hace ver al alumno los errores cometidos por comparación de las dos gráficas.

- En la representación de intensidad aparece en la pantalla la energía de la señal recibida por el micrófono, pudiendo entrenar, entre otros aspectos, el acento.
- En la representación del tono fundamental se visualiza la frecuencia de vibración de las cuerdas vocales; por ejemplo, permite entrenar la discriminación entre frases afirmativas e interrogativas, que tienen distinto punto de elevación del tono.
- En la representación de sonoridad gráfica, se dejan sin colorear ausencias de sonido típicas de las oclusivas p,t,k. Distingue por colores los sonidos sonoros y los fricativos sordos.

El control de videojuegos por voz es muy útil para estimular a los más pequeños y para conseguir que los mayores continúen trabajando en sus ratos de ocio. Existen varios juegos controlados por los tres parámetros que el sistema entrena cuya filosofía es la de desplazar un móvil, controlado por la voz, por una trayectoria previamente establecida por el profesor.

Isoton también está siendo empleado para entrenamiento de aspectos suprasegmentales tanto a nivel pre-lingüístico como a nivel lingüístico, así como en la adquisición y corrección de fonemas.

Conclusiones

Con la aparición del ordenador personal se ha pasado de utilizar equipos destinados a una aplicación concreta, a programas que convierten nuestro ordenador en una solución eficaz a muchos problemas. Nos encontramos con ordenadores enormemente flexibles que incorporan multitud de periféricos con una única utilidad. Parecería mucho más lógico hacer también periféricos configurables por programa, pequeños ordenadores para incorporar a nuestro Pc.

La tarjeta de voz desarrollada en nuestro Departamento es un intento de flexibilizar periféricos que tradicionalmente han sido rígidos. ¿Por qué comprarse un periférico para síntesis de voz, otro para reconocimiento, otro para realizar audiometrías, etc.? Mejor será equipar nuestro ordenador con uno solo, que permita estas aplicaciones y otras muchas más.

Siendo lógico que nuestro objetivo sea el de sacar el máximo partido a esta potente máquina hardware, los trabajos futuros se deben encaminar en dos direcciones.

Por un lado optimizar y ampliar las aplicaciones ya desarrolladas (las mencionadas en el texto y otras en uso pero que no han sido tratadas en este breve espacio). Así ha ocurrido con el audiómetro presentado, que es una evolución totalmente compatible con la versión primera realizada, pero con unas mayores prestaciones, o con la versión bajo Windows desarrollada tanto para el programa PcVox como para Isoton.

Y por otro lado, el conectar con los usuarios lo suficiente como para que sean ellos los que nos planteen las necesidades que todavía no cubre nuestra tarjeta, y que serían puntos de partida ideales para nuevas aplicaciones.

Agradecimientos

Desde estas líneas queremos expresar nuestro agradecimiento a nuestra propia universidad y sobre todo al Instituto Nacional de Servicios Sociales (INSERSO) del Ministerio de Asuntos Sociales, ya que los proyectos desarrollados pertenecen al Convenio entre el INSERSO y la Universidad Politécnica de Madrid.

Bibliografía

Gómez Serra, Sara. "Sistema de exploración audiométrica completa sobre Pc". Proyecto Fin de Carrera, E.T.S.I. de Telecomunicación de Madrid, 1995.

Berrojo Martín, Miguel Angel. "Sistema de visualización de parámetros de voz en un entorno PC: PcVox". Proyecto Fin de Carrera, E.T.S.I. de Telecomunicación de Madrid, 1990.

Poncela González, Javier. "El sistema de visualización de parámetros de voz WPcVox: implementación de nuevas prestaciones". Proyecto Fin de Carrera, E.T.S.I. de Telecomunicación de Madrid, 1994.

Martínez Olalla, Rafael. "Desarrollo de un entorno de animación interactiva para la ayuda a deficientes auditivos". Proyecto Fin de Carrera, E.T.S.I. de Telecomunicación de Madrid, 1995.