



ALEX NABAUM

El lenguaje de los teléfonos

Los móviles podrán reconocer el habla gracias a “chips” más rápidos y a un mejor software.

MARA E. VATZ

Los teléfonos móviles y las PDA inalámbricas tienen un problema típico: o bien carecen de teclado o el que tienen es muy pequeño. Esto hace que teclear cualquier cosa que no sea un número de teléfono sea una tarea tediosa. La solución está en camino: dispositivos móviles capaces de reconocer el habla

Algunos teléfonos móviles ya utilizan el reconocimiento de voz como alternativa a mensajes tecleados en tareas simples como marcar un número, pero pronto podremos dictar un mensaje de texto a nuestro teléfono, preguntar a nuestro coche

una dirección, o indicar a nuestro MP3 que queremos escuchar a Los Beatles.

La verdad es que incluso hoy los teléfonos móviles de gama alta son capaces de utilizar un sofisticado software de reconocimiento de voz que podría conducir a olvidarnos de los teclados. “El problema fundamental a la hora de introducir información en dispositivos móviles es el interface, y ese problema puede ser solucionado por la voz”, afirma Rich Geruson, director ejecutivo de VoiceSignal, una compañía de tecnología de voz.

Mientras compañías como IBM y Dragon Systems, han estado vendiendo software de reconocimiento de voz para ordenadores durante más de una década, los teléfonos móviles con capacidades limitadas de reconocimiento de voz, sólo comenzaron a aparecer hace algunos años. Y hasta ahora tales dispositivos han sido fundamentalmente equipos dependientes del hablante, es decir, que sólo funcionan bien para sus usuarios principales, y tienen que ser entrenados para reconocer palabras concretas.

La disponibilidad de procesadores más rápidos y de un software más eficiente permite ahora sistemas independientes del hablante que pueden reconocer la voz de cualquier usuario y que no requieren entrenamiento. Estos sistemas pueden discernir entre miles, y no sólo docenas, de sustantivos y están diseñados para trabajar aún cuando el hablante se encuentre en un entorno ruidoso, como pudiera ser un coche en movimiento.

Para los ingenieros de VoiceSignal, la clave en este avance fue el cambio de enfoque de la precisión a la eficiencia. Los algoritmos de reconocimiento de voz

altamente precisos diseñados para los ordenadores son excesivamente complejos para los terminales móviles. Los algoritmos tradicionales para estos dispositivos requerían una menor capacidad de proceso y estaban limitados a un pequeño vocabulario porque trabajaban mediante la comparación de la onda sonora correspondiente a una palabra concreta, con la onda sonora almacenada en la memoria del dispositivo.

En lugar de almacenar todas las ondas sonoras de cada palabra en su léxico, el sistema de VoiceSignal almacena información de fonemas, que son las unidades más pequeñas del habla. Cada fonema puede ser descrito de acuerdo a un conjunto de parámetros acústicos. El software mide la pronunciación que hace un usuario de estos parámetros acústicos y busca a continuación las palabras que casen. Estos valores paramétricos utilizan menos memoria que los archivos de audio, de manera que el software puede manejar un mayor vocabulario.

Esto está permitiendo aplicaciones que van más allá de la marcación por voz. Por ejemplo VoiceSignal ofrece un software que permite a los usuarios de un teléfono móvil seleccionar cualquier elemento del menú con una simple pronunciación. “Si intentamos enviar un mensaje a través de nuestro teléfono móvil, tenemos que realizar una media de diez pulsaciones sólo para llegar al espacio para el texto”, afirma Geruson. “Con nuestra tecnología sólo hay que decir enviar mensaje al móvil de John Smith y el cursor se planta en el espacio del mensaje listo para comenzar”. El primer

teléfono con estas capacidades salió al mercado el pasado verano. Próximamente esta compañía presentará un software para teléfonos que permitirá a los usuarios dictar textos o correos, lo que será particularmente útil en Asia. “Si pensamos que es difícil teclear en teclados occidentales, imaginemos cual es el problema en Japón o China”, afirma Geruson.

Mientras tanto, investigadores de la compañía ScandSoft están comenzando a utilizar el reconocimiento de voz en coches. Un conjunto compuesto por

cerca. En los laboratorios de investigación de Mitsubishi Electric, se está desarrollando un software que utiliza la voz para simplificar las tareas de búsqueda de información. En lugar de buscar a lo largo de las miles de entradas de canciones en un MP3, por ejemplo, un usuario puede seleccionar una canción determinada tan sólo mencionándola por su título, o el del grupo, o el del álbum. “Pensamos que una de las cosas para las cuales la voz era importante era para seleccionar algo”, afirma el investigador de Mitsubishi Meter Wolf.

Los algoritmos de reconocimiento de voz dejan paso a un software que puede comprender a cualquier usuario, sin necesidad de entrenamiento.

micrófono, altavoz y la máquina de voz de ScandSoft proporciona a los conductores un interface de manos libres para sus móviles. Un teléfono equipado con un transmisor Bluetooth puede ser instalado en cualquier parte del coche y los conductores pueden utilizar órdenes de voz para marcar, aceptar o rechazar llamadas, ajustar el volumen o controlar las opciones del menú, y todo ello sin apartar las manos del volante o los ojos de la calzada.

Aunque la industria de los equipos móviles ha sido la primera en utilizar el reconocimiento de voz, los fabricantes de electrónica de consumo les siguen de

Sin embargo, y a pesar de estos avances, aún no está claro si podremos utilizar el reconocimiento de voz para cualquier cosa. Los usuarios de teléfonos móviles pueden sentirse incómodos dictando correos en público. Igualmente sabemos que siempre querrán teclados para introducir información sensible como, por ejemplo, el número de su tarjeta de crédito. Rich Geruson predice que esta tecnología transformará algún día la forma en que utilizamos los teléfonos móviles. Tan pronto como una minoría comience a utilizar esta tecnología, afirma, “su uso se extenderá y pronto estará en todas partes”.

LA VOZ ES LA ORDEN

El software de reconocimiento de voz de los teléfonos móviles más recientes funciona comparando la información de ondas sonoras, como timbre y duración, con parámetros almacenados de todas las palabras de su vocabulario. Si el software encuentra una palabra que se corresponda, produce una acción predefinida como por ejemplo marcar un número telefónico.

