

ADQUISICION DE DATOS PARA EL PROYECTO  
RECONOCIMIENTO DE FONEMAS POR COMPUTADOR

ING. TANIA PEREZ  
ESCUELA POLITECNICA NACIONAL

RESUMEN

En este artículo se describen brevemente los procedimientos de adquisición de datos para el proyecto RECONOCIMIENTO DE FONEMAS POR COMPUTADOR tanto en los aspectos de hardware como de software.

ABSTRACT

In this paper the hardware and software procedures for the data acquisition for the project PHONEME RECOGNITION BY COMPUTER are briefly described.

El ingreso de datos al computador es el primer paso para el análisis y despliegue gráfico de los fonemas que se lleva a cabo posteriormente mediante el uso de diferentes técnicas y algoritmos.

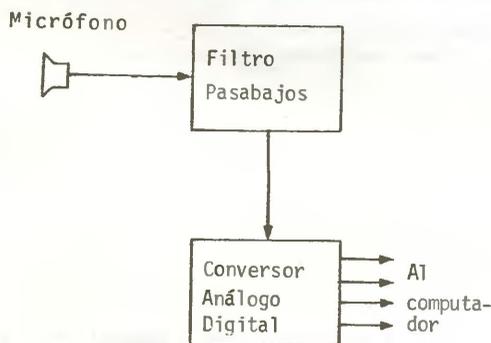


Fig. 1.

En términos generales, se puede considerar que los elementos básicos para la adquisición de datos del proyecto en mención son: un micrófono que capta la voz y entrega una señal análoga, un filtro pasabajos que limita el ancho de banda de la señal a procesarse y el convertor análogo - digital que posibilita el ingreso de datos al computador. Ver fig. 1

El micrófono utilizado debe ser de alta calidad para evitar, en lo posible la distorsión de la onda y garantizar una respuesta de frecuencia suficiente.

FILTRO PROGRAMABLE \*

El diagrama de bloques del filtro utilizado se muestra en la fig. 2.

El diseño ha sido realizado en base del circuito integrado MF6-50 que es un filtro activo pasabajos de alta precisión tipo Butterworth de sexto orden (atenuación 36 dB por octava). Han sido utilizados 2 circuitos MF6-50 en cascada logran-

ING. GUALBERTO HIDALGO  
ESCUELA POLITECNICA NACIONAL

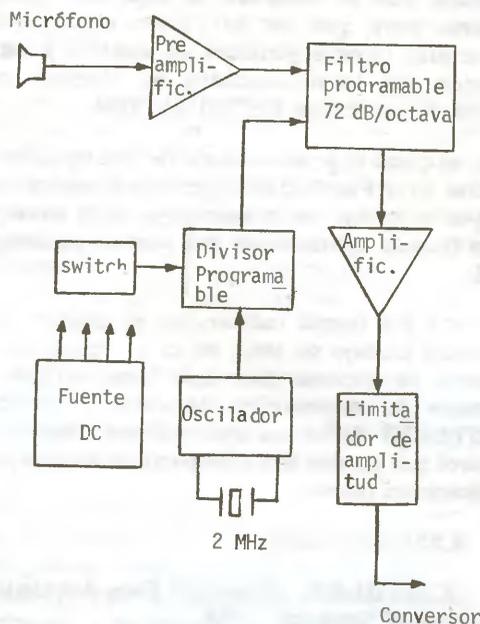


Fig. 2.

dose una atenuación de 72 dB por octava. La frecuencia de corte para este filtro tiene una relación de 1:50 con la frecuencia de reloj entregada al circuito. Para el presente caso la frecuencia base de reloj se obtiene desde un oscilador de 2 MHz diseñado con cristal. A continuación, mediante un divisor de frecuencia programable se pueden obtener varias frecuencias de reloj que a su vez permiten obtener las siguientes frecuencias de muestreo: 8 KHz, 12 KHz, 16 KHz, 24 KHz, 32KHz y 48 KHz. A su vez las frecuencias de corte son 1/3 de las de muestreo.

De todas estas frecuencias la que se ha utilizado hasta el momento es la de 12 KHz para el muestreo, lo que permite obtener una frecuencia de corte de 4KHz. La máxima amplitud de señal que entrega el filtro está condicionada por los rangos de amplitud que acepta la tarjeta de conversión análogo-digital utilizada, que en este caso es de +/-10 voltios. Para mayor seguridad se utiliza un limitador de amplitud bipolar con diodos zener de 9.6 voltios.

TARJETA DE CONVERSION ANALOGO DIGITAL (METRABYTE uCDAS-16G)

Cuando se necesitaba adquirir una tarjeta de conversión análogo-digital de esta naturaleza, no fue fácil encontrar una adecuada para el computador PS-2/80 del que dispone el proyecto. Pues el mercado ofrecía una gran variedad de tarjetas de conversión análogo-digital diseñadas para

computadores del tipo PC, XT y AT pero no para el bus con estructura de microcanal que utilizan los sistemas PS-2 (modelos 50 hasta 80).

La tarjeta MetraByte modelo uCDAS-16G es una tarjeta de expansión entrada / salida de alta velocidad y de ganancia programable para conversión análogo-digital. Una visión general de su estructura se muestra en la fig. 3

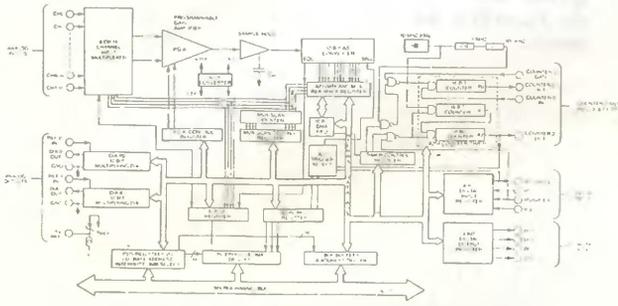


Fig. 3.

IBM PS/2 and Microchannel Bus are registered trademarks of International Business Machines Corporation.

El convertor utilizado es el HI-674A de 12 bits que da una máxima frecuencia de conversión de 70 KHz para ganancia unitaria.

La conversión análogo-digital se inicia por un comando de software, y cuando la conversión ha finalizado los datos se transfieren por hardware a través del DMA (direct memory access). Todos estos modos de operación son seleccionados por el registro de control de la tarjeta y son compatibles con el software utilizado.

La configuración de los canales A/D se selecciona durante el proceso de instalación. Para este caso se seleccionaron canales con entrada diferencial, pues con esta configuración se consigue un nivel de rechazo al ruido superior que en los canales de un solo terminal, se seleccionó además escala bipolar por su mayor versatilidad para el manejo de señales.

#### SOFTWARE UTILIZADO

Para el muestreo de la señal se utiliza uno de los programas utilitarios que forma parte del software adquirido con la tarjeta de conversión. Este programa es el PROG63 que muestrea la señal análoga que ingresa desde el micrófono a una frecuencia de 12 KHz y transfiere los datos a la memoria en forma directa a través del DMA, luego configura un arreglo, en el que ubica los datos para que estos puedan ser utilizados. Se realizaron algunas modificaciones a este programa para que, en lugar de conformar el arreglo en memoria, los datos sean transferidos directamente al disco de almacenamiento.

Este programa da la posibilidad de tomar 32000 muestras en la grabación de cada archivo. Como cada muestra ocupa 2 bytes, el almacenamiento en el disco es de 64 Kbytes por archivo. La limitación del nú-

mero de muestras a 32000 se debe al hecho de que las subrutinas de trabajo de la tarjeta están escritas en Basica, lo cual limita la memoria administrada a 64 Kbytes, razón por la que el tiempo máximo de muestreo es de 2 segundos 2/3 para cada archivo.

Dentro de este mismo programa se hace una llamada a la subrutina GRAFSEM1 programada en QB que extrae del disco de almacenamiento las 32000 muestras correspondientes a un archivo y luego despliega en pantalla la onda grabada, en bloques sucesivos de 1000 muestras cada uno.

#### GRAFIZACION EN PAPEL DE LAS ONDAS DE VOZ

La grafización de las ondas de voz tanto en el dominio del tiempo como en el de la frecuencia es premisa fundamental para el desarrollo del software de reconocimiento de fonemas.

El análisis minucioso de estos gráficos, sobre todo en el dominio del tiempo han servido de base para la implementación de los algoritmos de reconocimiento de vocales, de búsqueda de periodicidad y de segmentación, entre otros.

El dispositivo utilizado para este propósito es el Plotter Hewlett Packard 7475A versión 001 que es un grafizador vectorial que produce gráficos de excelente calidad a alta velocidad en papel o transparencias de 4 dimensiones diferentes, que se eligen con ayuda de dos switches ubicados en el panel posterior del plotter. El 7475 grafiza con una aceleración aproximada de 2g y una máxima velocidad de 38.1 centímetros/segundo.

La versión 001 de este plotter está equipada con el interfase RS-232-C/CCITT a través del cual se conecta directamente al sistema PS-2/80.

Las líneas de comunicación del interfase RS-232-C se conectan al plotter a través de un conector standard de 25 pines ubicado en su zona posterior y puede operar en la configuración de tres terminales (transmisión, recepción, referencia). Se tiene además posibilidad de chequeo de paridad, pero el sistema disponible no lo utiliza.

#### VELOCIDAD

El plotter está diseñado para operar en modo asincrónico con velocidades 75, 110, 150, 200, 300, 600, 1200, 2400, 4800 y 9600 baudios. Se pueden elegir valores intermedios utilizando un reloj externo conectado al pin 17 del conector. La velocidad de envío de datos del sistema PS-2/80 es de 9600 baudios, valor que se configura con ayuda de los switches B1, B2, B3 y B4 ubicados detrás del plotter.

#### HANDSHAKING

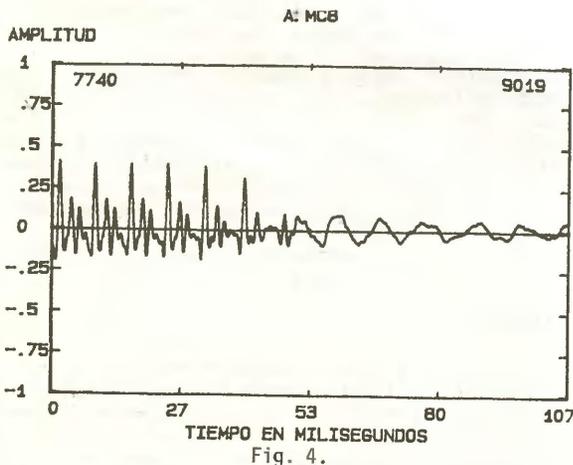
El 7475 utiliza un almacenamiento intermedio de entrada (buffer) de 1024 bytes para sincronizar el procesamiento de datos con la velocidad a la cual se está recibiendo. La presencia de este almacenamiento asegura que el computador y el plotter transfieran información uno a

otro de tal forma que los datos no se pierdan ni sean interpretados erróneamente. Este es el propósito del handshaking. El plotter es capaz de manejar 4 métodos de handshaking para prevenir la sobrecarga del buffer y la resultante pérdida de datos. Sin embargo son las características del computador y sus requerimientos los que determinan el método de handshake a utilizarse.

Para el caso del sistema PS-2/80, éste utiliza el método de handshake de línea física (hardware) que consiste en la utilización de una conexión física en el pin 20 del cable RS-232-C para controlar el handshaking. El plotter controla la secuencia de intercambio de datos generando un voltaje en el pin 20 del conector (CD line) para indicar al computador cuando enviar el siguiente bloque de datos. Si hay suficiente espacio en el almacenamiento intermedio del plotter para aceptar y almacenar otro bloque de datos, el plotter pone la línea CD de "data terminal ready" (pin 20) en estado alto. Si el espacio no es suficiente pone la línea CD en estado bajo. Por el monitoreo de esta línea el computador sabe cuando puede o no transferir el siguiente bloque de datos.

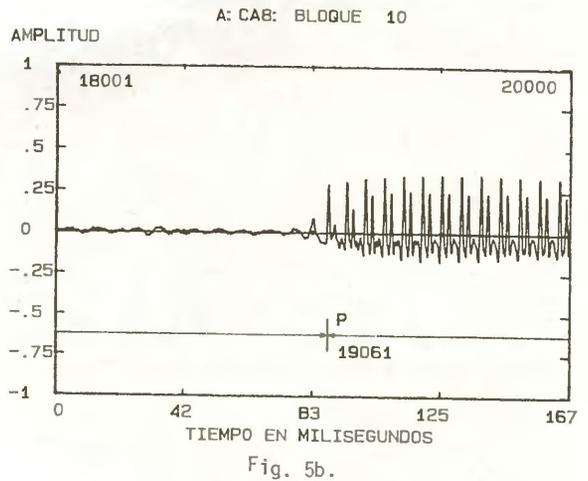
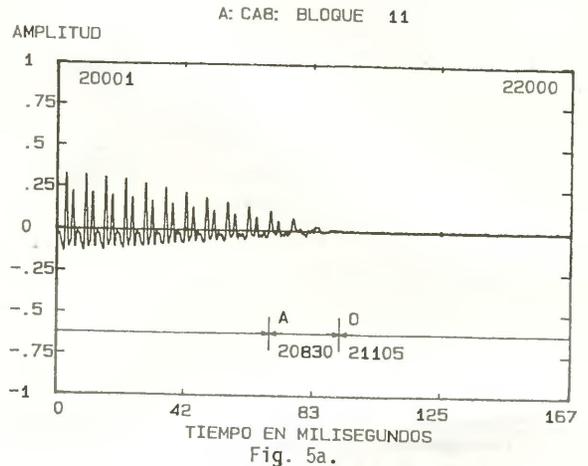
#### SOFTWARE UTILIZADO PARA LA GRAFICACION

Todos los programas desarrollados para obtener los diferentes gráficos en el plotter, han sido escritos en Basica. El programa más frecuentemente utilizado ha sido el GRAPH1 que permite extraer del disco de almacenamiento diversos números de muestras que son guardados en memoria en arreglos de 6400, 12800, 15360, 16640, 19200 o 25600 muestras, según se haya elegido. Posteriormente se grafiza en papel en bloques sucesivos de 1280 muestras; número que se considera idóneo ya que permite una buena definición para el análisis de la onda, tal como se puede apreciar en el ejemplo de la figura 4. Es importante anotar que el programa posibilita continuidad en la grafización. Con solo dar el dato del número de muestras que ya ha sido extraído, la sucesión de gráficos continúa automáticamente. En la parte superior de gráfico se anota también el número de la primera y última muestras correspondientes a ese bloque.



Con el desarrollo del algoritmo de seg-

mentación automática, los últimos archivos tomados se han grafizado incluyendo la nueva información correspondiente a la segmentación de la voz analizada, que aparece en la zona inferior del gráfico con los indicativos de los diferentes segmentos encontrados dentro de cada bloque de 2000 muestras. Para este propósito utiliza los literales "A" para segmentos aperiódicos, "P" para segmentos periódicos y "O" para segmentos oclusivos. Como se podrá apreciar del ejemplo de las figuras 5a y 5b también se indicará el número de muestra en el que comienza cada segmento.



\*El diseño y construcción de este filtro fueron realizados por el Ing. Erwin Barriga A.

#### Referencias

- [1] uCDAS-16G MANUAL, (C) 1989 MetraByte Corporation.
- [2] INTERFACING AND PROGRAMMING MANUAL, HP 7475A, Graphics Plotter, 1983, 1984, Hewlett-Packard Company.

PEREZ, TANIA

Ingeniero en Electrónica y Telecomunicaciones graduada en el Instituto Bonch Bruyevich, Leningrado, 1977. Actualmente se desempeña como profesora principal a tiempo completo en la Escuela Politécnica Nacional y realiza estudios de postgrado en Computación e Informática en la misma Institución. Participa en el PROYECTO CONUEP 88-01, RECONOCIMIENTO DE FONEMAS POR COMPUTADOR.

HIDALGO, GUALBERTO

Ingeniero en Electrónica y Telecomunicaciones graduado en la Escuela Politécnica Nacional en el año de 1974. Obtuvo el título de Master en Ingeniería de Comunicaciones en el Imperial College de Londres, Inglaterra, en 1979. Actualmente se desempeña como profesor a tiempo completo en la Escuela Politécnica Nacional y es estudiante externo de la Universidad de Londres. Dirige el PROYECTO CONUEP 88-01, RECONOCIMIENTO DE FONEMAS POR COMPUTADOR.