

021.3845363
7629



ESCUELA POLITECNICA NACIONAL

FACULTAD DE INGENIERIA ELECTRICA

TESIS DE GRADO

"TECNICAS DIGITALES PARA LA DETECCION
DEL TONO DE LA VOZ"

Tesis Previa a la Obtención del Título
de Ingeniero en Electrónica
y Telecomunicaciones

CARLOS ENRIQUE TOBON MEJIA

JULIO, 1982

002010

Certifico que el presente trabajo
ha sido elaborado en su totalidad
por el Señor Carlos E. Tobón M.

A handwritten signature in cursive script, reading "Gualberto Hidalgo", written over a horizontal line.

Ing. Gualberto Hidalgo
Director de Tesis

DEDICO ESTA TESIS

A la ilustre memoria de mi padre,
OSCAR TOBON B., como testimonio
de mi eterna admiración.

A la consagración de mi madre y a
su infatigable ejemplo profesional.

A mis hermanas y a toda mi familia
motivo de mi permanente deseo de
superación moral e intelectual.

A mi novia, Sylvia, amante compañe-
ra y amiga, con indecible amor

I N D I C E

PAG.

INTRODUCCION

CAPITULO I : LA VOZ HUMANA.

| | |
|--|----|
| 1.1. Proceso y mecanismo de producción de la voz | 1 |
| 1.2. Fonética Inglesa | 7 |
| 1.3. Fonética Española | 27 |
| 1.4. Parámetros de la voz | 33 |

CAPITULO II : METODOS PARA LA DETECCION DEL TONO DE LA VOZ

| | |
|---|-----|
| 2.1. Métodos de detección | 37 |
| 2.2. Técnicas de procesamiento en paralelo para detección | 42 |
| 2.3. Detección del tono por reducción de información ... | 66 |
| 2.4. Detección de tono utilizando la función promedio de la magnitud de la diferencia | 85 |
| 2.5. Métodos autocorrelativos de detección de tono | 108 |
| 2.6. Evaluación de los métodos | 128 |

CAPITULO III : PROGRAMAS DIGITALES PARA LA DETECCION DEL TONO

| | |
|---|-----|
| 3.1. Introducción | 132 |
| 3.2. Estructura de la biblioteca de programas | 133 |
| 3.3. Programa para ingreso de datos | 139 |

| | |
|---|-----|
| 3.4. Programa para la detección por procesamiento en paralelo | 164 |
| 3.5. Programa para la detección por reducción de información | 181 |

CAPITULO IV : RESULTADOS Y CONCLUSIONES

| | |
|--|-----|
| 4.1. Resultados | 196 |
| 4.2. Conclusiones | 248 |
| 4.3. Metodología de uso de los programas | 254 |

APENDICE "A" : LISTADOS DE LOS PROGRAMAS

BIBLIOGRAFIA

I N T R O D U C C I O N

INTRODUCCION

Este trabajo de tesis tiene como objetivo, presentar un estudio teórico de varios métodos para la detección del período de tono de señales de voz, en los dominios del tiempo y de la frecuencia; y el desarrollo de la metodología y programas de dos de estos métodos para futuros análisis de las características de la voz española, en el computador digital.

Para un mejor entendimiento de los algoritmos que se analizarán, es necesario tener un ligero conocimiento de los diferentes rasgos y parámetros de la voz; además de ciertos fundamentos matemáticos y un conocimiento de las técnicas de programación, especialmente en el lenguaje BASIC, que es el utilizado por el computador TEKTRONIX. modelo 4051 del Laboratorio de Control de nuestra facultad.

Los dos métodos de detección de tono, para los que se han desarrollado los programas, para su utilización experimental en diferentes campos de investigación de la voz, son: el de procesamiento en paralelo y el de reducción de información, ambos en el dominio del tiempo. Se han escogido estos dos métodos por ser los que presentan una mayor velocidad de ejecución, por necesitar una menor capacidad de memoria para su desarrollo y por ser en el momento uno de los métodos más exactos para la detección del tono. Sin embargo,

se presentan elementos de otros métodos para detección de tono, para poder tener una mejor apreciación de las diversas técnicas existentes.

En el Primer Capítulo se hace una introducción en el campo de la voz humana, en lo referente al proceso de producción de la voz. Además se ingresa al estudio de la fonética inglesa y española. para un mejor entendimiento de los mecanismos de producción de la voz. Para terminar se verán los diferentes parámetros de la voz.

En el Segundo Capítulo se tratan los diferentes aspectos teóricos de los diferentes métodos que se estudian dentro de esta tesis. En este Capítulo se presentan una gran variedad de figuras, como ayuda para una fácil comprensión de los diferentes algoritmos que se tratan.

En el Tercer Capítulo se realiza el desarrollo de la metodología para la implementación de los diferentes programas, para resolver los problemas que se plantean para la ejecución de los algoritmos de detección de tono, en base a los tópicos teóricos tratados en el capítulo anterior.

En el Cuarto Capítulo se presentan los diferentes resultados obtenidos durante las pruebas realizadas con los programas desarrollados, que pasan a formar parte de la biblioteca de programas. Además, se exponen las conclusiones derivadas del trabajo que se ha realizado.

Debo manifestar que la motivación que tuve para la realización de este trabajo, fue la de colaborar en el proyecto de análisis de voz planteado por el Ingeniero Gualberto Hidalgo. Queda pues este trabajo para su utilización futura en la detección del período de tono de señales que se podrán ingresar en forma más sencilla, práctica y real por medio del interfase con la computadora, que se podrá realizar por medio del sistema de adquisición de datos para el análisis de fonemas, que se está desarrollando simultáneamente como trabajo de tesis. por dos de mis compañeros. Este sistema dará la posibilidad de ingresar hasta 64 K-words (palabras de 12 bits) a una frecuencia de muestreo programable entre 8 KHz y 48 KHz.

Deseo manifestar mi agradecimiento a la Escuela Politécnica Nacional, a mis maestros y a todas aquellas personas que han hecho posible la realización de este trabajo. Quiero agradecer en especial al Ingeniero Gualberto Hidalgo por su colaboración durante el desarrollo teórico de esta tesis y a los Ingenieros Patricio Burbano y Efraín del Pino quienes con sus conocimientos y experiencia me guiaron durante el desarrollo de esta tesis.

Carlos E. Tobón M.

C Á P I T U L O P R I M E R O

L a V o z H u m a n a

- 1.1 Proceso y mecanismo de producción de la voz.
- 1.2 Fonética Inglesa
- 1.3 Fonética Española
- 1.4 Parámetros de la voz

C A P I T U L O I

La Voz Humana

Para poder aplicar las técnicas de procesamiento digital de señales a problemas de comunicación de la voz, es necesario entender los fundamentos principales del proceso de producción de la voz.

Antes de discutir las diferentes técnicas para detección de tono, es necesario considerar los diferentes tipos de sonidos que conforman la voz humana. Para ello se hará una breve introducción a la fonética castellana e inglesa y una discusión del lugar y manera de articulación para cada una de las clases de fonemas. Los tópicos a considerarse incluyen propagación del sonido en el canal bucal, analogías a líneas de transmisión y el comportamiento en estado estacionario del sistema bucal en la producción de un sonido sostenido. Esta teoría provee las bases para aproximar la señal de la voz a la señal de salida de un sistema lineal variable en el tiempo (canal bucal) excitado por ruido aleatorio o por pulsos de secuencia cuasiperiódica. Esta aproximación se aplica para obtener modelos discretos en el tiempo para representar señales muestreadas de la voz.

1.1. Proceso y Mecanismo de Producción de la Voz

Las señales de la voz se componen de una secuencia de sonidos, cuya transmisión sirve como una representación simbólica de la información. El ordenamiento de estos símbo-

los del sonido, está gobernado por las reglas del lenguaje. El estudio de estas reglas y sus implicaciones en la comunicación humana es del dominio de la lingüística, y el estudio y clasificación de los sonidos de la voz se llama fonética. Vale la pena discutir las principales clases de sonidos de la voz, antes de proceder a una discusión detallada de los modelos matemáticos de detección de tono.

La Figura 1.1 es una radiografía que pone en evidencia los rasgos importantes de un sistema bucal humano.(1) El canal bucal, indicado por la línea punteada, comienza en la abertura entre las cuerdas bucales o glotis, y termina en los labios. El canal bucal está conformado por la farínge (conexión entre el esófago y la boca) y la boca o cavidad oral. La longitud total promedio del canal bucal es aproximadamente 17 cm. La sección de área del canal bucal, determinada por la posición de la lengua, labios, cuñada y velo del paladar varía entre cero y aproximadamente 20cm^2 . El canal nasal comienza en el velo del paladar y termina en las fosas nasales. Cuando el paladar está en posición baja en el canal nasal está acústicamente acoplado al canal bucal para producir los sonidos nasales de la voz.

En el estudio del proceso de producción de la voz, es conveniente abstraer los rasgos importantes del sistema físico, de tal manera que lleve a un modelo matemático real

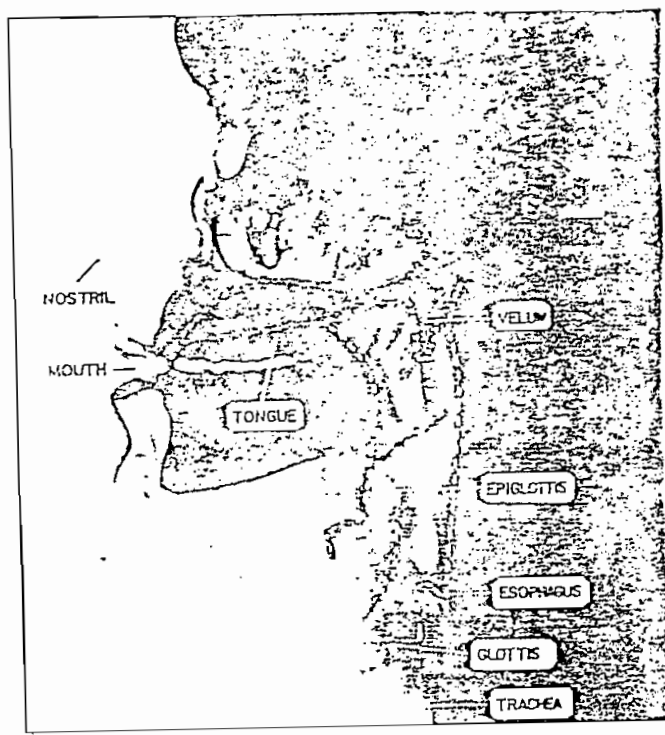


Fig. 1.1 Radiografía del aparato bucal humano.

utilizable. En la Figura 1.2 se muestra un diagrama esquemático del sistema bucal. Este diagrama incluye el sistema subglotal constituido por los pulmones, bronquios y traquea. El sistema subglotal sirve como una fuente de energía para la producción de la voz: La voz es simplemente la onda acústica que es radiada desde este sistema cuando el aire es expulsado desde los pulmones y el flujo de aire resultante es perturbado por una contracción en alguna parte del canal bucal. Como ejemplo de una onda de voz, en la Figura 1.3a se muestra la forma de onda de la expresión inglesa "should we chase", hablada por un locutor masculino. Los rasgos generales de esta onda pueden ser analizados por medio de un estudio detallado del mecanismo de producción de la voz.

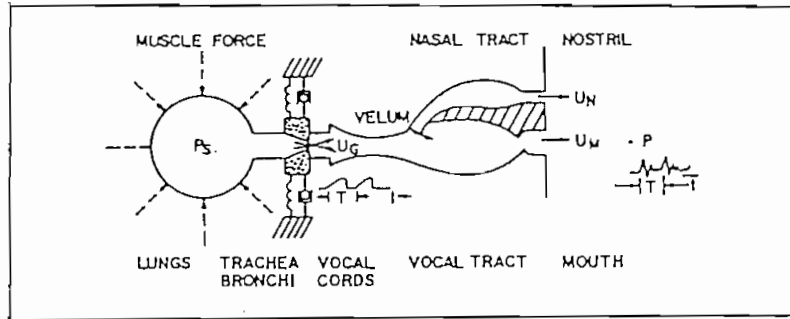


Fig. 1.2 Diagrama esquemático del aparato bucal.

Los sonidos de la voz pueden ser clasificados en tres clases de acuerdo a su modo de excitación. La primera clase son los sonidos "con voz", los que son producidos al forzar aire a través de la glotis con las cuerdas bucales ajustadas a cierta tensión, para que vibren en una oscilación relajada, produciendo pulsos cuasiperiódicos de aire que excitan al canal bucal. En la Figura 1.3b los segmentos con voz son /U/, /d/, /w/, /i/y/e/. La segunda clase con los sonidos "fricativos o sin voz", que son generados al formar una contracción en algún punto del canal bucal (usualmente hacia el final de la boca), y forzando el aire a través de la contracción a una alta velocidad suficiente para producir turbulencia. Esto crea una fuente de ruido de espectro ancho para excitar el canal bucal. El segmento marcado \int en la figura 1.3a es la "sh" fricativa. La tercera clase son los sonidos "explosivos", resultado de realizar una oclusión completa (usualmente hacia el frente del canal bucal), realizando presión detrás de la oclusión, y suspendiéndola abruptamente. De este tipo de excitación es el sonido marcado /t \int / al comienzo de la cuarta línea de la figura 1.3a. Nótese el intervalo de

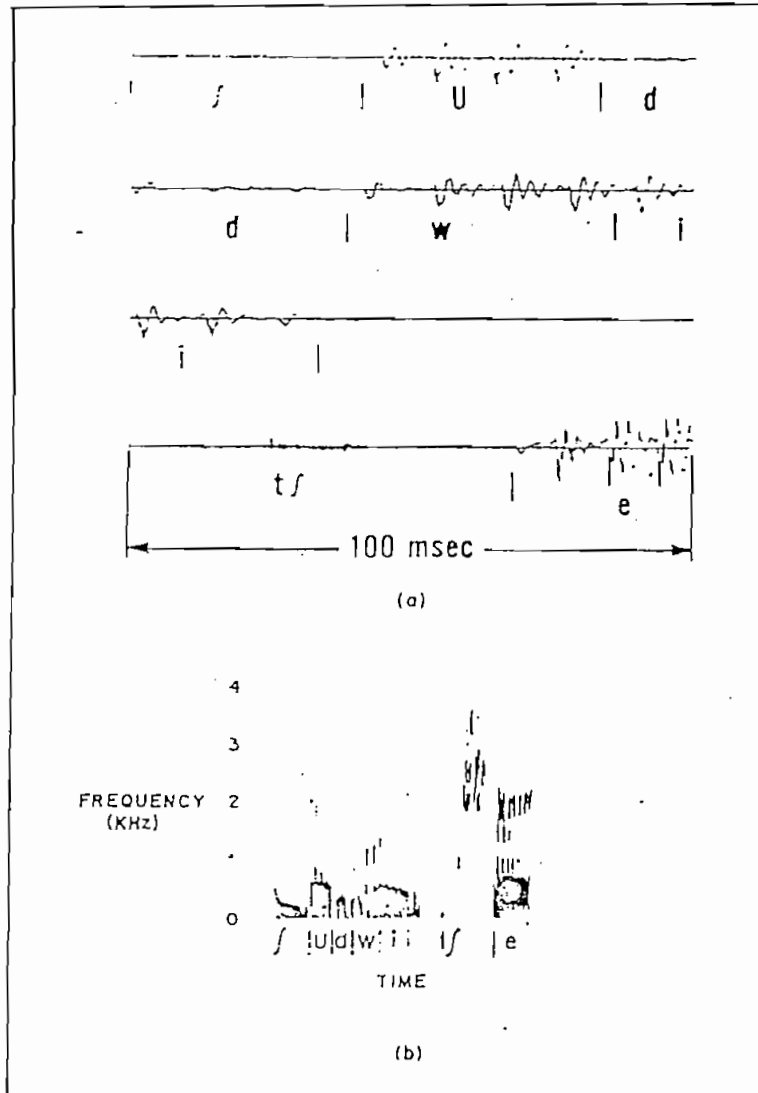


Fig. 1.3 (a) Forma de onda de la expresión "should we chase";

(b) Espectrograma correspondiente

amplitud muy pequeña al final del tercer renglón que precede al inicio de la forma de onda del sonido. Este intervalo corresponde al tiempo de cierre del canal bucal.

En la Figura 1.2 se muestran el canal bucal y el canal nasal como tubos de sección de área no uniforme. A medida que el sonido se propaga a lo largo de estos tubos, el espectro de frecuencia toma forma de acuerdo a la selectividad

de frecuencia del tubo. Este efecto es muy similar al observado en instrumentos de viento. En el contexto de producción de la voz, las frecuencias de resonancia del tubo del canal bucal se llaman "formantes de frecuencia" o simplemente "formantes". Los formantes de frecuencia dependen de la forma y dimensión del canal bucal; cada forma se caracteriza por un grupo de formantes de frecuencia. Diferentes sonidos se forman variando la forma del canal bucal. Por lo tanto, las propiedades espectrales de la señal de la voz varían con el tiempo conforme va variando la forma del canal bucal.

Las características del espectro variable en el tiempo de la señal de la voz pueden ser graficados utilizando un espectrógrafo de sonido. Este aparato produce un patrón bidimensional llamado "espectrograma", en el cual la dimensión vertical corresponde a la frecuencia y la dimensión horizontal al tiempo. La oscuridad del patrón es proporcional a la energía de la señal. Por lo tanto, las frecuencias de resonancia del canal bucal aparecen como bandas oscuras en el espectrograma. Las regiones con voz están caracterizadas por una apariencia estriada, debida a la periodicidad de la onda en el tiempo: mientras las regiones sin voz están compactamente oscuras. El espectrograma de la figura 1.3b está graficado en correspondencia con el gráfico de la Figura 1.3a de tal forma que los rasgos del dominio del tiempo y del dominio de la frecuencia puedan ser correlacionadas. (2)

El espectrógrafo de sonido ha sido una de las herramientas principales en la investigación de la voz, y aunque aparatos más flexibles pueden ser generados utilizando técnicas de procesamiento digital, sus principios básicos siguen siendo muy usados (3).

1.2 Fonética Inglesa

La mayoría de los idiomas, incluyendo el español y el inglés, pueden ser descritos en términos de un grupo de sonidos distintivos llamados "fonemas".

En particular, para el inglés americano, se definen cerca de 42 fonemas incluyendo vocales, diptongos, semivocales y consonantes. La lingüística es la ciencia que estudia los rasgos distintivos o características de los fonemas. (4.5) Para nuestro propósito es suficiente considerar la caracterización acústica de los diferentes sonidos, incluyendo el lugar y la manera de articulación, formas de onda y su caracterización espectrográfica.

En la Tabla 1.1 se indican las diferentes clases en que se dividen los fonemas en inglés americano. Las cuatro clases principales de sonidos son vocales, diptongos semivocales y consonantes. Cada una de estas clases se dividen en subclases que están relacionadas con la manera y lugar de articulación del sonido dentro del canal bucal.

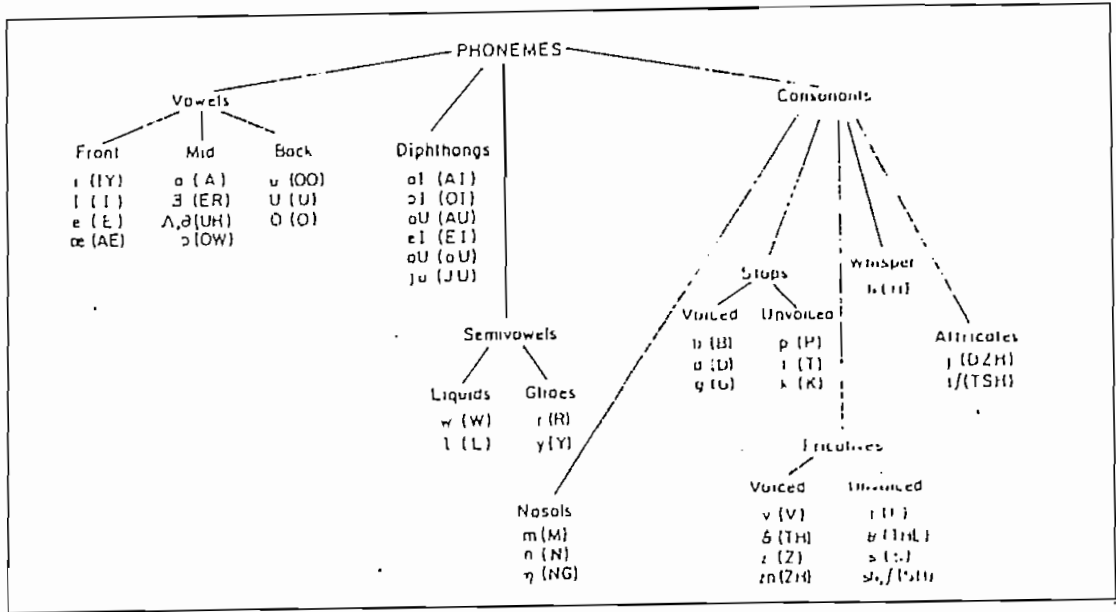


Tabla 1.1 Fonemas del Inglés Americano

Cada uno de estos fonemas pueden ser clasificados ya sea como sonido continuo o discontinuo. Los sonidos continuos se producen por una configuración fija (no variable en el tiempo) del canal bucal excitado por una fuente apropiada. El grupo de sonidos continuos incluye las vocales, los fricativos (con voz y sin voz), y los nasales. El resto de sonidos (diptongos, semivocales, puntuales y africativas) se producen por una configuración variable del canal bucal. A estas se les clasifica como no continuos o discontinuos.

"Las vocales" se producen al excitar el canal bucal en estado estacionario con pulsos de aire cuasiperiódicos causados por la vibración de las cuerdas bucales. La forma en que la sección transversal de área varía a lo largo del

canal bucal determina las frecuencias resonantes del canal (formantes) y de este modo se produce el sonido. La dependencia de la sección transversal de área de la distancia a lo largo del canal se llama "función de área" del canal bucal. La función de área para una vocal se determina primordialmente por la posición de la lengua, pero las posiciones de la guíada, labios y el paladar también influyen en el sonido resultante. Por ejemplo, para formar la vocal /a/ como en "father", el canal bucal se abre en la parte frontal y es contraído en la parte posterior por el cuerpo de la lengua. En contraste, la vocal /i/ como en "eve" se forma alzando la lengua contra el paladar, causando una contracción en la parte posterior del canal bucal. Por lo tanto, cada vocal puede ser caracterizada por la configuración del canal bucal (función de área) que se use en su producción. Es obvio que ésta caracterización es un poco imprecisa, pues existen diferencias inherentes entre los canales bucales de los seres humanos. En la Tabla 1.2 se indican los valores promedios de las tres primeras frecuencias formantes de las vocales para locutores masculinos. Aunque existe una gran variación en las formas de las vocales, la información de la Tabla 1.2 sirve como una caracterización de las vocales.

Las formas de onda acústica y los espectrogramas para cada una de las vocales del inglés se indican en la Figura 1.4. Los espectrogramas muestran un patrón de resonancia diferente para cada vocal. Las ondas acústicas muestran

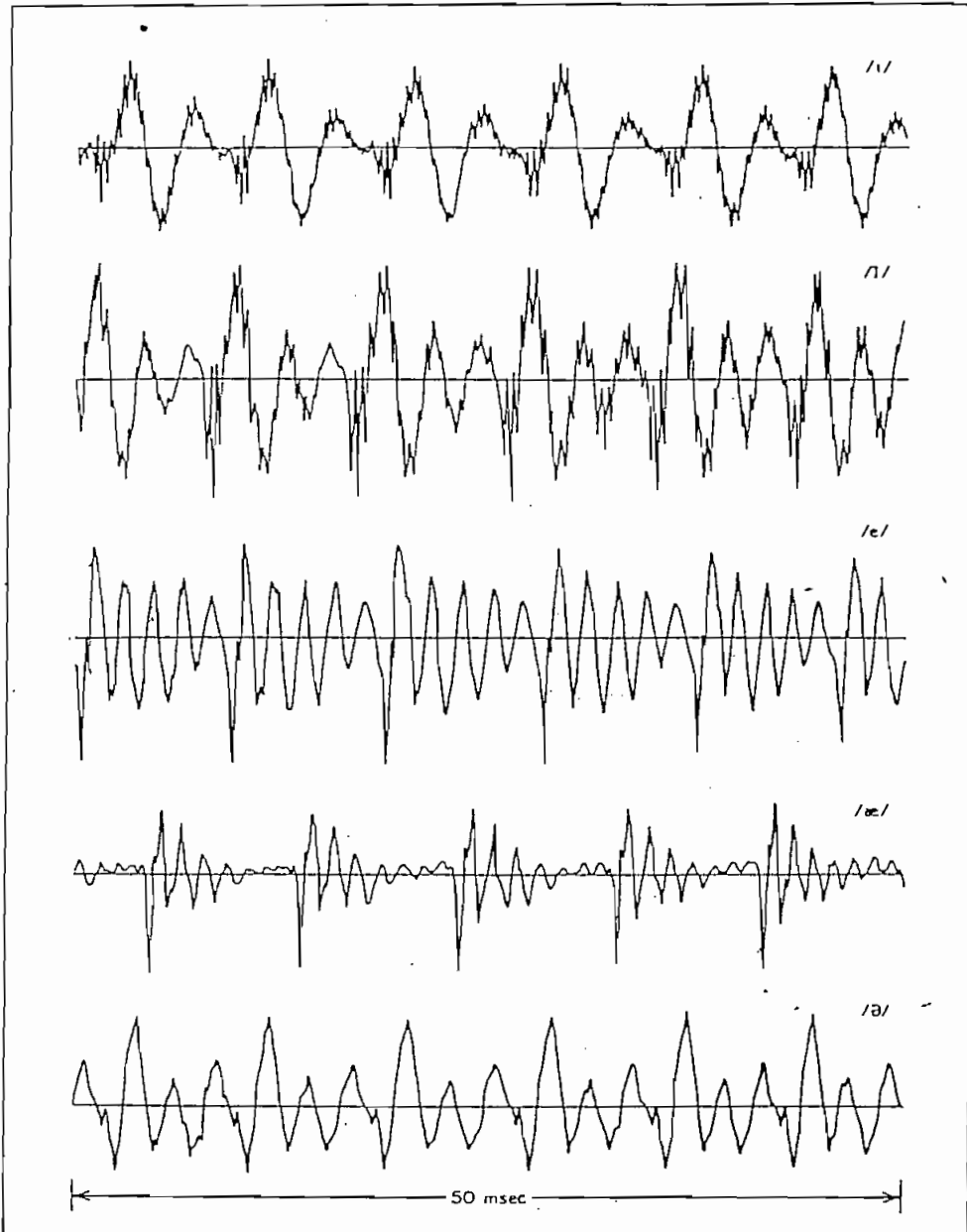


Fig. 1.4 Formas de ondas y espectrogramas de las diferentes vocales inglesas

las características de periodicidad de los sonidos con voz. Por ejemplo. la vocal /i/ muestra una oscilación amortiguada de baja frecuencia, sobre la cual se superpone una oscilación de alta frecuencia relativamente fuerte. Esto es con-

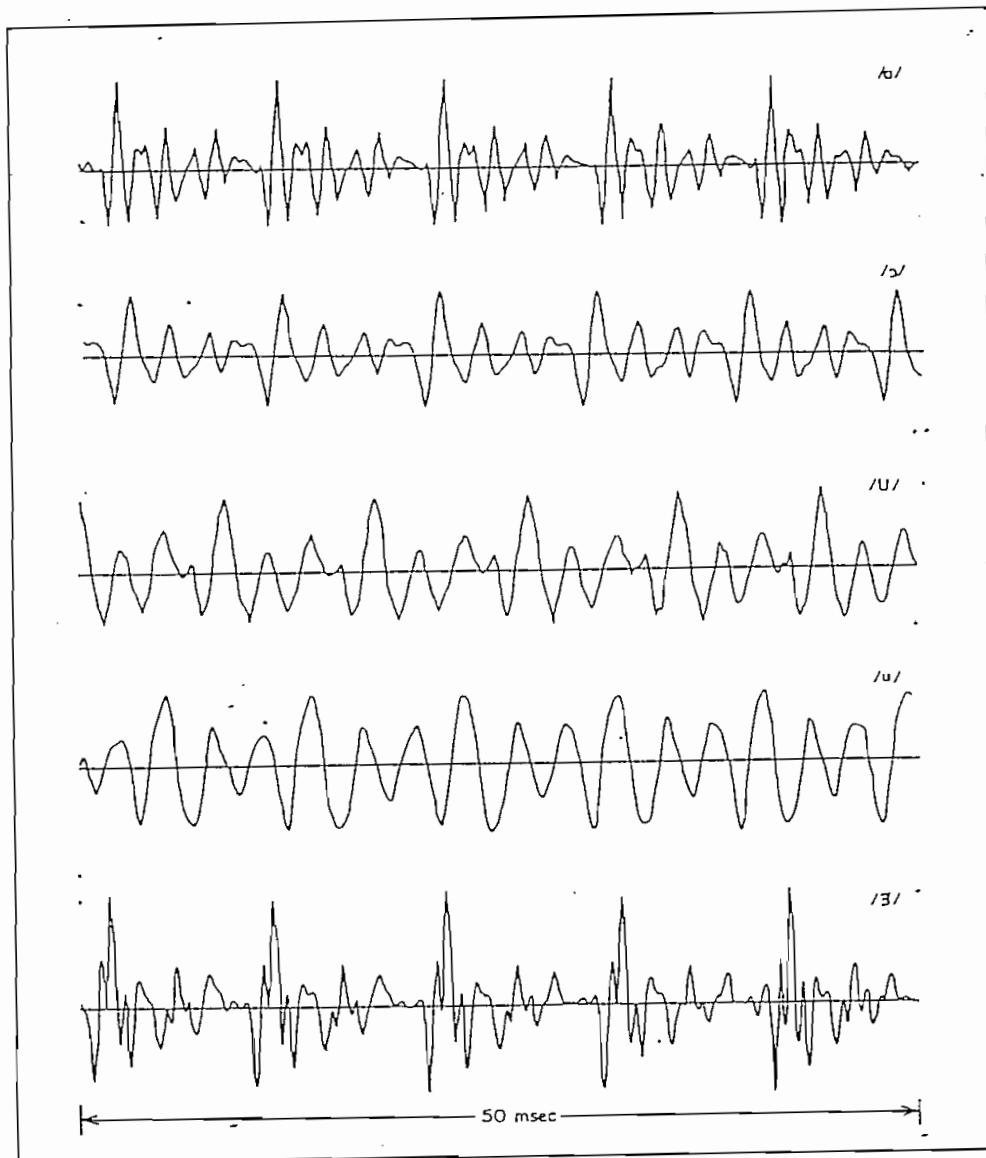


Fig. 1.4 Formas de ondas y espectrogramas de las diferentes vocales inglesas.

sistente con una primera formante baja, y una segunda y tercera formantes altas, correspondientes a la Tabla 1.2. En contraste la vocal /u/ muestra poca energía de alta frecuencia como consecuencia de primera y segunda formantes de baja frecuencia.

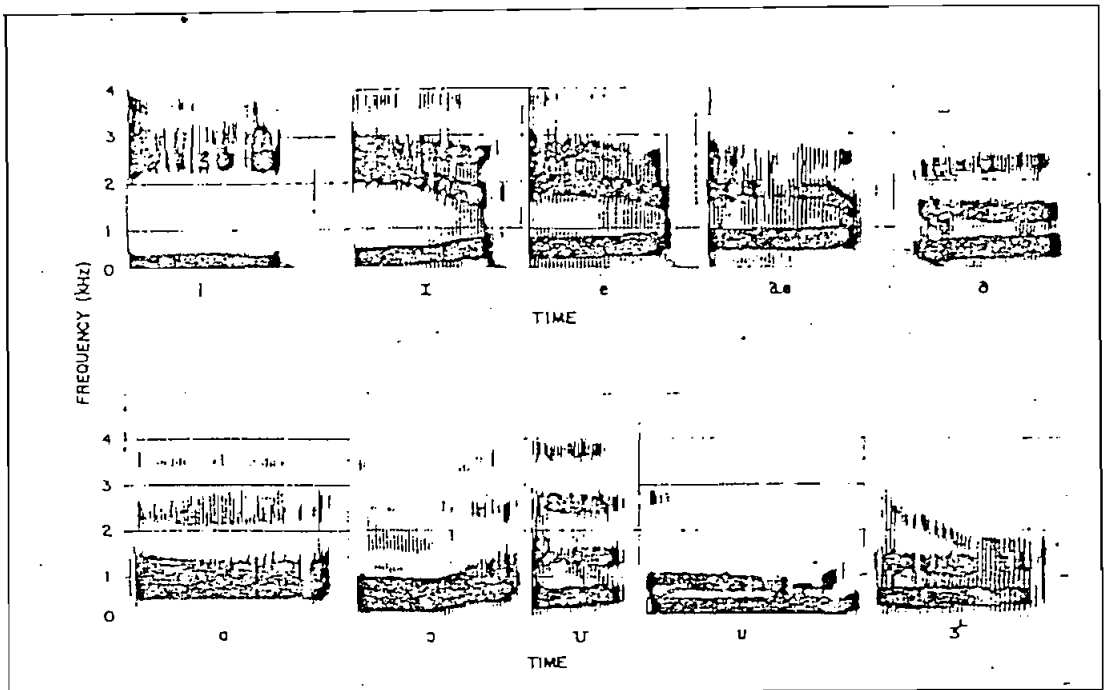


Fig. 1.4 Formas de ondas y espectrogramas de las diferentes vocales inglesas.

Existe un poco de ambigüedad y discrepancia en cuanto qué es un "diptongo". Una definición razonable es que un diptongo es un monosílabo deslizante, que comienza en o cerca de la posición para otra vocal. De acuerdo a esta definición, en inglés americano hay seis diptongos: /eI/ (como en bay), /oU/ (como en boat). /aI/ (como en buy). /aU/ (como en how), /oI/ (como en boy), y /ju/ (como en you).

Los diptongos se producen por la variación suave del canal bucal entre las configuraciones de las vocales del diptongo. Para ilustrar este punto, la Figura 1.6 muestra el plano de las medidas de la segunda formante versus la primera formante para los diptongos. Las flechas de esta figura indican la dirección de circulación de las formantes a medida que se incrementa el tiempo. Los círculos punteados en

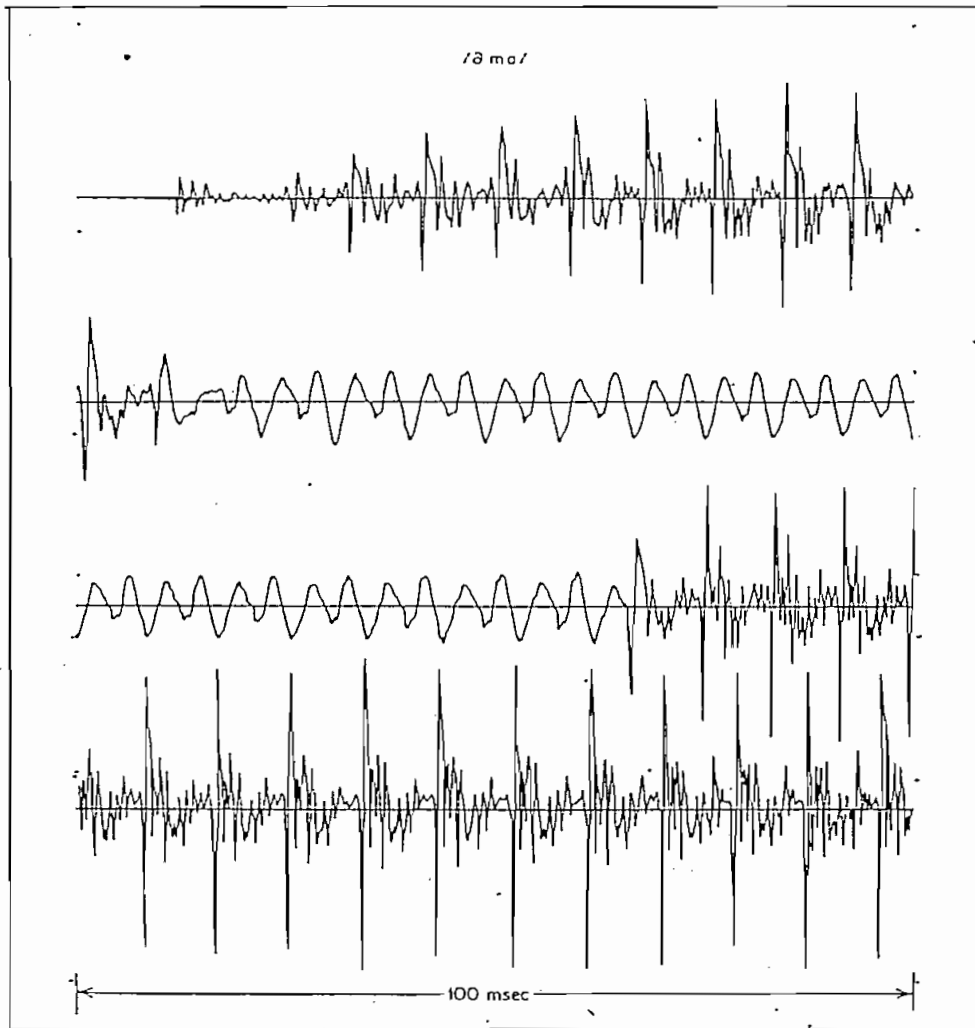


Fig. 1.5 Formas de onda y espectrogramas de las expresiones /UH-M-A/ y /UH-N-A/

| Formantes de Frecuencia de las Vocales | | | | |
|--|----------------|----------------|----------------|----------------|
| Símbolo escrito para la vocal | Palabra típica | F ₁ | F ₂ | F ₃ |
| IY | (beet) | 270 | 2290 | 3010 |
| I | (bit) | 390 | 1990 | 2550 |
| E | (bet) | 530 | 1840 | 2480 |
| AE | (bat) | 660 | 1720 | 2410 |
| UH | (but) | 520 | 1190 | 2390 |
| A | (hot) | 730 | 1090 | 2440 |
| OW | (bought) | 570 | 840 | 2410 |
| U | (foot) | 440 | 1020 | 2240 |
| OO | (boot) | 300 | 870 | 2240 |
| ER | (bird) | 490 | 1350 | 1690 |

Tabla 1.2 Valores promedios de los Formantes de Frecuencia de las vocales inglesas

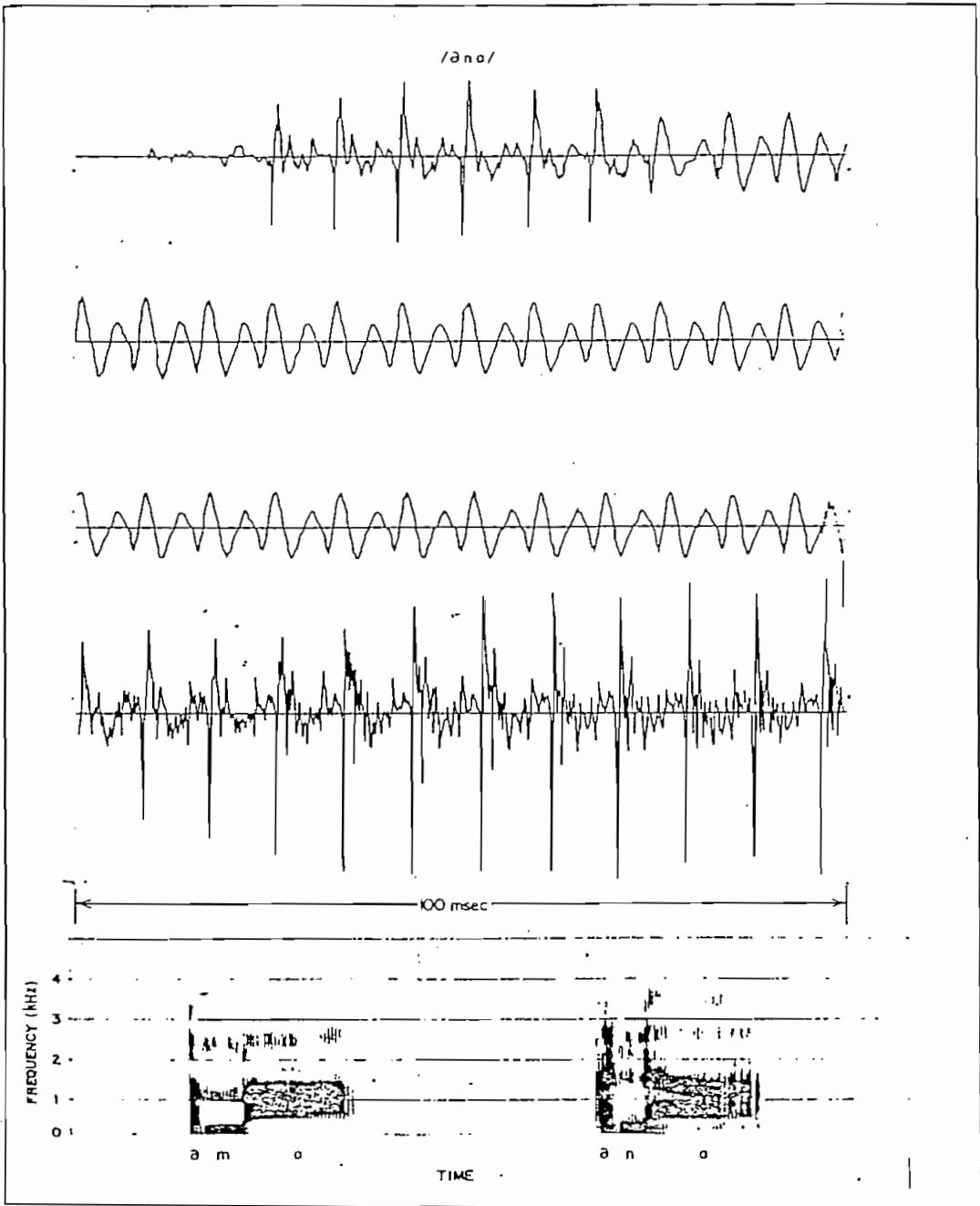


Fig. 1.5 Formas de onda y espectrogramas de las expresiones /UH-M-A/ y /UH-N-A/.

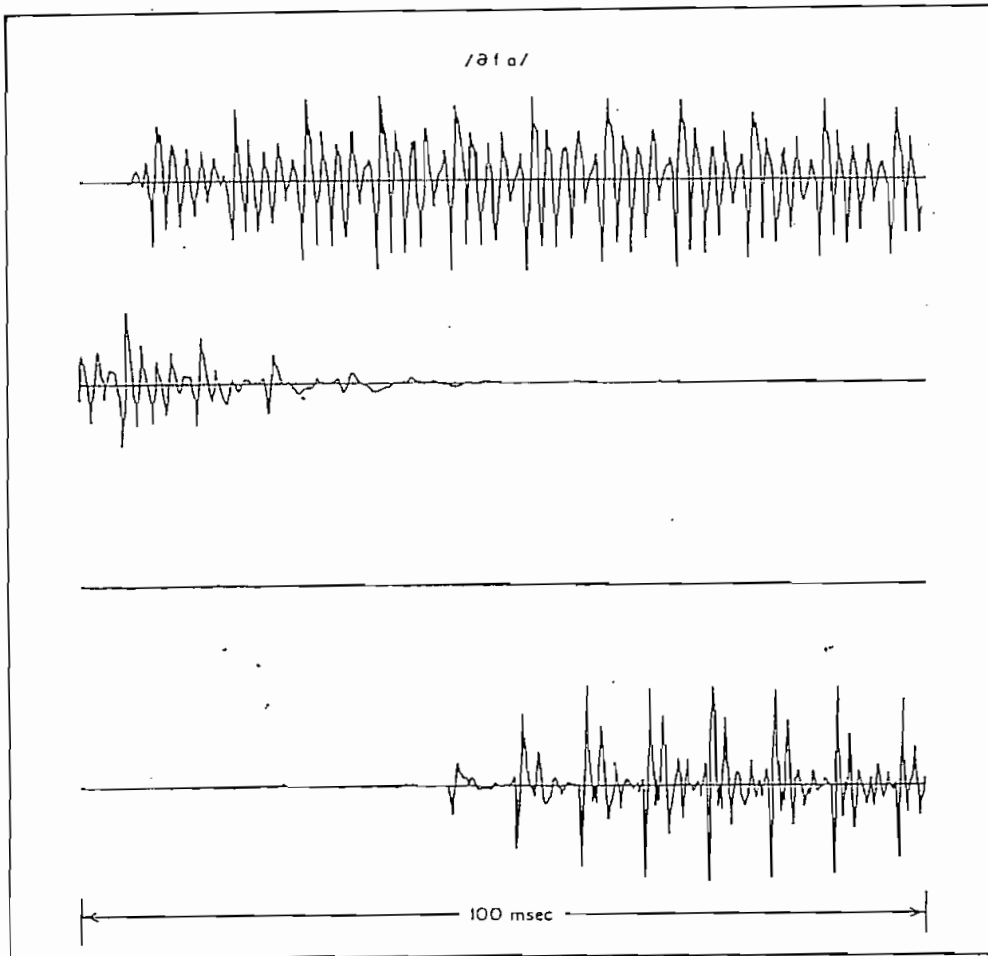


Fig. 1.6 Formas de onda y espectrogramas de las expresiones /UH-F-A/ /UH-S-A/ y /UH-SH-A/.

esta figura indican las posiciones promedio de las vocales. Basados en esta información y en otras medidas, los diptongos se pueden caracterizar por una función de área del canal bucal, variable en el tiempo, que varíe entre dos configuraciones de vocales. (6).

El grupo de sonidos /w/, /l/, /r/, y /y/ es un poco difícil de caracterizar. Estos sonidos se llaman "semivocales" porque su naturaleza es parecida a la de las vocales. Son ge-

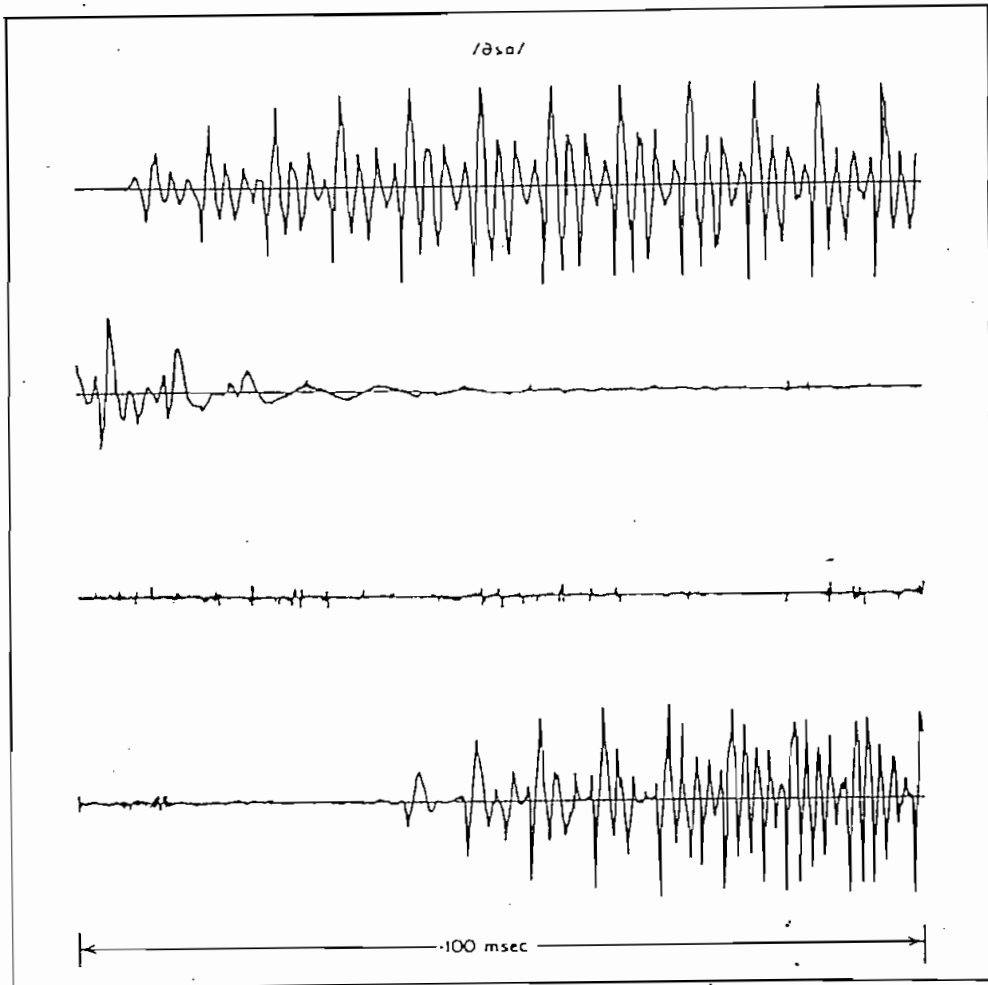


Fig. 1.6 Formas de onda y espectrogramas de las expresiones /UH₂F₂A/ UH₂S₂A/ y /UH₂SH₂A/.

neralmente caracterizados por una transmisión deslizante en el canal bucal (función de área) entre fonemas adyacentes; por lo tanto las características acústicas de estos sonidos están influenciadas por el contexto en que ocurren. Un ejemplo de la semivocal /w/ está indicado en la Figura 1.3.

Las "consonantes nasales" /m/./n/, y /ŋ/ se producen con una excitación glotal y el canal bucal totalmente contraído en algún punto a lo largo del camino oral. El velo del paladar se baja, de tal forma que el aire fluya a través

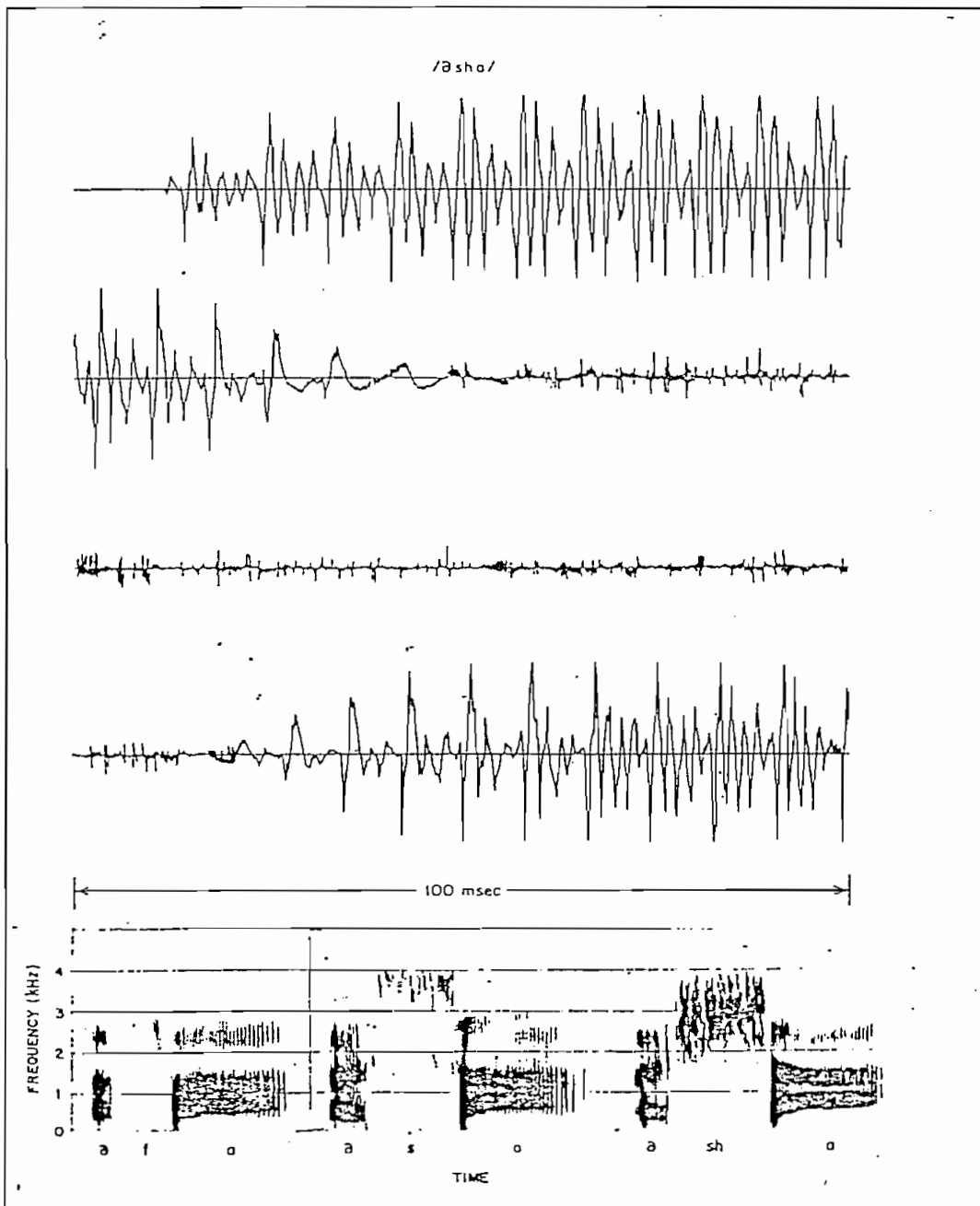


Fig. 1.6 Formas de onda y espectrogramas de las expresiones /UH-F-A/UH-S-A/ y /UH-SH-A/.

del canal nasal, siendo el sonido emitido por las fosas nasales. La cavidad oral, aunque está contraída hacia la parte frontal, se mantiene acústicamente acoplada a la faringe. Así la boca sirve como una cavidad resonante que atrapa la energía acústica a ciertas frecuencias naturales. En lo que concierne a la irradiación del sonido, estas frecuencias resonantes de la cavidad oral, aparecerán como antiresonancias, o ceros de transmisión de sonido. Además, las consonantes nasales y las vocales nasalizadas (esto es algunas vocales que siguen o preceden a consonantes nasales) son caracterizadas por resonancias que son de espectro más ancho, o más amortiguadas, que aquellas de las vocales. El ensanchamiento de las resonancias nasales se debe a que la superficie interna del canal nasal es circunvalado de tal forma que la cavidad nasal tiene una razón de superficie de área a sección de área transversal grande.

Las tres consonantes nasales se distinguen por el lugar a lo largo del canal oral en donde se produce una contracción total. Para la /m/ la contracción se produce en los labios, para la /n/ la contracción es justo atrás de los dientes, y para la /ŋ/ la contracción es justo delante del velo del paladar. En la Figura 1.5 se indican las formas de ondas típicas y espectrogramas para dos consonantes nasales en el contacto vocal-nasal-vocal. Se ve claro que las formas de onda de /m/ y /n/ son parecidas. Los espectrogramas mues-

tran una concentración de energía de baja frecuencia, con un rango medio de frecuencias que no contiene picos prominentes. Esto es por la combinación particular de resonancias y antiresonancias que resultan del acoplamiento de los canales nasal y oral (7).

Los "fricativos sin voz" /f/, /e/, /s/ y /sh/ son producidos por la excitación del canal bucal, por medio de un flujo de aire que se vuelve turbulento en la región de la contracción en el canal oral. La localización de la contracción sirve para determinar que sonido fricativo se produce. Para la /f/ la contracción se produce cerca de los labios, para la /e/ se produce cerca de los dientes, para la /s/ se da cerca de la mitad del canal oral, y para la /sh/ se produce cerca de la parte posterior del canal oral.

Por lo tanto, el sistema para producir fricativos sin voz, consiste en una fuente de ruido en el lugar de la contracción, que separa el canal bucal de las dos cavidades. El sonido se emite desde los labios, esto es desde la cavidad frontal, La cavidad posterior sirve, como en el caso de los nasales, para atrapar energía y así introducir antiresonancias a la salida bucal. (2,8).

La Figura 1.6 muestra las formas de onda y espectrogramas de los fricativos /f/, /s/ y /sh/. La naturaleza no-periódica de la excitación fricativa es obvia en los gráficos. Las diferencias espectrales entre ellos pueden apreciarse al comparar los tres espectrogramas.

Los "fricativos con voz" /v/, /th/, /z/ y /zh/ son los equivalentes de los fricativos sin voz /f/, /θ/, /s/ y /sh/ respectivamente, por lo que el lugar de la contracción para cada uno de los fonemas correspondientes es idéntico. Los fricativos con voz difieren notoriamente de sus equivalentes sin voz, en que para su producción entran dos fuentes de excitación (ruido y sonido de la glotis). Para los fricativos con voz las cuerdas bucales están vibrando, por lo que una de las fuentes de excitación está en la glotis.

De cualquier modo, desde que el canal bucal esté contraído en algún punto hacia la glotis, el flujo de aire se convierte en turbulencia en los alrededores de la contracción. Por lo que el espectro de fricativos con voz debe tener dos componentes diferentes. Estos rasgos de excitación pueden observarse en la Figura 1.7 que muestra las formas de onda y espectros para varios fricativos con voz. La similitud del fricativo sin voz /f/ con el fricativo con voz /v/, puede verse comparando los espectrogramas de las Figuras 1.6 y 1.7.

Las "consonantes oclusivas con voz" /b/, /d/, y /g/ son transitorios, sonidos discontinuos que se producen por la formación de presión detrás de una contracción total en alguna parte del canal oral, y liberándola repentinamente. Para la /b/ la contracción es en los labios, para la /d/ la

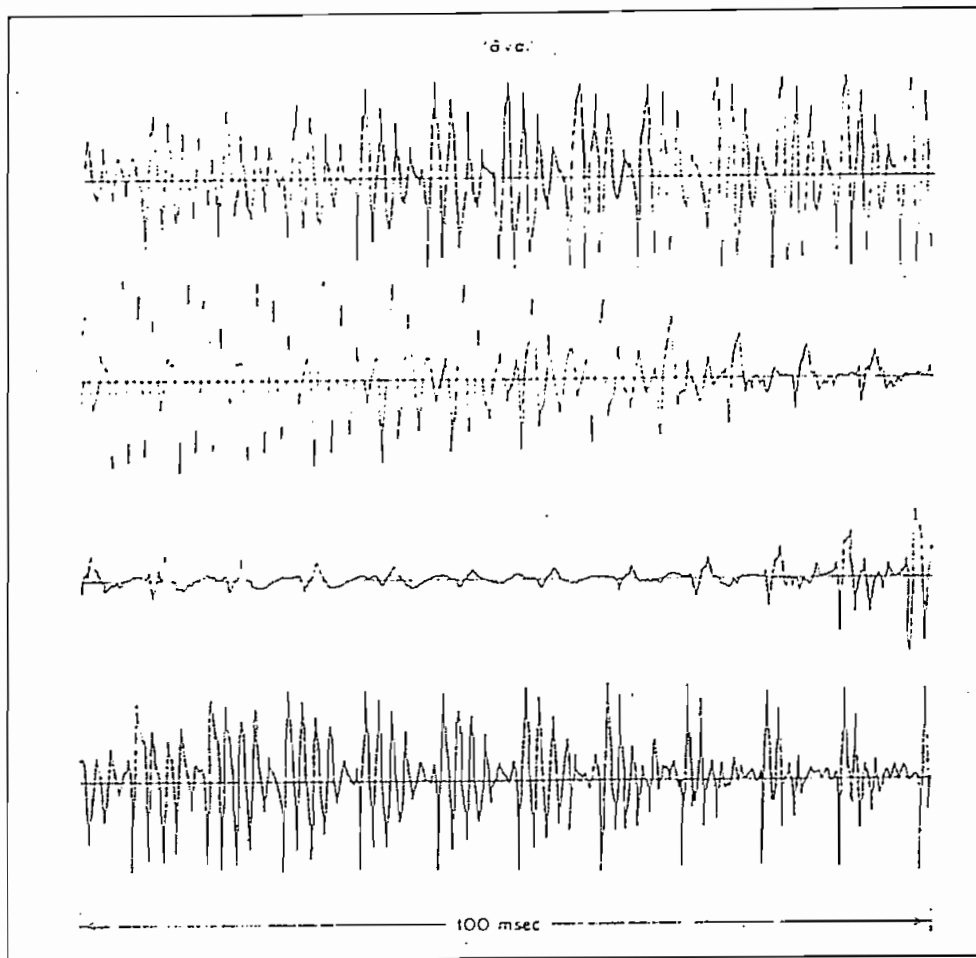


Fig. 1.7 Formas de onda y espectrogramas de las expresiones /UH-V-A/ y /UH-ZH-A/.

contracción es detrás de los dientes, y para la /g/ es cerca del velo del paladar. Durante el período en que hay una contracción total en el canal, no se emite ningún sonido de los labios. Sin embargo, con frecuencia a través de las paredes de la garganta. Esto ocurre cuando las cuerdas bucales son capaces de vibrar, aunque el canal bucal esté cerrado en algún punto.

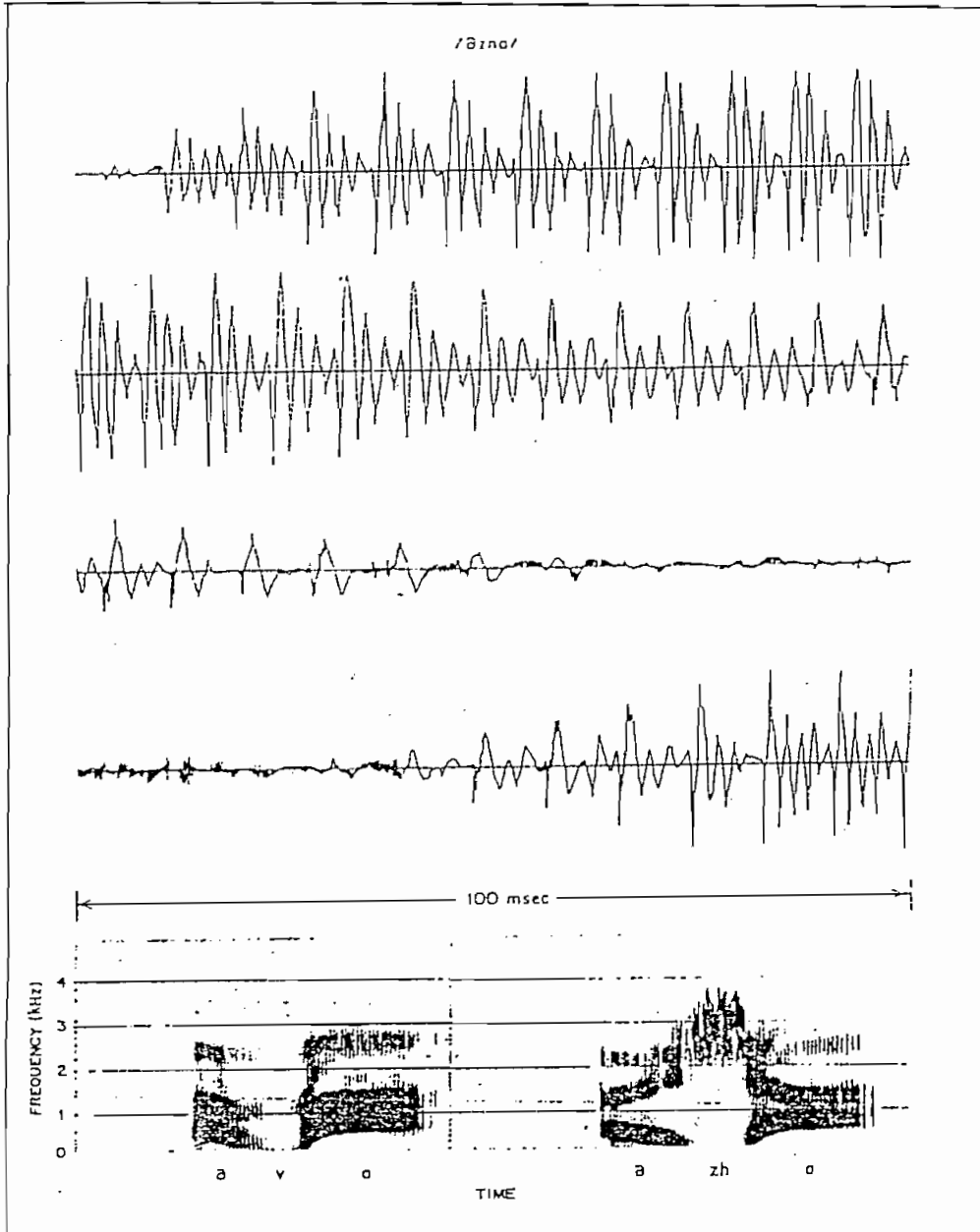


Fig. 1.7 Formas de onda y espectrogramas de las expresiones /UH-V-A/ y /UH-ZH-A/.

Como los sonidos oclusivos son de naturaleza dinámica, sus propiedades están influenciadas por la vocal que sigue a la consonante. Por la cual, la forma de onda para consonantes oclusivas da poca información acerca de la consonante en particular. En la Figura 1.8 se muestra la forma de onda y el espectrograma de la sílaba /UH-B-A/. La /b/ muestra unos pocos rasgos distintivos, excepto para la excitación con voz, y carece de energía de alta frecuencia. (9).

Las "consonantes oclusivas sin voz" /p/, /t/ y /k/ son similares a sus equivalentes con voz /b/, /d/ y /g/ con una excepción. Durante el período de cierre total del canal, a medida que se forma presión, las cuerdas bucales no vibran. A continuación del período de cierre, a medida que se suelta la presión, hay un pequeño intervalo de fricción (debido a la turbulencia del aire que se escapa) seguido por un período de aspiración (el aire fluye desde la glotis excitando los resonantes del canal bucal) antes de que la excitación con voz inicie.

La Figura 1.9 muestra las formas de onda y espectrogramas de las consonantes puntuales sin voz /p/ y /t/. El tiempo durante el cual se forma presión, puede observarse claramente. También, se ve que el tiempo y contenido de frecuencia del ruido de fricción y aspiración varía notablemente de acuerdo a la consonante oclusiva.

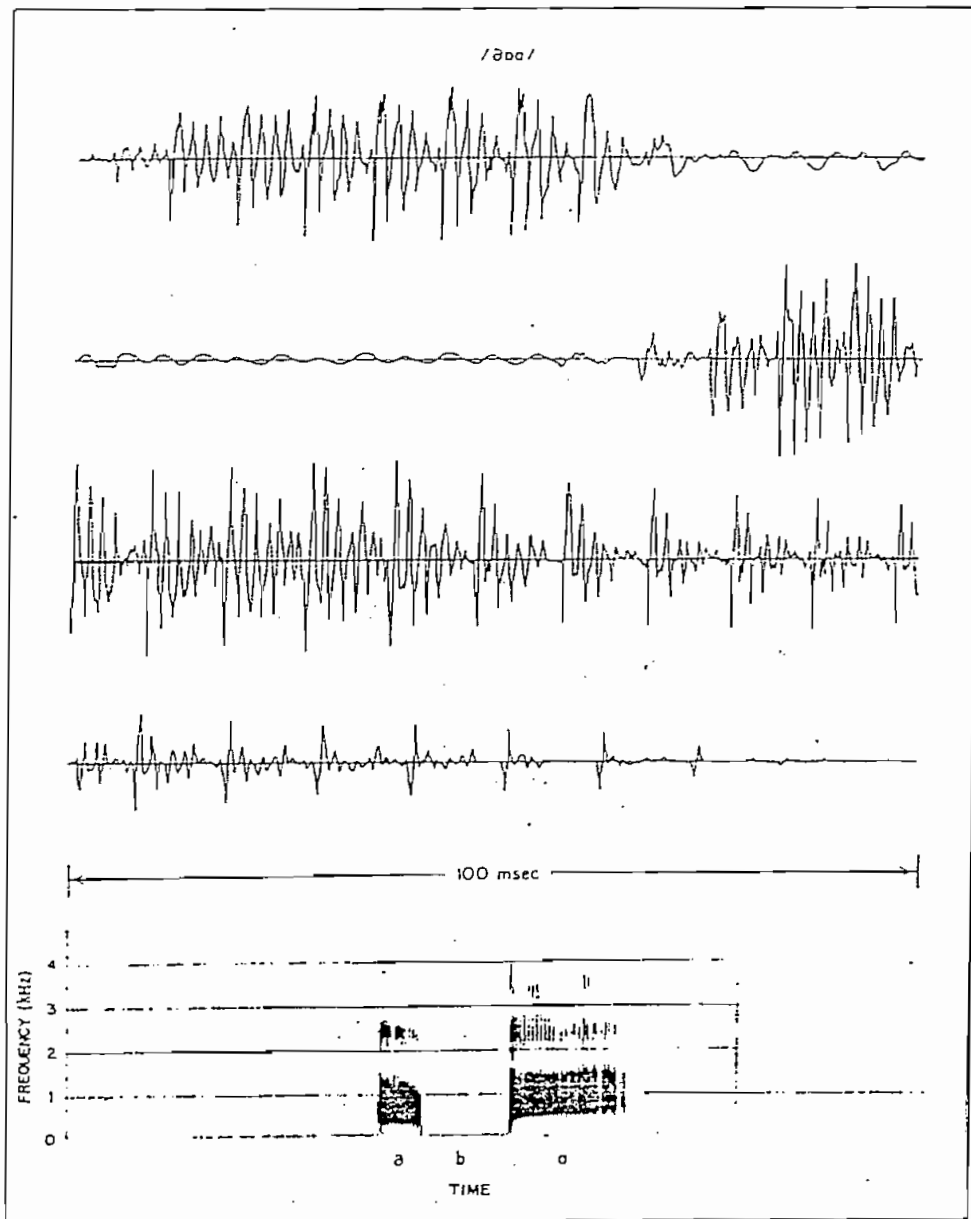


Fig. 1.8 Formas de onda y espectrogramas de la expresión /θba/

Las consonantes del inglés americano restantes son las "africativas" /tʃ/ y /dʒ/, y el fonema /h/. La africana

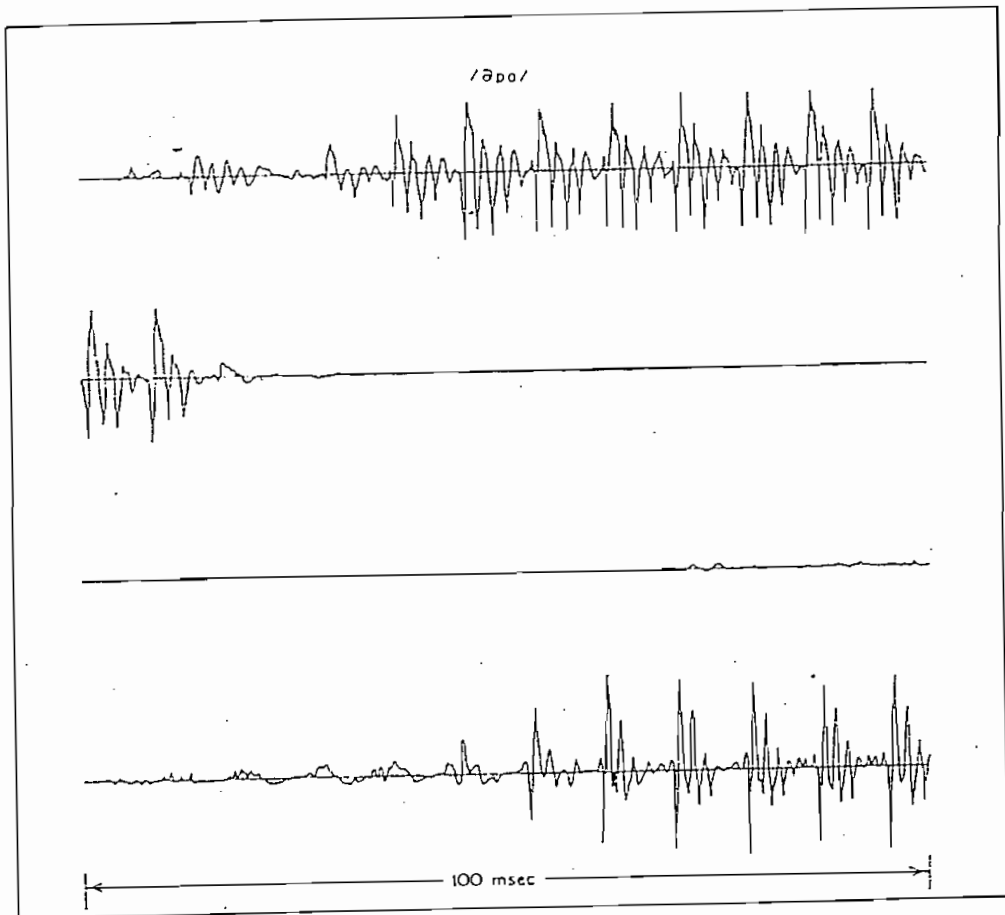


Fig. 1. 9 Formas de onda y espectrogramas de las expresiones /UH-P-A/ y /UH-T-A/.

sin voz /tʃ/ es un sonido dinámico, que puede ser modelado como la concatenación de la oclusiva /t/ y la fricativa /ʃ/. La africativa con voz /dʒ/ puede ser modelada como la concatenación de la puntual /d/ y la fricativa /ʒ/. Finalmente, el fonema /h/ es producido por la excitación del canal bucal, por medio de un flujo de aire, sin vibración de las cuerdas bucales, pero produciéndose turbulencia en la glotis. Las características de la /h/ son invariablemente las de la vocal que la sigue, puesto que el canal bucal asume la posición de la vocal que sigue, durante la producción de la /h/. (10)

002010

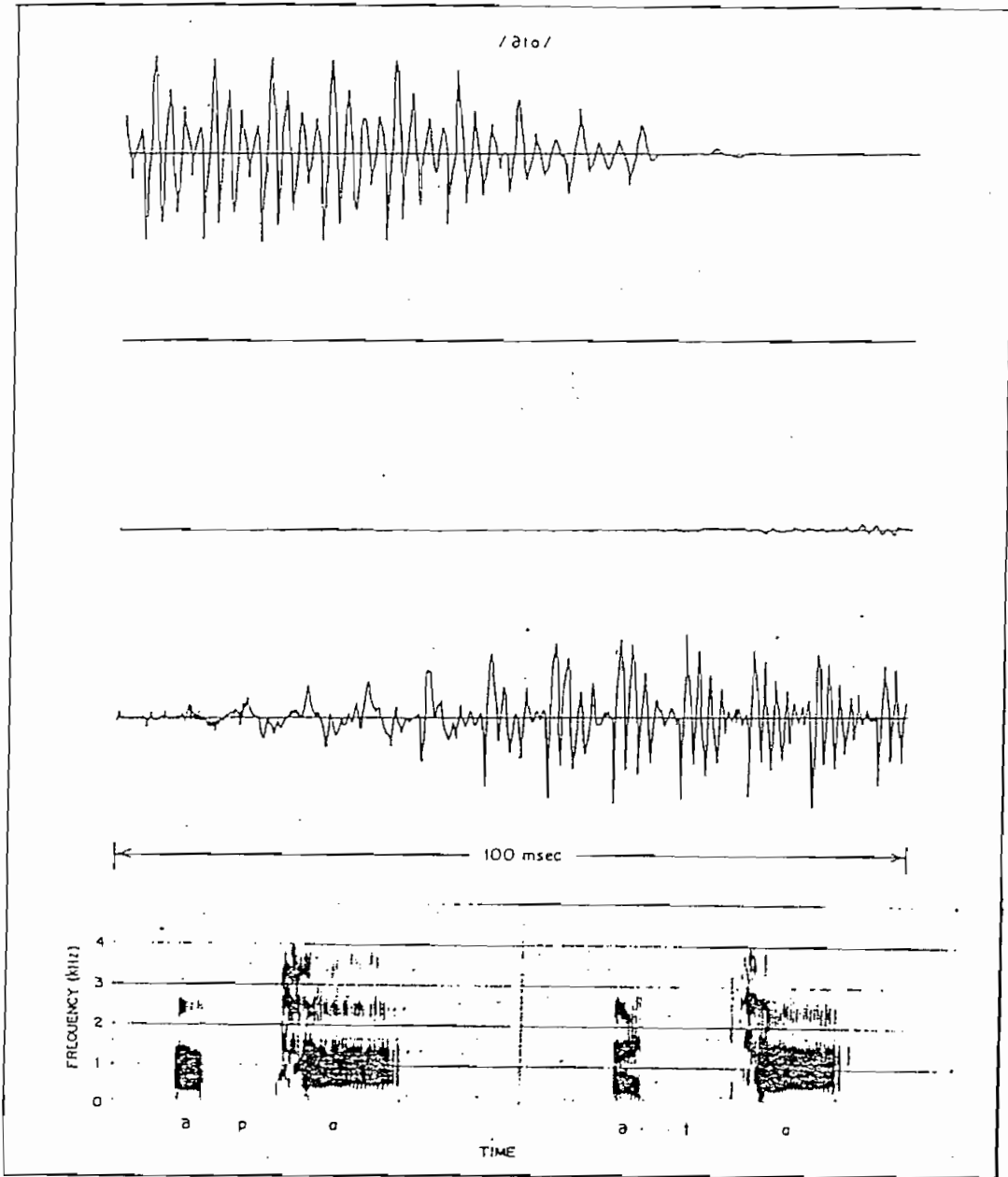


Fig. 1.9 Formas de ondas y espectógramas de las expresiones /UH-P-A/ y /UH-T-A/.

1.3 Fonética Española

A continuación se realiza una breve descripción de la fonética española y sus características.

Los "fonemas vocálicos" son los fonemas que por sí solos, aisladamente o combinados entre sí, pueden formar palabras o sílabas. Los demás fonemas, incapaces de formar por sí solos, sin el concurso de una vocal, palabras o sílabas, son "fonemas consonánticos". Habrá que estudiar los fonemas, para mayor claridad, por separado, y establecer un sistema vocálico y otro consonántico; aunque no debe olvidarse que están íntimamente relacionadas. Algunos fonemas vocálicos están tan emparentados con otros consonánticos que llegan a neutralizar su contraste en algunas posiciones dentro de la palabra.

El español utiliza fonológicamente dos de las propiedades articulatorias y acústicas que sirven para la distinción de los fonemas vocálicos entre sí. Una, es el grado de abertura, que condiciona la mayor o menor frecuencia del primer formante de la vocal. La otra es la configuración de la cavidad bucal según la posición de la lengua y los labios, reflejada en la mayor o menor frecuencia del segundo formante de la vocal (timbre).

El español distingue tres grados de abertura: abertura máxima, media y mínima. El grado de abertura máxima o de /a/ tiene su primer formante situado en unos 700 Hz de fre-

cuencia. El grado de abertura media de /e,o/ tiene en primer formante a unos 500 Hz. Y el grado de abertura mínima o de /i,u/ tiene el primer formante bajo los 400 Hz.

Según la forma y tamaño del canal bucal y su diferente timbre, tenemos dos fonemas graves de localización posterior /u/ y /o/, un fonema medio /a/ y dos fonemas agudos /e/ y /i/ de localización anterior. La /u/ tiene un segundo formante a 700 Hz, la /o/ un segundo formante a 1000 Hz, la /a/ tiene el segundo formante a 1500 Hz, la /e/ un segundo formante a 1800 Hz, y la /i/ un segundo formante a 2000 Hz. Se trata, pues, de un sistema vocálico triangular:

u i
 o e
 a

Fonológicamente, estos fonemas se definen así:

/a/, fonema vocal de abertura máxima de timbre neutro; /e/, fonema vocal de abertura media de timbre agudo (posición anterior); /o/, fonema vocal de abertura media de timbre grave (posición posterior); /i/, fonema vocal de abertura mínima de timbre agudo; y /u/, fonema vocal de abertura mínima de timbre grave.

La serie /e,i/, anterior y aguda, y la posterior y grave /o,u/, son por su articulación palatales sin labializar y velares con labialización, respectivamente.

Estas cinco vocales presentan variantes de articulación, según su posición en la palabra. El fonema /a/, pre-

senta bajo ciertas condiciones la articulación palatal /a/, o velar /a/; que no son más que variantes combinatorias. (11) En el espectrograma se observa que en estos casos lo que varía sobre todo es el segundo formante. En el caso de la /a/ sube en la escala de frecuencias, acercándose a la altura normal del fonema /e/. En el caso de la /a/ en cambio, disminuye su frecuencia, aproximándose a la del fonema /o/.

El fonema /e/ también varía fonéticamente como /e/ abierta y como /e/ cerrada, según los sonidos con que está en contacto. En los espectrogramas, el segundo formante de /e/ está a 1700 Hz, casi la frecuencia de las variantes palatales de /a/; en cambio el segundo formante de /e/ está a 1900 Hz, casi la frecuencia de la /i/.

Paralelamente, el fonema /o/ presenta dos matices: /o/ abierta y /o/ cerrada. El segundo formante de /o/ está a 1100 Hz, frecuencia de las variantes velares de /a/; en cambio el segundo formante de /o/ está a 800 Hz, casi la frecuencia de la /u/.

Los fonemas /i/, /u/ son realizados fonéticamente como abiertos o cerrados, según condiciones fijas debidas a los sonidos vecinos o su posición en la sílaba. Ninguno de los dos matices pasa de ser una simple variación combinatoria de los fonemas únicos /i/, /u/.

En realidad para el hablante español los fonemas diferenciados, es decir abierto o cerrado, se confunden en su sentimiento lingüístico en una sola forma. Las cinco vocales presentan además una variación fonética especial en posición débil, condicionada por el acento de la palabra. Esta variante tampoco tiene valor distintivo alguno, puesto que no sirve para la diferenciación de las palabras.

La situación de los "diptongos" es más compleja. Existen seis decrecientes: /a₁i/, /a₁u/, /e₁i/, /e₁u/, /o₁i/, /o₁u/, y ocho crecientes: /ja/, /je/, /jo/, /ju/, /wa/, /we/, /wi/, /wo/. Por ejemplo en aire, causa, seis, reuma, sois, hacia, tierra, adiós, viuda, cuarto, cuerda, cuita, menguó.

Los diptongos se dividen en monofonemáticos y combinaciones de los dos fonemas diferentes. De la primera regla práctica que dió Trubetzkoy para determinar la naturaleza monofonemática de dos sonidos sucesivos, según la cual tales sonidos, en ciertas circunstancias, no forman parte de dos sílabas distintas, solo los ocho diptongos crecientes y los diptongos /au/, /eu/, /ou/ son monofonemáticos. Los tres diptongos restantes quedan desprovistos de valor monofonemático ya que cuando van seguidos de vocal en la cadena hablada, sus elementos pueden repartirse entre sílabas sucesivas.

Para concluir podemos decir que los diptongos no son realización fonética de fonemas particulares únicos, sino de dos fonemas contiguos diferentes. Las semivocales y semiconsonantes son sólo variantes combinatorias de los fonemas vocales /i/, /u/. (12)

Estas mismas consideraciones valen para los triptongos, o combinaciones de tres fonemas dentro de una misma sílaba, como en sentenciáis, que el triptongo /jái/ es la combinación de los fonemas /i/, /a/, /i/.

Los "sonidos consonánticos" españoles se dividen en:

Oclusivas: /p/, /b/; /t/; /t/, /d/; /k/, /g/.

Africadas: /ç/, /j/.

Fricativas: /θ/, /f/; /ø/, /z/, /ð/; /s/, /z/; /s/, /z/, /ɾ/; /y/; /x/, /g/, /w/.

Laterales: /l̄/, /l̄/, /l/, /l̄/.

Vibrantes: /r/, /r̄/.

Nasales: /m/, /m̄/, /n̄/, /n̄/, /n/, /n̄/, /ŋ/.

Fonológicamente, varios de estos sonidos son sólo variantes combinatorias de fonemas. En realidad el sistema fonológico de las consonantes españolas nos ofrece los siguientes fonemas: /p/, /b/, /t/, /d/, /k/, /g/, /ç/, /f/, /ø/, /s/, /x/, /y/, /m/, /n̄/, /n̄/, /l/, /l̄/, /r/, /r̄/.

Los rasgos diferenciales entre fonemas consonánticos son: diferencia líquida/no-líquida, diferencia nasal/oral, diferencia de localización, diferencia interrumpida/continua, diferencia sonora/sorda, y la diferencia floja/tensa. La diferencia líquida /no-líquida depende de sí los fonemas consonánticos tienen características vocálicas o si son sin rasgos vocálicos, respectivamente. La diferencia nasal/oral

depende de si en la realización de los fonemas interviene el accesorio resonador de las fosas nasales o sin esa resonancia nasal. Las diferencias de localización son de cuatro órdenes: labial, dental, alveolo-palatal y velar, que se reducen a la doble oposición densa/difusa, esto es, fonemas con predominio de la cavidad bucal y fonemas con predominio de la cavidad faríngea; y grave/aguda, es decir, fonemas con cavidad indivisa y fonemas con cavidad bucal contraída. La diferencia interrumpida/continua, sirve para oponer las oclusivas a las fricativas y dentro de las líquidas, las vibrantes a las laterales. La diferencia sonora/sorda y la floja/tensa sirven como complemento para la distinción de los fonemas. (13)

Según las propiedades distintivas, los fonemas consonánticos pueden clasificarse en cinco series fundamentales:

Líquidas: /l/, /l/, /r/, /r/.

Nasales: /m/, /n/, /n/.

Sonoras: /b/, /d/, /y/, /g/.

Fricativas: /f/, /o/, /s/, /x/

Oclusivas: /p/, /t/, /c/, /k/

(sordas) } orales } no líquidas

Aunque esto es todo lo que diremos sobre la lingüística y la fonética, esto no quiere decir que se minimice su importancia, especialmente en las áreas de reconocimiento y síntesis de la voz.

1.4 Parámetros de la Voz

Los procesos para el reconocimiento de la voz, se caracterizan principalmente por: la discriminación entre períodos con voz y sin voz, la detección de picos de sílabas y la clasificación de núcleos de sílabas, la asunción de límites de sílabas y la clasificación en grupos consonánticos.

La primera etapa de este proceso se realiza por medio de un sistema de transformación de la señal acústica de la voz en ciertos parámetros de tiempo variable, que son una representación básica de las sensaciones auditivas tales como: altura, tono, aspereza, timbre y duración subjetiva. La variación temporal de estos parámetros encierra la información relevante de la voz. Las demás etapas del proceso de reconocimiento de la voz son: la selección de parámetros dominantes y un proceso de reconocimiento final basado en las reglas de la fonética y de la lingüística.

En nuestro estudio nos ocuparemos únicamente de las leyes básicas que controlan las sensaciones auditivas esenciales; esto es, los parámetros intensidad, tono, aspereza, timbre y duración subjetiva.

La "intensidad" se define como el atributo de la sensación auditiva, según la cual se pueden ordenar los sonidos en una escala que se extiende de suave (piano) a fuerte (forte). Esta se compone, en una forma compleja, de elementos

temporales y cierta parte espectral. Para sonidos estacionarios, la altura total se obtiene como la integral de la altura específica, que se extiende a lo largo de la distribución de la banda crítica de la escala. Este principio es válido también para sonidos no estacionarios. El patrón de banda crítica para altura específica se establece como una función del tiempo.

El "tono" es el atributo de la sensación auditiva, según el cual, los sonidos se pueden ordenar en una escala que se extienda de bajo a alto. El caso más simple de detección de tono es el de detección de un solo tono puro. En este caso, la sensación de tono está ambigüamente relacionada por medio de la escala media con la frecuencia de tono. Por ejemplo, cuando las vocales /u/ e/i/ se producen con una misma frecuencia fundamental de 120 Hz, producen un mismo tono con voz que también se produce por un tono puro de 120 Hz. Pero las dos vocales difieren en el criterio bajo/alto, la vocal /i/ se percibe como más alta que la vocal /u/.

Dentro de un sonido se puede ajustar un solo tono puro, de tal forma que su tono corresponda al de la formante de la vocal. El tipo de tono que se relaciona con componentes espectrales, como tonos y formantes puros, se llama "tono espectral". El otro tipo de tono que no está directamente relacionado con componentes sinusoidales o formantes, sino que es dependiente de un patrón de componentes parciales, se llama "tono virtual". La detección de tono virtual se considera como un procedimiento para el reconocimiento del patrón auditivo.

Tanto el tono espectral como el tono virtual pueden ser deducidos del patrón de altura específica. En nuestro estudio nos ocuparemos fundamentalmente de las técnicas para la detección del tono espectral.

La "aspereza" se define como el atributo de la sensación auditiva, según el cual el sonido puede ordenarse en una escala que se extienda desde puro hasta áspero. La aspereza se produce por la fluctuación de sonidos de cualquier tipo. A razones muy bajas (bajo 20 Hz) y muy alta (sobre 200 Hz) la aspereza es muy pequeña, en cambio a razón de 75 Hz la producción de aspereza es óptima.

La aspereza producida por un sonido de considerable ancho de banda está compuesta por ciertas porciones de banda ancha, que son idénticas a las bandas auditivas críticas. Las variaciones rápidas de la señal de la voz, tales como: la estructura periódica de tiempo de los segmentos con voz de tono bajo, la envolvente temporal de ciertos fonemas como la /r/, y el comienzo de consonantes oclusivas; se representan muy bien con un modelo de parámetros relacionados con la aspereza.

El "timbre" se define como el atributo de la sensación auditiva, según la cual dos sonidos de apariencia similar y con la misma altura y tono son distinguibles. La "agudeza" es un componente importante del timbre, por lo que ha sido estudiado aisladamente. La agudeza es un atributo auditivo relacionado con la densidad, compactez y brillantez.

La "duración subjetiva" se diferencia de la duración física no sólo en forma cualitativa, sino también en forma cuantitativa. Por ejemplo, las pausas entre impulsos de sonidos sucesivos se perciben como mucho más cortos que simples impulsos de sonidos, aunque sean iguales en su duración física. Además, cuando la duración física es menor a 800 ms., la duración subjetiva de tonos puros depende de la frecuencia de tono. Un impulso de tono de 3,2 KHz de 25 ms. de duración física produce una duración subjetiva correspondiente a la producida por una pausa de duración física de 100 ms.

C A P Í T U L O S E G U N D O

Métodos Para La Detección del Tono de La Voz

- 2.1 Métodos de detección
- 2.2 Técnicas de Procesamiento en Paralelo para Detección.
- 2.3 Detección del tono por reducción de información.
- 2.4 Detección del tono utilizando la función promedio de la magnitud de la diferencia.
- 2.5 Métodos autocorrelativos de detección del tono
- 2.6 Evaluación de los métodos.

C A P I T U L O I I

Metodos Para La Deteccion Del Tono De La Voz

2.1 Método para la Detección de Tono

La detección de tono o la estimación de la frecuencia fundamental es uno de los problemas más importantes dentro del procesamiento de la voz. Los detectores de tono son un componente esencial dentro de varios sistemas procesadores de voz. Aparte de proveer información acerca de la naturaleza de la fuente de excitación para producción de voz, el contorno de tono de una expresión sirve para reconocimiento del locutor, para sistemas de verificación y para casi todos los sistemas para análisis y síntesis de la voz (decodificadores de voz).

Por la gran importancia de la detección de tono, se han propuesto una gran variedad de algoritmos para su detección. Todos los sistemas propuestos tienen sus limitaciones, y se puede afirmar que ninguno de los esquemas para detección de tono actuales, puede brindar resultados satisfactorios dentro de un ancho rango de locutores, aplicaciones y medios de operación (10).

La medición exacta del período de tono de la señal de la voz es frecuentemente muy difícil por muchas razones. Una razón es que la forma de onda de excitación glotal no es

un tren de pulsos periódicos perfecto. Si bien encontrar el período de una onda periódica es sencillo, la medición del período de una onda de voz que varía tanto en período como en la estructura de la forma de onda dentro de un período, puede ser un poco difícil. La segunda dificultad en la medición del período de tono es la interacción entre el canal bucal y la excitación glotal. En ciertos instantes las formantes del canal bucal pueden alterar significativamente la estructura de la forma de onda glotal, de tal forma que el período de tono actual sea difícil de detectar. Generalmente estas interacciones son más perjudiciales para la detección de tono, durante movimientos rápidos de las articulaciones cuando las formantes están cambiando rápidamente. El tercer problema en la medición del tono es el de la dificultad de definir exactamente el inicio y terminación de cada período de tono durante segmentos con voz. La determinación de las localizaciones del comienzo y final del período de tono son un tanto arbitrarias. Por ejemplo, basados en la forma de onda acústica, algunos criterios para definir el comienzo y final del período incluye el valor máximo durante el período, el cruce de cero previo al máximo, etc. El único requisito para estas mediciones es que los valores sean consistentes período a período, de tal manera que se pueda definir exactamente las ubicaciones del comienzo y el final de cada período de tono. Las discrepancias en las mediciones del período de tono se deben no sólo a la cuasi-periodicidad de la onda de la voz, sino también a que las mediciones de los picos depen-

den de la estructura formante durante el período de tono, mientras que los cruces de cero dependen de las formantes, el ruido y de cualquier nivel de corriente continua dentro de la forma de onda acústica. Una cuarta dificultad en la detección de tono es la distinción entre segmentos sin voz y segmentos con voz de bajo nivel. En muchos casos las transiciones entre segmentos sin voz y segmentos con voz de bajo nivel son muy artificiales, debido a la dificultad de señalarlas con precisión.

Además de las dificultades para la medición del período de tono enumeradas, ocurren complicaciones adicionales cuando se estudia la extracción de tono de una voz que se ha transmitido a través de un sistema telefónico. Los efectos que produce un sistema telefónico sobre la voz incluyen filtrado lineal, procesamiento no lineal y la introducción de ruido en la señal de la voz. En lo referente al filtrado lineal, el sistema telefónico actúa como un filtro pasabanda de frecuencia baja de corte de aproximadamente 200 Hz y de frecuencia alta de corte de aproximadamente 3200 Hz, que puede atenuar significativamente la frecuencia fundamental de tono y varias de las armónicas altas de tono. Esto trae como consecuencia una mayor dificultad en la detección de la periodicidad. La contribución no lineal del sistema telefónico a la señal de la voz, dependiendo del sistema de transmisión, puede incluir distorsión de fase, modulación de amplitud de la señal de la voz, superposición entre dos o más mensajes y recortado o distorsión de sonidos de nivel extremadamente alto. Debido a todos estos

efectos la línea telefónica hace mucho más difícil la detección del período de tono.

Como resultado de las numerosas dificultades en la medición del tono; se han desarrollado una gran variedad de métodos sofisticados para la detección de tono. Básicamente un detector de tono es un aparato que realiza una decisión con voz/sin voz y provee de medidas del período de tono durante períodos con voz. Sin embargo, algunos algoritmos de detección de tono sólo determinan el período durante los segmentos con voz y dejan la decisión con voz/sin voz a otras técnicas.

Los algoritmos para la detección de tono pueden ser divididos aproximadamente en tres grupos, un grupo que utiliza primordialmente las propiedades del dominio del tiempo para señales de la voz, un grupo que utiliza principalmente las propiedades del dominio de la frecuencia de la señal de la voz, y un grupo que utiliza tanto las propiedades en el dominio del tiempo como las propiedades en el dominio de la frecuencia de la señal de la voz.

Los detectores de tono en el dominio del tiempo operan directamente en la forma de onda de la voz para la estimación del período de tono. Para estos detectores de tono las mediciones más comunes que se hacen son las de picos y valles, cruce de cero y las mediciones de autocorrelación. La suposición básica que se hace en todos estos casos es que si una señal cuasiperiódica ha sido adecuadamente procesada para minimizar los efectos de la estructura formante, entonces simplemente las mediciones en el dominio del tiempo

darán buenas estimaciones del período.

Los detectores de tono en el dominio de la frecuencia usan la propiedad de que si la señal es periódica en el dominio del tiempo, entonces el espectro de frecuencia de la señal consistirá de una serie de impulsos a la frecuencia fundamental y sus armónicas. Por lo tanto, se pueden realizar mediciones sencillas en el espectro de frecuencia de la señal o en una versión transformada no linealmente del espectro, para estimar el período de la señal.

Los detectores de tono de la clase híbrida incorporan rasgos tanto del dominio del tiempo como del dominio de la frecuencia para la detección de tono. Por ejemplo, un detector de tono híbrido puede usar las técnicas del dominio de la frecuencia para proveer una forma de onda espectralmente aplanada y luego utilizar las mediciones autorrelativas para la estimación del período de tono (15).

Los principales métodos para detección de tono propuestos son cuatro detectores de tono en el dominio del tiempo, dos en el dominio de la frecuencia y dos detectores híbridos. Las técnicas en el dominio del tiempo que se han propuesto son:

- 1) Técnicas de Procesamiento en Paralelo (PPROC) de L. Rabiner y B. Gold (16).
- 2) Método de Reducción de Información (DARD) de Neil J. Miller (17).

- 3) Función Promedio de la Magnitud de la Diferencia (AMDF) de Ross, Shaffer, Cohen y Manley (18).

Las técnicas para detección de tono en el dominio de la frecuencia que se han propuesto son:

- 4) Método del Cepstrum (CEP) de A. Michael Noll. (19).

- 5) Método de las Armónicas Corruptas del Espectro de Potencia (GCD) de Sreenivas y Rao (20).

Las técnicas híbridas para la detección de tono que se han propuesto son:

- 6) Métodos de autocorrelación usando Recortado (AUTO) de Man Mohan Sondhi (21)

- 7) Técnica de Filtrado Inverso Simplificado (SIFT de John D. Markel (22).

De entre estos métodos se analizarán cuatro y se desarrollarán los programas para los dos algoritmos de tono más efectivos.

2.1 Técnicas de Procesamiento en Paralelo para Detección de Tono:

En este Capítulo se discutirá el método de procesamiento en paralelo, propuesto por Gold y modificado por Gold y Rabiner. Las razones para discutir este método en particular son; ha sido utilizado con éxito en una gran variedad de aplicaciones, se basa totalmente en procesamiento en el dominio del tiempo, se puede implementar en un computador de pro-

pósitos generales para que opere rápidamente. Este ilustra el principio básico del procesamiento en paralelo en el procesamiento de la voz (23).

Los principios básicos de este algoritmo son los siguientes:

- 1) Se procesa la señal de la voz para crear un número de trenes de impulsos que contengan la periodicidad de la señal original y que descarte los rasgos que sean irrelevantes para el proceso de detección de tono.
- 2) Este proceso permite el uso de detectores de tono simples para la estimación del período de cada tren de impulsos.
- 3) Las estimaciones de varios de estos detectores se combinan lógicamente para encontrar el período de la onda de voz.

Este algoritmo computacional para la estimación del período de tono de la voz en el dominio del tiempo y dos modificaciones recientes del algoritmo, se discutirán detalladamente.

Estimaciones en Paralelo

En 1960 la idea de diseñar un estimador o detector del período de tono, basados en las técnicas de procesamiento en paralelo, fue desarrollada por B. Gold en el laboratorio del MIT en Massachusetts. La idea básica en el procesamiento

en paralelo era que un mejoramiento en la exactitud, podía obtenerse por medio de la combinación adecuada de las salidas de más de un estimador de período de tono.

El procesamiento en paralelo parecía apropiado, debido a su similitud con el proceso humano asociado con la estimación del período de tono, es decir la inspección visual de la onda de la voz.

Minsky en una publicación interna del MIT reportó experimentos en procesamiento paralelo para la estimación del período de tono. Gill desarrollo mediciones del período de tono basado en la unión y correlación de tres sistemas en paralelo para formar un detector de tono. Estos dos esquemas seguían un procesamiento de las mediciones en paralelo del período de tono un tanto rudimentario. Por lo contrario, los esquemas que se van a describir tienden a simplificar los detectores en paralelo, pero utilizan un algoritmo relativamente sofisticado para procesar las salidas de los estimadores de tono (24).

El primer esquema de procesamiento en paralelo desarrollado por Gold consta de un programa de computador en el que se usa tres estimadores de período de tono en paralelo. Este esquema procesa la banda total de la onda de la voz y utiliza las ubicaciones de los picos de la señal y pruebas de regularidad para la estimación de los períodos de tono. Un algoritmo combinacional relativamente elemental

se usa para determinar la estimación final del período de tono.

Luego de este primer intento, se desarrolló en el Laboratorio Lincoln un estimador de período de tono, basado en la combinación de seis estimadores de tono simples y en una lógica más adecuada. Este esquema de procesamiento en paralelo difiere sustancialmente del anterior en casi todos los aspectos. En lugar de la voz de banda completa, sólo se procesa la onda filtrada con un filtro pasa-bajo; las mediciones de los detectores individuales son de pico a pico y de pico a valle; y se utiliza un detector de coincidencias sofisticado como parte del algoritmo de decisión final. Este esquema se simuló en un computador digital y se implementó como parte de la estructura del "hardware" de un aparato. Este aparato ha sido utilizado satisfactoriamente en varios sistemas para codificación de la voz, pero generalmente se le ha considerado como muy complejo y costoso para usos generales. Aunque aparentemente, los últimos avances en componentes de circuitos integrados están eliminando este problema.

Recientemente se ha visto la necesidad de esquemas para la estimación del período de tono, más rápidos y eficientes, para utilizarse en computadores digitales en la investigación de los contornos del tono de la voz. Para poder procesar una gran cantidad de información en un computador digital, se necesita operar los algoritmos a tiempo real o cerca

a éste. Se han inventado poderosas formas de medición del período de tono basadas en el análisis del espectro de la voz, pero para la simulación computacional de estas técnicas se necesita 100 o más veces del tiempo real. Debido a este problema, la búsqueda de un programa computable más rápido y eficiente, llevó a las dos versiones modificadas del algoritmo de Gold que serán analizadas más adelante.

Algoritmo Original de Tono

El algoritmo original propuesto por Gold y Rabiner se muestra en la Figura 2.1. Este algoritmo se puede dividir convenientemente en cuatro etapas:

- 1) Filtrado de la señal de la voz.
- 2) Generación de seis funciones de los picos de la señal de la voz filtrada.
- 3) Seis estimadores simples de período de tono idénticos, cada uno operando sobre una de las seis funciones anteriores.
- 4) Computación final del período de tono, basados en el examen de los resultados de cada uno de los estimadores simples de período de tono.

La onda de la voz se muestrea a una velocidad suficiente para dar un tiempo de resolución adecuado. Por ejemplo, un muestreo a 10 KHz permite que se determine el período dentro de un $T = 10^{-4}$ segundos. (10).

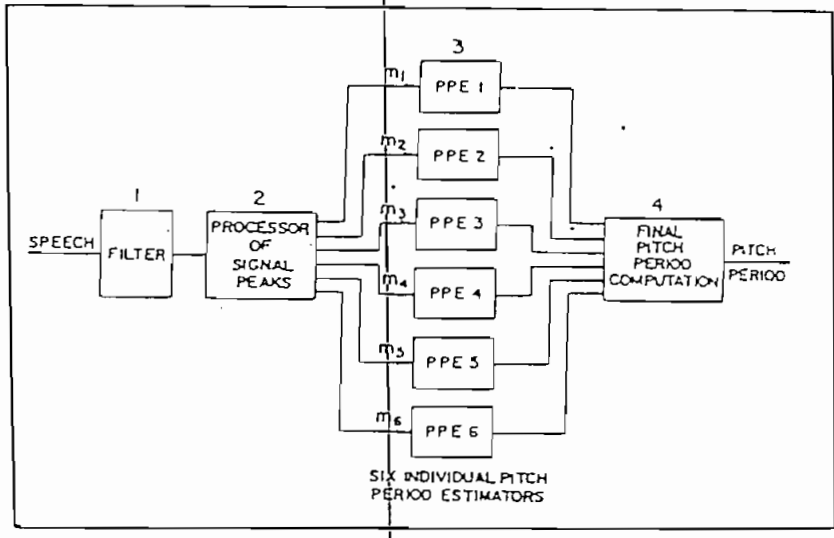


Fig. 2.1 Diagrama de bloques de un detector de tono por procesamiento en paralelo en el dominio del tiempo.

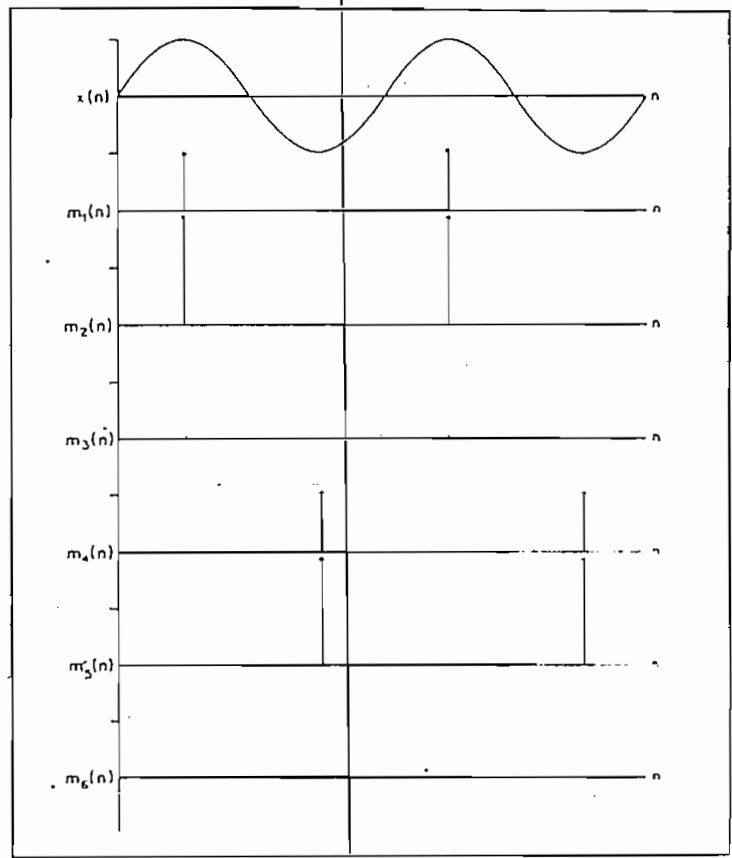


Fig. 2.2 Trenes de impulsos generados para los picos y valles de una señal de entrada sinusoidal $x(n)$.

El propósito principal del filtro pasa-bajo es el de seleccionar aproximadamente la región del primer formante. Ninguna otra información es necesaria, pues los picos generados por las formantes mayores, sólo tienden a reducir la exactitud de la detección de tono. Si la voz de entrada contiene la frecuencia fundamental, se puede utilizar un filtro pasa-bajo con 36 dB por octava y corte sobre los 600 Hz, aunque un punto de corte preciso y un diseño específico del filtro no son críticos. Por supuesto que se debe tener cuidado de eliminar el ruido de 60 Hz y 120 Hz; por ejemplo, con un filtro pasa-alto de cerca de 100 Hz de punto de corte.

En caso de que no esté presente la frecuencia fundamental, porque la voz ha sido transmitida a través de un cable telefónico, se puede utilizar un filtro pasa-banda desde 300 Hz hasta 900 Hz, para que al menos se procesen dos armónicas superiores de la voz.

Para las aplicaciones con frecuencia fundamental presente, donde la más alta frecuencia fundamental a ser procesada sea menor que 600 Hz, puede reducirse el punto de corte del filtro pasa-bajo. Sin embargo, si la frecuencia de corte se reduce aproximadamente a 250 Hz, aparece un cambio notable en las estimaciones del período de tono y esto produce una codificación de voz de calidad áspera. También

en el caso de que no exista presencia de la frecuencia fundamental y de que se use elementos no lineales para la generación de la frecuencia fundamental, se nota la calidad áspera de la codificación de la voz.

La segunda etapa indicada en la Figura 2.1, consiste en la ubicación de los "picos" y "valles" (máximos y mínimos) y en la generación de varios trenes de impulsos a partir de las ubicaciones y amplitudes de los picos y valles de la señal filtrada. Cada tren de impulsos consta de impulsos positivos que ocurren en las ubicaciones ya sea de picos o de valles de la señal filtrada. Los seis casos utilizados por Gold y Rabiner de acuerdo a las Figuras 2.2 y 2.3 son:

- 1) $m_1(n)$: Un impulso de amplitud igual a la del pico ocurre en la posición de cada pico.
- 2) $m_2(2)$: Un impulso igual a la diferencia entre la amplitud del pico y la amplitud del valle precedente ocurre en la posición de cada pico.
- 3) $m_3(n)$: Un impulso igual a la diferencia entre la amplitud del pico y la amplitud del pico precedente ocurre en la posición de cada pico. Si esta diferencia es negativa el impulso se iguala a cero.
- 4) $m_4(n)$: Un impulso igual a la amplitud del valle cambiada de signo ocurre en la posición de cada valle.

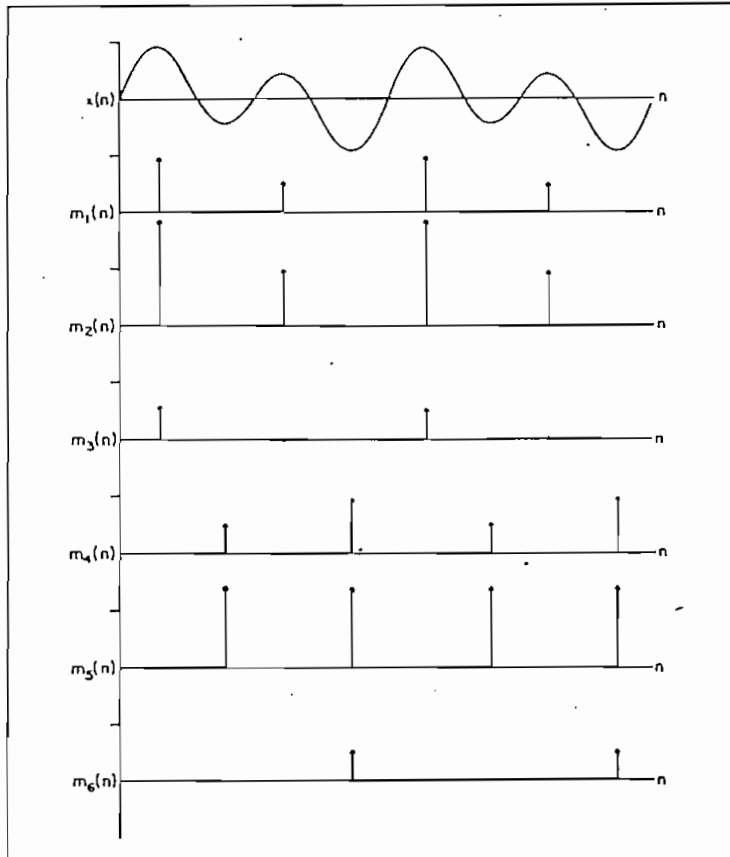


Fig. 2.3 Trenes de impulsos generados para los picos y valles de una señal de entrada $x(n)$.

- 5) $m_5(n)$: Un impulso igual a la amplitud del valle cambiada de signo más la amplitud del pico precedente ocurre en la posición de cada valle.
- 6) $m_6(n)$: Un impulso igual a la amplitud del valle cambiada de signo más la amplitud del valle precedente ocurre en la posición de cada valle. Si esta diferencia es negativa el impulso se iguala a cero.

Las figuras 2.2 y 2.3 muestran una sinusoidal pura y una fundamental débil con una segunda armónica fuerte, junto con sus respectivos trenes de impulsos generados de acuerdo a las definiciones anteriores. Los pulsos de amplitudes m_1 , m_2 y m_3 se generan en cada pico positivo de la señal filtrada, mientras que los pulsos de amplitudes m_4 , m_5 y m_6 se generan en cada pico negativo (valle) de la señal filtrada. Las mediciones m_1 y m_4 son medidas de un pico positivo o negativo, mientras que las mediciones m_2 , m_3 , m_5 y m_6 dependen de picos previos de la señal. Las mediciones m_2 y m_5 son medidas pico a valle y valle a pico, mientras que m_3 y m_6 son medidas pico a pico previo y valle a valle previo. Se debe observar que todos los trenes de impulsos son positivos, pues no se permiten valores negativos. Cuando se presenta un valor negativo, este valor se iguala a cero.

La elección de este grupo particular de mediciones se basa en dos casos extremos como los indicadas en las Figuras 2.2 y 2.3. Para el caso de la Figura 2.2 donde sólo la frecuencia fundamental está presente, las medidas m_3 y m_6 son poco útiles, pero en cambio las medidas m_1 , m_2 , m_4 y m_5 proveen de muy buenas indicaciones del período. Para el caso de la Figura 2.3 donde están presentes una frecuencia fundamental débil y una segunda armónica fuerte, las medidas m_3 y m_6 son probablemente correctas, mientras que las medidas m_1 , m_2 , m_4 y m_5 probablemente darán indicaciones incorrec-

tas de la mitad del período. Aunque las medidas m_1 , m_2 , m_4 y m_5 pueden fallar, ésto puede corregirse por medio de la computación final.

El propósito para la generación de estos trenes de impulsos es el de simplificar la estimación del período en una base de tiempo corta. En la tercera etapa de la Figura 2.1 se aplican los seis trenes de pulsos a los seis detectores individuales de tono. La manera de operación de los estimadores simples de período de tono se puede apreciar en la Figura 2.4. Cada tren de impulsos se procesa con un sistema no lineal variable en el tiempo llamado "circuito de ventana exponencial para la detección de picos". (16).

En esencia cada estimador simple de período de tono es un circuito detector de picos. Cuando a la entrada del estimador se detecta un pulso de amplitud suficiente, la salida se pone al valor del pulso detectado durante un intervalo de blanqueo, durante el cual ningún pulso puede ser detectado. Al final de este intervalo de blanqueo, la salida comienza a caer exponencialmente. Cuando un pulso exceda al nivel de la caída exponencial de salida, se repite el proceso. Los valores de los intervalos de caída y blanqueo dependen de las más recientes estimaciones del período de tono.

El tiempo de blanqueo y la constante exponencial de cada detector son funciones de las estimaciones suavizadas del período de tono, P_{av} , del detector respectivo. El valor

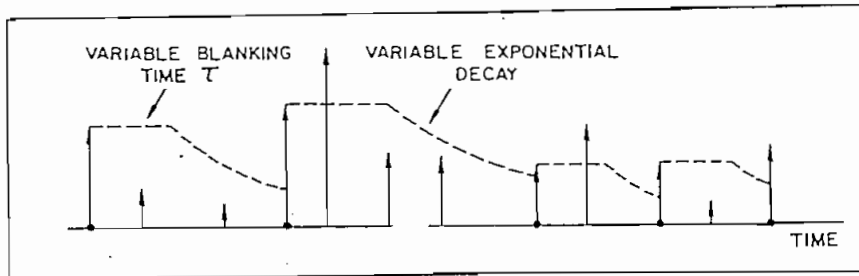


Fig. 2.4 Forma de operación de cada estimador del período de tono.

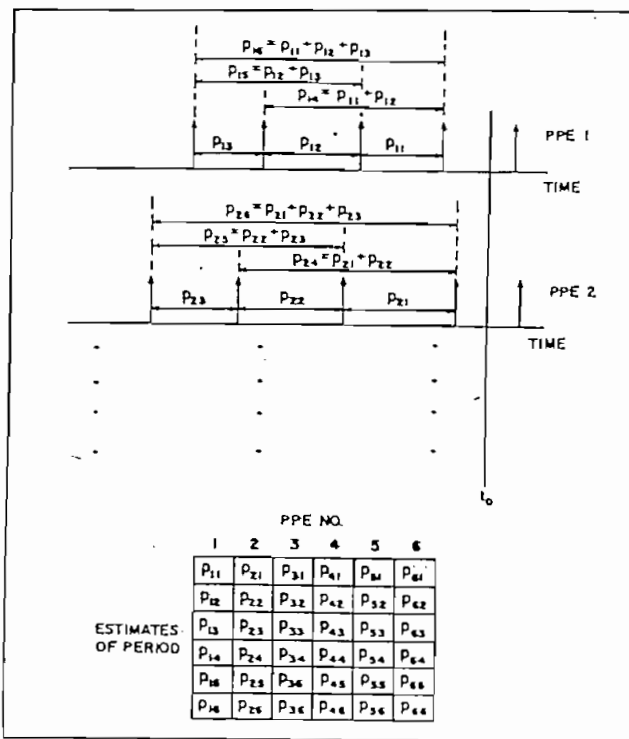


Fig. 2.5 Estimación final del período de tono en base a la matriz de estimaciones formada con los resultados de los seis estimadores del período de tono.

P_{av} se determina de la expresión:

$$P_{av}(n) = (P_{av}(n-1) + P_{new}) / 2$$

donde P_{new} es la estimación más reciente del período de tono, $P_{av}(n)$ es la estimación corriente suavizada del período de tono y $P_{av}(n-1)$ es la estimación suavizada previa del período de tono. Cada vez que se detecta un nuevo pico, el valor de P_{av} se recalcula de acuerdo al nuevo valor.

Para prevenir valores extremos de la constante de tiempo de blanqueo o de la constante de tiempo de amortiguación, el valor P_{av} se limita entre 4 mseg y 10 mseg. Dentro de estos límites, la dependencia de la constante de tiempo de blanqueo " τ " y de la constante de tiempo de amortiguación " β " con el valor P_{av} está dada por las expresiones

$$\tau = 0.4 P_{av}$$

$$\beta = P_{av} / 0.695.$$

El resultado del proceso descrito es la suavización del tren de impulsos, produciéndose una secuencia de pulsos cuasi-periódicos como se puede observar en la Figura 2.4. La longitud de cada uno de estos pulsos es una estimación del período de tono. El período de tono se estima periódicamente midiendo las longitudes de los pulsos a lo largo del intervalo de muestreo.

La técnica anterior se aplica a cada uno de los seis trenes de impulsos, por lo que se obtienen seis estimaciones del período de tono. Estas seis estimaciones se combinan con dos de las estimaciones más recientes para cada uno de los seis detectores de tono. Estas estimaciones se comparan y el valor que más veces haya ocurrido se declara

como el período de tono en ese tiempo. Este procedimiento correspondiente a la cuarta etapa de la Figura 2.1, proporciona muy buenas estimaciones del período de segmentos con voz. Para segmentos sin voz existe inconsistencia entre las estimaciones. Cuando se detecta esta inconsistencia, el segmento se clasifica como sin voz. El proceso completo se repite periódicamente para producir una estimación del período de tono y una clasificación con voz/sin voz en función del tiempo.

Esta computación final del período de tono se realiza por medio de un computador de propósitos especiales, con una memoria y un control del "hardware" para dirigir toda la operación. La estimación del período de tono en cualquier tiempo " t_0 " se realiza de acuerdo al proceso correlacionado con la Figura 2.5 que se verá a continuación.

Primero se forma una matriz de 6X6 con las estimaciones del período de tono. Las columnas de la matriz representan los detectores individualmente y las filas son las estimaciones del período respectivamente. Las tres primeras filas son las tres estimaciones del período de tono más recientes. La cuarta fila es la suma de la primera y segunda fila, la quinta fila es la suma de la segunda y tercera fila, y la sexta fila es la suma de las primeras tres filas. La técnica para la formación de la matriz se ilustra en la Figura 2.5. La razón por la que se forman estas tres últimas

fílas de la matriz es que a veces los detectores individuales pueden indicar la segunda o la tercera armónica de la señal en lugar de la fundamental; la información de las tres últimas fílas en este caso será más correcta que las tres estimaciones del período de tono más recientes.

Como segundo paso, se compara cada una de las entradas o valores de la primera fila con los otros 35 valores de la matriz y se va contando el número de coincidencias. El valor de P_{i1} ($i = 1, 2, 3, 4, 5, 6$) con el mayor número de coincidencias es la estimación final del período de tono.

Para determinar si dos estimaciones del período de tono son coincidentes, parece más apropiado observar su razón que su diferencia. Sin embargo, el cálculo de la razón puede ser muy aproximado para evitar la división computacional. Debido a que dentro de varias partes de la voz existen variaciones apreciables entre mediciones sucesivas del período de tono, es de gran utilidad el incluir valores de umbral para la definición de coincidencia. También se debe tratar de seleccionar, para cada computación completa de un período de tono, el valor de umbral que de la respuesta más consistente.

Sólo la más reciente estimación del período de tono de cada detector puede ser candidato para período final de tono. Este candidato es una de las seis elecciones posibles para período de tono definitivo. Para determinar el resulta-

do, cada candidato se compara numéricamente con las 35 estimaciones de período de tono restantes,

Hay que repetir el procedimiento en forma total para los cinco candidatos restantes. Al final de todas estas operaciones la estimación final del período de tono que tenga el mayor número de coincidencias será el período de tono resultante. Para realizar esta operación computable de todo el algoritmo, se necesitan un total de 6×35 mediciones de coincidencia.

Es interesante ver que el algoritmo de procesamiento en paralelo para la detección de tono descrito, puede operar efectivamente hasta con un nivel de ruido acústico muy alto.

Aunque la descripción anterior del algoritmo parece muy compleja, en realidad este esquema para la detección de tono puede ser implementado eficientemente ya sea en un dispositivo de "hardware" de propósitos especiales o en un simple computador de propósitos generales. Por lo que una operación cercana al tiempo real, dentro de un factor de 2 veces el tiempo real, se hace posible en los computadores que existen actualmente.

En las pruebas que se han realizada con este algoritmo para la detección de tono, se han utilizado como señal de entrada una señal de voz sintética muestreada. La ventaja de utilizar voz sintética es que los valores verdaderos de

período de tono se conocen exactamente, ya que se generan artificialmente, con lo que se puede obtener una medida de la exactitud del algoritmo. La desventaja de la voz sintética es que se genera de acuerdo a un modelo simple, por lo que durante la ejecución del algoritmo no aparecerán las propiedades inusuales de la voz natural. (25).

Modificaciones al Algoritmo de Tono

El tiempo aproximado de ejecución del algoritmo original de tono es cerca de 50 veces el tiempo real en un computador moderadamente rápido como el TX-2 del laboratorio Lincoln. Debido a la necesidad de un aumento en la velocidad del programa, se realizaron dos modificaciones del algoritmo original.

La primera modificación se basa en la necesidad de una mayor velocidad del programa y en una limitación del rango de frecuencia fundamental a aproximadamente 220 Hz, esto es para una implementación dirigida al análisis de la voz masculina. Se realizaron fundamentalmente tres modificaciones de las condiciones del algoritmo original.

Primero se cambió el filtro pasa-bajo y el muestreo previo de la señal por un filtro análogo. Se usó un filtro de 36 dB por octava de atenuación, con un punto de corte a baja frecuencia de 70 Hz y un punto de corte a alta frecuencia de 600 Hz.

Se utilizó un tiempo de blanqueo fijo y una constante de amortiguación independiente de la estimación suavizada del período de tono. Los valores utilizados fueron para el tiempo de blanqueo 3 mseg y para el tiempo de la constante de amortiguación 5.2 mseg. Estos valores se escogieron para obtener resultados adecuados a una frecuencia fundamental promedio de 125 Hz.

Por último, se utilizó sólo un grupo de medidas coincidentes, basadas en las diferencias entre las estimaciones del período de tono y no en sus razones. Debido a que el rango permitido de período de tono es menor al utilizado en el algoritmo original, se observó experimentalmente que una sólo medida de coincidencia daba buenos resultados a lo largo del intervalo de interés. La medida de coincidencia de ± 300 mseg que se utilizó en esta modificación, permitió la consideración de un rango de períodos de tono de 5-14 mseg correspondientes a valores de " F_0 " (frecuencia fundamental de tono) desde 70 Hz hasta 200 Hz. Al utilizar una sólida medida de coincidencia se elimina el procedimiento de búsqueda complicado de la estimación corriente del período de tono.

Estas modificaciones se programaron en el algoritmo original y se realizaron varias pruebas para voces sintéticas y reales. El programa se escribió en FORTRAN IV para un computador GE 635 y se observó que su operación era mucho más

rápida. El tiempo de ejecución del programa fué cerca de 1.3 a 1.5 veces el tiempo real. Estos rasgos y la exactitud en la estimación del período de tono son los puntos fuertes de éste algoritmo, que lo hacen atractivo para la detección de tono.

Para aquellos proyectos en que la extracción de la " F_0 " de contorno es de importancia, la velocidad computacional de este algoritmo modificado es mucho más rápida comparada con técnicas para la detección de tono como las de cepstrum y autocorrelación que se verán más adelante.

Unos ejemplos de la estimación de frecuencia fundamental a partir de la primera modificación del algoritmo se indican en las Figuras 2.6 y 2.7. En la Figura 2.6 se muestra un ejemplo de estimación de frecuencia fundamental a partir de voz sintética. La razón para utilizar voz sintética es que los valores de " F_0 " se conocen de antemano, con lo que se puede medir la exactitud del algoritmo. Se puede apreciar que las estimaciones del período de tono están dentro de ± 2 Hz los valores originales, aunque a veces la diferencia es mayor. También se puede observar que al comienzo y ocasionalmente al final de un segmento con voz, no existe una estimación del período de tono debido a que el número de coincidencias es pequeño. La indicación de un segmento sin voz aparece hasta que un segmento con voz tenga suficiente duración, para que el número de coincidencias exceda el valor umbral que indique un segmento con voz. Esto ocurre usualmente después de 20-30 mseg de haber comenzado el segmento con voz.

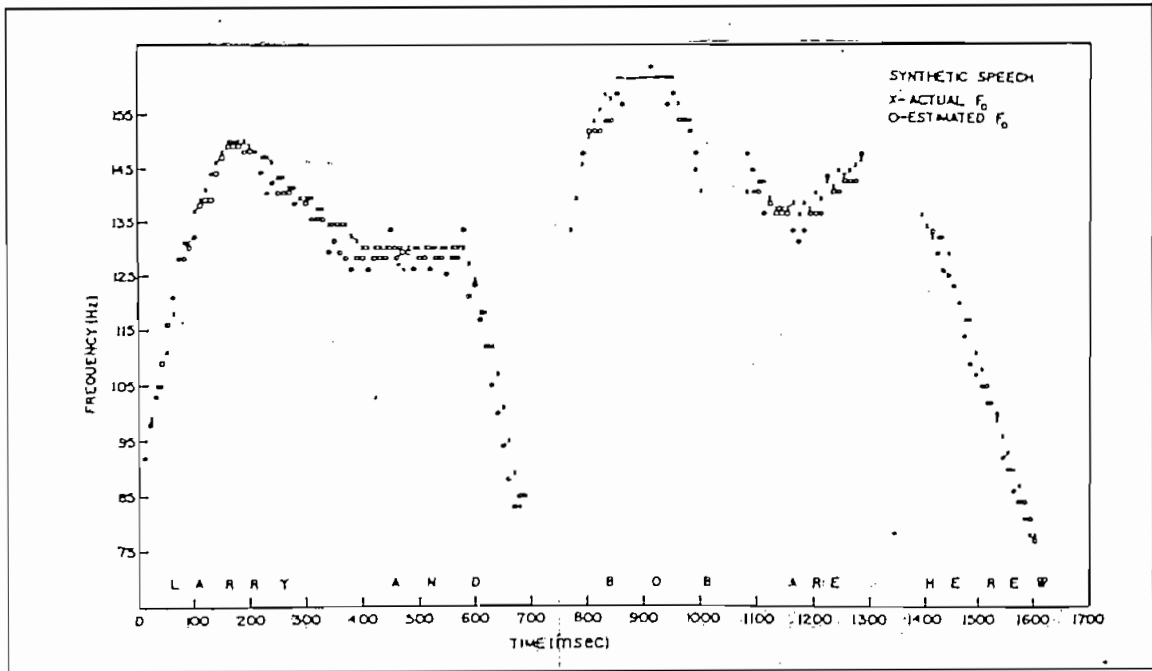


Fig. 2.6 Comparación entre las estimaciones de Frecuencia Fundamental obtenidas utilizando el algoritmo modificado y los valores verdaderos de voz sintética.

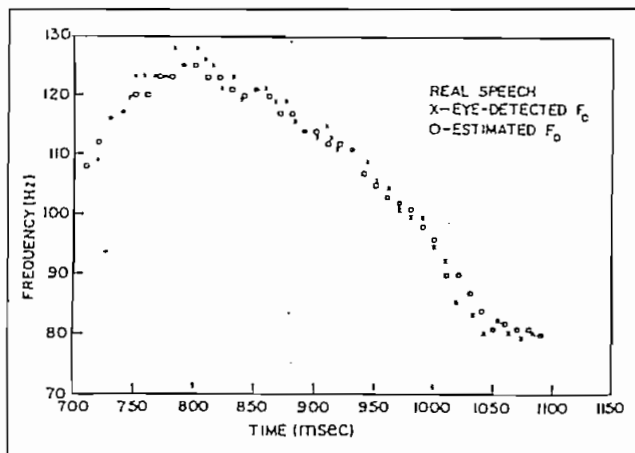


Fig. 2.7 Comparación entre las estimaciones de Frecuencia fundamental obtenidas usando el algoritmo modificado y valores detectados por el ojo humano para la voz natural.

En la Figura 2.7 se muestra una comparación entre la estimación de tono a partir del algoritmo modificado y la detección visual de una voz real. Se puede observar claramente que las detecciones visuales están aproximadamente de acuerdo con las estimaciones de tono a lo largo del intervalo. No se puede realizar una estimación cuantitativa en base a esta figura debido a que la exactitud en la detección de períodos de tono en forma visual no es muy buena. Por ejemplo, en el laboratorio Lincoln para una prueba con el codificador de voz se observó que la extracción del período de tono por medio del algoritmo original era favorable en proporción de 2 a 1 con respecto a la detección visual.

La segunda modificación al algoritmo original de tono se basa únicamente en la necesidad de medir frecuencias fundamentales bajo los 300 Hz. Para ello se realiza una simplificación del algoritmo original de acuerdo a los siguientes puntos:

- 1) El filtrado de la señal se realiza de una manera más selectiva.
- 2) El número de estimadores de período de tono se reduce a cuatro.
- 3) La constante de tiempo de blanqueo y la constante de atenuación se ajustan como en la primera modificación.

- 4) Se toman dos en vez de tres estimaciones de tono de cada uno de los estimadores de período de tono.
- 5) Se elimina el sistema de la Figura 2.6, pues sólo una de las columnas indicadas es necesaria.

Los detectores de picos de la Figura 2.8 solamente localizan y miden las amplitudes de los picos negativos y positivos de las dos ondas incidentes. Los períodos de tono se estiman a partir de los cuatro trenes de pulsos, por medio de cuatro detectores de tono idénticos que operan en la misma forma que los del algoritmo original de la Figura 2.1. La única diferencia es que se les hace trabajar con un tiempo de blanqueo fijo de 2 mseg y una constante de amortiguamiento ajustada para que baje a la mitad de su valor inicial al cabo de 5 mseg. También, cada detector se fija a un valor de aproximadamente 20 dB bajo la señal máxima de amortiguamiento permisible si el tiempo de amortiguamiento excede los 16 mseg.

La memoria de almacenamiento básica para la computación final, consiste de 12 estimaciones del período de tono, tres por cada detector, como se ilustra en la Figura 2.9. Por lo que para esta modificación del algoritmo existen solo cuatro candidatos a período de tono. La selección se realiza de nuevo en base al candidato más popular, o sea

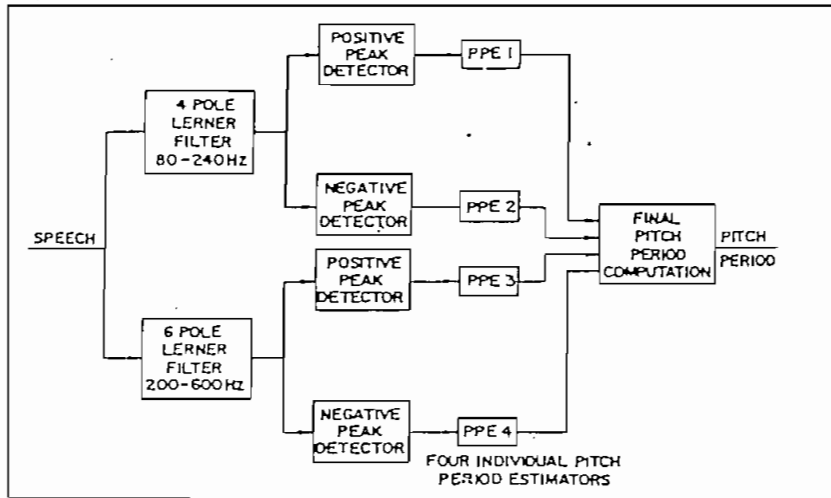


Fig. 2.8 Diagrama de bloques para la segunda modificación del algoritmo de estimación del período de tono por procesamiento en paralelo.

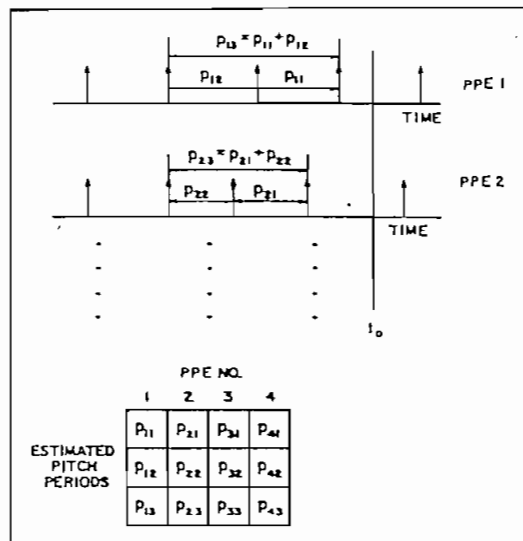


Fig. 2.9 Estimación Final del período de tono para la segunda modificación del algoritmo.

el que tiene el mayor número de coincidencias; aunque ahora ya no se le resta a éste número el valor identificativo de la columna, sino que la coincidencia se define por medio de la desigualdad.

$$|p_c - p_i| < \frac{1}{8} (p_i)$$

Uno de los rasgos principales de este algoritmo es la combinación de las salidas de los estimadores de período de tono para lograr resultados de alta precisión. El algoritmo original funciona bien inclusive cuando la voz se genera de un medio ruidoso. También, da buenos resultados en ausencia de la frecuencia fundamental.

El tiempo de ejecución para el programa de la primera modificación del algoritmo es mucho menor que el de otras técnicas para estimación del período de tono. Esta modificación es tan efectiva como el algoritmo original y funciona bien en presencia de una frecuencia fundamental menor de 220 Hz.

La segunda modificación del algoritmo original lleva a una realización del "hardware" mucho más simple que la del algoritmo original y provee de un funcionamiento efectivo en presencia de una frecuencia fundamental menor a 300 Hz.

Para concluir, los detalles particulares de estos métodos no son tan importantes como los principios básicos que se utilizan. Primero se debe notar que la señal de la

voz se procesa para obtener un grupo de trenes de impulsos que retengan sólo los rasgos esenciales de periodicidad. - Debido a esta simplificación de la estructura de la señal, basta con un estimador de tono muy simple para realizar buenas estimaciones del período de tono. Finalmente, la combinación de varias estimaciones del tono aumenta la precisión en la estimación final. Por lo tanto, la simplificación en el procesamiento de la señal implica el incremento en la complejidad de la lógica para la estimación del tono de la voz.

2.3. Detección de Tono por Reducción de Información

En este capítulo se verá un algoritmo para determinar la frecuencia fundamental de una señal de la voz muestreada por medio de la segmentación de la señal en períodos de tono. La segmentación se realiza por la identificación de las muestras de la onda correspondientes al inicio de cada período de tono.

La segmentación se realiza en tres fases. Primero se utilizan mediciones de cruce de cero y de energía, para construir la estructura informativa a partir de los muestreos para período de tono. Luego, se reduce el número de candidatos dentro de la estructura utilizando segmentación silábica, estimaciones burdas del período de tono y funciones de discriminación. Finalmente, los candidatos para período de tono restantes se corrigen para compensar los errores introducidos durante el proceso de reducción de in-

formación.

Este proceso para la detección de tono tiene las siguientes características: computa los valores que identifican el inicio de cada período de tono; permite el análisis de voces masculinas y femeninas por medio de la detección de frecuencias de tono en un rango de 50 a 500 Hz; y su procesamiento requiere menos de $20 \cdot N$ operaciones computacionales, donde "N" es el número de muestreos de la señal de la voz.

Este algoritmo se desarrolló para utilizarlo en un sistema de reconocimiento sincrónico de tono de la voz. Sin embargo, la determinación del período de tono es esencial para la comprensión del tiempo de sincronismo de tono, en el estudio de prosódicos y en la comprensión de varios sistemas de banda ancha. (1,2,3).

La mayoría de esquemas para detección de tono de voz presentan deficiencias en su ejecución computacional. Generalmente requieren cientos de computaciones para cada muestreo de la señal. El algoritmo por reducción de información elimina este problema por medio de la siguiente estrategia de la reducción de información.

Construcción de la Estructura Informativa

En la primera parte del algoritmo, a partir de la señal de voz muestreada se construye una estructura informativa. En esta estructura informativa se almacenan las muestras de la señal que se declaren "candidatos" a período de

tono. También se guarda en esta estructura informativa cierta información que se requerirá en las etapas posteriores del proceso.

La señal de la voz se muestrea inicialmente a razón de 10 o 20 KHz, con lo que se mantiene la inteligibilidad de la señal. Sin embargo, la gran cantidad de muestreos que se incluyen en esta presentación hacen que el procesamiento computacional de la señal sea demasiado costoso. El presente algoritmo utiliza una forma alternativa de presentación de la señal. La señal se describe por medio de elementos llamados "ciclos de excursión", que consisten de la parte de la onda entre dos cruces de cero consecutivos. Aunque esta descripción de la señal no preserva la inteligibilidad de la señal, sí preserva suficiente información de la señal para la detección del tono. Como se pueden apreciar existe mucho menos ciclos de excursión dentro de una señal que muestras de ella, lo que presenta un gran ahorro computacional.

Cada "marca" o indicador de la posición está contenida en un ciclo de excursión. En las etapas posteriores del algoritmo se identifican las marcas por medio de las características del ciclo de excursión al que pertenecan. Por lo que la entrada a la estructura informativa se realiza describiendo los ciclos de excursión.

Cada período de tono consta de un pequeño número de ciclos de excursión. Al primer ciclo de excursión

que ocurre en un período de tono se le puede llamar "ciclo principal". El comienzo de un período de tono se define como la primera muestra diferente de cero de un ciclo principal. De acuerdo a esta definición, la identificación del ciclo principal de un período de tono permite la identificación de su comienzo. Por lo tanto, los ciclos de excursión almacenados en la estructura informativa se pueden limitar a aquellos que sean posibles ciclos principales.

Para limitar la estructura se debe analizar numéricamente cada ciclo de excursión como posible ciclo principal. Este análisis se basa en dos propiedades de las señales de voz. La primera es que la forma de onda de la voz se caracteriza generalmente por tener mayores amplitudes al comienzo del período de tono que cerca de su fin. Esta propiedad se puede observar en la Figura 2.10. Como el ciclo principal es el primer ciclo de excursión en un período de tono, por esto contiene la amplitud máxima del período de tono. De acuerdo a esto, cada ciclo de excursión puede caracterizarse por su amplitud. La segunda propiedad es que las porciones o segmentos sin voz tienen intervalos de tiempo entre cruces de cero mas cortos que los intervalos para segmentos con voz.

Por lo tanto, los ciclos principales tienden a tener grandes amplitudes y larga duración, consecuentemente tienen una energía considerable. Por lo cual, los ciclos

de excursión pueden ser analizados de acuerdo a su energía. Para este proceso se suman las muestras dentro de cada ciclo de excursión, computándose el valor aproximado de la energía. Las magnitudes de las sumas de los ciclos principales, generalmente exceden a las sumas computadas de segmentos sin voz. Se estima un valor umbral que exceda las sumas computadas de los segmentos sin voz. Las sumas computadas para cada ciclo de excursión se van comparando con este valor de umbral; si sólo aquellos ciclos de excursión, en que su suma exceda al valor de umbral, se guardan en la estructura informativa. Esta eliminación de los ciclos de excursión de segmentos sin voz ayuda a reducir más la información a procesarse. En las figuras 2.11 y 2.12 se ilustra éste análisis de energía.

El número de posibles ciclos principales se puede reducir más utilizando otra de las propiedades de la señal de la voz. Los ciclos principales exhiben amplitudes extremas negativas, como se ilustra en la Figura 2.43(a) o amplitudes extremas positivas, como se ilustra en la Figura 2.13 (d) ; sin embargo la mayoría de amplitudes exhiben una misma polaridad dentro de una señal de voz dada. Esta polaridad se puede determinar al computar la amplitud máxima de la onda a lo largo de varios períodos de tono. La polaridad equivale al signo de la amplitud máxima. Si la polaridad es negativa se cambia de signo a todas las

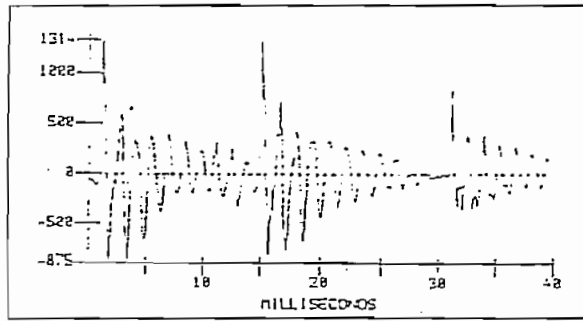


Fig. 2.10 Períodos de tono con mayores amplitudes cerca de su comienzo que cerca de su final.

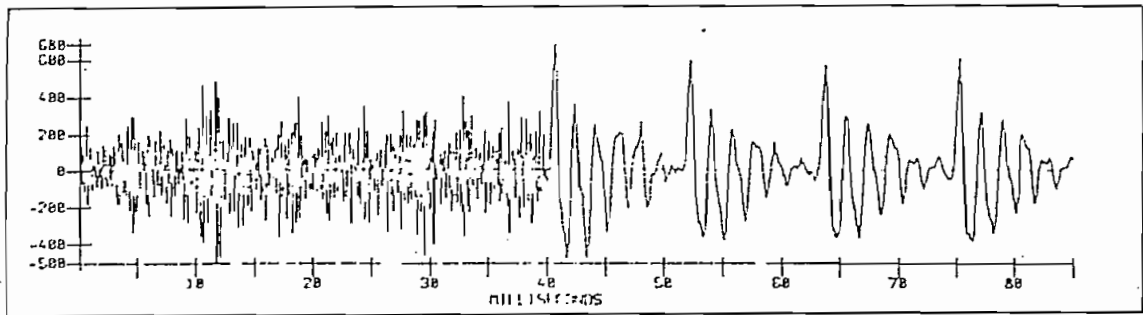


Fig. 2.11 Formas de onda "sin voz" y "con voz"

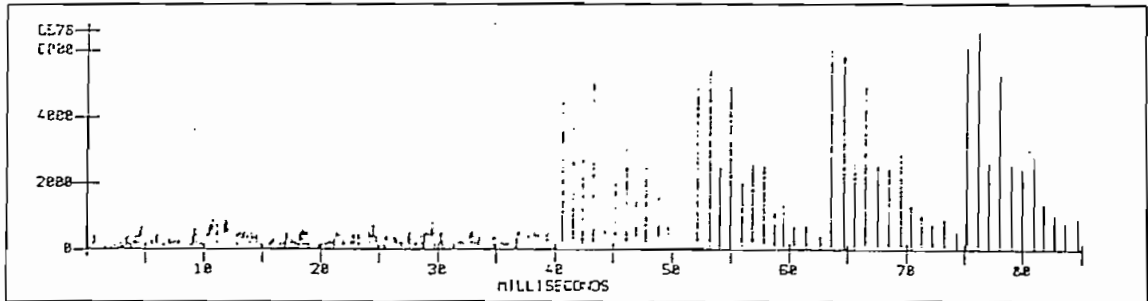


Fig. 2.12. Sumas de las muestras de los ciclos de excursión de la onda de la Fig. 2.11.

muestras, para hacerla positiva. Una vez que todos los ciclos principales se han hecho positivos, sólo se necesita considerar como posibles ciclos principales a aquellos ciclos de excursión que contengan muestreos positivos.

Una reducción adicional del número de ciclos de excursión posibles, se puede realizar limitando la frecuencia de tono. Las frecuencias de tono que pueda detectar el algoritmo, se limita a aquellas bajo los 500 Hz. Esta limitación permite la separación de cada marca del período de tono con al menos 2 mseg con respecto a las demás, Cuando dos ciclos de excursión, seleccionados como posibles ciclos principales, ocurren en un tiempo de menos de 2 mseg, uno de ellos puede ser eliminado. Sólo aquel ciclo de excursión con el mayor sumatorio de amplitud se retiene en la estructura informativa. Mientras las marcas esten separadas con al menos 2 mseg, no pueden existir más de 500 entradas almacenadas por cada segundo de señal muestreada.

La estructura de datos contiene tres entradas de información por cada ciclo de excursión en consideración como ciclo principal. La primera entrada es el número del primer muestreo diferente de cero del ciclo de excursión. Este número de muestreo será utilizado como marca del período de tono de aquellos ciclos de excursión subsiguientes identificados como ciclos principales. La segunda entrada de caracterización de cada ciclo de excursión representa el valor de la amplitud máxima del ciclo de excursión. La

tercera entrada es el número de muestras de este valor máximo. Estas dos últimas entradas se utilizan en las etapas posteriores del algoritmo para aislar los ciclos principales.

La estructura informativa se construye examinando cada muestreo de la señal de la voz. Las muestras negativas están contenidas en ciclos de excursión negativos, por lo cual no se computan. En cambio un muestreo positivo se suma a las demás muestras de su ciclo de excursión, para computar su suma de amplitudes. También se usa cada muestreo positivo para actualizar el valor máximo y la suma. La detección de un muestreo negativo después de un cruce de cero determina la finalización del ciclo de excursión.

En la determinación o final de cada ciclo, se compara el valor de la suma con un valor de umbral, para analizar un ciclo de excursión como un posible ciclo principal. Los ciclos de excursión para los que las sumas de las amplitudes excede al valor de umbral, se examinan con respecto al precedente ciclo principal para determinar su intervalo de tiempo. Cuando el intervalo de tiempo es mayor que 2 mseg, el ciclo principal precedente se introduce a la estructura informativa. Cuando este intervalo de tiempo es menor que 2 mseg, se comparan las sumas de las amplitudes de los dos ciclos de excursión con el mayor sumatorio de amplitud, se considera como un presunto ciclo principal. (4)

Todo este algoritmo para la construcción de la estructura informativa requiere aproximadamente de la mitad del tiempo de computación cuando se muestrea a 20 KHz. Pero la utilización de un muestreo burdo produce varias distorsiones causadas por el ruido ambiental y los segmentos de fricativos de la voz. Estas distorsiones afectan los cruces de cero, que son de mucha importancia en la construcción de la estructura informativa. Esta distorsión puede corregirse por medio de un procesamiento de la señal.

Las condiciones ambientales pueden introducir componentes de baja frecuencia en la señal muestreada, distorsionando el eje del cero efectivo. Para remover esta distorsión se preprocesa la señal pasándola a través de un filtro pasa-alto con frecuencia de corte a 200 Hz.

La componente fricativa frecuentemente dificulta la localización del cruce de cero. Esta componente usualmente tiene una energía significativa sobre los 1000 Hz. Para atenuar la distorsión introducida por la componente fricativa, se procesa la señal pasándola a través de un filtro pasa-bajo con frecuencia de corte a 900 Hz.

Estos dos preprocesamiento no requieren de filtros de frecuencias de corte precisas, ni de características críticas de pasa-banda, ni de atenuación crítica.

Reducción de la Estructura Informativa

En la primera parte del algoritmo se construyó una estructura informativa que contiene a aquellos ciclos de excursión que se consideraran como presuntos ciclos principales de los demás ciclos de excursión. Durante el aislamiento de los ciclos principales no se utiliza toda la estructura informativa, pues realmente la estructura se divide en regiones de voz continua. Estas regiones con voz se dividen además en intervalos correspondientes a sílabas. Para cada uno de estos intervalos silábicos se determina una estimación de la frecuencia promedio de tono. Esta estimación se utiliza para efectuar el aislamiento del ciclo principal dentro del respectivo intervalo silábico.

La división de la estructura informativa en regiones de voz continua se realiza por medio de la localización de aquellas entradas o valores pertenecientes a la estructura que indiquen una discontinuidad en la voz. La frecuencia de tono de señales con voz normalmente no baja de 50 Hz. Por esto, dos ciclos principales consecutivos dentro de un intervalo de voz continua no pueden estar separados con más de 20 msec, Un intervalo mayor a 20 msec entre dos ciclos principales consecutivos indica una discontinuidad en la voz. Como la estructura informativa se construyó en forma secuencial, es decir en el orden de ocurrencia de cada ciclo de excursión, y esta contiene la información co-

respondiente al tiempo de inicio de cada ciclo de excursión ; se concluye que una discontinuidad de la voz puede ser detectada examinando sólo las entradas adyacentes de la estructura informativa.

Algunas regiones de voz continua tienen menos de dos entradas en la estructura informativa. Estas regiones se asumen como sin voz y se desechan, pues la voz humana normalmente contiene al menos dos períodos de tono dentro de cada intervalo con voz. Este proceso de aislamiento de regiones con voz continua, también permite realizar la decisión con voz/sin voz, gracias a la discriminación que se realiza de las regiones con voz.

Según las reglas del lenguaje, una sílaba aparece siempre que ocurra un fonema sonoro entre dos menos sonoros. Por esto la división de regiones con voz en intervalos silábicos requiere la medición de la sonoridad. En este algoritmo se utiliza la amplitud de los datos contenidos en la estructura informativa como una aproximación de la sonoridad. Cada entrada de la estructura informativa contiene el valor máximo correspondiente a un ciclo de excursión y el tiempo de ocurrencia de este máximo. A partir de esta información se dividen las regiones con voz en intervalos silábicos; primero se ubica el núcleo de cada sílaba

y luego se determinan los límites silábicos entre núcleos adyacentes. (29)

Cada región continua con voz contiene por lo menos un núcleo silábico, Si la región tiene una o varias sílabas, se puede aislar fácilmente el núcleo primario silábico que corresponde a aquella entrada de la estructura informativa de amplitud máxima. El núcleo primario y el núcleo secundario se definen por medio de las siguientes propiedades: cada uno ocurre al menos a 80 msec de otro núcleo silábico; la amplitud asociada con una entrada identificada como un núcleo silábico excede a todas las amplitudes a 60 msec de éste.

Para la determinación del núcleo secundario se examina la estructura informativa a ambos lados del núcleo principal. Estas entradas adyacentes al núcleo principal se examinan hasta localizar una que cumpla con las propiedades de amplitud y tiempo expuestas. En la Figura 2.14 se pueden observar las amplitudes almacenadas en la estructura informativa para una región continua con voz; las flechas señalan las amplitudes de las entradas correspondientes a núcleos silábicos.

El límite silábico entre dos núcleos consecutivos se localiza construyendo la envolvente de las amplitudes entre los núcleos. Para la construcción de la envolvente se

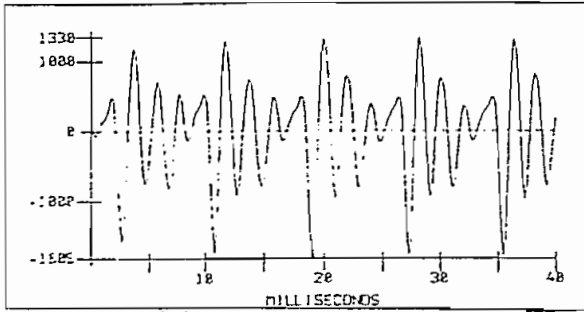


Fig. 2.13(a) Señal con ciclos principales negativos

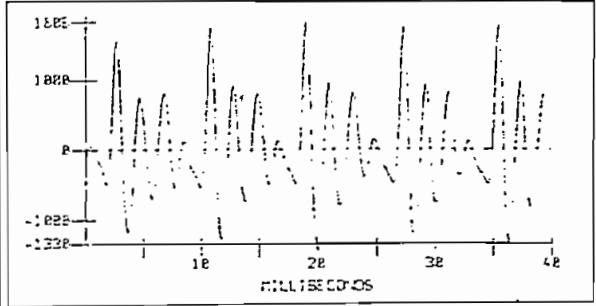


Fig. 2.13 (b) Señal con ciclos principales positivos

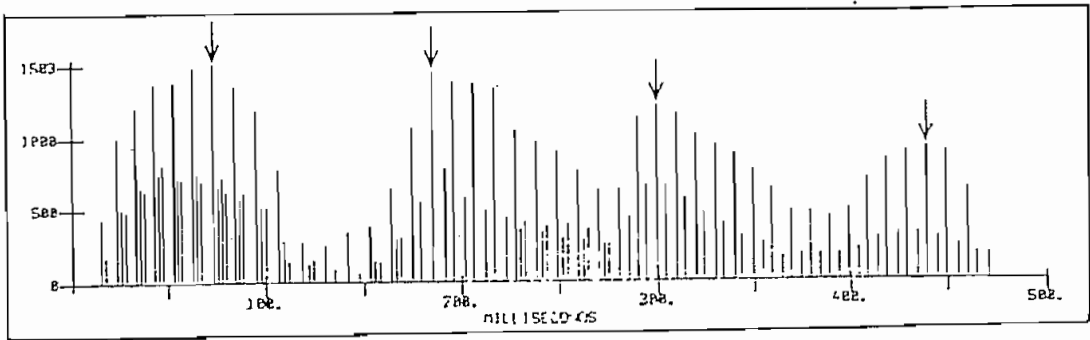


Fig. 2.14 Amplitudes de la estructura de datos de una región continua con voz. Las flechas indican los valores identificados como núcleos silábicos.

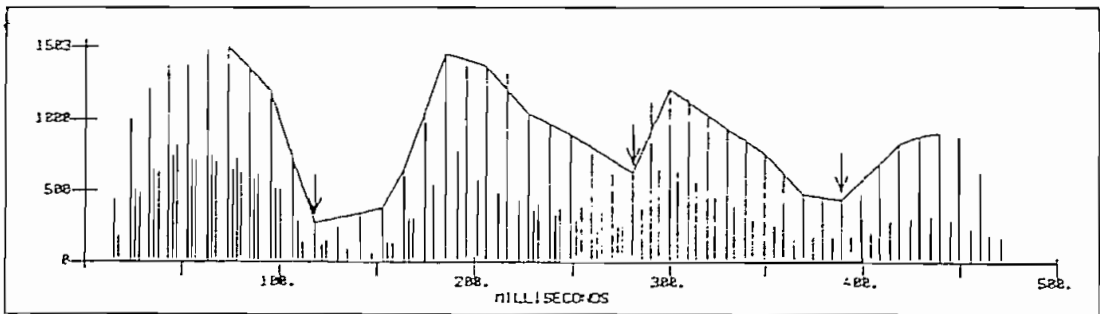


Fig. 2.15 Envoltura de la estructura de datos. Las flechas indican las entradas identificadas como límites silábicos.

segmenta el intervalo de tiempo entre los núcleos en intervalos de 20 mseg. Para cada segmento se determina la amplitud que representará la amplitud de la envolvente durante el segmento respectivo. Esta amplitud se determina localizando la amplitud máxima del segmento. Una vez que se haya construido la envolvente, se aísla el segmento de amplitud más pequeña. La entrada de la estructura informativa correspondiente a la menor amplitud de la envolvente es el límite silábico que se buscaba.

En la Figura 2.15 se puede apreciar la envolvente de las amplitudes correspondiente a una región continua con voz; las flechas señalan las amplitudes de aquellas entradas o datos de la estructura informativa que han sido aisladas como límites silábicos. Cuando hay más de un intervalo silábico en una región continua con voz, hay que modificar los intervalos silábicos de tal manera que los intervalos adyacentes queden separados con 40 mseg. Esto resulta muy efectivo en la reducción de errores de concatenación.

Para cada intervalo silábico se computa una estimación de la frecuencia promedio de tono, para facilitar el proceso de aislamiento o selección de los ciclos principales. Esta estimación se realiza construyendo una envolvente lineal discreta para cada sílaba. El primer paso para la construcción de la envolvente es la división del intervalo silábico en segmentos de 20 mseg. Luego se analizan las entradas de la estructura informativa correspondientes a estos segmentos

para localizar la entrada que tenga la amplitud máxima en cada segmento. Para finalizar se conectan las entradas cuyas amplitudes excedan el 90% del valor de su envolvente se declaran ciclos principales. Como consecuencia, los intervalos de tiempo entre estas entradas son frecuentemente períodos de tono.

Para algunas sílabas muy cortas la estimación de la frecuencia de tono podría ser errónea. Por esto, cada estimación debe ser comparada con las estimaciones de las sílabas anteriores, para cerciorarse de la precisión de la estimación. Para comparar esta estimación con las anteriores, primero se debe determinar una frecuencia media de tono a partir de las estimaciones previas. Debido a que el rango de frecuencia de tono para un locutor dado no se expande a más de una octava, las estimaciones de tono que sean más del doble o menos de la mitad de la frecuencia media de tono serán erróneas. Como los errores más comunes en la estimación del tono son la duplicación y la división para dos de la frecuencia, entonces se deben dividir para dos o duplicar estas estimaciones de la frecuencia de tono hasta estar dentro de un rango aceptable de error.

Este proceso de aislamiento de los ciclos principales se realiza en dos etapas. En la primera etapa se aíslan aquellos ciclos principales que sean fácilmente identificables, basados en la propiedad de que cada ciclo principal de

La primera parte del proceso actual de aislamiento se aplicó a los intervalos silábicos dentro de la región continua de voz. La segunda parte del proceso, en cambio, opera sobre toda la región con voz, aislando cualquier ciclo principal que no haya sido detectado en la primera etapa. La mayoría de los ciclos principales dentro de una región con voz se aíslan en la primera parte del proceso. Esto implica que varios de los intervalos de tiempo entre estos ciclos principales corresponden a estimaciones de frecuencias de tono incorrectas. Las frecuencias de tono deducidas durante la identificación de los ciclos principales, se pueden utilizar para dividir la región con voz en intervalos relacionados con la frecuencia de tono, para una mejor determinación de las frecuencias de tono. Estos intervalos constan de aquellos períodos de tono calculados cuyas frecuencias de tono difieren de las de los períodos adyacentes en menos del 30%. Estos intervalos simples se componen de un sólo período de tono; mientras los intervalos simples e intervalos compuestos se componen de más de un período de tono. Las frecuencias de tono de intervalos compuestos son generalmente correctas. Aunque ocasionalmente estos intervalos compuestos son el resultado de la duplicación o división para dos del período de tono.

En las etapas anteriores del algoritmo se identificaron los ciclos principales de una señal a partir de la magnitud de las amplitudes de los ciclos de excursión. En algunos intervalos de voz, puede suceder que el segundo ciclo

de excursión tenga una amplitud mayor que la del primer ciclo de excursión. De acuerdo a esto, se puede cometer ocasionalmente un error al determinar los ciclos principales como la amplitud máxima de cada ciclo de excursión. En este caso, como las entradas están separadas entre sí por lo menos en 2 mseg los períodos de tono erróneos serán 2 mseg más largos o 2 mseg más cortos que el valor real.

El proceso que se ha descrito en este capítulo es efectivo solamente con voces grabadas directamente que no contengan distorsión de fase. La fase de la señal es crítica para la determinación de las marcas del período de tono. Por esto, este algoritmo es incapaz de aislar las posiciones iniciales de cada período de tono para señales con distorsión de fase. Sin embargo, el algoritmo podría retener su habilidad para la detección correcta de las frecuencias de tono. Esto se debe a que para la determinación de frecuencias de tono correctas, se requiere solamente que la separación de los ciclos de excursión con las amplitudes máximas de la señal coincida con la separación de los períodos de tono.

Si se restringe el rango de locutores a procesar, se simplificarían bastante los cálculos del algoritmo. Si se procesaran únicamente voces masculinas, esto permitiría una separación entre las entradas de la estructura informativa de 4 o 5 mseg en vez de 2 mseg; con lo que se reduciría el número de datos a procesar. En cambio, si se procesaran úni-

camente voces femeninas, esto permitiría la restricción de las frecuencias de tono a un rango sobre los 140 Hz.

Este algoritmo puede correrse en tiempo real a una velocidad media de computación utilizando una frecuencia de muestreo de 20 KHz. Un incremento del tiempo de procesamiento permitiría un proceso adicional de verificación de las marcas de período de tono y de las frecuencias de tono deducidas. La incrementación del tiempo de computación permitiría un proceso independiente que utilizaría las sumas en vez de las amplitudes de los ciclos de excursión para la detección del período de tono. Este proceso es aplicable debido a que las sumas de los ciclos principales frecuentemente exceden a las sumas de los demás ciclos de excursión dentro de cada período de tono. Los resultados de los dos procesos, tanto el de las amplitudes como el de las sumas, podrían entonces ser comparados, y se retendría solamente la información del método que produjera las frecuencias de tono más suavizadas.

Un incremento del tiempo de computación, podría también utilizarse para examinar la envolvente de las amplitudes sincrónicas de tono construídas a partir de las amplitudes de los ciclos principales. Cuando existe duplicación de las frecuencias de tono u ocurren ciclos principales extras, esta envolvente exhibe discontinuidades.

2.4 Detección de tono utilizando la función promedio de la magnitud de la diferencia

La función promedio de la magnitud de la diferencia (AMDF) es una variación de la función de autocorrelación (ACF). Esta función es una herramienta muy útil en el procesamiento de señales de voz. La AMDF no es más que una variación del análisis autocorrelativo, donde en vez de correlacionar con varios retardos la señal de entrada de voz o sea formando multiplicaciones y sumatorios para cada valor de retardo, se forma una señal de diferencia entre la señal retrasada y la señal original de la voz, a la cual se le va tomando la magnitud de la diferencia para cada valor de retardo.

En este capítulo, se describirá un método para utilizar la función promedio de la magnitud de la diferencia y un algoritmo lógico de decisión asociado, para la estimación del período de tono de sonidos con voz. Las principales razones para la implementación de este método son:

- 1) La función AMDF brinda una medición simple que permite una buena estimación del período de tono.
- 2) No hay que realizar operaciones multiplicativas.
- 3) Sus características dinámicas son fáciles de implementar con sólo 16 bits.
- 4) La naturaleza de sus operaciones permite su implementación en un procesador programable o en un computador de propósitos especiales.

La implementación de este tipo de detección de tono puede realizarse tanto en tiempo real como en una simulación en tiempo no real.

Análisis de Autocorrelación y Correlación Cruzada

Matemáticamente, la función de autocorrelación de un segmento "L" de voz digitalizada se define de acuerdo a la siguiente expresión:

$$ACF_r = \frac{1}{L} \sum_{j=1}^L S_j \circ S_{j-r}, \quad r = 0, 1, \dots, r_{\max}$$

donde:

S_j = j-ésimo muestreo del vector de la onda de voz

L = tamaño del segmento de voz

r = valor del retardo

r_{\max} = valor máximo del retardo ($r_{\max} = L$)

Las principales propiedades de la función de autocorrelación son:

- 1) La función de autocorrelación de una señal periódica es también periódica con el mismo período.
- 2) La función de autocorrelación es una función par.
- 3) Tiene su valor máximo para $r = 0$.
- 4) El valor $ACF(0)$ equivale a la energía para señales determinísticas o a la potencia promedio para señales periódicas o aleatorias.

Estas propiedades convierten a la función de autocorrelación en una base o herramienta para la estimación de la periodicidad de todo tipo de señales, incluyendo a las señales de voz (10)

La función de autocorrelación de una señal con voz puede utilizarse para la detección de tono, como se verá en los capítulos posteriores. Generalmente no es necesario computar toda la función de autocorrelación para cada segmento "L" de voz; usualmente se computan valores de retardo o retraso dentro de un rango de 3 a 15 mseg aproximadamente. Debido a que las frecuencias de tono generalmente caen dentro de un rango de 70 a 300 Hz, correspondientes al rango de investigación de 3 a 15 mseg, se puede eliminar el procesamiento excesivo de información en la detección de tono por autocorrelación de cada intervalo "L" para retardos que están fuera de este rango.

En realidad solo se necesita la computación de la mitad de la función de autocorrelación ya que la función es par. A la porción resultante de estas limitaciones, se le examina en busca de su pico máximo. Asumiendo que se satisfacen los criterios de voz de la lógica de tono, se toma como período de tono la posición o el valor de retardo con respecto al retardo cero u origen verdadero de la función de autocorrelación del pico máximo.

Se puede realizar una reducción en el procesamiento de la información al computar sólo una porción del análisis del intervalo de tiempo "L", donde a $L' < L$ se le aplica una autocorrelación de cruce con todo el intervalo de tiempo "L",

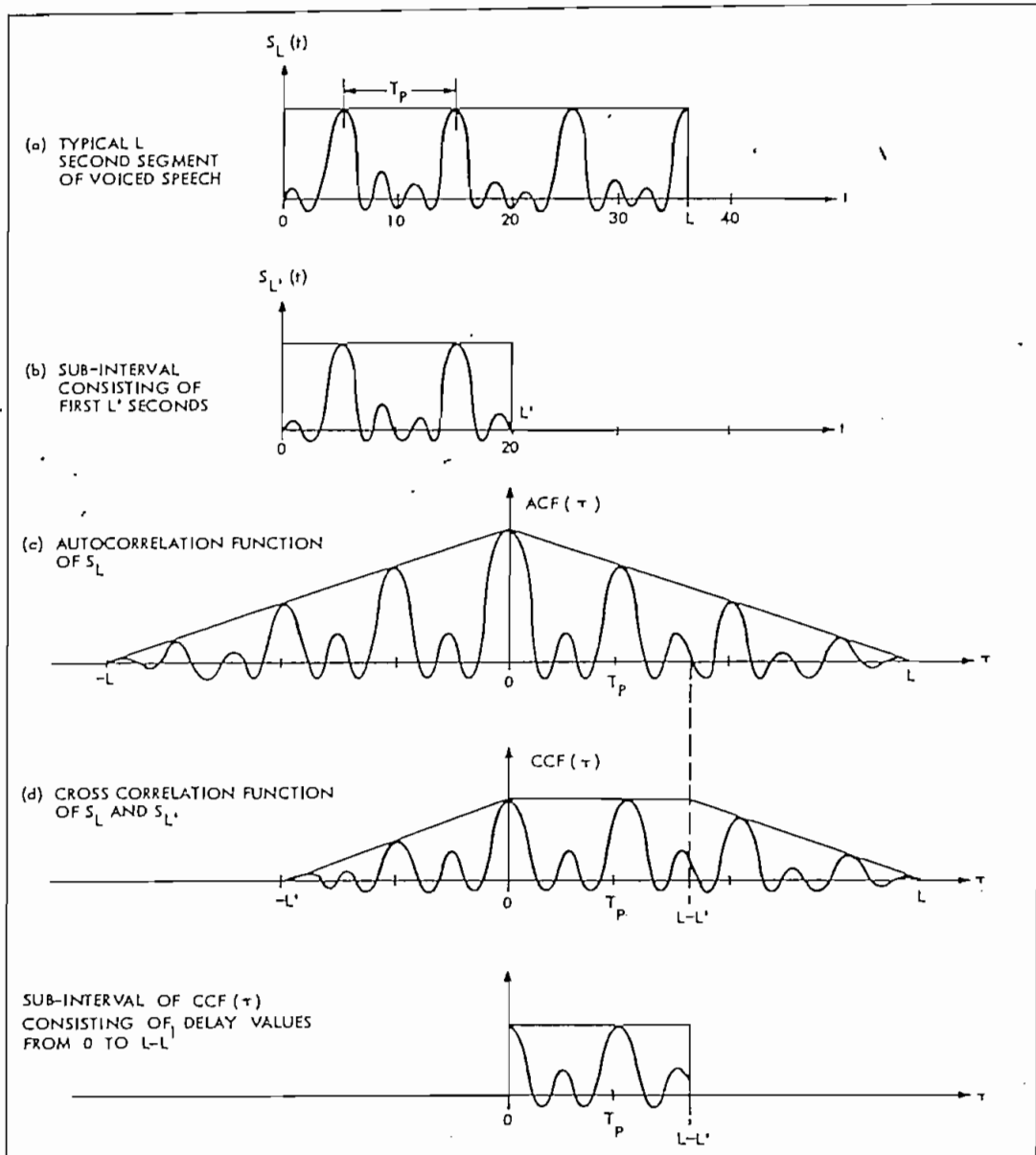


Fig. 2.16 Comparación entre las funciones de autocorrelación y la función de autocorrelación cruzada.

como se puede apreciar en la Figura 2.16. En esta Figura se puede apreciar la función de autocorrelación cruzada (CCF)

la cual es muy útil para la detección del período de tono de la voz.

La expresión matemática de la función de autocorrelación cruzada para un segmento "L" de voz digitalizada es:

$$CCF_r = \frac{1}{L'} \sum_{j=1}^{L'} S'_j \cdot S_{j-r}, \quad r = 0, 1, \dots, r_{\max}$$

donde

S_j = j-esimo muestreo del vector de la onda de voz.

$$(S_j) = (S_1, S_2, \dots, S_{L'})$$

S'_j = j-esimo muestreo del subvector de la onda de

$$\text{voz. } (S'_j) = (S_1, S_2, \dots, S_{L'})$$

L' = porción del segmento de voz

r = valor del retardo.

r_{\max} = valor máximo del retardo ($r_{\max} \leq L - L'$)

Para la función de autocorrelación cruzada la longitud "L" se escoge de acuerdo al período de tono que se espera obtener. Esta longitud podría ser aproximadamente igual a dos períodos del promedio normal para locutores masculinos, esto es 100 Hz. Bajo estas condiciones "L" sería cerca de 20 mseg y un valor típico de L sería 36 mseg, con lo que se tendría un rango para la investigación del período de tono de $L - L'$ igual a 16 mseg. En la práctica se ha visto que "L'" puede tomar valores hasta de 8 o 9 mseg y "L" hasta de 23 mseg.

Una ventaja de este método de autocorrelación cruzada es que el tamaño de los picos de correlación tiende a mantenerse constante como una función del tiempo de retardo. Con respecto a esto, en el método de autocorrelación se puede observar un decrecimiento lineal en el tamaño de los picos de correlación en función del tiempo de retardo. Por otro lado, la constancia de los valores del tamaño de los picos de correlación de la Figura 2.16 ocurre debido a que siempre se realiza una transmisión completa de información entre los dos segmentos correlacionados en cruce por medio del método de la función de correlación de cruce (FF). Este no es el caso del método de autocorrelación (ACF) en el que la transmisión de la información cae linealmente en función del tiempo de retardo. Los picos de correlación son más fáciles de encontrar durante la detección de tono por el método de correlación cruzada (QCF), pues tienden a mantenerse más grandes.

En realidad si L y L' se escogieran cuidadosamente, el método de correlación cruzada minimizaría sus requerimientos computacionales, manteniendo suficiente información para la estimación del período de tono con buena precisión. (18)

Función Promedio de la Magnitud de la Diferencia (AMDF)

Como se puede apreciar en el subíndice anterior, para la computación del proceso de detección de tono por me-

dio de la función de autocorrelación se necesitan demasiadas operaciones matemáticas. Por esta razón, para la medición del período de tono de segmentos con voz, se utiliza una variación del análisis autocorrelativo analizado. Esta variación es la función promedio de la magnitud de la diferencia (AMDF), la cual se implementa con la resta, la suma y el valor absoluto de los valores de operación; en contraste con las operaciones de suma y multiplicación que se utilizaban con la función de autocorrelación.

La función promedio de la magnitud de la diferencia para un segmento con voz, se define de acuerdo a la expresión

$$D_r = \frac{1}{L} \sum_{j=1}^L \left| S_j - S_{j-r} \right|, \quad r = 0, 1, \dots, r_{\max}$$

donde

S_j = j-ésimo muestra del vector de la onda de voz

$(S_j) = (S_1, S_2, \dots, S_L)$

S_{j-r} = muestreos de la señal realizados con un retardo de "r" segundos.

Las barras verticales de la expresión sacan la magnitud de la diferencia $S_j - S_{j-r}$. La señal de diferencia D_r se forma al realizar el sumatorio de las magnitudes de las diferencias entre los valores originales de la voz y los valores de la onda retrasada de la voz. La diferencia entre las señales es siempre igual a cero cuando el retardo es cero o sea cuando no existe retraso entre las señales. Se ha ob-

servado también que para retrasos correspondientes al período de tono de sonidos con voz de estructura cuasiperiódica, aparecen valores de diferencia nulos.

La función promedio de la magnitud de la diferencia para una secuencia de muestreos (S_k) se define por medio de la siguiente expresión.

$$D_n = \frac{1}{N} \sum_{k=0}^{N-1} |S_k - S_{k-n}|$$

donde el valor de retardo "n" varía en un rango desde $-(N-1)$ hasta $+(N-1)$, es decir que para generar la función AMDF completa $n = -(N-1), \dots, -1, 0, 1, 2, 3, \dots, +(N-1)$. El rango de variación del sumatorio es desde $k = n$ hasta $k = N-1$ para $n \geq 0$. Para $n < 0$, el rango de variación del sumatorio es desde $k=0$ hasta $k = N-1+n$. De acuerdo a esta definición se puede decir que D_n es una función par. (31)

La expresión anterior de la función AMDF puede aproximarse de la siguiente manera

$$D_n = \frac{1}{N} \sum_k |S_k - S_{k-n}| \approx B_n \left(\frac{1}{N} \sum_k (S_k - S_{k-n})^2 \right)^{1/2}$$

En esta expresión el coeficiente " B_n " es un factor de escala. Para el caso de una secuencia Gaussiana se puede determinar analíticamente un valor de " B_n " que permita una igualdad en el valor promedio entre la magnitud promedio y el valor de las sumas rms. Para los otros tipos de distribuciones, el valor de " B_n " se puede determinar experimentalmente, al examinar un gran número de secuencias. El valor de " B_n " depende de la función de densidad de probabilidad asociada de S_k y S_{k-n} . Como esta función de densidad de probabilidad asociada de S_{k-n} varía generalmente con el valor de retardo "n",

el coeficiente " β_n " será entonces una función del valor " n ". Normalmente el valor de " β_n " varía entre 0,6 y 1,0 dependiendo de la secuencia de muestreos.

Si se expande la expresión aproximada de la función AMDF, la expresión de D_n queda de la siguiente forma:

$$D_n \approx \beta_n \left(\frac{1}{N} \sum_k S_k^2 + \frac{1}{N} \sum_k S_{k-n}^2 - \frac{2}{N} \sum_k S_k \cdot S_{k-n} \right) 1/2$$

Ahora se define la función de autocorrelación de una secuencia de muestreos (S_k) como

$$R_n = \frac{1}{N} \sum_k S_k \cdot S_{k-n}$$

De acuerdo a esta expresión el tercer término de la expresión de D_n es igual a $2R_n$. Si se asume que la secuencia (S_k) corresponde a un proceso estacionario, se pueden expresar los sumatorios de los dos primeros términos de la expresión de D_n como simples funciones de autocorrelación (ACF) evaluadas para $n = 0$. Esto es

$$R_0 = \frac{1}{N} \sum_k S_k^2 \approx \frac{1}{N} \sum_k S_{k-n}^2$$

Si se reemplazan estas dos últimas expresiones en la ecuación de D_n , se obtiene la expresión definitiva de la función AMDF

$$D_n \approx \beta_n (2(R_0 - R_n)) 1/2$$

Las propiedades de la función promedio de la magnitud de la diferencia (AMDF) se pueden caracterizar a partir de la última expresión de D_n . Específicamente, el valor de la

función AMDF es igual a cero cuando no existe retardo ($n = 0$) y varía como la raíz cuadrada de la diferencia de las funciones de autocorrelación (ACF) respectivas,

El valor de D_n aparecerá como nulo en aquellos puntos donde R_n sea muy grande comparada con R_0 . Esto ocurre cuando se toma la secuencia (S_k) de un sonido con voz que contenga dos o más períodos de tono. La separación entre estos valores nulos es igual al período de tono.

En la figura 2.17 se pueden apreciar los gráficos de las funciones R_n , $(R_0 - R_n)$ y $(R_0 - R_n)^{1/2}$ para una función de autocorrelación (ACF) periódica. En el gráfico en que se comparan la diferencia con la raíz cuadrada de la diferencia, se puede apreciar que el efecto que produce la raíz cuadrada es el de reducir el ancho de los valores nulos. Este efecto es muy útil para incrementar la precisión en la medición del período de tono. Por lo tanto, un ancho reducido de los valores nulos permite una mayor precisión en la determinación de sus posiciones.

Se puede afirmar que la función promedio de la magnitud de la diferencia (AMDF) es una variación del análisis de autocorrelación (ACF), donde en vez de correlacionar la señal de entrada de la voz para varios retardos por medio de multiplicaciones y sumas, se forma una señal de diferencia entre la señal retardada de voz y la señal original de voz a la cual se le saca el valor de la magnitud absoluta para cada valor de retraso. A diferencia de la función de

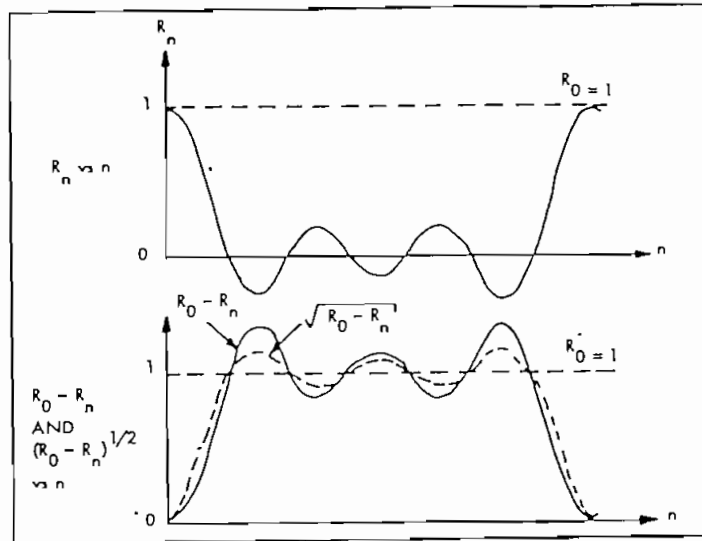


Fig. 2.17 Gráficos de R_n ; $R_0 - R_n$ y $(R_0 - R_n)^{1/2}$ para una función de autocorreclación periódica

autocorreclación de la función de correlación cruzada, la función promedio de la magnitud de la diferencia (AMDF) no requiere dentro de sus cálculos de operaciones multiplicativas, lo cual es una ventaja para aplicaciones en tiempo real.

Para cada valor de retraso o retardo, se realiza la computación a lo largo de una ventana de "L" muestreos, en un proceso similar al que se utilizó para la obtención de la función de correlación cruzada de la figura 2.16. Para poder generar todo el rango de valores de retardo, se "diferencia en cruce" la ventana por medio de un análisis de intervalo completo. La ventaja de este método es que los ta-

maños de los valores nulos tienden a mantenerse constantes en función del valor del retardo. Esto se debe a que se produce una transmisión completa de información entre los dos segmentos que se diferencian en cruce. (18).

En los detectores de tono de este tipo, el factor que limita su precisión es la imposibilidad de separar completamente la estructura, de los efectos de la envolvente espectral. Por esta razón, se debe utilizar un proceso de decisión lógico y las características principales de la voz junto con la ~~función AMDF~~, para una mejor detección del período de tono.

Proceso lógico para la detección de tono

En la Figura 2.18 se puede apreciar el diagrama de flujo para una detección del período de tono por medio de la función promedio de la magnitud de la diferencia (AMDF). En este diagrama de flujo se ve el grupo de instrucciones lógicas desarrolladas para la respectiva extracción de información del período de tono a partir de la función AMDF. La complejidad de este grupo de instrucciones lógicas es comparable a la lógica de un detector de tono por autocorrelación (ACF).

Existen cinco caminos lógicos diferentes, los cuales se escogen en base a las tres más recientes decisiones

lógicas con voz/sin voz (VUV). El parámetro "LOGIC" se obtiene al tratar las tres decisiones consecutivas como un número binario, de acuerdo a la siguiente ecuación.

$$\text{LOGIC} = \text{VUV}(n) + 2.\text{VUV}(n-1) + 4.\text{VUV}(n-2)$$

donde

$$\text{VUV}(n) = 0, \text{ si el enesimo intervalo fue sin voz.}$$

$$\text{VUV}(n) = 1, \text{ si el enesimo intervalo fue con voz.}$$

El rango de valores que puede tomar "LOGIC" varía de 0 a 7, es decir que según el diseño lógico de tono puede haber ocho condiciones posibles. Los valores de umbral que se utilizan en este diagrama de flujo se determinaren empíricamente al examinar la información de las voces de varios locutores diferentes.

En el camino "A", la decisión VUV es "sin voz", por lo cual la lógica pregunta si esta decisión debe ser cambiada a una decisión "con voz". El cambio se justifica cuando se presenta una fuerte onda periódica dentro del intervalo de análisis.

En el camino "B", la decisión VUV es "con voz". Normalmente el tono deber ser igual a la posición mínima de la función AMDF dentro del rango de búsqueda. Sin embargo, puede producirse una decisión "sin voz" si el valor del máximo de la función AMDF no es suficientemente fuerte o si la razón del valor máximo para el valor mínimo es menor que el valor de umbral especificado.

En el camino "C", las decisiones VUV para los intervalos "n" y "(n-1)" son "con voz", pero la decisión para el intervalo "(n-2)" es "sin voz". Esto significa que para una indicación de presencia del primer impulso de un intervalo "con voz", el extractor de tono debe cambiar a decisión VUV "con voz" y colocar como nuevo valor de tono el valor mínimo de la función AMDF.

En el camino "D", se extiende en un intervalo más a la señal "con voz" cuando la decisión VUV indique una señal "sin voz" después de un largo lapso de señal "con voz". Si para el nuevo intervalo se detecta una decisión "sin voz", la onda de voz tendrá una amplitud tan bajo que no será de importancia para el sintetizador. En cambio es de mucha importancia si el extractor de tono estaba errado, pues el intervalo resulta ser en realidad "con voz". Como se puede apreciar, por medio de esta extensión se elimina la posibilidad de que se presente un intervalo "sin voz" en medio de un sonido "con voz".

El camino "E" representa el camino normal para sonidos sostenidos "con voz", en el cual se utiliza un proceso para determinar la magnitud y posición del período de tono verdadero. Se define una ventana de rastreo de ± 12 muestreos con respecto a la última medición del período de tono, para realizar la búsqueda lógica del valor mínimo. A este mínimo se lo compara con el valor mínimo de todo el rango

de búsqueda de la función AMDF. Normalmente la posición de este mínimo corresponderá al período de tono, pero la lógica del proceso puede cambiar el valor del período de tono por la posición del mínimo fuera de la zona de búsqueda, si la amplitud del valor mínimo fuera del rango de rastreo resulta menor que la mitad de la amplitud del valor mínimo rastreado. En el caso de que existan frecuencias de tono muy altas, se tendrá una mayor cantidad de mínimos en la función en la función AMDF. Para este caso, sólo se podría cambiar el período de tono si la amplitud del valor mínimo fuera del rango de rastreo fuera menor que un octavo de la amplitud del valor mínimo rastreado dentro de la ventana. En este camino, también hay un camino para cambiar la decisión VUV de "con voz" a "sin voz" y para extender el intervalo para chequear el valor previo del período de tono.

Para todos los intervalos "sin voz" (UV) y para el primer intervalo de los segmentos "con voz" (V), se invierten en el tiempo los muestreos de la señal de entrada. Este procedimiento se realiza para superar un problema serio relacionado con el primer impulso de un sonido "con voz". Para este intervalo de análisis, la onda tiene una parte "sin voz" y la otra parte "con voz", con la ventana de rastreo en la parte "sin voz". La función de correlación de estos muestreos no existirá para ninguno de los valores de información. Al

invertir la función de tiempo se ubican las muestras de la parte "con voz" de la onda de voz en la ventana de integración. Si el intervalo es básicamente "con voz", la función AMDF tendrá un mínimo de pequeña amplitud; si en cambio menos del 50% del intervalo es "con voz", la decisión lógica será una salida "sin voz".

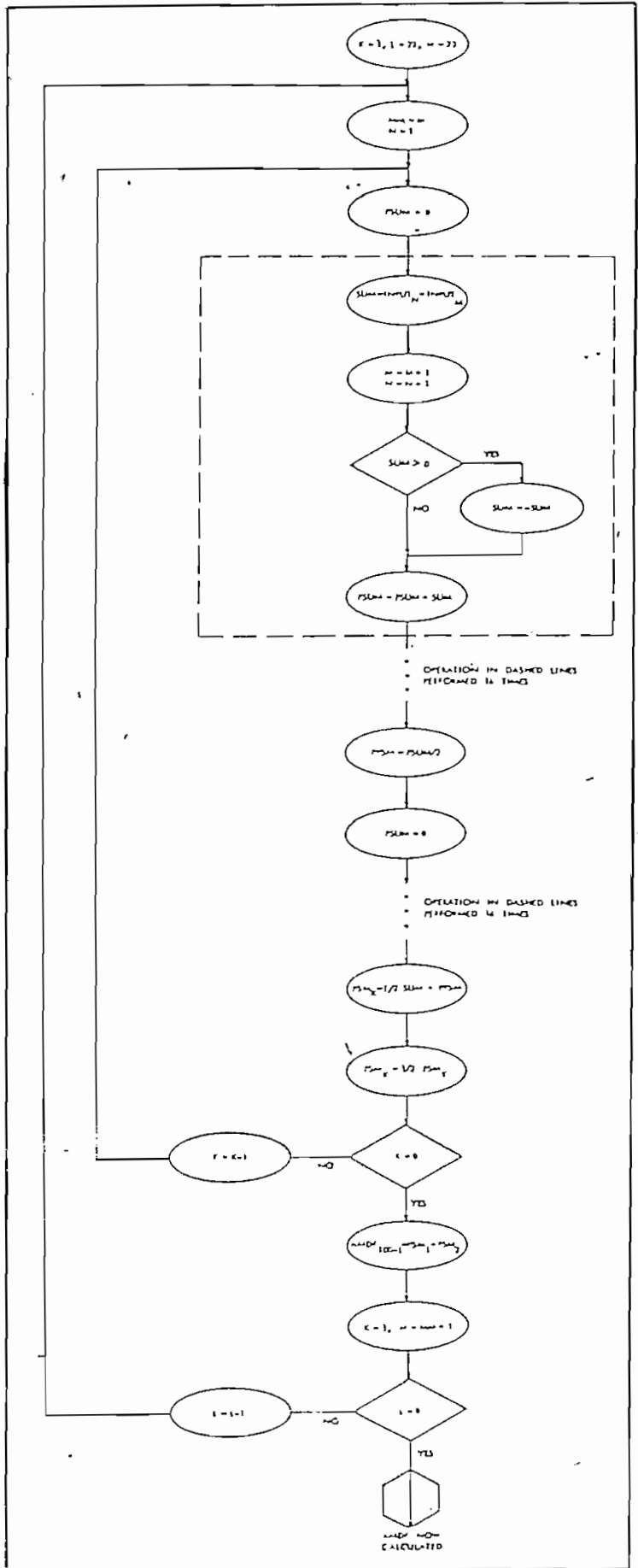
Para todos los intervalos "con voz" excepto aquellos correspondientes al primer impulso de un sonido "con voz", no se invertirán en el tiempo las muestras durante el análisis de la ventana de rastreo. En el final de un sonido "con voz" se realiza la misma inversión en el tiempo, para que la parte periódica "con voz" se incluya en la ventana de rastreo. Este proceso permite una mayor precisión en la detección del período de tono de un sonido "con voz". (18).

Implementación en Tiempo Real

La habilidad para implementar el algoritmo para la detección de tono a partir de la función AMDF en tiempo real, depende directamente del número de operaciones necesarias y de la velocidad computacional de la máquina que se utilice.

En la Figura 2.19 se encuentra el diagrama de flujo del grupo de instrucciones que permiten la generación de las muestras de una función AMDF en tiempo real en un procesador de señales GTE Sylvania. Para este aparato en particular, una instrucción de multiplicación lleva 750 nseg, mientras

Fig. 2.19 Diagrama de flujo para la generación de una Función AMDF en tiempo real.



que instrucciones de suma, resta y de manipulación de información toman generalmente de 250 a 375 nseg; cada una. Para cada grupo de operaciones dentro del bloque punteado de la Figura 2.19 se requiere un tiempo de 1,875 μ seg; en este tiempo se incluyen 500 nseg para cada instrucción de carga (load), resta, examen (test) y negación, y un tiempo adicional de 375 nseg para almacenar la suma parcial.

Para el análisis de cada intervalo se necesitan un número de computaciones igual al número de muestreos de la ventana de búsqueda, para poder calcular un sólo punto de la función AMDF. Por ejemplo para un segmento de voz de 9 mseg muestreado a 7040Hz se tienen 64 muestreos, lo que implica la necesidad de 64 computaciones. El intervalo de tiempo será entonces igual a $64 \times 1,875 \mu$ seg o sea 120 μ seg. Como se generan 77 muestreos por cada porción de la función AMDF, el tiempo total que se requiere es cerca de 9,24 mseg. Estos 77 muestreos son lo adecuado para permitir el análisis para la detección de tono en un rango de 70 Hz a 300 Hz si se muestrea a 7040 Hz.

También se debe incluir dentro de esta estimación de tiempo, el tiempo que requieren los lazos de control para mantener una cierta precisión en la generación de la función AMDF. Para esto se requieren 250 μ seg aproximadamente. El tiempo total para el intervalo, incluyendo unos 0,2 mseg

para la lógica de decisión, equivale a la estimación del tiempo de corrido u operación para la detección del período de tono en este aparato. Los tiempos para la generación de la función AMDF y para la ejecución de la lógica de decisión en varios sistemas actuales son de 10 mseg y 0,25 mseg respectivamente.

Para tener una buena precisión al realizar el sumatorio correspondiente a la expresión de la función AMDF, se divide este sumatorio en cuatro sumatorios de 16 términos cada uno para el cálculo de cada punto de la función AMDF. La expresión dividida queda así:

$$D_r = \frac{1}{4} \sum_{j=1}^{16} |s_j - s_{j-r}| + \frac{1}{4} \sum_{j=17}^{32} |s_j - s_{j-r}| \\ + \frac{1}{4} \sum_{j=33}^{48} |s_j - s_{j-r}| + \frac{1}{4} \sum_{j=49}^{64} |s_j - s_{j-r}|$$

donde $r = 0, 1, \dots, r_{\max}$

La información de entrada se escala de tal forma que el valor máximo que se pueda obtener sea $\pm (2^{10} - 1)$. Debido a que cada muestreo es el resultado de la combinación de dos muestreos de entrada, se requiere un factor de 2 que da un máximo igual a $2(2^{10} - 1) = 2^{11} - 2$. Como el número de muestreos dentro de la ventana de búsqueda es de 64, la exactitud que se requiere para el cálculo de cada punto de la función AMDF es igual a $(2^6) (2^{11} - 2) = 2^{17} - 2^7$; con lo que se sobrepasaría la capacidad de un computador de 16-bit. En cambio si se utiliza esta división o segmentación de la expresión,

la exactitud que se requiere para el sumatorio de 16 puntos de la función AMDF es igual a $(2^4) (2^{11} - 2) = 2^{15} - 2^5$. Luego se combinan las cuatro sumas parciales para generar un punto de la función AMDF. Este proceso si es calculable sin sobrepasar la capacidad de un computador de 16-bit.

Existen varios métodos para reducir el tiempo de ejecución o corrido del algoritmo para la detección de tono por medio de la función AMDF. Una posibilidad es generar la función AMDF en un aparato "hardware" externo y luego introducir los resultados al computador a través de un canal I/O (entrada/salida). La lógica para la detección del período de tono de la señal AMDF requiere un diseño de "hardware" muy complejo, pero esto representaría una menor carga del procesador; por lo tanto, lo indicado sería el realizar el proceso lógico dentro del procesador. El circuito de interface hacia un generador de la función AMDF externo puede ser un poco sencillo. Para la entrada, el generador de la función AMDF usaría el convertidor A/D (análogo/digital) del procesador; y para la salida, se podría transmitir la función AMDF hacia el procesador central a la frecuencia de muestreo. Otra forma de meter la información es utilizando un conversor A/D de magnitud y signo, con el que se eliminaría la instrucción de examen de la magnitud del bloque punteado de la Figura 2.19. Esto reduciría el tiempo de generación de la función AMDF de 9,24 mseg a cerca de 6,80 mseg.

Otra posibilidad para una operación más rápida es la de desarrollar instrucciones especiales para reducir el número de computaciones. Una instrucción que sumada a la magnitud del acumulador dentro de un registro de precisión de 20-bit en un ciclo de instrucciones, podría reducir el tiempo del lazo interior y el tiempo de escalamiento de las sumas intermedias. Este tipo de instrucción podría reducir el tiempo de generación de la función AMDF de 9.24 mseg a cerca de 5,44 mseg.

Otra posibilidad interesante incluye ideas como la de limitar la banda de la señal de entrada a 1000 Hz y realizar un nuevo muestreo a un 1/5 de la frecuencia de muestreo original. Este puede brindar una reducción del número de computaciones necesarias para obtener la función AMDF. Se han logrado buenos resultados con un esquema en el cual se genera cada segundo o tercer punto de la función AMDF, mientras que se restringe la lógica de decisión para poder ignorar los puntos de la función que no se hayan calculado. En estos dos últimos casos, se debe notar que al reducir la información es muy posible que se pierda en precisión. (22).

Al simular la lógica para la detección de tono mediante la función AMDF, la decisión con voz/sin voz puede alterarse, con lo que se producirían errores en la medición del período de tono. El caso en que se producen los errores más graves es cuando una decisión "con voz" se cambia a decisión "sin voz" por alguna degradación de la señal.

La introducción de ruido en la señal de voz, también

produce una generación de errores en la medición del período de tono. Estos errores dependen de la voz del locutor, pero consisten en la duplicación del período de tono al comienzo o en la parte central de sonidos "con voz". Normalmente se puede realizar una buena detección del período de tono por medio de la función AMDF utilizando una relación de señal a ruido S/N igual a cero dB. Sin embargo, se debe realizar un estudio de un rango amplio de sonidos y locutores para poder determinar la relación de señal a ruido más baja que se pueda utilizar para la detección del período de tono utilizando la función promedio de la magnitud de la diferencia (AMDF). (32).

Como ya se ha visto la implementación de los métodos de autocorrelación (ACF) y de correlación cruzada (CCF) requiere de multiplicaciones en vez de las instrucciones de suma y magnitud que se utilizan en el método de detección de tono por medio de la función AMDF. Para ciertas computadoras con operaciones multiplicativas veloces, no hay mucha diferencia entre el tiempo para generar la función ACF o la función CCF en comparación con el tiempo para generar la función AMDF. Sin embargo, el costo del "hardware" para aumentar la velocidad de multiplicación es considerable.

También se debe anotar que para mantener un rango dinámico grande asociado a las operaciones multiplicativas se necesita un mayor escalamiento que para el proceso equivalente de las sumas. Para concluir, se puede afirmar que en

el uso del algoritmo AMDF para la detección del período de tono de la voz, no se requieren operaciones multiplicativas ni se tiene ninguna restricción del rango dinámico. Esto representa una ventaja sobre los métodos de ACF y CCF con respecto a la complejidad del computador para su implementación.

2.5 Métodos de detección de tono usando la función de autocorrelación.

Anteriormente, se pudo apreciar como la función de autocorrelación provee de una representación adecuada, a partir de la cual se pueden desarrollar esquemas para la detección del período de tono en función del tiempo. En este capítulo, se verán varios detalles relacionados con la implementación de detectores de tono basados en la función de autocorrelación.

La mayor limitación de la representación autocorrelativa es que retiene demasiada información de la señal de la voz. En realidad basta con un intervalo de valores de la función de autocorrelación entre $0 \leq r \leq 10$, para poder realizar una buena estimación de la función de transferencia del canal bucal.

Para un mejor procesamiento de la señal de la voz, se suele hacer que la periodicidad sea más prominente, al mismo tiempo que se suprimen aquellos rasgos que puedan distorsionar la señal. Las técnicas que realizan este

tipo de procesamiento de una señal se llaman "aplanadoras de espectro", pues su objetivo es el de remover los efectos producidos por la función de transferencia del canal bucal, haciendo que cada armónica se ponga a un mismo nivel de amplitud.

Los tres métodos para la detección de tono que se verán en este capítulo utilizan la técnica de aplanamiento del espectro. El procedimiento común a los tres métodos es el siguiente:

- 1) las armónicas de la frecuencia fundamental se igualan en amplitud.
- 2) Se sincronizan en fase las armónicas entre sí, con lo que se obtiene un tren de impulsos con picos de amplitud grande.
- 3) Se determina la ubicación de los picos de mayor amplitud (máximos).
- 4) El intervalo de tiempo entre estos pulsos corresponderá a la estimación del período de tono.
- 5) La decisión con voz/sin voz se basa en la presencia o ausencia del tren de pulsos respectivamente.

Los tres métodos para la detección de tono que se describirán son:

- 1) Autocorrelación del espectro aplanado.
- 2) Recortado central del espectro y autocorrelación.
- 3) Recortado de tres niveles y autocorrelación.

Las amplitudes de las armónicas de una señal no se pueden igualar tan fácilmente, debido a la gran variedad de formantes y frecuencias fundamentales que existen. Para esto se utiliza un sistema dinámico que se adapte a las variaciones del espectro en el tiempo. En la figura 2.20 (a) se puede observar el diagrama de bloques de un sistema de este tipo. La señal de la voz se filtra a través de "n" filtros pasa-banda de aproximadamente 100 Hz, para expandir el ancho de banda de la señal. Las salidas " F_i " de los filtros van a un rectificador de onda completa y a un suavizador para dar una estimación " A_i " de la amplitud de la señal " F_i ". La señal " S_i " es la señal " F_i " retrasada para compensar el retardo que se introduce al suavizar la señal. La señal " $C_i = S_i/A_i$ " equivale a la señal " F_i " normalizada en amplitud. La suma de las señales " C_i " es igual al espectro aplanado de la señal. En la Figura 2.21 se puede observar el espectro aplanado de una señal de voz. Como era de esperarse, los intervalos con voz presentan un tren de impulsos, mientras que en los intervalos sin voz está ausente el tren de impulsos (21).

Si las amplitudes de las armónicas se sincronizaran en fase, se facilitaría la detección del período de tono de la señal. La sincronización de estas amplitudes se puede realizar con el sistema de la Figura 2.20(b). En la Figura

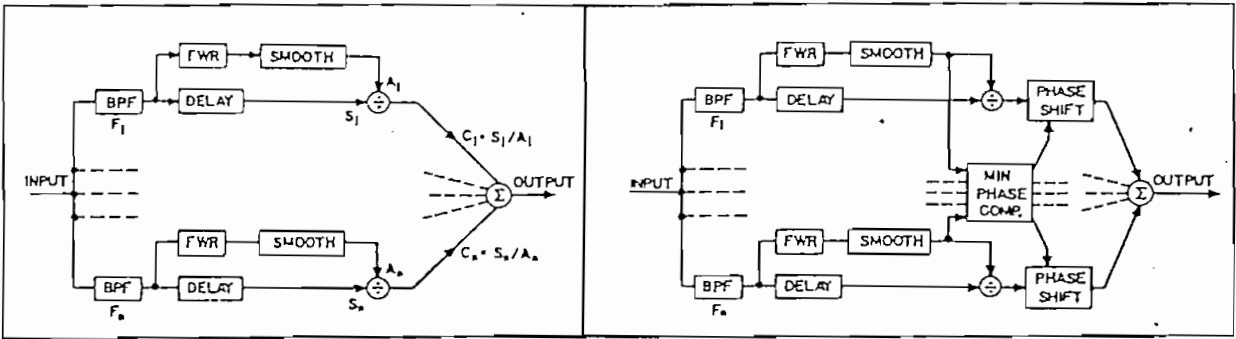


Fig. 2.20 (a) Esquema del aplanador de espectro.

Fig. 2.20 (b) Esquema del aplanador de espectro con mínima compensación de Fase.

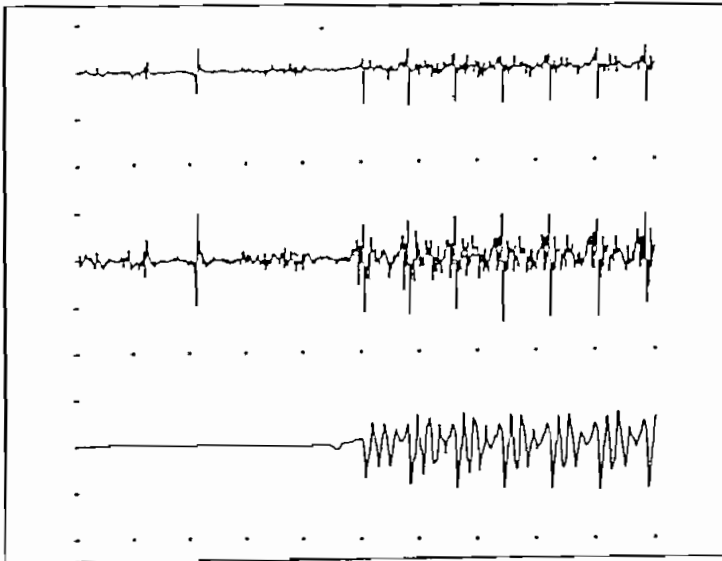


Fig. 2.21 Ejemplos de espectro aplanado y mínima compensación de Fase (señal superior), espectro aplanado (señal de la mitad) y onda original de voz (señal inferior).

2.21 se puede observar un espectro aplanado y sincronizado.

Detección de Tono por Autocorrelación del Espectro Aplanado.

En este método se sincronizan las amplitudes de las armónicas por medio de la autocorrelación de la señal del espectro aplanado. Aunque en otros métodos se ha utilizado la función de autocorrelación para la detección del período de tono, lo que distingue a este método es la forma en que se preprocesa la señal. El preproceso de aplanamiento del espectro elimina las formantes de la señal, permitiendo que la función de autocorrelación no tenga picos debido a estas formantes.

Para la computación de la función de correlación se aísla un segmento de 30 mseg del espectro aplanado de la señal de la voz, al cual se le multiplica por una ventana "hamming". La computación de la función de autocorrelación para este segmento, se realiza para un retardo de 15 mseg y se normaliza con respecto a un retardo igual a cero. A continuación se selecciona un nuevo intervalo de 30 mseg, a 15 mseg del segmento anterior, para el cual se repite todo el proceso. Este proceso se realiza en forma sucesiva hasta obtener una señal como la de la Figura 2.22 (a). En esta figura, para poder apreciar mejor la onda, se rectificó en media onda y se sacó la raíz cuadrada de la función de correlación.

Generalmente los picos de la función de correlación correspondientes al período de tono son fáciles de detectar. Sin embargo, para asegurarse de que no se pierdan los picos durante una transmisión rápida del canal bucal, se utiliza

un algoritmo de decisión. Este algoritmo y el método para generar parámetros para la decisión con voz/sin voz se describirá después de analizar los otros dos métodos para la detección del período de tono.

Detección de Tono por Recortado Central y Autocorrelación.

Aunque se han propuesto numerosas técnicas de aplanamiento de espectro la más aceptada es la de recortado central propuesta también por Sondhi. El recortado central de una señal de voz se obtiene a partir de la transformación no lineal indicada en la Figura 2.23(a). La técnica para obtener la señal recortada centralmente se ilustra en la figura 2.23(b) en la parte superior se observa el segmento de voz para el cual se va a computar la función de autocorrelación. Para este segmento, se determina la amplitud máxima " A_{\max} " para cada intervalo de 5 msec y se fija un valor de recortado " C_L " igual a " $\pm k.A_{\max}$ ". Normalmente "k" se escoge igual a 0,3 aproximadamente. En la parte inferior de la figura 2.23(b) se puede observar que la salida del recortado central es igual a la entrada menos el nivel de recorte " C_L "; y para los valores bajo el nivel de recorte la salida se iguala a cero. En este caso, los picos se convierten en pulsos constituidos por la parte de cada pico, que exceda al valor de recorte. A esta señal se le computa por medio de la función de autocorrelación (33).

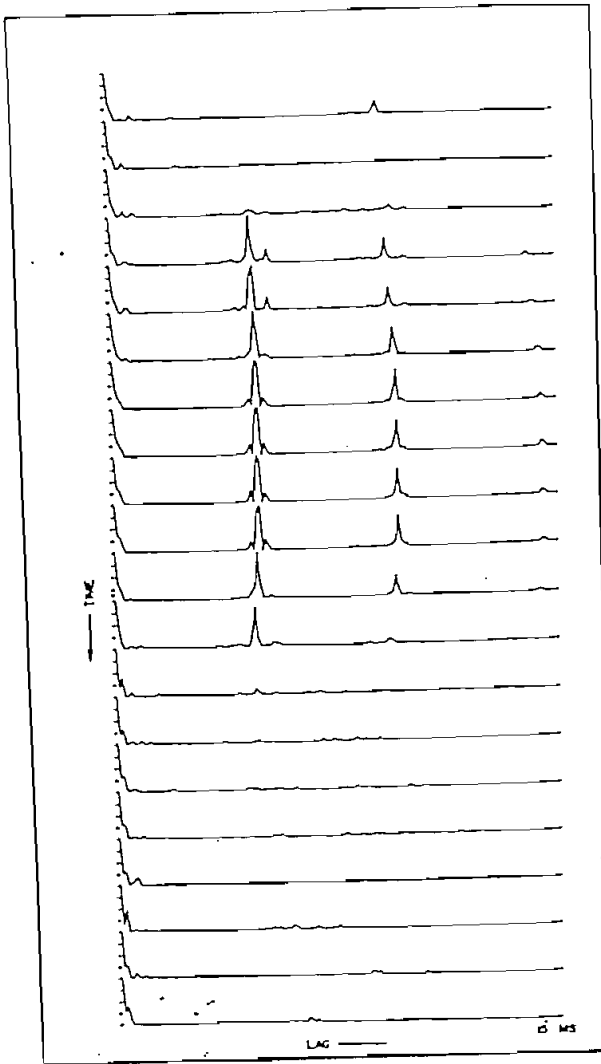


Fig. 2.22 (a) Función de correlación con espectro aplanado.

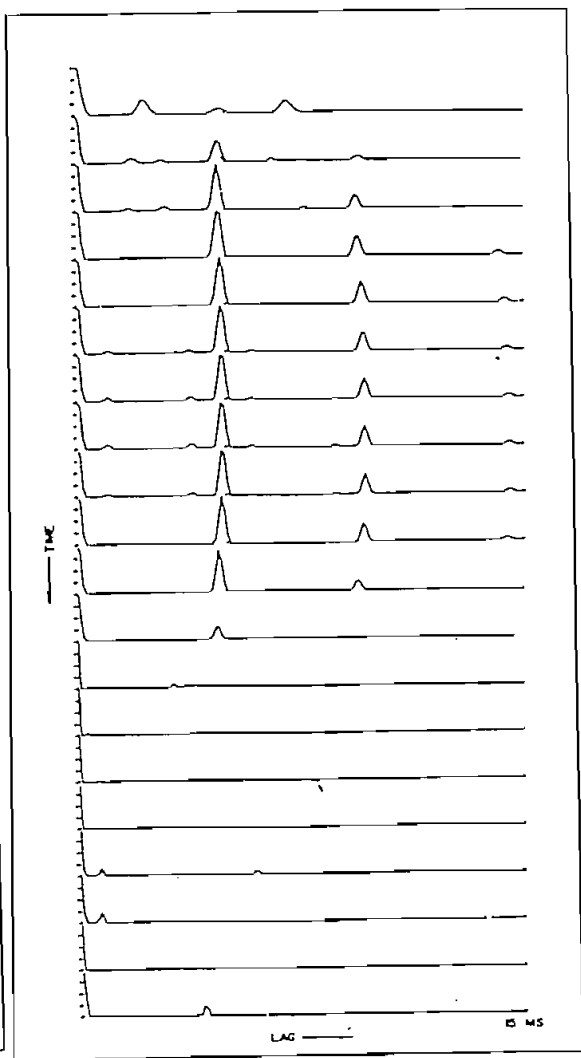


Fig. 2.22 (b) Función de autocorrelación con recortado central del espectro.

En la Figura 2.24 se indica el efecto que produce el proceso de recortado central en la computación de la función de autocorrelación. En la parte superior se muestra un segmento de señal con voz muestreado a 10 KHz. Para la función

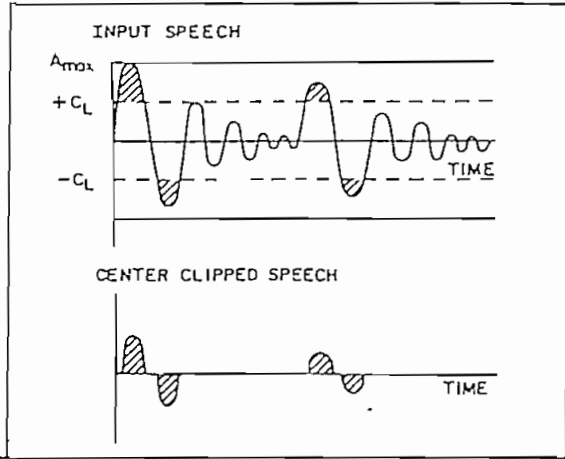


Fig. 2.23(a) Función de recortado central

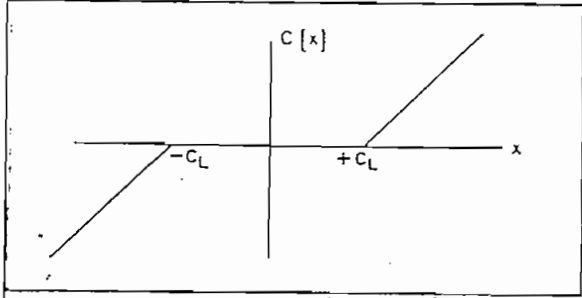


Fig. 2.23(b) Ilustración del proceso de recortado central.

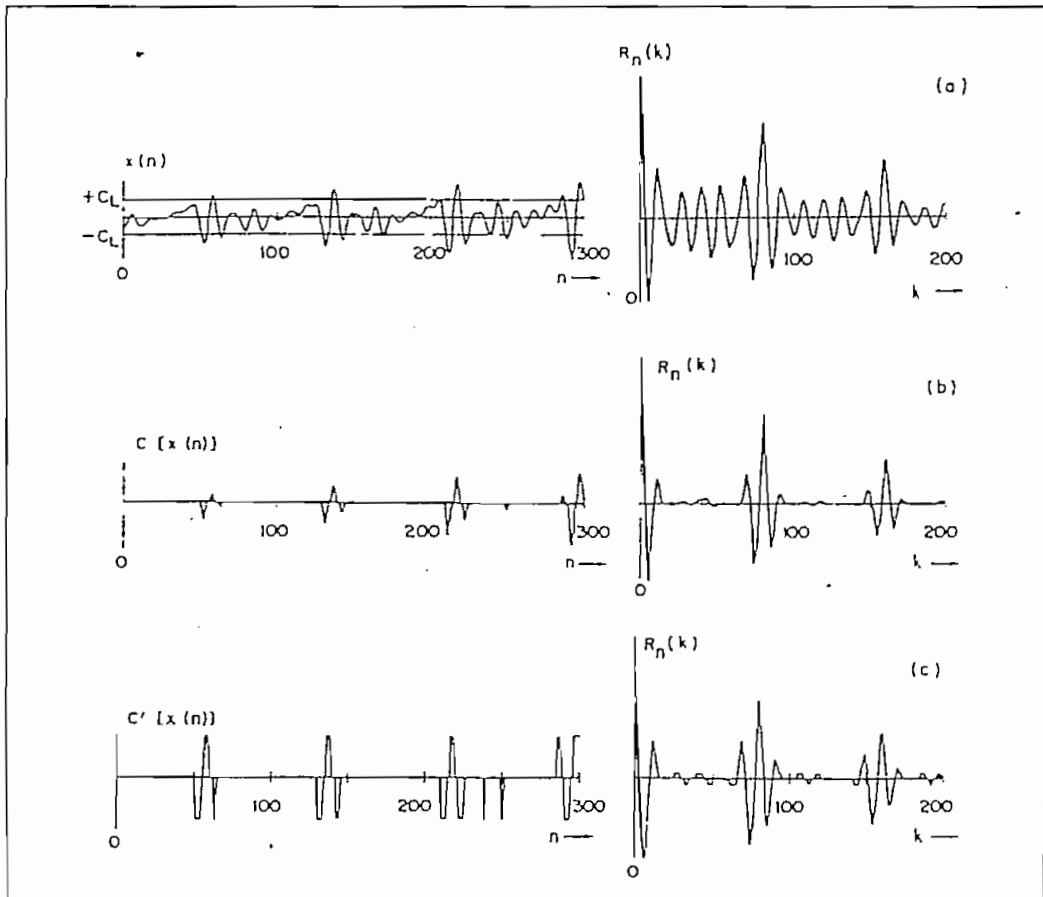


Fig. 2.24 Ejemplos de ondas y sus Funciones de correlación, (a) sin recortado (b) con recortado central y (c) con recortado a tres niveles.

de autocorrelación de este segmento se puede observar un pico o máximo muy fuerte que indica su período de tono. Los demás picos se le atribuyen a las oscilaciones amortiguadas del canal bucal. En la parte central se puede observar el recortado central para una señal de voz, en el cual se utilizó un 68% de la amplitud máxima durante los primeros 100 muestreos. Se puede notar que únicamente quedan graficados varios pulsos separados por el período original de tono. La función de autocorrelación resultante de esta señal, tiene muy pocos picos que puedan crear confusión durante la detección del período de tono.

Por lo tanto, entre más grande sea el nivel de recorte menos pulsos aparecerán a la salida, y menos picos extraños aparecerán en la función de autocorrelación. Esto se puede ver claramente en la Figura 2.25 donde aparecen las funciones de autocorrelación para tres niveles de recortado central diferentes. Se puede apreciar que a medida que el nivel de recorte decrece, aparecen más picos a lo largo del proceso, lo que implica que la función de autocorrelación se haga muy compleja. La indicación más clara del período de tono se obtiene para el nivel de recorte más alto que se permita, pues existe una limitación de este valor. Cuando la amplitud de la señal varía apreciablemente al comienzo o final de un segmento con voz, si se utiliza un nivel de recorte muy alto, se puede perder la onda si ésta cae por debajo de este nivel de recorte.

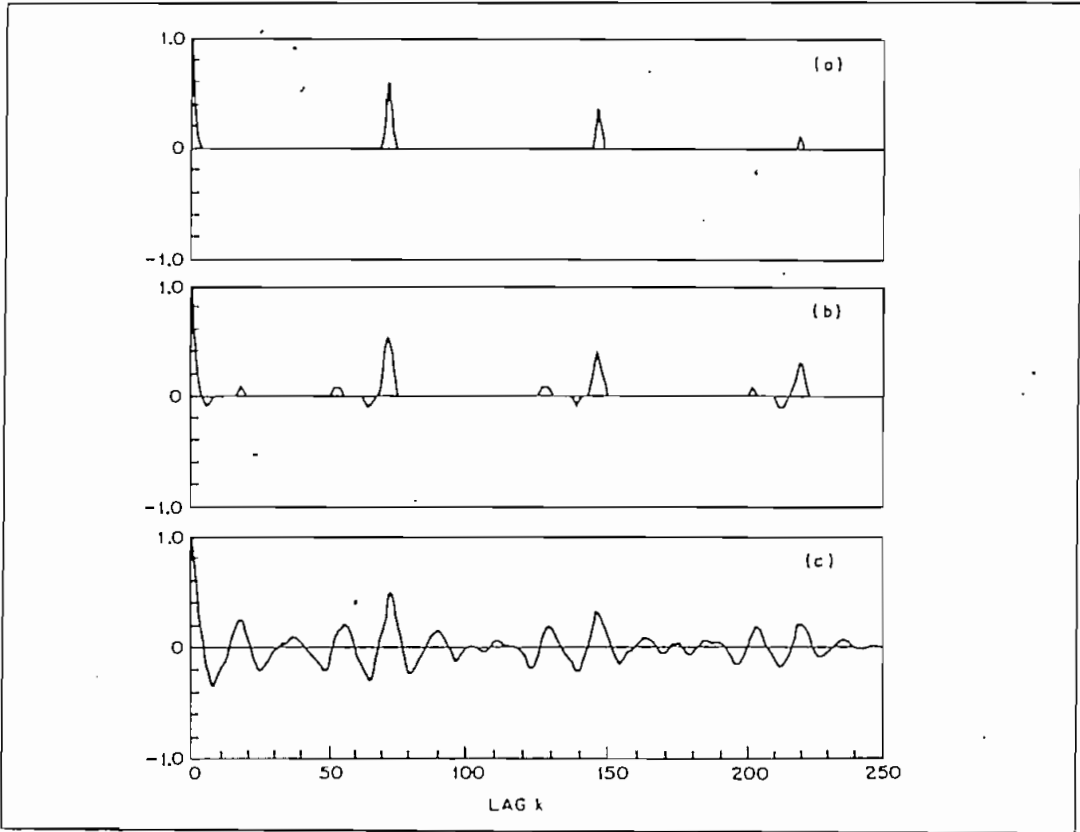


Fig. 2.25 Funciones de autocorrelación para una onda de voz recortada centralmente: (a) $C_L = 80\%$; (b) $C_L = 64\%$ y (c) $C_L = 48\%$.

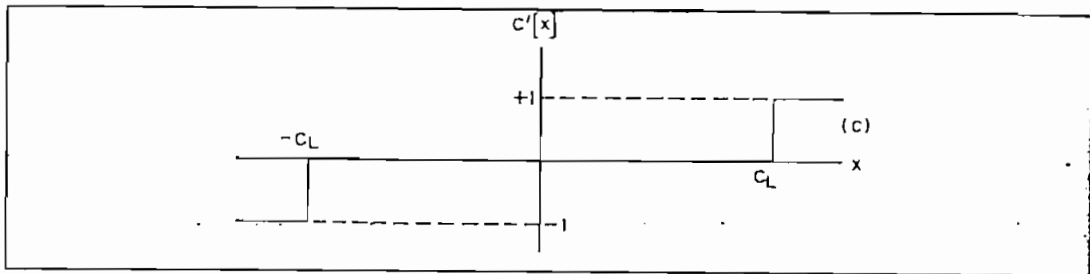


Fig. 2.26 Función de recortado central a 3 niveles.

Como se ha visto, la técnica de recortado central es un proceso bastante simple y muy eficiente para eliminar las formantes de una señal de voz. Con esta técnica, la función de autocorrelación y el algoritmo de decisión que se describirá más adelante, se puede lograr una muy buena detección del período de tono de una señal de voz.

Por lo menos en un tipo de situación, como la de un segmento con voz casi sinusoidal, este proceso funciona mejor que el de aplanamiento del espectro. Esto ocurre para el sonido /i/ hablado por una mujer y filtrado por un filtro pasa-alto con punto de corte a 200 o 300 Hz. Esta situación no es inusual si la voz ha viajado a través de una línea telefónica. En este tipo de casos, la exactitud de métodos como el cepstrum y el de aplanamiento del espectro, depende de la presencia de un gran número de armónicas lo que los hace susceptibles a errores. En cambio en el método de recortado central la ausencia de un gran número de armónicas no representa ningún problema. (34).

Detección de Tono por Recortado de Tres Niveles y Autocorrelación.

Los problemas que acarrearán los picos extraños en la función de autocorrelación, pueden solucionarse por medio del recortado central de la señal antes de computar la función de autocorrelación. Sin embargo, el problema principal

de la representación autocorrelativa es el gran número de computaciones que requiere para el proceso de detección de tono. Una simple modificación de la función de recortado central permite reducir el número de computaciones necesarias para representar la función de autocorrelación sin que se degenera la señal para la detección de tono. A esta modificación se le llama "recortado de tres niveles" y corresponde a la función de la Figura 2.26. En esta figura se puede observar que cuando $x(n) > C_L$ la salida de la función de recortado es igual a +1, y si $x(n) < C_L$ entonces la salida es igual a -1; fuera de estos límites la función es igual a cero. En la parte inferior de la Figura 2.24 se puede observar la onda resultante de la función de recortado de tres niveles para un segmento con voz.

Por medio de la técnica de recortado de tres niveles se enfatiza la importancia de aquellos picos que excedan el nivel de recorte, es decir se eliminan la mayoría de los picos extraños, con lo cual se logra una indicación de la periodicidad precisa. La computación de la función de autocorrelación con un recortado de tres niveles para una señal de voz es muy sencilla. La expresión que describe la función de autocorrelación de una señal recortada a tres niveles es

$$R_n(k) = \sum_{m=0}^{N-k-1} \bar{y}(n+m) \cdot y(n+m+k)$$

donde el producto $y(n+m)$ y $(n+m+k)$ sólo puede tomar tres valores diferentes.

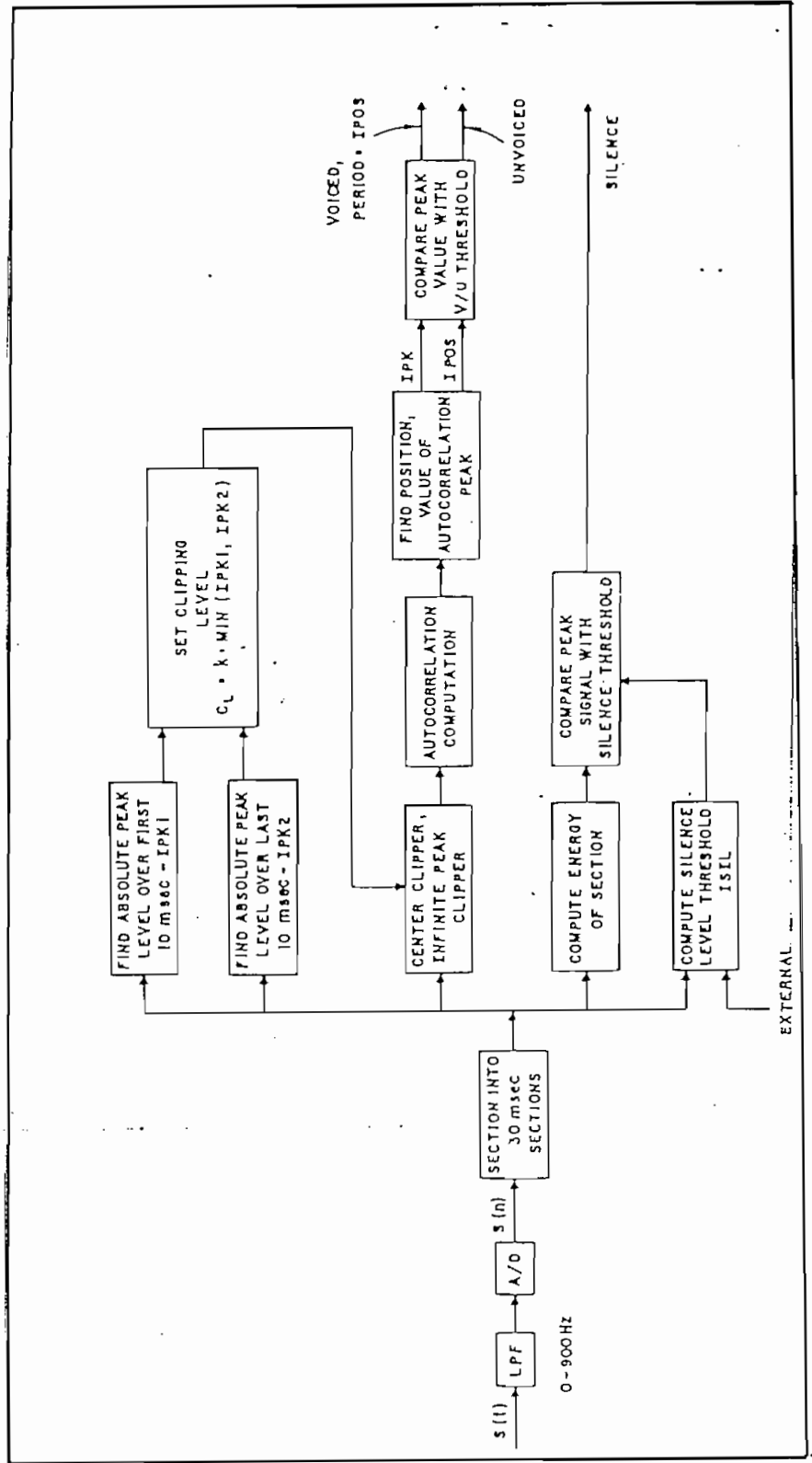
$$\begin{aligned} y(n+m) \text{ y } (n+m+k) &= 0 && ; \text{ si } y(n+m) = 0 \quad \text{o} \quad y(n+m+k) = 0 \\ &= +1 && ; \text{ si } y(n+m) = y(n+m+k) \\ &= 1 && ; \text{ si } y(n+m) \neq y(n+m+k) \end{aligned}$$

de acuerdo a estas expresiones, se puede ver que para implementar esta función con un dispositivo "hardware" se requiere de un simple circuito combinatorial y de un contador "up-down" para ir acumulando el valor de la función de autocorrelación para cada valor de "k".

A partir de esta función recortada a tres niveles, se puede implementar la función amplitud de la magnitud de la diferencia (AMDF), para realizar la computación del período de tono de una señal de voz. Combinaciones de técnicas para la detección del período de tono, como ésta, se utilizan frecuentemente en situaciones específicas (35).

En la Figura 2.27 se puede apreciar el diagrama de bloques para la implementación digital en "hardware", del algoritmo para la detección de tono por medio de la función de autocorrelación y la técnica de recortado de tres niveles. Los pasos que se siguen para la ejecución de este algoritmo son:

Fig. 2.27 Diagrama de bloques de un detector de tono con autocorrelación recortada.



- 1) Se filtra la señal de voz con un filtro pasa-bajo análogo de frecuencia de corte igual a 900 Hz y se muestrea la señal a 10 KHz.
- 2) Se seleccionan segmentos de voz de 30 msec de longitud o sea de 300 muestreos, separados entre sí por intervalos de 20 msec.
- 3) Se computa la magnitud promedio por medio de una ventana rectangular de 100 muestreos. Se comparan las amplitudes de los picos con respecto a un valor de umbral, que se determina midiendo el nivel de la amplitud de los picos de una señal de ruido a lo largo de 50 msec. Si la amplitud del pico de la señal es mayor que el valor de umbral, el segmento se clasifica como "con voz" y se continúa en la ejecución del algoritmo; de lo contrario el segmento se clasificara como "sin voz".
- 4) Se determina el nivel de recorte como un porcentaje (por ejemplo 68%) del mínimo valor entre las amplitudes absolutas de los valores máximos de la señal entre los 100 primeros y 100 últimos muestreos del segmento de voz.
- 5) Con este nivel de recorte se procesa la señal de voz aplicando la técnica de recortado de tres niveles y computando la función de autocorrelación a lo largo de varios períodos de tono.

- 6) Se localiza el pico de mayor amplitud de la función de autocorrelación y se lo compara con respecto a un valor de umbral igual al 30% de $R_n(0)$. Si la amplitud del pico es menor que el valor de umbral, el segmento se clasifica como "con voz" y el período de tono respectivo es igual a la ubicación o posición del pico de amplitud máxima.

A partir de este esquema del algoritmo se pueden implementar los algoritmos para las técnicas de autocorrelación del espectro aplanado y de recortado central del espectro, realizando pequeñas variaciones (36).

En la Figura 2.28 se pueden observar las ondas de salida de los tres algoritmos expuestos en este capítulo. En estas ondas se pueden apreciar algunos puntos dispersos que corresponden a errores debidos a algún pico de mayor amplitud que la del pico correspondiente al período de tono. También se puede notar que los períodos de tono promedios entre 100 y 150 muestreos presentan una caída debida a la atenuación de los picos del período de tono causada por la computación de la función de autocorrelación de un segmento sin voz. Por lo tanto, los picos de la función de autocorrelación producidos por la respuesta del canal bucal son de mayor amplitud que aquellos que se deben a la periodicidad.

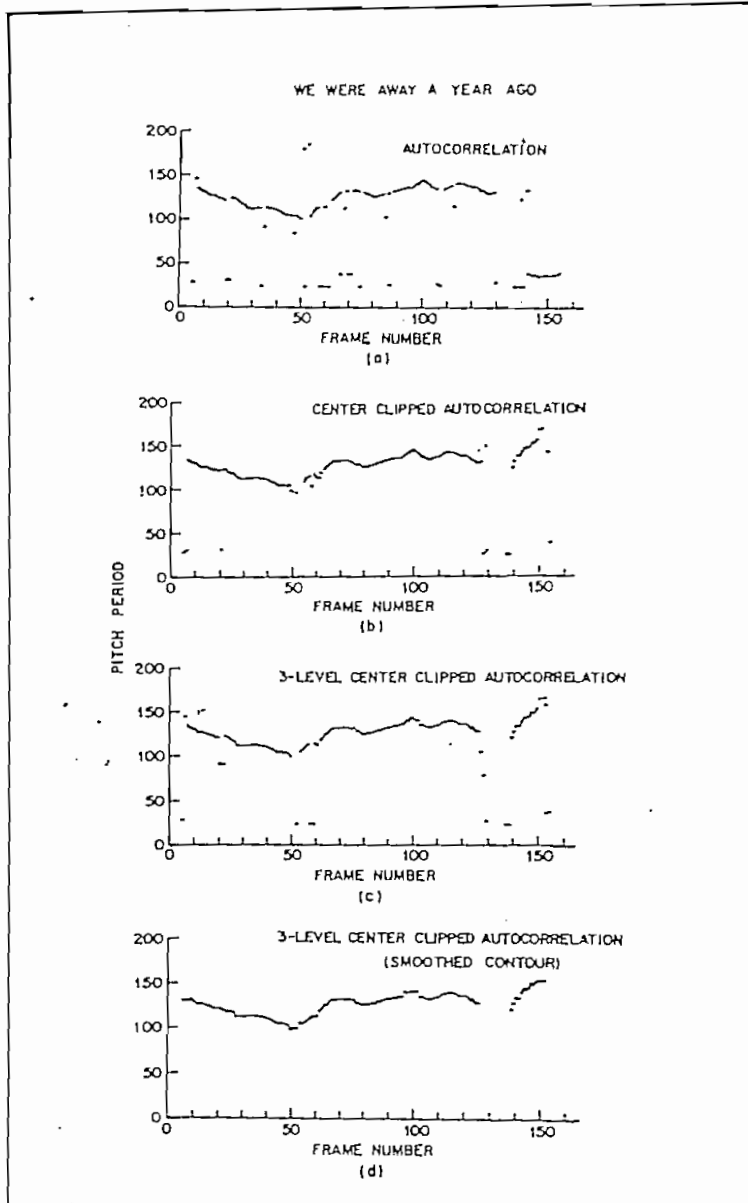


Fig. 2.28 Salidas de un detector de tono autocorrelativo; (a) sin recortado; (b) con recortado central; (c) con recortado central a tres niveles y (d) salida de (c) suavizada en forma no lineal.

En la parte superior de la Figura 2.28 se puede observar la onda resultante del algoritmo de autocorrelación; la siguiente onda corresponde al algoritmo para detección del período de tono por medio del recortado central de la función de autocorrelación; la tercera onda corresponde a la detección del período de tono por medio del recortado de tres niveles de la función de autocorrelación y en la parte inferior aparece la versión suavizada de la onda recortada a tres niveles y autocorrelacionada. En esta última onda se puede apreciar que por medio del recortado y el filtrado para su suavización, se han eliminado la mayoría de los errores de las ondas superiores.

Algoritmo de Decisión

El algoritmo de decisión que se expone a continuación está directamente relacionado con el algoritmo expuesto para la detección del período de tono por medio del recortado a tres niveles de la función de autocorrelación. Este algoritmo de decisión forma parte de las técnicas de autocorrelación del espectro aplanado y la de recortado central de la función de autocorrelación, para la detección del período de tono de la voz.

Por medio de este algoritmo de decisión se escogen los picos de correlación que correspondan a períodos de tono de intervalos con voz y además realiza la decisión con voz/sin voz. Las principales etapas de este algoritmo de decisión son:

- 1) Se seleccionan segmentos de voz de 30 mseg, separados entre sí por intervalos de 15 mseg, como base del algoritmo.
- 2) Se pasa la función de autocorrelación por una ventana Hamming, para enfatizar las amplitudes de los picos de la señal.
- 3) Se define un valor de umbral, a partir del cual se ubica el primer pico de la función de correlación que exceda a este valor de umbral. La posición de este pico corresponde al período de tono para el correspondiente intervalo de 15 mseg.
- 4) La decisión del período de tono para cada intervalo se compara con las de los intervalos posterior y precedente. Si el intervalo precedente y el intervalo posterior son "con voz", entonces el intervalo en análisis se declara como "con voz". Si los intervalos anterior y precedente son "sin voz", el intervalo en análisis se declara como "sin voz". Si los períodos de tono precedente y posterior son aproximadamente iguales y si el período de tono que se analiza difiere de estos, con más del 60%, entonces el período de tono para el intervalo respectivo es igual al promedio de los períodos de

tono adyacentes. Si dos intervalos sucesivos presentan una gran diferencia con respecto al período de tono del intervalo precedente, esta diferencia se acepta como una variación genuina de la voz.

- 5) Si se detecta un pico dentro de un intervalo de 15 mseg, el valor de umbral para los siguientes intervalos se divide para dos, para una región de \pm 1 mseg alrededor de este pico. El valor de umbral vuelve a ser el mismo si el período de tono cambia o si el intervalo "con voz" termina.

Por medio de este algoritmo de decisión se pueden eliminar los errores de duplicación del período de tono, de división para dos del período de tono, de segmentos "con voz" espurea y de segmentos "sin voz" espurea. (21).

Los métodos para la detección del período de tono expuestos en este capítulo, pueden soportar un filtrado considerable de tipo pasa-alto, sin que se degenera el proceso de detección. Estos algoritmos son capaces de tolerar un ruido de gran ancho de banda sin perder en la exactitud de la detección del período de tono.

Para concluir, la simulación en un computador de cualquiera de estos algoritmos para la detección del período de tono no es muy adecuada, pues los tiempos de computación son muy superiores a los de las técnicas descritas en los primeros capítulos; pero esto no significa que se pueda mejorar por medio de un intenso estudio del algoritmo y de los

modelos de "hardware" existentes. Hay que anotar también que a partir de la función de autocorrelación se han enunciado numerosos algoritmos para la estimación del período de tono a parte de los que se han estudiado y que sin ninguna duda muchos más se enunciarán en el futuro.

2.6 Evaluación de los métodos

Los cuatro métodos para detección de tono estudiados en este capítulo, corresponden a los métodos que presentan mayores velocidades de ejecución de acuerdo a un estudio realizado por un grupo de ingenieros de la IEEE. Según este estudio aparecen los algoritmos clasificados de acuerdo a su velocidad así:

- 1) DARD - (reducción de información)
- 2) PPROC -(procesamiento en paralelo)
- 3) AMDF - (función promedio de la magnitud de la diferencia).
- 4) AUTO- (función de autocorrelación).

Cabe anotar que los otros métodos mencionados, además de éstos, son mucho más lentos.

De acuerdo a este análisis se han escogido los dos primeros algoritmos para su implementación en programas para el computador, ya que por tratarse de un minicomputador con una capacidad de memoria muy limitada para este tipo de algoritmos y de una velocidad lenta con respecto a computado-

res de mayor capacidad, se tiene que evitar la utilización de algoritmos que impliquen grandes operaciones matemáticas como es el caso de los algoritmos AMDF y AUTOOC que se basan fundamentalmente en un análisis espectral de la función de autocorrelación o de una variación de esta.

Entre los métodos de detección de tono por reducción de información y el de procesamiento en paralelo, existe una relación de velocidad de ejecución de 1 a 1.5 a favor del primero. Es decir que el algoritmo de detección por reducción de información realiza un procesamiento de la señal de voz que le permite ejecutar el algoritmo para tres segmentos de longitudes iguales, mientras el método de procesamiento en paralelo lo haga tan solo para dos de ellos.

Sin embargo, se debe anotar que uno de los problemas más complejos en la comparación y evaluación de los diferentes detectores del período de tono es el de escoger un criterio objetivo para ello. Básicamente el problema es que un criterio adecuado para una cierta aplicación puede no servir para otro tipo de aplicación.

Existen varias características para formarse un criterio con respecto a la elección de un algoritmo adecuado para la detección del período de tono. Entre ellas están las siguientes:

- 1) Exactitud en la estimación del período de tono.
- 2) Exactitud en la decisión "con voz"/"sin voz".
- 3) La calidad de las mediciones.
- 4) La velocidad de operación.
- 5) La complejidad del algoritmo.
- 6) La posibilidad de implementación de su "hardware"
- 7) El costo de su implementación en "hardware"

Dependiendo de la aplicación específica que se le vaya a dar, se deben considerar uno o más de esos factores para formarse un criterio acertado con respecto a cuales utilizar.

En este trabajo los factores determinantes van a ser la exactitud en la estimación del período de tono, la velocidad de operación y la complejidad del algoritmo. Con respecto a la exactitud en las estimaciones del período de tono, la mayoría de autores coinciden que entre estos cuatro métodos existe una equivalencia entre los resultados obtenidos para una variedad de locutores.

Con respecto a la complejidad de los algoritmos, ya se ha anotado, que los dos primeros métodos son más convenientes por el hecho de que a pesar de que los algoritmos son un poco extensos, no son tan complejos como el desarrollo matemático ya sea de la función de autocorrelación o de la función AMDF, que como ya se ha visto no es más que una variación de esta primera. Para el desarrollo de este trabajo este factor resulta determinante ya que la capacidad real del computador Textronix 4051, es tan solo de 30 KBytes. Es-

to implica que para el desarrollo de un programa se tendrá que tratar de evitar la mayor cantidad de operaciones matemáticas sin perder exactitud en la estimación del período de tono. Permitiéndose al mismo tiempo el procesar la mayor cantidad de datos que sea posible.

C A P I T U L O T E R C E R O

Programas Digitales Para La Detección Del Tono

3.1 Introducción.

3.2 Estructura de la Biblioteca de Programas.

3.3 Programa para Ingreso de Datos.

3.4 Programa para la Detección del Tono por Procesamiento
en Paralelo.

3.5 Programa para la Detección del tono por Reducción
de información.

C A P Í T U L O I I I

Programas digitales para la detección del Tono

3,1 Introducción

En este capítulo, se tratará de explicar en forma detallada la metodología que se ha seguido para el desarrollo de los programas, para resolver los algoritmos de detección de tono.

Cada uno de los algoritmos de tono se tratarán en forma separada, ya que sus fundamentos son totalmente independientes y además por que físicamente están separados. Para poder realizar las pruebas para comprobar el funcionamiento de estos dos programas, se vió la necesidad de desarrollar un programa de ingresos de datos que permitiera varias posibilidades de ingreso, útiles para el presente trabajo o para aplicaciones futuras.

En este trabajo se han tomado en cuenta las necesidades para futuras aplicaciones, razón por la cual se ha tratado de adaptar una parte de este programa de ingreso a las condiciones de operación que se tendrán con el interfase que se está desarrollando para este computador del Laboratorio de Control.

Debo anotar que durante el desarrollo de estos programas, se han tratado de optimizar al máximo la utilización de memoria real, la cual es muy pequeña (30KBytes), para lograr tiempos de ejecución lo más rápidos posibles.

En cuanto al idioma utilizado en la programación de este trabajo es el "BASIC", que es lenguaje utilizado por el computador Tektronix 4051, cuyas características se discutirán más adelante.

3.2. Estructura de la Biblioteca de Programas.

El equipo de computación que se ha utilizado es el Tektronix 4051, disponible en el Laboratorio de Control de esta Facultad, Este equipo tiene dos unidades de disco una impresora y un sistema de grafización (plotter). La capacidad de memoria que se tiene es de 30 KByte, la cual es muy pequeña si se tiene en cuenta de que se va a trabajar con matrices o vectores de muchos elementos. Por esta razón se ve la necesidad de separar el problema en tres programas diferentes, para evitar tener toda la memoria ocupada sólo por el listado de los programas. Estos tres programas van ser manejados por un programa de control o programa maestro, el cual se va a ocupar de cargar el programa que se requiera en determinado momento, borrando previamente todo lo que este en la memoria a excepción del propio programa maestro.

Los tres programas en que se ha dividido el trabajo son:

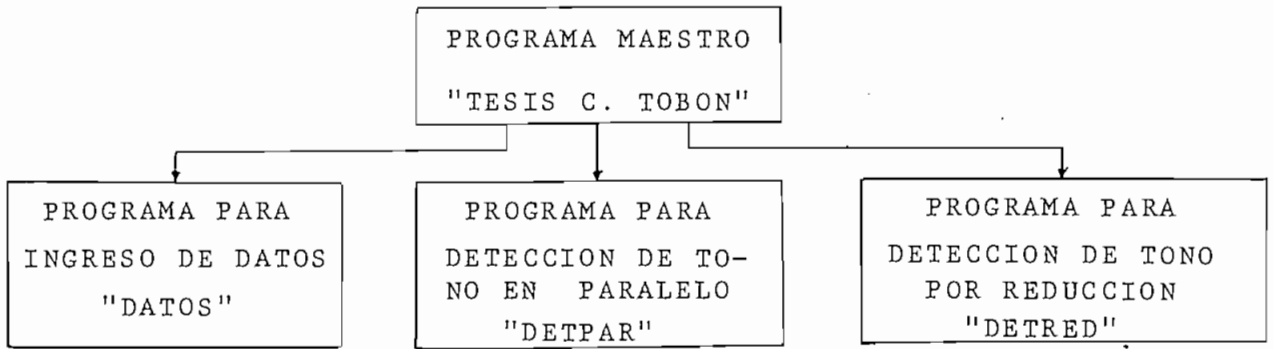
1. Programa de Ingreso de Datos
 - a) Ingreso Manual de las Muestras de Voz
 - b) Ingreso desde un Archivo de Disco
 - c) Ingreso por medio de Interpolación

2. Programa de Detección por Procesamiento en Paralelo.
3. Programa de Detección por Reducción de Información.

Estos tres programas van a estar siempre almacenados en un Floppy Disk de 630 KBytes de memoria y solo se cargarán en memoria, cuando sean requeridos por el usuario, por medio del programa maestro. Esto presenta un gran ahorro de memoria si se toma en cuenta que cada uno de estos programas ocupa aproximadamente 14 Kbytes, es decir mas de una tercera parte de la capacidad total de memoria real. Si se toma en cuenta que para realizar la detección de tono de un segmento de voz, se necesitarían el programa de ingreso de datos y uno de los programas para detección de tono; con estos dos programas en memoria se tendría prácticamente copada la capacidad de memoria del equipo, sin dejar más que unos 6 Kbyte para la ejecución de los programas. En cambio teniendo un programa a la vez, quedan unos 12 Kbytes para la ejecución del algoritmo.

A continuación se da la capacidad o espacio de memoria que ocupa cada programa:

| | |
|-------------------|-------------|
| 1. TESIS / CTOBON | 1214 Bytes |
| 2. DATOS | 14287 Bytes |
| 3. DETPAR | 18493 Bytes |
| 4. DETRED | 21940 Bytes |



Fif. 3.1 Estructura de la Biblioteca de Programa

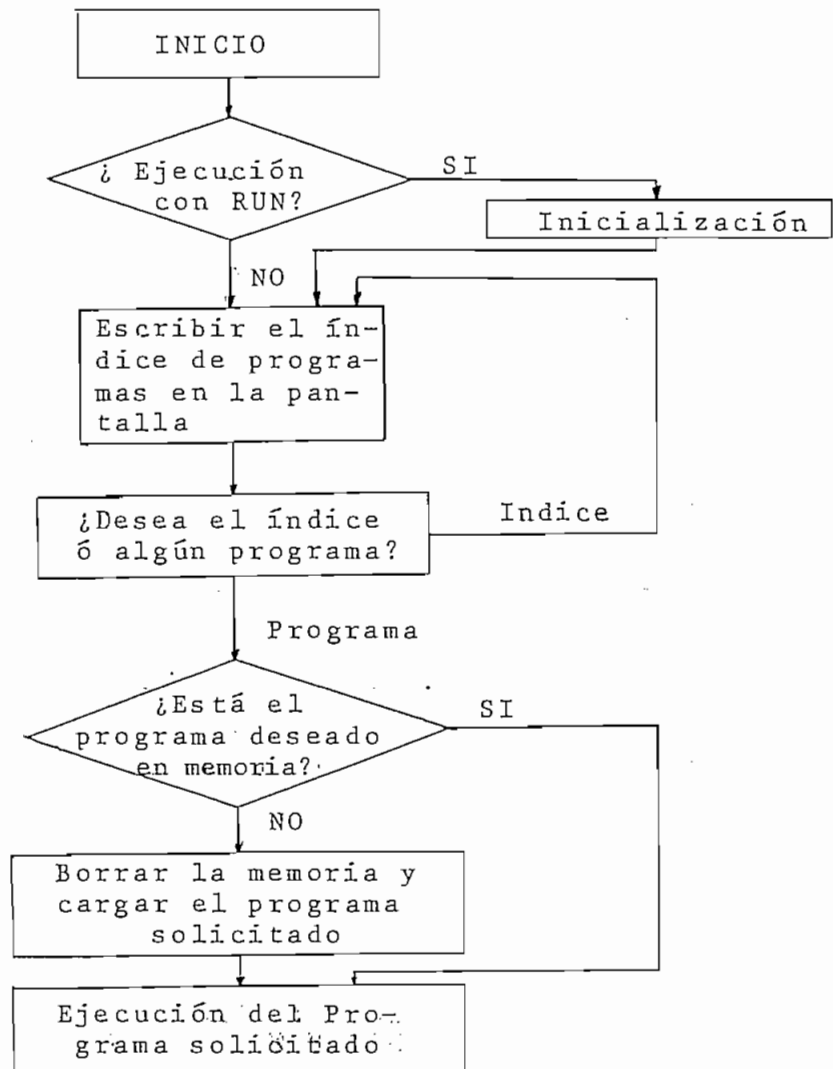


Fig. 3.2. Diagrama de flujo del programa maestro: TESIS/C. TOBON

En el disco de tesis, se encuentran almacenados, aparte de los programas ya mencionados, todos los archivos de trabajo que se han utilizado durante la ejecución de los programas. También están almacenados en este disco todos los archivos con los datos y valores de voz almacenados para la realización de las diferentes pruebas.

En cuanto al lenguaje "BASIC" utilizado, es un lenguaje de programación que brinda ciertas ventajas al usuario en relación con otro lenguaje como el "FORTRAN": En este lenguaje se tiene ventajas como la de dimensionar con variables, tener variables literales dimensionables, marcar los límites de los lazos con variables o expresiones, entre otras.

Al utilizar esta versión de "BASIC" hay que prestar especial atención en no utilizar en forma repetida variables del programa principal en las subrutinas, pues no existe independencia entre ellas, como sucede en otros lenguajes como por ejemplo el FORTRAN.

Hay que aclarar también, que cuando se está operando bajo el comando de un programa maestro, tampoco va a haber independencia entre las variables de los diferentes programas; ya que la carga de los diferentes programas se está realizando por medio de la instrucción "APPEND", la cual borra el listado del programa que haya estado en memoria a partir de una determinada línea de programa, pe-

ro sin alterar el contenido de las diferentes variables almacenadas en la memoria. Esta instrucción además realiza la carga del programa que se le haya ordenado, a partir del número de línea que se haya indicado.

En la Figura 3.1. se puede apreciar la configuración general de la biblioteca de programas, que permitirá la resolución de los algoritmos planteados en el capítulo anterior.

Debo aclarar que en este capítulo se explicarán los programas en forma concisa, poniendo énfasis en sus características principales. Para ello se presentará la metodología seguida, un diagrama de flujo y una lista de variables utilizadas con su respectivo significado. Para completar, se presentan en el APENDICE A todos los listados de los programas desarrollados en este trabajo.

El programa maestro que maneja automáticamente el funcionamiento de los demás programas de la biblioteca, se llama " TESIS / CTOBON".

Este programa funciona por medio del teclado del computador, borrando el programa que esté en memoria, para enseguida cargar otro en lugar del anterior. Esta operación se puede realizar cuantas veces se desee, siguiendo un orden lógico preestablecido. Las posibilidades que se tienen con este programa son:

TECLA 1 --- INDICE DE PROGRAMAS

TECLA 2 --- INGRESO DE DATOS

TECLA 3 --- DETECCION POR PROCESAMIENTO EN PARALELO.

TECLA 4 --- DETECCION POR REDUCCION DE INFORMACION.

Al oprimir la tecla 1, se llama el índice de programas, es decir la lista de programas disponibles. Las demás teclas sirven para llamar o cargar el respectivo programa que se indique en el índice.

Este programa realiza las siguientes funciones:

1. Chequea si está o no en memoria el programa que se requiera.
2. Si el programa está en memoria, va a ejecutarlo.
3. Si el programa no está en memoria, realiza los siguientes pasos:
 - a) Borra la memoria a partir de la línea 1010.
 - b) Carga el programa deseado a partir de la línea 1010.
 - c) Va a ejecutar dicho programa.

El programa maestro, sólo necesita inicializar ciertas variables para el funcionamiento de todos los programas, al ser ejecutado por primera vez.

En el proceso de borrado y carga de los diferentes programas hay que anotar que el programa maestro va a permanecer, casi en su totalidad, siempre en memoria. Esto permite que pueda ser llamado a ejecución en cualquier momento, ya sea por un programa o por el usuario.

Los nombres de las variables utilizadas en este programa y sus significados son:

| <u>NOMBRE</u> | <u>SIGNIFICADO</u> |
|---------------|---|
| UØ | Cantidad que sirve para inicializar el funcionamiento de este programa. |
| U1 | Número de la unidad de discos a ser utilizados. |
| 01 | Número del programa que se desea ejecutar |
| 02 | Número del programa que esté en memoria. |

En la Figura 3.2 se indica el diagrama de flujo del programa maestro " TESIS / CTOBON".

3.3. Programa para Ingreso de Datos (DATOS)

En el desarrollo de este programa se ha tratado de implementar un programa para diferentes propósitos, el cual pueda ser de mucha utilidad para su utilización de un futuro. Durante su desarrollo se han considerado situaciones como la de su sobredimensionamiento de la memoria, es decir la posibilidad de que se llegue a terminar la capacidad real de memoria. Esta situación puede ser muy común, si se considera que tan solo con una matriz o vector

2000 elementos se puede llegar a la capacidad máxima de memoria real, pues este programa de ingreso ocupa unos 14 Kbytes de memoria. Si se considera el resto de variables y las operaciones que tiene que realizar el programa, se llega al caso de saturación de la memoria.

En el análisis de segmentos de voz, se trabaja con segmentos de por lo menos 50 mseg, los cuales se muestrean a una frecuencia aproximadamente de 10 KHz. Esto representa un vector de 500 elementos, cada uno de los cuales va a ocupar 8 bytes. Esto significa que este vector ocuparía 4 Kbytes mas 18 bytes que utiliza para su definición interna. Como se puede apreciar la pequeña capacidad de memoria de este computador sería una limitación durante este trabajo, es por esto que en este programa se ha considerado un análisis de memoria en los casos que ha sido necesario.

Este programa de ingreso tiene tres opciones diferentes, estas son:

- 1) Ingreso manual de las muestras
- 2) Ingreso desde un archivo de disco
- 3) Ingreso por medio de interpolación.

En la Figura 3.3 se puede apreciar un diagrama de flujo en forma general del programa : DATOS".

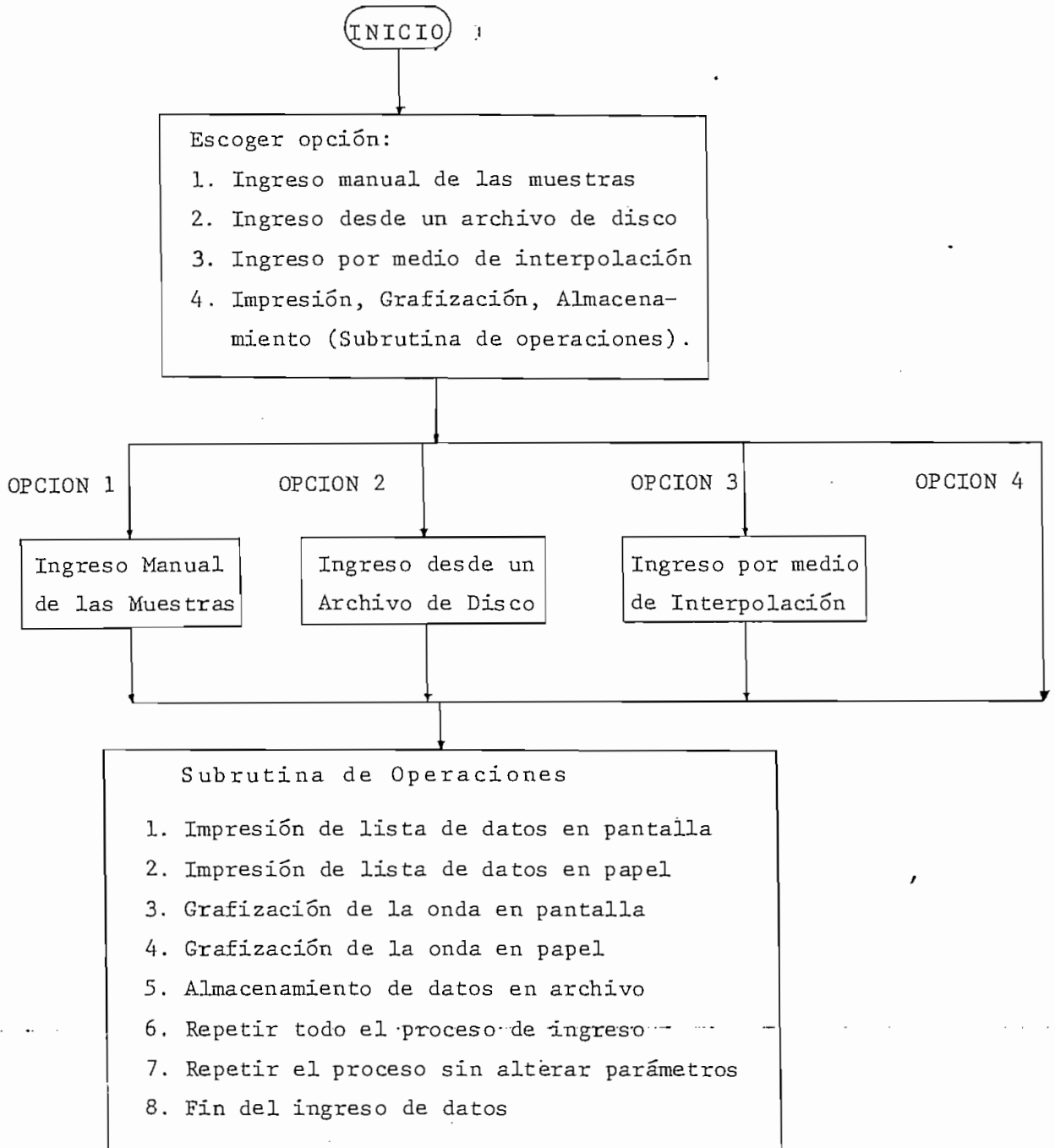


Fig. 3.3 Diagrama de flujo general del programa de ingreso de datos "DATOS".

Además, permite realizar las siguientes operaciones con las muestras de voz ya ingresadas:

- 1) Impresión de lista de datos en la pantalla
- 2) Impresión de lista de datos en papel
- 3) Grafización de la onda en pantalla
- 4) Grafización de la onda en papel
- 5) Almacenamiento de datos en archivo
- 6) Repetir el proceso de ingreso sin alterar parámetros.

Ingreso Manual de Muestras

El ingreso manual de los datos puede ser aconsejable en el caso de tener un segmento de voz bastante corto pues lo contrario éste es un trabajo bastante tedioso, si se toma en cuenta que se tendrá que analizar una gran variedad de segmentos de voz y probablemente para diferentes personas.

El usuario al escoger esta opción de ingreso, tendrá que disponer de la siguiente información que le será solicitada por el programa:

- 1) El número de muestras al ingresar
- 2) La frecuencia de muestreo en KHz.
- 3) Los valores numéricos de las muestras, normalizados entre - 10 y 10.

El programa se encarga entonces de analizar si la capacidad de memoria real disponible es suficiente para rea

lizar el respectivo ingreso. En caso de que no sea suficiente, se encarga de escribir en pantalla una indicación para redimensionamiento. Esto se debe a que no es práctico el ponerse a ingresar unos 2000 valores en forma manual. Para el caso de tener una gran cantidad de muestras es mucho más aconsejable el utilizar la opción de ingreso por interpolación, la cual es mucho más rápida.

En caso de que la capacidad de memoria sea suficiente, irán saliendo mensajes en la pantalla solicitando los valores numéricos de las muestras de voz en forma "secuencial", es decir en orden de ocurrencia en el tiempo. El ingreso "secuencial" de las muestras es una condición "necesaria" del programa. El ingreso se realizará hasta que salga un mensaje de "FIN DE INGRESO DE DATOS". En este instante el programa se encarga de almacenar las muestras de voz en un archivo de trabajo "CMUESTREO/A1" en el disco. Este archivo de trabajo sirve para almacenar los datos a utilizarse en la subrutina de operaciones, como en los programas para detección del tono del segmento de voz ingresado. Este almacenamiento se realiza para evitar tener que ocupar toda la memoria real con todas las muestras de voz, pues mucho más conveniente tener únicamente los valores necesarios en memoria en determinado instante.

Luego de almacenar los datos en disco, el programa va a la subrutina de operaciones donde se presenta un índice de las posibilidades de operación para el usuario. De esta

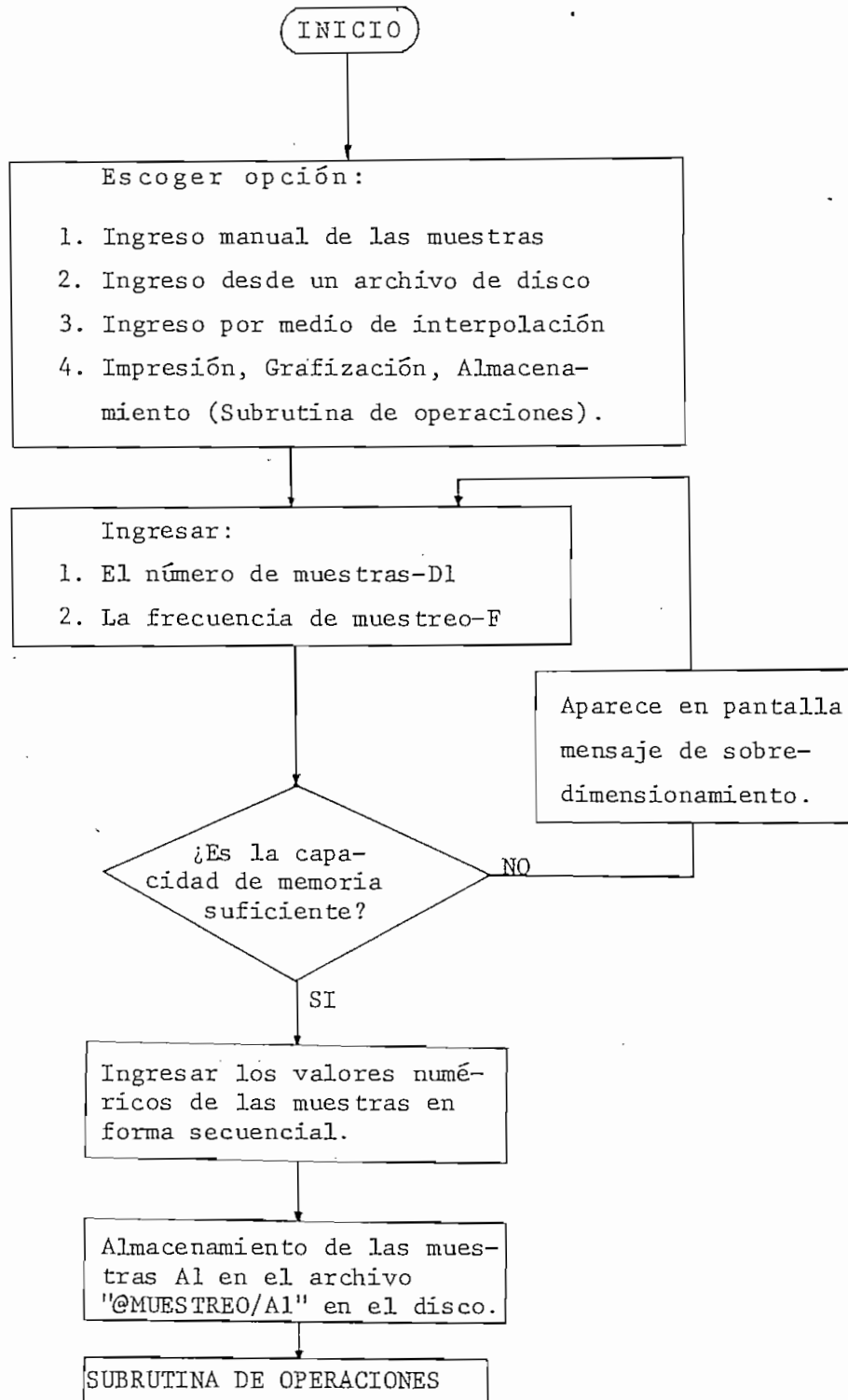


Fig. 3.4 Diagrama de flujo del programa de ingreso manual de las muestras

subrutina se hablará más adelante. En caso de ordenar desde esta subrutina un fin del programa de ingreso de datos, automáticamente saldrá el índice de programas dando la posibilidad al usuario de ordenar la ejecución de cualquiera de los programas.

Debo aclarar que en caso de que usuario no tenga más información que la longitud del segmento de voz en milisegundos, simplemente debe darse como dato una frecuencia de muestreo (típicamente 10 KHz) entre por lo menos unos 7 KHz y unos 60 KHz como máximo. Una frecuencia menor a 7KHz significaría un muestreo demasiado espaciado, el cual no aseguraría muy buenos resultados en la detección del tono del segmento de voz correspondiente. Con una frecuencia mayor a los 60 KHz se tendría una cantidad apreciable de datos que en vez de ayudar a conseguir mejores resultados, dificultaría las operaciones dentro del programa, aumentando considerablemente el tiempo de ejecución de los programas.

En la Figura 3.4 se puede apreciar el diagrama de flujo para el ingreso manual de las muestras.

Ingreso desde un Archivo en el Disco

Este tipo de ingreso de datos va a ser el más utilizado en la práctica, ya que en futuro cercano se van a tener almacenados en disco infinidad de señales de voz, como parte de la tesis que se está desarrollando para un sistema de adquisición de datos. Además, como parte de este trabajo de tesis se va a contar con un conjunto de archivos de datos bajo una biblioteca común llamada "@VOZ/".

Este tipo de ingreso de datos es muy útil, también, en caso de querer repetir la detección de tono para algún segmento de voz que se haya almacenado anteriormente en un archivo de disco.

El usuario al escoger este tipo de ingreso, deberá conocer la siguiente información:

- 1) El número de la unidad de disco en que esté el archivo.
- 2) El nombre del archivo en el disco (opcional).
- 3) El número de muestras en el archivo (opcional).
- 4) La frecuencia de muestreo (opcional).

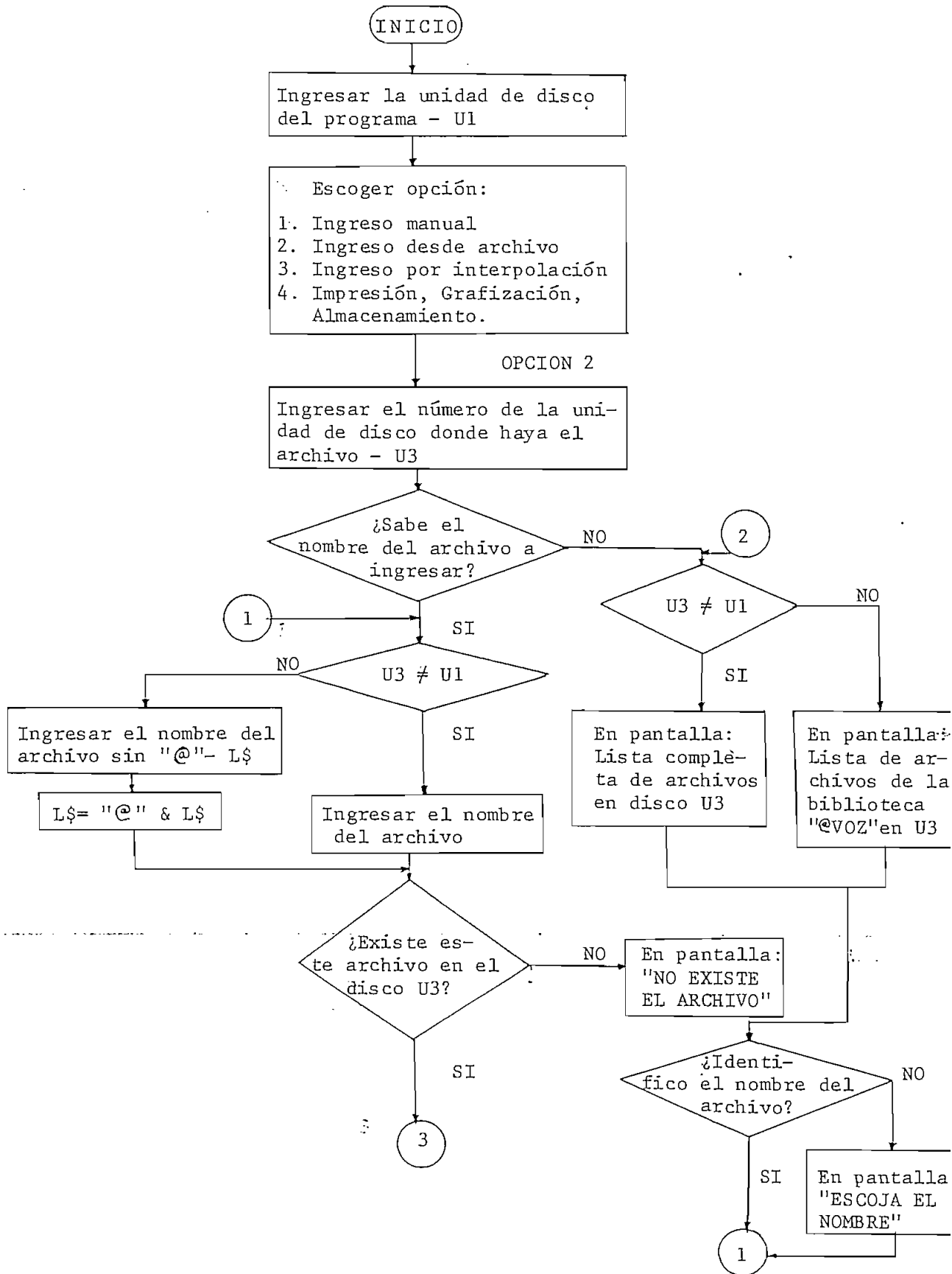
En caso de conocerse a ciencia cierta el nombre del archivo, el programa hará que aparezca una lista completa de los archivos en el disco seleccionado. Si el disco o unidad de disco seleccionada es la que tiene los programas, es decir el trabajo de tesis, aparecerá una lista de todos los archivos del conjunto o biblioteca "@VOZ/". De lo contrario, aparecerá una lista completa de todos los archivos en el correspondiente disco. De esta forma se podrá identificar o seleccionar un archivo de voz para realizar la detección del tono.

En caso de haber seleccionado un archivo de voz en el disco de la tesis, hay que anotar que todos estos archivos tiene almacenados como primeros dos datos el número de muestras en el archivo y la frecuencia a la que fueron muestreadas, respectivamente. Si en cambio se selecciona algún otro disco con archivos con muestras de voz, habría nece-

sariamente que saber el número de muestras en el archivo y la frecuencia a la que hayan sido muestreadas, a no ser que el archivo esté con la misma notación utilizada durante el desarrollo de esta tesis.

Esta claro que en caso de utilizar algún disco con muestras de voz, las muestras dentro del archivo a utilizarse deben encontrarse almacenadas en forma "secuencial".

Luego de este proceso de identificación del archivo de datos a ingresarse, el programa procede a analizar el número de elementos del archivo que pueda cargar o almacenar simultáneamente en la memoria real. Si este número es mayor o igual al número de elementos en el archivo, es decir a la dimensión del vector de datos "A1", se colocan las banderas $D4=1$ y $D5=1$ que le indicarán al programa, durante su desarrollo, el tipo de funcionamiento con el que debe operar. De lo contrario, las banderas se pondrán así: $D4=0$ y $D5=0$. La bandera $D4$ le indica al programa, durante todo su funcionamiento, que hay que trabajar por etapas realizando transferencias entre el disco y la memoria. De esta forma se va a ir calculando o realizando el programa por partes y al culminar cada parte se almacenará la información requerida en un archivo de trabajo en el disco. Este archivo es el "@MUESTREO/A1", que no sólo realiza esta función de almacenamiento de datos parciales, sino que sirve para realizar la transferencia de los datos a los programas para la detección del tono de la voz.



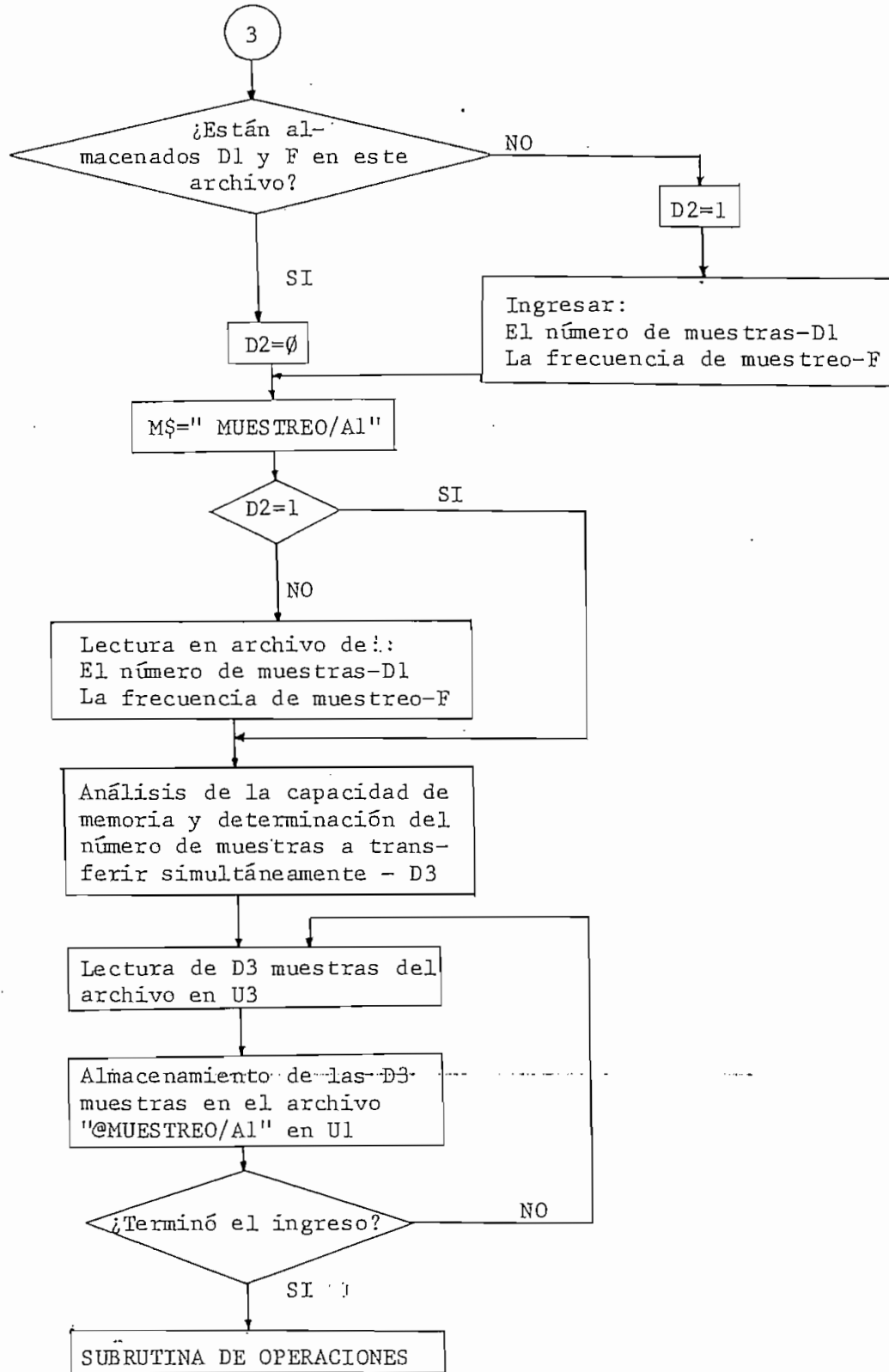


Fig. 3.5 Diagrama de flujo del programa de ingreso desde un archivo en disco.

La bandera D5 sirve para indicar la culminación del proceso de calculos-almacenamiento en archivo, es decir marca la culminación de las transferencias de información entre la memoria real y el disco para un determinada tarea.

Cuando las banderas son iguales a cero, también se van a guardar los datos en el archivo de trabajo "@MUESTREO/A1". La primordial diferencia entre estos dos modos de operación radica en el tiempo de ejecución, pues entre más transferencias haya que realizar más demorada va a ser la ejecución del programa.

En fin este programa de ingreso de datos desde un archivo de voz en el disco, se reduce a la lectura de los datos desde el archivo donde se encuentren, para almacenarlos o escribirlos en el archivo de trabajo "@MUESTREO/A1". Al terminar esta tarea el programa converge a las subrutina de operaciones, donde por medio de un índice que aparece en la pantalla, se le brinda al usuario la posibilidad de escoger el tipo de operación que éste desee.

En la Fig. 3-5 se puede apreciar el diagrama de flujo correspondiente al ingreso de datos desde un archivo en disco. Debo anotar que para este tipo de funcionamiento, se puede tener una gran cantidad de datos que únicamente estaría limitada por la capacidad de memoria del disco. Otra limitación puede ser también, el tiempo que tomaría este programa en procesar una cantidad muy grande de datos.

Ingreso por medio de Interpolación

Durante el desarrollo de este trabajo de tesis, se ha visto la necesidad de implementar un programa para ingreso de datos, que facilite el ingreso de gran cantidad de valores. Una forma sería el ingresar los datos desde un archivo en disco, el problema es que en el momento no se dispone de esta facilidad. La otra posibilidad es el ingresar manualmente las muestras de voz, utilizando gráficos de señales de voz. Esto resulta muy poco eficiente ya que hay que ampliar estos gráficos para poder realizar un muestreo a unos 10 KHz, lo que significaría que en un segmento de unos 50mseg ampliado a una longitud de 1 metro, habría que leer 500 muestras separadas entre sí por 2 milímetros. Como se puede ver este trabajo representaría horas de mediciones para tener una lista de valores numéricos de las muestras de un gráfico de una señal de voz.

Luego de haber realizado este trabajo gráfico, habría que ingresar manualmente el listado de los valores numéricos de las muestras, lo cual representaría otra buena cantidad de tiempo. Además de estos inconvenientes, hay que tomar en cuenta que este conjunto de mediciones ópticas tendrían ciertos errores.

Por estos motivos se ha desarrollado un programa para ingresar datos por medio de la técnica de interpolación. Esta técnica de interpolación consiste en tomar únicamente la

información estrictamente necesaria de un gráfico, para ingresarla al computador y por medio de un programa de interpolación, calcular todos los puntos de la curva que sean de interés para el presente algoritmo.

Analizando las diferentes curvas que se tiene a disposición para este trabajo, se pudo observar que la característica predominante en estos gráficos es poligonal. Es decir que la mayoría de estas curvas se pueden clasificar como "curvas poligonales", razón por la cual se ha utilizado una "interpolación lineal" para la obtención de los datos deseados.

Para poder realizar la interpolación lineal de las muestras de voz, es necesario ingresar todos aquellos puntos de quiebre de la curva, es decir los máximos y los mínimos de la curva, además de los puntos de inflexión o cambio de la curvatura. Todos estos puntos de quiebre deben ser ingresados, ya sea manualmente o por medio del grafizador o plotter del sistema de computación.

Para este ingreso de datos para la interpolación se han considerado dos posibilidades: el ingreso manual de las coordenadas de los puntos de quiebre y el ingreso de las coordenadas de los puntos de quiebre por medio del grafizador. En el primer caso, se tiene que realizar un trabajo gráfico previo para encontrar los valores numéricos a ingresar para la interpolación. En el segundo caso, basta con disponer del gráfico de la curva, pues por medio de un visor o lu-

pa que tiene el grafizador, se pueden ir ingresando los puntos de interés, es decir los puntos de quiebre. Para este segundo caso, se ha desarrollado un programa de ingreso, por medio del cual solo hay que ir ubicando los puntos de quiebre con ayuda del visor del grafizador y oprimiendo la tecla de RETURN. Para iniciar este ingreso hay que colocar previamente los límites físicos del gráfico por medio de las teclas de SET del grafizador o plotter.

Los límites físicos que se deben fijar con las teclas SET, deben corresponder a un valor menor al mínimo valor de la curva (límite inferior) y a un valor mayor al máximo valor (límite superior) de la curva a ingresarse. Para que la curva sea simétrica con respecto al origen de coordenadas, las amplitudes de estos dos límites deben ser iguales en valor absoluto.

El usuario en caso de escoger el ingreso de datos por medio de interpolación con ingreso manual de las esquinas o puntos de quiebre, tendrá que tener la siguiente información:

- 1.) La amplitud para el tiempo cero (valor inicial)
- 2.) La amplitud para el tiempo L_1 , donde L_1 es igual a la longitud del segmento de voz en milisegundos (valor final).
- 3.) La lista con las coordenadas de las esquinas o puntos de quiebre de la curva.
- 4.) Conocer el número total de esquinas que vaya a ingresar, incluyendose el valor inicial y

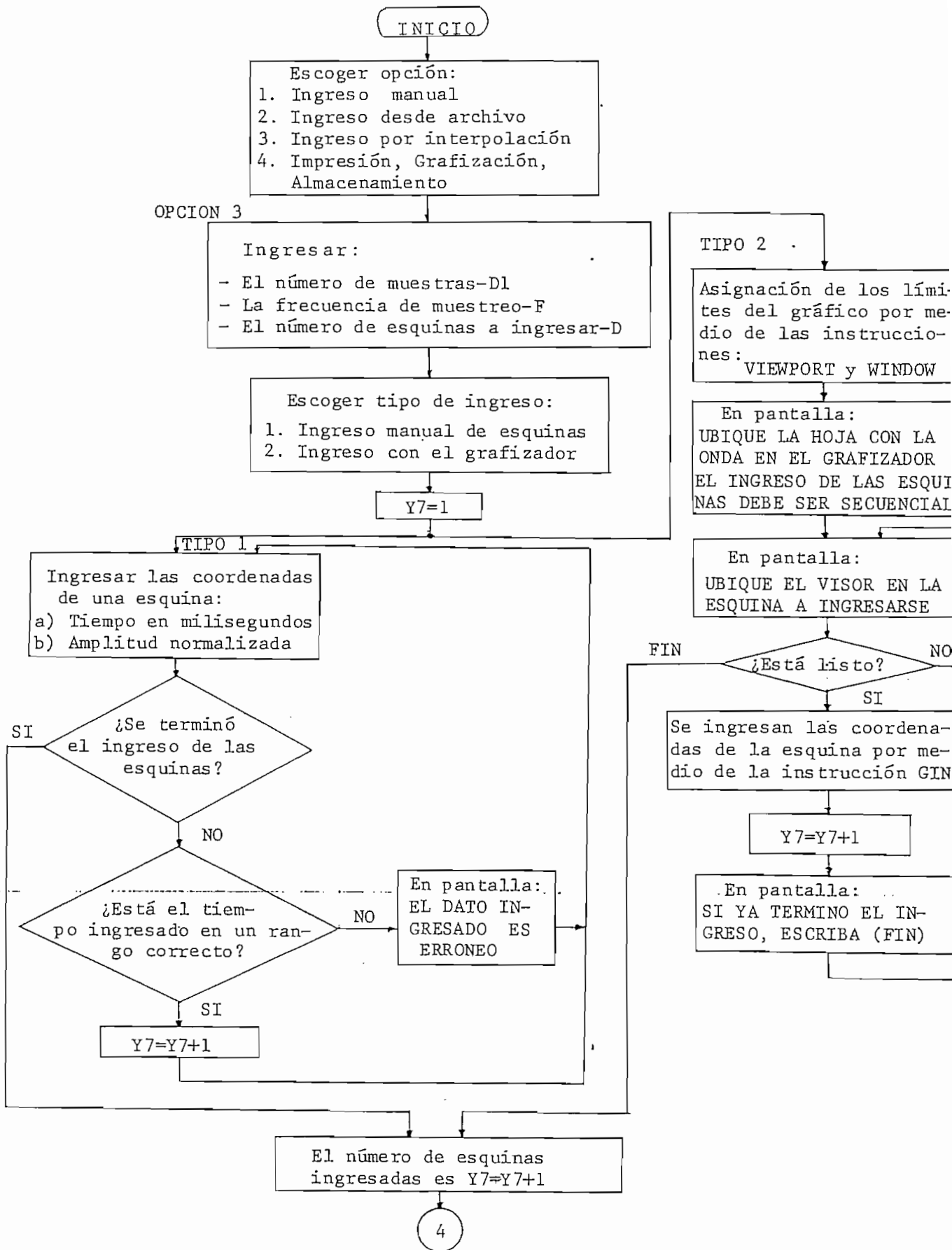
el valor final.

- 5) El número de muestras a interpolarsen.
- 6) La frecuencia a la que se desee muestrear la onda.

Después de obtener toda esta información el programa procederá a un redimensionamiento de los vectores con las coordenadas de las esquinas, en caso de que se hayan sobredimensionado por el usuario. Con estos nuevos vectores se hará la interpolación, después de haber realizado el análisis de la capacidad de memoria disponible. Tras terminar cada etapa de interpolación que se hagan necesarias, el programa procederá a almacenar las muestras de voz en el archivo de trabajo "MUESTREO/Al", para finalmente ir a la subrutina de operaciones.

El usuario para poder utilizar el programa de ingreso de datos por medio de interpolación, con el ingreso de las esquinas con el grafizador, debe tener la siguiente información:

- 1) Disponer de un gráfico claro de la curva en papel.
- 2) Conocer un número mayor o igual al número de esquinas a ingresarse.
- 3) Tener señalados en la hoja del gráfico los límites inferior y superior de la curva, para fijar el SET.
- 4) El número de muestras a interpolarsen.



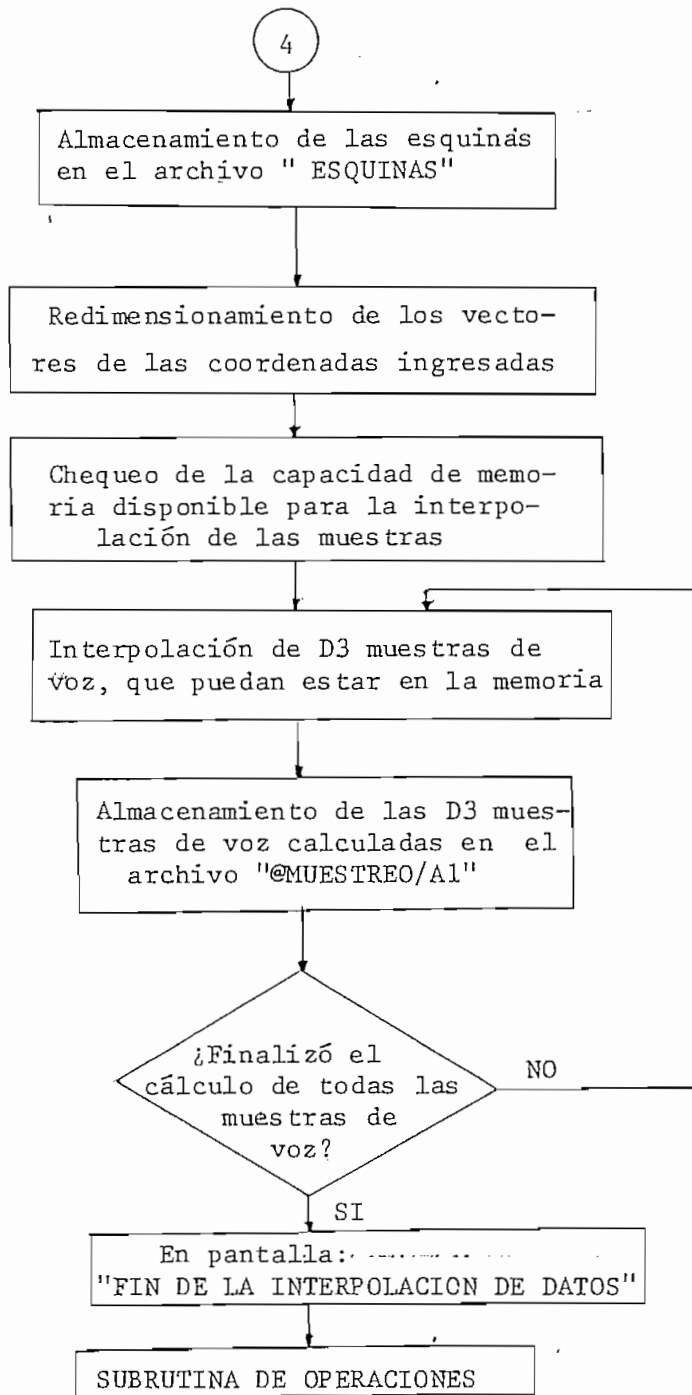


Fig. 3-6 Diagrama de flujo del programa de ingresos de datos por medio de interpolación.

5) La frecuencia a la que se desee muestrear la onda.

La forma de operación de este programa es similar a la descrita anteriormente para el otro caso de interpolación.

Para estos dos casos de ingreso por medio de interpolación debe anotarse, que en caso de que las ondas a ingresarse no sean de tipo poligonal, habría que ver si tomando más valores intermedios se lograría una buena interpolación. De lo contrario habría que utilizar otro tipo de interpolación como la "seccional cúbica".

Una condición necesaria para todos los tipos de ingreso de datos descritos anteriormente, es que el ingreso se realice en forma secuencial. Para el ingreso manual de las esquinas para la interpolación, debo aclarar que las amplitudes deben estar normalizadas entre -10 y 10, para fines gráficos. En la Fig. 3.6 se puede apreciar el diagrama de flujo para el ingreso de datos por medio de la interpolación seccional lineal.

Subrutina de Operaciones

Por medio de esta subrutina se ha tratado de brindar al usuario el siguiente conjunto de facilidades, las cuales le aparecerán en pantalla al finalizar cualquier tipo de ingreso de datos, así:

1) Impresión de lista de datos en pantalla

- 2) Impresión de lista de datos en papel
- 3) Grafización de la onda en pantalla
- 4) Grafización de la onda en papel
- 5) Almacenamiento de datos en archivo
- 6) Repetir todo el proceso de ingreso
- 7) Repetir el proceso sin alterar parámetros
- 8) Fin del ingreso de datos

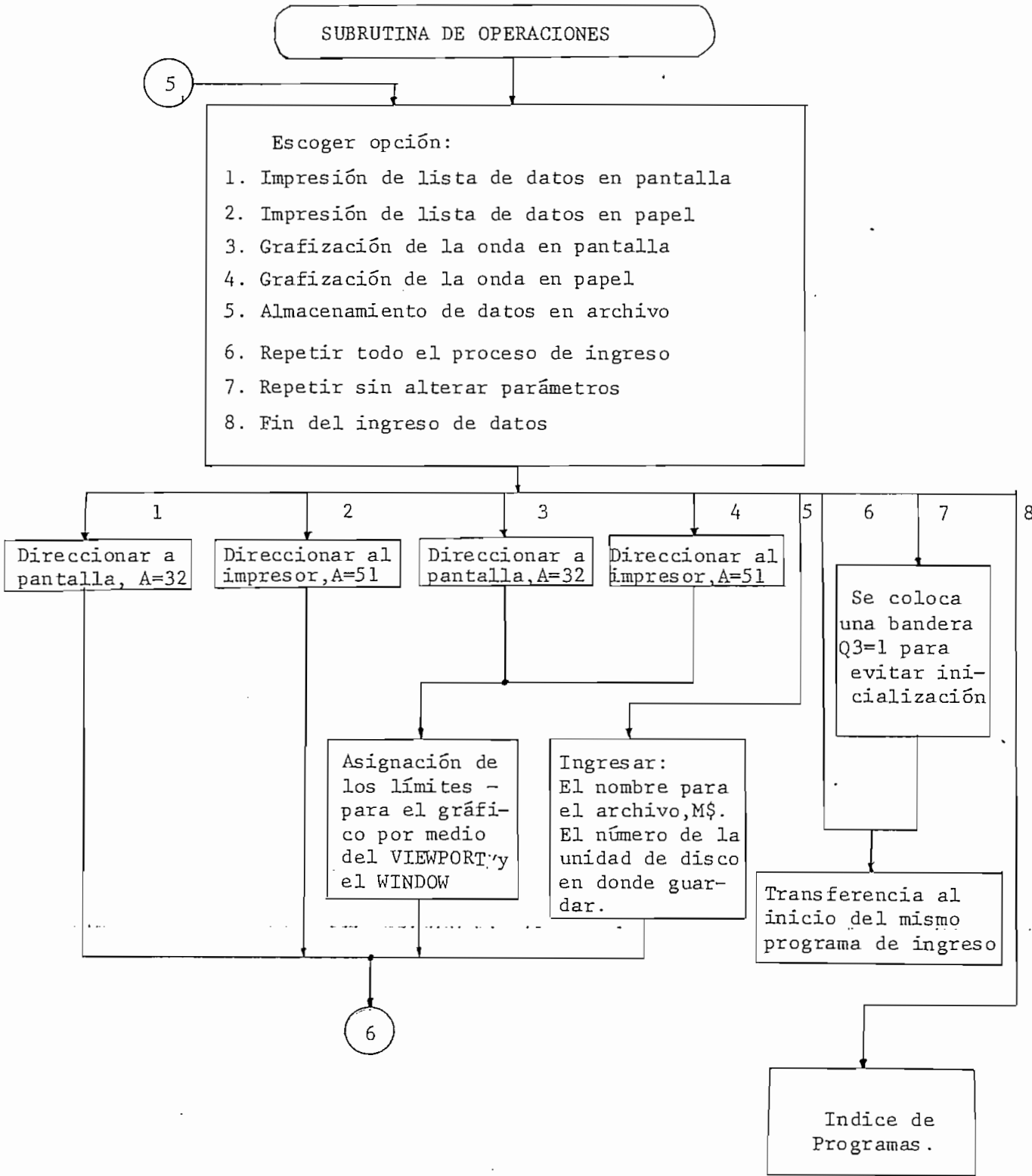
Las técnicas de programación utilizadas para el desarrollo de estos programas no se van a analizar dentro de esta tesis. Para poder tener una buena apreciación de la onda de voz a ser analizada, se dispone ya sea de una lista de valores numéricos de la curva o de su gráfico. En caso de que el usuario desee un gráfico de la onda de voz, éste puede ser total o de algún intervalo de tiempo de especial interés. Esto permite segmentar en varios gráficos una onda para una mejor apreciación. Cada gráfico va a estar identificado por un nombre o título que debe ingresar el usuario en el momento que le sea solicitado por el computador. Además en la parte inferior de cada gráfico aparecerá escrita ~~la longitud o duración total del segmento de voz grafizado.~~

Debo anotar que para grafizar las diferentes ondas utilizadas para las pruebas de los dos algoritmos de tono, se ingresaron todos los valores numéricos correspondientes a las amplitudes de las curvas en forma normalizada, entre -10 y 10. Se realizó esta normalización de las curvas para tener una mejor apreciación de la magnitud de las diferen-

tes letras con respecto a las demás.

El numeral 5 de la subrutina de operaciones, permite al usuario el almacenamiento de los datos ingresados en un archivo, al cual él debe dar un nombre determinado para su futura identificación. El usuario debe tomar en cuenta que cuando los programas de ingreso almacenan automáticamente los datos en el archivo de trabajo "@MUESTREO/A1", es sólo para su utilización interna; pues al realizar un nuevo ingreso se va a alterar el contenido anterior del archivo. Debo anotar que el nombre que el usuario desee dar al archivo debe ser de un máximo de 10 caracteres, comenzando necesariamente con una letra. Este nombre puede estar conformado por letras, números y el separador "/"; ningún otro caracter puede ser utilizado dentro del nombre de un archivo. Este programa hace que cualquier nuevo archivo de datos ingrese a la biblioteca "@VOZ/", razón por la cual el nombre completo del archivo quedará precedido por el nombre de la biblioteca de archivo de voz "@VOZ/", para facilitar la identificación de cualquier archivo de datos de voz a cualquier usuario en el futuro. Por ejemplo, si se quiere identificar a un segmento de voz con el nombre "LETRA/J", entonces el archivo que contenga estos datos tendrá el nombre "@VOZ/LETRA/J".

Los numerales 6, 7 y 8 no son más que transferencias a diferentes partes del programa, para la realización de un trabajo solicitado por el usuario. Estas transferencias permiten la repetición de un determinado ingreso o la culmi-



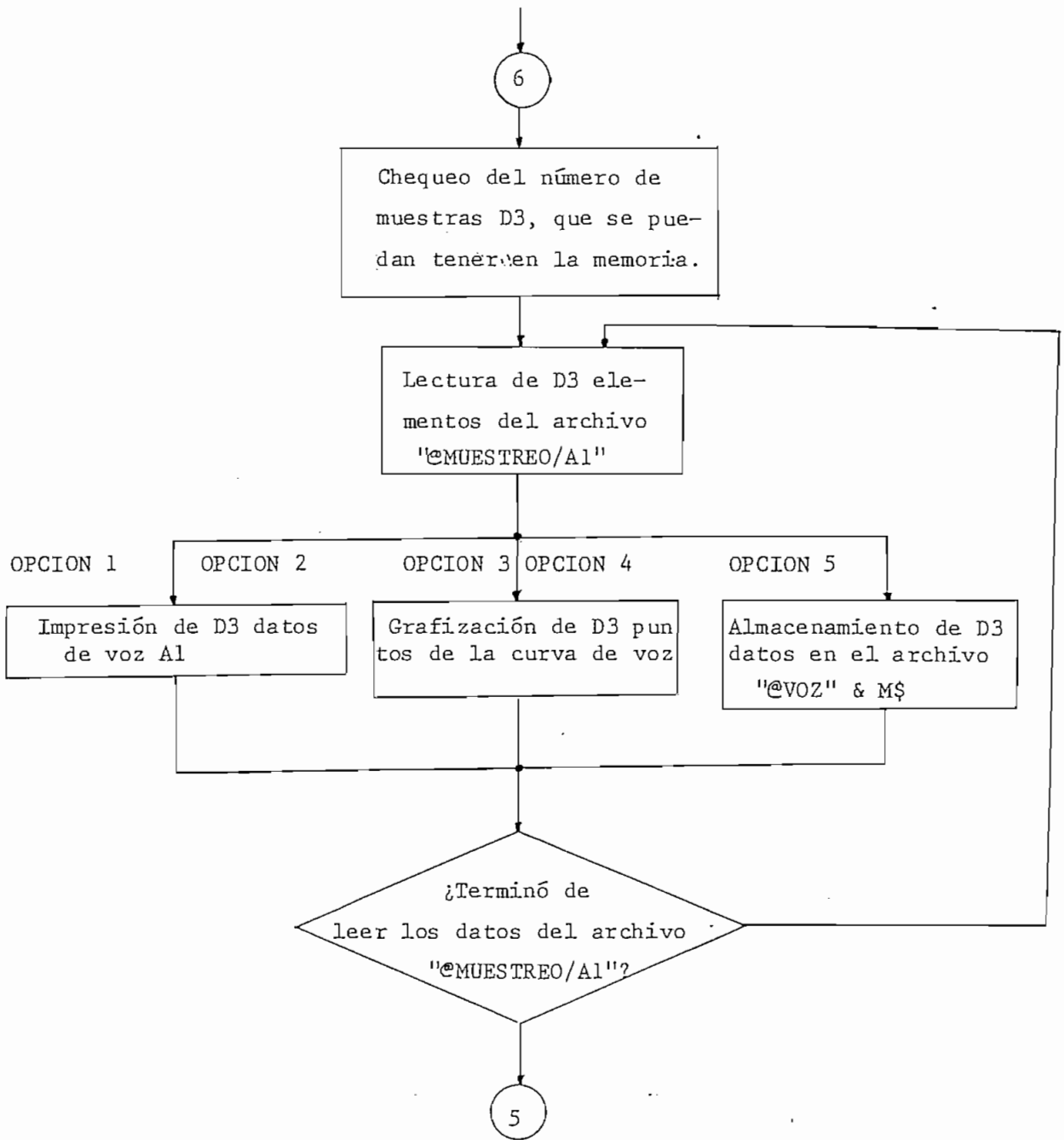


Fig. 3.7 Diagrama de flujo general de la subrutina de operaciones del programa de ingreso "DATOS".

nación del programa de ingreso de datos. Al culminar el programa de ingreso se realiza una transferencia automática al índice de programas, para que el usuario tenga la posibilidad de escoger el siguiente programa a ejecutarse.

En la Fig. 3.7 se puede apreciar el diagrama de flujo correspondiente a la subrutina de operaciones del programa "DATOS". En caso de que se necesite ver las características de estos programas de ingreso en forma más detallada, se debe chequear el estado del programa "DATOS", el cual se encuentra entre la página 2 y la página 10 del APENDICE "A".

A continuación se brinda una lista de las principales variables utilizadas en este programa y sus significados.

| <u>NOMBRE</u> | <u>SIGNIFICADO</u> |
|---------------|---|
| O1 | Número del programa que se desea ejecutar. |
| O2 | Número del programa que está en memoria. |
| Q | Número del tipo de ingreso deseado. |
| K4 | Capacidad de memoria reservada para operaciones simples. |
| A1 | Vector de las muestras de voz. |
| U3 | Número de la unidad de disco donde está el archivo deseado. |
| U1 | Número de la unidad de disco en la que está el programa. |
| M\$ | Nombre del archivo en el que se van a almacenar datos. |
| L\$ | Nombre del archivo del que se van a leer datos. |
| D2 | Bandera que indica si están almacenados en un archivo |

NOMBRE

SIGNIFICADO

| | |
|------------|---|
| | el número de muestras y la frecuencia de muestreo. |
| Q2 | Número que indica el tipo de ingreso de datos para la interpolación. |
| A | Dirección de la unidad deseada (impresor, grafizador, pantalla). |
| D | Valor mayor o igual al número de esquinas a ingresarse. |
| X4 | Coordenada de tiempo del punto inferior para la interpolación seccional lineal. |
| Y4 | Coordenada de amplitud del punto inferior para la interpolación seccional lineal. |
| X5 | Coordenada de tiempo del punto superior para la interpolación seccional lineal. |
| Y5 | Coordenada de amplitud del punto superior para la interpolación seccional lineal. |
| X6 | Valor a interpolarse entre los valores X4 y X5. |
| P | Vector con las coordenadas de tiempo de las esquinas. |
| P1 | Vector con las coordenadas de amplitud de las esquinas. |
| P2 | Valor de la pendiente de la recta de interpolación. |
| Y7 | Número de esquinas ingresadas. |
| Q1 | Número del tipo de operación deseada. |
| Q3 | Bandera para repetir el proceso de ingreso sin alterar parámetros. |
| QØ | Número del último tipo de ingreso utilizado. |
| X1 | Límite inferior de tiempo para el gráfico. |
| X2 | Límite superior de tiempo para el gráfico. |
| Y1 | Límite inferior de la ventana para fines gráficos. |

| <u>NOMBRE</u> | <u>SIGNIFICADO</u> |
|---------------|--|
| Y2 | Límite superior de la ventana para fines gráficos. |
| XØ | Magnitud horizontal de un caracter en unidades gráficas. |
| YØ | Magnitud vertical de un caracter en unidades gráficas. |
| D1 | Número de muestras de voz. |
| F | Frecuencia de muestreo en KHz . |
| L1 | Longitud del segmento de voz en milisegundos. |
| C1 | Cantidad de memoria que ocupa el vector A1. |
| U2 | Número de la unidad de disco en la que se quiere almacenar un segmento de voz. |
| D3 | Número de elementos que se puede procesar simultáneamente en la memoria. |
| D5 | Bandera que indica la finalización o no de las transferencias disco-memoria o viceversa. |
| D4 | Bandera que indica la necesidad o no de realizar transferencias entre la memoria y el disco. |
| N | Contador del número de transferencias. |

3.4. Programa para la Detección del Tono por
el Procesamiento en Paralelo (DETPAR)

Los fundamentos básicos utilizados para el desarrollo de este programa son los vistos en el capítulo 2, referentes al método de detección de tono planteado por Rabiner y Schafer. Por medio de este programa se ha tratado de desarrollar un programa que permita trabajar en la mejor forma posible, de acuerdo a las posibilidades del sistema de computación

utilizado. Cabe anotar que este algoritmo se desarrolló por sus autores sin considerar las grandes limitaciones de memoria real y de velocidad que tiene el sistema de computación del Laboratorio de Control.

Debo aclarar que durante la realización de este trabajo he desarrollado algunos mecanismos necesarios para la optimización en el uso de la memoria real de este computador. Para ello se ha hecho necesario estar evaluando permanentemente la capacidad de memoria disponible para un proceso determinado, para establecer las etapas en que se pueda realizar, el número de elementos que se puedan procesar simultáneamente y las transferencias a disco que se hagan necesarias, ya sea para la lectura o escritura de datos o resultados. Estas continuas transferencias a disco que se han hecho necesarias para la implementación de este algoritmo, han repercutido notoriamente en el tiempo de ejecución del programa, pues la velocidad de la unidad periférica de discos es mucho más lenta que la del procesador del computador.

A continuación se brinda una lista de los pasos que se han seguido para la implementación del algoritmo para la detección por procesamiento en paralelo.

1. Inicialización de las diferentes constantes y banderas necesarias durante el desarrollo de este programa.
2. Análisis de la capacidad de memoria real disponible, para determinar si se pueden tener todos los elementos del archivo de trabajo que contiene los datos ("@MUESTREO/A1") en la

memoria, dejando espacio para las matrices que contendrán las coordenadas de los picos (máximos) y los valles (mínimos) del segmento de voz.

3. Lectura de los datos de archivo que se puedan contener en memoria para la determinación de los picos y valles de la señal.
4. Determinación de los picos y valles de la señal. Estos valores se van colocando en matrices en forma de coordenadas, para no tener vectores prácticamente llenos de ceros para los valores que no sean ni picos ni valles.
5. Almacenamiento de los picos y valles determinados para los datos leídos del archivo. Los vectores A2, A3, A4 y A5 correspondientes a las coordenadas de los picos y los valles de la señal se almacenan en los archivos "@POSICION/PICOS" y "@POSICION/VALLES", respectivamente.
6. Se chequea si ya se han leído todos los datos de voz del archivo de trabajo, Si no se han leído todos, se repite el proceso desde el numeral 3. Si en cambio ya se terminó de leer los datos y por lo tanto de determinar los picos y los valles, se continúa con el siguiente paso.
7. Se realiza la impresión en pantalla o en papel de los máximos y mínimos de la señal, si ha sido solicitada por el usuario.
8. Análisis de la capacidad de memoria real disponible para determinar cuantos picos y valles pueden tenerse simultáneamente en memoria, para la generación de los trenes de impulsos m_1 , m_2 , m_3 , m_4 , m_5 y m_6 .

9. Lectura de los picos y valles de los archivos que se puedan contener en la memoria, para la generación de los trenes de impulsos.
10. Cálculo de los seis trenes de impulsos de acuerdo a las siguientes expresiones::

$$M1(I) = \text{ABS}(A2(I))$$

$$M2(I) = A2(I) - N1(2)$$

$$M3(I) = A2(I) - N1(1)$$

$$M4(I) = \text{ABS}(A3(I))$$

$$M5(I) = -A3(I) + N1(1)$$

$$M6(I) = -A3(I) + N1(2)$$

Donde: si $M3(I)$ menor a cero, implica que $M3(I) = \emptyset$

si $M6(I)$ es menor a cero, entonces $M6(I) = \emptyset$

El significado de estas variables puede verse en la lista de variables para este programa.

11. Almacenamiento de los vectores M1, M2, M3, M4, M5 y M6 en los archivos "@TREN/UNO", "@TREN/DOS", "@TREN/TRES", "@TREN/ CUATRO", "@TREN/CINCO" y "@TREN/ SEIS", respectivamente.
12. Se chequea si ya se han leído todos los picos y valles de los archivos respectivos en el disco. Si no se han leído todos, se repite el proceso desde el numeral 9. Si ya se leyeron todos los picos y valles, se puede continuar con el proceso.
13. Análisis de la capacidad de memoria disponible, para determinar cuántos de los trenes de impulsos pueden ser procesados para su estimación del período de tono, dejando la memoria necesaria para los cálculos y resultados.

14. Lectura de archivos del número de trenes de impulsos que se puedan tener simultáneamente en la memoria para su estimación del período de tono.
15. Estimación de los períodos de tono para cada tren de impulsos de acuerdo al algoritmo planteado en el capítulo anterior.
16. Impresión de los valores numéricos de cada tren de impulsos y de sus respectivas estimaciones del período de tono, si a sido solicitada por el usuario. Esta impresión puede ser en pantalla o papel.
17. Almacenamiento de las estimaciones de tono para cada tren de impulsos en el archivo "@ESTIMACION/TONO", en forma secuencial.
18. Se repite este proceso desde el numeral 14 hasta que se hayan calculado las estimaciones del período de tono, para todos los trenes de impulsos.
19. Análisis de la capacidad de memoria disponible, para determinar la cantidad de estimaciones del período de tono que se puedan tener simultáneamente en la memoria, para la realización del análisis estadístico del número de coincidencias.
20. Determinación del número de coincidencias entre las estimaciones del período de tono, de acuerdo al algoritmo de procesamiento en paralelo del capítulo 2.
21. Impresión de los períodos de tono finales y del tiempo de ejecución del programa de detección del tono.

Para poder realizar un profundo análisis de las téc-

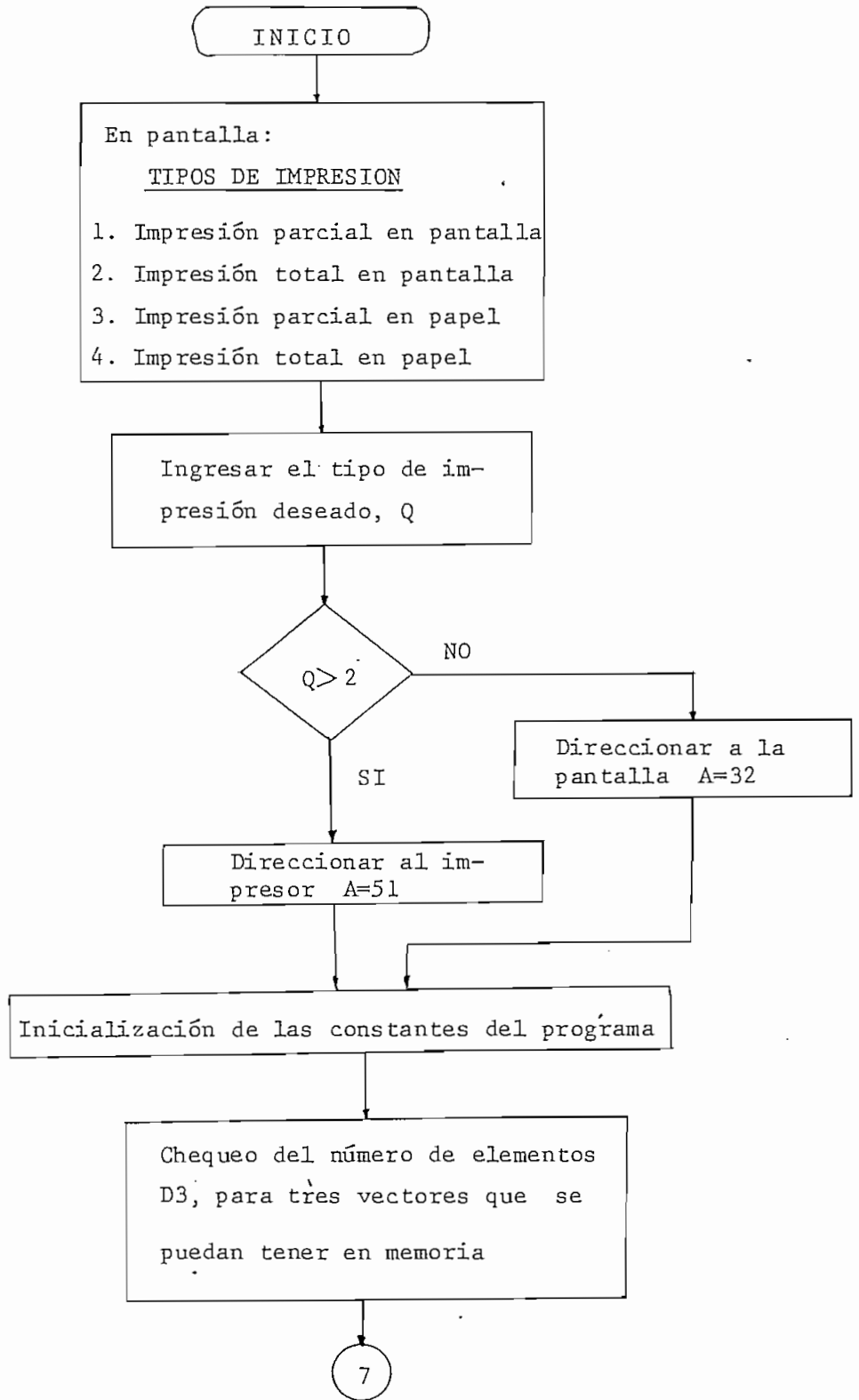


Fig. 3.8 Diagrama de flujo del programa de detección de tono por procesamiento en paralelo "DETPAR".

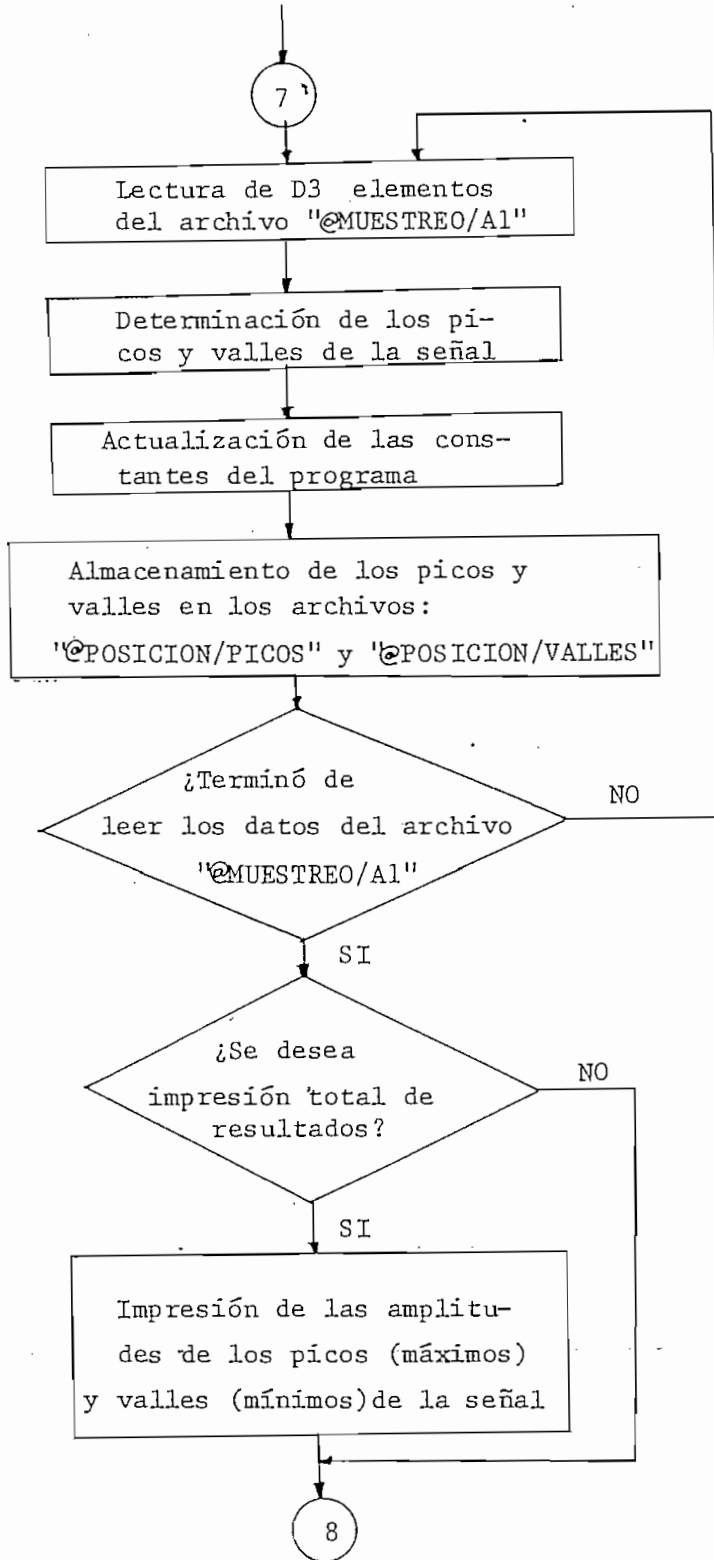


Fig. 3.8 Diagrama de flujo del programa de detección de tono por procesamiento en paralelo "DETPAR".

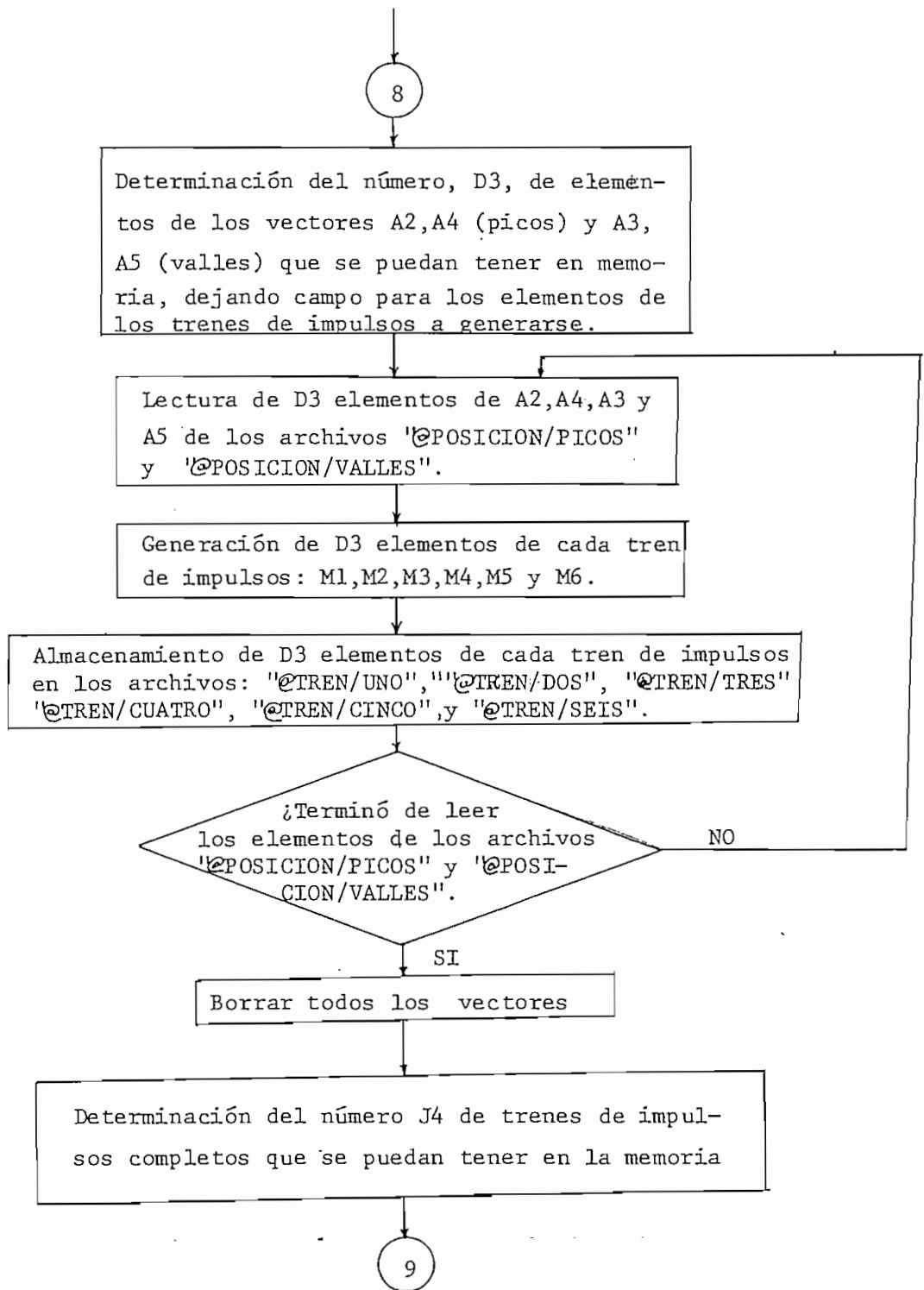


Fig. 3.8 Diagrama de flujo del programa de detección de tono por procesamiento en paralelo "DET PAR".

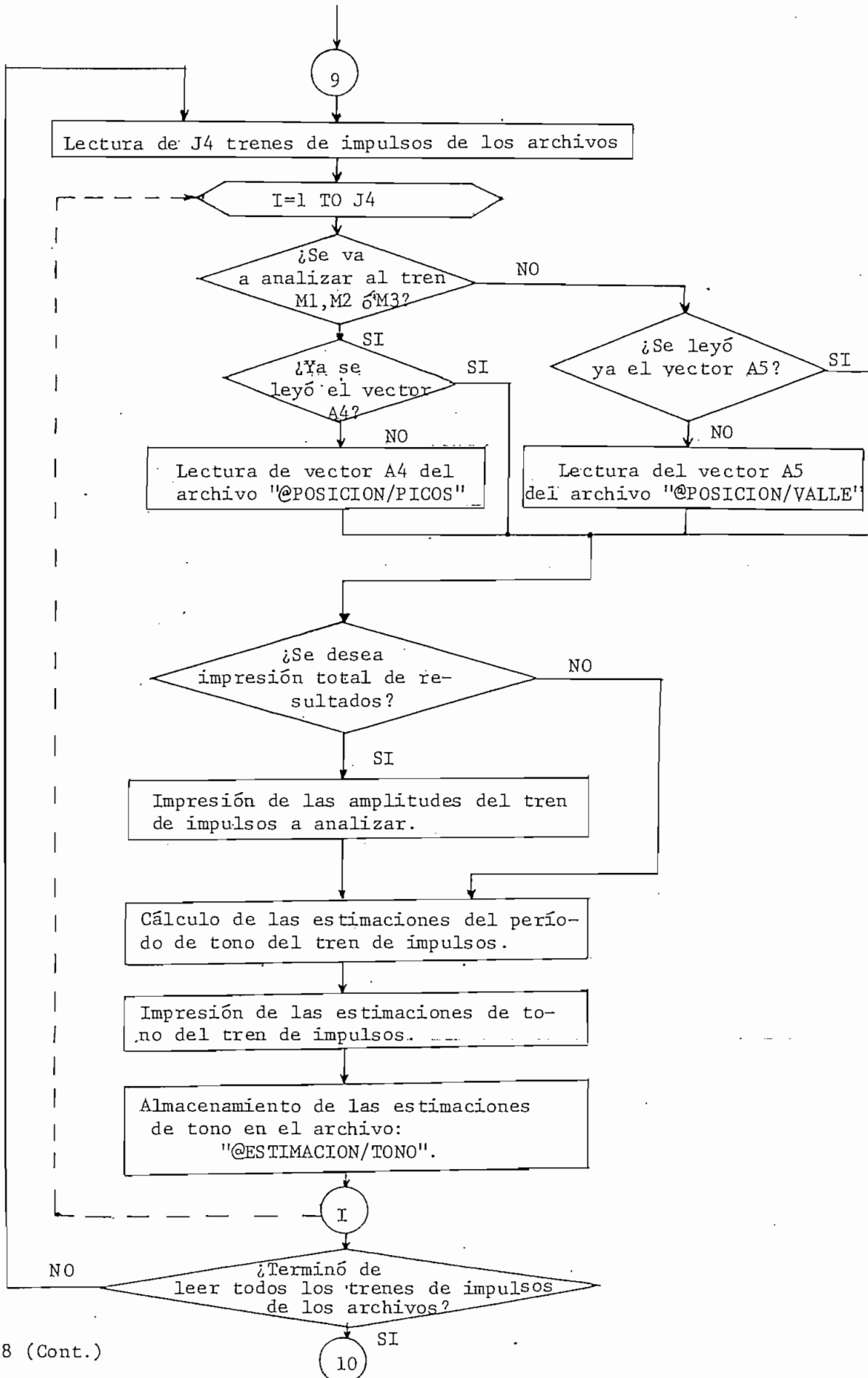


Fig. 3.8 (Cont.)

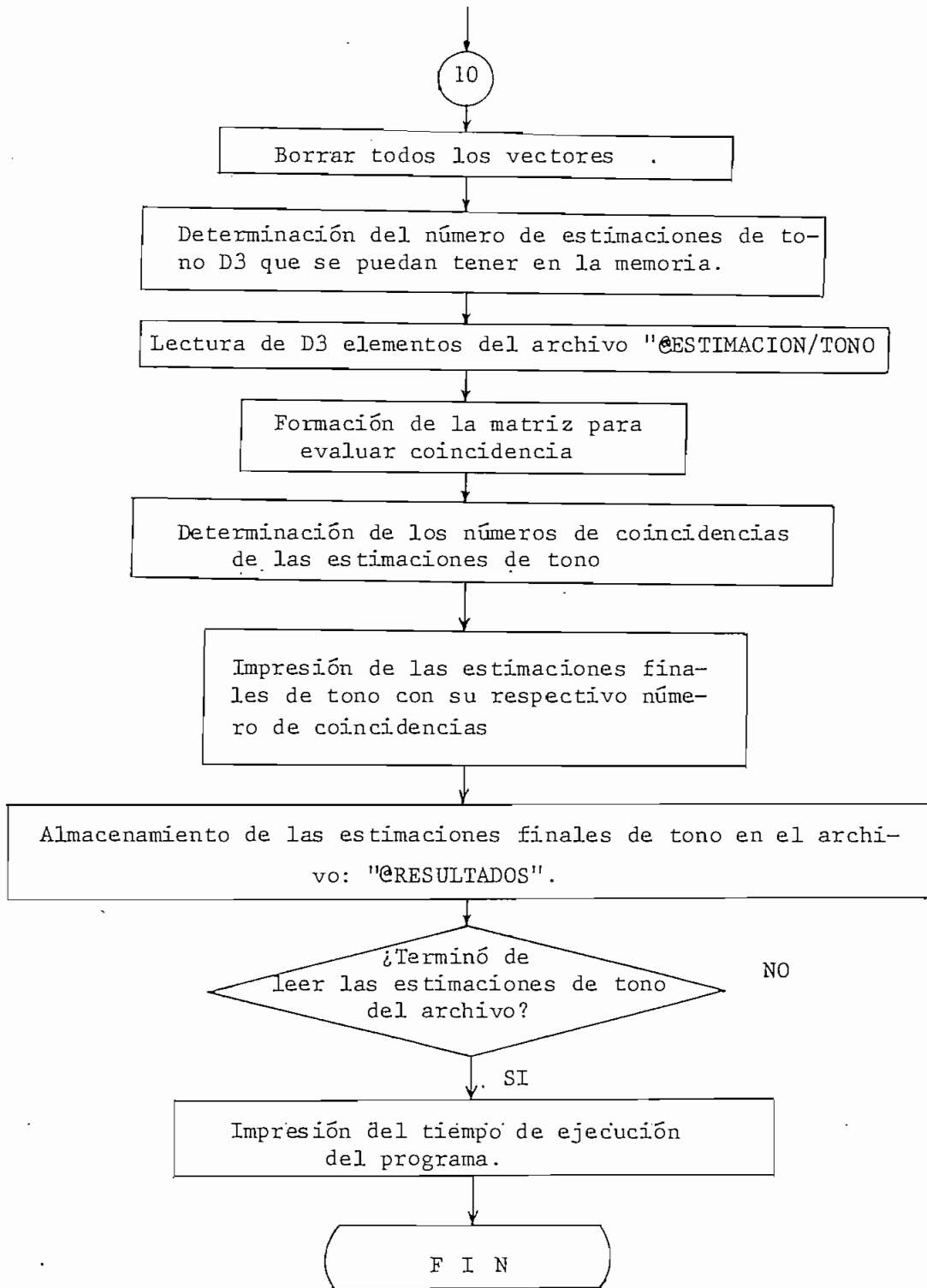


Fig. 3.8 Diagrama de flujo del programa de detección de tono por procesamiento en paralelo "DETPAR".

nicas de programación utilizadas, se debe tener un buen conocimiento del algoritmo y unos buenos conocimientos de programación en "BASIC". Para este tipo de análisis se dispone del listado completo de este programa entre la página 11 y la página 23 del Apéndice "A".

En la Figura 3.8 se puede apreciar el diagrama de flujo correspondiente al programa "DETPAR" para la detección del período de tono por procesamiento en paralelo.

A continuación se da una lista de las principales variables utilizadas en este programa y sus significados.

| <u>NOMBRE</u> | <u>SIGNIFICADO</u> |
|---------------|---|
| 01 | Número del programa que se desea ejecutar. |
| 02 | Número del programa que está en memoria. |
| Q | Número de impresión deseado. |
| A | Dirección de la unidad de impresión deseada. |
| J1 | Número de máximos o picos del segmento de voz. |
| J2 | Número de mínimos o valles del segmento de voz. |
| P | Porcentaje para considerar una amplitud como "con voz". |
| K1 | Valor mínimo del período de tono de la voz. |
| K2 | Margen de seguridad para el cálculo del máximo número de períodos en un segmento de voz. |
| K3 | Constante de coincidencia. |
| K4 | Capacidad de memoria reservada para operaciones simples. |
| G2 | Número de vectores a considerar para un análisis de la capacidad de memoria. |
| A1 | Vector de las amplitudes de las muestras de voz. |
| A2 | Vector con las amplitudes de los picos de la señal. |

| <u>NOMBRE</u> . . . | <u>SIGNIFICADO</u> |
|---------------------|--|
| A3 | Vector con las amplitudes de los valles de la señal. |
| A4 | Vector con las posiciones de los picos de la señal. |
| A5 | Vector con las posiciones de los valles de la señal. |
| N1(1) | Máximo precedente del tren de picos A2. |
| N1(2) | Mínimo precedente del tren del valle A3. |
| J3 | Número común de picos y valles. |
| M1 | Vector con las amplitudes del tren de impulsos m_1 . |
| M2 | Vector con las amplitudes del tren de impulsos m_2 . |
| M3 | Vector con las amplitudes del tren de impulsos m_3 . |
| M4 | Vector con las amplitudes del tren de impulsos m_4 . |
| M5 | Vector con las amplitudes del tren de impulsos m_5 . |
| M6 | Vector con las amplitudes del tren de impulsos m_6 . |
| C1 | Cantidad de memoria que va a ocupar un determinado vector. |
| J4 | Número de trenes de impulsos que se pueden procesar simultáneamente. |
| E | Vector con las estimaciones de tono para cada tren de impulsos. |
| D | Vector con el número de estimaciones de tono para cada tren de impulsos. |
| M | Vector para operaciones en la subrutina de cálculo. |
| B1 | Valor de umbral de amplitud correspondiente a un segmento "con voz" . |
| N3 | Contador de estimaciones de tono para cada tren de impulsos. |
| C1 | Estimación corriente suavizada del período de tono. |
| C2 | Constante de tiempo de blanqueo. |

| <u>NOMBRE</u> | <u>SIGNIFICADO</u> |
|---------------|---|
| C3 | Constante de amortiguación. |
| C4 | Número de muestreos de blanqueo. |
| C5. | Posición final de una estimación del período de tono. |
| C8 | Amplitud de la exponencial decreciente del estimador del período de tono. |
| A7 | Amplitud del pulso detectado como inicio de un período de tono. |
| N4 | Mínimo número de estimaciones de tono común a los 6 trenes de impulsos. |
| B | Matriz para la determinación del número de coincidencias. |
| E9 | Período de tono final correspondiente al máximo número de coincidencias. |
| C9 | Número máximo de coincidencias de una estimación de tono. |
| D1 | Número de muestras de voz. |
| F | Frecuencia de muestreo en KHz. |
| D3 | Número de elementos que se pueden procesar simultáneamente en la memoria. |
| D4 | Bandera que indica la necesidad o no de realizar transferencias entre la memoria y el disco. |
| D5 | Bandera que indica la finalización o no de las transferencias entre disco y memoria o viceversa. |
| N | Contador del número de transferencias. |

Como se puede apreciar en esta lista y en el listado del programa, existen variables que se han utilizado más de

una vez durante el programa y para propósitos diferentes. Esto se debe a que por las circunstancias de capacidad de memoria que se han presentado, durante el desarrollo del programa para este algoritmo, hubo que ahorrar o disminuir el número de variables a utilizarse.

Este programa permite al usuario cuatro opciones para la impresión de resultados; estas son:

- 1) Impresión parcial en pantalla.
- 2) Impresión total en pantalla.
- 3) Impresión parcial en papel.
- 4) Impresión total en papel.

A continuación se presenta un ejemplo del tipo de impresión total de resultados, para la Prueba 5.1 correspondiente a la detección de tono por procesamiento en paralelo del segmento de voz de la "LETRA/A1". Este tipo de impresión puede ser solicitada por el usuario al oprimir los números 2 o 4 para impresión. Para poder establecer la diferencia con la impresión parcial de resultados, hay que referirse a la Prueba 5.1 que se presenta entre los resultados generales en la página 216 del capítulo 4.

TREN DE IMPULSOS GENERADOS M1

| | | | | | |
|----------|---|-------|----------|---|-------|
| M1(7) | = | 3,350 | M1(21) | = | 0,783 |
| M1(39) | = | 6,568 | M1(50) | = | 2,765 |
| M1(61) | = | 2,027 | M1(76) | = | 2,418 |
| M1(86) | = | 3,447 | M1(99) | = | 0,615 |
| M1(117) | = | 6,664 | M1(142) | = | 2,276 |
| M1(168) | = | 3,112 | M1(182) | = | 0,392 |
| M1(200) | = | 6,526 | M1(226) | = | 2,041 |
| M1(253) | = | 3,278 | M1(268) | = | 0,657 |
| M1(286) | = | 6,600 | M1(310) | = | 2,017 |
| M1(335) | = | 3,290 | M1(350) | = | 0,261 |
| M1(369) | = | 7,319 | M1(390) | = | 2,024 |
| M1(414) | = | 3,441 | M1(428) | = | 0,051 |
| M1(445) | = | 6,784 | M1(464) | = | 1,829 |
| M1(467) | = | 2,222 | | | |

ESTIMACIONES DE TONO DEL TREN DE IMPULSOS M1

| | | | | | |
|--------|---|-------|--------|---|-------|
| E1(1) | = | 7,800 | E1(2) | = | 8,300 |
| E1(3) | = | 8,600 | E1(4) | = | 8,300 |

TREN DE IMPULSOS GENERADOS M2

| | | | | | |
|----------|---|--------|----------|---|-------|
| M2(7) | = | 3,350 | M2(21) | = | 0,649 |
| M2(39) | = | 11,556 | M2(50) | = | 0,595 |
| M2(61) | = | 4,876 | M2(76) | = | 1,411 |
| M2(86) | = | 6,096 | M2(99) | = | 0,899 |
| M2(117) | = | 11,768 | M2(142) | = | 5,603 |
| M2(168) | = | 6,404 | M2(182) | = | 0,891 |
| M2(200) | = | 11,207 | M2(226) | = | 5,370 |
| M2(253) | = | 6,572 | M2(268) | = | 0,577 |
| M2(286) | = | 11,188 | M2(310) | = | 5,363 |
| M2(335) | = | 5,568 | M2(350) | = | 1,082 |
| M2(369) | = | 12,048 | M2(390) | = | 5,343 |
| M2(414) | = | 6,711 | M2(428) | = | 1,346 |
| M2(445) | = | 11,293 | M2(464) | = | 5,124 |
| M2(467) | = | 0,536 | | | |

ESTIMACIONES DE TONO DEL TREN DE IMPULSOS M2

| | | | | | |
|--------|---|-------|--------|---|-------|
| E2(1) | = | 7,800 | E2(2) | = | 8,300 |
| E2(3) | = | 8,600 | E2(4) | = | 8,300 |

TREN DE IMPULSOS GENERADOS M3

| | | | | | |
|----------|---|-------|----------|---|-------|
| M3(7) | = | 3,350 | M3(21) | = | 0,000 |
| M3(39) | = | 7,351 | M3(50) | = | 0,000 |
| M3(61) | = | 4,792 | M3(76) | = | 0,000 |
| M3(86) | = | 5,865 | M3(99) | = | 0,000 |
| M3(117) | = | 7,279 | M3(142) | = | 0,000 |
| M3(168) | = | 0,836 | M3(182) | = | 0,000 |
| M3(200) | = | 6,918 | M3(226) | = | 0,000 |
| M3(253) | = | 1,237 | M3(268) | = | 0,000 |
| M3(286) | = | 7,256 | M3(310) | = | 0,000 |
| M3(335) | = | 1,273 | M3(350) | = | 0,000 |
| M3(369) | = | 7,580 | M3(390) | = | 0,000 |
| M3(414) | = | 1,417 | M3(428) | = | 0,000 |
| M3(445) | = | 6,733 | M3(464) | = | 0,000 |

ESTIMACIONES DE TONO DEL TREN DE IMPULSOS M3

E3(1) = 7.800 E3(2) = 8.300
E3(3) = 8.600 E3(4) = 8.300

TREN DE IMPULSOS GENERADOS M4

| | |
|------------------|------------------|
| M4(18) = 1.432 | M4(26) = 4.988 |
| M4(49) = 3.359 | M4(51) = 2.849 |
| M4(71) = 3.829 | M4(78) = 2.649 |
| M4(97) = 1.513 | M4(105) = 5.104 |
| M4(129) = 3.327 | M4(153) = 3.293 |
| M4(180) = 1.283 | M4(187) = 4.682 |
| M4(213) = 3.329 | M4(237) = 3.294 |
| M4(265) = 1.233 | M4(273) = 4.588 |
| M4(297) = 3.346 | M4(323) = 2.278 |
| M4(348) = 1.343 | M4(356) = 4.729 |
| M4(379) = 3.319 | M4(401) = 3.270 |
| M4(426) = 1.295 | M4(434) = 4.509 |
| M4(456) = 3.295 | M4(465) = 1.685 |
| M4(478) = 3.617 | |

ESTIMACIONES DE TONO DEL TREN DE IMPULSOS M4

E4(1) = 7.900 E4(2) = 16.800
E4(3) = 8.300 E4(4) = 7.800

TREN DE IMPULSOS GENERADOS M5

| | |
|-------------------|------------------|
| M5(18) = 4.782 | M5(26) = 4.205 |
| M5(49) = 9.927 | M5(51) = 0.084 |
| M5(71) = 5.856 | M5(78) = 0.231 |
| M5(97) = 4.960 | M5(105) = 4.490 |
| M5(129) = 9.991 | M5(153) = 5.569 |
| M5(180) = 4.395 | M5(187) = 4.290 |
| M5(213) = 9.854 | M5(237) = 5.335 |
| M5(265) = 4.511 | M5(273) = 3.932 |
| M5(297) = 9.945 | M5(323) = 4.295 |
| M5(348) = 4.632 | M5(356) = 4.468 |
| M5(379) = 10.638 | M5(401) = 5.294 |
| M5(426) = 4.735 | M5(434) = 4.560 |
| M5(456) = 10.079 | M5(465) = 0.143 |
| M5(478) = 5.839 | |

ESTIMACIONES DE TONO DEL TREN DE IMPULSOS M5

E5(1) = 8.000 E5(2) = 8.400
E5(3) = 8.400 E5(4) = 8.200

TREN DE IMPULSOS GENERADOS M6

| | | | |
|------------|-------|------------|-------|
| M6(18) = | 1.432 | M6(26) = | 3.556 |
| M6(49) = | 0.000 | M6(51) = | 0.000 |
| M6(71) = | 0.980 | M6(78) = | 0.000 |
| M6(97) = | 0.000 | M6(105) = | 3.591 |
| M6(129) = | 0.000 | M6(153) = | 0.000 |
| M6(180) = | 0.000 | M6(187) = | 3.398 |
| M6(213) = | 0.000 | M6(237) = | 0.000 |
| M6(265) = | 0.000 | M6(273) = | 3.355 |
| M6(297) = | 0.000 | M6(323) = | 0.000 |
| M6(348) = | 0.000 | M6(356) = | 3.387 |
| M6(379) = | 0.000 | M6(401) = | 0.000 |
| M6(426) = | 0.000 | M6(434) = | 3.215 |
| M6(456) = | 0.000 | M6(465) = | 0.000 |
| M6(478) = | 5.302 | | |

ESTIMACIONES DE TONO DEL TREN DE IMPULSOS M6

E6(1) = 0.000

ESTIMACIONES FINALES DEL PERIODO DE TONO

| | | | |
|---------|------------|-----|---|
| E(1) = | 0.000 MSEG | N = | 0 |
| E(2) = | 8.300 MSEG | N = | 1 |
| E(3) = | 0.000 MSEG | N = | 0 |
| E(4) = | 8.300 MSEG | N = | 1 |

N - NUMERO DE COINCIDENCIAS DEL PERIODO DE TONO

TIEMPO DE EJECUCION = 251 SEGUNDOS

3.5. Programa para la Detección del Tono por Reducción de Información - (DETRED)

El algoritmo que se ha desarrollado por medio de este programa es el enunciado por Miller, para la detección del período de tono por reducción de información. Aunque este algoritmo se basa en una reducción de la estructura informativa, es decir en una reducción de los datos a procesarse, se ha hecho necesario considerar el mismo tipo de análisis de memoria en la elaboración de este programa, debido a la pequeña capacidad de memoria disponible.

Debo aclarar que para este algoritmo se han desarrollado estas técnicas de transferencia de información entre el disco y la memoria, para todas aquellas situaciones propicias. En algunas partes de este algoritmo no se hizo posible este tipo de operación, ya sea por lo impredecible de la cantidad de datos que se necesitaban para un determinado cálculo o porque sencillamente se necesitaba de toda la información en forma simultánea en la memoria real y no de una parte.

Sin embargo, existe la posibilidad de realizar un programa para el cual se realice la lectura de cualquier dato necesario desde un archivo en disco. El problema fundamental que presenta este tipo de implementación, es que el tiempo de ejecución del programa sería tan grande que el programa no sería de utilidad para ningún tipo de aplicación.

A continuación se da una lista de los pasos que realiza este programa para la detección del período de tono de segmentos de voz.

1. Inicialización de las diferentes constantes y banderas necesarias durante el desarrollo del programa.
2. Análisis de la capacidad de memoria real disponible, para determinar si se pueden tener todos los elementos del archivo de trabajo que contiene los datos (@MUESTREO/A1) en la memoria, dejando memoria para las matrices A2, A3, A4 y S en las que se va a guardar la información de los ciclos de excursión.
3. Lectura de los datos de voz que se puedan tener simultáneamente en la memoria, para la determinación de los ciclos de excursión.
4. Determinación del primer muestreo diferente de cero "A2", de la amplitud máxima "A3" y de su número de muestreo "A4" para cada ciclo de excursión del segmento de voz.
5. Determinación de la polaridad "N1" de la onda de voz, es decir de la máxima amplitud en valor absoluto. (signo)
6. Cálculo de las sumas de las amplitudes de los muestreos "S" para cada ciclo de excursión.
7. Impresión de las amplitudes máximas de los ciclos de excursión en pantalla o papel, si ha sido solicitada por el usuario.
8. Almacenamiento de los vectores A2, A3, A4 y S correspondientes a los ciclos de excursión. Estos vectores se almacenan en los archivos "@CICLOS/A2", "@CICLOS/A3", "@CICLOS/A4" y "@CICLOS/SUM" respectivamente.
9. Se chequea si ya se han leído todos los datos de voz del archivo de trabajo. Si no se han acabado de leer, se repite el

proceso desde el numeral 3. Cuando se haya acabado de leer todos los datos se continúa con el proceso.

10. Análisis de la capacidad de memoria disponible para determinar el número de elementos de los archivos de A2, A3, A4 y S que se puedan tener simultáneamente en la memoria, para el cálculo de los candidatos a ciclos principales.
11. Lectura de una igual cantidad de elementos de los archivos "@CICLOS/A2", "@CICLOS/A3", "@CICLOS/A4" y "@CICLOS/S".
12. Se eliminan todos los ciclos de excursión negativos, después de haber realizado el correspondiente análisis de la polaridad de la onda de voz.
13. Se eliminan los ciclos de excursión correspondientes a segmentos "sin voz".
14. Se realiza un chequeo de la separación entre los candidatos a ciclos principales.
15. Impresión de las amplitudes máximas de los candidatos a ciclos principales en pantalla o papel, si ha sido solicitada por el usuario.
16. Almacenamiento de los candidatos a ciclos principales en los archivos de datos "@CICLOS/PRIN/A2", "@CICLOS/PRIN/A3" y "@CICLOS/PRIN/A4".
17. Se chequea si ya se han leído todos los elementos de los archivos "@CICLOS/", Si no se han acabado de leer, se repite el proceso desde el numeral 11. Si en cambio, ya se ha terminado la lectura de los elementos de estos archivos, se continúa con el proceso.
18. Como para los futuros cálculos del algoritmo es impredecible el número de elementos de los vectores A2, A3 y A4 que se

requerirán simultáneamente en la memoria, hay que leer completamente estos archivos de datos "@CICLOS/PRIN/A2", "@CICLOS/PRIN/A3" y "@CICLOS/PRIN/A4".

19. Determinación del número de regiones continuas del segmento de voz.
20. Ubicación de los núcleos silábicos primarios de cada región continua de voz.
21. Ubicación de los núcleos silábicos secundarios de cada región continua de voz.
22. Impresión de los núcleos silábicos de la región continua que se está analizando. Esta impresión puede ser en pantalla o papel.
23. Ubicación de los límites silábicos de la región continua de voz.
24. Impresión de los límites silábicos de la región continua en pantalla o papel.
25. Segmentación de cada sílaba para la formación de la envolvente lineal necesaria para la selección de los ciclos principales de cada sílaba de la región continua de voz.
26. Impresión de los ciclos principales de la región continua de voz. Esta impresión va a ser en pantalla o papel de acuerdo a las necesidades del usuario.
27. Almacenamiento de los vectores C2, C3 y C4, correspondientes a la información de los ciclos principales de la región continua, en los archivos de datos "@CICLOPRIN/C2", "@CICLOPRIN/C3" y "@CICLOPRIN/C4".
28. Lectura de estos archivos de datos para el cálculo de las

estimaciones de tono de la región continua de voz.

29. Cálculo de la frecuencia promedio de tono de cada sílaba de la región continua de voz.
30. Impresión de las estimaciones del período de tono de la región continua de voz.
31. Corrección de errores de duplicación y de división para dos de la frecuencia.
32. Cálculo de la frecuencia media de tono de la región continua para la eliminación de los ciclos principales mayores a un determinado ciclo principal en un intervalo alrededor de éste.
33. Cálculo de los períodos de tono del nuevo conjunto de ciclos principales de la región continua de voz.
34. Impresión de las estimaciones finales del período de tono de la región continua de voz.
35. Si ya se han analizado todas las regiones continuas de voz finaliza el programa. De lo contrario, se repite el proceso desde el numeral 20.

Durante el desarrollo de este programa se ha tratado de realizar un programa con la menor cantidad de limitaciones posibles. Sin embargo, debido a la limitada capacidad de este sistema de computación, debo anotar que las limitaciones de memoria que puedan asomar durante la ejecución del programa van a depender únicamente de la capacidad de memoria real que tenga el sistema que se utilice.

En el APENDICE "A" entre la página 24 y la página 38 se presenta un listado completo de este programa, para un análisis más detallado de las técnicas de programación utiliza-

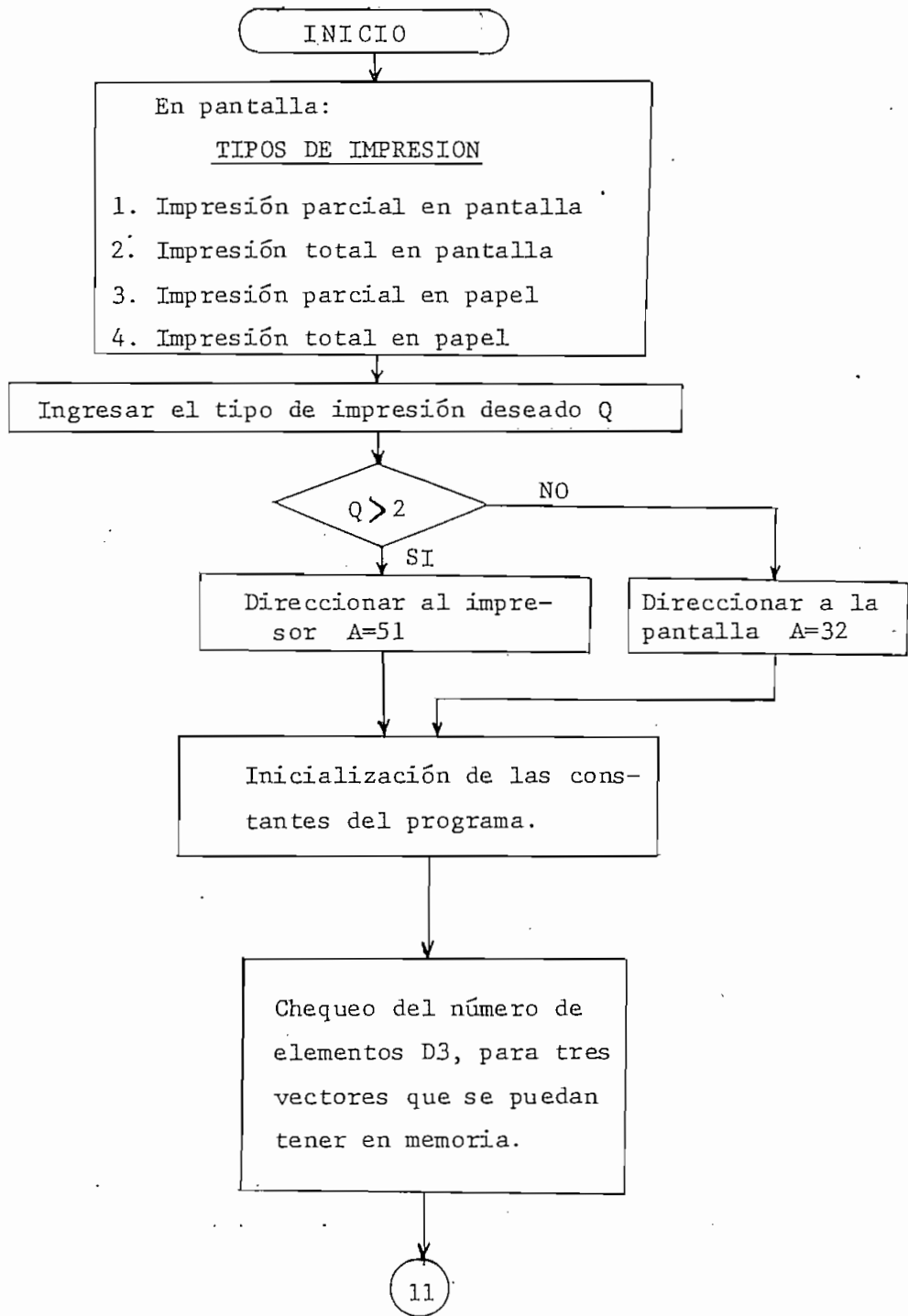


Fig. 3.9 Diagrama de flujo del programa de detección por reducción de información "DETRED"

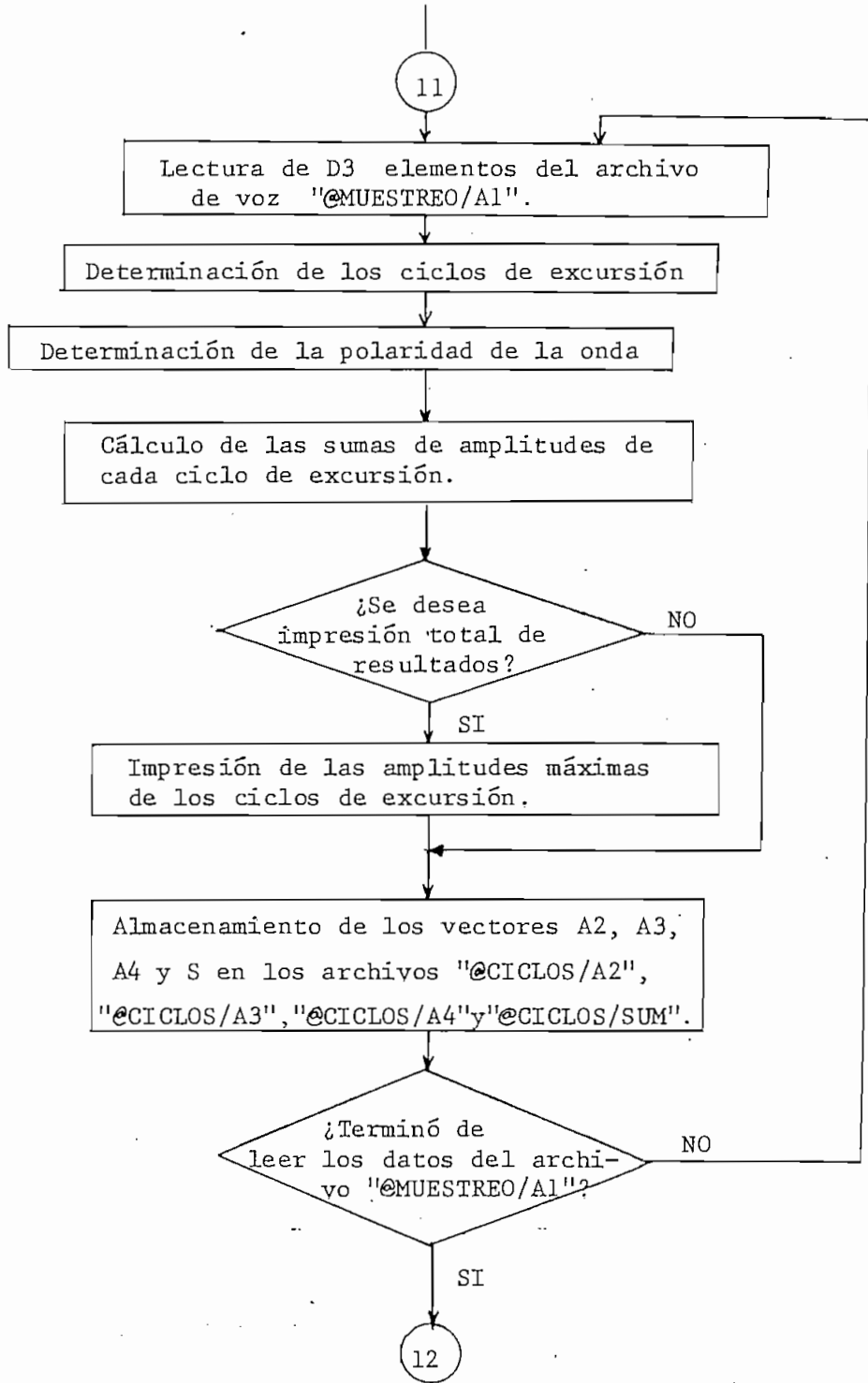


Fig. 3.9 Diagrama de flujo del programa de detección por reducción de información .DETRED".

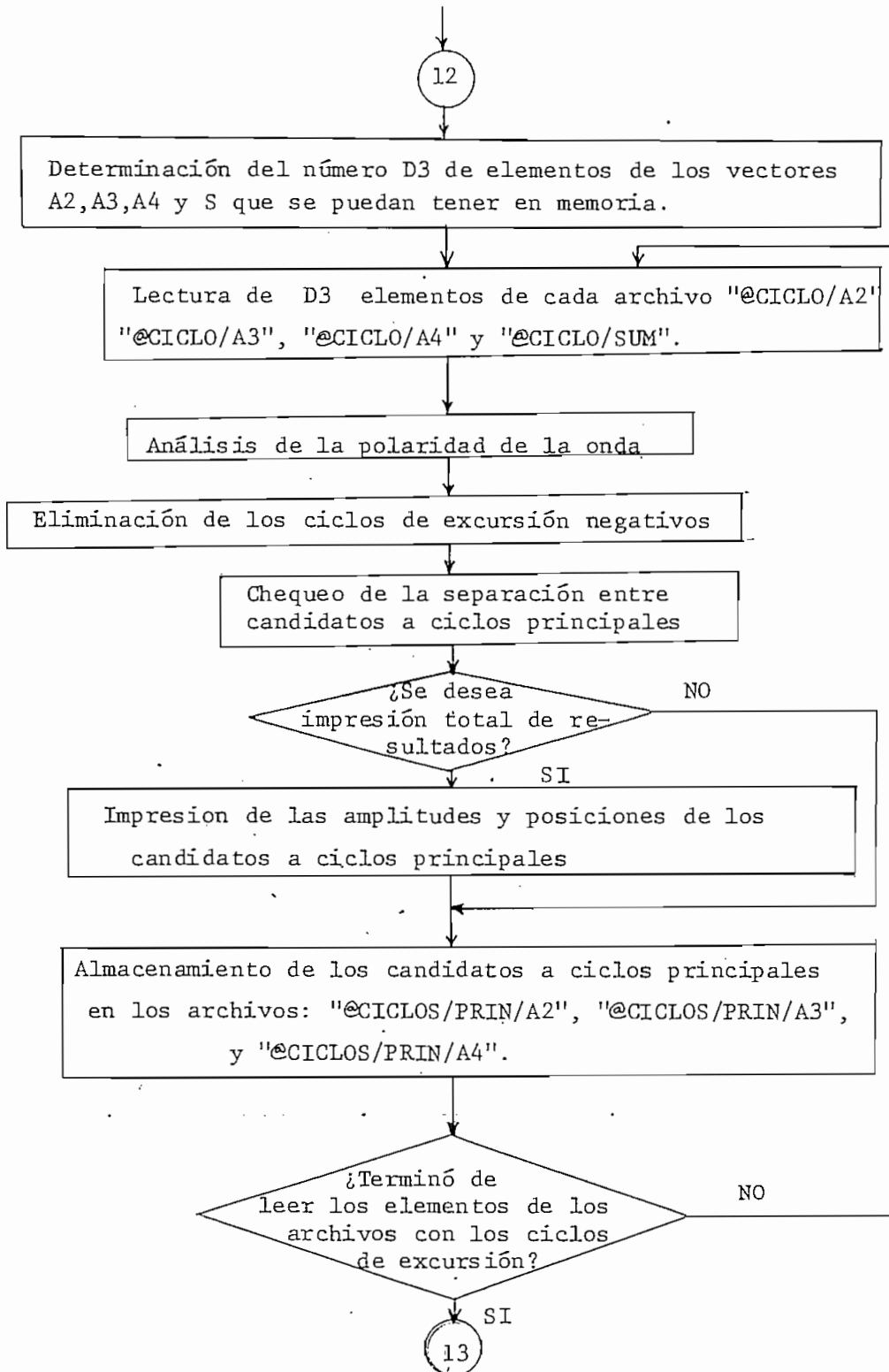


Fig. 3.9 Diagrama de flujo del programa de detección por reducción de información "DETRED".

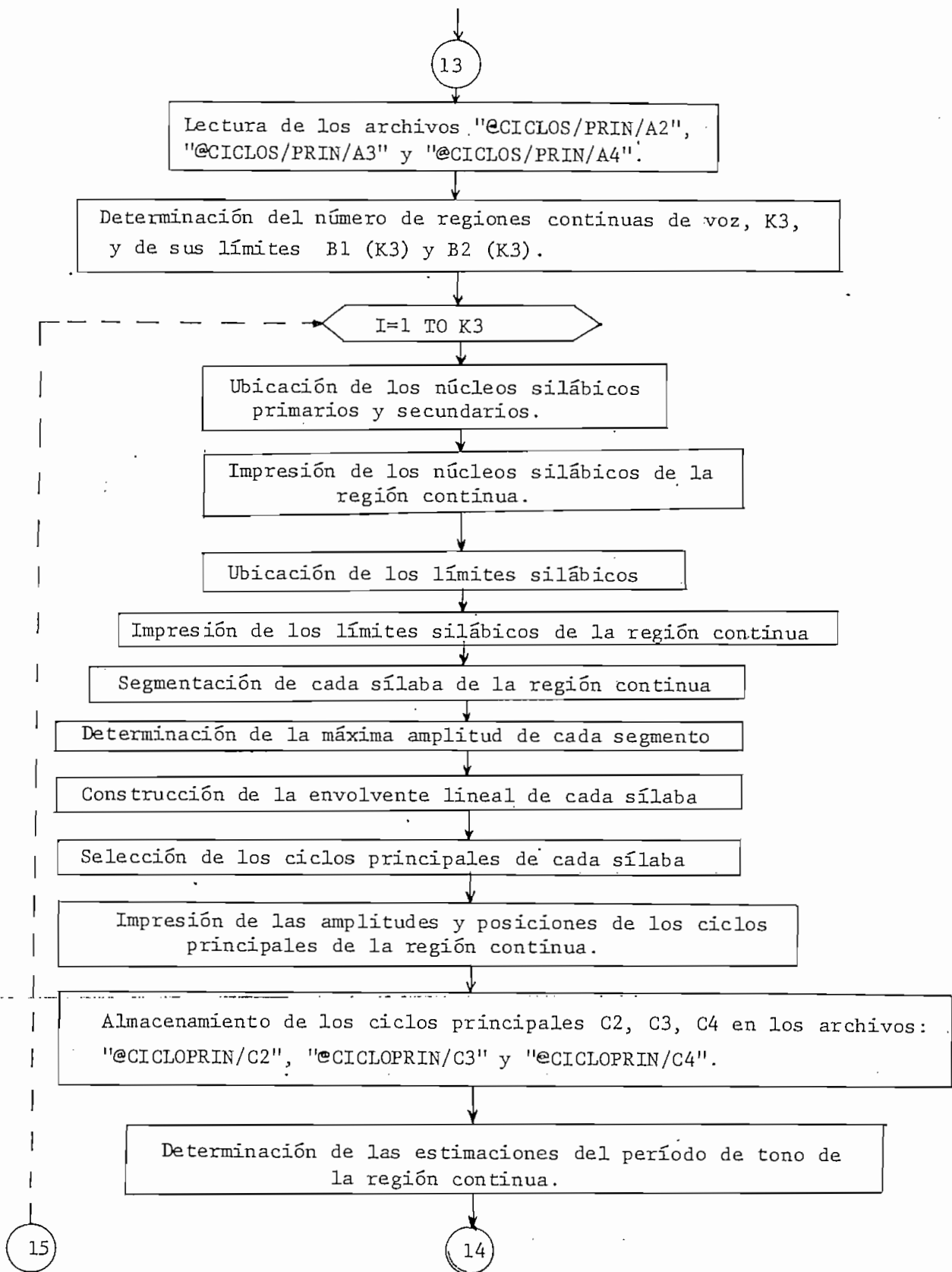


Fig. 3.9 (Cont.)

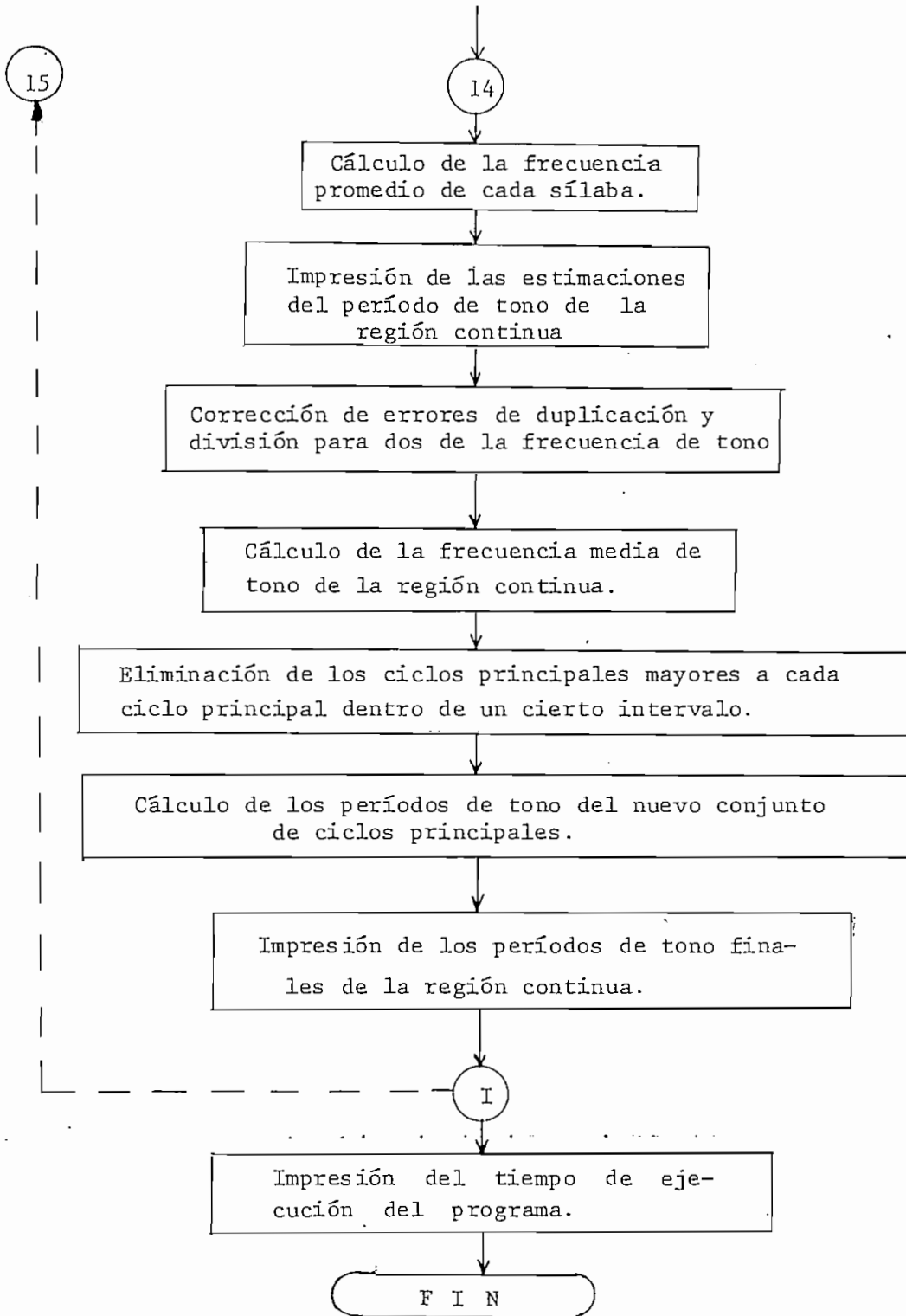


Fig. 3.9 Diagrama de flujo del programa de detección por reducción de información "DETRED".

das. Además en la Fig. 3.9 se puede observar el diagrama de flujo general de este programa.

A continuación se da una lista de las principales variables utilizadas en este programa y sus significados.

| <u>NOMBRE</u> | <u>SIGNIFICADO</u> |
|---------------|--|
| O1 | Número del programa que se desea ejecutar. |
| O2 | Número del programa que está en la memoria. |
| Q | Número del tipo de impresión deseado. |
| A | Número de la dirección del periférico para impresión. |
| L1 | Longitud del segmento de voz.. |
| L2 | Valor de umbral de la suma de un segmento "sin voz". |
| L3 | Valor de umbral de separación entre dos posibles ciclos principales. |
| L4 | Valor límite de separación entre regiones continuas. |
| L5 | Valor de umbral de separación entre el núcleo primario y el núcleo secundario de una región continua. |
| L6 | Constante de distancia alrededor de un núcleo silábico. |
| L7 | Valor de segmentación entre núcleos silábicos. |
| L8 | Intervalo de comparación alrededor de un ciclo principal. |
| K4 | Capacidad de memoria reservada para operaciones de ejecución del programa. |
| C2 | Número de vectores a considerarse en el análisis de la memoria. |
| A1 | Vector con las amplitudes de las muestras de voz. |

NOMBRE

SIGNIFICADO

| | | |
|-----|------|---|
| A2 | | Vector con los números de muestras de las primeras amplitudes diferentes a cero de cada ciclo de excursión. |
| A3 | | Vector con las amplitudes máximas de cada ciclo de excursión. |
| A4 | | Vector con los números de muestras o posiciones de las amplitudes máximas de cada ciclo de excursión. |
| S | | Vector con las sumas de las amplitudes de cada ciclo de excursión. |
| N1 | | Amplitud máxima del segmento de voz. |
| D6 | | Cantidad de ciclos de excursión en un segmento de voz. |
| D7 | | Número de muestras real de una amplitud del segmento de voz. |
| T | | Vector para operaciones de transferencia entre la memoria y el disco. |
| TØ | | Bandera que indica la creación o no de un archivo. |
| T1 | | Bandera que indica lectura o escritura en archivo. |
| T2 | | Número del archivo con que se desea trabajar. |
| T3 | | Bandera para ordenar la apertura o no de un archivo. |
| T4 | | Bandera para ordenar la rebobinación o no de un archivo. |
| T\$ | | Nombre del archivo a operar. |
| K3 | | Número de regiones continuas del segmento de voz. |
| D7 | | Máximo número de regiones continuas que puedan existir. |
| B1 | | Muestreo inicial de una región continua. |
| B2 | | Muestreo final de una región continua. |

NOMBRE

SIGNIFICADO

| | | |
|----|------|--|
| D2 | | Número de candidatos a ciclos principales. |
| MØ | | Amplitud del núcleo silábico primario. |
| NØ | | Número de muestra del núcleo silábico primario. |
| M | | Vector con las amplitudes de los núcleos silábicos de la región continua. |
| N | | Vector con los números de muestras o posiciones de los núcleos silábicos de la región continua. |
| K5 | | Número de núcleos silábicos de una región continua. |
| B3 | | Número total de núcleos primarios y secundarios de una región continua. |
| M1 | | Vector con las amplitudes de los límites silábicos de la región continua. |
| N1 | | Vector con los números de muestras o posiciones de los límites silábicos de la región continua. |
| M2 | | Vector con las amplitudes máximas de cada segmento de la envolvente lineal. |
| N2 | | Vector con los números de las muestras o posiciones de las amplitudes máximas de cada segmento de la envolvente. |
| I2 | | Número de segmentos entre dos núcleos silábicos. |
| I4 | | Número de segmentos dentro de una sílaba, es decir entre límites silábicos. |
| K5 | | Número de ciclos principales dentro de una sílaba. |
| B4 | | Vector con la <u>cantidad de ciclos principales</u> en cada sílaba de una región continua. |

| <u>NOMBRE</u> | <u>SIGNIFICADO</u> |
|---------------|--|
| C2 | Vector con las posiciones de las primeras amplitudes diferentes a cero de cada ciclo principal. |
| C3 | Vector con las amplitudes máximas de cada ciclo principal. |
| C4 | Vector con las posiciones de las amplitudes máximas de cada ciclo principal. |
| E | Matriz con las estimaciones de tono del segmento de voz. |
| E1 | Vector con las estimaciones promedio de frecuencia de cada sílaba. |
| E2 | Frecuencia media de tono de la región continua. |
| K6 | Número de períodos de tono en una sílaba. |
| B5 | Vector con los números de períodos de tono en cada sílaba de una región continua. |
| C1 | Cantidad de memoria que ocupan unos vectores en la memoria. |
| D1 | Número de muestras del segmento de voz. |
| F | Frecuencia de muestreo de las muestras de voz. |
| D3 | Número de elementos de uno o varios vectores que se pueden tener simultáneamente en la memoria. |
| D4 | Bandera que indica si se necesita realizar transferencias de memoria a disco o viceversa. |
| D5 | Bandera que indica la culminación o no de las transferencias entre la memoria y el disco. |
| N | Contador del número de transferencias que se vayan realizando. |

Como se puede apreciar, en este programa también hubo la necesidad de utilizar varias variables en forma repetida para diferentes propósitos a lo largo del programa.

Este programa permite al usuario cuatro posibilidades de impresión de resultados, estas son:

- 1) Impresión parcial en pantalla
- 2) Impresión total en pantalla
- 3) Impresión parcial en papel
- 4) Impresión total en papel

A continuación se presenta un ejemplo del tipo de impresión de resultados en forma total para la Prueba 5.2. correspondiente a la detección de tono por reducción de información del segmento de voz de la "LETRA/A1". Este tipo de impresión de resultados puede ser solicitada por el usuario al oprimir las clases 2 o 4 de impresión. Para que el usuario pueda diferenciar con el tipo de impresión parcial de resultados, debe referirse a la Prueba 5.2. que se presenta entre los resultados generales en la página:217 del capítulo 4.

CICLOS PRINCIPALES DE LA REGION CONTINUA 1

| | | | | | |
|----------|---|-------|----------|---|-------|
| C3(39) | = | 6.568 | C3(117) | = | 6.664 |
| C3(200) | = | 6.526 | C3(286) | = | 6.600 |
| C3(445) | = | 6.784 | | | |

ESTIMACIONES DE TONO DE LA REGION CONTINUA 1

| | | | | | |
|----------|---|-------|----------|---|--------|
| E(1, 1) | = | 8.000 | E(1, 2) | = | 8.300 |
| E(1, 3) | = | 8.600 | E(1, 4) | = | 15.900 |

PERIODOS DE TONO DE LA REGION CONTINUA 1

| | | | |
|----------|---|--------|------|
| E(1, 1) | = | 8.000 | MSEG |
| E(1, 2) | = | 8.300 | MSEG |
| E(1, 3) | = | 8.600 | MSEG |
| E(1, 4) | = | 15.900 | MSEG |

TIEMPO DE EJECUCION = 162 SEGUNDOS

CAPITULO CUARTO

Resultados y Conclusiones

- 4.1 Resultados
- 4.2 Conclusiones
- 4.3 Metodología de uso de los programas.

C A P I T U L O I V

Resultados y Conclusiones

4.1 Resultados

A lo largo de esta tesis se han ido desarrollando los programas necesarios para poder realizar el ingreso de una señal de voz, para poder realizar la detección del período de tono por medio de los dos programas desarrollados para los algoritmos de procesamiento en paralelo y de reducción de información.

Para la utilización de los programas que se han implementado, se hizo necesario disponer de los gráficos de varios segmentos de voz ya sea de corta duración o larga duración. Para esto he utilizado los gráficos de varios de los segmentos de voz de letras o expresiones inglesas correspondientes a las figuras del Capítulo I.

Debo aclarar que no fue posible conseguir gráficos de señales de voz española ni listas de valores de alguna expresión española. Lo ideal habría sido disponer de archivos de datos en disco, circunstancia que aún no es una realidad en nuestra facultad.

Para tener una buena apreciación del funcionamiento de los programas, decidí realizar el ingreso de datos y las pruebas respectivas para las diez vocales de la lengua inglesa y para dos expresiones inglesas de 400 milisegundos

de duración cada una. Cada segmento de voz de una vocal inglesa utilizado tiene una duración de 50 milisegundos. De esta forma se dispone de segmentos de voz de diferentes longitudes.

Primeramente para el ingreso de todas estas señales de voz se utilizó el programa para el ingreso de datos por medio de interpolación, donde se utilizó un ingreso de las esquinas con ayuda del grafizador. Previo al trabajo en el computador fue necesario ampliar los gráficos de los segmentos de voz a ser ingresados para facilitar la ubicación de los puntos de inflexión o quiebre con la ayuda del visor óptico del grafizador.

Para que las amplitudes de los diferentes gráficos quedarán relacionadas entre sí, se tuvo especial cuidado en establecer una misma dimensión física por medio de las teclas de SET. Debo anotar que el ingreso de un segmento de 50 milisegundos por medio de este método requiere por lo menos de unos 15 minutos.

El gráfico de los segmentos de voz de 400 milisegundos de duración, se dividió en cuatro gráficos de 100 milisegundos para tener una mejor apreciación de la onda de voz.

Este tipo de ingreso de datos va a tener errores de observación como son los de paralelaje. Sin embargo, se puede apreciar al comparar los gráficos de las siguientes páginas con las figuras del primer capítulo que los resultados obtenidos por medio de este programa son muy buenos. Además

debe considerarse que los valores de las amplitudes tienen un significado relativo entre si para la detección del tono. Es decir que no importa el valor numérico de las muestras sino su relación de amplitud con respecto a las demás muestras. Sin embargo, al realizar la medición de los diferentes valores numéricos de un gráfico de estos y compararlas con los valores medidos para la figura correspondiente se pudo observar que los errores cometidos eran del orden del 2% en los pocos casos que se presentaron variaciones.

Para el muestreo de todas estas señales de voz ingresadas se utilizó una frecuencia de muestreo de 10 HKz. Todas estas señales de voz ingresadas fueron grafizadas y almacenadas en la biblioteca de archivos de datos "@VOZ" para las posteriores pruebas a realizarse.

Para la realización de las diferentes pruebas de detección de tono se realizó el ingreso de las muestras de cada segmento de voz desde el archivo correspondiente en el disco. Este tipo de ingreso es muy rápido lo que hace más eficiente el funcionamiento de todo el proceso de detección.

Cada prueba se dividió en dos partes, una primera parte en la cual se realizó la detección de tono por medio del programa de procesamiento en paralelo. La segunda parte de cada prueba corresponde a la detección de tono con el programa de reducción de información.

Previo a la realización de todas las pruebas definitivas se llevó a cabo un proceso experimental para esta-

blecer los valores más adecuados de las diferentes constantes utilizadas en los dos programas.

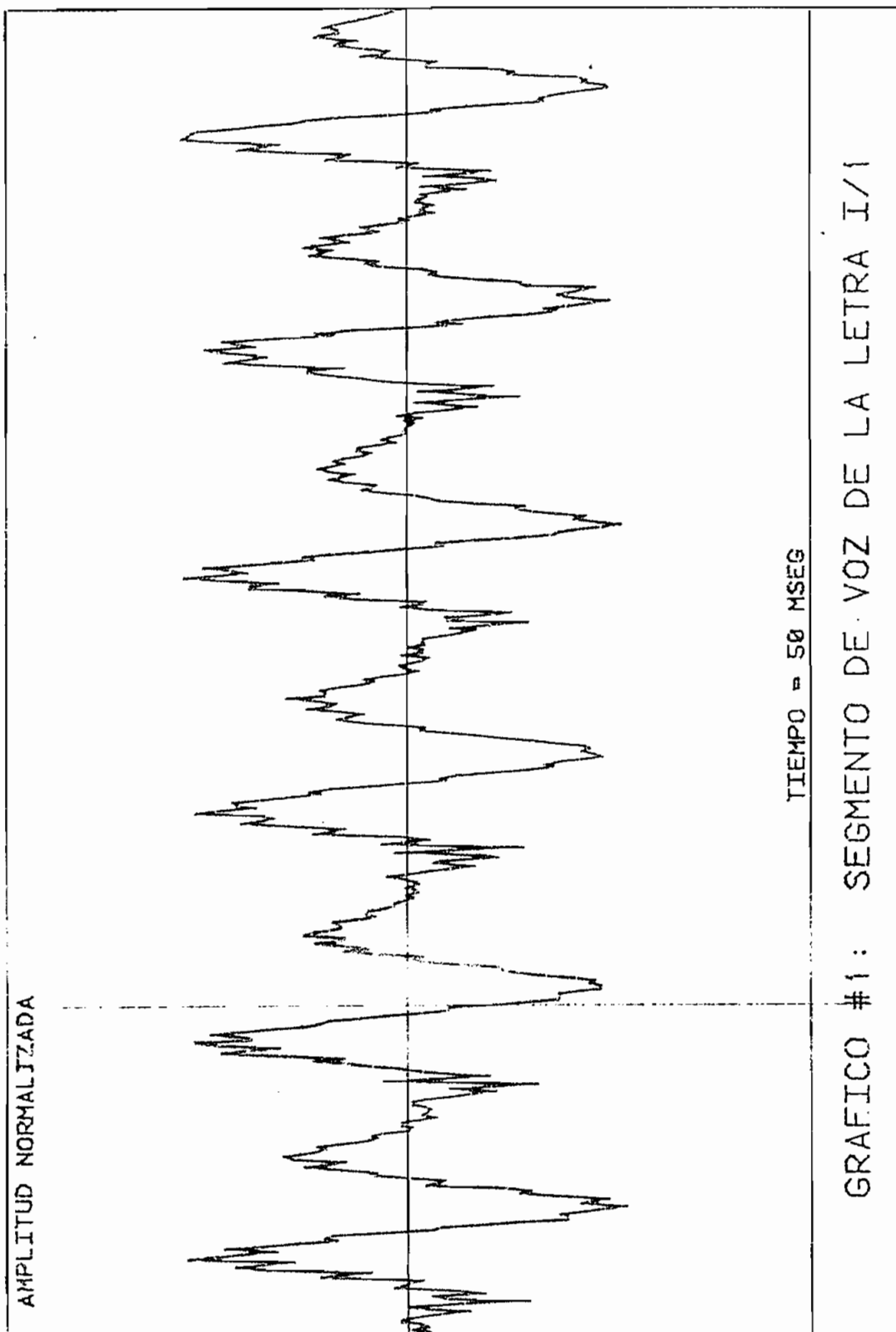
La constante de memoria K4 se varió entre 1KByte y 5KByte, estableciéndose que el mínimo valor que se le podía dar para tener un funcionamiento satisfactorio de los programas era de 2KBytes. Con este valor de memoria se logro la realización de todas las pruebas sin que apareciera ningún problema de saturación de la memoria (Memory Full).

Para el programa de detección por procesamiento en paralelo se variaron las constantes P, K1 y K3 hasta establecer sus valores más propicios. El valor mínimo del período de tono K1 no alteraba los resultados mayormente, debido a que no estaban analizando las voces de diferentes locutores. Por esta razón se ha dejado $K1=4$ milisegundos, es decir igual a su valor típico. La constante de coincidencia K3 se varió entre 0.1 milisegundos y 0.4 milisegundos estableciéndose que a medida que se aumentaba el número de coincidencias de los períodos de tono de las diferentes pruebas. Para realizar una selección relativamente estricta del período de tono se ha fijado una constante de coincidencia $K3=0.2$. El porcentaje para considerar una amplitud como significativa dentro de la detección de tono P es el parámetro que más se ha variado para poder establecer o determinar el valor que brinde la mayor cantidad de eliminaciones de los errores de duplicación y división para dos del período de tono. Después de probar con diferentes valores se estableció que para $P=0.7$ se obtenían los mejores resultados.

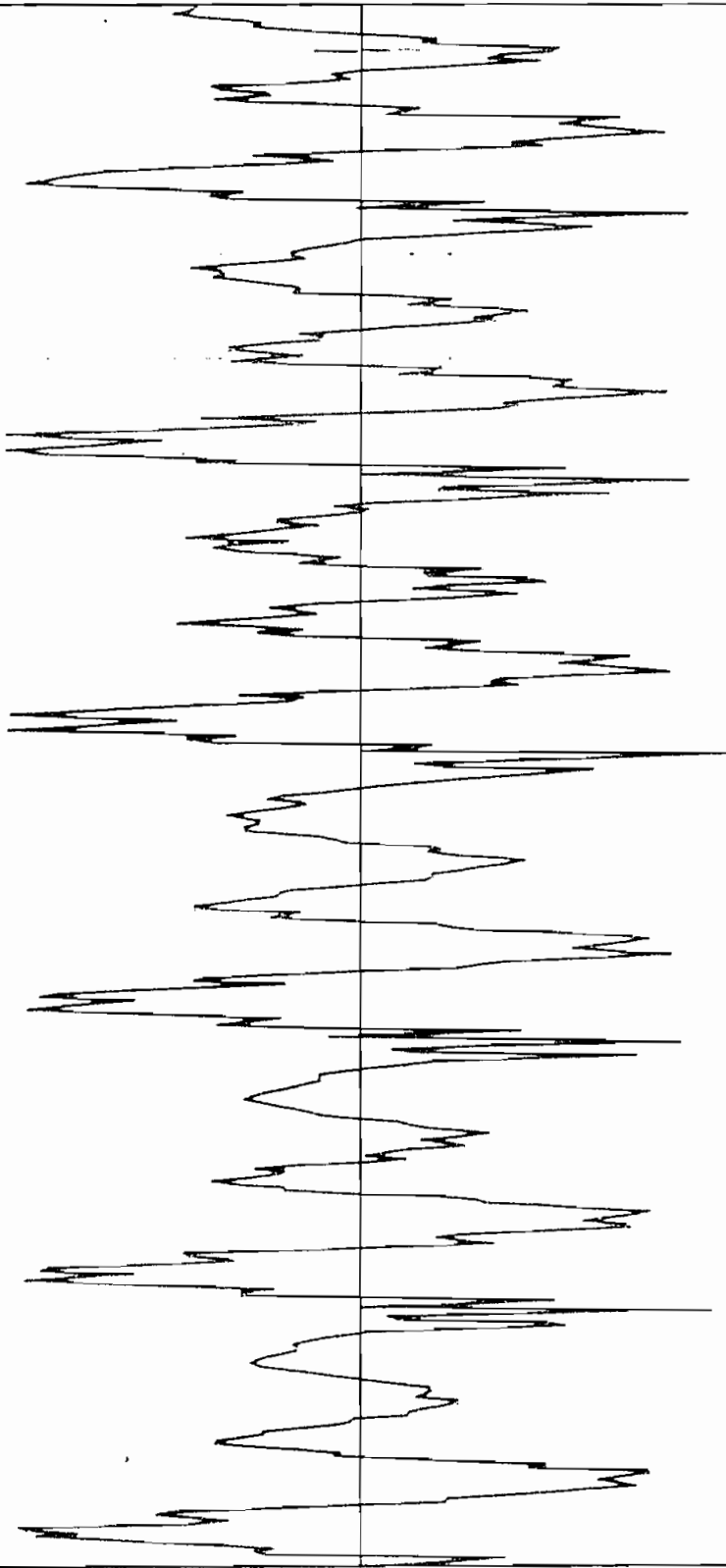
Para el programa de detección por reducción de información sólo hubo que variar el porcentaje del valor de la envolvente para declarar una entrada de información como un ciclo principal. De acuerdo al algoritmo este valor debía ser del 90% , pero al realizar las diferentes pruebas con el programa se pudo observar que este valor era muy alto. Por esta razón fue necesario declarar como ciclos principales a todas aquellas entradas con amplitudes mayores o iguales al 80% del valor de la envolvente lineal evitando que se eliminarán algunos ciclos principales lo que conduciría a errores de duplicación del período de tono de algunas regiones de los segmentos de voz.

Para ambos programas, las estimaciones de tono iguales a cero representan valores de estimaciones de tono "sin voz" o eliminadas por el algoritmo.

En las siguientes páginas se presentan todas las pruebas de voz realizadas con los diferentes valores de las constantes establecidos durante el desarrollo de este trabajo de tesis. En cada caso se presentan primero el gráfico de ~~la señal obtenido por medio del programa de grafización.~~ A continuación se presentan la impresión parcial de resultados de la detección de tono por procesamiento en paralelo (caso 1) y la impresión parcial de los resultados de la detección de tono por reducción de información (caso 2). La numeración de cada prueba esta relacionada con la numeración del gráfico respectivo. En cada caso aparece impreso el tiempo de ejecución del respectivo programa.



AMPLITUD NORMALIZADA



TIEMPO = 50 MSEG

GRAFICO #2: SEGMENTO DE VOZ DE LA LETRA I/2

ESCUELA POLITECNICA NACIONAL
FACULTAD DE INGENIERIA ELECTRICA
TRABAJO DE TESIS - JUNIO 1982
PROGRAMADO POR: CARLOS E. TOBON
DETECCION DEL PERIODO DE TONO DEL SEGMENTO DE VOZ: LETRA I/2
PRUEBA #2.1 FECHA DE REALIZACION: 01-JUL-82
=====

DETECCION DE TONO POR PROCESAMIENTO EN PARALELO

ESTIMACIONES DE TONO DEL TREN DE IMPULSOS M1

E1(1) = 8.300 E1(2) = 8.700
E1(3) = 8.900 E1(4) = 8.900
E1(5) = 8.500

ESTIMACIONES DE TONO DEL TREN DE IMPULSOS M2

E2(1) = 31.800

ESTIMACIONES DE TONO DEL TREN DE IMPULSOS M3

E3(1) = 0.000

ESTIMACIONES DE TONO DEL TREN DE IMPULSOS M4

E4(1) = 7.200 E4(2) = 17.800
E4(3) = 19.800

ESTIMACIONES DE TONO DEL TREN DE IMPULSOS M5

E5(1) = 17.100 E5(2) = 17.500
E5(3) = 8.700

ESTIMACIONES DE TONO DEL TREN DE IMPULSOS M6

E6(1) = 17.100

ESTIMACIONES FINALES DEL PERIODO DE TONO

E(1) = 8.300 MSEG N = 1
E(2) = 8.700 MSEG N = 1
E(3) = 8.900 MSEG N = 1
E(4) = 8.900 MSEG N = 2
E(5) = 0.000 MSEG N = 0

N - NUMERO DE COINCIDENCIAS DEL PERIODO DE TONO

TIEMPO DE EJECUCION = 275 SEGUNDOS

ESCUELA POLITECNICA NACIONAL
FACULTAD DE INGENIERIA ELECTRICA
TRABAJO DE TESIS - JUNIO 1982
PROGRAMADO POR: CARLOS E. TOBON
DETECCION DEL PERIODO DE TONO DEL SEGMENTO DE VOZ: LETRA I/2
PRUEBA #2.2 FECHA DE REALIZACION: 01-JUL-82
=====

DETECCION DE TONO POR REDUCCION DE INFORMACION

NUCLEOS SILABICOS DE LA REGION CONTINUA 1

M (358) = 8.434

LIMITES SILABICOS DE LA REGION CONTINUA 1

M1(4) = 8.133 H1(491) = 4.473

CICLOS PRINCIPALES DE LA REGION CONTINUA 1

C3(13) = 8.133 C3(93) = 7.977
C3(180) = 7.931 C3(269) = 8.380
C3(358) = 8.434 C3(443) = 7.967

ESTIMACIONES DE TONO DE LA REGION CONTINUA 1

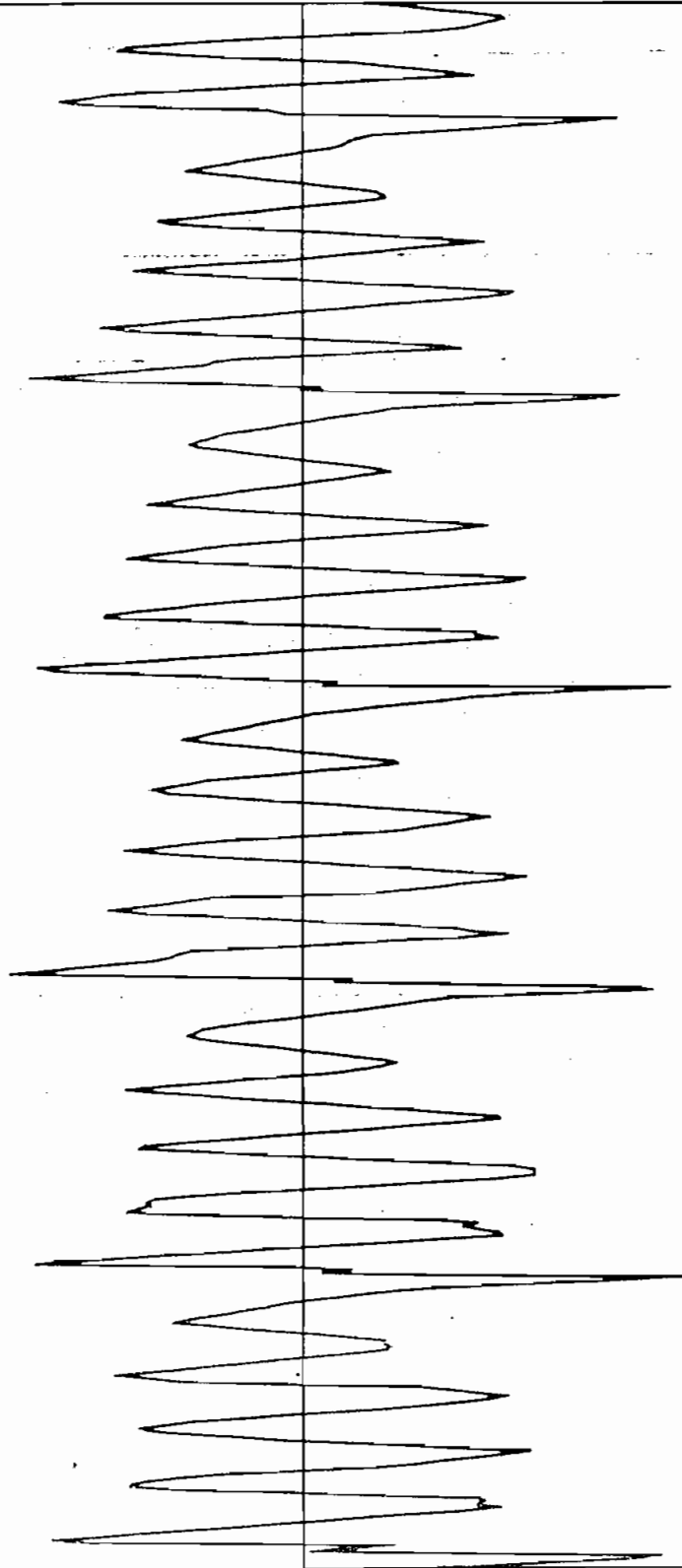
E(1, 1) = 8.400 E(1, 2) = 8.600
E(1, 3) = 9.100 E(1, 4) = 8.900
E(1, 5) = 8.400

PERIODOS DE TONO DE LA REGION CONTINUA 1

E(1, 1) = 8.400 MSEG
E(1, 2) = 8.600 MSEG
E(1, 3) = 9.100 MSEG
E(1, 4) = 8.900 MSEG
E(1, 5) = 8.400 MSEG

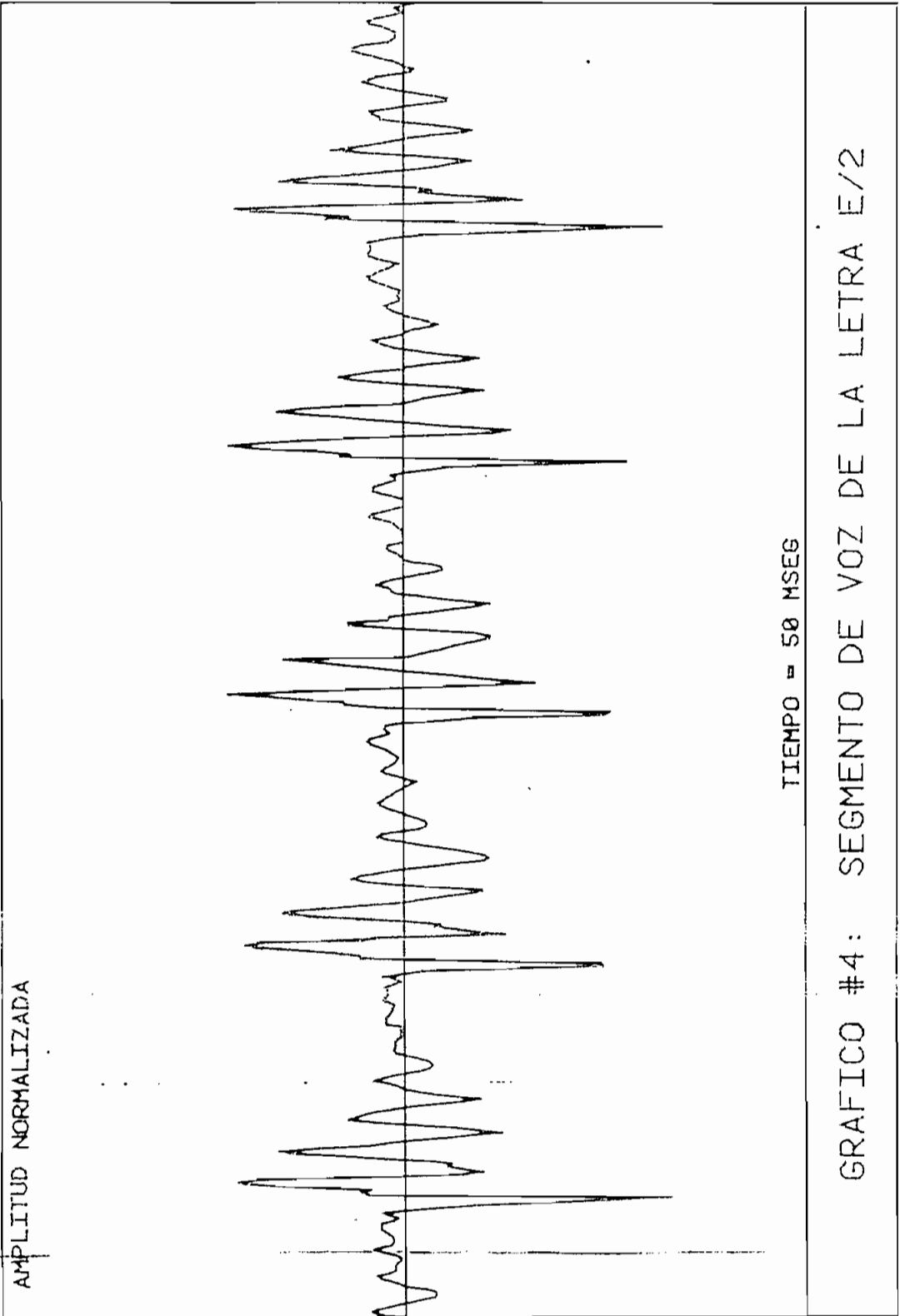
TIEMPO DE EJECUCION = 136 SEGUNDOS

AMPLITUD NORMALIZADA

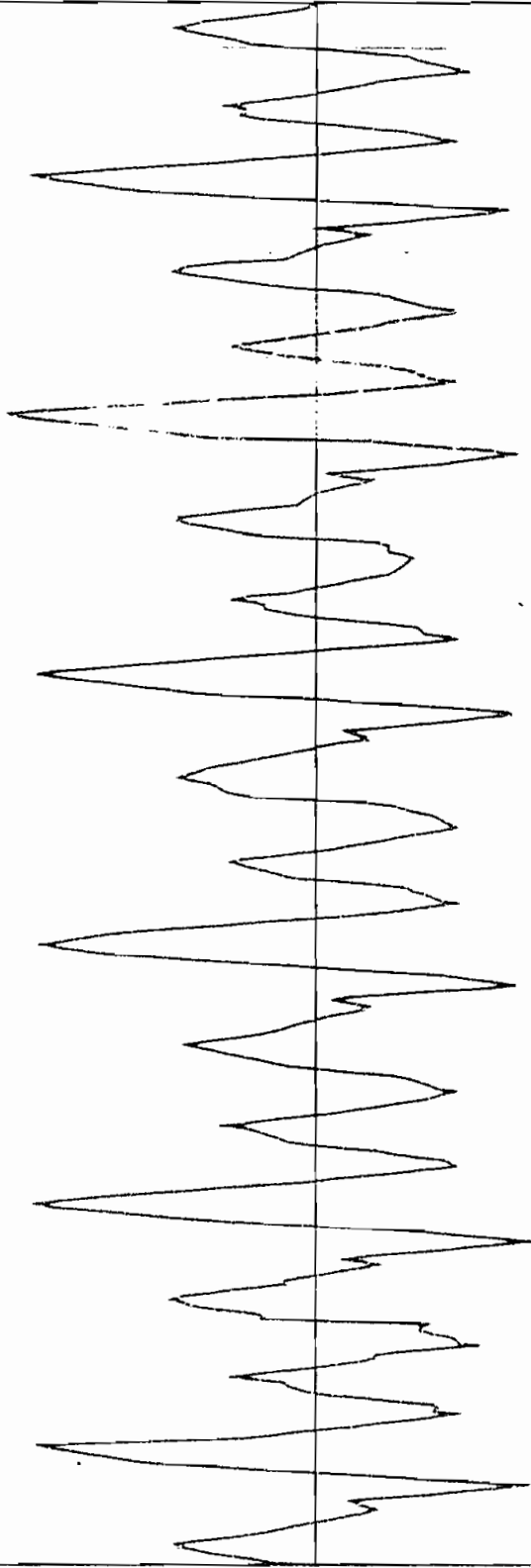


TIEMPO = 50 MSEG

GRAFICO #3: SEGMENTO DE VOZ DE LA LETRA E/1



AMPLITUD NORMALIZADA



TIEMPO = 50 MSEG

GRAFICO #5: SEGMENTO DE VOZ DE LA LETRA A/1

ESCUELA POLITECNICA NACIONAL
FACULTAD DE INGENIERIA ELECTRICA
TRABAJO DE TESIS - JUNIO 1982
PROGRAMADO POR: CARLOS E. TORON
DETECCION DEL PERIODO DE TONO DEL SEGMENTO DE VOZ: LETRA A/1
PRUEBA #5.2 FECHA DE REALIZACION: 01-JUL-82
=====

DETECCION DE TONO POR REDUCCION DE INFORMACION

NUCLEOS SILABICOS DE LA REGION CONTINUA 1

M (369) = 7.319

LIMITES SILABICOS DE LA REGION CONTINUA 1

M1(30) = 6.568 M1(486) = 3.378

CICLOS PRINCIPALES DE LA REGION CONTINUA 1

C3(39) = 6.568 C3(117) = 6.664
C3(200) = 6.526 C3(286) = 6.600
C3(369) = 7.319 C3(445) = 6.784

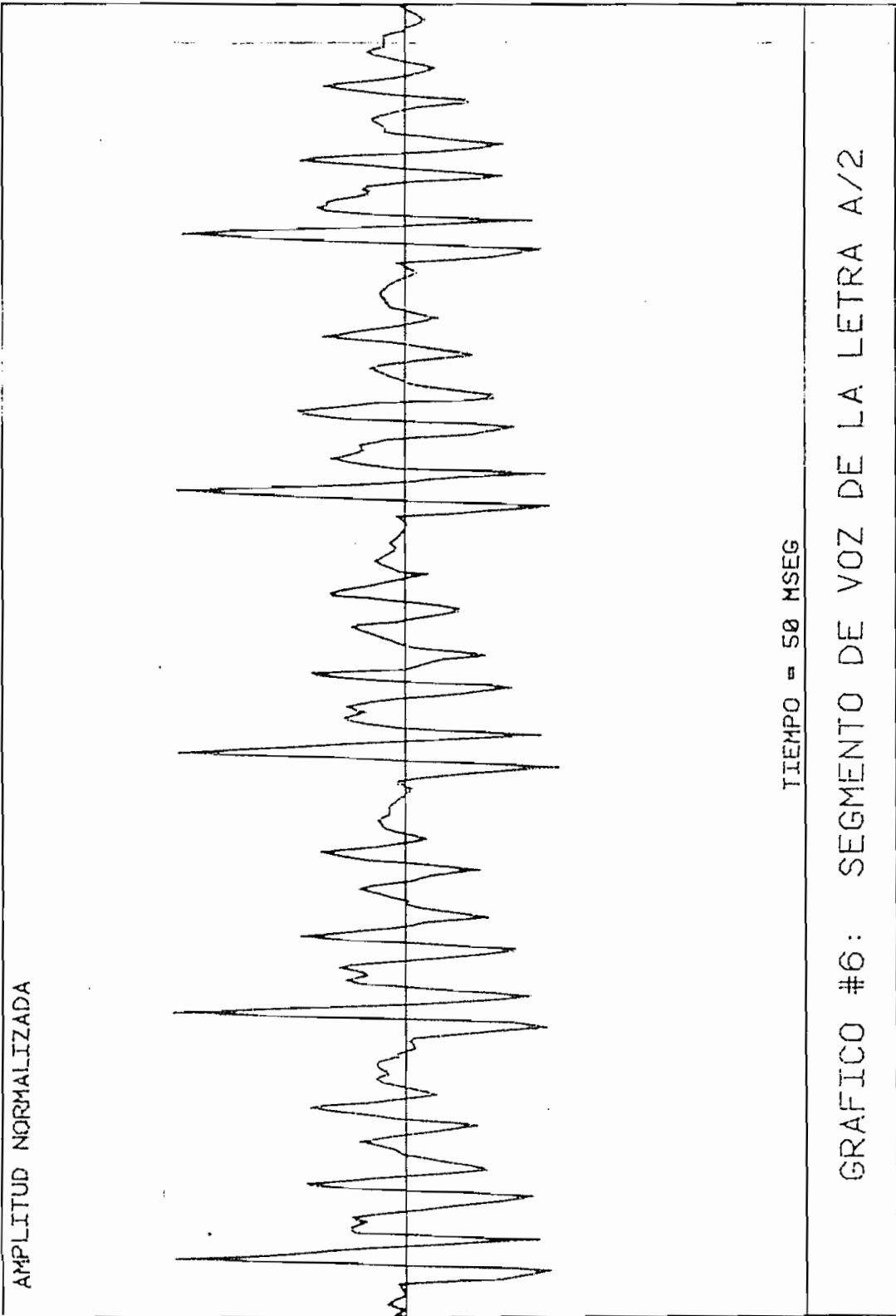
ESTIMACIONES DE TONO DE LA REGION CONTINUA 1

E(1, 1) = 8.000 E(1, 2) = 8.300
E(1, 3) = 8.600 E(1, 4) = 8.200
E(1, 5) = 7.700

PERIODOS DE TONO DE LA REGION CONTINUA 1

E(1, 1) = 8.000 MSEG
E(1, 2) = 8.300 MSEG
E(1, 3) = 8.600 MSEG
E(1, 4) = 8.200 MSEG
E(1, 5) = 7.700 MSEG

TIEMPO DE EJECUCION = 126 SEGUNDOS



ESCUELA POLITECNICA NACIONAL
FACULTAD DE INGENIERIA ELECTRICA
TRABAJO DE TESIS - JUNIO 1982

PROGRAMADO POR: CARLOS E. TOBON

DETECCION DEL PERIODO DE TONO DEL SEGMENTO DE VOZ: LETRA A/2
PRUEBA #6.1

FECHA DE REALIZACION: 01-JUL-82

DETECCION DE TONO POR PROCESAMIENTO EN PARALELO

ESTIMACIONES DE TONO DEL TREN DE IMPULSOS M1

E1(1) = 9.500 E1(2) = 9.900
E1(3) = 9.900 E1(4) = 9.700

ESTIMACIONES DE TONO DEL TREN DE IMPULSOS M2

E2(1) = 9.500 E2(2) = 9.900
E2(3) = 9.900 E2(4) = 9.700

ESTIMACIONES DE TONO DEL TREN DE IMPULSOS M3

E3(1) = 9.500 E3(2) = 9.900
E3(3) = 9.900 E3(4) = 9.700

ESTIMACIONES DE TONO DEL TREN DE IMPULSOS M4

E4(1) = 9.400 E4(2) = 9.900
E4(3) = 11.100 E4(4) = 8.500

ESTIMACIONES DE TONO DEL TREN DE IMPULSOS M5

E5(1) = 9.400 E5(2) = 9.900

ESTIMACIONES DE TONO DEL TREN DE IMPULSOS M6

E6(1) = 2.900 E6(2) = 29.300

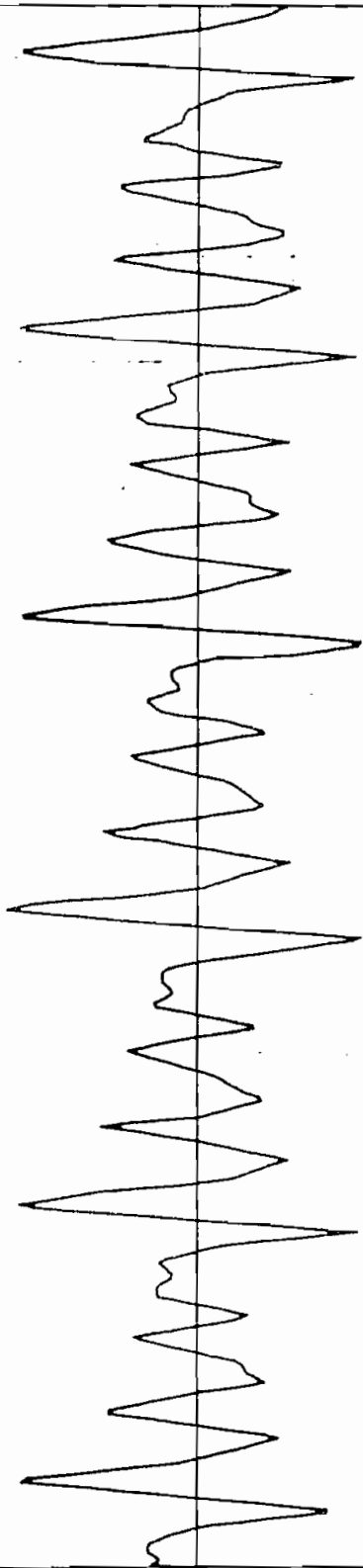
ESTIMACIONES FINALES DEL PERIODO DE TONO

E(1) = 9.500 MSEG N = 1
E(2) = 9.900 MSEG N = 1
E(3) = 9.900 MSEG N = 1
E(4) = 9.700 MSEG N = 2

N - NUMERO DE COINCIDENCIAS DEL PERIODO DE TONO

TIEMPO DE EJECUCION = 184 SEGUNDOS

AMPLITUD NORMALIZADA



TIEMPO = 50 MSEG

GRAFICO #7: SEGMENTO DE VOZ DE LA LETRA O/1

ESCUELA POLITECNICA NACIONAL
FACULTAD DE INGENIERIA ELECTRICA
TRABAJO DE TESIS - JUNIO 1982

PROGRAMADO POR: CARLOS E. TOBON

DETECCION DEL PERIODO DE TONO DEL SEGMENTO DE VOZ: LETRA O/1
PRUEBA #7.1

FECHA DE REALIZACION: 01-JUL-82

DETECCION DE TONO POR PROCESAMIENTO EN PARALELO

ESTIMACIONES DE TONO DEL TREN DE IMPULSOS M1

E1(1) = 9.000 E1(2) = 9.400

E1(3) = 18.500 E1(4) = 8.800

ESTIMACIONES DE TONO DEL TREN DE IMPULSOS M2

E2(1) = 9.000 E2(2) = 9.400

E2(3) = 9.400 E2(4) = 9.100

E2(5) = 8.800

ESTIMACIONES DE TONO DEL TREN DE IMPULSOS M3

E3(1) = 9.000 E3(2) = 9.400

E3(3) = 9.400 E3(4) = 9.100

ESTIMACIONES DE TONO DEL TREN DE IMPULSOS M4

E4(1) = 9.100 E4(2) = 9.400

E4(3) = 9.400 E4(4) = 9.100

E4(5) = 8.900

ESTIMACIONES DE TONO DEL TREN DE IMPULSOS M5

E5(1) = 9.000 E5(2) = 9.500

E5(3) = 9.300 E5(4) = 9.000

ESTIMACIONES DE TONO DEL TREN DE IMPULSOS M6

E6(1) = 9.100 E6(2) = 9.400

E6(3) = 9.400 E6(4) = 9.100

ESTIMACIONES FINALES DEL PERIODO DE TONO

E(1) = 9.000 MSEG N = 5

E(2) = 9.000 MSEG N = 5

E(3) = 18.500 MSEG N = 3

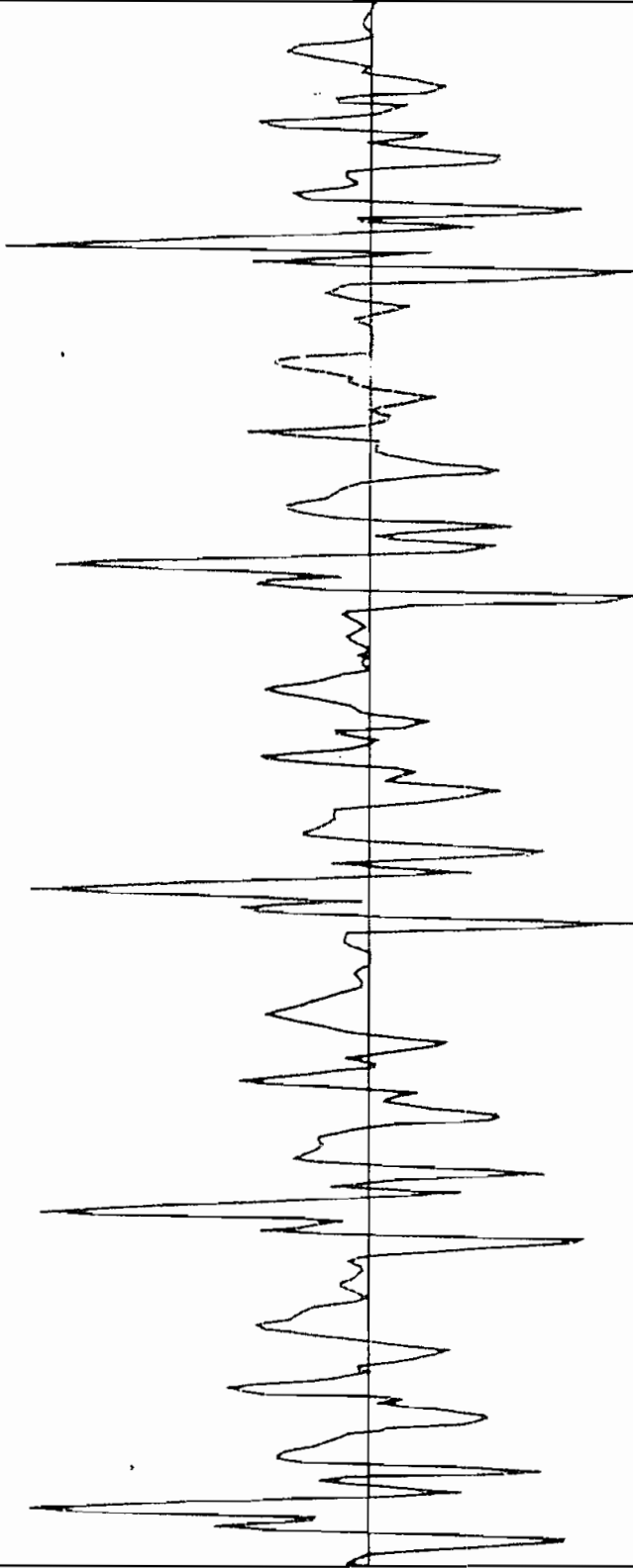
E(4) = 9.400 MSEG N = 2

E(5) = 0.000 MSEG N = 0

N - NUMERO DE COINCIDENCIAS DEL PERIODO DE TONO

TIEMPO DE EJECUCION = 164 SEGUNDOS

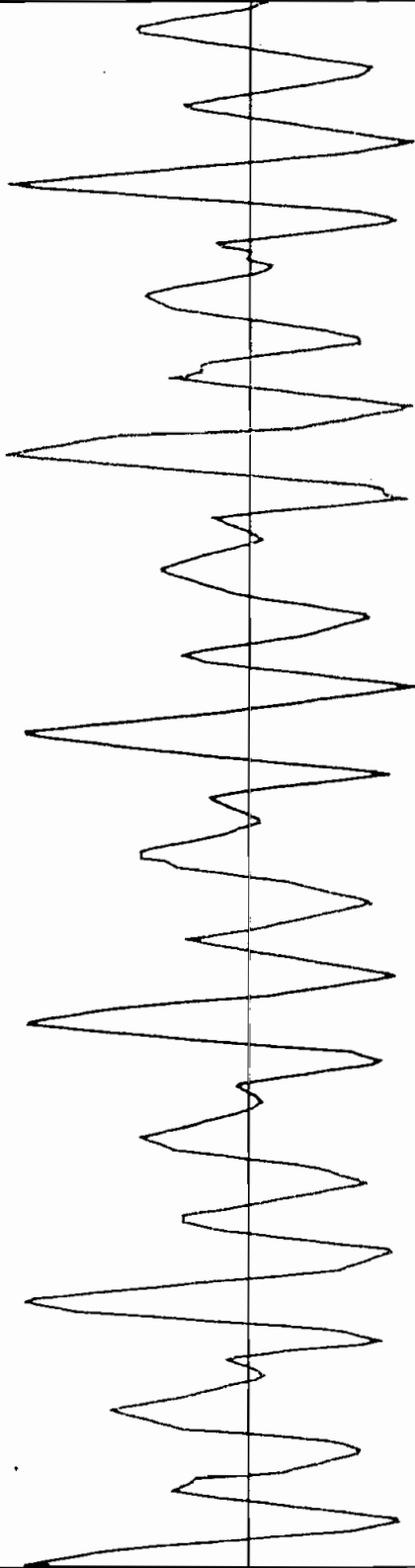
AMPLITUD NORMALIZADA



TIEMPO = 50 MSEG

GRAFICO #8: SEGMENTO DE VOZ DE LA LETRA O/2

AMPLITUD NORMALIZADA



TIEMPO = 50 MSEG

GRAFICO #9: SEGMENTO DE VOZ DE LA LETRA U/1

ESCUELA POLITECNICA NACIONAL
FACULTAD DE INGENIERIA ELECTRICA
TRABAJO DE TESIS - JUNIO 1982
PROGRAMADO POR: CARLOS E. TOBON
DETECCION DEL PERIODO DE TONO DEL SEGMENTO DE VOZ: LETRA U/1
PRUEBA #9.2 FECHA DE REALIZACION: 01-JUL-82
=====

DETECCION DE TONO POR REDUCCION DE INFORMACION

NUCLEOS SILABICOS DE LA REGION CONTINUA 1

M (356) = 5.738

LIMITES SILABICOS DE LA REGION CONTINUA 1

M1(22) = 1.804 M1(485) = 2.662

CICLOS PRINCIPALES DE LA REGION CONTINUA 1

C3(25) = 1.804 C3(51) = 3.242
C3(86) = 5.264 C3(175) = 5.215
C3(268) = 5.275 C3(356) = 5.738
C3(442) = 5.680

ESTIMACIONES DE TONO DE LA REGION CONTINUA 1

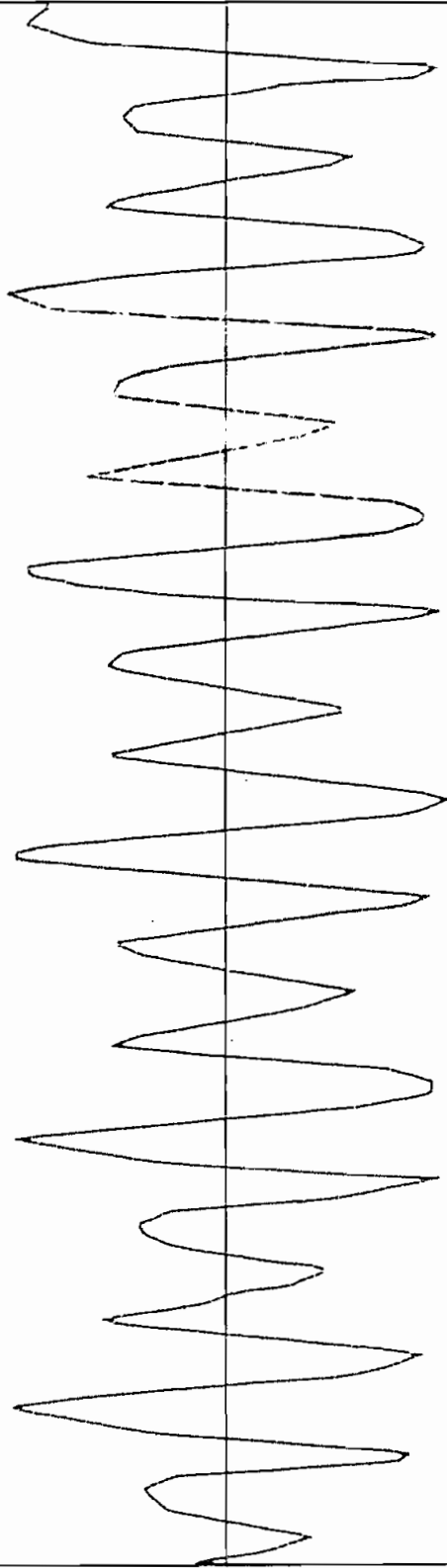
E(1, 1) = 2.200 E(1, 2) = 3.500
E(1, 3) = 9.000 E(1, 4) = 9.000
E(1, 5) = 9.100 E(1, 6) = 8.600

PERIODOS DE TONO DE LA REGION CONTINUA 1

E(1, 1) = 9.000 MSEG
E(1, 2) = 9.000 MSEG
E(1, 3) = 9.100 MSEG
E(1, 4) = 8.600 MSEG

TIEMPO DE EJECUCION = 126 SEGUNDOS

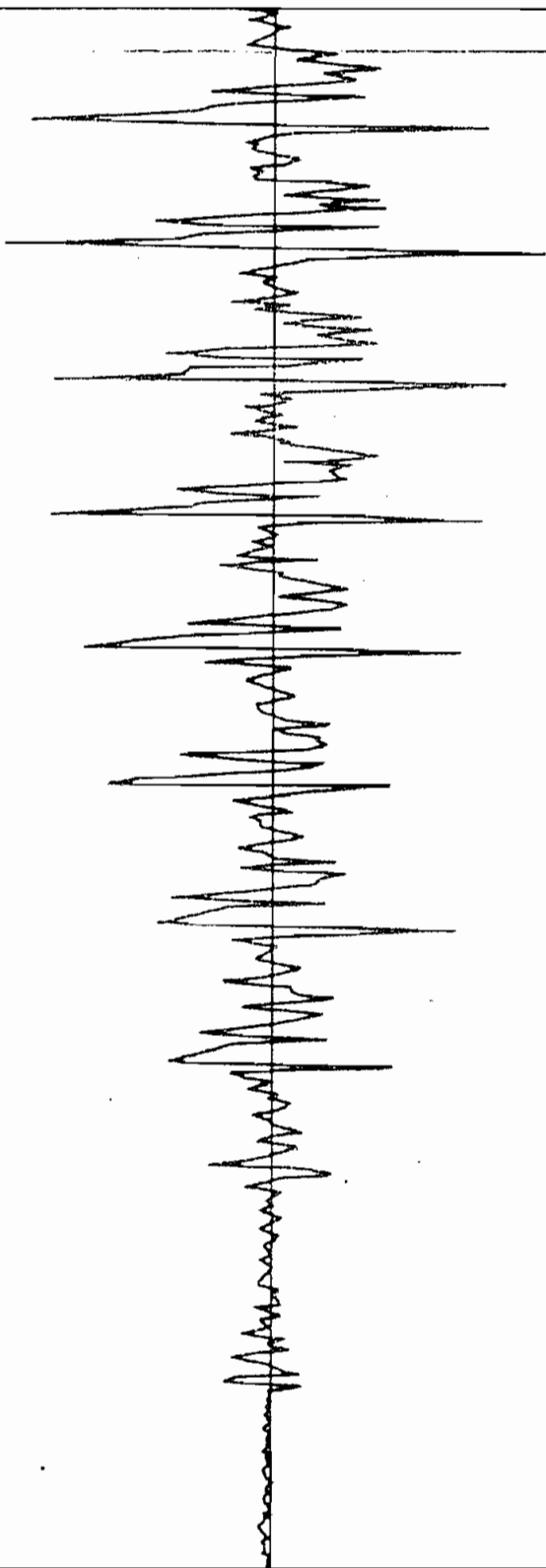
AMPLITUD NORMALIZADA



TIEMPO = 50 MSEG

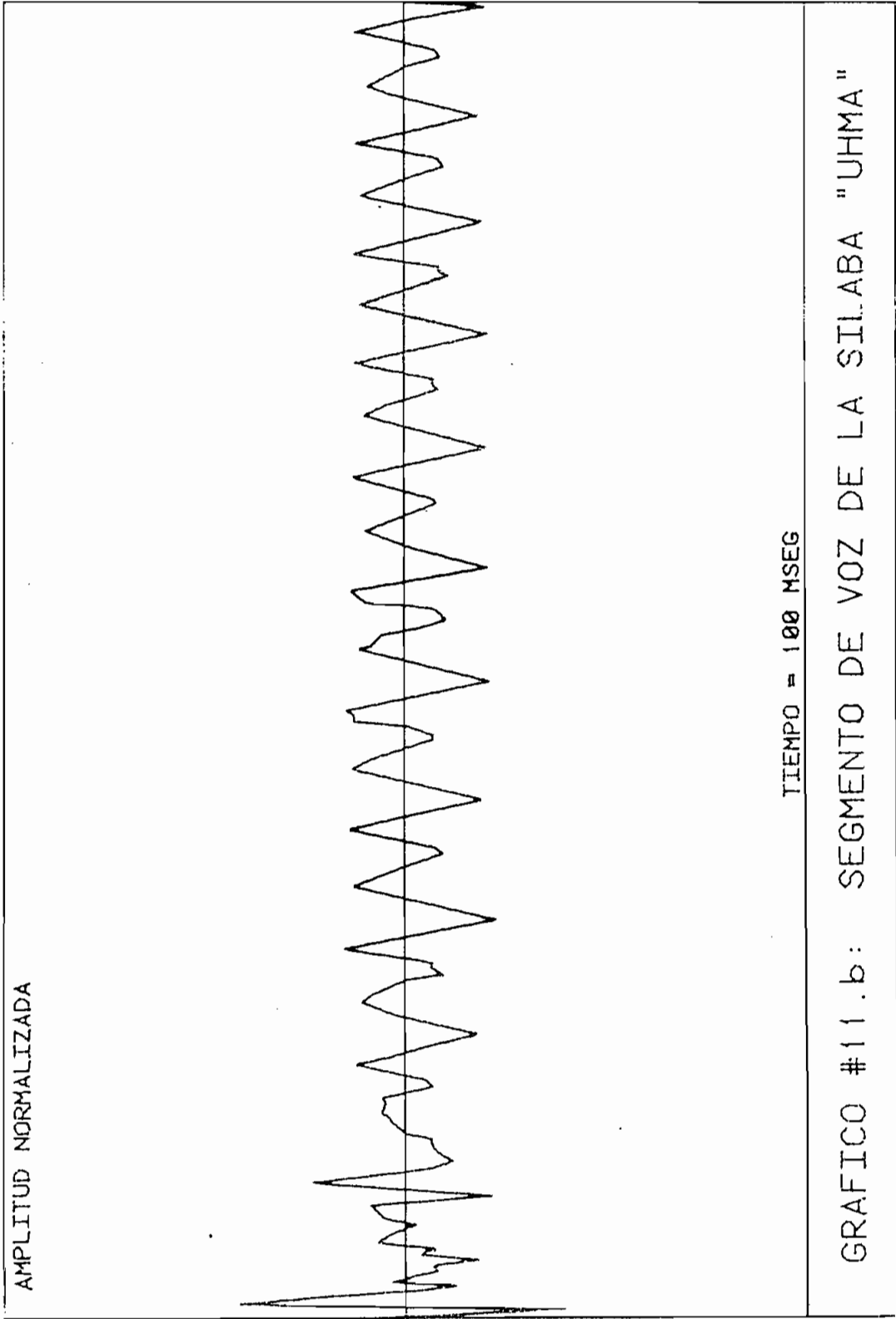
GRAFICO #10: SEGMENTO DE VOZ DE LA LETRA U/2

AMPLITUD NORMALIZADA

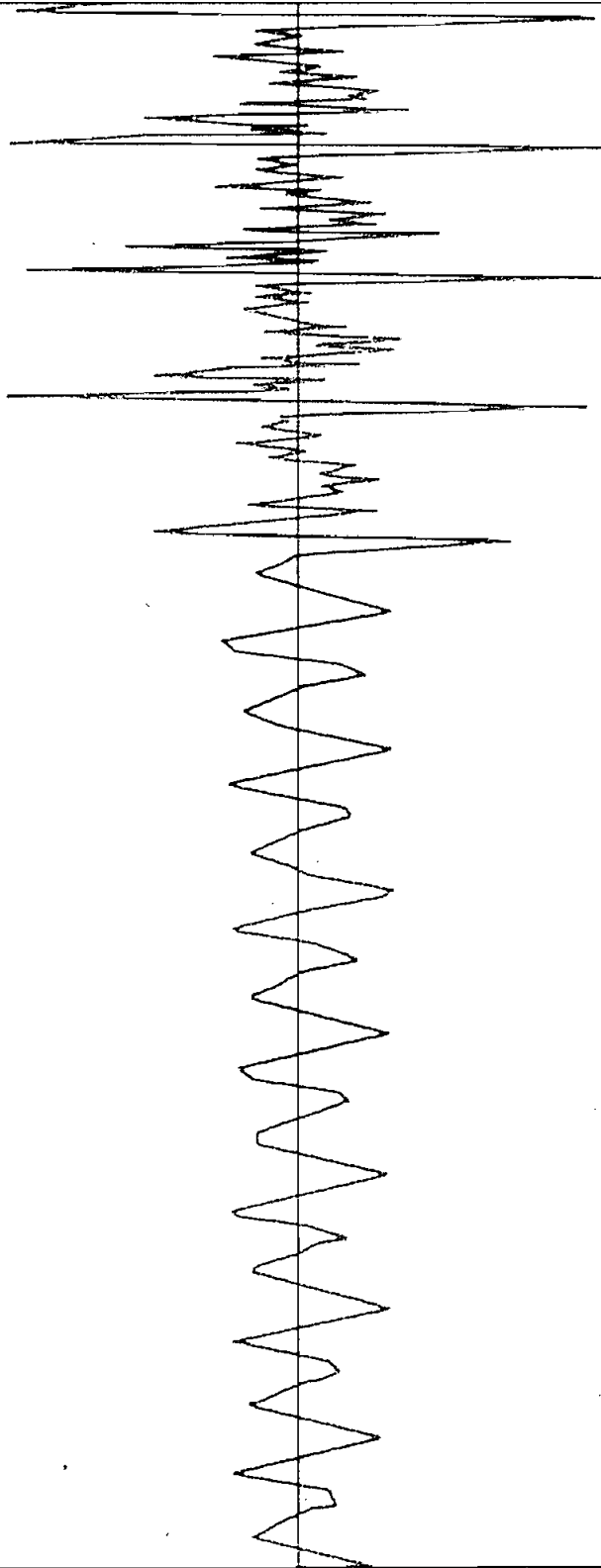


TIEMPO = 100 MSEG

GRAFICO #11.a: SEGMENTO DE VOZ DE LA SILABA "UHMA"

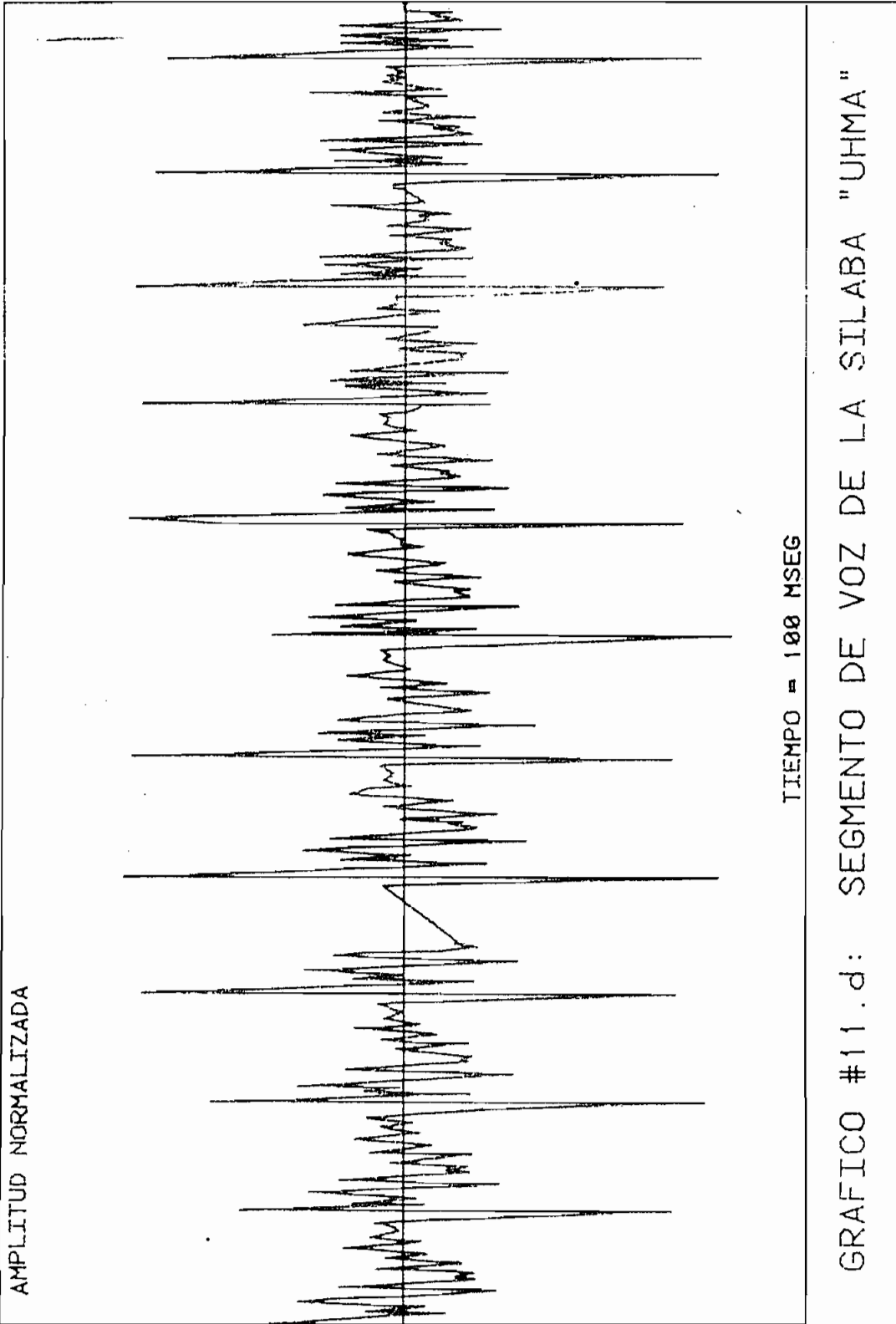


AMPLITUD NORMALIZADA



TIEMPO = 100 MSEG

GRAFICO #11.c: SEGMENTO DE VOZ DE LA SILABA "UHMA"



ESTIMACIONES DE TONO DEL TREN DE IMPULSOS M5

| | | | |
|----------|--------|----------|--------|
| E5(1) = | 17.100 | E5(2) = | 8.400 |
| E5(3) = | 72.800 | E5(4) = | 9.100 |
| E5(5) = | 8.600 | E5(6) = | 8.500 |
| E5(7) = | 8.000 | E5(8) = | 8.200 |
| E5(9) = | 8.800 | E5(10) = | 8.400 |
| E5(11) = | 8.700 | E5(12) = | 9.000 |
| E5(13) = | 9.100 | E5(14) = | 8.900 |
| E5(15) = | 8.900 | E5(16) = | 4.400 |
| E5(17) = | 6.700 | E5(18) = | 10.000 |
| E5(19) = | 8.300 | E5(20) = | 26.200 |
| E5(21) = | 8.300 | E5(22) = | 4.000 |
| E5(23) = | 4.800 | E5(24) = | 18.200 |
| E5(25) = | 24.400 | E5(26) = | 10.300 |
| E5(27) = | 8.700 | | |

ESTIMACIONES DE TONO DEL TREN DE IMPULSOS M6

| | | | |
|----------|-------|----------|--------|
| E6(1) = | 8.200 | E6(2) = | 8.300 |
| E6(3) = | 8.200 | E6(4) = | 9.700 |
| E6(5) = | 8.300 | E6(6) = | 8.300 |
| E6(7) = | 8.800 | E6(8) = | 8.900 |
| E6(9) = | 9.300 | E6(10) = | 34.700 |
| E6(11) = | 8.700 | | |

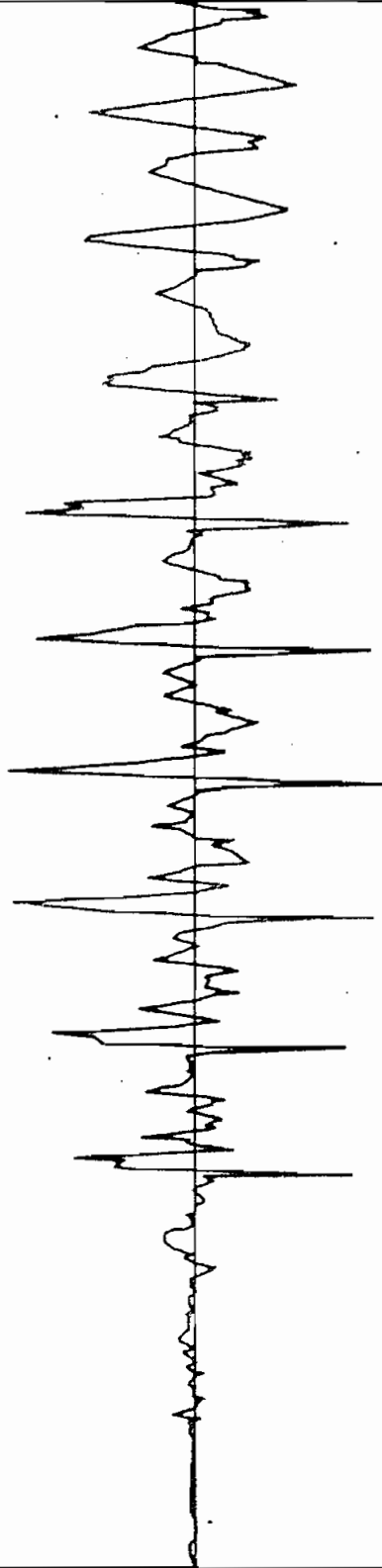
ESTIMACIONES FINALES DEL PERIODO DE TONO

| | | | | |
|---------|---------|------|-----|---|
| E(1) = | 165.200 | MSEG | N = | 2 |
| E(2) = | 0.000 | MSEG | N = | 0 |
| E(3) = | 0.000 | MSEG | N = | 0 |
| E(4) = | 0.000 | MSEG | N = | 0 |
| E(5) = | 0.000 | MSEG | N = | 0 |
| E(6) = | 8.400 | MSEG | N = | 1 |
| E(7) = | 0.000 | MSEG | N = | 0 |
| E(8) = | 0.000 | MSEG | N = | 0 |
| E(9) = | 0.000 | MSEG | N = | 0 |
| E(10) = | 17.900 | MSEG | N = | 1 |
| E(11) = | 0.000 | MSEG | N = | 0 |
| E(12) = | 8.800 | MSEG | N = | 1 |
| E(13) = | 17.200 | MSEG | N = | 1 |
| E(14) = | 0.000 | MSEG | N = | 0 |
| E(15) = | 0.000 | MSEG | N = | 0 |
| E(16) = | 0.000 | MSEG | N = | 0 |
| E(17) = | 0.000 | MSEG | N = | 0 |
| E(18) = | 0.000 | MSEG | N = | 0 |
| E(19) = | 0.000 | MSEG | N = | 0 |
| E(20) = | 0.000 | MSEG | N = | 0 |
| E(21) = | 0.000 | MSEG | N = | 0 |
| E(22) = | 0.000 | MSEG | N = | 0 |
| E(23) = | 0.000 | MSEG | N = | 0 |
| E(24) = | 0.000 | MSEG | N = | 0 |
| E(25) = | 0.000 | MSEG | N = | 0 |
| E(26) = | 0.000 | MSEG | N = | 0 |
| E(27) = | 0.000 | MSEG | N = | 0 |

N - NUMERO DE COINCIDENCIAS DEL PERIODO DE TONO

TIEMPO DE EJECUCION = 1554 SEGUNDOS

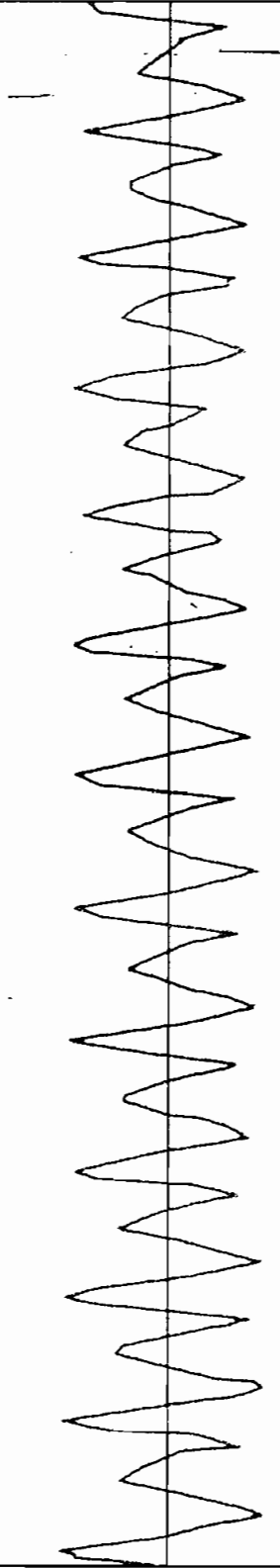
AMPLITUD NORMALIZADA



TIEMPO = 100 MSEG

GRAFICO #12.a: SEGMENTO DE VOZ DE LA SILABA "UHNA"

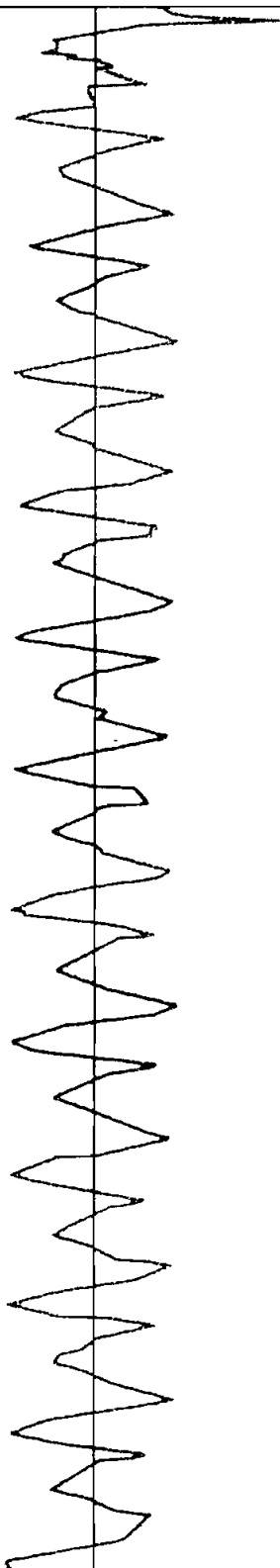
AMPLITUD NORMALIZADA



TIEMPO = 100 MSEG

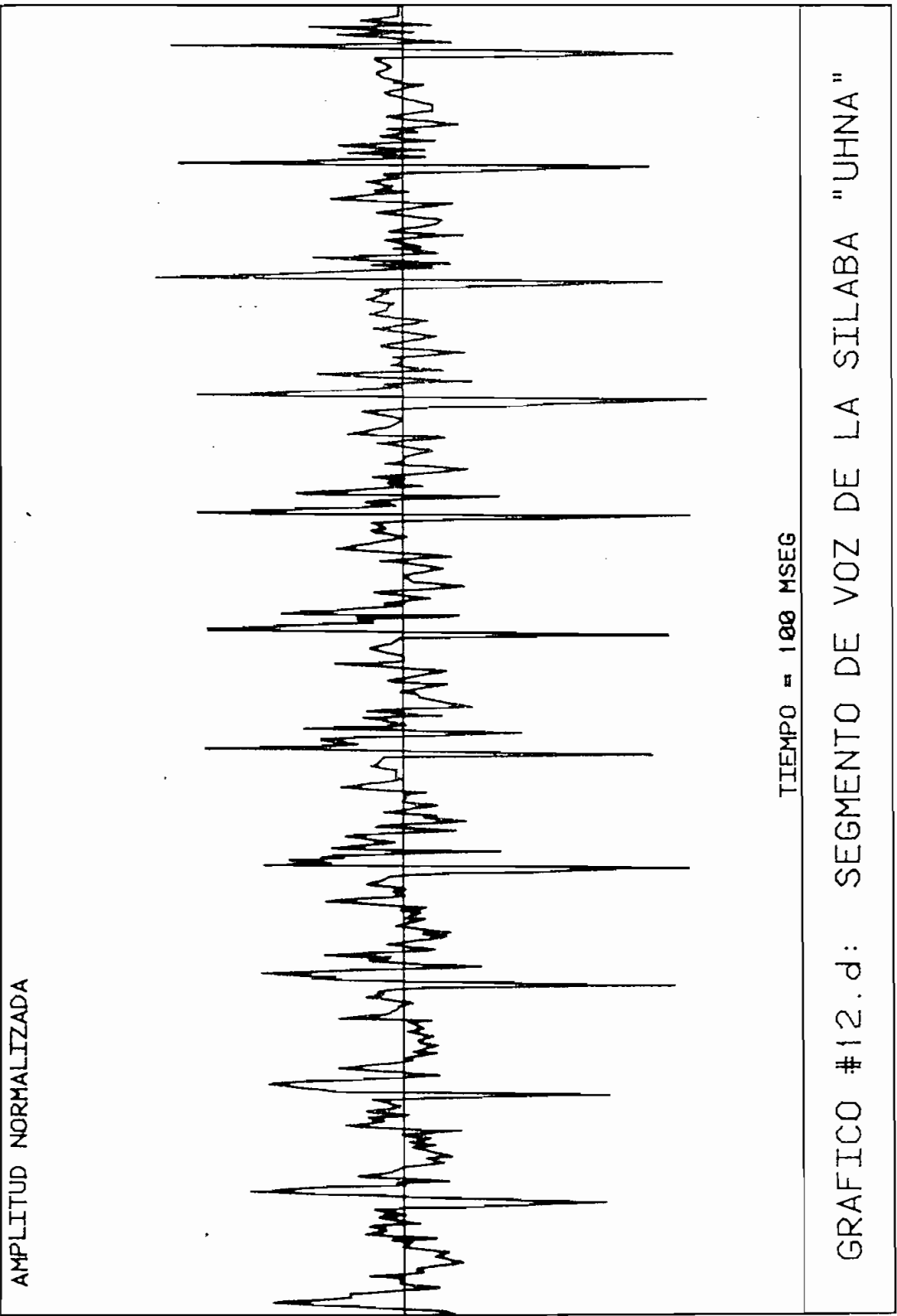
GRAFICO #12.b: SEGMENTO DE VOZ DE LA SILABA "UHNA"

AMPLITUD NORMALIZADA



TIEMPO = 100 MSEG

GRAFICO #12.c: SEGMENTO DE VOZ DE LA SILABA "UHNA"



ESCUELA POLITECNICA NACIONAL

FACULTAD DE INGENIERIA ELECTRICA

TRABAJO-DE TESIS - JUNIO 1982

PROGRAMADO POR: CARLOS E. TORON

DETECCION DEL PERIODO DE TONO DEL SEGMENTO DE VOZ: SILAB/UHNA

PRUEBA #12.1

FECHA DE REALIZACION: 28-JUN-82

DETECCION DE TONO POR PROCESAMIENTO EN PARALELO

ESTIMACIONES DE TONO DEL TREN DE IMPULSOS M1

| | | | |
|----------|--------|----------|--------|
| E1(1) = | 8.300 | E1(2) = | 8.300 |
| E1(3) = | 16.400 | E1(4) = | 41.900 |
| E1(5) = | 8.000 | E1(6) = | 16.500 |
| E1(7) = | 74.900 | E1(8) = | 4.600 |
| E1(9) = | 34.100 | E1(10) = | 20.800 |
| E1(11) = | 24.700 | E1(12) = | 12.800 |
| E1(13) = | 34.800 | E1(14) = | 7.200 |
| E1(15) = | 6.000 | E1(16) = | 4.900 |

ESTIMACIONES DE TONO DEL TREN DE IMPULSOS M2

| | | | |
|----------|--------|----------|--------|
| E2(1) = | 8.300 | E2(2) = | 82.800 |
| E2(3) = | 8.400 | E2(4) = | 8.500 |
| E2(5) = | 8.200 | E2(6) = | 16.400 |
| E2(7) = | 8.300 | E2(8) = | 16.800 |
| E2(9) = | 16.700 | E2(10) = | 16.900 |
| E2(11) = | 8.400 | E2(12) = | 8.900 |
| E2(13) = | 8.400 | E2(14) = | 8.500 |
| E2(15) = | 8.400 | E2(16) = | 16.300 |
| E2(17) = | 8.000 | E2(18) = | 8.600 |
| E2(19) = | 8.300 | E2(20) = | 8.500 |
| E2(21) = | 8.700 | E2(22) = | 27.200 |
| E2(23) = | 8.800 | E2(24) = | 17.700 |
| E2(25) = | 9.000 | | |

ESTIMACIONES DE TONO DEL TREN DE IMPULSOS M3

| | | | |
|----------|--------|----------|-------|
| E3(1) = | 9.100 | E3(2) = | 8.800 |
| E3(3) = | 8.900 | E3(4) = | 8.900 |
| E3(5) = | 17.700 | | |

ESTIMACIONES DE TONO DEL TREN DE IMPULSOS M4

| | | | |
|---------------------|-------------------|---------------------|-------------------|
| E4(1) = | 8.100 | E4(2) = | 8.300 |
| E4(3) = | 8.400 | E4(4) = | 69.500 |
| E4(5) = | 78.700 | E4(6) = | 37.900 |
| E4(7) = | 25.700 | E4(8) = | 21.500 |
| E4(9) = | 13.400 | E4(10) = | 39.500 |
| E4(11) = | 12.000 | E4(12) = | 5.200 |

ESTIMACIONES DE TONO DEL TREN DE IMPULSOS M5

| | | | |
|----------|--------|----------|--------|
| E5(1) = | 6.500 | E5(2) = | 8.300 |
| E5(3) = | 8.400 | E5(4) = | 36.500 |
| E5(5) = | 62.600 | E5(6) = | 8.400 |
| E5(7) = | 28.100 | E5(8) = | 4.500 |
| E5(9) = | 17.100 | E5(10) = | 20.500 |
| E5(11) = | 4.600 | E5(12) = | 25.900 |
| E5(13) = | 25.100 | E5(14) = | 8.100 |
| E5(15) = | 9.800 | E5(16) = | 15.600 |
| E5(17) = | 7.800 | E5(18) = | 4.400 |
| E5(19) = | 3.100 | E5(20) = | 20.800 |
| E5(21) = | 25.800 | E5(22) = | 4.700 |
| E5(23) = | 4.100 | | |

ESTIMACIONES DE TONO DEL TREN DE IMPULSOS M6

| | | | |
|----------|--------|----------|--------|
| E6(1) = | 9.000 | E6(2) = | 17.700 |
| E6(3) = | 9.100 | E6(4) = | 8.800 |
| E6(5) = | 26.300 | | |

ESTIMACIONES FINALES DEL PERIODO DE TONO

| | | | |
|---------|-------------|-----|---|
| E(1) = | 8.300 MSEG | N = | 2 |
| E(2) = | 8.300 MSEG | N = | 2 |
| E(3) = | 16.400 MSEG | N = | 1 |
| E(4) = | 0.000 MSEG | N = | 0 |
| E(5) = | 0.000 MSEG | N = | 0 |
| E(6) = | 0.000 MSEG | N = | 0 |
| E(7) = | 0.000 MSEG | N = | 0 |
| E(8) = | 0.000 MSEG | N = | 0 |
| E(9) = | 0.000 MSEG | N = | 0 |
| E(10) = | 0.000 MSEG | N = | 0 |
| E(11) = | 0.000 MSEG | N = | 0 |
| E(12) = | 0.000 MSEG | N = | 0 |
| E(13) = | 0.000 MSEG | N = | 0 |
| E(14) = | 0.000 MSEG | N = | 0 |
| E(15) = | 0.000 MSEG | N = | 0 |
| E(16) = | 0.000 MSEG | N = | 0 |
| E(17) = | 0.000 MSEG | N = | 0 |
| E(18) = | 0.000 MSEG | N = | 0 |
| E(19) = | 0.000 MSEG | N = | 0 |
| E(20) = | 0.000 MSEG | N = | 0 |
| E(21) = | 0.000 MSEG | N = | 0 |
| E(22) = | 0.000 MSEG | N = | 0 |
| E(23) = | 0.000 MSEG | N = | 0 |
| E(24) = | 0.000 MSEG | N = | 0 |
| E(25) = | 0.000 MSEG | N = | 0 |

ESCUELA POLITECNICA NACIONAL
FACULTAD DE INGENIERIA ELECTRICA
TRABAJO DE TESIS - JUNIO 1982
PROGRAMADO POR: CARLOS E. TORON
DETECCION DEL PERIODO DE TONO DEL SEGMENTO DE VOZ: SILAB/UHNA
PRUEBA #12.2 FECHA DE REALIZACION: 28-JUN-82
=====

DETECCION DE TONO POR REDUCCION DE INFORMACION

NUCLEOS SILABICOS DE LA REGION CONTINUA 1

M (502) = 4.401 M (3700) = 8.573

LIMITES SILABICOS DE LA REGION CONTINUA 1

M1(20) = 0.000 M1(1938) = 1.763
M1(3985) = 0.811

CICLOS PRINCIPALES DE LA REGION CONTINUA 1

| | |
|------------------|------------------|
| C3(20) = 0.000 | C3(194) = 0.479 |
| C3(335) = 3.513 | C3(418) = 4.162 |
| C3(502) = 4.401 | C3(587) = 4.130 |
| C3(668) = 3.595 | C3(867) = 2.195 |
| C3(1033) = 2.213 | C3(1116) = 2.229 |
| C3(1197) = 2.176 | C3(1277) = 1.891 |
| C3(1361) = 1.990 | C3(1447) = 2.056 |
| C3(1532) = 1.892 | C3(1613) = 1.800 |
| C3(1697) = 1.769 | C3(1778) = 1.794 |
| C3(1858) = 1.824 | C3(1938) = 1.763 |
| C3(2110) = 1.861 | C3(2197) = 1.807 |
| C3(2279) = 1.779 | C3(2360) = 1.636 |
| C3(2449) = 1.785 | C3(2534) = 1.704 |
| C3(2584) = 1.501 | C3(2620) = 1.838 |
| C3(2704) = 1.812 | C3(2752) = 1.623 |
| C3(2787) = 1.943 | C3(2991) = 4.343 |
| C3(3087) = 5.694 | C3(3170) = 5.805 |
| C3(3254) = 7.624 | C3(3344) = 8.044 |
| C3(3430) = 7.045 | C3(3521) = 7.471 |
| C3(3612) = 8.067 | C3(3700) = 8.573 |
| C3(3789) = 7.334 | C3(3877) = 6.945 |
| C3(3963) = 7.630 | |

ESTIMACIONES DE TONO DE LA REGION CONTINUA 1

| | | | |
|------------|--------|------------|--------|
| E(1, 1) = | 16,900 | E(1, 2) = | 14,200 |
| E(1, 3) = | 8,000 | E(1, 4) = | 8,500 |
| E(1, 5) = | 8,400 | E(1, 6) = | 8,400 |
| E(1, 7) = | 19,400 | E(1, 8) = | 16,300 |
| E(1, 9) = | 8,400 | E(1,10) = | 8,100 |
| E(1,11) = | 8,300 | E(1,12) = | 8,000 |
| E(1,13) = | 8,400 | E(1,14) = | 8,800 |
| E(1,15) = | 8,300 | E(1,16) = | 8,200 |
| E(1,17) = | 8,100 | E(1,18) = | 8,200 |
| E(1,19) = | 7,900 | | |

PERIODOS DE TONO DE LA REGION CONTINUA 1

| | |
|------------|-------------|
| E(1, 1) = | 16,900 MSEG |
| E(1, 2) = | 14,200 MSEG |
| E(1, 3) = | 8,000 MSEG |
| E(1, 4) = | 8,500 MSEG |
| E(1, 5) = | 8,400 MSEG |
| E(1, 6) = | 8,400 MSEG |
| E(1, 7) = | 19,400 MSEG |
| E(1, 8) = | 16,300 MSEG |
| E(1, 9) = | 8,400 MSEG |
| E(1,10) = | 8,100 MSEG |
| E(1,11) = | 8,300 MSEG |
| E(1,12) = | 8,000 MSEG |
| E(1,13) = | 8,400 MSEG |
| E(1,14) = | 8,800 MSEG |
| E(1,15) = | 8,300 MSEG |
| E(1,16) = | 8,200 MSEG |
| E(1,17) = | 8,100 MSEG |
| E(1,18) = | 8,200 MSEG |
| E(1,19) = | 7,900 MSEG |

TIEMPO DE EJECUCION = 900 SEGUNDOS

Debo aclarar que los tiempos de ejecución de los programas incluyen el tiempo requerido para la impresión de resultados y el ingreso del nombre del segmento de voz. Esto significa que estos tiempos de ejecución de los algoritmos son totales. Si se toma en cuenta que el computador es más lento para las operaciones de impresión, se puede concluir que el tiempo de ejecución del algoritmo es en realidad igual a las 2/3 partes del tiempo total de la ejecución.

En la página 39 del Apéndice "A" se dispone de una lista completa de los archivos contenidos en el disco correspondiente a esta tesis. Todos los archivos comienzan con "@VOZ/" corresponden a los diferentes segmentos de voz disponibles en disco.

4.2 Conclusiones

En base a los algoritmos expuestos y a las pruebas de los programas desarrollados para el computador, se concluye lo siguiente:

1. Los fundamentos de los algoritmos para detección del período de tono, están presentados en la forma más clara y ordenada, de modo que el lector puede interpretarlos con facilidad. La comprensión de ciertos aspectos requiere de un estudio previo de las diferentes características fonológicas de la voz humana. Para un mejor entendimiento se han evadido ciertos tópicos que en vez de aclarar la teoría podría confundir o desviar la atención del objetivo concreto

de este trabajo de tesis.

2. A partir de estos algoritmos se han podido desarrollar dos programas para detección del período de tono del segmento de voz, tales que puedan aplicarse a voces de diferentes locutores para el análisis de las características fonológicas de la voz.
3. Los programas para el computador digital se han implementado de tal forma que puedan ser utilizadas para cualquier tipo de aplicaciones, dentro de las limitaciones de memoria del equipo.
4. Las técnicas de análisis de memoria implementadas para estos programas permiten establecer la información necesarias para evadir una saturación de memoria (Memory Full) por medio de transferencias de información entre memoria y disco. en cualquier instante que se requiera.
5. El número de transferencias entre la memoria y el disco durante la ejecución de los programas, puede reducirse únicamente aumentando la capacidad de memoria real del sistema de computación.
6. Entre mayor sea el número de transferencias entre el disco y la memoria o viceversa, mayor va a ser el tiempo de ejecución de los programas.
7. En caso de presentarse una saturación de la memoria por disponer de un segmento de voz con un número excesivo de muestras, puede resolverse al problema de dos maneras di-

ferentes. Una sería disminuir la frecuencia de muestreo del segmento de voz, de tal forma que se reduzcan en lo posible el número de muestras a procesarse. Esto podría realizarse tomando en cuenta que no se vaya a perder considerablemente en la exactitud de los resultados, por el distanciamiento entre las muestras. La otra posibilidad es la de segmentar la onda de voz en segmentos de menor duración, permitiéndose la ejecución de los programas dentro del mismo grado de exactitud.

8. La frecuencia de muestreo de los segmentos de voz puede variar entre 7 KHz y unos 50 KHz. Estableciéndose como frecuencia ideal de muestreo la de 10 KHz, por el reducido número de muestras y la exactitud de la información que proporciona este tipo de muestreo.
9. Los tiempos de ejecución de los programas no son muy rápidos, esto se debe primordialmente a la pequeña capacidad de memoria de que se dispone, la que hace necesaria la realización de transferencias de información entre el disco y la memoria.
10. La relación entre los tiempos de ejecución de los programas es aproximadamente de 1 a 1.65, es decir que el programa de detección por reducción de información (DETRED) es más rápido que el de detección por procesamiento en paralelo (DETPAR).
11. La limitación de ambos programas va a depender únicamente de la capacidad de memoria real del computador. Con

los 30 KBytes de memoria disponibles actualmente en este computador, se va a permitir el procesamiento de segmentos de voz de una duración máxima de 400 o 500 milisegundos muestreados a 10 KHz; dependiendo del número de máximos y mínimos que tenga la onda, en el caso del "DETPAR" y del número de ciclos de excursión que tenga la onda, para el caso del "DETRED".

12. En el programa de detección por procesamiento en paralelo un aumento de la constante de coincidencia, produce una ventana de coincidencia más grande, lo que daría como resultado un aumento del número de coincidencias de ciertos períodos de tono.
13. Los errores de duplicación del período de tono presentes en dos de las pruebas realizadas con el programa de detección por procesamiento en paralelo, se deben a que para estos dos casos en particular el porcentaje $P = 0.7$ utilizado para todas las pruebas, alcanzó a eliminar una de las amplitudes máximas significativas para el cálculo del período de tono de la señal correspondiente. Esta situación pudo ser obviada reduciendo el porcentaje P de eliminación de amplitudes no significativas, para estos dos casos en particular, pero el objetivo de este trabajo era el de establecer una regla general sin caer en el análisis de situaciones particulares.
14. En las pruebas del programa de detección del tono por reducción de información, se presentaron dos casos de dupli-

cación del período de tono y un caso de división para dos del período de tono. Estos errores podían haber sido corregidos disminuyendo o aumentando el porcentaje de referencia con la amplitud de la envolvente considerado para la eliminación de los ciclos de excursión que no sean ciclos principales, respectivamente. En el primer caso, la amplitud de un ciclo de excursión que debía ser considerado como ciclo principal, fue menor que el 80% de la envolvente; mientras que en el segundo caso se consideró como ciclo principal a algún ciclo de excursión cuya amplitud superó a éste porcentaje para ser considerado erróneamente como un ciclo principal.

15. Los resultados gráficos obtenidos para el ingreso de los segmentos de voz por medio de la técnica de interpolación, a partir de los puntos de inflexión o quiebre ingresados con el grafizador, son totalmente satisfactorios. Esto se puede apreciar claramente comparando los gráficos de la sección 4.1 con sus correspondientes figuras del capítulo 1.
16. Los resultados obtenidos para las diferentes pruebas realizadas con los dos programas de detección de tono son muy satisfactorios, ya que se ha conseguido una precisión bastante aceptable, para los diferentes segmentos de voz analizados.
17. En caso de que para una aplicación específica se necesite

una muy buena exactitud en los resultados de los programas de detección de tono, se debe realizar una regresión lineal de los valores que se obtengan para mejorar la exactitud de éstos. Si no basta con este proceso para mejorar los resultados, se debe realizar una suavización de los resultados utilizando ya sea una ventana rectangular o una ventana de Hamming, para posteriormente repetir la regresión lineal para la corrección de los errores.

18. Quedan implementados estos programas para la detección del período de tono, los cuales deberán ser utilizados para realizar pruebas del sistema para adquisición de datos, para el análisis digital de fonemas, que se está desarrollando y como un elemento para futuros análisis de la fonología española.
19. Como complemento de este trabajo de tesis, dejo estructurada una biblioteca de archivos con datos de voz inglesa, la cual podrá ser utilizada en los diferentes trabajos de proyectos y tesis que se están desarrollando y que se vayan a desarrollar en un futuro, relacionados con estos aspectos de comunicación o análisis de señales de voz.

Para terminar me permito sugerir que se continúe con el desarrollo de estos proyectos relacionados con el análisis de señales de voz, hasta que se tengan los elementos suficientes para el desarrollo de sistemas codificadores de voz, para la transmisión de señales de voz en función de sus parámetros.

4.3. Metodología de Uso de los Programas

Para la utilización de los diferentes programas desarrollados en este programa de tesis se deben seguir los siguientes pasos:

1. Prenda el computador, las unidades de disco, el impresor y el grafizador si van a ser utilizados.
2. Coloque el disco en una de las unidades. En caso de ir a utilizar otro disco con archivos de datos colóquelo en la otra unidad.
3. Inicialice el sistema de reloj del computador por medio del teclado, escribiendo:

```
CALL "SETTIM","DD-MMM-AA ¯ HH:MM:SS"
```

luego oprima la tecla RETURN ,

donde:

DD : día

MMM : tres primeras iniciales del mes en Inglés.

AA : año

¯ : espacio en blanco

HH : hora

MM : minutos

SS : segundos (opcional)

4. Carque el disco de los programas por medio del teclado, escribiendo:

```
CALL "MOUNT", No. , X$
```

luego oprima la tecla RETURN,

donde:

No. : número de la unidad de disco en la que se ha
colocado el disco con los programas.

5. Cargue a la memoria del computador el programa maestro por medio del teclado, escribiendo:
OLD "TESIS7CTOBON"
luego oprima la tecla RETURN.
6. Haga correr el programa escribiendo:
RUN
y oprima la tecla RETURN.
7. Continúe la ejecución siguiendo las instrucciones que le vaya dando el computador a través de la pantalla. Cada vez que tenga que dar un comando o ingresar un dato al computador, estos deben ser tecleados y enseguida se debe oprimir la tecla RETURN, excepto cuando salga el índice de programas.
8. En caso de que salga el índice de programas en la pantalla, después de escoger el programa deseado, sólo hay que oprimir la tecla deseada. Estas teclas son aquellas que están ubicadas en la parte superior izquierda del teclado. Al final de cualquier programa o en caso de alguna interrupción, se puede llamar al índice de programas oprimiendo la TECLA 1.
9. Cada vez que aparezca una pequeña "F" en forma intermitente, en la esquina superior izquierda de la pantalla, esto significa que la pantalla está llena, razón por la cual se debe oprimir la tecla HOME PAGE.

10. Si disminuye la intensidad de la pantalla, basta con oprimir la tecla (SHIFT), para que recupere su luminosidad.
11. Si en algún momento se desea interrumpir la ejecución del programa, basta con oprimir una o dos veces la tecla BREAK. Al oprimir una sola vez esta tecla, el computador va a terminar la tarea que esté ejecutando, para luego pararse. En cambio, si se oprime dos veces, el computador termina la instrucción en que está para luego pararse. Cuando se oprima dos veces esta tecla pueden surgir errores al tratar de ejecutar de nuevo el programa, debido a que se cerraron algunos archivos que se suponían abiertos.
12. En caso de que aparezca en la pantalla un mensaje de MEMORY FULL, esto significa que se ha saturado la memoria real del computador. Para reiniciar la ejecución, lo más adecuado es repetir el proceso de escrito entre los numerales 5 al 7. A lo mejor corre sin problemas el programa; de lo contrario se está utilizando una cantidad excesiva de datos. Las soluciones para este problema están expuestas en las conclusiones de esta tesis.
13. Antes de apagar el computador, se debe escribir por medio del teclado lo siguiente:
DEL ALL
y se debe oprimir la tecla HOME PAGE.

A P E N D I C E

```

1 U0=-1
4 GO TO 100
8 Q1=1
9 GO TO 1000
12 Q1=2
13 GO TO 1000
16 Q1=3
17 GO TO 1000
100 REM      *****          TESIS/CTOBON          *****
101 IF U0<>-1 THEN 400
110 PRINT 'JJ INGRESE UNIDAD DONDE ESTA EL DISCO:  ':
120 INPUT U1
130 CALL 'UNIT',U1
140 U0=0
400 PRINT 'LJ      TECNICAS DIGITALES PARA LA DETECCION DEL TONO DE VOZ'
410 PRINT '      *****'
420 PRINT 'JJJ      TECLA 1 -- INDICE DE PROGRAMAS'
430 PRINT 'J       TECLA 2 -- INGRESO DE DATOS'
440 PRINT 'J       TECLA 3 -- DETECCION POR PROCESAMIENTO EN PARALELO'
450 PRINT 'J       TECLA 4 -- DETECCION POR REDUCCION DE INFORMACION'
460 PRINT 'JJ      ESCOJA TECLA GGGG'
470 END
500 REM      ***** LINEA 500 *****
510 REM      *** CARGA AUTOMATICA DE PROGRAMAS ***
520 DATA 'DATOS','DETFAR','DETRER'
530 RESTORE 520
540 FOR J=1 TO Q1
550 READ A#
560 NEXT J
570 DELETE 1010,20000
580 J=MEMORY
590 APPEND A#:1000
1000 REM      ***** LINEA 1000 *****
1010 REM      ***** CARGA AUTOMATICA DE PROGRAMAS ***
1020 Q2=0
1030 IF Q1<>Q2 THEN 500
1040 END

```

```

1000 REM ***** PROGRAMA PARA INGRESO DE DATOS (DATOS) *****
1010 Q2=1
1020 IF Q1<>Q2 THEN 500
1030 PRINT "LJI INGRESO DE DATOS"
1040 PRINT "I *****"
1050 PRINT "JJJ CLASE 1 -- INGRESO MANUAL DE LAS MUESTRAS"
1060 PRINT "J CLASE 2 -- INGRESO DESDE UN ARCHIVO DE DISCO"
1070 PRINT "J CLASE 3 -- INGRESO POR MEDIO DE INTERPOLACION"
1080 PRINT "J CLASE 4 -- IMPRESION,GRAFIZACION,ALMACENAMIENTO"
1090 PRINT "JJGGG OPRIMA EL NUMERO DE CLASE DESEADO: ";
1100 INPUT Q
1110 REM *** INICIALIZACION DE CONSTANTES Y BANDERAS ***
1120 K4=2000
1130 Q3=0
1140 D2=0
1150 Q1=0
1160 GO TO Q OF 1190,1370,1880,1170
1170 Q=Q0
1180 GO TO 2840
1190 REM ***** PROGRAMA DE INGRESO MANUAL DE DATOS *****
1200 PRINT "LJI INGRESO MANUAL DE DATOS"
1210 PRINT "I *****"
1220 GOSUB 4540
1230 IF D4=1 THEN 1270
1240 PRINT "JJGGG ESTA DIMENSION ES EXCESIVA PARA LA CAPACIDAD REAL"
1250 PRINT "JJGGG REDUZCA EL NUMERO DE MUESTRAS A INGRESAR"
1260 GO TO 1220
1270 DELETE A1
1280 DIM A1(D3)
1290 PRINT "LJJ INGRESE LOS DATOS DE VOZ SECUENCIALMENTE"
1300 PRINT "GGG *****"
1310 INPUT A1
1320 PRINT "JJGGG FIN DEL INGRESO DE DATOS"
1330 M$="@MUESTREG/A1"
1340 N=1
1350 GOSUB 5250
1360 GOSUB 2840
1370 REM ***** PROGRAMA DE INGRESO DE DATOS DESDE UN ARCHIVO *****
1380 PRINT "LJ INGRESO DE DATOS DESDE UN ARCHIVO"
1390 PRINT " *****"
1400 PRINT "JJJGG INGRESE LA UNIDAD DE DISCO EN QUE HAY EL ARCHIVO;"
1410 INPUT U3
1420 CALL "MOUNT",U3,X$
1430 CALL "UNIT",U3
1440 PRINT "JGG SABE EL NOMBRE DEL ARCHIVO DE DATOS (SI O NO): ";
1450 INPUT A$
1460 IF A$="NO" OR A$="N" THEN 1760
1470 IF U3<>U1 THEN 1550
1480 PRINT "JGG INGRESE EL NOMBRE DEL ARCHIVO (SIN '@'); ";
1490 INPUT L$
1500 IF ASC(L$)<>64 THEN 1530
1510 PRINT "J ERROR: NOMBRE NO DEBE TENER '@', REPITAG"
1520 GO TO 1470

```



```

1530 L$='@'&L$
1540 GO TO 1570
1550 PRINT 'JGG  INGRESE EL NOMBRE DEL ARCHIVO:  ';
1560 INPUT L$
1570 CALL 'FILE',U3,L$,Y$
1580 IF Y$="" THEN 1750
1590 PRI 'JJ  ESTAN ALMACENADOS EN ESTE ARCHIVO EL NUMERO DE MUESTREOS
1600 PRI 'J  Y LA FRECUENCIA DE MUESTREO DE LA ONDA DE VOZ?(SI O NO): '
1610 INPUT A$
1620 IF A$='SI' OR A$='S' THEN 1650
1630 D2=1
1640 GOSUB 4540
1650 M$='@MUESTREO/A1'
1660 N=1
1670 CALL 'UNIT',U3
1680 GOSUB 5480
1690 N=N-1
1700 CALL 'UNIT',U1
1710 GOSUB 5250
1720 IF D5<>1 THEN 1670
1730 IF D4=1 THEN 1860
1740 GO TO 1870
1750 PRINT 'JGG  NO EXISTE ESE NOMBRE DE ARCHIVO EN EL DISCOJJ'
1760 PRINT 'JJI'
1770 IF U3<>U1 THEN 1800
1780 DIRECTORY 0,'@VOZ#'
1790 GO TO 1810
1800 DIRECTORY 0,'@#'
1810 PRINT 'JJJGG  IDENTIFICO YA EL NOMBRE DEL ARCHIVO (SI O NO):  ';
1820 INPUT A$
1830 IF A$='SI' OR A$='S' THEN 1470
1840 PRINT 'JGG  ESCOJA CUALQUIER NOMBRE DE ARCHIVO'
1850 GO TO 1470
1860 D2=0
1870 GOSUB 2840
1880 REM      ***** PROGRAMA DE INGRESO DE DATOS POR INTERPOLACION *****
1890 PRINT 'LJI  INGRESO DE DATOS POR INTERPOLACION'
1900 PRINT 'I *****'
1910 PRINT 'JJJ      1 -- INGRESO MANUAL DE LAS ESQUINAS'
1920 PRINT 'J      2 -- INGRESO DE ESQUINAS CON EL GRAFIZADOR'
1930 PRINT 'JJJGG  OPRIMA EL NUMERO DEL METODO DESEADO:  ';
1940 INPUT Q2
1950 A=1
1960 Y1=0
1970 Y2=100
1980 GOSUB 4540
1990 PRINT 'JJGG  ASEGURESE DE INCLUIR LA ESQUINA INICIAL F(Y7)=0'
2000 PRINT 'JGG  Y LA ESQUINA FINAL F(Y7)=';L1
2010 PRI 'JJJ  INGRESE UN VALOR IGUAL O MAYOR AL NUMERO DE ESQUINAS:
2020 INPUT D
2030 REM      *** PROCESO DE INGRESO DE LAS ESQUINAS ***
2040 GOSUB Q2 OF 2410,4870
2050 IF Q2=1 THEN 2070

```

```

2060 GOSUB 2650
2070 M$='@MUESTREO/A1'
2080 N=1
2090 X4=0
2100 Y4=P1(1)
2110 I=1
2120 J=0
2130 D6=0
2140 DELETE A1
2150 DIM A1(D3)
2160 J=J+1
2170 X5=P(J)
2180 Y5=P1(J)
2190 GO TO 2220
2200 I=0
2210 I=I+1
2220 X6=I/F+D6
2230 IF X6>X5 THEN 2340
2240 P2=(Y5-Y4)/(X5-X4)
2250 A1(I)=P2*(X6-X5)+Y5
2260 IF I<D3 THEN 2210
2270 D6=X6
2280 GOSUB 5250
2290 IF D5=0 THEN 2200
2300 PRINT 'JJGG FIN DE LA INTERPOLACION DE DATOS'
2310 PRINT 'JJGG OPRIMA [RETURN] PARA CONTINUAR'
2320 INPUT A$
2330 GO TO 2390
2340 X4=X5
2350 Y4=Y5
2360 IF J<Y7 THEN 2160
2370 X5=L1
2380 GO TO 2230
2390 DELETE P,P1
2400 GOSUB 2840
2410 REM      ***** SUBROUTINA DE INGRESO MANUAL DE ESQUINAS *****
2420 Y7=1
2430 PRINT 'LJI INGRESO MANUAL DE LAS ESQUINAS'
2440 PRINT 'I *****'
2450 PRINT 'JJJGG INGRESE LAS COORDENADAS DE LAS ESQUINAS'
2460 DELETE P,P1
2470 DIM P(D),P1(D)
2480 PRINT 'JJGG INGRESE EL TIEMPO [mseg]: P(';Y7;') = ';
2490 INPUT P(Y7)
2500 PRINT 'JGG INGRESE LA AMPLITUD EN Y: P1(';Y7;') = ';
2510 INPUT P1(Y7)
2520 IF P(Y7)=-1 THEN 2610
2530 IF P(Y7)<0 THEN 2590
2540 IF P(Y7)>L1 THEN 2590
2550 Y7=Y7+1
2560 IF Y7>D THEN 2610
2570 PRINT 'JJGG PARA TERMINAR TIEMPO = (-1)'
2580 GO TO 2480

```

```

2590 PRINT 'JJGG EL DATO P(Y7) INGRESADO ES ERRONEO'
2600 GO TO 2480
2610 PRINT 'JJJGG FIN DEL INGRESO DE LAS ESQUINAS'
2620 Y7=Y7-1
2630 GOSUB 4950
2640 RETURN
2650 REM ***** SUBROUTINA DE INGRESO DE ESQUINAS CON GRAFIZADOR *****
2660 Y7=1
2670 PRINT 'LJJI INGRESO DE ESQUINAS CON EL GRAFIZADOR'
2680 PRINT 'I *****'
2690 PRINT 'JJJGG UBIQUE LA HOJA CON LA ONDA EN EL GRAFIZADOR'
2700 PRINT 'JJGG EL INGRESO DE LAS ESQUINAS DEBE SER SECUENCIAL'
2710 PRINT 'JGG UBIQUE EL VISOR EN LA ESQUINA (PUNTO A INGRESAR)'
2720 PRINT 'JGG ESTA LISTO (SI O NO): ';
2730 INPUT A$
2740 IF A$='FIN' OR A$='F' THEN 2810
2750 IF A$='NO' OR A$='N' THEN 2720
2760 DIM P(D),P1(D)
2770 GIN @1:P(Y7),P1(Y7)
2780 Y7=Y7+1
2790 PRINT 'JJGG SI YA TERMINO EL INGRESO, ESCRIBA [FIN]'
2800 GO TO 2710
2810 Y7=Y7-1
2820 GOSUB 4950
2830 RETURN
2840 REM ***** SUBROUTINA DE OPERACIONES *****
2850 PRINT 'LJIGGG OPERACIONES REALIZABLES'
2860 PRINT 'I *****'
2870 PRINT 'JJJ 1 -- IMPRESION DE LISTA DE DATOS EN PANTALLA'
2880 PRINT 'J 2 -- IMPRESION DE LISTA DE DATOS EN PAPEL'
2890 PRINT 'J 3 -- GRAFIZACION DE LA ONDA EN PANTALLA'
2900 PRINT 'J 4 -- GRAFIZACION DE LA ONDA EN PAPEL'
2910 PRINT 'J 5 -- ALMACENAMIENTO DE DATOS EN ARCHIVO'
2920 PRINT 'J 6 -- REPETIR TODO EL PROCESO DE INGRESO'
2930 PRINT 'J 7 -- REPETIR EL PROCESO SIN ALTERAR PARAMETROS'
2940 PRINT 'J 8 -- FIN DEL INGRESO DE DATOS'
2950 PRINT 'JJJGG OPRIMA EL NUMERO DE LA OPERACION DESEADA: ';
2960 INPUT Q1
2970 Q3=0
2980 GOSUB Q1 OF 3050,3050,3360,3360,4350,1120,3000,3020
2990 GO TO 2840
3000 Q3=1
3010 GO TO 1160
3020 Q0=Q
3030 DELETE A1,P,P1
3040 GO TO 400
3050 REM ***** SUBROUTINA DE IMPRESION *****
3060 J1=0
3070 PRINT 'JJ INGRESE EL NOMBRE DEL SEGMENTO DE VOZ: ';
3080 INPUT N$
3090 GO TO Q1 OF 3120,3100
3100 A=51
3110 GO TO 3130

```

```

3120 A=32
3130 PRINT @A:'LJ          LISTA DE MUESTREOS DEL SEGMENTO DE VOZ:  ';N$
3140 PRINT @A:'          *****', 'JJJ'
3150 IF D4=1 THEN 3190
3160 L$='@MUESTREO/A1'
3170 N=1
3180 GOSUB 5480
3190 FOR I=1 TO D3 STEP 2
3200 J=J1+I
3210 IF I<D3 THEN 3250
3220 PRINT @A; USING 3230:'A1(',J,') = ',A1(I)
3230 IMAGE 10X,3A,4D,6A,3D,3D
3240 GO TO 3290
3250 PRINT @A; USING 3260:'A1(',J,') = ',A1(I)
3260 IMAGE 10X,3A,4D,6A,3D,3D,S
3270 PRINT @A; USING 3280:'A1(',J+1,') = ',A1(I+1)
3280 IMAGE 10X,3A,4D,6A,3D,3D
3290 NEXT I
3300 IF D5=1 THEN 3330
3310 J1=J1+D3
3320 GO TO 3180
3330 PRINT 'JJJGG  OPRIMA [RETURN] PARA CONTINUAR'
3340 INPUT A$
3350 RETURN
3360 REM      ***** SUBROUTINA DE GRAFIZACION *****
3370 Y1=15
3380 Y2=95
3390 PAGE
3400 IF G1=3 THEN 3440
3410 A=1
3420 PRINT 'JJGG  UBIQUE LA HOJA EN EL PLOTTER Y OPRIMA [RETURN]'
3430 GO TO 3460
3440 A=32
3450 PRINT 'JGG  INGRESE EL TITULO DEL GRAFICO: '; 'J';
3460 INPUT N$
3470 PRINT 'JJGG  CUANDO ACABE DE ANALIZAR EL GRAFICO OPRIMA [RETURN]'
3480 PRINT 'JJ  INGRESE LOS LIMITES DE TIEMPO PARA EL GRAFICO'
3490 PRINT 'J  LIMITE INFERIOR MAYOR O IGUAL A CERO:  ';
3500 INPUT X1
3510 PRINT 'J  LIMITE SUPERIOR MENOR O IGUAL A  ';L1;';  ';
3520 INPUT X2
3530 J1=0
3540 GOSUB 4870
3550 PAGE
3560 IF A=32 THEN 3590
3570 VIEWPORT 0,150,Y1,Y2
3580 GO TO 3600
3590 VIEWPORT 0,130,Y1,Y2
3600 WINDOW X1,X2,-10,10
3610 MOVE @A:X1,0
3620 IF D4=1 THEN 3670
3630 L$='@MUESTREO/A1'
3640 N=1

```

```

3650 GOSUB 5480
3660 I=0
3670 I=I+1
3680 X6=I/F+J1
3690 IF X6<X1 THEN 3720
3700 IF X6>X2 THEN 3750
3710 DRAW @A:X6,A1(I)
3720 IF I<D3 THEN 3670
3730 J1=X6
3740 IF D5<>1 THEN 3650
3750 GOSUB 4870
3760 AXIS @A:
3770 L=LEN(N$)
3780 MOVE @A:L1/2,-10
3790 SCALE 1,1
3800 XO=1.792
3810 YO=2.816
3820 IF A=1 THEN 3850
3830 RMOVE @A:-L/2*XO,-4*YO
3840 GO TO 3870
3850 RMOVE @A:-L/2*1.5*XO,-4*YO
3860 PRINT @1,17:1.5*XO,1.5*YO
3870 PRINT @A:N$
3880 IF A=32 THEN 3900
3890 PRINT @1,7:
3900 GOSUB 4870
3910 MOVE @A:0,-10
3920 Y1=0
3930 GOSUB 4870
3940 GIN @A:X3,Y3
3950 MOVE @A:X3,Y3
3960 SCALE 1,1
3970 RMOVE @A:0,-YO
3980 RDRAW @A:0,-1.5*YO
3990 RMOVE @A:0,0.75*YO
4000 IF A=32 THEN 4030
4010 RDRAW @1:150,0
4020 GO TO 4040
4030 RDRAW @32:130,0
4040 RMOVE @A:0,-0.75*YO
4050 RDRAW @A:0,1.5*YO
4060 Y1=15
4070 GOSUB 4870
4080 MOVE @A:L1/2,-10
4090 V$='TIEMPO ='
4100 T$=STR(X2-X1)
4110 T$=V$&T$
4120 T$=T$&' MSEG'
4130 DELETE T
4140 T=LEN(T$)
4150 SCALE 1,1
4160 RMOVE @A:-T/2*XO,-1.5*YO
4170 PRINT @A:T$

```

```

4180 GOSUB 4870
4190 MOVE @A:0,10
4200 SCALE 1,1
4210 RMOVE @A:0:0,75*Y0
4220 PRINT @A:" AMPLITUD NORMALIZADA"
4230 Y1=0
4240 Y2=100
4250 GOSUB 4870
4260 MOVE @A:0,10
4270 DRAW @A:0,-10
4280 DRAW @A:L1,-10
4290 DRAW @A:L1,10
4300 DRAW @A:0,10
4310 IF A=1 THEN 4340
4320 INPUT A$
4330 PAGE
4340 RETURN
4350 REM ***** SUBROUTINA DE ALMACENAMIENTO EN ARCHIVO *****
4360 Q0=Q
4370 Q=2
4380 GOSUB 4670
4390 IF D4=1 THEN 4450
4400 L$="@MUESTREO/A1"
4410 N=1
4420 CALL "UNIT",U1
4430 GOSUB 5480
4440 N=N-1
4450 CALL "UNIT",U2
4460 GOSUB 5250
4470 IF D5<>1 THEN 4420
4480 Q=Q0
4490 PRINT "JGG EL SEGMENTO DE VOZ QUEDA ALMACENADO EN: ";M$
4500 PRINT "JJ OPRIMA [RETORN] PARA CONTINUAR"
4510 INPUT A$
4520 CALL "UNIT",U1
4530 RETURN
4540 REM ***** SUBROUTINA DIMENSION *****
4550 IF Q3=1 THEN 4660
4560 PRINT "JJGG INGRESE LA DIMENSION DE LA MATRIZ DE VOZ A1(D1)"
4570 PRINT "JGG EL NUMERO DE MUESTREOS ES: D1 = ";
4580 INPUT D1
4590 PRINT "JGG LA FRECUENCIA DE MUESTREO EN (KHz) ES: F = ";
4600 INPUT F
4610 L1=D1/F
4620 PRINT "JGG LA LONGITUD DEL SEGMENTO DE VOZ EN (mseg) ES: L1 = ";L1
4630 C1=8*D1+18
4640 IF Q=3 THEN 4660
4650 GOSUB 5140
4660 RETURN
4670 REM ***** SUBROUTINA NOMBRE DEL SEGMENTO DE VOZ *****
4680 PRINT "LJJGG INGRESE EL NOMBRE DEL SEGMENTO DE VOZ: "
4690 PRINT "JGG (MAXIMO 10 CARACTERES): ";
4700 INPUT M$

```

```

4710 PRINT 'JGG  INGRESE EL NUMERO DE LA UNIDAD EN DONDE GUARDAR:  ';
4720 INPUT U2
4730 CALL 'MOUNT',U2,X$
4740 CALL 'UNIT',U2
4750 M$='@VQZ/'&M$
4760 CALL 'FILE',U2,M$,X$
4770 IF X$='' THEN 4860
4780 PRINT 'J  EL ARCHIVO: ';M$;' YA EXISTE EN EL DISCO'
4790 PRINT 'J  IESEA DESTRUIRLO? (SI O NO):  ';
4800 INPUT X$
4810 IF X$='NO' OR X$='N' THEN 4840
4820 KILL M$
4830 GO TO 4860
4840 PRINT 'LJGG  INGRESE EL NUEVO NOMBRE DEL SEGMENTO:  ';
4850 GO TO 4700
4860 RETURN
4870 REM      **** SUBROUTINA WINDOW: VIEWPORT ****
4880 IF Q3=1 THEN 4940
4890 IF A=32 THEN 4920
4900 VIEWPORT 0,150,Y1,Y2
4910 GO TO 4930
4920 VIEWPORT 0,130,Y1,Y2
4930 WINDOW 0,L1,-10,10
4940 RETURN
4950 REM      **** SUBROUTINA DE DIMENSIONAMIENTO DE P Y P1 ****
4960 CLOSE 1
4970 KILL '@ESQUINAS'
4980 CREATE '@ESQUINAS';200,0
4990 OPEN '@ESQUINAS':1,'F',X$
5000 DIM P(Y7),P1(Y7)
5010 WRITE #1:P,P1
5020 DELETE P,P1
5030 DIM P(Y7),P1(Y7)
5040 CALL 'REWIND',1
5050 FOR I=1 TO Y7
5060 READ #1:P(I)
5070 NEXT I
5080 FOR J=1 TO Y7
5090 READ #1:P1(J)
5100 NEXT J
5110 CLOSE 1
5120 GOSUB 5140
5130 RETURN
5140 REM      **** SUBROUTINA DE CHEQUEO DE MEMORIA ****
5150 D4=1
5160 IF MEMORY-C1>K4 THEN 5220
5170 D3=INT((MEMORY-K4-18)/8)
5180 D5=0
5190 IF D3=>D1 THEN 5220
5200 D4=0
5210 GO TO 5240
5220 D3=D1
5230 D5=1

```

```

5240 RETURN
5250 REM      **** SUBROUTINA DE ALMACENAMIENTO EN ARCHIVO ****
5260 IF N<>1 THEN 5320
5270 CLOSE 1
5280 KILL M$
5290 CREATE M$;200,0
5300 OPEN M$;1,'F',X$
5310 WRITE #1;D1,F
5320 DIM A1(D3)
5330 WRITE #1:A1
5340 IF N=0 THEN 5460
5350 IF D4=1 THEN 5460
5360 IF Q=2 THEN 5380
5370 N=N+1
5380 IF N*D3<D1 THEN 5430
5390 D3=D1-(N-1)*D3
5400 N=0
5410 DELETE A1
5420 DIM A1(D3)
5430 IF Q<>2 THEN 5470
5440 N=N+1
5450 GO TO 5470
5460 D5=1
5470 RETURN
5480 REM      **** SUBROUTINA DE LECTURA DE ARCHIVOS ****
5490 IF N<>1 THEN 5620
5500 IF D2=1 THEN 5610
5510 D5=0
5520 CLOSE 2
5530 OPEN L$;2,'R',X$
5540 READ #2;D1,F
5550 GO TO Q1 OF 5580,5560,5580,5560,5580
5560 PRINT 'JJ EL NUMERO DE MUESTREOS EN EL ARCHIVO ES:  D1 =  ';D1
5570 PRINT 'J  LA FRECUENCIA DE MUESTREO ES:           F =  ';F
5580 L1=D1/F
5590 C1=8*D1+18
5600 GOSUB 5140
5610 IF D4=1 THEN 5650
5620 IF D3*N<D1 THEN 5660
5630 D3=D1-(N-1)*D3
5640 N=0
5650 D5=1
5660 DELETE A1
5670 DIM A1(D3)
5680 READ #2:A1
5690 N=N+1
5700 RETURN

```



```

1000 REM   *** PROGRAMA PARA DETECCION DE TONO EN PARALELO (DIETPAR) ***
1010 Q2=2
1020 IF Q1<>Q2 THEN 500
1030 CALL 'TIME',Z$
1040 W$=SEG(Z$,1,9)
1050 X$=SEG(Z$,11,2)
1060 Z1=VAL(X$)*3600
1070 X$=SEG(Z$,14,2)
1080 Z1=Z1+VAL(X$)*60
1090 X$=SEG(Z$,17,2)
1100 Z1=Z1+VAL(X$)
1110 PRINT 'LJJ  INGRESE EL NOMBRE DEL SEGMENTO DE VOZ: ';
1120 INPUT V$
1130 PRINT 'JJ  INGRESE EL NUMERO DE PRUEBA A REALIZAR:  ';
1140 INPUT P0
1150 PRINT 'LJJI  TIPOS DE IMPRESION'
1160 PRINT 'I  *****'
1170 PRINT 'JJJ      CLASE 1 -- IMPRESION PARCIAL EN PANTALLA'
1180 PRINT 'J      CLASE 2 -- IMPRESION TOTAL EN PANTALLA'
1190 PRINT 'J      CLASE 3 -- IMPRESION PARCIAL EN PAPEL'
1200 PRINT 'J      CLASE 4 -- IMPRESION TOTAL EN PAPEL'
1210 PRINT 'JJGGG      OPRIMA EL NUMERO DE CLASE DESEADA:  ';
1220 INPUT Q
1230 IF Q>2 THEN 1260
1240 A=32
1250 GO TO 1270
1260 A=51
1270 PAGE
1280 PRINT @A:' ESCUELA POLITECNICA NACIONAL'
1290 PRINT @A:' FACULTAD DE INGENIERIA ELECTRICA'
1300 PRINT @A:' TRABAJO DE TESIS - JUNIO 1982'
1310 PRINT @A:' PROGRAMADO POR: CARLOS E. TOBON'
1320 PRINT @A:' DETECCION DEL PERIODO DE TONO DEL SEGMENTO DE VOZ: ';V$
1330 PRINT @A:' PRUEBA #';P0;
1340 PRINT @A:' FECHA DE REALIZACION: ';W$
1350 PRI @A:' =====';
1360 PRINT @A:'=====';
1370 PRINT @A:'J      DETECCION DE TONO POR PROCESAMIENTO EN PARALELO'
1380 PRINT @A:' *****'
1390 J1=0
1400 J2=0
1410 I6=0
1420 F=0.7
1430 K1=4
1440 K2=5
1450 K3=0.2
1460 K4=2000
1470 N=1
1480 C2=3
1490 L$='@MUESTRO/A1'
1500 GOSUB 7240
1510 REM   *** DETERMINACION DE LOS PICOS Y VALLES ***
1520 DELETE A2,A3,A4,A5,N1,N2

```

```

1530 DIM A2(D3/2),A3(D3/2),A4(D3/2),A5(D3/2),N1(2),N2(2)
1540 I=1
1550 K=0
1560 J=0
1570 IF A1(I)<A1(I+1) THEN 1740
1580 IF A1(I)>A1(I+1) THEN 1650
1590 IF I=D3-1 THEN 1620
1600 I=I+1
1610 GO TO 1570
1620 PRINT 'SEGMENTO SIN VOZ'
1630 END
1640 REM   *** DETERMINACION DE LOS MINIMOS ***
1650 N1(2)=A1(I)
1660 N2(2)=I+D6
1670 IF I=D3 THEN 1830
1680 I=I+1
1690 IF N1(2)=>A1(I) THEN 1650
1700 J=J+1
1710 A3(J)=N1(2)
1720 A5(J)=N2(2)
1730 REM   *** DETERMINACION DE LOS MAXIMOS ***
1740 N1(1)=A1(I)
1750 N2(1)=I+D6
1760 IF I=D3 THEN 1830
1770 I=I+1
1780 IF N1(1)<=A1(I) THEN 1740
1790 K=K+1
1800 A2(K)=N1(1)
1810 A4(K)=N2(1)
1820 GO TO 1650
1830 REM   *** ACTUALIZACION DE LAS CONSTANTES ***
1840 N=N-1
1850 J1=J1+K
1860 J2=J2+J
1870 D6=D6+D3
1880 GOSUB 7400
1890 IF D5<>1 THEN 1500
1900 CLOSE
1910 IF D4<>1 THEN 1960
1920 D3=J1
1930 DELETE A1
1940 IF Q/2-INT(Q/2)=0.5 THEN 2640
1950 GO TO 2150
1960 DELETE A1,A2,A3,A4,A5
1970 IF Q/2-INT(Q/2)=0.5 THEN 2640
1980 OPEN '@POSICION/PICOS';7,'R',X#
1990 C2=2
2000 D1=J1
2010 N=1
2020 GOSUB 7120
2030 IF D4=1 THEN 2060
2040 IF D3*N<D1 THEN 2080
2050 D3=D1-(N-1)*D3

```

```

2060 D5=1
2070 D4=0
2080 DELETE A2,A4
2090 DIM A2(D3),A4(D3)
2100 FOR J=1 TO D3
2110 READ #7:A2(J),A4(J)
2120 NEXT J
2130 N=N+1
2140 IF N<>2 THEN 2170
2150 PRINT @A:'J                                MATRIZ DE LOS MAXIMOS'
2160 PRINT @A:'                                *****'
2170 FOR I=1 TO D3 STEP 2
2180 IF I<D3 THEN 2220
2190 PRINT @A: USING 2200:'A2(',A4(I),') = ',A2(I)
2200 IMAGE 6X,3A,4D,6A,3D,3D
2210 GO TO 2260
2220 PRINT @A: USING 2230:'A2(',A4(I),') = ',A2(I)
2230 IMAGE 6X,3A,4D,6A,3D,3D,S
2240 PRINT @A: USING 2250:'A2(',A4(I+1),') = ',A2(I+1)
2250 IMAGE 7X,3A,4D,6A,3D,3D
2260 NEXT I
2270 IF D5<>1 THEN 2040
2280 IF D4<>1 THEN 2310
2290 D3=J2
2300 GO TO 2480
2310 DELETE A2,A4
2320 OPEN '@POSICION/VALLES';8,'R',X$
2330 D1=J2
2340 N=1
2350 GOSUB 7120
2360 IF D4=1 THEN 2390
2370 IF D3*N<D1 THEN 2410
2380 D3=D1-(N-1)*D3
2390 D5=1
2400 D4=0
2410 DELETE A3,A5
2420 DIM A3(D3),A5(D3)
2430 FOR J=1 TO D3
2440 READ #8:A3(J),A5(J)
2450 NEXT J
2460 N=N+1
2470 IF N<>2 THEN 2500
2480 PRINT @A:'J                                MATRIZ DE LOS MINIMOS'
2490 PRINT @A:'                                *****'
2500 FOR I=1 TO D3 STEP 2
2510 IF I<D3 THEN 2550
2520 PRINT @A: USING 2530:'A3(',A5(I),') = ',A3(I)
2530 IMAGE 6X,3A,4D,6A,3D,3D
2540 GO TO 2590
2550 PRINT @A: USING 2560:'A3(',A5(I),') = ',A3(I)
2560 IMAGE 6X,3A,4D,6A,3D,3D,S
2570 PRINT @A: USING 2580:'A3(',A5(I+1),') = ',A3(I+1)
2580 IMAGE 7X,3A,4D,6A,3D,3D

```

```

2590 NEXT I
2600 IF D5<>1 THEN 2370
2610 CLOSE
2620 IF D4=1 THEN 2640
2630 DELETE A3,A5
2640 PRINT 'JJJ FIN DE LA DETECCION DE LOS MAXIMOS Y MINIMOS DE A(D1)'
2650 REM . *** DETERMINACION DE LA CAPACIDAD DISPONIBLE DE MEMORIA ***
2660 L=0
2670 N1(1)=0
2680 I2=0
2690 N1(2)=0
2700 I6=0
2710 IF J1<J2 THEN 2740
2720 J3=J2
2730 GO TO 2750
2740 J3=J1
2750 IF D4<>1 THEN 2830
2760 J4=6
2770 C2=7
2780 I1=J3
2790 N=1
2800 GOSUB 7120
2810 IF D4=1 THEN 3100
2820 GOSUB 7120
2830 GOSUB 7590
2840 OPEN '@POSICION/PICOS';7,'U',X#
2850 OPEN '@POSICION/VALLES';8,'U',X#
2860 CALL 'REWIND',7
2870 CALL 'REWIND',8
2880 I2=1
2890 C2=8
2900 I1=J3
2910 N=1
2920 GOSUB 7120
2930 IF D4<>1 THEN 2960
2940 J4=6
2950 GO TO 2980
2960 IF N*D3<I1 THEN 2990
2970 I3=I1-(N-1)*I3
2980 I5=1
2990 DELETE A2,A3,A4,A5
3000 DIM A2(I3),A3(I3),A4(1),A5(1)
3010 FOR J=1 TO I3
3020 IF J=1 THEN 3060
3030 READ #7:A2(J),K7
3040 READ #8:A3(J),K7
3050 GO TO 3080
3060 READ #7:A2(1),A4(1)
3070 READ #8:A3(1),A5(1)
3080 NEXT J
3090 N=N+1
3100 IF A4(1)>A5(1) THEN 3130
3110 I7=0

```

```

3120 GO TO 3140
3130 D7=1
3140 REM   *** GENERACION DE LOS SEIS TRENES DE PULSOS ***
3150 DELETE M1,M2,M3,M4,M5,M6
3160 DIM M1(D3),M2(D3),M3(D3),M4(D3),M5(D3),M6(D3)
3170 FOR I=1 TO D3
3180 IF D7=1 THEN 3260
3190 M3(I)=A2(I)-N1(1)
3200 M2(I)=A2(I)-N1(2)
3210 N1(1)=A2(I)
3220 M1(I)=ABS(A2(I))
3230 IF M3(I)=>0 THEN 3250
3240 M3(I)=0
3250 IF D7=1 THEN 3330
3260 M6(I)=-A3(I)+N1(2)
3270 M5(I)=-A3(I)+N1(1)
3280 N1(2)=A3(I)
3290 M4(I)=ABS(A3(I))
3300 IF M6(I)=>0 THEN 3320
3310 M6(I)=0
3320 IF D7=1 THEN 3190
3330 NEXT I
3340 IF D4=1 THEN 4420
3350 WRITE #1:M1
3360 WRITE #2:M2
3370 WRITE #3:M3
3380 WRITE #4:M4
3390 WRITE #5:M5
3400 WRITE #6:M6
3410 IF D5<>1 THEN 2960
3420 DELETE M1,M2,M3,M4,M5,M6
3430 C1=(8*J3+18)*7
3440 IF MEMORY-C1=>K4 THEN 3510
3450 J4=INT((MEMORY-K4-8*J3+18)/(8*J3+18))
3460 IF J4=>1 THEN 3500
3470 PRINT 'JIGG EL NUMERO DE DATOS ES EXCESIVO'
3480 PRINT 'IGG REDIMENSIONE A1(D1) PARA REINICIAR EL PROCESO'
3490 END
3500 IF J4<=6 THEN 3520
3510 J4=6
3520 GO TO J4 OF 3530,3580,3670,3800,3970,4180
3530 DELETE M1
3540 DIM M1(J3)
3550 CALL 'REWIND',1
3560 READ #1:M1
3570 GO TO 4420
3580 DELETE M1
3590 DIM M1(J3)
3600 CALL 'REWIND',1
3610 READ #1:M1
3620 DELETE M2
3630 DIM M2(J3)
3640 CALL 'REWIND',2

```

```

3650 READ #2:M2
3660 GO TO 4420
3670 DELETE M1
3680 DIM M1(J3)
3690 CALL 'REWIND',1
3700 READ #1:M1
3710 DELETE M2
3720 DIM M2(J3)
3730 CALL 'REWIND',2
3740 READ #2:M2
3750 DELETE M3
3760 DIM M3(J3)
3770 CALL 'REWIND',3
3780 READ #3:M3
3790 GO TO 4420
3800 DELETE M1
3810 DIM M1(J3)
3820 CALL 'REWIND',1
3830 READ #1:M1
3840 DELETE M2
3850 DIM M2(J3)
3860 CALL 'REWIND',2
3870 READ #2:M2
3880 DELETE M3
3890 DIM M3(J3)
3900 CALL 'REWIND',3
3910 READ #3:M3
3920 DELETE M4
3930 DIM M4(J3)
3940 CALL 'REWIND',4
3950 READ #4:M4
3960 GO TO 4420
3970 DELETE M1
3980 DIM M1(J3)
3990 CALL 'REWIND',1
4000 READ #1:M1
4010 DELETE M2
4020 DIM M2(J3)
4030 CALL 'REWIND',2
4040 READ #2:M2
4050 DELETE M3
4060 DIM M3(J3)
4070 CALL 'REWIND',3
4080 READ #3:M3
4090 DELETE M4
4100 DIM M4(J3)
4110 CALL 'REWIND',4
4120 READ #4:M4
4130 DELETE M5
4140 DIM M5(J3)
4150 CALL 'REWIND',5
4160 READ #5:M5
4170 GO TO 4420

```

```

4180 DELETE M1
4190 DIM M1(J3)
4200 CALL 'REWIND',1
4210 READ #1:M1
4220 DELETE M2
4230 DIM M2(J3)
4240 CALL 'REWIND',2
4250 READ #2:M2
4260 DELETE M3
4270 DIM M3(J3)
4280 CALL 'REWIND',3
4290 READ #3:M3
4300 DELETE M4
4310 DIM M4(J3)
4320 CALL 'REWIND',4
4330 READ #4:M4
4340 DELETE M5
4350 DIM M5(J3)
4360 CALL 'REWIND',5
4370 READ #5:M5
4380 DELETE M6
4390 DIM M6(J3)
4400 CALL 'REWIND',6
4410 READ #6:M6
4420 L=L+1
4430 FOR L=L TO 6
4440 DELETE M
4450 DIM M(J3)
4460 GO TO L OF 4470,4500,4530,4560,4590,4620
4470 M=M1
4480 DELETE M1
4490 GO TO 4640
4500 M=M2
4510 DELETE M2
4520 GO TO 4640
4530 M=M3
4540 DELETE M3
4550 GO TO 4640
4560 M=M4
4570 DELETE M4
4580 GO TO 4640
4590 M=M5
4600 DELETE M5
4610 GO TO 4640
4620 M=M6
4630 DELETE M6
4640 GOSUB 5720
4650 GO TO J4 OF 4660,4670,4680,4690,4700,4710
4660 GO TO L OF 3620,3750,3920,4130,4380,4710
4670 GO TO L OF 4710,3880,4710,4340,4710,4710
4680 GO TO L OF 4710,4710,4300,4710,4710,4710
4690 GO TO L OF 4710,4710,4710,4340,4710,4710
4700 GO TO L OF 4710,4710,4710,4710,4380,4710

```

```

4710 NEXT L
4720 REM   *** DETERMINACION DEL MINIMO COMUN DE ESTIMACIONES ***
4730 DELETE M,M1,M2,M3,M4,M5,M6,A4,A5
4740 D6=0
4750 N5=1
4760 C9=0
4770 N4=D(1)
4780 FOR L=2 TO 6
4790 IF N4=>D(L) THEN 4810
4800 N4=D(L)
4810 NEXT L
4820 DIM B(7,6)
4830 PRINT @A:'JJ           ESTIMACIONES FINALES DEL PERIODO DE TONO'
4840 PRINT @A:'           *****'
4850 C2=8
4860 D1=N4
4870 GOSUB 7120
4880 N=1
4890 CALL 'REWIND',9
4900 IF D4=1 THEN 4930
4910 IF D3*N<D1 THEN 4940
4920 D3=D1-(N-1)*D3
4930 D5=1
4940 DELETE E
4950 DIM E(8,D3)
4960 FOR I=1 TO 6
4970 FOR J=1 TO D3
4980 IF J<(N-1)*D3 THEN 5020
4990 IF J>N*D3 THEN 5020
5000 READ #9:E(I,J)
5010 GO TO 5030
5020 READ #9:K7
5030 NEXT J
5040 NEXT I
5050 N=N+1
5060 REM   *** FORMACION DE LA MATRIZ PARA EVALUAR COINCIDENCIA ***
5070 FOR N3=1 TO D3
5080 FOR L=1 TO 6
5090 B(1,L)=E(L,N3)
5100 GO TO N3 OF 5140,5170
5110 B(2,L)=E(L,N3-1)
5120 B(3,L)=E(L,N3-2)
5130 GO TO 5190
5140 B(2,L)=E(L,D3)
5150 B(3,L)=E(L,D3-1)
5160 GO TO 5190
5170 B(2,L)=E(L,1)
5180 B(3,L)=E(L,D3)
5190 B(4,L)=B(1,L)+B(2,L)
5200 B(5,L)=B(2,L)+B(3,L)
5210 B(6,L)=B(3,L)+B(4,L)
5220 NEXT L
5230 REM   *** DETERMINACION DEL NUMERO DE COINCIDENCIAS ***

```



```

5240 E(8,N3)=0
5250 E(7,N3)=0
5260 FOR K=1 TO 6
5270 B(7,K)=-1
5280 FOR I=1 TO 6
5290 FOR J=1 TO 6
5300 IF B(I,J)=0 THEN 5330
5310 IF ABS(B(1,K)-B(I,J))>K3 THEN 5330
5320 B(7,K)=B(7,K)+1
5330 NEXT J
5340 NEXT I
5350 IF E(8,N3)=>B(7,K) THEN 5410
5360 E(8,N3)=B(7,K)
5370 E(7,N3)=B(1,K)
5380 IF C9=>B(7,K) THEN 5410
5390 C9=B(7,K)
5400 E9=B(1,K)
5410 NEXT K
5420 NEXT N3
5430 IF N5<>1 THEN 5490
5440 N5=N5+1
5450 CLOSE 8
5460 KILL '@RESULTADOS'
5470 CREATE '@RESULTADOS';200,0
5480 OPEN '@RESULTADOS';8,'U',X#
5490 FOR I=1 TO D3
5500 PRI @A: USI 5510:'E(*,I+D6,*) = *,E(7,I),' MSEG      N = *,E(8,I)
5510 IMAGE12X,2A,2D,6A,3D,3D,15A,2D
5520 WRITE #8:E(8,I),E(7,I)
5530 NEXT I
5540 IF D4=1 THEN 5570
5550 IF D5<>1 THEN 4910
5560 CLOSE
5570 PRI @A: USI 5580:'N - NUMERO DE COINCIDENCIAS DEL PERIODO DE TONO'
5580 IMAGE /,6X,47A
5590 DELETE A1,A2,A3,A4,A5,B,D,E,M,N,N1,N2
5600 CALL 'TIME',Z#
5610 X#=SEG(Z#,11,2)
5620 Z2=VAL(X#)*3600
5630 X#=SEG(Z#,14,2)
5640 Z2=Z2+VAL(X#)*60
5650 X#=SEG(Z#,17,2)
5660 Z2=Z2+VAL(X#)
5670 Z2=Z2-Z1
5680 PRINT @A: USING 5690:'TIEMPO DE EJECUCION = *,Z2,* SEGUNDOS'
5690 IMAGE /,6X,22A,4D,9A
5700 PRINT 'JJ FIN DE LA DETECCION DE TONO, OPRIMA TECLA DESEADA'
5710 END
5720 REM      ***** SUBROUTINA DE CALCULOS *****
5730 DELETE E
5740 DIM E(J3),D(6)
5750 N=1
5760 I7=0

```

```

5770 REM   *** LECTURA DE LAS MATRICES A4(J3),A5(J3) ***
5780 IF D2<>1 THEN 6030
5790 DELETE A2,A3
5800 IF L>3 THEN 5830
5810 CALL 'REWIND',7
5820 GO TO 5840
5830 CALL 'REWIND',8
5840 C2=1
5850 D1=J3
5860 GOSUB 7120
5870 IF D4=1 THEN 5900
5880 IF N*D3<D1 THEN 5910
5890 D3=D1-(N-1)*D3
5900 D5=1
5910 IF L>3 THEN 5950
5920 DELETE A4
5930 DIM A4(D3)
5940 GO TO 5970
5950 DELETE A5
5960 DIM A5(D3)
5970 FOR J=1 TO D3
5980 IF L>3 THEN 6010
5990 READ #7;K7,A4(J)
6000 GO TO 6020
6010 READ #8;K7,A5(J)
6020 NEXT J
6030 IF Q/2-INT(Q/2)=0,5 THEN 6300
6040 IF N<>1 THEN 6070
6050 PRINT @A:"J                               TREN DE IMPULSOS GENERADOS M";L
6060 PRINT @A:"                               *****"
6070 IF L>3 THEN 6190
6080 FOR I=1 TO D3 STEP 2
6090 IF I<D3 THEN 6130
6100 PRINT @A: USING 6110:"M",L,"(",A4(I),") = ",M(I+D7)
6110 IMAGE 6X,A,D,A,4D,6A,3D,3D
6120 GO TO 6170
6130 PRINT @A: USING 6140:"M",L,"(",A4(I),") = ",M(I+D7)
6140 IMAGE 6X,A,D,A,4D,6A,3D,3D,S
6150 PRINT @A: USING 6160:"M",L,"(",A4(I+1),") = ",M(I+1+D7)
6160 IMAGE 7X,A,D,A,4D,6A,3D,3D
6170 NEXT I
6180 GO TO 6290
6190 FOR I=1 TO D3 STEP 2
6200 IF I<D3 THEN 6240
6210 PRINT @A: USING 6220:"M",L,"(",A5(I),") = ",M(I+D7)
6220 IMAGE 6X,A,D,A,4D,6A,3D,3D
6230 GO TO 6280
6240 PRINT @A: USING 6250:"M",L,"(",A5(I),") = ",M(I+D7)
6250 IMAGE 6X,A,D,A,4D,6A,3D,3D,S
6260 PRINT @A: USING 6270:"M",L,"(",A5(I+1),") = ",M(I+1+D7)
6270 IMAGE 7X,A,D,A,4D,6A,3D,3D
6280 NEXT I
6290 D7=D3+D7

```

```

6300 REM   *** CALCULO DE LA AMPLITUD MAXIMA ***
6310 IF D4<>1 THEN 6330
6320 IF L<>3 THEN 6350
6330 D2=1
6340 GO TO 6360
6350 D2=0
6360 B1=0
6370 FOR I=1 TO D3
6380 B1=B1 MAX M(I)
6390 NEXT I
6400 B1=F*B1
6410 IF N<>1 THEN 6480
6420 REM   *** CALCULO DE LAS ESTIMACIONES DEL PERIODO DE TONO ***
6430 N3=1
6440 K=0
6450 C1=6
6460 C2=0.4*C1
6470 C3=C1/0.695
6480 FOR I=1 TO D3
6490 IF K=0 THEN 6680
6500 IF L<=3 THEN 6530
6510 C5=A5(I)
6520 GO TO 6540
6530 C5=A4(I)
6540 IF K=1 THEN 6770
6550 C8=A7*EXP(-C3*1.0E-3*(C5-C6-C4)/F)
6560 IF M(I)<=C8 THEN 6800
6570 E(N3)=(C5-C6)/F
6580 C1=(C1+E(N3))/2
6590 IF C1=>4 THEN 6610
6600 C1=4
6610 IF C1<=10 THEN 6630
6620 C1=10
6630 C2=0.4*C1
6640 C3=C1/0.695
6650 N3=N3+1
6660 K=1
6670 GO TO 6700
6680 IF M(I)<B1 THEN 6800
6690 K=K+1
6700 IF L<=3 THEN 6730
6710 C6=A5(I)
6720 GO TO 6740
6730 C6=A4(I)
6740 A7=M(I)
6750 C4=C2*F
6760 GO TO 6800
6770 IF C5-C6<C4 THEN 6800
6780 K=K+1
6790 GO TO 6550
6800 NEXT I
6810 IF L<>3 THEN 6830
6820 DELETE A4

```

```

6830 IF D4=1 THEN 6860
6840 N=N+1
6850 IF D5<>1 THEN 5880
6860 IF N3>1 THEN 6880
6870 E(1)=0
6880 D(L)=N3-1
6890 PRINT @A:"J          ESTIMACIONES DE TONO DEL TREN DE IMPULSOS M";L
6900 PRINT @A:"          *****"
6910 FOR J=1 TO N3-1 STEP 2
6920 PRINT @A: USING 6930:"E",L,"(",J,") = ",E(J)
6930 IMAGE 7X,A,D,A,2D,6A,3D,3D,S
6940 IF J<N3-1 THEN 6970
6950 PRINT @A:""
6960 GO TO 6990
6970 PRINT @A: USING 6980:"E",L,"(",J+1,") = ",E(J+1)
6980 IMAGE 8X,A,D,A,2D,6A,3D,3D
6990 NEXT J
7000 FOR K=N3 TO J3
7010 E(K)=0
7020 NEXT K
7030 REM   *** ALMACENAMIENTO DE ESTIMACIONES DE TONO ***
7040 IF D6=1 THEN 7090
7050 D6=1
7060 KILL "@ESTIMACION/TONO"
7070 CREATE "@ESTIMACION/TONO";200,0
7080 OPEN "@ESTIMACION/TONO";9,"U",X$
7090 DIM E(J3)
7100 WRITE #9:E
7110 RETURN
7120 REM   **** SUBROUTINA DE CHEQUEO DE MEMORIA ****
7130 D4=1
7140 C1=(8*D1+18)*C2
7150 IF MEMORY-C1>K4 THEN 7210
7160 D3=INT((MEMORY-K4-18)/(8*C2))
7170 D5=0
7180 IF D3=>D1 THEN 7210
7190 D4=0
7200 GO TO 7230
7210 D3=D1
7220 D5=1
7230 RETURN
7240 REM   **** SUBROUTINA DE LECTURA DE ARCHIVOS ****
7250 IF N>1 THEN 7320
7260 D5=0
7270 CLOSE
7280 OPEN L$;9,"R",X$
7290 READ #9:D1,F
7300 GOSUB 7120
7310 IF D4=1 THEN 7340
7320 IF D3*N<D1 THEN 7350
7330 D3=D1-(N-1)*D3
7340 D5=1
7350 DELETE A1

```

```

7360 DIM A1(D3)
7370 READ #9:A1
7380 N=N+1
7390 RETURN
7400 REM   *** ALMACENAMIENTO DE PICOS Y VALLES EN ARCHIVO ***
7410 IF N>1 THEN 7500
7420 CLOSE 7
7430 CLOSE 8
7440 KILL '@POSICION/PICOS'
7450 KILL '@POSICION/VALLES'
7460 CREATE '@POSICION/PICOS';200,0
7470 CREATE '@POSICION/VALLES';200,0
7480 OPEN '@POSICION/PICOS';7,'F',X$
7490 OPEN '@POSICION/VALLES';8,'F',X$
7500 DIM A2(K),A4(K),A3(J),A5(J)
7510 FOR I=1 TO K
7520 WRITE #7:A2(I),A4(I)
7530 NEXT I
7540 FOR I=1 TO J
7550 WRITE #8:A3(I),A5(I)
7560 NEXT I
7570 N=N+1
7580 RETURN
7590 REM   *** CREACION DE ARCHIVOS PARA LOS TRENES DE PULSOS ***
7600 KILL '@TREN/UNO'
7610 KILL '@TREN/DOS'
7620 KILL '@TREN/TRES'
7630 KILL '@TREN/CUATRO'
7640 KILL '@TREN/CINCO'
7650 KILL '@TREN/SEIS'
7660 CREATE '@TREN/UNO';200,0
7670 CREATE '@TREN/DOS';200,0
7680 CREATE '@TREN/TRES';200,0
7690 CREATE '@TREN/CUATRO';200,0
7700 CREATE '@TREN/CINCO';200,0
7710 CREATE '@TREN/SEIS';200,0
7720 OPEN '@TREN/UNO';1,'U',X$
7730 OPEN '@TREN/DOS';2,'U',X$
7740 OPEN '@TREN/TRES';3,'U',X$
7750 OPEN '@TREN/CUATRO';4,'U',X$
7760 OPEN '@TREN/CINCO';5,'U',X$
7770 OPEN '@TREN/SEIS';6,'U',X$
7780 RETURN

```

```

1000 REM   *** PROGRAMA DE DETECCION DE TONO POR REDUCCION (DETRED) ***
1010 Q2=3
1020 IF Q1<>Q2 THEN 500
1030 CALL "TIME",Z$
1040 W$=SEG(Z$,1,9)
1050 X$=SEG(Z$,11,2)
1060 Z1=VAL(X$)*3600
1070 X$=SEG(Z$,14,2)
1080 Z1=Z1+VAL(X$)*60
1090 X$=SEG(Z$,17,2)
1100 Z1=Z1+VAL(X$)
1110 PRINT "LJJ  INGRESE EL NOMBRE DEL SEGMENTO DE VOZ: ";
1120 INPUT V$
1130 PRINT "JJ  INGRESE EL NUMERO DE PRUEBA A REALIZAR: ";
1140 INPUT P0
1150 PRINT "LJJ
                                TIPOS DE IMPRESION"
1160 PRINT "
                                *****"
1170 PRINT "JJJ  CLASE 1 -- IMPRESION PARCIAL EN PANTALLA"
1180 PRINT "J    CLASE 2 -- IMPRESION TOTAL EN PANTALLA"
1190 PRINT "J    CLASE 3 -- IMPRESION PARCIAL EN PAPEL"
1200 PRINT "J    CLASE 4 -- IMPRESION TOTAL EN PAPEL"
1210 PRINT "JJGGG  OPRIMA EL NUMERO DE CLASE DESEADA: ";
1220 INPUT Q
1230 IF Q>2 THEN 1260
1240 A=32
1250 GO TO 1270
1260 A=51
1270 PAGE
1280 PRINT @A:" ESCUELA POLITECNICA NACIONAL"
1290 PRINT @A:" FACULTAD DE INGENIERIA ELECTRICA"
1300 PRINT @A:" TRABAJO DE TESIS - JUNIO 1982"
1310 PRINT @A:" PROGRAMADO POR: CARLOS E. TOBON"
1320 PRINT @A:" DETECCION DEL PERIODO DE TONO DEL SEGMENTO DE VOZ: ";V$
1330 PRINT @A:" PRUEBA #";P0;
1340 PRINT @A:"
                                FECHA DE REALIZACION: ";W$
1350 PRI @A:" =====
1360 PRINT @A:"=====
1370 PRINT @A:"JJ  DETECCION DE TONO POR REDUCCION DE INFORMACION"
1380 PRINT @A:"
                                *****"
1390 L2=0
1400 L3=2
1410 L4=20
1420 L5=80
1430 L6=60
1440 K4=2000
1450 N=1
1460 D7=0
1470 C2=3
1480 L$="@MUESTREC/A1"
1490 D6=0
1500 GOSUB 7560
1510 REM   *** DETERMINACION DE LOS CICLOS DE EXCURSION ***
1520 DELETE A2,A3,A4,S

```

```

1530 DIM A2(D3/2),A3(D3/2),A4(D3/2),S(D3/2)
1540 K=0
1550 N1=0
1560 FOR I=1 TO D3-1
1570 IF A1(I)*A1(I+1)>0 THEN 1640
1580 K=K+1
1590 A2(K)=I+1+D7
1600 S(K)=0
1610 A3(K)=A1(I+1)
1620 A4(K)=I+1+D7
1630 GO TO 1690
1640 IF K=0 THEN 1740
1650 IF ABS(A3(K))=>ABS(A1(I+1)) THEN 1690
1660 A3(K)=A1(I+1)
1670 A4(K)=I+1+D7
1680 REM *** DETERMINACION DE LA POLARIDAD ***
1690 IF ABS(N1)>ABS(A1(I+1)) THEN 1720
1700 N1=A1(I+1)
1710 REM *** CALCULO DE LAS SUMAS DE CADA CICLO DE EXCURSION ***
1720 S(K)=S(K)+ABS(A1(I+1))
1730 L2=L2+A1(I+1)
1740 NEXT I
1750 IF K=0 THEN 1990
1760 L2=L2/K
1770 D6=D6+K
1780 D7=D7+D3
1790 DIM A2(K),A3(K),A4(K),S(K)
1800 D2=0
1810 IF Q/2-INT(Q/2)=0.5 THEN 1950
1820 IF N<>1 THEN 1850
1830 PRINT @A;"J
1840 PRINT @A;"
1850 FOR I=1 TO K STEP 2
1860 IF I<K THEN 1900
1870 PRINT @A: USING 1880:"A3(",A4(I),") = ",A3(I)
1880 IMAGE 6X,3A,4D,6A,3D,3D
1890 GO TO 1940
1900 PRINT @A: USING 1910:"A3(",A4(I),") = ",A3(I)
1910 IMAGE 6X,3A,4D,6A,3D,3D,S
1920 PRINT @A: USING 1930:"A3(",A4(I+1),") = ",A3(I+1)
1930 IMAGE 6X,3A,4D,6A,3D,3D
1940 NEXT I
1950 IF D4<>1 THEN 1980
1960 D3=K
1970 GO TO 2340
1980 GOSUB 7710
1990 N=N+1
2000 IF D5<>1 THEN 1500
2010 T4=1
2020 DELETE A1,A2,A3,A4,S
2030 N=1
2040 C2=4
2050 D1=D6

```

```

2060 GOSUB 7390
2070 T0=0
2080 T1=1
2090 T3=0
2100 T2=1
2110 T$='@CICLOS/A2'
2120 GOSUB 7980
2130 DELETE A2
2140 DIM A2(D3)
2150 A2=T
2160 T2=2
2170 T$='@CICLOS/A3'
2180 GOSUB 7980
2190 DELETE A3
2200 DIM A3(D3)
2210 A3=T
2220 T2=3
2230 T$='@CICLOS/A4'
2240 GOSUB 7980
2250 DELETE A4
2260 DIM A4(D3)
2270 A4=T
2280 T2=4
2290 T$='@CICLOS/SUM'
2300 GOSUB 7980
2310 DELETE S
2320 DIM S(D3)
2330 S=T
2340 K1=0
2350 K3=0
2360 D6=0
2370 D4=0
2380 REM   *** ANALISIS Y DECISION CON RESPECTO A LA POLARIDAD ***
2390 FOR L=1 TO D3
2400 IF N1>0 THEN 2430
2410 A3(L)=-A3(L)
2420 REM   *** ELIMINACION DE LOS CICLOS DE EXCURSION NEGATIVOS ***
2430 IF A3(L)<0 THEN 2670
2440 REM   *** ELIMINACION DE LOS CICLOS DE EXCURSION SIN VOZ ***
2450 IF S(L)<L2*0.2 THEN 2670
2460 K1=K1+1
2470 S(K1)=S(L)
2480 A2(K1)=A2(L)
2490 A3(K1)=A3(L)
2500 A4(K1)=A4(L)
2510 REM   *** CHEQUEO DE LA SEPARACION ENTRE LOS CICLOS PRINCIPALES ***
2520 IF K1<2 THEN 2640
2530 IF K3=1 THEN 2660
2540 IF (A2(K1)-A2(K1-1))/F<L3 THEN 2610
2550 K2=K2+1
2560 A2(K2)=A2(K1-1)
2570 A3(K2)=A3(K1-1)
2580 A4(K2)=A4(K1-1)

```



```

2590 IF D6=1 THEN 2710
2600 GO TO 2670
2610 IF S(K1)>S(K1-1) THEN 2670
2620 K3=1
2630 GO TO 2550
2640 K2=0
2650 GO TO 2670
2660 K3=0
2670 NEXT L
2680 K1=K1+1
2690 D6=1
2700 GO TO 2550
2710 IF K2=0 THEN 3080
2720 D7=A2(K2)
2730 D2=D2+K2
2740 IF Q/2-INT(Q/2)=0.5 THEN 2880
2750 IF N<>1 THEN 2780
2760 PRINT @A:'J'                CANDIDATOS A CICLOS PRINCIPALES'
2770 PRINT @A:''                *****'
2780 FOR I=1 TO K2 STEP 2
2790 IF I<K2 THEN 2830
2800 PRINT @A: USING 2810:'A3(',A4(I),') = ',A3(I)
2810 IMAGE 6X,3A,4D,6A,3D,3D
2820 GO TO 2870
2830 PRINT @A: USING 2840:'A3(',A4(I),') = ',A3(I)
2840 IMAGE 6X,3A,4D,6A,3D,3D,S
2850 PRINT @A: USING 2860:'A3(',A4(I+1),') = ',A3(I+1)
2860 IMAGE 6X,3A,4D,6A,3D,3D
2870 NEXT I
2880 IF N<>1 THEN 2910
2890 T0=1
2900 T3=1
2910 T1=0
2920 DELETE T
2930 DIM T(K2),A2(K2),A3(K2),A4(K2)
2940 T2=5
2950 T$='@CICLOS/PRIN/A2'
2960 T=A2
2970 GOSUB 7980
2980 T2=6
2990 T$='@CICLOS/PRIN/A3'
3000 T=A3
3010 GOSUB 7980
3020 T2=7
3030 T$='@CICLOS/PRIN/A4'
3040 T=A4
3050 GOSUB 7980
3060 T4=0
3070 N=N+1
3080 IF D5<>1 THEN 2060
3090 CLOSE 9
3100 CLOSE 1
3110 CLOSE 2

```

```

3120 CLOSE 3
3130 CLOSE 4
3140 DELETE A1,S
3150 D3=D2
3160 T0=0
3170 T3=0
3180 T4=1
3190 T1=1
3200 T2=5
3210 T#='@CICLOS/PRIN/A2'
3220 GOSUB 7980
3230 DELETE A2
3240 DIM A2(D3)
3250 A2=T
3260 T2=6
3270 T#='@CICLOS/PRIN/A3'
3280 GOSUB 7980
3290 DELETE A3
3300 DIM A3(D3)
3310 A3=T
3320 T2=7
3330 T#='@CICLOS/PRIN/A4'
3340 GOSUB 7980
3350 DELETE A4
3360 DIM A4(D3)
3370 A4=T
3380 REM   *** DETERMINACION DEL NUMERO DE REGIONES CONTINUAS ***
3390 K3=1
3400 J=1
3410 D7=INT(D7/200)
3420 IF D7>0 THEN 3440
3430 D7=1
3440 DELETE B1,B2
3450 DIM B1(D7),B2(D7)
3460 B1(K3)=A2(1)
3470 FOR I=2 TO D2
3480 IF (A2(I)-A2(I-1))/F<=L4 THEN 3550
3490 IF J<2 THEN 3520
3500 B2(K3)=A2(I-1)
3510 K3=K3+1
3520 B1(K3)=A2(I)
3530 J=1
3540 GO TO 3560
3550 J=J+1
3560 NEXT I
3570 IF J<2 THEN 3600
3580 B2(K3)=A2(D2)
3590 GO TO 3610
3600 K3=K3-1
3610 DIM B1(K3),B2(K3)
3620 REM   *** UBICACION DE LOS NUCLEOS SILABICOS PRIMARIOS ***
3630 K4=0
3640 FOR I=1 TO K3

```

```

3650 J=K4
3660 J=J+1
3670 IF B1(I)=A2(J) THEN 3690
3680 IF J<D2 THEN 3660
3690 M0=0
3700 L=J-1
3710 L=L+1
3720 IF B2(I)<A2(L) THEN 3780
3730 IF M0=>A3(L) THEN 3760
3740 M0=A3(L)
3750 N0=A4(L)
3760 IF L<I2 THEN 3710
3770 GO TO 3800
3780 L=L-1
3790 STOP
3800 REM   *** UBICACION DE LOS NUCLEOS SILABICOS SECUNDARIOS ***
3810 D7=INT((N0-B1(I)-L5*F)/(L6*F)+(B2(I)-N0-L5*F)/(L6*F)+2)
3820 IF D7=>1 THEN 3840
3830 D7=0
3840 DELETE M,N
3850 DIM M(D7+1),N(D7+1)
3860 M(1)=M0
3870 N(1)=N0
3880 K5=1
3890 IF K5>D7 THEN 4070
3900 M(K5+1)=0
3910 N(K5+1)=0
3920 IF (N(K5)-B1(I))/F<L5 THEN 4070
3930 I1=J-1
3940 I1=I1+1
3950 IF N(K5)-L5*F<A4(I1) THEN 4000
3960 IF M(K5+1)=>A3(I1) THEN 4040
3970 M(K5+1)=A3(I1)
3980 N(K5+1)=A4(I1)
3990 GO TO 4040
4000 IF N(K5+1)+L6*F<A4(I1) THEN 4050
4010 IF M(K5+1)=>A3(I1) THEN 4040
4020 M(K5+1)=0
4030 GO TO 4050
4040 IF I1<L THEN 3940
4050 K5=K5+1
4060 GO TO 3890
4070 O=0
4080 DELETE C1,C2
4090 DIM C1(K5),C2(K5)
4100 FOR I2=1 TO K5
4110 IF M(I2)=0 THEN 4150
4120 O=O+1
4130 C1(O)=M(I2)
4140 C2(O)=N(I2)
4150 NEXT I2
4160 FOR I3=1 TO O
4170 M(I3)=C1(O+1-I3)

```

```

4180 N(I3)=C2(O+1-I3)
4190 NEXT I3
4200 DELETE C1,C2
4210 K5=0
4220 IF K5>D7 THEN 4420
4230 M(K5+1)=0
4240 N(K5+1)=0
4250 IF (B2(I)-N(K5))/F<L5 THEN 4420
4260 FOR I1=J TO L
4270 IF N(K5)+L5*F>A4(I1) THEN 4310
4280 IF M(K5+1)>A3(I1) THEN 4310
4290 M(K5+1)=A3(I1)
4300 N(K5+1)=A4(I1)
4310 NEXT I1
4320 I2=J-1
4330 I2=I2+1
4340 IF N(K5+1)-L6*F>A4(I2) THEN 4390
4350 IF N(K5)+L5*F<A4(I2) THEN 4400
4360 IF M(K5+1)=>A3(I2) THEN 4390
4370 M(K5+1)=0
4380 GO TO 4400
4390 IF I2<L THEN 4330
4400 K5=K5+1
4410 GO TO 4220
4420 B3=O-1
4430 FOR I3=0 TO K5
4440 IF M(I3)=0 THEN 4480
4450 B3=B3+1
4460 M(B3)=M(I3)
4470 N(B3)=N(I3)
4480 NEXT I3
4490 PRINT @A:"J      NUCLEOS SILABICOS DE LA REGION CONTINUA ";I
4500 PRINT @A:"      *****"
4510 FOR IO=1 TO B3 STEP 2
4520 IF IO<B3 THEN 4560
4530 PRINT @A: USING 4540:"M (' ,N(IO),') = ' ,M(IO)
4540 IMAGE 6X,3A,4D,6A,3D,3D
4550 GO TO 4600
4560 PRINT @A: USING 4570:"M (' ,N(IO),') = ' ,M(IO)
4570 IMAGE 6X,3A,4D,6A,3D,3D,S
4580 PRINT @A: USING 4590:"M (' ,N(IO+1),') = ' ,M(IO+1)
4590 IMAGE 6X,3A,4D,6A,3D,3D
4600 NEXT IO
4610 REM      *** UBICACION DE LOS LIMITES SILABICOS ***
4620 K4=J
4630 DELETE M1,N1
4640 DIM M1(B3+1),N1(B3+1)
4650 L7=20
4660 IF B3<2 THEN 4690
4670 IF B3<3 THEN 4690
4680 L7=40
4690 FOR K=1 TO B3-1
4700 I2=INT((N(K+1)-N(K))/(L7*F))+1

```

```

4710 IF I2=>1 THEN 4730
4720 I2=1
4730 DELETE M2,N2
4740 DIM M2(I2),N2(I2)
4750 FOR I1=1 TO I2
4760 M2(I1)=0
4770 I3=K4-1
4780 I3=I3+1
4790 IF N(K)+(I1-1)*L7*F>A4(I3) THEN 4850
4800 IF N(K+1)<A4(I3) THEN 4860
4810 IF N(K)+I1*L7*F<A4(I3) THEN 4860
4820 IF M2(I1)=>A3(I3) THEN 4850
4830 M2(I1)=A3(I3)
4840 N2(I1)=A4(I3)
4850 IF I3<L THEN 4780
4860 K4=I3
4870 NEXT I1
4880 M1(K+1)=M2(1)
4890 N1(K+1)=N2(1)
4900 FOR I1=2 TO I2
4910 IF M1(K+1)<=M2(I1) THEN 4940
4920 M1(K+1)=M2(I1)
4930 N1(K+1)=N2(I1)
4940 NEXT I1
4950 NEXT K
4960 REM *** CONSIDERACION DE LOS EXTREMOS COMO LIMITES SILABICOS ***
4970 M1(1)=A3(J)
4980 N1(1)=B1(I)
4990 M1(B3+1)=A3(L)
5000 N1(B3+1)=B2(I)
5010 PRINT @A:"J          LIMITES SILABICOS DE LA REGION CONTINUA ";I
5020 PRINT @A:"          *****"
5030 FOR IO=1 TO B3+1 STEP 2
5040 IF IO<B3+1 THEN 5080
5050 PRINT @A: USING 5060:"M1(',N1(IO),') = ',M1(IO)
5060 IMAGE 6X,3A,4D,6A,3D,3D
5070 GO TO 5120
5080 PRINT @A: USING 5090:"M1(',N1(IO),') = ',M1(IO)
5090 IMAGE 6X,3A,4D,6A,3D,3D,S
5100 PRINT @A: USING 5110:"M1(',N1(IO+1),') = ',M1(IO+1)
5110 IMAGE 6X,3A,4D,6A,3D,3D
5120 NEXT IO
5130 REM *** SEGMENTACION DE CADA SILABA ***
5140 L7=20
5150 I6=0
5160 I7=0
5170 K4=J
5180 DELETE B4
5190 DIM B4(B3)
5200 FOR K=1 TO B3
5210 K5=0
5220 I4=INT((N1(K+1)-N1(K))/(L7*F))+1
5230 IF I4=>1 THEN 5250

```

```

5240 I4=1
5250 REM *** DETERMINACION DE MAXIMA AMPLITUD DE CADA SEGMENTO ***
5260 DELETE M2,N2,C2,C3,C4
5270 DIM M2(I4+1),N2(I4+1),C2(L+1-J),C3(L+1-J),C4(L+1-J)
5280 FOR I3=1 TO I4
5290 M2(I3)=0
5300 N2(I3)=0
5310 I2=K4-1
5320 I2=I2+1
5330 IF N1(K)+(I3-1)*L7*F>A4(I2) THEN 5390
5340 IF N1(K+1)<A4(I2) THEN 5400
5350 IF N1(K)+I3*L7*F<A4(I2) THEN 5400
5360 IF M2(I3)=>A3(I2) THEN 5390
5370 M2(I3)=A3(I2)
5380 N2(I3)=A4(I2)
5390 IF I2<L THEN 5320
5400 NEXT I3
5410 REM *** INICIALIZACION DE EXTREMOS DE CADA ENVOLVENTE LINEAL ***
5420 X4=N1(K)
5430 Y4=M1(K)
5440 M2(I4+1)=M1(K+1)
5450 N2(I4+1)=N1(K+1)
5460 I1=K4-1
5470 REM *** FORMACION DE LA ENVOLVENTE LINEAL ***
5480 FOR I3=1 TO I4+1
5490 IF N2(I3)=0 THEN 5670
5500 X5=N2(I3)
5510 Y5=M2(I3)
5520 IF I1=>L THEN 5680
5530 I1=I1+1
5540 IF A4(I1)<X4 THEN 5630
5550 IF A4(I1)>X5 THEN 5650
5560 REM *** SELECCION DE LOS CICLOS PRINCIPALES DE CADA SILABA ***
5570 Y=(Y5-Y4)/(X5-X4)*(A4(I1)-X5)+Y5
5580 IF Y*0.8>A3(I1) THEN 5630
5590 K5=K5+1
5600 C2(K5)=A2(I1)
5610 C3(K5)=A3(I1)
5620 C4(K5)=A4(I1)
5630 IF I1<L THEN 5530
5640 GO TO 5680
5650 X4=X5
5660 Y4=Y5
5670 NEXT I3
5680 K4=I1
5690 B4(K)=K5
5700 IF K5=0 THEN 6080
5710 IF K<>1 THEN 5740
5720 PRINT @A:'J' ,          CICLOS PRINCIPALES DE LA REGION CONTINUA ';I
5730 PRINT @A:'          *****'
5740 FOR IO=1 TO K5 STEP 2
5750 IF IO<K5 THEN 5790
5760 PRINT @A: USING 5770:'C3(',C4(IO),') = ',C3(IO)

```

```

5770 IMAGE 6X,3A,4D,6A,3D,3D
5780 GO TO 5830
5790 PRINT @A: USING 5800:"C3(',C4(I0),' ) = ',C3(I0)
5800 IMAGE 6X,3A,4D,6A,3D,3D,S
5810 PRINT @A: USING 5820:"C3(',C4(I0+1),' ) = ',C3(I0+1)
5820 IMAGE 6X,3A,4D,6A,3D,3D
5830 NEXT I0
5840 IF D6<>0 THEN 5870
5850 T0=1
5860 T3=1
5870 T1=0
5880 DELETE T
5890 DIM T(K5),C2(K5),C3(K5),C4(K5)
5900 T2=1
5910 T$="@CICLOPRIN/C2"
5920 T=C2
5930 GOSUB 7980
5940 T2=2
5950 T$="@CICLOPRIN/C3"
5960 T=C3
5970 GOSUB 7980
5980 T2=3
5990 T$="@CICLOPRIN/C4"
6000 T=C4
6010 GOSUB 7980
6020 T0=0
6030 T3=0
6040 T4=0
6050 D7=D7+K5
6060 IF D6=>B4(K) THEN 6080
6070 D6=B4(K)
6080 NEXT K
6090 REM ***** ESTIMACIONES DEL PERIODO DE TONO ***
6100 DELETE E1,E,B5,C2,C3,C4
6110 DIM E1(B3),E(B3,D6-1),B5(B3),C2(D7),C3(D7),C4(D7)
6120 T0=0
6130 T3=0
6140 T4=1
6150 T1=1
6160 D3=D7
6170 T2=1
6180 T$="@CICLOPRIN/C2"
6190 GOSUB 7980
6200 C2=T
6210 T2=2
6220 T$="@CICLOPRIN/C3"
6230 GOSUB 7980
6240 C3=T
6250 T2=3
6260 T$="@CICLOPRIN/C4"
6270 GOSUB 7980
6280 C4=T
6290 K7=1

```

```

6300 FOR K=1 TO B3
6310 E1(K)=0
6320 K5=0
6330 FOR I1=K7 TO B4(K)-1
6340 IF C2(I1+1)-C2(I1)=0 THEN 6390
6350 K5=K5+1
6360 E(K,K5)=(C2(I1+1)-C2(I1))/F
6370 REM   *** CALCULO DE LA FRECUENCIA PROMEDIO DE CADA SILABA ***
6380 E1(K)=E1(K)+1000/E(K,I1)
6390 NEXT I1
6400 K7=B4(K)
6410 IF K5<>0 THEN 6460
6420 K5=1
6430 E(K,K5)=0
6440 E1(K)=0
6450 GO TO 6470
6460 E1(K)=E1(K)/K5
6470 IF E1(K)=>40 THEN 6490
6480 E1(K)=40
6490 PRINT @A:'J           ESTIMACIONES DE TONO DE LA REGION CONTINUA ';I
6500 PRINT @A:'           *****'
6510 IF B3=1 THEN 6540
6520 PRINT @A:'J           SILABA ';K
6530 PRINT @A:'           -----'
6540 FOR IO=1 TO K5 STEP 2
6550 IF IO<K5 THEN 6590
6560 PRINT @A: USING 6570:'E(',K,',',',IO,') = ',E(K,IO)
6570 IMAGE 6X,2A,2D,A,2D,6A,3D,3D
6580 GO TO 6630
6590 PRINT @A: USING 6600:'E(',K,',',',IO,') = ',E(K,IO)
6600 IMAGE 6X,2A,2D,A,2D,6A,3D,3D,S
6610 PRINT @A: USING 6620:'E(',K,',',',IO+1,') = ',E(K,IO+1)
6620 IMAGE 6X,2A,2D,A,2D,6A,3D,3D
6630 NEXT IO
6640 REM   *** CORRECCION DE ERRORES DE DUPLICACION Y DE DIVISION ***
6650 REM           PARA DOS DE LAS FRECUENCIAS
6660 IF K>1 THEN 6690
6670 E2=E1(K)
6680 GO TO 6770
6690 IF E1(K)<2*E2 THEN 6720
6700 E1(K)=E1(K)/2
6710 GO TO 6690
6720 IF E1(K)>E2/2 THEN 6750
6730 E1(K)=E1(K)*2
6740 GO TO 6720
6750 REM   *** CALCULO DE LA FRECUENCIA MEDIA DE TONO DE LA REGION ***
6760 E2=(E2*(K-1)+E1(K))/K
6770 REM   *** ELIMINACION DE CICLOS PRINCIPALES MAYORES EN UN ***
6780 REM   *** INTERVALO ALREDEDOR DE CADA CICLO PRINCIPAL ***
6790 L8=3*F*1000/(2*E1(K))
6800 FOR I1=1 TO B4(K)
6810 IF C3(I1)=0 THEN 6930
6820 I2=1

```



```

6830 I2=I2+1
6840 IF I1=I2 THEN 6920
6850 IF C4(I1)-L8/2>C4(I2) THEN 6920
6860 IF C4(I1)+L8/2<C4(I2) THEN 6930
6870 IF C3(I2)=>C3(I1) THEN 6900
6880 C3(I2)=0
6890 GO TO 6920
6900 C3(I1)=0
6910 GO TO 6930
6920 IF I2<B4(K)-1 THEN 6830
6930 NEXT I1
6940 REM   *** CALCULO DE LOS PERIODOS DE TONO DEL NUEVO CONJUNTO ***
6950 REM   *** DE CICLOS PRINCIPALES ***
6960 K6=0
6970 I3=0
6980 I3=I3+1
6990 IF C3(I3)=0 THEN 7050
7000 I4=I3+1
7010 IF C3(I4)=0 THEN 7070
7020 IF C2(I4)-C2(I3)=0 THEN 7050
7030 K6=K6+1
7040 E(K,K6)=(C2(I4)-C2(I3))/F
7050 IF I3<B4(K)-1 THEN 6980
7060 GO TO 7100
7070 IF I4>B4(K)-1 THEN 7100
7080 I4=I4+1
7090 GO TO 7010
7100 IF K6<>0 THEN 7130
7110 K6=1
7120 E(K,K6)=0
7130 B5(K)=K6
7140 DIM E(K,K6)
7150 PRINT @A: 'JJ           PERIODOS DE TONO DE LA REGION CONTINUA ';I
7160 PRINT @A: '           *****'
7170 IF B3=1 THEN 7200
7180 PRINT @A: 'J           SILABA ';K
7190 PRINT @A: '           -----'
7200 FOR IO=1 TO K6
7210 PRINT @A: USING 7220: 'E(',K,',',',IO,')    = ',E(K,IO), ' MSEG'
7220 IMAGE 15X,2A,2D,A,2D,7A,3D,3D,5A
7230 NEXT IO
7240 NEXT K
7250 NEXT I
7260 DELETE A1,A2,A3,A4,B1,B2,B4,B5,C1,C2,C3,C4,E,E1,M,N,M1,N1,S,T
7270 CALL 'TIME',Z$
7280 X$=SEG(Z$,11,2)
7290 Z2=VAL(X$)*3600
7300 X$=SEG(Z$,14,2)
7310 Z2=Z2+VAL(X$)*60
7320 X$=SEG(Z$,17,2)
7330 Z2=Z2+VAL(X$)
7340 Z2=Z2-Z1
7350 PRINT @A: USING 7360: 'TIEMPO DE EJECUCION = ',Z2, ' SEGUNDOS'

```

```

7360 IMAGE /,6X,22A,4D,9A
7370 PRINT 'JJ FIN DE LA DETECCION DE TONO, OPRIMA TECLA DESEADA'
7380 END
7390 REM ***** SUBROUTINA DE CHEQUEO DE MEMORIA *****
7400 IF N<>1 THEN 7510
7410 D4=1
7420 C1=(8*D1+18)*C2
7430 IF MEMORY-C1>K4 THEN 7490
7440 D3=INT((MEMORY-K4-18)/(8*C2))
7450 D5=0
7460 IF D3=>D1 THEN 7490
7470 D4=0
7480 GO TO 7510
7490 D3=D1
7500 D5=1
7510 IF D4=1 THEN 7540
7520 IF D3*N<D1 THEN 7550
7530 D3=D1-(N-1)*D3
7540 D5=1
7550 RETURN
7560 REM ***** SUBROUTINA DE LECTURA DE ARCHIVOS *****
7570 IF N>1 THEN 7640
7580 D5=0
7590 CLOSE
7600 OPEN L$;9,'R',X$
7610 READ #9:D1,F
7620 GOSUB 7390
7630 IF D4=1 THEN 7660
7640 IF D3*N<D1 THEN 7670
7650 D3=D1-(N-1)*D3
7660 D5=1
7670 DELETE A1
7680 DIM A1(D3)
7690 READ #9:A1
7700 RETURN
7710 REM *** ALMACENAMIENTO DE CICLOS DE EXCURSION ***
7720 IF N<>1 THEN 7750
7730 T0=1
7740 T3=1
7750 T1=0
7760 DELETE T
7770 DIM T(K)
7780 T2=1
7790 T$='@CICLOS/A2'
7800 T=A2
7810 GOSUB 7980
7820 T2=2
7830 T$='@CICLOS/A3'
7840 T=A3
7850 GOSUB 7980
7860 T2=3
7870 T$='@CICLOS/A4'
7880 T=A4

```

```

7890 GOSUB 7980
7900 T2=4
7910 T$='@CICLOS/SUM'
7920 T=S
7930 GOSUB 7980
7940 T0=0
7950 T3=0
7960 T4=0
7970 RETURN
7980 REM   *** SUBROUTINA DE MANEJO DE ARCHIVOS ***
7990 IF T0<>1 THEN 8020
8000 KILL T$
8010 CREATE T$;200,0
8020 IF T3<>1 THEN 8220
8030 T4=1
8040 GO TO T2 OF 8050,8070,8090,8110,8130,8150,8170,8190,8210
8050 OPEN T$;1,'U',X$
8060 GO TO 8220
8070 OPEN T$;2,'U',X$
8080 GO TO 8220
8090 OPEN T$;3,'U',X$
8100 GO TO 8220
8110 OPEN T$;4,'U',X$
8120 GO TO 8220
8130 OPEN T$;5,'U',X$
8140 GO TO 8220
8150 OPEN T$;6,'U',X$
8160 GO TO 8220
8170 OPEN T$;7,'U',X$
8180 GO TO 8220
8190 OPEN T$;8,'U',X$
8200 GO TO 8220
8210 OPEN T$;9,'U',X$
8220 IF T4<>1 THEN 8410
8230 GO TO T2 OF 8240,8260,8280,8300,8320,8340,8360,8380,8400
8240 CALL 'REWIND',1
8250 GO TO 8410
8260 CALL 'REWIND',2
8270 GO TO 8410
8280 CALL 'REWIND',3
8290 GO TO 8410
8300 CALL 'REWIND',4
8310 GO TO 8410
8320 CALL 'REWIND',5
8330 GO TO 8410
8340 CALL 'REWIND',6
8350 GO TO 8410
8360 CALL 'REWIND',7
8370 GO TO 8410
8380 CALL 'REWIND',8
8390 GO TO 8410
8400 CALL 'REWIND',9
8410 IF T1<>1 THEN 8630

```

```

8420 DELETE T
8430 DIM T(D3)
8440 GO TO T2 OF 8450,8470,8490,8510,8530,8550,8570,8590,8610
8450 READ #1:T
8460 GO TO 8810
8470 READ #2:T
8480 GO TO 8810
8490 READ #3:T
8500 GO TO 8810
8510 READ #4:T
8520 GO TO 8810
8530 READ #5:T
8540 GO TO 8810
8550 READ #6:T
8560 GO TO 8810
8570 READ #7:T
8580 GO TO 8810
8590 READ #8:T
8600 GO TO 8810
8610 READ #9:T
8620 GO TO 8810
8630 GO TO T2 OF 8640,8660,8680,8700,8720,8740,8760,8780,8800
8640 WRITE #1:T
8650 GO TO 8810
8660 WRITE #2:T
8670 GO TO 8810
8680 WRITE #3:T
8690 GO TO 8810
8700 WRITE #4:T
8710 GO TO 8810
8720 WRITE #5:T
8730 GO TO 8810
8740 WRITE #6:T
8750 GO TO 8810
8760 WRITE #7:T
8770 GO TO 8810
8780 WRITE #8:T
8790 GO TO 8810
8800 WRITE #9:T
8810 RETURN

```

CLASIFICACION DE LOS ARCHIVOS EN EL DISCO

1. Programas de tesis

SCRATCHLIB/TESIS/CTOBON

SCRATCHLIB/DATOS

SCRATCHLIB/DETPAR

SCRATCHLIB/DETRED

2. Duplicados de los programas de tesis

CALIS/TESIS

CALIS/DATOS

CALIS/DETPAR

CALIS/DETRED

3. Archivos de datos de segmentos de voz

VOZ/LETRA/I1

VOZ/LETRA/I2

VOZ/LETRA/E1

VOZ/LETRA/E2

VOZ/LETRA/A1

VOZ/LETRA/A2

VOZ/LETRA/O1

VOZ/LETRA/O2

VOZ/LETRA/U1

VOZ/LETRA/U2

VOZ/SILAB/UHMA

VOZ/SILAB/UHNA

4. Archivo de trabajo para uso en todos los programas.

MUESTREO/A1

5. Archivo para uso en el programa "DATOS"

ESQUINAS

6. Archivos para uso en el programa "DETPAR"

POSICION/PICOS

POSICION/VÁLLES

TREN/UNO

TREN/DOS

TREN/TRES

TREN/CUATRO

TREN/CINCO

TREN/SEIS

ESTIMACION/TONO

RESULTADOS

7. Archivos para uso en el programa "DETRED"

CICLOS/A2

CICLOS/A3

CICLOS/A4

CICLOS/SUM

CICLOS/PRIN/A2

CICLOS/PRIN/A3

CICLOS/PRIN/A4

CICLOPRIN/C2

CICLOPRIN/C3

CICLOPRIN/C4

B I B L I O G R A F I A

B I B L I O G R A F I A

- (1) H. Fletcher, Speech and Hearing in Communication, D. Van Nostrand Co., Nueva York, 1.972.
- (2) J.L. Flanagan, Speech Analysis, Synthesis and Perception, 2nda. edición, Springer - Verlag, Nueva York, 1.972.
- (3) W. Koenig, H.K. Dunn y L. Y. Lacy, The Sound Spectrograph, J. Acoust. Soc. Am., vol. 17, p. 19-49, 1.946.
- (4) R. Jakobson, C. Fant y M. Halle, Preliminaries to Speech Analysis, M.I.T., Cambridge, 1.963.
- (5) N. Chomsky y M. Halle, The Sound Pattern of English, Harper y Row, Nueva York, 1.968.
- (6) A. Holbrook y G Fairbanks, Diphthong Formants and Their Movements, J. of Speech and Hearing Research, vol. 5, p. 38-58, 1952.
- (7) O. Fujimura, Analysis of Nasal Consonants, J. Acoust. Soc. Am., vol. 34, p. 1865-1875, 1.962.
- (8) J.M. Heinz y K. N. Stevens, On the Propierties of Voiceless Fricative Consonants, J. Acoust. Soc. Am., vol. 33, p. 589, 1.961.
- (9) P. Delattre y A. Liberman, Acoustic Loei--and-Transitional-Cuesfor Consonants, J. Acoust. Soc. Am., vol. 27, p. 769, 1.955.
- (10) L.R. Rabiner y R. W. Schafer, Digital Processing of Speech Signals, Prentice Hall, Nueva Jersey, 1.978.
- (11) B. Malmberg, La Fonética, 5ta. Edición, Eudeba, Buenos Aires, 1.972.
- (12) J.D. Bowen y R. P. Stockwell, The Phonemic Interpretation of Semivowels in Spanish, en Langauge, p. 236-240, 1.955.

- (13) E. Alarcos Llorach, Fonología Española, 3era. edición, Gredos, Madrid, 1.961.
- (14) E. Zwicker y E. Terhardt, Automatic Speech Recognition using Psychoacoustic Models, J. Acoust. Soc. Am., vol. 65 No. 2, p. 487-489, 1.979.
- (15) L. Rabiner, M. Cheng, A. Rosenberg y C. McGonegal, A Comparative Performance Study of Several Pitch Detection Algorithms, IEEE Trans. Acoust., Speech and Signal Process., vol. ASSP-24, p. 399-417, 1.976.
- (16) B. Gold y L. Rabiner, Parallel Processing Techniques for Estimating Pitch Periods of Speech in the Time Domain, J. Acoust. Soc. Am., vol. 46, p. 442-448, 1.969.
- (17) N. J. Miller, Pitch Detection by Data Reduction, IEEE Trans. Acoust., Speech and Signal Process., vol. ASSP-23, No. 1, p. 72-79, 1.975.
- (18) M. Ross, H. Shaffer, A. Cohen, R. Freudberg y H. Manley, Average Magnitude Difference Function Pitch Extractor, IEEE Trans. Acoust., Speech and Signal Process., vol. ASSP-22, p. 353-362, 1.974.
- (19) A. M. Noll, Cepstrum Pitch Determination, J. Acoust. Soc. Am., vol. 41, p. 293-309, 1.967.
- (20) T. V. Screenivas y P. V. S. Rao, Pitch Extraction from Corrupted Harmonics of the Power Spectrum, J. Acoust. Soc. Am., vol. 65, No. 1, p. 223-228, 1.979.
- (21) M. Mohan Sondhi, New Methods of Pitch Extraction, IEEE Trans. Audio Electroacoust., vol. AU-16, p. 262-266, 1.968.
- (22) J. D. Markel, The SIFT Algorithm for Fundamental Frequency Estimation, IEEE Trans. Audio Electroacoust., vol. AU-20, p. 367-377, 1.972.
- (23) B. Gold, Computer Program for Pitch Extraction, J. Acoust. Soc. Am., vol. 34, No. 7, p. 916-921, 1.962.

- (24) J. S. Gill, Estimation of Larynx-Pulse Timing During Speech, Proc. Int. Congr. Acoust., vol. 4, Copenhagen, 1.962.
- (25) B. Gold, Description of a Computer Program for Pitch Detection, Proc. Int. Congr. Acoust., vol 4, Copenhagen, 1.962.
- (26) R. J. Scott y S. E. Gerber, Pitch-synchronous Time Comprehension of Speech, Proc. Conf. Speech Commun. and Process., AFCRL -72-0120, p. 198-201, Nueva York, 1.972.
- (28) A. V. Oppenheim, Speech Analysis-Synthesis System Based on Homomorphic Filtering, J. Acoust. Soc. Am., vol. 45, p. 458-465, 1.969.
- (29) T. S. Harms, General Phonetics, Univ. of Wisconsin Press, p. 72-75, Madison-Wisconsin, 1.969.
- (30) M. R. Schroeder, Period Histogram and Product Spectrum: New Methods for Fundamental-Frequency Measurement, J. Acoust. Soc. Am., vol. 43, p. 829-834, 1.968.
- (31) J. M. Wozencraft, Principles of Communication Engineering, I. M. Jacobs, Nueva York, 1.965.
- (32) H. L. Shaffer y C. Howard, Real-Time Generation of the Cross-Correlation Function and Difference Magnitude Function, GTE Sylvania, Needham, Massachusetts, 1.970.
- (33) R. B. Blackman y J. W. Tukey, The Measurement of Power Spectra, Dover, Nueva York, 1.959.
- (34) L. R. Rabiner, On the Use of Autocorrelation Analysis for Pitch Detection, IEEE Trans. Acoust., Speech and Signal Proc., vol. ASSP-25, No. 1, p. 24-33, 1.977.
- (35) J. J. Dubnosky, R. W. Schafer y L. R. Rabiner, Real-Time Digital Hardware Pitch Detector, IEEE Trans. Acoust., Speech and Signal Proc., vol. ASSP-24, No. 1, p. 2-8, 1.976.
- (36) L. R. Rabiner y R. W. Schafer, Digital Processing of Speech and Signals, Prentice Hall, p. 150-158, Englewood Cliffs, Nueva York, 1.965.