

Procesamiento Digital de Señales

Parcial 2

Gianfranco Fagioli, Gaspar Oberti, Franco Matzkin, Facundo Salmerón,
Darién Julián Ramírez, Iván Schweikofski y *el Creador de la Biblia*.

Distribúyase y mejórese.

Índice

1. Diseño de filtros digitales	3
1.1. ¿Qué diferencias y similitudes existen entre los conceptos de sistema y filtro?	3
1.2. Explique que es un filtro de compensación de fase, un filtro de fase lineal y un filtro adaptativo.	3
1.3. Ejemplifique dos casos de aplicación donde se requiera la utilización de filtrado estático y adaptativo respectivamente. Justifique su respuesta.	3
1.4. Liste las ventajas y desventajas de utilizar un filtro de Butterworth.	4
1.5. Liste las ventajas y desventajas de utilizar un filtro de Chebyshev tipo II.	4
1.6. Detalle los pasos para el diseño de un filtro de Chebyshev tipo II a partir de su prototipo analógico.	5
1.7. Describa sucintamente una técnica de diseño de filtros IIR y una de filtros FIR.	5
1.8. Apéndice	6
1.8.1. Comparativa de filtros	6
2. Identificación de Sistemas	6
2.1. Escriba una definición de identificación de sistemas.	6
2.2. ¿Cuáles son las limitaciones del método de predicción lineal?	6
2.3. Indique si las siguientes afirmaciones respecto a los métodos de identificación de sistemas basados en predicción lineal son correctas o no, justificando sus respuestas:	6
2.3.1. No pueden utilizarse en sistemas con entradas aleatorias.	6
2.3.2. Son los métodos más sencillos para identificar un sistema lineal de orden pequeño.	6
2.3.3. No sirven para encontrar los parámetros de un sistema ARMA.	7
2.3.4. Pueden extenderse fácilmente a los sistemas no lineales.	7
2.4. Deduzca la ecuación de Wiener-Hopf utilizada en el método de predicción lineal para el caso de señales determinísticas.	7
2.5. Deduzca la ecuación para el cálculo de la ganancia en el método de predicción lineal.	8
2.6. Describa los métodos de estimación de orden y liste sus ventajas relativas.	9
3. Procesamiento de la Voz	9
3.1. Describa y clasifique desde el punto de vista fenomenológico a la señal producida a nivel de las cuerdas vocales. ¿Qué características posee su espectro? Grafique.	9
3.2. La figura representa el modelo lineal de producción de la señal de voz:	10
3.2.1. ¿Qué tipos de señales de excitación se utilizan? ¿Qué parámetros definen cada tipo de excitación? ¿Cómo los estimarías?	10
3.2.2. ¿Qué modela el bloque denominado <i>sistema</i> y de que tipo es? ¿Cuántos parámetros se necesitan definirlo? ¿Cómo se estiman estos parámetros?	10
3.3. Proponga un método para calcular la frecuencia fundamental (F_0) de un tramo de señal de voz.	11
3.4. ¿A que se denominan <i>frecuencias formantes</i> ? ¿Cómo se denotan y que importancia tienen en la generación de la voz?	11
3.5. ¿Qué es la escala de mel y cómo se utiliza para el análisis de voz?	11
3.6. Describa las hipótesis en las que se basa el análisis cepstral del habla para separar la señal de excitación, de la respuesta del tracto vocal ¿Son realistas estas hipótesis?	11

3.7.	Considere la gráfica de una vocal sostenida $x[n]$ que se muestra en la Figura 1.(a). Indique, dibujando sobre la misma, un Período Fundamental (T_0) y estime (aproximadamente) su duración. La frecuencia de muestreo es de 8 kHz:	12
3.7.1.	Defina el <i>cepstrum real</i> $c[n]$	12
3.7.2.	A partir de la Figura 1.(b), correspondiente al <i>cepstrum real</i> del segmento de vocal mostrado en la Figura 1.(a), indique cómo estimar la Frecuencia Fundamental (F_0) del segmento y compare con el resultado obtenido en el punto anterior.	12
3.8.	Proponga un método para contar la cantidad de cruces por cero de un tramo de señal de voz. ¿Cómo haría para que el método fuera robusto al ruido?	12
3.9.	Indique las consideraciones especiales que deben tenerse en cuenta al aplicar el método de predicción lineal a señales de voz. ¿En qué casos aplicaría el método adaptativo de Widrow?	13
4.	Análisis Tiempo-Frecuencia	13
4.1.	Explique la necesidad de utilizar transformaciones tiempo-frecuencia.	13
4.2.	Defina la <i>Transformada de Fourier de Corta Duración</i> (STFT), en sus versiones continua y discreta.	13
4.3.	Explique al menos dos maneras de interpretar conceptualmente a la STFT. Utilice gráficos detallando las magnitudes y unidades de los ejes.	14
4.4.	Explique en que consiste el <i>principio de incertidumbre</i> y como se expresa analíticamente.	14
4.5.	¿Es posible interpretar la STFT en términos de una <i>base</i> ? En caso afirmativo ejemplifique graficando algunos elementos de la misma.	15
4.6.	Describa el algoritmo rápido para el cálculo de la Transformada Ondita Diádica.	15
4.7.	Defina la <i>Distribución de Wigner-Ville</i> en su versión continua. Resuma ventajas y desventajas.	15
4.8.	Esquematice y compare las particiones del plano tiempo-frecuencia que se obtienen mediante una representación que utilice una base de <i>deltas de Dirac</i> , la base de la <i>Transformada Discreta de Fourier</i> , la <i>Transformada de Fourier de Corta Duración</i> y la <i>Transformada Onditas Discreta Diádica</i>	16
5.	Problemas	17
5.1.	Problema 1	17
5.2.	Problema 2	18
5.3.	Problema 3	18
5.4.	Problema 4	18
5.5.	Problema 5	19
5.6.	Problema 6	19

1. Diseño de filtros digitales

1.1. ¿Qué diferencias y similitudes existen entre los conceptos de sistema y filtro?

1. Ambos relacionan la entrada y la salida mediante:
 - Una ecuación (digitales).
 - Hardware (analógicos).
2. Los filtros son un caso particular de sistema LTI (no introducen componentes frecuenciales).
3. Los filtros son un sistema pero se estudian desde la perspectiva de *¿qué se quiere dejar pasar o qué no?*.

$$\text{Filtro} \implies \text{Sistema}$$

1.2. Explique que es un filtro de compensación de fase, un filtro de fase lineal y un filtro adaptativo.

- Clasificación de filtros según su comportamiento en el tiempo:
 - *Filtro adaptativo*: Se adaptan al contexto variando en el tiempo los parámetros característicos (ancho de banda, frecuencia de corte, etc). El hecho de tener coeficientes variables del filtro, es necesario cuando se conocen *a priori* las características estadísticas de la señal a filtrar o cuando se conoce y se sabe que son variables en el tiempo. Se usan para analizar señales aleatorias no estacionarias.
- Clasificación de filtros según su fase:
 - *Filtro de fase lineal*: Modifican la fase de la señal de entrada de forma lineal, proporcional a la frecuencia, independiente del módulo.
 - *Filtro de compensación de fase*: No eliminan componentes frecuenciales, sólo modifican la fase. Tienen como finalidad compensar problemas de distorsión de fase que sean previos a la aplicación del filtro.

1.3. Ejemplifique dos casos de aplicación donde se requiera la utilización de filtrado estático y adaptativo respectivamente. Justifique su respuesta.

- a. *Filtro estático*: Siempre funcionan filtrando de la misma manera, osea es LTI (Lineal e Invariante en el Tiempo).
 - *Ejemplo*: un ECG que se transmite por radio frecuencia el cuál se sabe que va a poseer ruido desde la frecuencia 50Hz hasta 80Hz *siempre*, entonces aplico un filtro estático rechaza-banda para eliminar la misma.
 - *Ejemplo*: filtro analógico implementado electrónicamente para eliminar algunas componentes de ruido conocidas (rechaza banda) en la señal del monitor

- *Ejemplo:* se puede aplicar un filtro a una señal de conversación entre un adulto y un niño para eliminar la voz del adulto por ejemplo utilizando un filtro pasa bajo.
- b. *Filtro adaptativo:* Se adaptan al contexto variando en el tiempo los parámetros característicos, o sea es LTV (Lineal y Variante en el Tiempo). El hecho de tener coeficientes variables del filtro, es necesario cuando se conocen *a priori* las características estadísticas de la señal a filtrar o cuando se conoce y se sabe que son variables en el tiempo.
- Se usan para analizar señales aleatorias no estacionarias.

- *Ejemplo:* Cualquier sistema en el que las características del ruido no sean constantes, ya sea un ECG que va variando su ruido o un celular el cuál posee dos micrófonos, uno para tomar la voz y otro para tomar el ruido, las cuales son aleatorias y van cambiando en el tiempo, entonces el filtro adaptativo tiene que tomar ambas y a la señal de la voz restarle el ruido interpretado por el otro micrófono.
- *Ejemplo:* La transmisión de un fax donde la misma puede haber perturbaciones en la señal eléctrica en el envío, entonces se coloca un filtro adaptativo para compensar estas perturbaciones que puede haber en la transmisión.
- *Ejemplo:* Cuando se realiza un análisis de electrocardiograma la señal de la actividad eléctrica del corazón se ve afectada por la interferencia de la fuente eléctrica del aparato medidor. Una forma de eliminar correctamente este ruido es obtener la señal de la fuente y hacer variar los coeficientes del filtro según la frecuencia que tenga la señal de la fuente.
- *Ejemplo:* Cancelación de eco de una señal de voz mediante filtro adaptativo. Por ejemplo, en la comunicación telefónica cuando se utiliza altavoz, la señal producida por el parlante se introduce en el micrófono. Se utiliza un filtro adaptativo para limpiar la señal de micrófono haciendo variar los coeficientes del filtro según las frecuencias que tenga la señal de voz.

1.4. Liste las ventajas y desventajas de utilizar un filtro de Butterworth.

- (+) No posee distorsiones en la banda de paso ni en la de rechazo.
- (-) La banda de transición es más grande que para los otros filtros a un mismo grado.

1.5. Liste las ventajas y desventajas de utilizar un filtro de Chebyshev tipo II.

- (+) Posee una banda de transición más angosta con respecto a Butterworth.
 - (+) Se puede escoger su paridad para la banda de rechazo.
 - (+) Al aumentar su orden, se estrecha la banda de transición.
 - (-) Ya tiene definida su fase y no se puede controlar, es decir, no se puede llegar a una fase lineal a diferencia de los FIR (*Finite Impulse Response*).
 - (-) Puede llegar a ser inestable por ser un filtro recursivo.
 - (-) Se requiere una gran precisión numérica para evitar la inestabilidad.
 - (-) Posee ondulaciones en la banda de rechazo.
 - (-) Banda de paso monótonamente decreciente.
- Los filtros Chebyshev de tipo II difieren de los de tipo I en que en vez de presentar ondulaciones en la banda de paso, éstas se encuentran en la banda de rechazo (siempre respetando un margen de error determinado). Por lo tanto, en la banda de paso la señal pasará sin sufrir alteraciones en la magnitud. Dependiendo del problema en particular, éstas oscilaciones pueden importar en mayor o menor medida.

- Respecto al filtro de Butterworth, posee como ventaja que la banda de transición es más estrecha, pero como desventaja las oscilaciones en la banda de rechazo.
- Respecto al elíptico, la ventaja es que posee menos oscilaciones pero como desventaja posee una banda de transición mayor.

1.6. Detalle los pasos para el diseño de un filtro de Chebyshev tipo II a partir de su prototipo analógico.

1. Se realiza en diseño analógico del filtro pasa bajo normalizado en el plano s , se establece el orden del filtro según los requerimientos.
2. *Transformación en frecuencia*: se convierte en filtro pasa bajo, pasa banda o rechaza banda y escalar para cambiar la frecuencia de corte.
3. *Transformación conforme*: se pasa al dominio digital, es decir, del dominio s al dominio z . Se puede hacer con una transformación bilineal.

Los pasos 2 y 3 se pueden intercambiar.

1.7. Describa sucintamente una técnica de diseño de filtros IIR y una de filtros FIR.

a. Técnica de diseño de filtros IIR

- 1) **Especificación** de las características de la señal (tipo de fuente de la señal, frecuencia de muestreo, ancho de banda de la señal) y la respuesta en frecuencia y tolerancia (atenuación de la banda de paso, atenuación en la banda de rechazo, frecuencia de paso, frecuencia de rechazo, frecuencia de corte).
- 2) Seleccionar uno de los métodos para calcular los **coeficientes** a_k y b_k de un filtro IIR.
 - *Método 1*:
 - Diseño analógico (filtro pasa bajo normalizado).
 - Transformación en frecuencia (analógica, en s).
 - Transformación conforme (bilineal).
 - *Método 2*:
 - Diseño analógico (filtro pasa bajo normalizado).
 - Transformación conforme (bilineal).
 - Transformación en frecuencia (digital, en z).
 - *Método directo*:
 - Colocación de polos y ceros.
- 3) **Realización**, donde se convierte la función de transferencia en una estructura de filtro adecuada.
- 4) Se analizan los **errores** de cuantización, redondeo, número de bits, etc.
- 5) **Implementación** del filtro en *software* o *hardware*.

b. Técnica de diseño de filtros FIR (Método de Fourier y Ventaneo)

- 1) **Especificación de requerimientos**, módulo y fase de la respuesta en frecuencia.
- 2) **Muestreo** de la respuesta en frecuencia.
- 3) Pasaje al dominio temporal mediante la TDF inversa. No causal y de duración infinita (*sinc*).
- 4) Truncado temporal con ventanas teniendo en cuenta el efecto de Gibbs (relación entre el lóbulo central y los lóbulos laterales).
- 5) Corrección de amplitud multiplicando la señal por un valor que la aumenta o la disminuye.
- 6) Se causal el filtro realizando un corrimiento de la señal.

1.8. Apéndice

1.8.1. Comparativa de filtros

Para un mismo orden:

	Butterworth	Chebyshev Tipo 1	Chebyshev Tipo 2	Elíptico
Banda de transición	Grande	Media	Media	Chica
Ondulaciones	No posee	Banda de paso	Banda de rechazo	Bandas de paso y rechazo
Para orden par-impar	-	Formas diferentes	Formas diferentes	Formas diferentes
Monotonía decreciente	En todas las bandas	Banda de rechazo	Banda de paso	No posee

2. Identificación de Sistemas

2.1. Escriba una definición de identificación de sistemas.

Proceso mediante el cual se pueden determinar la estructura y parámetros característicos de un sistema cuyas características intrínsecas se desconocen. En algunos casos la estructura del sistema puede ser conocida o supuesta *a priori* y la identificación se reduce a la búsqueda de los parámetros.

2.2. ¿Cuáles son las limitaciones del método de predicción lineal?

1. Supone que el sistema es lineal.
2. Sirve sólo para determinar los coeficientes del sistema (no el orden).
3. Para que produzca una buena estimación se necesita un orden muy alto.
4. La ecuación de transferencia posee sólo polos.

2.3. Indique si las siguientes afirmaciones respecto a los métodos de identificación de sistemas basados en predicción lineal son correctas o no, justificando sus respuestas:

2.3.1. No pueden utilizarse en sistemas con entradas aleatorias.

Falso.

Para señales aleatorias estacionarias también se cumple $R_a = -r$.

Para señales aleatorias no estacionarias sólo funciona para procesos localmente estacionarios que pueden ser fragmentados en procesos estacionarios.

2.3.2. Son los métodos más sencillos para identificar un sistema lineal de orden pequeño.

Falso.

Existen métodos que permiten estimar fácilmente los parámetros característicos de sistemas lineales de orden 3 o menor. Consisten en estimular al sistema con un *delta de Dirac* o el escalón y estudiar la respuesta temporal.

Mediante Laplace se puede escribir

$$Y(s) = H(s)X(s)$$

y es la respuesta, x es la entrada y H la transferencia.

$X(s)$ del *delta de Dirac* es 1, $Y(s) = H(s)$ donde haciendo la transformada inversa de la salida obtenemos la transferencia del sistema.

Otra forma es estimular el sistema con senoidales de frecuencia variable en el rango de interés y analizar la respuesta en frecuencia de la salida.

2.3.3. No sirven para encontrar los parámetros de un sistema ARMA.

Falso.

Cualquier sistema ARMA puede ser aproximado mediante un sistema AR a través de

$$H(z) = G \frac{B(z)}{A(z)} \approx \frac{G}{C(z)}$$

$C(z)$ es de orden infinito pero se puede aproximar el sistema ARMA haciendo $C(z)$ del mismo orden que $A(z)$

$$H(z) = \frac{G}{1 + \sum_{k=1}^p a_k z^{-k}}$$

quedando la ecuación de recurrencia

$$s_n = \sum_{k=1}^p a_k s_{n-k} + G\mu_n$$

p : orden óptimo.

G : Ganancia.

μ_n : entrada en el instante n .

2.3.4. Pueden extenderse fácilmente a los sistemas no lineales.

Falso.

Eliminar el supuesto de linealidad implica costos computacionales altos en sus implementaciones. Estas técnicas no convencionales pueden utilizar algoritmos genéticos, redes neuronales, programación dinámica, búsqueda aleatorias y otras.

2.4. Deduzca la ecuación de Wiener-Hopf utilizada en el método de predicción lineal para el caso de señales determinísticas.

Sea s_n la salida del sistema para el instante n , se aproxima el sistema mediante

$$\hat{s}_n = - \sum_{k=1}^p a_k s_{n-k}$$

donde p es el orden óptimo de aproximación y a los coeficientes del sistema.

El error cuadrático de esta aproximación para la n -ésima muestra está dado por

$$e_n^2 = (s_n - \hat{s}_n)^2 = (s_n + \mathbf{a}^T \mathbf{s}_n)^2$$

$$e^2 = \sum_n (s_n + \mathbf{a}^T \mathbf{s}_n)^2 = \sum_n (s_n^2 + 2\mathbf{a}^T s_n \mathbf{s}_n + \mathbf{a}^T \mathbf{s}_n \mathbf{s}_n^T \mathbf{a})$$

Se quiere minimizar este error optimizando los coeficientes del sistema, por lo tanto se calcula el gradiente

$$\nabla_a e^2 = 0$$

$$\nabla_a e^2 = \sum_n 2s_n \mathbf{s}_n + 2 \left(\sum_n \mathbf{s}_n \mathbf{s}_n^T \right) \mathbf{a} = 0$$

$$\left(\sum_n \mathbf{s}_n \mathbf{s}_n^T \right) \mathbf{a} = - \sum_n s_n \mathbf{s}_n$$

$$(\text{Matriz de correlación}) \mathbf{a} = -\text{Vector de correlación}$$

$$\mathbf{R} \mathbf{a} = -\mathbf{r}$$

Esto conforma un sistema de ecuaciones lineales denominado *sistema de Wiener-Hopf*.

2.5. Deduzca la ecuación para el cálculo de la ganancia en el método de predicción lineal.

Sea s_n la salida del sistema en el tiempo n y \hat{s}_n la estimación del sistema AR dada por

$$\hat{s}_n = - \sum_{k=1}^p a_k s_{n-k} \quad p = \text{orden del sistema}$$

el error se puede ver como

$$e_n = s_n - \hat{s}_n = G \mu_n$$

Se debe asegurar que la energía total de la señal de entrada sea igual a la energía total de la señal del error.

Las entradas que interesan son la entrada impulsiva y la de ruido blanco. Para entrada impulsiva

$$h_n = - \sum_{k=1}^n a_k h_{n-k} + G \delta_n$$

A partir del sistema de Wiener-Hopf

$$\mathbf{R} \mathbf{a} = -\mathbf{r}$$

se puede escribir

$$\begin{aligned} \mathbf{r}^T \mathbf{a} &= -r_0 + G^2 \\ G^2 &= \mathbf{r}^T \mathbf{a} + r_0 = E_p \end{aligned}$$

Para el ruido blanco es la misma expresión.

2.6. Describa los métodos de estimación de orden y liste sus ventajas relativas.

1. Error de predicción final:

Utiliza la medición de los errores residuales en la resolución de los *sistema de Wiener-Hopf*. Se asume una entrada no correlacionada y con espectro plano.

El mínimo error promedio para el sistema de orden p está dado por

$$E_p = r_0 + \mathbf{r}^T \mathbf{a}$$

Error normalizado

$$V_p = \frac{E_p}{r_0}$$

V_p es monótono decreciente con p y acotado.

Como se utilizan secuencias infinitas, la pendiente V_p no se hace exactamente cero. Por eso se utiliza un umbral para la tolerancia.

$$1 - \frac{V_{p+1}}{V_p} < \gamma$$

(-) El orden obtenido es subjetivo

(+) Es fácil de calcular

El incremento de orden se detiene cuando se cumple la inecuación.

2. Criterio de Akaike:

Para el problema de todos polos y asumiendo una distribución gaussiana en la señal, se mide el error según

$$I_p = \log E_p + \frac{2p}{N_e}$$

El orden es único

No es el número efectivo de la muestra de la señal, es decir, contemplando los efectos del ventaneo de la señal. Se calcula mediante la relación de la energía de la ventana utilizada y la energía de la ventana cuadrada.

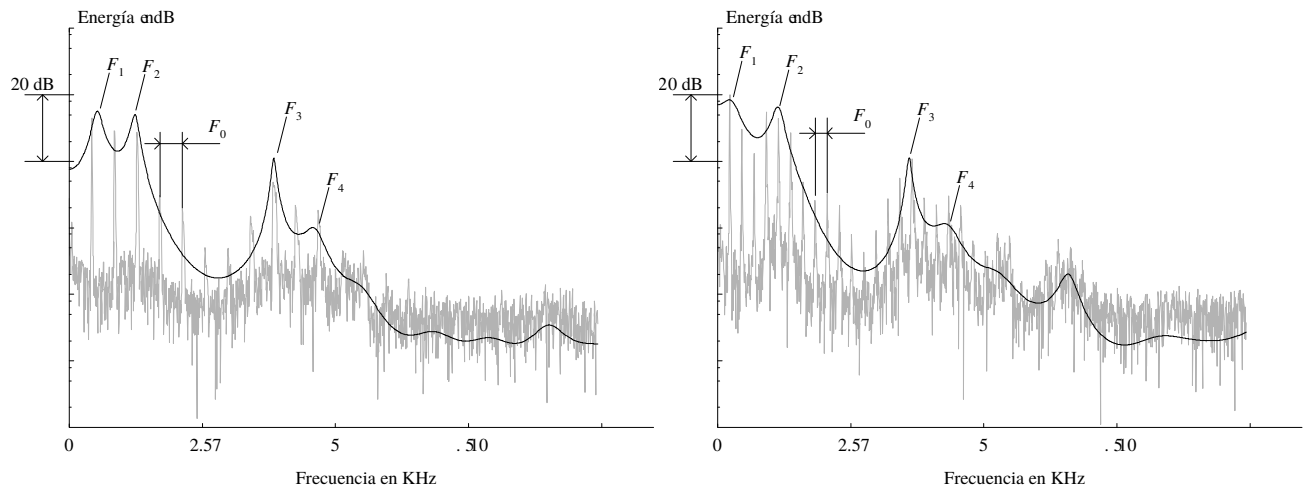
En el error de predicción final, a partir de cierto orden su amplitud relativa se mantiene constante a diferencia del método de Akaike el cual posee un mínimo en el orden óptimo. Akaike permite detectar más fácilmente el orden óptimo.

3. Procesamiento de la Voz

3.1. Describa y clasifique desde el punto de vista fenomenológico a la señal producida a nivel de las cuerdas vocales. ¿Qué características posee su espectro? Grafique.

Fenomenológicamente la señal es estacionaria por tramos y cuasi-periódica, ventanas de 10 a 30 milisegundos.

Su espectro es un tren de pulsos espaciados F_0 que es la frecuencia fundamental la cual posee información del hablante y de la entonación.



3.2. La figura representa el modelo lineal de producción de la señal de VOZ:

3.2.1. ¿Qué tipos de señales de excitación se utilizan? ¿Qué parámetros definen cada tipo de excitación? ¿Cómo los estimarías?

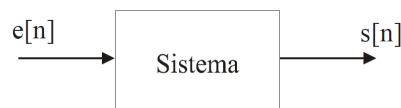
La señal de excitación corresponde al impulso glótico del tracto vocal. Éste posee una frecuencia F_0 denominada frecuencia fundamental.

El parámetro que define ésta excitación es la frecuencia fundamental F_0 .

Se puede estimar mediante:

1. Analizando los cuasi-períodos de la señal en el dominio temporal (restando dos picos y viendo la frecuencia de muestreo).
2. *Autocorrelación sin sesgo*: se calculan los coeficientes de autocorrelación $y[k] = \sum_n x[n]x[n-k]$.
3. Analizando el cepstrum real de la señal y observando el tiempo donde aparece el segundo pico.
4. *Clipping*: eliminar información de la parte central de la señal que no es necesaria para hallar F_0 .
5. *Autocorrelación pesada*: se hace *autocorrelación sin sesgo* y *clipping*, suavizar la estimación del tiempo para cada cuasi-período según $AP[\tau, q]$

3.2.2. ¿Qué modela el bloque denominado *sistema* y de que tipo es? ¿Cuántos parámetros se necesitan definirlo? ¿Cómo se estiman estos parámetros?



El sistema es la función de transferencia del tracto vocal.

Se define con las frecuencias formantes que pueden variar con el tiempo para producir los diferentes fonemas.

Éstas se pueden estimar hallando los máximos de la función de transferencia que se hubican después del primer máximo. La función de transferencia se halla calculando el cepstrum $IFFT(\log |IFFT(x[n])|)$ y tomando los 13 primeros coeficientes.

3.3. Proponga un método para calcular la frecuencia fundamental (F_0) de un tramo de señal de voz.

1. Analizando los cuasi-períodos de la señal en el dominio temporal (restando dos picos y viendo la frecuencia de muestreo).
2. *Autocorrelación sin sesgo*: se calculan los coeficientes de autocorrelación $y[k] = \sum_n x[n]x[n-k]$.
3. Analizando el cepstrum real de la señal y observando el tiempo donde aparece el segundo pico.
4. *Clipping*: eliminar información de la parte central de la señal que no es necesaria para hallar F_0 .
5. *Autocorrelación pesada*: se hace *autocorrelación sin sesgo* y *clipping*, suavizar la estimación del tiempo para cada cuasi-período según $AP[\tau, q]$

3.4. ¿A que se denominan *frecuencias formantes*? ¿Cómo se denotan y que importancia tienen en la generación de la voz?

Las *frecuencias formantes* son las frecuencias en las que se encuentran máximos locales en el dominio frecuencial de la respuesta al impulso del tracto vocal. El *primer* (F_1) y el *segundo* (F_2) pico son frecuencias que permiten identificar el *fonema sonoro* y los *demás picos* (F_3, F_4) permiten identificar la identidad del hablante.

Son picos de mayor intensidad en el espectro de sonido. Se producen por la resonancia del tracto vocal que modifica la amplitud de los armónicos del sonido. Frecuencias amplificadas por resonancia.

3.5. ¿Qué es la escala de mel y cómo se utiliza para el análisis de voz?

La *escala de mel* es una escala que relaciona la frecuencia percibida por el oído humano y la frecuencia física. Es decir, es una unidad de medida para la frecuencia percibida. Es no lineal ya que el oído es más sensible a los cambios de las frecuencias bajas que de las altas. Se determinó que $1000[mels] = 1000[Hz]$.

$$F_{mel} = 1000 \log_2 \left(1 + \frac{f_{Hz}}{1000} \right)$$

En el análisis de la voz, lo que se utiliza es un banco de filtros triangulares linealmente espaciados en la *escala de mel*, que se le aplican a la señal en el dominio de las frecuencias. Con esto se obtiene un vector de coeficientes de *mel*. Una vez obtenidos se les aplica la transformada de coseno (más rápida que Fourier) obteniendo los coeficientes ceptrales en *escala de mel* (CCEM) los cuales tienen información referida a la envolvente de los coeficientes de mel.

3.6. Describa las hipótesis en las que se basa el análisis cepstral del habla para separar la señal de excitación, de la respuesta del tracto vocal ¿Son realistas estas hipótesis?

La señal de excitación del tracto vocal y la respuesta del tracto vocal son independientes y varían de forma diferente, por lo que se pueden separar en el cepstrum.

La duración de la respuesta al impulso del sistema termina antes que el segundo pico del pulso de la señal de excitación.

la hipótesis inicial para poder realizar un análisis cepstral es que las componentes a evaluar tienen que distinguirse mucho en frecuencia, es decir, una tener frecuencia muy alta y la otra muy baja. Por eso funciona perfectamente para el análisis del habla.

La segunda hipótesis es que un ciclo de una las componentes tiene que terminar cuando empieza la otra, no se tienen que superponer.

3.7. Considere la gráfica de una vocal sostenida $x[n]$ que se muestra en la Figura 1.(a). Indique, dibujando sobre la misma, un Período Fundamental (T_0) y estime (aproximadamente) su duración. La frecuencia de muestreo es de 8 kHz:

El período es de $60 \left[\frac{\text{muestras}}{\text{seg}} \right]$ (observación).

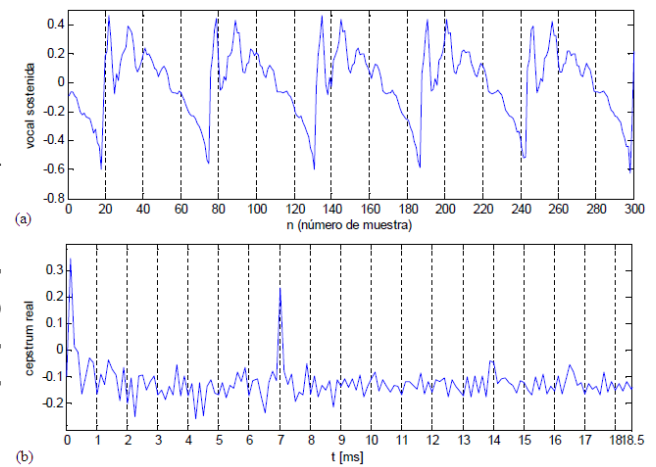
$$T_0 = \frac{60[\text{muestras}]}{8000[\text{Hz}]} = 7,5 \times 10^{-3}[\text{seg}]$$

3.7.1. Defina el *cepstrum real* $c[n]$.

$$\begin{aligned} c[n] &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \log |x(e^{j\omega})| e^{j\omega} d\omega \\ &= \text{IDFT}(\log |DTFT(x[n])|) \end{aligned}$$

Tiene que ser complejo para poder reconstruir la señal.

3.7.2. A partir de la Figura 1.(b), correspondiente al *cepstrum real* del segmento de vocal mostrado en la Figura 1.(a), indique cómo estimar la Frecuencia Fundamental (F_0) del segmento y compare con el resultado obtenido en el punto anterior.



$$F_0 = \frac{1}{T_0} = \frac{1}{7 \times 10^{-3}[\text{seg}]} = 142,85[\text{Hz}]$$

3.8. Proponga un método para contar la cantidad de cruces por cero de un tramo de señal de voz. ¿Cómo haría para que el método fuera robusto al ruido?

La cantidad de cruces por cero nos da una estimación de la frecuencia de la señal. Se puede obtener verificando la cantidad de veces que la misma cambia de signo.

$$CPC = \frac{1}{2} \sum_{n=1}^{N-1} (\text{signo}(x[n]) - \text{signo}(x[n+1]))$$

Es robusto al ruido ya que éste se le suma a la señal y no influye significativamente en la *cantidad de cruces por cero*. Lo que puede ocurrir es que aumente un poco el número de los mismos cuando por ejemplo la señal pasa de positivo a negativo y cerca de ese instante el ruido oblique a la misma a subir y volver a bajar. Lo mismo puede ocurrir en el caso contrario pero a pesar de esto sigue siendo *un buen estimador de la frecuencia de la señal*.

3.9. Indique las consideraciones especiales que deben tenerse en cuenta al aplicar el método de predicción lineal a señales de voz. ¿En qué casos aplicaría el método adaptativo de Widrow?

El método de predicción lineal supone:

1. Un modelo AR (todos polos).
2. Un sistema lineal.
3. Una señal estacionaria o cuasi-estacionaria.

Lo aplicaría cuando se quiere analizar una señal no estacionaria. En Widrow, el *error cuadrático medio* (ECM) se aproxima mediante el *error cuadrático instantáneo* (ECI):

$$\nabla \xi_n^2 = \frac{\partial e_n}{\partial a_n} = \frac{\partial}{\partial a_n} (s_n + \mathbf{a}_n^T \mathbf{s}_n)$$

Remplazando en a_{n+1} .

$$a_{n+1} = a_n - 2\mu s_n \hat{s}_n \quad a_n \in [-0,5, 0,5]$$

μ determina la velocidad de convergencia.

4. Análisis Tiempo-Frecuencia

4.1. Explique la necesidad de utilizar transformaciones tiempo-frecuencia.

Estas transformaciones son necesarias ya que la TDF no otorga información sobre los tiempos en los que ocurren las distintas frecuencias de la señal.

4.2. Defina la *Transformada de Fourier de Corta Duración* (STFT), en sus versiones continua y discreta.

■ Continua

Dada una ventana real simétrica $g(t) = g(-t)$, se genera una base al desplazarla y modularla con una frecuencia ξ .

$$g_{u,\xi}(t) = e^{i\xi t} g(t - u) \quad \text{Átomo tiempo-frecuencia}$$

Si $\|g_{u,\xi}(t)\| = 1$ la STFT (*Transformada de Fourier de Tiempo Corto*) es

$$Sf(u, \xi) = \int_{-\infty}^{\infty} f(t) g_{u,\xi}^*(t) dt = \int_{-\infty}^{\infty} f(t) g(t - u) e^{-i\xi t} dt$$

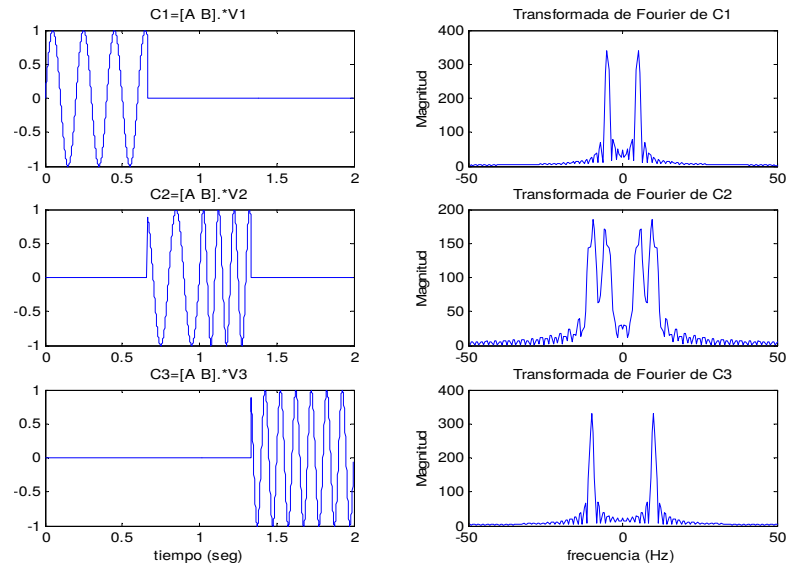
■ Discreta

$$g_{m,l}[n] = g[n - m] e^{\frac{-i2\pi ln}{N}} \quad \text{Átomo tiempo-frecuencia}$$

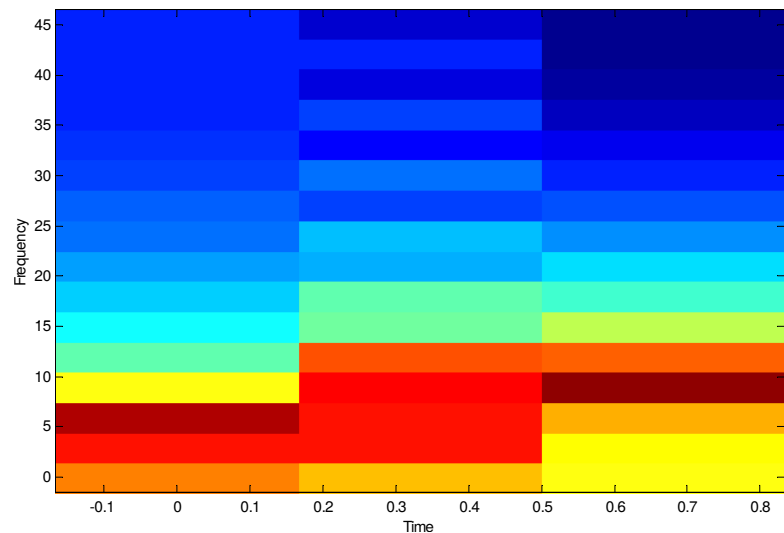
$$Sf[m, l] = \langle f[n], g_{m,l}[n] \rangle = \sum_{n=0}^{N-1} f[n] g[n - m] e^{\frac{-i2\pi ln}{N}}$$

4.3. Explique al menos dos maneras de interpretar conceptualmente a la STFT. Utilice gráficos detallando las magnitudes y unidades de los ejes.

1. Se puede interpretar como la *Transformada de Fourier* aplicada al producto de una ventana por la función en cierto intervalo de tiempo.



2. Se puede representar mediante un espectrograma en el cual el eje de las abscisas representa el tiempo y el de las ordenadas las frecuencias. En el interior se representa la energía de la transformada para cada instante de tiempo y en cada frecuencia.



4.4. Explique en que consiste el *principio de incertidumbre* y como se expresa analíticamente.

Consiste en que si se incrementa la resolución frecuencial disminuye la resolución temporal. Lo mismo ocurre para el caso contrario.

El principio de la incertidumbre de Heisenberg asegura que la varianza temporal σ_t y la varianza frecuencial σ_ω de una función $f(t) \in L^2\mathbb{R}$ (espacio de Hilbert de dimensión infinita de señales de energía finita) satisfacen la siguiente inecuación:

$$\sigma_t \sigma_\omega \geq \frac{1}{2}$$

En la STFT, la función $g(t)$ permanece igual, sólo se desplaza en el tiempo por lo que la resolución es uniforme en tiempo y frecuencia.

$$\begin{aligned}\Delta t &= \text{Tamaño de ventana} = T_{vent} \\ \Delta f &= \frac{1}{\Delta t} = \frac{1}{T_{vent}} \\ \implies \Delta_t \Delta_f &= 1\end{aligned}$$

4.5. ¿Es posible interpretar la STFT en términos de una *base*? En caso afirmativo ejemplifique graficando algunos elementos de la misma.

La STFT no es una base ya que para serlo debe ser un conjunto de vectores *generador* y *l.i.*

- Si las ventanas no se solapan pierdo información de la señal (no es *generador*).
- Si las ventanas se solapan hay información redundante (no es *l.i.*).

El hecho de que la STFT no sea una base significa que no puedo reconstruir de manera exacta la señal original, ya que tendrá una combinación lineal de infinitas soluciones.

4.6. Describa el algoritmo rápido para el cálculo de la Transformada Ondita Diádica.

A la señal original se la filtra con un filtro *pasa alto* g y un *pasa bajo* h cuyas frecuencias de corte se establecen siempre a la mitad. Luego, con la salida de ambos se realiza un submuestreo (para cada salida de los filtros se descarta la mitad de las muestras). Lo que queda proveniente del filtro *pasa alto* son los coeficientes de detalle de nivel 1.

A lo que salió del filtro *pasa bajo* se le aplica el mismo procedimiento, aplicar g y h , obtener los coeficientes de detalle nivel 2 y repetir hasta llegar a a_j .

Notar que como se está submuestreando, la frecuencia que puede tomar la señal se va reduciendo y los filtros cortan siempre por la mitad.

Cuando se llega a p_i se obtiene el coeficiente de detalle de nivel p y el coeficiente de aproximación de nivel p .

Analíticamente:

$$\begin{aligned}a_{j+1}[p] &= \sum_{n=-\infty}^{\infty} h[n-2p]a_j[n] \\ d_{j+1}[p] &= \sum_{n=-\infty}^{\infty} g[n-2p]a_j[n]\end{aligned}$$

y para la reconstrucción:

$$a_j[p] = \sum_{n=-\infty}^{\infty} h[p-2n]a_{j+1}[n] + \sum_{n=-\infty}^{\infty} g[p-2n]d_{j+1}[n]$$

4.7. Defina la *Distribución de Wigner-Ville* en su versión continua. Resuma ventajas y desventajas.

Las STFT y las Wavelets calculan la correlación de la señal con familias de átomos tiempo-frecuencia, por lo que la resolución tiempo-frecuencia de éstas queda limitada a la resolución de sus átomos.

La distribución de Wigner-Ville presenta la densidad de energía en tiempo-frecuencia calculada mediante la correlación de la señal con ella misma desplazada en tiempo y frecuencia.

La aplicación de esta distribución es imitada debido a la existencia de términos de interferencia.

- *Continua:*

$$P_v f(\mu, \xi) = \int_{-\infty}^{\infty} f\left(\mu + \frac{\tau}{2}\right) f^*\left(\mu - \frac{\tau}{2}\right) e^{-i\xi\tau} d\tau$$

- *Discreta:*

$$P_v f[n, k] = \sum_{p=-N}^{N-1} f\left[n + \frac{p}{2}\right] f^*\left[n - \frac{p}{2}\right] e^{-\frac{i2\pi kp}{N}}$$

$\frac{p}{2}$ son muestras intermedias, es necesario interpolar.

Ventaja: Muy buena resolución frecuencial.

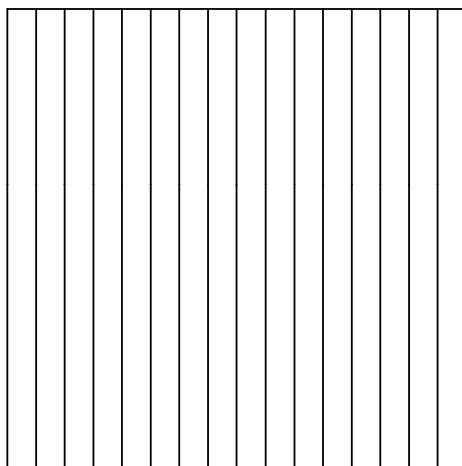
Desventaja: Al ser una transformación cuadrática en f , aparecen términos que provocan interferencia.

$$P_v f = P_v f_1 + P_v f_2 + P_v[f_1, f_2] + P_v[f_2, f_1] \quad \text{términos de interferencia}$$

Tiene interferencias cruzadas entre f_1 y f_2 , y ambas interferencias entre las frecuencias positivas y negativas de una misma f .

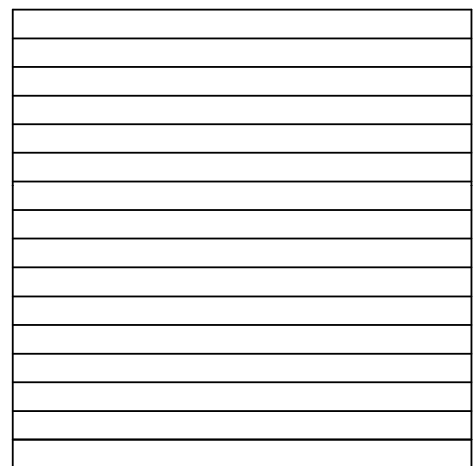
4.8. Esquematice y compare las particiones del plano tiempo-frecuencia que se obtienen mediante una representación que utilice una base de deltas de Dirac, la base de la Transformada Discreta de Fourier, la Transformada de Fourier de Corta Duración y la Transformada Onditas Discreta Diádica.

1. Deltas de Dirac:



Dirac

2. Transformada Discreta de Fourier:



Fourier

3. Transformada de Fourier de Tiempo Corto:

5.2. Problema 2

Un grupo de ornitólogos está interesado en desarrollar una herramienta que les permita clasificar el canto de los pájaros a fin de monitorear en forma automática las poblaciones silvestres y sus variaciones debido a las corrientes migratorias. Se sabe que el canto de los pájaros está formado por sílabas y que éstas constituyen los bloques de construcción elementales, en forma similar a como se da en el habla humana. En muchos tipos de canciones la mayoría de las sílabas pueden aproximarse como uno o varios pulsos sinusoidales breves que varían en amplitud y frecuencia. Sin embargo la variabilidad espectral de las canciones es varios ordenes de magnitud más rápida que en los humanos, requiriendo de una alta resolución temporal, del orden de 1 milisegundo. La energía espectral está típicamente concentrada en una estrecha área entre 1 y 6 kHz, la cual varía para cada especie. Se le solicita diseñar una herramienta que permita extraer la información relevante de archivos de sonido grabados adecuadamente, y detectar la especie a la cual corresponde, de entre las 14 especies en estudio.

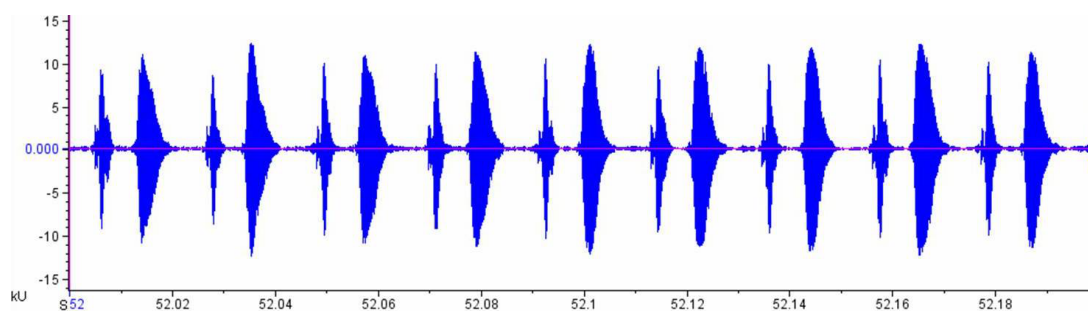


Figura 1: Registro típico de la señal acústica del canto de un pájaro en función del tiempo (mins).

5.3. Problema 3

Un grupo de biólogos marinos está interesado en desarrollar una herramienta que les permita clasificar el canto de las ballenas a fin de monitorear en forma automática las poblaciones salvajes y sus variaciones debido a las corrientes oceánicas. Se sabe que el canto está formado por sílabas y que éstas constituyen los bloques de construcción elementales, en forma similar a como se da en el habla humana. En muchos tipos de canciones la mayoría de las sílabas pueden aproximarse como uno o varios pulsos sinusoidales breves que varían en amplitud y frecuencia. Sin embargo, la variabilidad espectral de las canciones es varios ordenes de magnitud más rápida que en los humanos, requiriendo de una alta resolución temporal, del orden de 1 milisegundo. La energía espectral está típicamente concentrada en una estrecha área entre 1 y 6 kHz, la cual varía para cada especie. Se le solicita diseñar una herramienta que permita extraer la información relevante de archivos de sonido grabados adecuadamente, y detectar la especie a la cual corresponde, de entre las 14 especies en estudio.

5.4. Problema 4

A finales del siglo XX, un grupo de matemáticos israelíes propuso la existencia de un código secreto en la Biblia. El código fue *descubierto* mediante una computadora en la versión hebrea del Antiguo Testamento (Torah), eliminando los espacios entre palabras, y convirtiendo el texto en una única palabra hebrea de 304.805 letras. Un programa exploraba esta larga tira en busca de las palabras y frases que se le ingresaban. El proceso comienza con la primera letra, primero de corrido, luego saltando de a una letra, luego de a dos, y así sucesivamente hasta terminar. Seguidamente, se rehace el mismo proceso comenzando desde la segunda letra, y luego desde las demás hasta terminar. Se le solicita que proponga un algoritmo para el programa que realiza la búsqueda dentro del texto.

5.5. Problema 5

Se desea diseñar un sistema que permita convertir una imagen digital en una señal audible (en el rango de 0-4 kHz), para poderla transmitir por una línea telefónica analógica, de tal manera que posteriormente pueda reconstruirse nuevamente la imagen original. Proponga un algoritmo que permita realizar ambas tareas y realice todas las consideraciones necesarias para su implementación.

5.6. Problema 6

En la sala del consejo directivo de la FICH se han realizado registros de voz mediante micrófonos ubicados frente a cada uno de los consejeros participantes y se quiere rescatar con mayor fidelidad la locución de uno de ellos en particular. Proponga un sistema que permita filtrar esta voz suponiendo que la señal registrada por todos los otros micrófonos puede ser considerada como el ruido de fondo que también se capta en el micrófono de interés.