



UNIVERSIDAD DE BUENOS AIRES
FACULTAD DE INGENIERÍA
Año 2016 - 2^{do} Cuatrimestre

SEÑALES Y SISTEMAS (85.05)

TRABAJO PRÁCTICO ESPECIAL

TEMA: Análisis y procesamiento de la señal de habla

FECHA: 23 de noviembre de 2016

INTEGRANTES:

Manso, Juan

- #96133

<juanmanso@gmail.com>

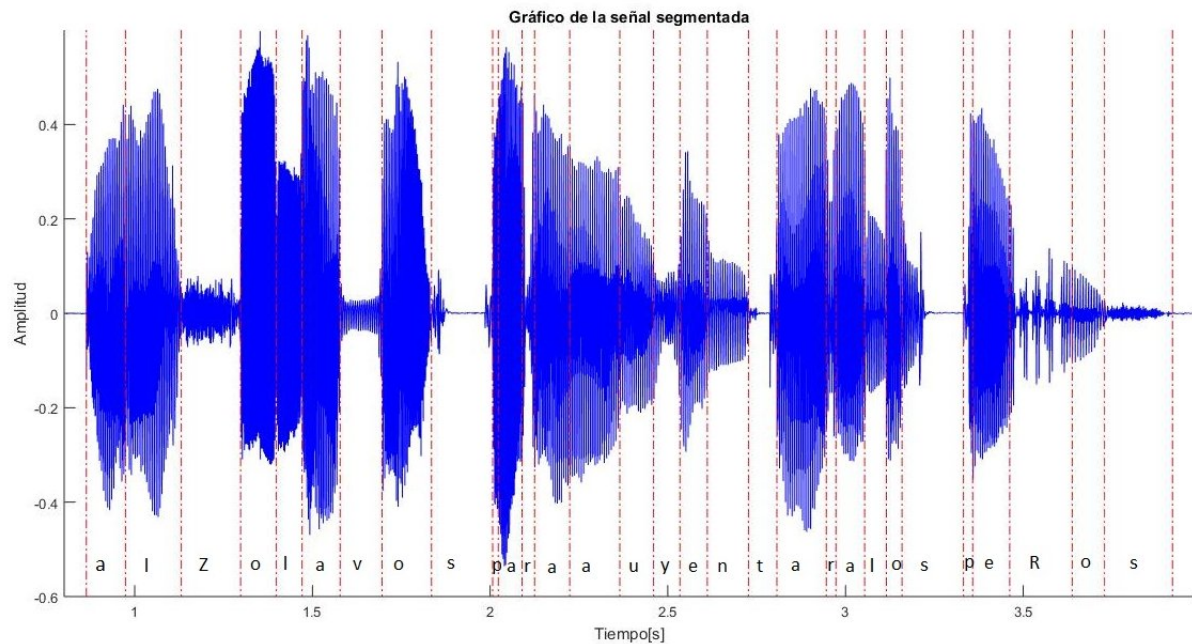
Resumen

El siguiente trabajo práctico tiene como objetivo hacer uso de técnicas y herramientas de análisis de señales y sistemas, aplicándolas al análisis y procesamiento de la señal de habla.

1. Resolución de los ejercicios

1.1. Ejercicio 1

Grafique la señal de voz del archivo `hh2.wav`, ubicando en ella porciones de señales que se o correspondan con fonemas sonoros y sordos. Segmentar y etiquetar en forma aproximada cada uno de los fonemas presentes en la señal.

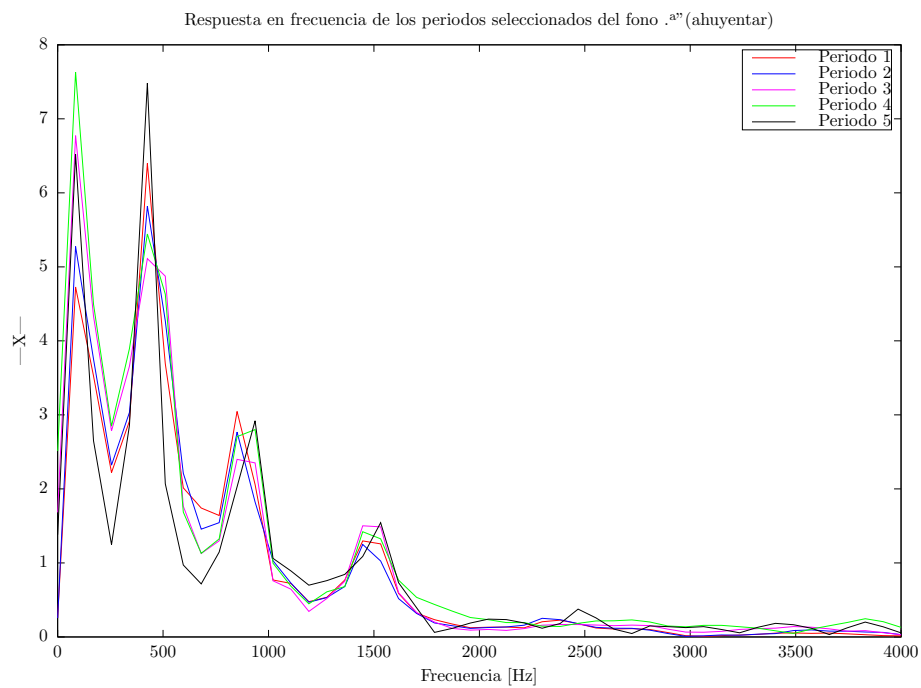
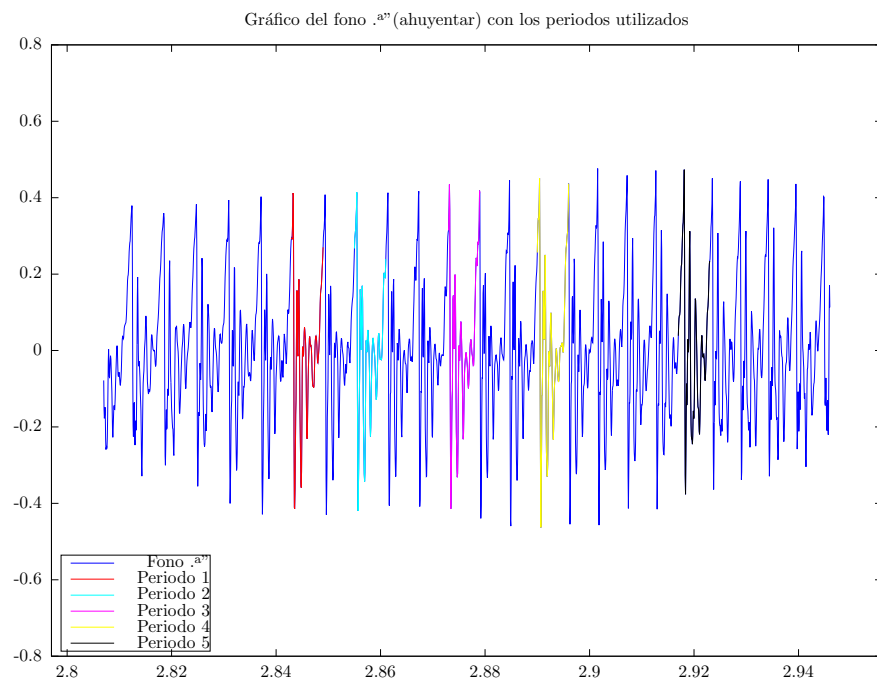


1.2. Ejercicio 2

Con la segmentación realizada en el ejercicio 1 de la señal `hh2.wav`, encuentre los coeficientes de Fourier de un período del segmento de señal correspondiente a un fono [a]. Repetir el cálculo para varios períodos de la vocal.

A continuación se presenta el gráfico del fono [a] seleccionado, que corresponde a la última [a] de la palabra *ahuyentar*. En ella también se presentan los periodos seleccionados para el cálculo de coeficientes.

Con los periodos seleccionados previamente, se obtuvieron los coeficientes de Fourier, representados en el siguiente gráfico



1.3. Ejercicio 3

Reconstruya la señal temporal a partir de los coeficientes calculados. Escuche y compare las distintas reconstrucciones correspondientes a coeficientes de Fourier tomados de distintos períodos. Compárelas también con la señal original. ¿Qué observación se puede hacer sobre la periodicidad de los fonemas vocálicos?

1.4. Ejercicio 4

Grabe la misma frase del ejercicio 1. Mencione las diferencias entre ambas señales.

1.5. Ejercicio 5

Grafique los espectrogramas de banda angosta de los segmentos de señal correspondientes a tres vocales presentes en la señal `hh2.wav`. Compare y analice las diferencias.

1.6. Ejercicio 6

Genere diez ciclos del tren de pulsos glóticos según los modelos de Rosenberg. Tomar una frecuencia $F_0 = 200$ Hz, y fases de apertura y cierre de 40 % y 16 %, respectivamente, de la duración de un pulso. Considerar una amplitud máxima de 1. A los efectos de la simulación, considerar una frecuencia de muestreo de 16 kHz. Estimar su espectro de amplitud y explicar su contenido.

1.7. Ejercicio 7

$$H_n(z) = \frac{1}{(1 - p_n \cdot z^{-1})(1 - p_n^* \cdot z^{-1})} \quad (1)$$

$$p_n = e^{\frac{-2 \cdot \pi \cdot B}{F_s}} \cdot e^{j \cdot \frac{2 \cdot \pi \cdot F_n}{F_s}} \quad (2)$$

Utilizando las ec. 1 y 2, generar un modelo de tracto vocal para cada uno de los siguientes conjuntos de valores de parámetros, que se corresponde con una vocal emitida por una locutora. Grafique diagrama de polos y ceros, y la respuesta en frecuencia del sistema

	F_1	B_1	F_2	B_2	F_3	B_3	F_4	B_4
a	830	110	1400	160	2890	210	3930	230
e	500	80	2000	156	3130	190	4150	220
i	330	70	2765	130	3740	178	4366	200
o	546	97	934	130	2966	185	3930	240
u	382	74	740	150	2760	210	3380	180

para cada vocal, y compare.

1.8. Ejercicio 8

Utilizando los resultados de los dos último ejercicios, sintetice un segundo de las cinco vocales. Escuche y grafique. Haga un análisis en frecuencia, y en tiempo-frecuencia.

1.9. Ejercicio 9

Aplicando el método descripto en la introducción, estime el contorno de la frecuencia fundamental de la voz en el archivo `hh2.wav`. Grafique en forma sincrónica con la onda.

1.10. Ejercicio 10

Con el resultado del ejercicio 9, aplique el método PSOLA para aumentar y disminuir un 10 %, 20 % y 30 % los valores de frecuencia en el contorno de F_0 de la voz en el archivo `hh2.wav`.

Hay que hacer algo automático, un programa que busque los máximos y eso. Hacer lo que dice RODri. No hacerlo si no se llega

1.11. Ejercicio 11

Repita el ejercicio 10 pero aplicando el método para ajustar la velocidad del habla.

No tiene nada que ver con el 10. Hay que cambiar la velocidad. O sea lo que hizo rodri. Podría cambiar la frecuencia de muestreo.

1.12. Ejercicio 12

Repita el ejercicio 8, pero variando la frecuencia fundamental desde 200 Hz a 300 Hz en forma lineal. Escuche la onda resultante, ¿cómo se percibe el cambio en la frecuencia fundamental? Estime el F_0 resultante y compárelo con el teórico.

1.13. Ejercicio 13

Repita el ejercicio 12, pero variando la frecuencia fundamental desde 200 Hz a 100 Hz.

1.14. Ejercicio 14

A la señal de vocales generada en el ejercicio 8 realícele un filtrado con el objetivo de eliminar su frecuencia fundamental. Puede utilizar la herramienta fdatoool para diseñar el filtro. Justifique el filtro implementado. Grafique ambas señales, haga un análisis en frecuencia y compare. ¿Perceptualmente se percibe alguna diferencia? ¿Porqué?

2. Código fuente

tp.m

```
1 % TP de Señales y Sistemas
2 close all;
3 clear all;
4
5 graphics_toolkit('gnuplot');
6
7 % Macros
8
9 F0=200;
10 TP=0.4/F0;
11 TN=0.16/F0;
```

```

12 P0=1;
13
14 bool_print=0;
15 bool_graph=0;
16
17 print_color="-color"; % Si quiere imprimirse color, -color. Si no, -mono
18 format_out_name=".tex"; % Configuro el formato de impresión
19 format_print_config="-depslatex"; % Idem
20 print_path="."; % Ruta de salida
21
22 % Arreglos
23 [y, Fs]= wavread('hh2.wav');
24
25
26
27 %%% Ejercicio 1 %%%
28
29 % Etiquetado de los fonemas
30 % Alzó la voz para ahuyentar a los perros
31 % Formato: fonema_letra=[principio, fin]
32 %# Comentario: n* de muestra = tiempo * Fs;#
33
34 limites_fonemas=[0.864, 0.974, 1.131, 1.298, 1.398, 1.471, 1.578, 1.696,
35 1.835, 2.007, 2.023, 2.090, 2.125, 2.224, 2.365, 2.460, 2.534, 2.611,
36 2.727, 2.807, 2.946, 2.973, 3.054, 3.115, 3.159, 3.332, 3.358, 3.462,
37 3.638, 3.729, 3.920] .*Fs;
38 fonema_string=['a';'l';'Z';'o';'l';'a';'v';'o';'s';'p';'a';'r';'a';'a';'h';
39 'u';'y';'e';'n';'t';'a';'r';'a';'l';'o';'s';'p';'e';'R';'o';'s'];
40
41 % Gráfico de la señal
42 if(bool_graph)
43 figure(1)
44 hold on
45 plot((1:length(y))./Fs,y,'b')
46 i=1;
47 % Gráfico de los límites de los fonemas
48 while(i<=length(limites_fonemas))
49 line([limites_fonemas(i), limites_fonemas(i)]./Fs, get(gca, 'ylim')
50 , 'LineStyle','-.', 'Color','r');
51 if(i!=1)
52 text((limites_fonemas(i-1)+limites_fonemas(i))/2, -0.55,
53 fonema_string(i));
54 end
55 i=i+1;
56 end
57 xlabel('Tiempo[s]')
58 ylabel('Amplitud')
59 title('Gráfico de la señal segmentada')
60 axis([limites_fonemas(1)/Fs-0.1, limites_fonemas(end)/Fs+0.1, -0.8,
61 0.6]);
62
63 if(0)
64 if(bool_print)
65 print([print_path, 'Ejercicio_1', format_out_name], format_print_config,
66 print_color);
67 endif
68 end

```

```

62 end
63
64
65
66 %%% Ejercicio 2 %%%
67 % En este caso, se toma el segmento 'fonema_a6=[2.807, 2.946].*Fs'
68 % que representa la última 'A' de la palabra 'AHUYENTAR'. Se eligió
69 % este segmento dado a que está bien separado de los otros fonemas.
70
71 % Por inspección utilizando el WAVESURFER, se aproxima que el periodo
72 % del fono 'A' está entre los 6 y 7 segundos.
73
74 % Declaración de los 5 periodos a examinar
75 seg_a_1=[2.843*Fs:2.849*Fs];
76 seg_a_2=[2.855*Fs:2.861*Fs];
77 seg_a_3=[2.873*Fs:2.879*Fs];
78 seg_a_4=[2.890*Fs:2.896*Fs];
79 seg_a_5=[2.917*Fs:2.923*Fs];
80
81 if(bool_graph)
82 figure(2)
83     hold on
84     plot((limites_fonemas(20):limites_fonemas(21))./Fs,y((2.807*Fs:2.946*Fs
85     )), '-b');
86     plot(seg_a_1./Fs, y(seg_a_1), '-r');
87     plot(seg_a_2./Fs, y(seg_a_2), '-c');
88     plot(seg_a_3./Fs, y(seg_a_3), '-m');
89     plot(seg_a_4./Fs, y(seg_a_4), '-y');
90     plot(seg_a_5./Fs, y(seg_a_5), '-k');
91     title('Gráfico del fono "a" (ahuyentar) con los periodos utilizados');
92     legend('Fono "a"', 'Periodo 1', 'Periodo 2', 'Periodo 3', 'Periodo 4',
93     'Periodo 5', 'location', 'SouthWest');
94     axis([limites_fonemas(20)./Fs-0.01, limites_fonemas(21)./Fs+0.01]);
95 if(bool_print)
96     print([print_path, 'graf_fono_a', format_out_name], format_print_config,
97     print_color);
98 endif
99
100 dft_a(1,:)=fft(y(seg_a_1));
101 dft_a(2,:)=fft(y(seg_a_2));
102 dft_a(3,:)=fft(y(seg_a_3));
103 dft_a(4,:)=fft(y(seg_a_4));
104 dft_a(5,:)=fft(y(seg_a_5));
105
106 l_dft=length(dft_a(1,:));
107 f_plot=linspace(0,Fs/4,l_dft/2);
108
109 if(bool_graph)
110 figure(3)
111     hold on
112     stem(f_plot,abs(dft_a(1,1:48)), '-r');
113     stem(f_plot,abs(dft_a(2,1:48)), '-b');
114     stem(f_plot,abs(dft_a(3,1:48)), '-m');
115     stem(f_plot,abs(dft_a(4,1:48)), '-g');
116     stem(f_plot,abs(dft_a(5,1:48)), '-k');

```

```

116     title('Respuesta en frecuencia de los periodos seleccionados del fono "
117           a" (ahuyentar)');
118     legend('Periodo 1', 'Periodo 2', 'Periodo 3', 'Periodo 4', 'Periodo 5')
119     ;
120     xlabel('Frecuencia [Hz]');
121     ylabel('|X|');
122     if(bool_print)
123         print([print_path, 'graf_coef', format_out_name], format_print_config,
124               print_color);
125     endif
126     endif
127
128     %%% Inspección %%%
129     % Primer pico = 85.11 Hz
130     % Segundo pico = 425.5 Hz
131     % Tercer pico = 851.1 Hz o 936.2 Hz
132     % Cuarto pico = 1447 Hz o 1532 Hz
133
134     %figure(90)
135     %     plot(abs(dft_a(1,1:48)))
136     %return;
137
138     % Cálculo de amplitudes de los picos máximos
139     %%% Comentario: harmonic(i,j) tiene el i-ésimo armónico
140     %%%                de la j-ésima segmentación.
141     rango_maximos=[1,4; 4,10; 10,15; 15,25];
142     i=1;
143     while(i<5)
144         k=1;
145         while(k<6)
146             harmonic(i,k)= max(dft_a(k,(rango_maximos(i,1):rango_maximos(i,2)))
147                               );
148             k=k+1;
149         end
150         i=i+1;
151     end
152
153     %%% Ejercicio 3 %%%
154     % Reconstrucción de la señal temporal a partir de los coeficientes
155     rec_a=abs(ifft(harmonic));
156
157     % Periodizo la señal
158
159     for i=1:5
160         rec_a=[rec_a rec_a];
161     end
162
163     % Gráfico de las síntesis
164     if(bool_graph)
165         figure(80)
166         plot(1:length(rec_a),rec_a(1,:))
167         figure
168         plot(1:length(rec_a),rec_a(2,:))
169         figure
170         plot(1:length(rec_a),rec_a(3,:))

```



```

170 figure
171 plot(1:length(rec_a),rec_a(4,:))
172 figure
173 plot(1:length(rec_a),sum(rec_a))
174 endif
175
176 % Gráfico utilizando TODOS los coeficientes
177 a=abs(ifft(dft_a(1,:)));
178
179 if(bool_graph)
180 figure
181     hold on
182     plot(1:length(a),a);
183     plot(1:length(y(seg_a_1)),(-1).*y(seg_a_1),'r')
184 endif
185
186
187
188 %%% Ejercicio 5 %%%
189 fonema_a6=[2.807, 2.946].*Fs;
190 fonema_u1=[2.365, 2.460].*Fs;
191 fonema_o1=[1.298, 1.398].*Fs;
192 fonema_e2=[3.358, 3.462].*Fs;
193
194 % Acá defino las ventanas y demás
195 t_ventana = 15e-3;
196 l_ventana = t_ventana * Fs;
197 ventana = hamming(l_ventana);
198 overlap = round(l_ventana/2);
199
200 if(bool_graph)
201 figure(5)
202 spectrogram(y(fonema_a6(1):fonema_a6(2)),ventana,overlap,[],Fs,'yaxis');
203 title('Fono A');
204
205 figure(6)
206 spectrogram(y(fonema_u1(1):fonema_u1(2)),ventana,overlap,[],Fs,'yaxis');
207 title('Fono U');
208
209 figure(7)
210 spectrogram(y(fonema_o1(1):fonema_o1(2)),ventana,overlap,[],Fs,'yaxis');
211 title('Fono O');
212
213 figure(8)
214 spectrogram(y(fonema_e2(1):fonema_e2(2)),ventana,overlap,[],Fs,'yaxis');
215 title('Fono E');
216
217 end
218 %%% Ejercicio 6 %%%
219 t=0:1/Fs:1/F0;
220 pt=P0/2*(1-cos(t./TP.*pi)).*(t<=TP) + (TP<t).*(t<=TP+TN).*P0.*cos((t-TP)./
    TN.*(pi/2));
221
222 % Gráfico del pulso glótico
223 if(bool_graph)
224     figure
225         hold on
226         plot(t,pt)

```

```

227 end
228 % Genero el tren de pulsos de largo 10
229 tren_pt=pt;
230 for i=1:(10-1)
231     tren_pt=[tren_pt pt];
232 endfor
233
234 if(bool_graph)
235     figure
236     plot(1:length(tren_pt),tren_pt);
237 endif
238
239 tren_pt_dft=abs(fft(tren_pt));
240
241 if(bool_graph)
242     figure
243     plot(tren_pt_dft)
244 endif
245
246
247
248 %%% Ejercicio 7 %%%
249 % Cargo la tabla
250 resonancia=[ 830 1400 2890 3930;
251             500 2000 3130 4150;
252             330 2765 3740 4366;
253             546 934 2966 3930;
254             382 740 2760 3380];
255
256 bandwidth=[ 110 160 210 230;
257            80 156 190 220;
258            70 130 178 200;
259            97 130 185 240;
260            74 150 210 180];
261
262
263 % Calculo los polos para cada par (resonancia,bandwidth)
264 p = exp(-2*pi.*bandwidth./Fs) .* exp(1j*2*pi.*resonancia./Fs);
265 conj_p=conj(p);
266
267 z=0:8000;
268 z=2.*pi*z./length(z);
269 H_z=ones(rows(p),1);
270
271 if(0)
272 for i=1:(rows(p)-1)
273     H_z(i)=H_z(i).*1./((1-p(:,i)./z).*(1-conj_p(:,i)./z));
274 endfor
275 end
276
277 %%% No entiendo lo del cálculo de coeficientes PREGUNTAR
278
279 % Gráfico de polos y ceros para cada vocal (vocal=subplot)
280
281 if(!bool_graph)
282     figure
283     subplot(231)      % Vocal='a'
284     hold on

```

```
285     zplane([], [p(1,:) conj_p(1,:)]);
286     title('Polos para la vocal a');
287     subplot(232)
288     hold on
289     zplane([], [p(2,:) conj_p(2,:)]);
290     title('Polos para la vocal e');
291     subplot(233)
292     hold on
293     zplane([], [p(3,:) conj_p(3,:)]);
294     title('Polos para la vocal i');
295     subplot(234)
296     hold on
297     zplane([], [p(4,:) conj_p(4,:)]);
298     title('Polos para la vocal o');
299     subplot(235)
300     hold on
301     zplane([], [p(5,:) conj_p(5,:)]);
302     title('Polos para la vocal u');
303 endif
304
305
306
307 %%% Ejercicio 9 %%%
308
309 %% quefrecencia utilizando cepstrum()
```