

**INSTITUTO POLITÉCNICO NACIONAL
ESCUELA SUPERIOR DE CÓMPUTO**



Teoría de Comunicaciones y Señales.

**Proyecto Final
Detección de vocales.**



Grupo: 3CM5

Alumnos:

Ibarra Soto Alejandro
Mendoza Sánchez Marco Antonio.

Profesora:

Arzate Gordillo Jacqueline.

Frecuencia fundamental de las 5 vocales.

El idioma español tiene 5 sonidos vocálicos bien diferenciados entre ellos: /a/, /e/, /i/, /o/ y /u/. Aunque en algunas variantes el número de fonemas es mayor, en este proyecto sólo se analizan los anteriormente mencionados. En las posteriores tablas se muestran los datos de media y variación típica de los formantes F1 y F2 en las 5 vocales del español para 16 mujeres y 16 hombres respectivamente.

	F1	σ	F2	σ
[i]	240,75	23,2	2841,7	248
[i]	240,58	29,9	2828,25	236,42
/i/	240,75	24	2834,9	241,23
[é]	501,75	41,3	2292	167,5
[e]	481,5	45,7	2214	111,3
/e/	491,6	38,5	2252,08	134,1
[á]	661,5	24,1	1156,4	107,58
[a]	665,9	64,8	1179,16	113,4
/a/	663,75	43,4	1167,8	106,1
[ó]	510,75	70	967,5	149
[o]	510,75	60	994,5	101,9
/o/	510,75	68,5	981	112,8
[ú]	249,75	36,9	630	82,1
[u]	236,25		627,75	
/u/	243	29,5	628,8	63,5

	F1	σ	F2	σ
[i]	268,28	45,3	2342,15	158,22
[i]	260,68	36,4	2294,09	154,6
/i/	264,5	37,8	2317,5	154,3
[é]	449,71	66,6	2052,7	147,6
[e]	454,96	63,9	1935,28	107
/e/	453,8	60,8	1995,01	113,2
[á]	665,68	39,8	1220,4	103,7
[a]	648,84	38,3	1211,59	81,8
/a/	657,28	38,4	1215	87,5
[ó]	475,8	58,4	900,46	175,4
[o]	473,3	61,6	895,18	95,7
/o/	474,5	52	888,4	103
[ú]	291,09	39,4	685,12	57,5
[u]	283,5	23,1	653,06	40,6
/u/	293,5	37,7	669,08	44,2

Desarrollo.

Para la práctica se utilizó la paquetería sounddevice del lenguaje de programación Python 3, este módulo proporciona enlaces para la biblioteca PortAudio y algunas funciones convenientes para reproducir y grabar matrices numpy que contienen señales de audio.

En nuestro caso únicamente se utilizó la función de grabar, de manera que se obtuviera una matriz con el resultado.

```
rec = sd.rec(int(d*fm), fm, 1)
```

Posteriormente se le aplicó la FFT a la matriz, algoritmo visto en la práctica número 3.

```
trans = FFT(rec)
magns = eMag(trans)
```

En este proyecto se utilizó una frecuencia de muestreo de 16,384, para facilitar el proceso de la FFT, durante dos segundos.

Lo siguiente fue obtener el espectro de magnitud de la FFT donde los índices serían la frecuencia (eje X) y los valores por cada índice la magnitud (eje Y).

```
def eMag(afft):
    global fm
    res = np.array([])
    for i in range(fm):
        res = np.append(res,abs(complex(afft[i])))
    return res
```

Muestras iniciales.

Primeramente, se realizaron muestras de la frecuencia fundamental (es decir la frecuencia de magnitud máxima en el espectro de magnitud)

	A					PromF			E					PromF	
Fundamental	245	234	243	123	243				228	246	247	247	248		
F1	246	235	244	124	242	1091	218.2		226	247	246	246	247	1212	242.4
F2	244	233	242	246	245	1210	242		227	245	248	248	249	1217	243.4
F3	247	236	245	122	244	1094	218.8		225	244	245	245	246	1205	241
F4	243	232	247	248	241	1211	242.2		229	242	244	244	245	1204	240.8
F5	240	237	246	245	240	1208	241.6		224	243	249	249	244	1209	241.8
sumas F	1220	1173	1224	985	1212	Suma	Prom		1131	1221	1232	1232	1231	Suma	Prom
Prom	244	234.6	244.8	197	242.4	1162.8	232.56		226.2	244.2	246.4	246.4	246.2	1209.4	241.88

Sin embargo, los datos obtenidos no son los suficientes como para poder obtener un rango útil que permita diferenciar las vocales.

De estas pruebas podemos concluir que **la frecuencia fundamental no es suficiente para poder diferenciar a las vocales entre si**. Sin embargo no es posible descartar aún a las formantes, bajo esta idea se tomaron más muestras de las formantes, primeramente pronunciándolas de golpe para la formante 15 y 10:

Formante	A	E	I	O	U
15	357	375	356	348	372
	355	375	352	350	354
	360	366	322	359	358
	350	360	333	345	352
	350	369	344	356	392
10	315	325	342	336	340
	317	330	336	324	346
	312	344	337	332	334
	322	320	332	320	334
	312	333	343	316	330

Al no lograr obtener tampoco un rango útil se tomaron pruebas pronunciando las vocales en un tono sostenido por 2 segundos, de las formantes 10, 50 y 1000, intentando incluir en las tablas todos los rangos de valores que podía obtener la vocal.

Formante	A	E	I	O	U
10	268	260	275	255	285
	270	246	273	261	292
	268	268	269	254	281
	270	265	270	266	285
	268	261	278	267	312
		262	276	256	280
					296
50	435	415	415	403	422
	421	411	433	435	431
	429	413	408	434	422
	425	428	417	447	429
	417	444	413	428	465
	391	426	429	470	465
			453		463
					461
100	612	626	539	582	630
	589	579	608	568	691
	610	606	650	592	669
	587	693	550	581	646
	602	578	653	606	699
	589	571	665	542	682
		537	562	596	645
				614	654
1000	3325	3412	3405	3300	3439
	3354	3353	3434	3359	3482
	3230	3377	3229	3393	3470
	3346	3411	3382	3340	3365
	3340	3390	3515	3413	3325
	3368	3318	3348	3339	3422
		3543			

Sin embargo el rango en estos valores tampoco es útil, por lo que se volvió a tomar muestras del formante 1000 en vocales pronunciadas de golpe, esto con el objetivo de tener más dígitos y hacer un rango más grande.

Formante	A	E	I	O	U
1000	10887	9114	8379	8440	6312
	10880	10065	8364	8855	6262
	12033	10486	7318	10140	6218
	11163	10050	8017	8588	6459
	12384	9926	7440	8383	5634
	108	91	73	84	56
	11	100	83	88	63
Prom	11469.4	9928.2	7903.6	8881.2	6177

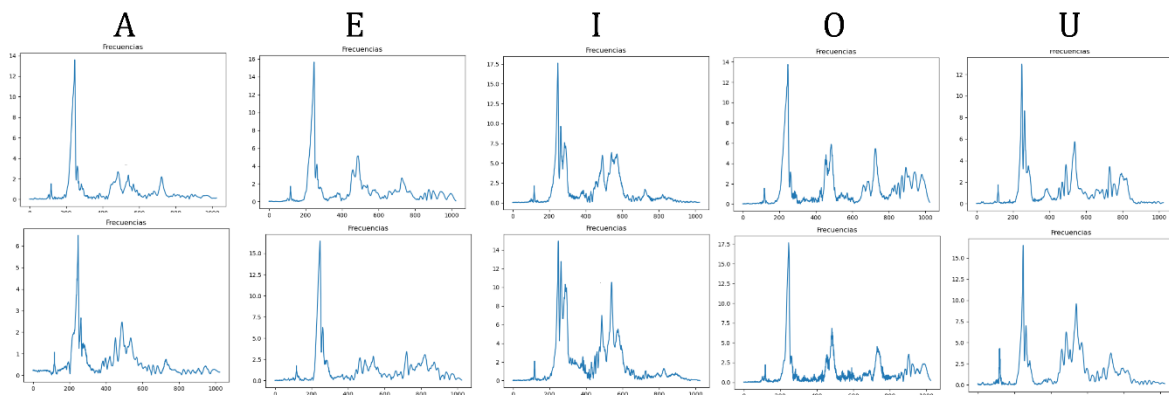
Finalmente se obtuvo un rango útil y se creó el ProyectoSeV1.py, el cual tuvo un desempeño bastante decepcionante.

Por lo tanto, se puede concluir que el **mediante el uso y análisis de las formantes por si solas tampoco es posible diferenciar a las vocales entre si.**

Frecuencias 400-600hz y 600-800hz.

Claramente es necesario analizar un dato extra aparte de las formantes, y el único dato que se posee es la magnitud.

Para analizarla fue necesario utilizar las gráficas del espectro de magnitud de cada letra, pero para esto es indispensable bajar la frecuencia de muestreo de 16384 a 1024 hz, obteniendo los siguientes resultados:



Podemos observar como claramente las señales son bastante diferentes entre las vocales, centrándonos en los rangos de 400 a 600hz y 600 a 800 hz.

Podemos notar que solamente las letras EOU poseen una magnitud significativa en el rango de 600-800, pero de esas 3 la U es la más grande y hay una clara diferencia entre la magnitud de E y O en el rango de 400-600.

Por otro lado, es bastante notable la diferencia entre la magnitud de A e I en el rango de 400 – 600.

Conclusión.

Es posible, mediante el uso de los rangos 400-600hz y 600-800hz diferenciar las vocales de la siguiente manera:

EOU de AI -> A de I

EO de U -> E de O

Se procedió a tomar las muestras de las magnitudes en ese rango de frecuencias.

PAR E O (1024 hz)			
E		O	
6 -- 8	4 -- 6	6 -- 8	4 -- 6
3.91	7.26	6.73	9.32
3.77	10.22	6.15	9.21
3.21	7.91	6.53	8.49
4.4	5.24	5.71	3.96
3.34	4.12	6.21	7.81
2.95	6.21	6.68	6.8
2.38	6.9	5.6	6.51
2.38	4.12	5.6	3.96
4.4	10.22	6.73	9.32

Podemos notar que el modelo de Pares de magnitudes es útil y válido, sin embargo estas pruebas siguen siendo en una frecuencia de 1024 hz, si se aumenta la frecuencia, la magnitud también aumentará, permitiéndonos un rango de magnitudes más grande y más seguro, entonces se realizaron las pruebas de 16384 hz.

PAR A E I O U (16384 hz) Hombres									
A		I		U		E		O	
6 -- 8	4 -- 6	6 -- 8	4 -- 6	6 -- 8	4 -- 6	6 -- 8	4 -- 6	6 -- 8	4 -- 6
8.65	70.06	38.06	83.02	162	210	20.55	119.14	94.13	201.51
13.97	62.95	39.61	94.14	196	229	69.54	175.61	103.41	204.2
16.73	72.25	29.99	96.55	270	181	62.24	120.17	83.82	181.23
11.92	44.72	39.18	53.25	227	255	53.83	136.98	88.5	150.28
19.5	64.39	49.63	105.3	259	259	77.18	122.77	72.45	108.87
14.38	61.8	36.26	87.115	203	233	42.19	127.05	113.42	162.37
24.69	67.62	39.4	54.22	204	287	42.2	86.78	74.52	120.6
8.65	44.72	29.99	53.25	162	181	20.55	86.78	72.45	108.87
24.69	72.25	49.63	105.3	270	287	77.18	175.61	113.42	204.2

PAR A E I O U (16384 hz) Mujeres									
A		I		U		E		O	
6 -- 8	4 -- 6	6 -- 8	4 -- 6	6 -- 8	4 -- 6	6 -- 8	4 -- 6	6 -- 8	4 -- 6
14	69	63	126	200	221	21	143	237	154
21	79	63	146	80	381	38	122	218	73
13	95	46	192	174	270	35	125	320	92
11	88	53	180	284	284	23	132	242	59
42	165	30	138	158	193	21	135	242	63
11	69	30	126	80	193	21	122	218	59
42	165	63	192	284	381	38	143	320	154

En el modelo de hombres, claramente existe la posibilidad de que la letra E entre en el rango de magnitudes que le correspondería a la A e I por lo tanto es necesario agregar las siguientes modificaciones al modelo de hombres.

EOU de AEI -> A de E, E de I

EO de U -> E de O

El modelo de mujeres permaneció sin modificaciones, solo se excluyeron las pruebas en rojo del análisis pues salían completamente del rango.

Resultando en el siguiente código:

Hombres.
<pre>if np.amax(magns[600:800])<25: if np.amax(magns[400:600])<75: letra = 'a' ltrEntry.set("A") else: letra = 'e' ltrEntry.set("E") elif np.amax(magns[600:800])<50: if np.amax(magns[400:600])<105: letra = 'i' ltrEntry.set("I") else: letra = 'e' ltrEntry.set("E") else: if np.amax(magns[400:600])>205: letra = 'u' ltrEntry.set("U") else: if np.amax(magns[600:800])<60: letra = 'e' ltrEntry.set("E") else: letra = 'o' ltrEntry.set("O")</pre>
Mujeres.
<pre>if np.amax(magns[600:800])<40: if np.amax(magns[400:600])<100: letra = 'a' ltrEntry.set("A") else: letra = 'e' ltrEntry.set("E") elif np.amax(magns[600:800])<70: letra = 'i' ltrEntry.set("I") else: if np.amax(magns[400:600])>190: letra = 'u' ltrEntry.set("U") else: letra = 'o' ltrEntry.set("O")</pre>

Pruebas

Posee un comboBox que permite al usuario seleccionar su género, un botón para grabar la voz, otorga la posibilidad de mostrar la gráfica del espectro de magnitud, botones limpiar y salir.

Las 5 vocales

Proyecto Final

Soy un hombre ▾

- Presione el botón escuchar.
- La letra identificada se muestra en area de texto correspondiente.
- Presione el botón espectro de magnitud para desplegar la gráfica.

Escuchar

Fundamental: Hz

Letra: Espectro de magnitud

- Hable mientras el botón " Escuchar " se encuentre en resaltado en blanco.

Limpiar Salir

Mientras el programa está escuchando el botón “escuchar” resalta en blanco, de manera que el usuario puede ver hasta cuando lo escuchará el programa.

Las 5 vocales

Proyecto Final

Soy un hombre ▾

- Presione el botón escuchar.
- La letra identificada se muestra en area de texto correspondiente.
- Presione el botón espectro de magnitud para desplegar la gráfica.

Escuchar

Fundamental: Hz

Letra: Espectro de magnitud

- Hable mientras el botón " Escuchar " se encuentre en resaltado en blanco.

Limpiar Salir

Al finalizar de procesar la señal muestra la frecuencia fundamental y la letra encontrada.

Las 5 vocales

Proyecto Final

Soy un hombre ▾

- Presione el botón escuchar.
- La letra identificada se muestra en area de texto correspondiente.
- Presione el botón espectro de magnitud para desplegar la gráfica.

Escuchar

Fundamental: 248 Hz

Letra: A

Espectro de magnitud

Limpia Salir

- Hable mientras el botón " Escuchar " se encuentre en resaltado en blanco.

El usuario puede observar el espectro de magnitud.

