

Diseño de un sistema biométrico Para procesar señales de voz

Carlos Solano, Jorge Carrillo, Emiliano Martinez, and Leonardo Millan

Abstract—Durante el transcurso de esta investigación acerca de los dispositivos de voz y el cómo estos han evolucionado desde los años 90, cuál será el futuro de esta tecnología; también se tocarán temas más a fondo de los estándares para el uso adecuado de la información que está en el internet, enfatizando también en la controversias que han tenido este tipo de tecnologías con el público general y cuales son regulaciones que tienen esta clase de dispositivos para salvaguardar la integridad de los usuarios. Ya que al usar la voz para interactuar con estos dispositivos se puede haber afectado su seguridad. Por otra parte, se verá acerca de cómo graficar la voz y el audio en general y para que nos pudiera ayudar para identificar y reconocer personas y que componentes se ocuparan para desarrollar este dispositivo. Por otra parte, se investigará más a fondo los sistemas de voz en general y cómo estos han interactuado con la sociedad, para luego desarrollar el código y así poder demostrar cómo se puede identificar y reconocer la voz de una persona mediante señales.

Index Terms—LaPlace, IEEE, Señales, Analisis, Biometricos, Journal.

1 INTRODUCTION

EN los últimos años los dispositivos de voz han ido teniendo una mayor relevancia en nuestra vida cotidiana, por su consecuencia es cada vez se hace más importante el entendimiento de cómo funcionan este tipo de dispositivos y cómo las señales de audio interactúan con este. Por eso se desarrollará un código en el que se buscará poder reconocer la voz de una persona entre un grupo de 4 personas.

Septiembre 6, 2022

1.1 Preguntas Detonantes

1. ¿Cómo ha evolucionado el estado de la tecnología de los dispositivos por voz en los últimos 30 años y cuál es la perspectiva proyectada para estos dispositivos en los próximos 20 años? En 1990 se lanzó Dragon Dictate como el primer software comercial de reconocimiento de voz, el cual cuesta 9,000, alrededor de 18,890 en 2021, donde los usuarios tenían que hacer una pausa entre cada palabra.[2] En 1992, la compañía celular AT&T introdujo el servicio de procesamiento de llamadas a través de reconocimiento de voz de Bell Labs. Ahora conocido como VCRP que procesa alrededor de 1.200 millones de transacciones de voz al año. Igualmente Mike Cohen quien se unió a google en 2004, fue quien descubrió una gran importancia en la innovación de la informática personal, en donde al inicio pudo llevar a cabo la tecnología Google Search (2007) de reconocimiento de voz al público, para poder reciclar los datos de voz de millones de usuarios en la red, entre ellos materiales de aprendizaje automático, para impulsar una mejor calidad . [2]

Mientras tanto apple y microsoft a comienzos de la década de los 2010, gracias a las redes neuronales recurrentes y la memoria a corto y largo plazo, pudieron dar como resultado un salto en este tipo de tecnología, donde ha sido bastante impulsado por el surgimiento y desarrollo del bajo costo computacional y avances algorítmicos masivos.[2]

Ahora si bien sobre la base de la evolución en esta tecnología y su respuesta respecto a las expectativas de los usuarios y la tecnología de reconocimiento de voz, se ha logrado nuevos avances durante esta última media década. Como por ejemplo se crean soluciones para optimizar las diferentes calidades de audio y los exigentes requisitos de hardware, que en consecuencia esto hace que los reconocimientos de voz para la búsqueda de información y dentro de la industria del Internet de las cosas formen parte de la vida cotidiana.[2]

2. ¿Cuáles son las normas y estándares para el adecuado manejo de la información digital que transita a través del internet? Confidencialidad: la información no se pone a disposición ni se revela a individuos, entidades o procesos no autorizados.[4] Integridad: Tener siempre con exactitud y completitud la información y métodos del proceso.[4] Disponibilidad: Acceso y uso de la información y los sistemas de tratamiento de la misma por parte de las personas o procesos cuando lo necesiten.[4]

3. En años recientes ¿Cuáles han sido los aspectos que han generado polémica con relación al uso de dispositivos activados por voz? Bueno, si bien es sabido, una de las polémicas más conocidas en la actualidad y a nivel mundial, es como es que estas compañías, que ofrecen estos servicios que son comandos a través de dispositivos activados mediante la voz, es muy probable y en muchos casos casi seguro, que nos “roben” información, para mejorar su base de datos y así de esta manera, que nuestros dispositivos tengan un mejor funcionamiento, sin embargo esto puede

• Ingeniería en Robotica y Sistemas Digitales, Instituto Tecnológico y de Estudios Superior de Monterrey, Guadalajara, Jal, 45201.

• Alumnos referentes a la universidad

Manuscrito creado Septiembre 8, 2022;

ser contradictorio, si es que se tratan de temas confidenciales u oficiales ante la ley, debido a la privacidad de la persona en el que no se le da su consentimiento para que estos datos sean reproducidos o almacenados.[3]

4. ¿Cuáles serían las regulaciones que debe cumplir el dispositivo físico para su implementación en al menos dos regiones geográficas del mundo? Algunas de las regulaciones que consideramos pertinentes para que dicho dispositivo físico, en la implementación de uso diario como debido a que la voz es un dato que nos identifican como personas, requieren protección jurídica y por lo tanto es importante precisar que la persona tiene derecho a impedir que su voz sea reproducida de cualquier manera si está, no otorga su consentimiento. Esto aplicaría para cualquier región geográfica del mundo, en cuanto a su implementación física, creemos que, como este es una solución realizada mediante software, como tal no debe de cumplir con alguna regulación más que las ya anteriormente mencionadas en relación con la seguridad de la persona.[5]

Si bien cualquier organización que lleve a cabo tratamiento de datos biométricos, debe tener en cuenta lo establecido en la Ley Orgánica de protección de datos y garantía de derechos digitales. Esto en cuanto a España, ahora bien, según INAI(2019)“El responsable o terceros que intervengan en cualquier fase del tratamiento de datos personales deberán guardar confidencialidad respecto de los datos personales, obligación que subsistirá aun después de finalizar su relación con el titular, o en el caso del encargado o de un empleado, con el responsable.”[7]

5. ¿Cómo se podría reconocer e identificar tu huella vocal para interactuar con alguna aplicación como XYZ? Debido a que como obtener una imagen se puede llegar a considerar relativamente sencillo y por lo tanto hasta una huella digital se podría copiar de una imagen, mientras que un molde 3D, de reconocimiento de señales mediante un sensor ya no es tan simple, debido a que requiere un entendimiento y manejo más complejo de éstas y es así como cualquier aplicación podría identificar una huella vocal, a través del análisis de estas señales y comparación en una base de datos previamente guardada.[6]

6. ¿De qué manera se podría usar y procesar las señales de voz para identificar y reconocer personas?

La forma en la que se podría procesar señales para identificar y reconocer personas sería mediante el análisis de las señales de voz. En este análisis, obtendremos los valores de distintos parámetros como lo son: Asimetría, curtosis, desviación estándar, varianza y media. Una vez obtenidos estos datos, pueden ser analizados para encontrar distintos patrones que nos permiten identificar a una persona. Los patrones se pueden usar como condicionales dentro del código para poder identificar a la persona de manera exitosa.

7. ¿Qué elementos y dispositivos electrónicos se tendrían que incorporar para cumplir la tarea de diseñar un sistema que procese la señal de voz?

Micrófono, algoritmo que procese las señales, Computadora

1.2 Investigación

En nuestra actualidad, post pandemia en donde el contacto físico es un poco más distante, está nueva tecnología de voz

está recibiendo un nuevo impulso. Tenemos como ejemplo los sistemas activados por la voz, entre ellos el de Google, Alexa y Siri, que son los principales que generaron un crecimiento en su rendimiento estos últimos años. Esto debido a que, dentro de la pandemia, nos pudimos percatar que además de responder preguntas y realizar compras, también se pueden usar para controlar aparatos electrónicos en los hogares, negocios o establecimientos médicos. “Los sistemas de voz ya han hecho importantes incursiones en los espacios inteligentes del hogar y el control mediante comandos hablados”, detalló el analista Jonathan Collins, de ABI Research. Esto estima que los envíos de aparatos de control por voz para dispositivos inteligentes del hogar, aumentaron sus ventas en un 30% a nivel global, esto es un tipo de incentivo para seguir desarrollando campos de investigación para estas nuevas tecnologías que se presentan. Siendo los asistentes de voz los que más acaparan el mercado de los hogares, estos dispositivos sirven para operar operaciones cotidianas mediante comandos de voz, estos dispositivos en materia de software utilizan inteligencia artificial y “Machine Learning” para reconocer nuestra voz y hacer las actividades que le pedimos de manera más eficiente.

1.3 Desarrollo y Simulación

1.3.1 Experimento 1

Para el desarrollo de la modelación de las gráficas, se utilizó, la herramienta de software Matlab, para así poder obtener la información arrojada por medio de nuestros datos, que en esta situación, es modelado a partir de la información obtenido por cincuenta muestras de audio individuales, para así poder obtener, datos estadísticos puntuales, como lo son la Media, Varianza, la desviación estándar y la curtosis de cada registro y así poder obtener de cada muestra, mediante la Transformada rápida de Fourier el Espectro de amplitud, con respecto a la frecuencia de la señal

```

1 clear
2 Fs=8000;
3
4 recObj = audiorecorder(Fs,16,1);
5 disp('Start speaking. ');
6 recordBlocking(recObj, 5);
7 disp('End of Recording. ');
8
9 %play(recObj);
10 y = getaudiodata(recObj);
11 tamaño_y = length(y);
12 t = 0:(tamaño_y - 1)/Fs;
13 subplot(2,1,1);
14 plot(t,y);
15 Media = mean(y)
16 Varianza = var(y)
17 DesviaciónSTD = std(y)
18 Asimetría = skewness(y)
19 Curtosis = kurtosis(y)
20 Dispersion = (Varianza/DesviaciónSTD)
21 x = fft(y);
22
23 P2 = abs(x/tamaño_y);
24 P1 = P2(1:tamaño_y/2+1);
25 P1(2:end-1) = 2*P1(2:end-1);
26
27

```

Figura 1.Código de MATLAB. Experimento 1.

Resultados Aquí podemos observar los resultados en cuanto al espectro de fourier, la frecuencia de las ondas vocales y su variación, con respecto al tiempo y amplitud

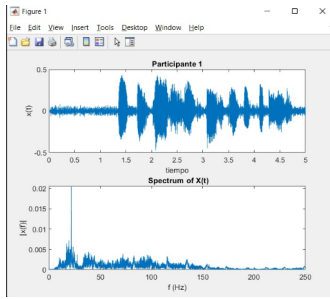


Figura 2. Señal en tiempo y frecuencia.

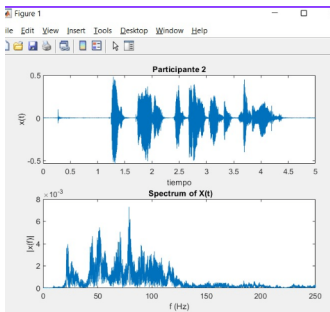


Figura 3. Señal en tiempo y frecuencia.

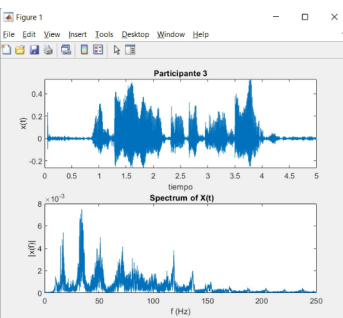


Figura 4. Señal en tiempo y frecuencia.

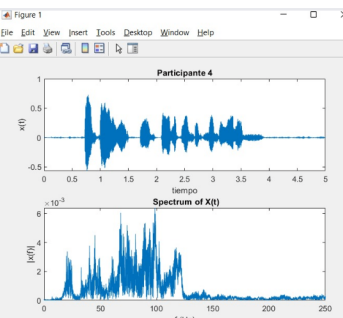


Figura 5. Señal en tiempo y frecuencia.

1.3.2 Experimento 2

El desarrollo de este experimento por otra parte, consiste en obtener datos estadísticos de cada muestra por integrante del equipo para así poder realizar un análisis de cada dato obtenido y que está misma solución, pueda ser implementada en nuestro algoritmo de reconocimiento de voz.

```
1 clear
2 mediaA=[];
3 VarianzaA=[];
4 DesviacionSTDA=[];
5 AsimetriaA=[];
6 CurtosisA=[];
7 DispersionA=[];
8 for i= 1:1
9     Fs=8000;
10    recObj = audioread("leo" + i + ".mp4");
11    y = (recObj);
12    tamano_y= length(y);
13    t= 0:(tamano_y - 1)/5;
14    subplot(2,1,1);
15    plot(t,y);
16    title(' Participante 4 ');
17    xlabel('tiempo');
18    ylabel('x(t)');
19    xlim([0 5]);
20    media=mean(y);
21    mediaA(i)=mean(y);
22    mediaProm=mean(media);
23    mediaA=mediaA';
24    Varianza = var(y);
25    VarianzaA(i) = var(y);
26    VarianzaA=VarianzaA';
```

Figura 5 Código de Matlab Experimento 2.

```
28 DesviacionSTD= std(y);
29 DesviacionSTDA=DesviacionSTDA';
30 AsimetriaA(i) = skewness(y);
31 AsimetriaA=AsimetriaA';
32 CurtosisA (i)= kurtosis(y);
33 CurtosisA=CurtosisA';
34 DispersionA(i) = (Varianza/DesviacionSTD);
35 DispersionA=DispersionA';
36 x = fft(y);
37 P2 = abs(x/tamano_y);
38 P1 = P2(1:tamano_y/2+1);
39 P1(2:end-1) = 2*P1(2:end-1);
40 f = Fs*(0:(tamano_y/2))/tamano_y;
41 subplot(2,1,2);
42 plot(f,P1);
43 xlim([0 250]);
44 title(' Spectrum of X(t)');
45 xlabel('f (Hz)');
46 ylabel('|x(f)|');
47 end
48
```

Figura 5 Código de Matlab Experimento 2.

```
30 f = Fs*(0:(tamano_y/2))/tamano_y;
31 subplot(2,1,2);
32 plot(f,P1);
33
34 title('Single-Sided Amplitude Spectrum of X(t)');
35 xlabel('f (Hz)');
36 ylabel('|P1(f)|');
37
38 end
39 %plot(Fs,abs(x));
40
```

Figura 5.1. Anexo Código de MATLAB. Experimento 2

	media	varianza	desviacion std	asimetria	curtosis	dispersion
audio 1	-3.97348785396533e-05	0.00531060955375509	0.07287392917174163	0.491889648112718	7.45117585727715	0.07387392917174163
audio 2	-0.000101028691453065	0.00438341168592109	0.0662073386107695	0.670925776318944	7.67902785177609	0.0662073386107695
audio 3	-0.000108908989273645	0.00432135413955830	0.0657370073821307	0.4334443205429060	6.00475601820132	0.0657370073821307
audio 4	-0.000143847862354756	0.00308763564153253	0.055664974740403	0.481514708785738	7.35058331327184	0.055664974740403
audio 5	-5.34955171697549e-05	0.00907304891090457	0.0952525335353546	0.272005398982152	6.84606952491234	0.0952525335353546
audio 6	-0.00020798753134148	0.02330736365225	0.15143669846983	0.21175258042840	6.527657841065	0.15143669846983
audio 7	-9.312815177968e-05	0.0276196469340371	0.166191604300690	0.195298991968144	6.02008651189731	0.166191604300690
audio 8	-0.000140356186891587	0.0264459632089303	0.16262148580476	0.265707052649449	6.11404505763103	0.16262148580476
audio 9	-0.000239639148487596	0.0289434820122414	0.170127840203305	-0.0102936600606887	6.50947603922634	0.170127840203305
audio 10	-0.000133858210052629	0.0190731090029266	0.138105427130604	0.155815372756443	6.35086712994845	0.138105427130604
audio 11	-0.000293543567466438	0.0214795267038866	0.146508952997995	0.0470934127462412	6.52733080000445	0.146508952997995
audio 12	-0.00019438800560557	0.0103920153858879	0.10235216777733	0.00464483981368943	6.2791023800224	0.10235216777733
audio 13	-0.000158862611013096	0.0221387581889394	0.148790988265215	-0.0361833382845853	6.06996836056522	0.148790988265215
audio 14	-0.000189063507487714	0.0255270795932529	0.1597719612323654	-0.000785771444456508	6.18218955423748	0.1597719612323654
audio 15	-0.000196003843947957	0.0103063965562904	0.101520424330725	0.200283038267987	6.24142765428192	0.101520424330725
audio 16	-5.34955171697549e-05	0.00907304891090457	0.0952525335353546	0.272005398982152	6.84606952491234	0.0952525335353546
audio 17	-0.000250417960049502	0.0115759100501251	0.107591403235226	0.25730395830434	7.80025440159437	0.107591403235226
audio 18	-5.29772052392525e-05	0.012941315066147	0.113759906971093	0.1933053938181852	7.63352062781606	0.113759906971093
audio 19	-0.000229905030017028	0.010854693183487	0.104185852294583	0.21234208604720	7.55736403131090	0.104185852294583
audio 20	-0.00010947266600371	0.0116609069158843	0.107985679216664	0.165464665448691	6.78307199462281	0.107985679216664
audio 21	-0.000118704808455677	0.0117252668083556	0.108283271138046	0.111465507868223	6.58224025720875	0.108283271138046
audio 22	-0.000214999739649341	0.0286056695996423	0.169132106944963	0.020452258736588	6.35308106302403	0.169132106944963
audio 23	-0.000140794610408187	0.0253484956780492	0.159212102808899	0.0770510364275251	6.33685873113077	0.159212102808899
audio 24	-0.000145327891858002	0.0247339822751650	0.157270411314923	0.077971494841504	6.5995563993160	0.157270411314923
audio 25	-0.00017371178501838	0.0202103298450476	0.142163039658963	0.0062643278610132	7.56329876176231	0.142163039658963

Figura 5. Resultados Estadísticos participante 1.

adurio 26	-0.00029160588183518	0.023220770004720	0.152715667738723	0.088471665121600	6.60427976709079	0.152715673738723
adurio 27	-5.9916902053674e-05	0.0168013942479493	0.126260122884417	-0.00810777090317291	6.75176572112714	0.126260122884417
adurio 28	-0.0002658347015985	0.0200451699437503	0.141580966036224	0.0035805464181411	6.24842836608465	0.141580966036224
adurio 29	-0.000432948734410619	0.0286092490535432	0.169142688424600	0.0817675036388447	6.89874237388447	0.169142688424600
adurio 30	-0.000182926765145869	0.0155582628350340	0.124732765834616	0.271617538430466	6.98661605356435	0.124732765834616
adurio 31	-0.000261213335942856	0.023553985629011	0.153473022047199	0.12453399141319	6.52704543883880	0.153473022047199
adurio 32	-0.000138113025194316	0.0180077278798414	0.134191529377377	0.253390327709936	6.75038132159888	0.134191529377377
adurio 33	-0.000144089984417005	0.013531898590395	0.128576430927443	0.211804370254636	6.60024610311199	0.128576430927443
adurio 34	-0.00018852526239443	0.0131842511700366	0.114822694490404	0.245886201180482	6.52761890615980	0.114822694490404
adurio 35	-0.00028088607478784	0.0141729368157372	0.119050144123127	0.190947398434648	7.71223768795547	0.119050144123127
adurio 36	-5.6418194452923e-05	0.0125309167544925	0.111941577416492	0.15695732151482	6.90001845725246	0.111941577416492
adurio 37	-0.00017861217194548	0.0207178171092097	0.14397158481119	0.09991825327112	6.61032766129793	0.14397158481119
adurio 38	-0.00019506055687348	0.0145021565052752	0.120424937212254	0.105402415136852	6.6303137461550	0.120424937212254
adurio 39	-2.0281058819281e-05	0.0197096471306838	0.140391050749981	0.154953167429847	6.1664171897506	0.140391050749981
adurio 40	-0.000107783547897344	0.0195029921567917	0.13963113666655	0.19757995454413	6.08080915641453	0.13963113666655
adurio 41	-0.0001717121613728	0.01478352389505	0.119072898841804	0.291275853811945	6.9200604816362	0.119072898841804
adurio 42	-0.000182521680665200	0.017193172126326	0.131123499521724	0.1439163462536165	6.96997813010295	0.131123499521724
adurio 43	-9.7927327821306e-05	0.0242865711902957	0.155841493801541	0.191028097575240	6.05062023153153	0.155841493801541
adurio 44	-0.000182926765145869	0.0155582628350340	0.124732765834616	0.271617538430466	6.98661605356435	0.124732765834616
adurio 45	-0.00025794202889433	0.0316115748385521	0.177796442142558	0.205293614235739	6.56814062945868	0.177796442142558
adurio 46	-6.5712696089570e-05	0.0298846464857101	0.177827631858574	0.202967416725870	6.4609218896063	0.177827631858574
adurio 47	-0.000127055528449282	0.022085687663131	0.158727065470211	0.21841149102511	6.80155478996068	0.158727065470211
adurio 48	-0.00040866570199893	0.0290260696773988	0.170371974448260	0.0884396823990704	7.1913097309294	0.170371974448260
adurio 49	-6.9180661704274e-05	0.0224928193020880	0.149796062430269	0.14801701584121	6.87501829815818	0.149796062430269
adurio 50	-0.00024624630161639	0.0161142075769376	0.126941748754842	0.301250824173321	6.27202691888895	0.126941748754842
promedios	-2.4625e-04	0.01838162732733	0.13232897722167	0.182343091920537	6.87234913678806	0.13232897722167

Figura 5.1 Resultados Estadísticos participante 1.

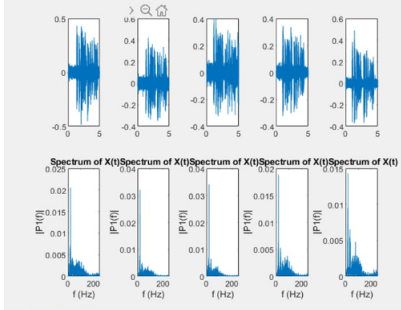


Figura 5.2 Resultados Estadísticos participante 1.

	media	varianza	desviacion std	asimetria	curtosis	dispersion
1						
2	audio 1	-0.00001393	0.073	0.4943	7.4719	0.073
3	adurio 2	-8.78E-05	0.0044	0.0663	7.6785	0.0663
4	adurio 3	-1.48E-04	0.0043	0.0658	6.0008	0.0658
5	audio 4	-1.44E-04	0.0031	0.0555	4.8033	0.0555
6	adurio 5	-1.72E-04	0.0038	0.062	5.5935	0.062
7	audio 6	-3.62E-05	0.0087	0.0934	-3.887	11.4853
8	adurio 7	-6.27E-05	0.0105	0.1025	-0.4592	11.2421
9	adurio 8	-3.51E-05	0.0104	0.102	-0.3351	12.8154
10	audio 9	-1.13E-05	0.0104	0.1019	-0.3487	11.549
11	adurio 10	-1.40E-05	0.0096	0.0979	-0.4671	14.5929
12	audio 11	-5.65E-05	0.0106	0.1028	-0.4989	13.1025
13	adurio 12	-1.05E-05	0.0109	0.1045	-0.5517	11.9285
14	adurio 13	-3.57E-05	0.0104	0.1018	-0.4525	14.0254
15	adurio 14	-2.37E-05	0.0092	0.096	-0.4628	14.0254
16	adurio 15	-2.62E-05	0.0088	0.0937	-0.5023	13.1429
17	adurio 16	-1.56E-05	0.0122	0.1103	-0.5491	13.658
18	adurio 17	-8.53E-06	0.0108	0.1039	-0.4854	12.097
19	adurio 18	-5.19E-05	0.0126	0.1124	-0.5036	10.957
20	adurio 19	-2.62E-05	0.0088	0.0937	-0.5023	13.1429
21	adurio 20	-2.87E-05	0.0057	0.0754	-0.3379	12.3167
22	adurio 21	-2.37E-05	0.006	0.0773	-0.422	10.4125
23	adurio 22	-4.17E-06	0.0067	0.0817	-0.4678	13.0354
24	adurio 23	-2.30E-06	0.0066	0.081	-0.5057	12.6847
25	adurio 24	-2.02E-05	0.0072	0.0846	-0.5526	12.3679
26	adurio 25	-3.25E-05	0.0053	0.0725	-0.5403	14.9096
27	adurio 26	-4.09E-06	0.0075	0.0868	-0.4168	10.8

Figura 5.3 Resultados Estadísticos participante 2.

29	adurio 28	-3.17E-05	0.0062	0.0789	-0.2987	10.9912	0.0789
30	adurio 29	-2.07E-05	0.0057	0.0757	-0.3759	12.1402	0.0757
31	adurio 30	-4.47E-05	0.0041	0.064	-0.4618	11.5561	0.064
32	adurio 31	-2.35E-06	0.0044	0.0666	-0.395	10.8809	0.0666
33	adurio 32	-4.11E-05	0.0056	0.0746	-0.4447	11.6186	0.0746
34	adurio 33	-1.62E-05	0.0046	0.0679	-0.5607	13.2368	0.0679
35	adurio 34	-2.60E-05	0.0052	0.0723	-0.441	12.3154	0.0723
36	adurio 35	-2.20E-05	0.0073	0.0852	-0.5361	11.0961	0.0852
37	adurio 36	-3.39E-05	0.0051	0.0715	-0.4951	11.7777	0.0715
38	adurio 37	-3.16E-05	0.0048	0.0695	-0.4541	11.172	0.0695
39	adurio 38	-5.78E-05	0.0044	0.0664	-0.4506	14.0339	0.0664
40	adurio 39	-3.80E-05	0.0054	0.0737	-0.4003	10.8267	0.0737
41	adurio 40	-5.48E-05	0.0053	0.0726	-0.3633	10.5766	0.0726
42	adurio 41	-4.78E-06	0.0044	0.0667	-0.445	11.5258	0.0667
43	adurio 42	-2.52E-05	0.0042	0.0657	-0.5222	14.2452	0.0647
44	adurio 43	-2.88E-05	0.0043	0.0667	-0.324	14.4189	0.0657
45	adurio 44	-3.10E-05	0.0044	0.0865	-0.3817	12.921	0.0667
46	adurio 45	-2.52E-05	0.0075	0.0647	-0.366	12.7073	0.0865
47	adurio 46	-6.22E-05	0.0054	0.0737	-0.5436	13.9796	0.0737
48	adurio 47	-1.98E-05	0.0049	0.0703	-0.4423	10.5724	0.0703
49	adurio 48	-3.39E-05	0.0051	0.0715	-0.4951	11.7777	0.0715
50	adurio 49	-3.87E-05	0.0047	0.0685	-0.4071	15.0462	0.0685
51	adurio 50	-3.87E-05	0.0045	0.0674	-0.54	13.2689	0.0674
52	promedio	-3.65329E-05	0.006682	0.08043	-0.352758	11.905452	0.08043

Figura 5.4 Resultados Estadísticos participante 2.

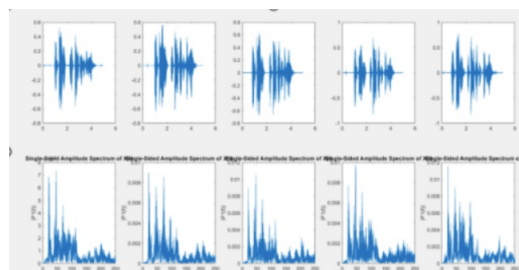


Figura 5.5 Resultados Estadísticos participante 2.

	Media	Var	Des STD	Asimetria	Curvosis	Dependencia
A1	2.42E-04	0.0007	0.027	-0.031	3.41E-04	0.0007477461
A2	0.00144410	0.00422127132	0.0640747461	2.00379939	16.11726	0.0484747461
A3	4.00E-05	0.00477206749	0.0080484766	1.64677793	16.1748093	0.0080484766
A4	1.02E-04	0.0038631621	0.0039205065	1.94605842	15.774427	0.0039205065
A5	6.40E-05	0.00364632670	0.0037075363	1.82684642	13.547642	0.0037075363
A6	3.40E-05	0.000002679	6.30E-05	0.0339	7.7722	0.0002
A7	2.77E-05	0.000002378	0.0016	-0.4227	7.7391	0.0002
A8	1.90E-05	0.000000432	0.0002	-0.482	6.6009	0.0002
A9	8.01E-07	0.000004627	0.0007	-0.343	7.6652	0.0002
A10	-1.90E-01	0.00000569	0.0002	-0.3914	6.1415	0.0002
A11	2.17E-07	0.000000332	0.0002	-0.1473	5.9869	0.0004
A12	7.75E-07	0.000000807	0.0004	-0.5779	7.5551	0.0003
A13	1.68E-07	0.000001624	0.0007	-0.3439	5.3353	0.0004
A14	0.000000086	0.0002	-0.272	5.9756	0.0007	
A15	-3.10E-07	0.000007965	0.0002	-0.5391	4.9376	0.0004
A16	-1.70E-08	0.000000373	6.30E-05	-0.3451	6.7584	0.0004
A17	1.90E-08	0.000000373	6.30E-05	-0.3451	6.7584	0.0004
A18	1.90E-08	0.000000373	6.30E-05	-0.3451	6.7584	0.0004
A19	0.00000128	0.00000448	0.0007	-0.343	6.9652	0.0004
A20	0.000000319	0.000000448	0.0002	-0.4911	7.817	0.0002
A21	-0.000000208	0.000000777	0.0002	-0.1473	5.9869	0.0004
A22	0.000001422	0.00000148	6.30E-05	-0.119	6.6009	0.0004
A23	-0.00000343	0.000000368	6.30E-05	-0.5652	7.7027	0.0007
A24	-0.00000343	0.000000475	6.30E-05	-0.1445	7.4888	0.0004
A25	6.00000128	0.000000448	0.0002	-0.143	6.2259	0.0007
A26	0.0000001628	0.000001786	0.0007	-0.4657	5.6529	0.0004
A27	-0.000001584	0.000000208	0.0002	-0.8044	6.9009	0.0005
A28	-0.000000055	0.000000272	0.0007	-0.5656	6.1814	0.0002
A29	0.000001195	0.000000332	0.0004	-0.5174	6.0751	0.0004
A30	0.0000000771	0.000000313	0.0019	-1.186	11.4769	0.0003
A31	0.000000708	0.000000448	0.0019	0.0096	10.1424	0.0002
A32	0.000000292	0.000000014	0.0001	-0.8352	6.6649	0.0004
A33	0.000001998	0.000000079	0.0002	-0.1466	7.1888	0.0002
A34	0.00000182	0.000000403	0.0001	-1.2752	10.8485	0.0004
A35	-0.000000808	0.000000768	6.30E-05	-1.1919	6.6111	0.0002
A36	-0.000004008	0.000000427	0.0004	-1.0342	8.762	0.0004
A37	-0.00000188	0.000000768	6.30E-05	-0.119	6.1824	0.0003

Figura 5.6 Resultados Estadísticos participante 3.

36	A35	-0.000000808	0.0000057
----	-----	--------------	-----------

Figura 5.9 Resultados Estadísticos participante 4.

36	A35	4.54E-05	0.0075	0.076	1.064	9.991	0.0546
37	A36	4.72E-06	0.0065	0.0713	0.991	12.654	0.05621
38	A37	3.81E-05	0.0012	0.0622	1.0018	14.6126	0.078
39	A38	2.38E-06	0.0045	0.0614	0.911	13	0.0678
40	A39	6.46E-05	0.0078	0.0671	1.0065	14.642	0.056
41	A40	1.71E-05	0.0061	0.0715	0.9991	12.661	0.0518
42	A41	6.60E-05	0.0051	0.0691	0.92	11.6542	0.054
43	A42	4.45E-06	0.0034	0.061	1.054	10.56	0.083
44	A43	4.01E-05	0.0091	0.0614	1.015	9.166	0.0891
45	A44	-1.41E-06	0.0016	0.071	1.061	10.562	0.095
46	A45	4.51E-05	0.0045	0.0618	1.054	14.653	0.051
47	A46	4.10E-06	0.0015	0.07	1	13.46	0.056
48	A47	3.02E-05	0.0025	0.0631	1.0067	9.56	0.064
49	A48	5.14E-06	0.0016	0.0656	0.91	12.2	0.055
50	A49	-1.84E-05	0.0075	0.061	1	15.6	0.059
51	A50	2.48E-06	0.0091	0.0709	0.951	11.562	0.054
52	Promedio	1.625E-05	0.00487	0.068018	1.0107992	12.582804	11.4437678

Figura 5.10 Resultados Estadísticos participante 4.

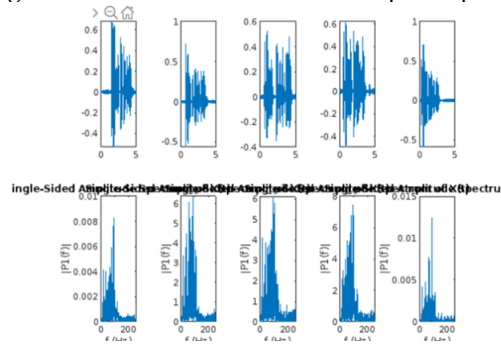


Figura 5.11 Resultados Estadísticos participante 4.

1.3.3 Experimento 3

En el tercer experimento se tomó los datos recolectados del experimento previo y se les hizo un análisis. Se encontró que los datos que más nos ayudaban a identificar personas eran la curtosis y la asimetría. Se usaron estos valores como parámetros para identificar a cada uno de los participantes.

```

if ( curtosisProm>5 && curtosisProm<7)
    fprintf("la voz es de participante 1 ")%JORGE
elseif ( curtosisProm>10 && curtosisProm <13 && asimetriaProm < 0)
    fprintf ("la voz es del participante 2")%CARLOS
elseif(curtosisProm>10 && curtosisProm < 13 && asimetriaProm >0)
    fprintf ("La voz es del participante 4")%LEO
elseif(curtosisProm>15)
    fprintf("voz de participante 3")%EMI
end

```

Figura 6.Código Experimento 3 1.

```

Command Window
>> experimento2_limpio
La voz es del participante 4>> experimento2_limpio
la voz es de participante 1 >> experimento2_limpio
la voz es del participante 2>> experimento2_limpio
fx voz de participante 3>>

```

Figura 6.Resultado Experimento 3 1.

2 CONCLUSIONES INDIVIDUALES

2.1 Conclusión Experimento 1

En esta parte del experimento, podemos observar y analizar una sola voz, con esta podemos encontrar las medias, varianzas desviaciones, curtosis etc, con estas gráficas podemos observar que es una señal aperiódica, discreta. En la gráfica de espectro podemos ver como se tiene una frecuencia más alta y comienza a bajar la intensidad de la señal. En esta parte se pudo apreciar la señal de una voz, la señal que cada persona tiene.

-Carlos Solano.

2.2 Conclusión Experimento 2

En la segunda parte del proyecto se analizaron varias veces la misma frase para obtener un promedio, aquí me di cuenta de que la señal aun siendo grabada en el mismo lugar, sin ruido de fuera la señal de la voz aún puede cambiar, si bien con los datos obtenidos se puede analizar mi voz, aún existe un rango de diferencia entre los datos ya que por distancia, micrófono, o cualquier mínimo cambio al grabar la voz, nos puede cambiar todo nuestro análisis.

-Carlos Solano.

2.3 Conclusión Experimento 3

En esta tercer parte observe como existe una diferencia entre las señales de cada uno de mis compañeros, la diferencia entre los datos que nos puede ayudar a diferenciar a las personas, ya que las gráficas en general se ven muy similares, para poder lograr un buen reconocimiento de voz se necesitan de un buen sistema que analice todos los datos en un instante y los vaya comprobando con una base de datos, esto para poder diferenciar de quien es la voz, la complejidad de la señal de voz crea la existencia de un detector de voz.

-Carlos Solano.

2.4 Conclusión Experimento 1

En el experimento uno desarrollamos un programa en Matlab que nos permitía capturar una señal de voz y poder manipularla. Logramos graficar la señal con respecto al dominio del tiempo en la cual se puede observar que es periódica y discreta. Después obtuvimos diferentes datos usando funciones de Matlab como lo son: Media, varianza, desviación estándar, asimetría, curtosis y dispersión. Después tomamos la señal y mediante el espectro de Fourier la pasamos del dominio del tiempo a la frecuencia, donde gráficamente se podía observar que la señal cambiaba.

-Jorge Carrillo.

2.5 Conclusión Experimento 2

En el segundo experimento repetimos el mismo proceso que en el primero, pero aumentamos el número de muestras. Utilizamos 50 muestras de las cuales obtuvimos todos los datos para poder sacar los promedios y poderlos usar como parámetros a futuro. Usé arreglos para poder guardar todos los datos y después hice un reacomodo para poder extraerlos Excel de manera más sencilla. Saqué los promedios por medio de Matlab ya que era más sencillo que en Excel. En este experimento pude ver varias gráficas de mis muestras y encontrar similitudes entre ellas. Las gráficas de frecuencia son en las que más similitudes encontré.

-Jorge Carrillo.

2.6 Conclusión Experimento 3

En el experimento tres hicimos algo más complejo que en el dos, tomamos los promedios de todos los datos extraídos de las señales y buscamos patrones de cada participante. Se hizo un rango de la curtosis y de la asimetría para poder usarlos como identificadores en el momento de que el programa analizara las señales. Con el uso de estos rangos fue posible identificar a los distintos participantes.

-Jorge Carrillo.

2.7 Conclusión Experimento 1

Para este primer experimento tuvimos que determinar la transformada de Fourier en tiempo discreto para poder graficar su espectro, sin embargo, antes de eso se tuvo que determinar la manera más óptima para poder obtener la señal de audio, que el software en este caso matlab, fuera capaz de almacenar y detectar dicha señal, en conjunto con la función de la transformada rápida de Fourier la cual es una formula proveniente de Laplace se logró el análisis de las muestras para éste primer experimento

-Emiliano Martinez.

2.8 Conclusión Experimento 2

Ahora bien, para continuar con el proceso de determinar un modelo óptimo para el desarrollo de un acercamiento a predicción de voces, fue necesario tomar un análisis de dichas muestras con datos orientados a la estadística, para así poder ver que datos tienen más variabilidad y viabilidad para empezar a desarrollar una propuesta de análisis de datos biométricos, que nos ayudan a identificar lo que caracteriza la voz de cada participante, independiente de que se haya dicho la misma frase.

-Emiliano Martinez.

2.9 Conclusión Experimento 3

Finalmente para este último experimento, en donde pudimos analizar los datos de voz tanto con datos de dispersión, como gráficamente y darnos una idea de cómo funcionaría un reconocimiento de voz automático, es decir con machine learning, esto desde sus bases, que es el cálculo de la transformada de Fourier, es decir con respecto a las muestras y de esta forma determinar la predicción de ciertos datos específico de voces, o bien lo que llamamos datos biométricos, nosotros optamos por considerar dos variables, para mejorar la exactitud de predicción, la primera y la más importante que fue la curtosis y la segunda la asimetría dentro del conjunto de los datos.

-Emiliano Martinez.

2.10 Conclusión Experimento 1

En la primera parte del experimento pudimos ver y observar las medias, varianzas desviaciones, curtosis y dispersión de una voz, ya recabado la información creamos las gráficas de cada punto y con esto pudimos observar cómo se van desarrollando alrededor del tiempo. Ya creada la gráfica podemos ver la señal de una voz y que cada voz tiene una señal diferente.

-Leonardo Millan.

2.11 Conclusión Experimento 2

En la segunda parte del experimento ahora se analizó la voz de la misma persona pero varias veces diciendo la misma palabra, esto nos sirvió para poder sacar un promedio, y como resultado se obtuvo que aunque se grabaron en las mismas condiciones se obtuvieron resultados distintos entre cada audio, por lo que concluí que pese sigue siendo la misma voz hay otros factores que deben ser tomados en cuenta.

-Leonardo Millan.

2.12 Conclusión Experimento 3

Ya en la última parte puede llegar a observar que existe una diferencia entre las señales de mi voz con la de mis compañeros, pero más que nada por los datos recabados ya que a la hora de graficar las señales tienden a ser parecidas, por lo que en mi conclusión si quiere tener la mejor exactitud para diferenciar las voces es que los audios recabados estén lo más libre de ruido y sea lo más parecido cada uno, y tener un número de datos para así poder conseguir un mejor promedio.

-Leonardo Millan.

3 CONCLUSIÓN GENERAL

En este reto se logró entender más a fondo cómo es el funcionamiento de señales y cómo estos nos ayudan para poder entender cómo se comportan las señales de audio. Fue necesario aplicar la transformada de Fourier para poder completar el análisis de las señales. Este reto nos ayudó a ampliar nuestro conocimiento de matlab, ya que aprendimos a hacer código para poder grabar voz y también importar archivos de audio.

4 CONCLUSIÓN GENERAL JORGE

En este proyecto tuve la oportunidad de poder aplicar lo aprendido en clase de manera práctica. Con los experimentos pude analizar las señales de voz recolectadas y pude obtener distintos parámetros de una señal. Analice estos parámetros y encontré patrones que se usaron como parámetros para identificar a las personas. Gracias a la transformada de Fourier, logré pasar mi señal de voz del dominio del tiempo al de frecuencia.

5 CONCLUSION GENERAL CARLOS

A lo largo de este proyecto se aplicaron y aprendieron diferentes métodos de análisis de señales, esto viendo la señal de voz de una persona, buscando el comportamiento de la señal en el tiempo y en la frecuencia, encontrando diferentes puntos de análisis como lo son la media, varianza, dispersión, curtosis, asimetría, etc. Al crear una base de datos con más señales de voz de la misma persona y así mismo de otras personas, podemos encontrar algunas diferencias entre cada señal, esto nos ayuda a poder hacer una pequeña identificación aprovechando esas diferencias entre las señales, si bien es un proceso complejo el cual requiere una base de datos grande y un análisis estadístico más complejo, con pocos datos y pocas voces se logró obtener un reconocimiento de voz.

6 CONCLUSION GENERAL LEO

En conclusión este reto, alcance a conocer más a fondo cómo es que funcionan las señales y cómo estas interactúan con nuestro alrededor, también cómo estas son volátiles ya que con un simple cambio en la prueba podría dar un resultado distinto. También se consiguió entender más a fondo cómo entender más a detalle el funcionamiento de estas y como los modelos matemáticos nos ayudan al entendimiento de estas, tales como la transformada de Laplace y la Transformada de Fourier.

7 CONCLUSION GENERAL EMILIANO

Dentro de esta unidad de formación, pude obtener conocimiento a cerca, de cómo es que funciona un algoritmo de señales, gracias a las matemáticas y como es que el enfoque a los dispositivos de voz ha ido evolucionado a través de la historia, gracias a esto tenemos cada vez nueva y mejor tecnología, donde su base es LaPlace para pasar a la transformada de Fourier y de esta manera poder analizar diferentes tipos de señales, como continuas o discretas, sin embargo en el curso nos enfocamos al análisis de señales discretas y su tratamiento con los datos en un tiempo finito, así como de forma compleja

REFERENCIAS

- 1 Normas ISO sobre gestión de seguridad de la información. (s/f). Mec.es. Recuperado el 5 de septiembre de 2022, de http://descargas.pntic.mec.es/mentor/visitas demoSeguridadInformatica/normas_iso_sobre_gestin_de_seguridad_de_la_informacin.html
- 2 Greame. J(2021).The evolution of speech recognition technology. Recuperado de. <https://www.techradar.com/news/theevolutionofspeech-recognitiontechnology>
- 3 Gonzalez.R(2019).Los errores de privacidad y polémicas que ensombrecen a los asistentes digitales de voz. Recuperado de. <https://rga.cl/2019/08/09/los-errores-de-privacidad-y-polemicas-que-ensombrecen-a-los-asistentes-digitales-de-voz/>
- 4 Contact Center(2021). ¿Es una huella digital más segura que una huella de voz? Recuperado. <https://contactcenterhub.es/es-una-huella-digital-mas-segura-que-una-huella-de-voz-2021-02-32116/>
- 5 CIO México(2018). El futuro llama a la puerta: el uso del reconocimiento de voz para la autenticación. recuperado de <https://cio.com.mx/futuro-llama-puerta-uso-reconocimiento-voz-autenticacion/>
- 6 Bussiness Insider(2020). La tecnología de voz tiene un impulso gracias a la pandemia del coronavirus. Recuperado de <https://businessinsider.mx/tecnologia-de-voz-impulso-gracias-pandemia-coronavirus/>
- 7 INAI.(2018).GUÍA para el Tratamiento de Datos Biométricos.https://home.inai.org.mx/wp-content/documentos/DocumentosSectorPrivado/GuiaDatosBiometricos_Web_Links.pdf