

Treball de recerca

Comparació d'ones

Aniol Juanola Vilalta

Tutor: Josep Kostopoulos

Segon de Batxillerat A

Institut Pere Alsius i Torrent

Besalú - Banyoles, Octubre del 2020

Abstract

The aim of this project is to compare two sound files of the same length.

Due to the limited number of open-source alternatives available online, I modified a free piece of software by the name of «scape-xcorr» so that it returns an image of the spectrum of the files by using the Fourier Transform. Because sound can also be represented as an image, a script which compares the output images by using image processing techniques such as the difference and the conversion of colour spaces has been written.

Both real life and synthetic sound files have been recorded and created to verify the objectivity and the reliability of the tool, which has proven to be effective.

To depict the results, correlation matrices have been used in order to both group the data and visually represent the difference between the other files by using colour scales. Also, an interactive user interface was implemented to interact with the spectrum appearance.

Sinopsis

El objetivo de este proyecto es comparar dos archivos de sonido de la misma duración.

Debido a la escasez de alternativas abiertas disponibles en Internet, modifiqué un programa llamado «scape-xcorr» para que devuelva una imagen del espectro de los archivos utilizando la Transformada de Fourier. Ya que el sonido puede ser representado como una imagen, se ha escrito un script que compara las imágenes generadas anteriormente entre ellas mediante técnicas de procesado de imágenes como la diferencia y la conversión de espacio de color.

Se han grabado y creado tanto archivos de la vida real como sintéticos para verificar la objetividad y la fiabilidad de la herramienta, que ha demostrado ser efectiva.

Para representar los resultados se han utilizado matrices de correlación para agrupar los datos en grupos y mostrar visualmente la diferencia entre los demás archivos utilizando escalas de colores. También se ha implementado una interfaz de usuario para interactuar con la apariencia del espectro.

Índex de continguts

1. Pròleg.....	5
2. Introducció.....	7
3. Part teòrica.....	8
3.1. Senyals.....	8
» Freqüència d'una ona pura.....	8
» Fase d'una ona pura.....	9
» Amplitud d'una ona pura.....	9
3.2. Combinació dels senyals.....	12
3.3. Captura del senyal.....	14
3.3.1. Adquisició.....	14
3.3.2. Tractament digital de les mostres.....	15
» Resolució i profunditat de bits.....	15
» Mostratge.....	16
3.3.3. Teorema de Nyquist.....	16
3.4. Transformada de Fourier.....	17
3.4.1. Visualització de cascada.....	19
3.4.2. FFTW.....	19
3.5. Espai de color.....	19
3.6. Resta d'imatges.....	21
3.7. Matrius de correlació.....	25
3.7.1. Desviació estàndard.....	26
4. Part pràctica.....	27
4.1. Introducció.....	27
4.2. Objectiu.....	27
4.3. Experimentació.....	27
4.3.1. Recerca d'eines existents.....	27
4.3.2. Creació d'arxius sintètics.....	28
4.3.3. Modificació i variacions dels arxius.....	28
4.3.4. Gravació de senyals reals.....	30
4.3.5. Modificació del software per exportar una imatge del FFTW.....	31
4.3.6. Comparació dels resultats.....	35
4.3.7. Organització i anàlisi dels resultats.....	36
» Tractament de les dades.....	37
4.4. Resultats de les diferents proves.....	38
4.5. Aplicacions a la vida real.....	42
5. Conclusions.....	44
Bibliografia.....	46
6. Annex.....	48

1. Pròleg

El treball de recerca és una de les fites més importants del batxillerat. Em considero una persona a la que li agraden els reptes, i més a aquesta escala. Tota la idea i plantejament del treball va començar ja fa temps, quan el meu avi em va regalar un transmissor antic que ell utilitzava per anar a cacera amb els seus companys. Tot i que en aquell moment no vaig saber què fer-ne, sabia que algun dia n'havia de treure profit d'alguna forma.

La idea va sorgir quan vaig pensar que seria interessant veure la potència de l'aparell, ja que funcionava a la perfecció i tenia curiositat per provar-lo i utilitzar-lo. La primera idea que vaig tenir va ser de crear uns arxius d'àudio i enviar-los amb el meu transmissor per rebre'ls des d'un altre lloc remot. Volia comprovar quines dificultats físiques i tècniques tenia, què afectava més a la qualitat del so i veure la diferència amb un transmissor modern. Per aconseguir-ho, m'hauria d'examinar de l'examen de radioaficionat, un títol que dura tota la vida i que mai se sap quan un necessitarà. No obstant, la situació va empitjorar considerablement degut a la pandèmia: la obtenció del títol i totes aquestes proves exteriors que comporten desplaçaments i moviments es van complicar.

Va ser en aquell moment en què vaig començar a fer recerca de les eines que existien en el món del software per a poder comparar els sons, cosa que ja donava per feta i que em pensava que seria el més mínim dels meus problemes. Mai m'hauria imaginat que en aquell moment tot el treball faria un tomb inesperat. Ja que no hi havia quasi cap eina de codi obert disponible, vaig veure que podia aprofitar l'ocasió per aprofundir en el món de la programació, que m'agrada des de petit i que és el que vull seguir estudiant més endavant. Així doncs, el meu nou objectiu va ser crear un software i veure quin resultat podia aconseguir.

Un dels mètodes que em va ajudar més en tot el treball, especialment en la memòria escrita, va ser el mètode Feynman. El mètode és molt simple: escriure un document en què hi ha coneixements llunyans per a moltes persones i fer que es pugui seguir mitjançant explicacions senzilles i concises. Amb molts exemples i frases simples he intentat que sigui un treball que gairebé tothom pugui comprendre i ampliar si ho desitja. Tot i que en la part pràctica s'ha creat una eina que és útil, realment això és el principi ja que es poden seguir fent moltes proves i milers de comparacions per a realitzar nous descobriments i aprofundir en aquest àmbit.

Una limitació que ha tingut el treball és el nivell de matemàtiques que implica. Tot i que per realitzar el treball en sí no es necessita aquest nivell, és veritat que hi ha un rerefons de matemàtiques complicat darrere la Transformada de Fourier que ni quan hagi acabat el curs de batxillerat tindrè. No obstant, això no ha estat cap impediment per a la realització de la part pràctica perquè l'ordinador ho calcula sense haver de tenir-ne un coneixement molt extens.

L'àmbit de la programació trobo que és molt captivant perquè la nostra societat cada vegada és més digital i sense els programadors no existiria i no evolucionaria (també és cert que si les xarxes socials i certs aparells no existissin viuríem en un món amb probablement molta més privacitat i llibertat). Durant el procés de pràctica, els errors i impediments que he anat trobant m'han permès aprendre moltíssim sobre àmbits com la física, la programació i sobre informàtica en general, que ni sabia que existien. M'ha servit per a construir una base pel que vull estudiar a la Universitat, una enginyeria informàtica. És cert que sempre he tingut aquesta vocació i curiositat per aquest tema, ja que des dels cinc anys que remeno i utilitzo ordinadors i aparells semblants. És per aquesta raó per la qual he triat aquest àmbit i aquest tipus de part pràctica, perquè des de petit ja sabia que era el que volia estudiar i forma part de mi.

S'ha mirat de seguir la guia del «moodle» per tal que s'incloguessin tots els continguts que es demanen, però degut a la naturalesa del treball certs apartats han estat unificats per tal de millorar-ne la llegibilitat. Especialment ha estat així a la introducció, on hi ha els objectius i la metodologia junts.

Un últim aspecte a remarcar és que m'hauria agradat afegir més referències a articles acadèmics al llarg del treball, però degut a l'alt preu que s'ha de pagar per poder accedir al ser contingut no ha estat possible. Crec que el coneixement, especialment el dels articles acadèmics recents que és potser el més interessant, hauria de ser gratuït i disponible per a tothom, per poder col·laborar amb la recerca i que tothom tingui les mateixes possibilitats.

2. Introducció

L'objectiu d'aquest treball de recerca és comparar dos fitxers d'àudio qualssevol per veure si s'assemblen o no, amb la intenció de trobar una aplicació per a la vida real. Tot i que la semblança és un criteri humà i per tant subjectiu, el mètode utilitzat és totalment objectiu a les dades obtingudes. La idea principal és que per a comparar dos arxius els podem restar entre ells; del resultat d'aquesta resta, que és simplement el que és diferent dels dos arxius, se'n pot extreure el percentatge de similitud i elaborar conclusions.

Els fitxers d'àudio es poden expressar en dues formes, en forma d'amplitud segons el temps i en forma d'amplitud segons les freqüències. Com que la comparació en base el temps és massa complexa de tractar, s'agafa l'arxiu de so i es transforma en base de freqüències mitjançant la transformada de Fourier. Aquesta transformada recull les dades d'un interval de temps i les converteix al domini de les freqüències, de tal forma que el so es representa d'una altra manera sense perdre informació. Per exemple, si apuntem l'instant (hora) en què un grill comença a cantar és una representació en el domini del temps, però si diem que el grill canta cada dos segons és una representació en el domini de les freqüències. Les dues coses representen exactament el mateix concepte.

En la part pràctica, es modifica un programa ja existent que utilitza la FFTW¹ per generar l'espectre dels diferents arxius de so que es volen comparar i l'exporta en format d'imatge. Un cop s'han generat aquestes imatges, s'utilitzen tècniques de processament d'imatge per fer-ne la resta. El resultat és una imatge de les parts de l'espectre que són diferents. Per tant, com més fosca sigui aquesta imatge resultant més semblants són els arxius, i com més llum tingui menys semblants seran.

El sistema que s'ha implementat té varies limitacions, entre elles: només es poden utilitzar arxius que tinguin la mateixa duració i que estiguin alineats en fase per aconseguir bons resultats.

Un cop totes les comparacions i les restes han estat realitzades, se n'extreu un valor que determina el percentatge de similitud entre 0 i 1. Per entendre els resultats, es creen matrius de correlació que ens ajuden a entendre les dades visualment i a comprovar que el programa funciona amb la percepció d'hom.

El treball s'ha estructurat de la següent forma: en primer lloc es troba la part teòrica, que està escrita de forma lineal perquè els coneixements es puguin anar adquirint des dels més bàsics al més complexes; en segon lloc, la part pràctica s'ha separat per clarificar encara més els conceptes amb els quals s'ha experimentat.

¹ The Fastest Fourier Transform in the West.

3. Part teòrica

3.1. Senyals

Quan nosaltres escoltem amb les orelles rebem senyals, és a dir, un conjunt de moltes ones de so que a simple vista no tenen cap forma (veure Figure 1) i que el nostre cervell interpreta. Aquests senyals estan composts per moltes ones pures (de forma sinusoidal (veure Figure 2)), que són senyals simples que tenen una sola freqüència i amplitud.

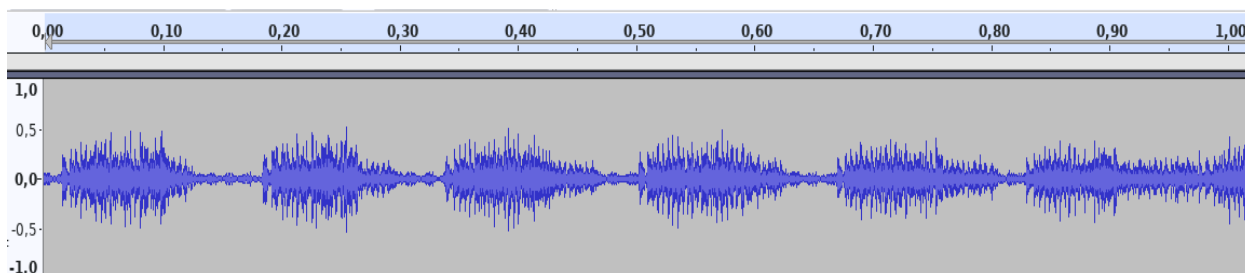


Figure 1: Representació en el temps del senyal del cant d'una cigala, formada per moltes ones diferents. [1]

Les formes d'ona o «waveforms» en anglès són la representació gràfica del senyal en funció del temps, és a dir, la variació de l'amplitud (alçada, eix vertical) de l'ona durant el temps (eix horitzontal) [2] [3] [4] [5] [6] [7]. L'amplitud es pot mesurar amb diferents unitats depenent del camp d'estudi, per exemple els ordinadors utilitzen voltatge però en el camp d'acústica es treballa amb decibels. Totes les ones pures tenen tres atributs bàsics: la **freqüència**, l'**amplitud** i la **fase**.

» Freqüència d'una ona pura

La freqüència és el nombre de vegades que el cicle (el pujar i baixar, veure Figure 2) d'una ona es repeteix per segon [8] [9]. Es mesura en Hertz (Hz) [10], que equival a un cicle o repetició per segon. Com més cicles hi ha per segon, més alta és la freqüència. El **període** (T), que és l'invers de la freqüència, és el temps que es tarda a realitzar un cicle en segons [8].

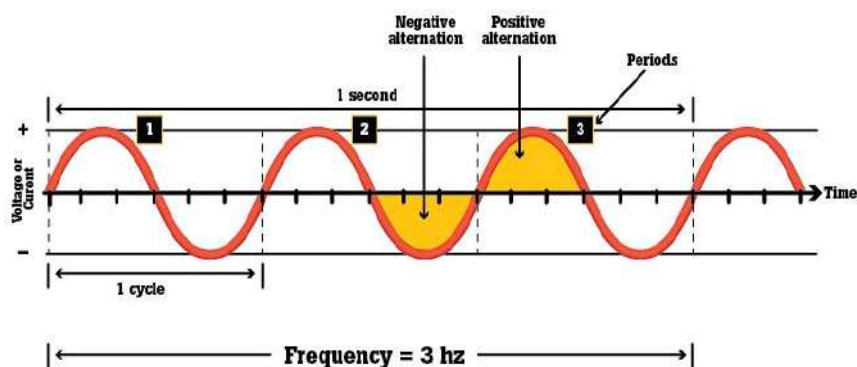


Figure 2: Representació gràfica de la freqüència en una ona. L'exemple mostra una ona de 3Hz, per tant el cicle de l'ona es repeteix 3 vegades cada segon. [11]

» Fase d'una ona pura

La fase és el desplaçament positiu (dreta) o negatiu (esquerra) d'una ona respecte el temps zero o inicial (quan l'ona es comença a emetre o es comença a mesurar). S'utilitza sobretot per a comparar dues ones de la mateixa freqüència que estan desplaçades en el temps. S'expressa amb la lletra ϕ i es pot mesurar amb segons, distància o angles (360° correspon a un cicle, 180° mig cicle, sempre en el sentit de les agulles del rellotge) [13]. Quan no hi ha separació i els pics queden perfectament solapats es diu que les ones estan en **fase** (veure Figure 3 per exemples).

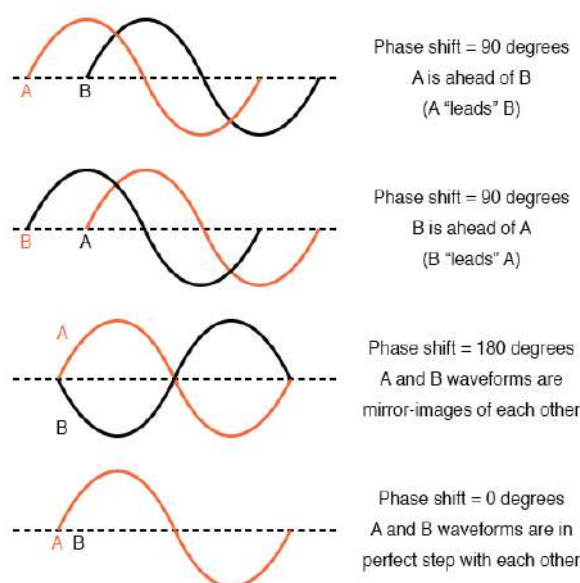


Figure 3: Quatre exemples de la fase. A l'últim exemple les ones estan en fase. [12]

Un cas especial és quan dues ones exactament iguals es troben en fase oposada, és a dir, en el cas del tercer exemple de la Figure 3 o el de la Figure 10. En aquest cas, el resultat de la suma de les dues ones és silenci, ja que les dues ones es cancel·len entre elles.

» Amplitud d'una ona pura

L'amplitud és el valor que mesura l'altitud de l'ona en el gràfic [14]. Es pot mesurar en decibels (volum d'àudio), volts, desplaçament (metres), etc.

En el cas del so el que passa és que viatja per un medi i fa que les seves molècules es moguin; per tant com més amplitud té una ona de so amb més energia es mouen aquestes partícules. L'amplitud té valors positius i negatius que oscil·len per sobre i per sota de la posició de repòs (o referència), és a dir, el silenci (veure Figure 4).

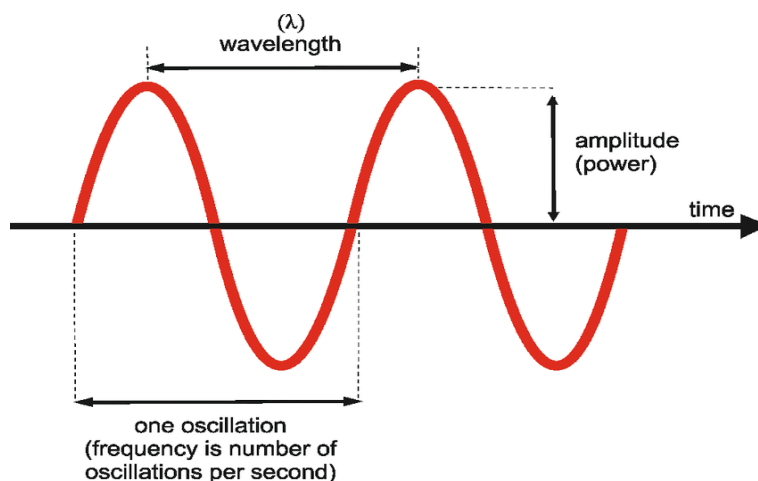


Figure 4: Representació de l'amplitud (eix Y), una oscil·lació i la longitud d'ona o "wavelength" (λ). L'eix de color negre és la posició de repòs [14].

La **longitud de l'ona** (λ) és la distància física entre un punt del cicle de l'ona i el mateix punt del següent cicle (veure Figure 5). [15]

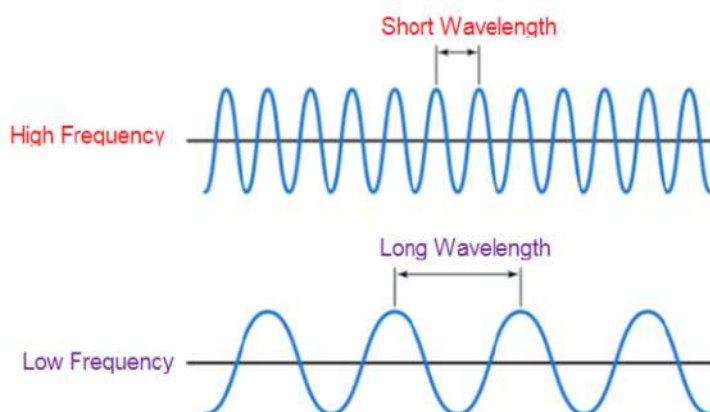


Figure 5: Dos exemples de la longitud d'ona. El primer és una ona d'alta freqüència, que té una longitud d'ona curta. El segon és una ona de baixa freqüència que té una longitud d'ona llarga. [15]

Per calcular-la, s'utilitza la següent fórmula:

$$\text{Longitud d'ona} = \lambda = \frac{\text{Velocitat del medi}}{\text{Freqüència}}$$

La **velocitat del medi**, és a dir la velocitat a la que el so es propaga en un material concret, canvia segons el medi pel qual viatja (veure 1). Segons la freqüència de l'ona i el material obtenim resultats diferents. Per exemple, si un sintetitzador² reproduïx una freqüència de 550Hz per un altaveu per l'aire (346 m/s) té una longitud d'ona de 0,63m aproximadament. Si aquesta ona viatgés pel ferro (5950 m/s), tindria una longitud d'ona de 10,81m aproximadament.

² Instrument musical que produeix sons a partir de combinacions d'ones pures.

Taula 1: Velocitat del so viatjant per diferents medis a una temperatura de 25°C [16].

Estat del medi	Medi	Velocitat del so (m/s)	Longitud d'ona (λ) a 550Hz
Sòlids	Alumini	6420	$\approx 11,67\text{m}$
	Ferro	5950	$\approx 10,81\text{m}$
	Vidre de rellotge	5100	$\approx 9,27\text{m}$
	Vidre flint	3980	$\approx 7,24\text{m}$
Líquids	Aigua destil·lada	1498	$\approx 2,72\text{m}$
	Mercuri	1450	$\approx 2,64\text{m}$
	Etanol	1207	$\approx 2,19\text{m}$
Gasos	Hidrogen	1339	$\approx 2,43\text{m}$
	Aire	346	$\approx 0,63\text{m}$
	Diòxid de carboni	269	$\approx 0,49\text{m}$

El mateix efecte passa amb les ones de llum, on el medi per on viatgen afecta la seva velocitat i longitud d'ona, però amb unes velocitats molt més altes (veure Figure 6).

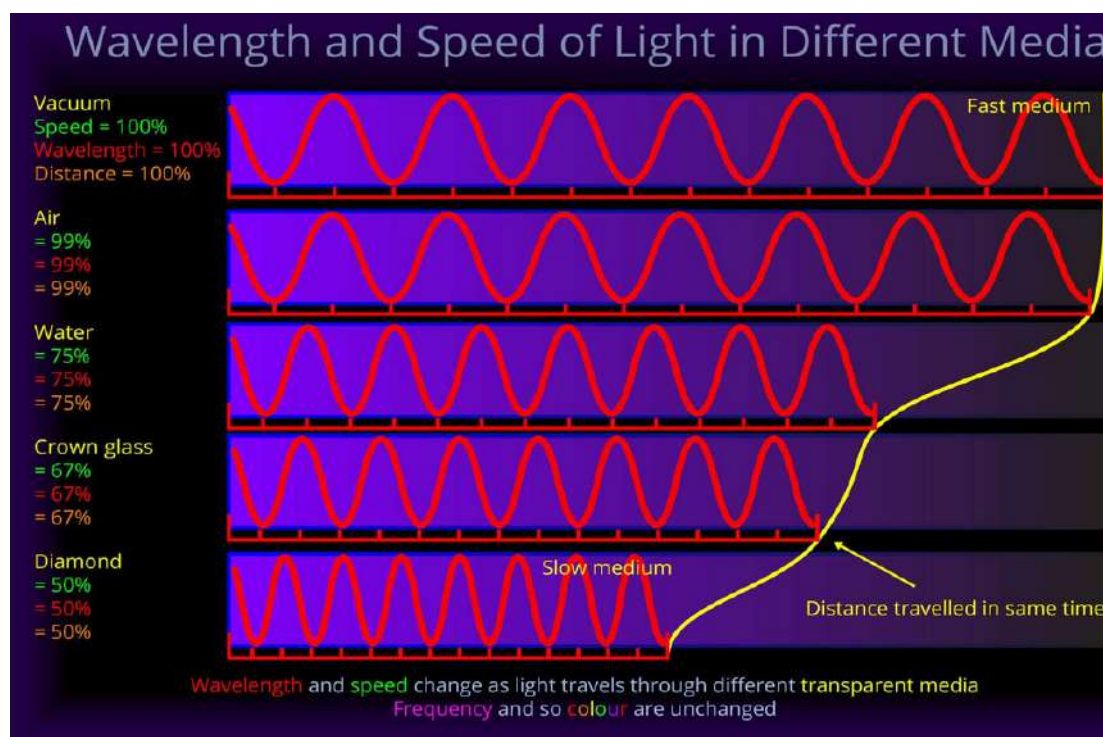


Figure 6: Gràfic que mostra la velocitat, la longitud d'ona i la distància que recorre la llum en el mateix període de temps en medis diferents. La freqüència es manté constant, i per tant el color de la llum tampoc varia. [17]

Decibels

Els decibels (dB) són una de les unitats que s'utilitzen per mesurar l'amplitud, sobretot en el camp d'acústica. Expressa el nombre d'increments de l'amplitud respecte un nivell de referència, que és 0 dB i es fixa en el llindar de la percepció humana. És una unitat logarítmica que és 10 vegades el logaritme de base 10 de la relació entre l'amplitud del senyal i l'amplitud base ($10 \cdot \log_{10}x$, veure la Figure 7). De forma

simplificada, podem mirar els decibels com al nombre de 0s que té l'amplitud. El nombre de 0 que suposa l'increment és el valor en decibels dividit per 10 (10 decibels suposen un 0 al darrere, 20 decibels suposen dos zeros, etc.).

dB	power ratio
100	10 000 000 000
90	1 000 000 000
80	100 000 000
70	10 000 000
60	1 000 000
50	100 000
40	10 000
30	1 000
20	100
10	10
6	3.981
3	1.995 (~2)
1	1.259
0	1
-1	0.794
-3	0.501 (~1/2)
-6	0.251
-10	0.1
-20	0.01
-30	0.001
-40	0.000 1
-50	0.000 01
-60	0.000 001
-70	0.000 000 1
-80	0.000 000 01
-90	0.000 000 001
-100	0.000 000 000 1

Figure 7: Taula on es mostren els diferents resultats de la unitat logarítmica. La potència escala amb proporció de x . [18]

3.2. Combinació dels senyals

Els senyals són la suma de moltes ones sinusoidals diferents. La suma de dos o més ones pures dona com a resultat una ona composta, senyal o «waveform». Per fer aquesta suma s'agafa el valor de l'amplitud a cada instant de l'ona i es suma amb el valor de l'amplitud de l'altra ona del mateix instant. Els següents exemples ho clarifiquen:

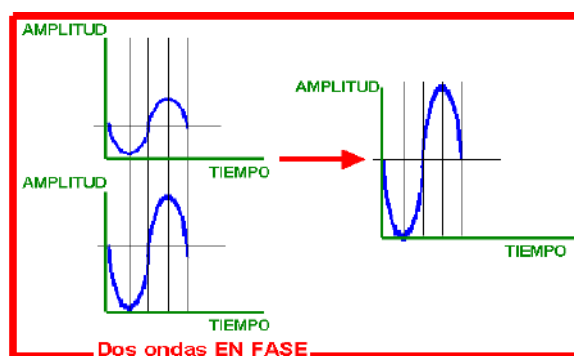


Figure 8: Exemple de dues ones en fase. Com que la freqüència de les ones és la mateixa, simplement augmenta l'amplitud, donant com a resultat una ona més forta. [13]

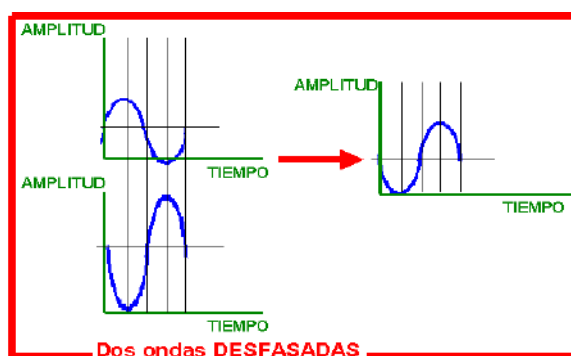


Figure 9: Exemple de dues ones desfasades 180° o mitja ona. Com que la segona ona té una amplitud superior a la primera, l'ona resultant queda invertida però té exactament la mateixa freqüència.[13]

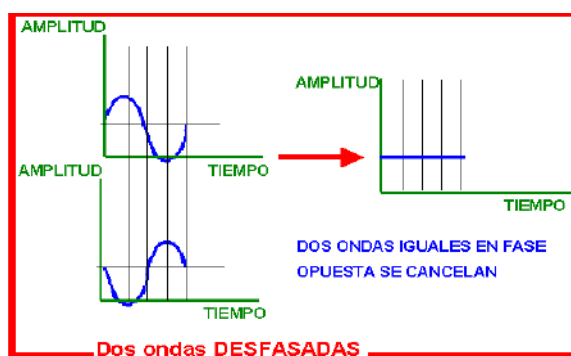


Figure 10: Exemple de dues ones iguals desfasades 180° o mitja ona. Com que tenen la mateixa amplitud positiva en l'ona superior i negativa en la inferior, es cancel·len entre elles i el resultat és silenci.[13]

Harmònics

Un harmònic d'una ona és una altra ona amb la freqüència multiplicada per un nombre enter positiu (2, 3, 4, mai valors decimals) de l'ona original o fonamental (veure Figure 11) [19]. S'ordenen de manera ordinal³; d'aquesta forma el segon harmònic és el doble de la freqüència fonamental, el tercer és el triple, etcètera. S'utilitzen molt perquè els senyals solen estar formats per diferents harmònics de les freqüències pures que els formen.

³ Seguint un ordre i unes posicions (primer, segon, tercer, quart...).

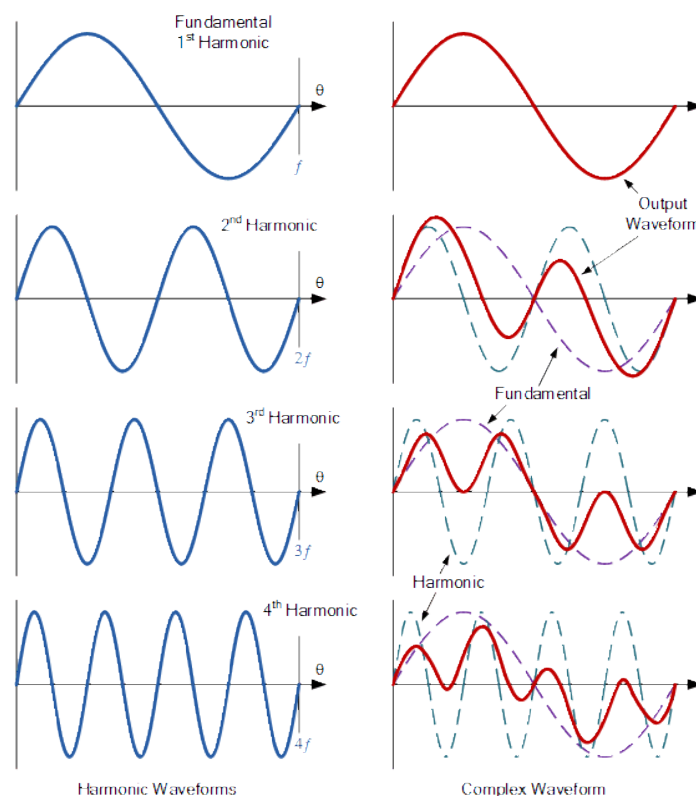


Figure 11: Suma d'una ona fonamental amb quatre harmònics diferents. A la dreta es veu el resultat de la suma (el resultat és l'ona vermella, les ones discontinues són els dos harmònics).[20]

3.3. Captura del senyal

Per capturar els diferents senyals que els humans percebem utilitzem l'oïda, un sistema natural complex que està format per la orella (redirigeix els senyals cap a l'oïda interna) i una membrana flexible que vibra i capta els senyals. Aquest aire entra fins uns petits ossets que amplifiquen el so i a dins l'oïda interna hi tenim la còclea [21] que estimula les cèl·lules nervioses que envien la informació al nostre cervell [22].

No obstant, els ordinadors utilitzen una manera completament diferent, que funciona recollint les dades dels sensors que s'hi connecten. Les plaques base d'ordinador disposen d'un xip que apunta el voltatge proporcionat per un micròfon cada X milers de vegades per segon [23].

3.3.1. Adquisició

Els ordinadors poden utilitzar diferents sensors per recollir dades. Els micròfons, per exemple, transformen les ones de so a un senyal elèctric. Tot i que n'hi ha de varis tipus, la idea general és que consten d'una membrana sensible on hi impacten les ones de so. Aquests petits cops d'aire fan que aquesta membrana es mogui (veure Figure 12) i mitjançant electrònica transforma aquest moviment en una corrent elèctrica que l'ordinador enregistra varis milers vegades per segon [24] [25].

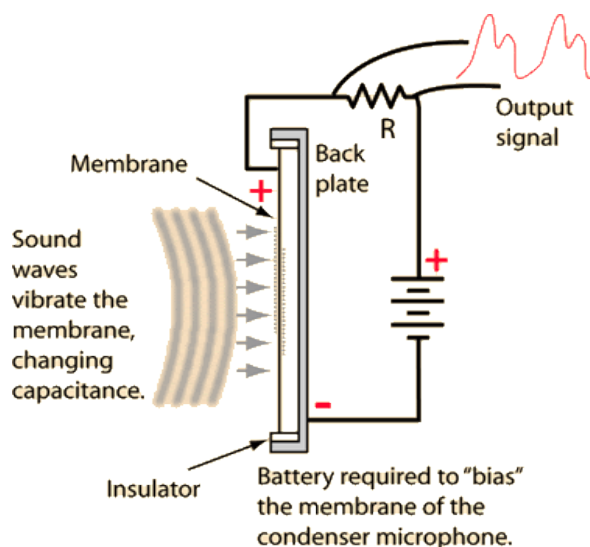


Figure 12: Esquema de funcionament d'un micròfon de condensador, on veiem la membrana i l'impacte de les ones de so, que produeixen les diferències de voltatge. La bateria és necessària per obtenir resultats [24].

3.3.2. Tractament digital de les mostres

» Resolució i profunditat de bits

Les mostres que els xips de la placa base recullen tenen més o menys precisió segons la **resolució** que s'utilitza. La resolució és el nombre de valors possibles que pot prendre una sola mostra (de quin número a quin número va l'escala) i s'expressa amb bits (veure 2). El nombre de valors normalment s'expressa en base 2^n , on el nombre de bits de l'escala és n (incrementa exponencialment). En els casos d'àudio, s'utilitza sobretot 16 bits per casos normals (qualitat CD) i 24 bits si es tracta d'una gravació de molt alta qualitat (qualitat estudi).

Taula 2: Diferents valors segons el nombre de bits de la resolució (2^n) [26]

Nombre de bits (n)	Nombre de valors de l'escala (2^n , per mostra)	Valors amb base 10 (per mostra)
1	2	De 0 a +1
2	4	de -2 a +1
4	16	de -8 a +7
8	256	de -128 a +127
11	2048	de -1024 a +1023
12	4096	de -2048 a +2047
16	65,536	de -32,768 a +32,767
18	262,144	de -131072 a +131071
20	1,048,576	de -524,288 a +524,287
24	16,777,216	de -8,388,608 a +8,388,607
32	4,294,967,296	de -2,147,483,648 a +2,147,483,647

Nombre de bits (n)	Nombre de valors de l'escala (2 ⁿ , per mostra)	Valors amb base 10 (per mostra)
48	281,474,976,710,656	de -140,737,488,355,328 a +140,737,488,355,327
64	18,446,744,073,709,551,616	de -9,223,372,036,854,775,808 a +9,223,372,036,854,775,807

Aquest nombre de bits correspon a la qualitat (la resolució de l'eix vertical de cada mostra) que s'agafa. S'anomena profunditat de bits («bit depth» en anglès)[27], i com més baixa és, més gran és l'error de la mesura de cada mostra (veure Figure 14). Es pot contrarestar amb un nombre de mostres per segon més alt.

» Mostratge

Per «mostrar» el senyal, el que el xip fa és anotar el valor de l'amplitud i de l'instant en què s'ha mesurat (de forma implícita), normalment varis milers de vegades per segon. Com més alta sigui la **frequència de mostratge** (el nombre de vegades en què es recullen les mostres) [29] [30], més precís i semblant al senyal real és el resultat. Es pot veure la freqüència de mostratge com la resolució de l'eix horitzontal. Per aconseguir un resultat 100% exacte a la realitat s'hauria de recollir un nombre de mostres infinit cada segon, cosa que és físicament impossible.

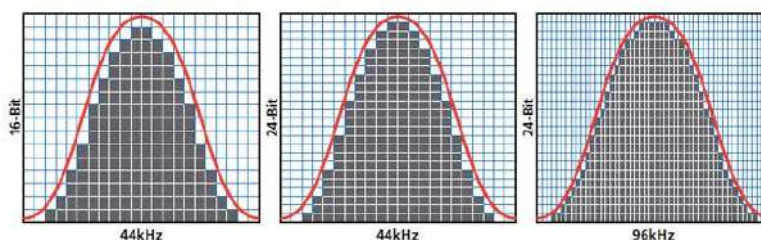


Figure 14: Tres exemples de mostreig. El primer té la freqüència de mostreig i la profunditat de bits més baixa dels tres, resultant en la qualitat més baixa. L'últim exemple és el que té més precisió, ja que té una freqüència de mostres dues vegades més gran i té molta més profunditat de bits. [31]

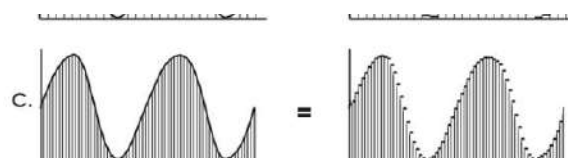


Figure 13: Tres exemples de mostreig digital. El primer cas té un nombre més baix de mostres per segon que l'últim, per tan el resultat és menys precís. L'últim resultat és el més acurat dels tres, ja que te una resolució més alta. [28]

3.3.3. Teorema de Nyquist

El Teorema de Nyquist [32] resumidament diu que es necessita recollir un nombre de mostres de com a mínim el doble de la freqüència pura màxima que es pretén capturar. Com que l'oïda humana pot sentir entre 20 i 20.000 Hz, normalment s'utilitza una freqüència de mostreig (o nombre de mostres que

s'adquireixen per segon) de 44100Hz perquè és més del doble de 20.000Hz i evita problemes amb l'[al·liasing](#)⁴.

3.4. Transformada de Fourier

Explicació simplificada

La Transformada de Fourier [34] és la part «màgica» del treball. Fourier estableix que qualsevol senyal compost es pot dividir en un nombre X d'ones pures o sinusoidals amb « Y » amplitud/importància. Si es considera un senyal com a una llista de freqüències que tenen un temps associat, el que fa la Transformada és expressar la mateixa informació però amb la quantitat de cada freqüència (és a dir, canvia el [domini](#) de la informació, veure Figure 15).

Un exemple visual pot ser un piano: cada divisió de freqüències que Fourier fa es pot representar com una tecla de piano amb la seva nota, i la quantitat de la freqüència representa el volum de cada tecla (les tecles amb un volum més alt serien les ones pures). En tocar totes les tecles del piano alhora amb el seu volum corresponent s'escoltaria el senyal original.

Un altre exemple podria ser amb un instrument de vent qualssevol. Si es connectessin 22050 flautes a un compressor (la meitat de 44100Hz, la freqüència de Nyquist amb la que es capturen la majoria dels senyals), cadascun amb una nota diferent (una nota per a cada freqüència). El que retornaria la transformada seria la quantitat d'aire que cada instrument hauria de tocar perquè se sentís el senyal real.

Tot i que no és del tot correcte, la fase es podria arribar a entendre com el retràs o l'avançament en el qual s'ha de començar a tocar cada flauta.

Explicació més profunda

La Transformada de Fourier [34] és la part central de la part pràctica del treball. Fourier estableix que qualsevol senyal compost es pot dividir en X nombre d'ones pures o sinusoidals. La Transformada de Fourier, mitjançant una operació matemàtica, descompon qualsevol senyal en les seves ones pures que el formen (veure Figure 15).

⁴ L'al·liasing és el fenomen que es produeix quan un senyal es mostreja un nombre finit de vegades, és a dir, sempre. Com que el mostreig no és infinit no es pot representar un rang infinit de freqüències. Quan es graven freqüències que estan fora d'aquest rang (que està determinat per la freqüència de mostreig), aquestes freqüències s'interpreten i es converteixen a freqüències que sí que estan a dins el rang, de forma que no es captura el senyal real. L'al·liasing, doncs, és quan una freqüència s'interpreta com una altra freqüència degut a una limitació de rang de freqüències. [33]

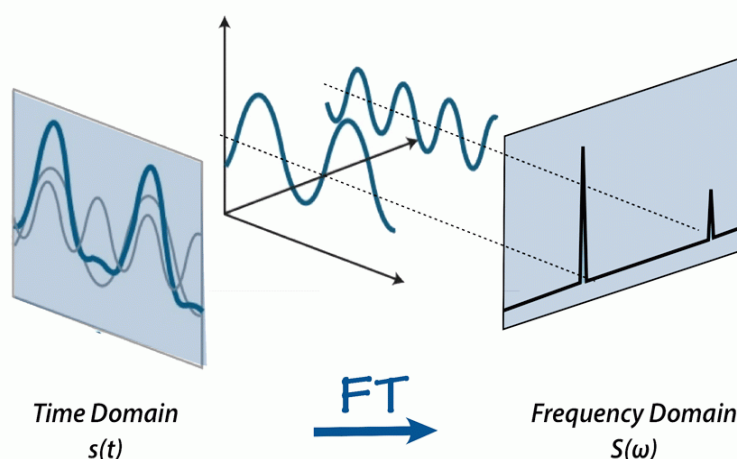


Figure 15: La Transformada de Fourier canvia qualsevol senyal (domini del temps) al domini de les freqüències (l'espectre). Els dos pics que es veuen al gràfic de la freqüència representen les dues ones pures que formen el senyal compost. [35]

El resultat de l'operació matemàtica torna una llista de vectors (un per a cada freqüència, en forma complexa $[a+bi]$), on el mòdul del vector representa la quantitat d'aquella freqüència en el senyal original (amplitud) i l'angle del vector respecte l'eix horitzontal representa la fase de la freqüència que s'analitza. Per tant, els vectors que tenen el mòdul més gran són els de les freqüències pures que formen el senyal compost. No obstant, cal destacar que la resta de freqüències gairebé sempre tindran un valor degut al [soroll](#) de fons que tenen tots els senyals.

La freqüència màxima que la Transformada expressarà es determina segons la freqüència de mostreig (veure 3.3.2 i 3.3.3).

Soroll

El soroll és el conjunt de les freqüències aleatòries (veure Figure 16) que no volem rebre d'un senyal. Aquestes freqüències sempre hi són presents i interfereixen amb els circuits del micròfon o del receptor i empitjoren la qualitat de la gravació. El soroll sol ser una combinació de soroll atmosfèric, altres senyals que es creuen amb el que volem escoltar i interferències que generen els aparells electrònics del nostre voltant [36]. Si el soroll no existís, es podria rebre qualsevol senyal dèbil a gairebé qualsevol distància. És la quantitat de soroll de l'ambient el que determina què podem rebre i la seva qualitat.

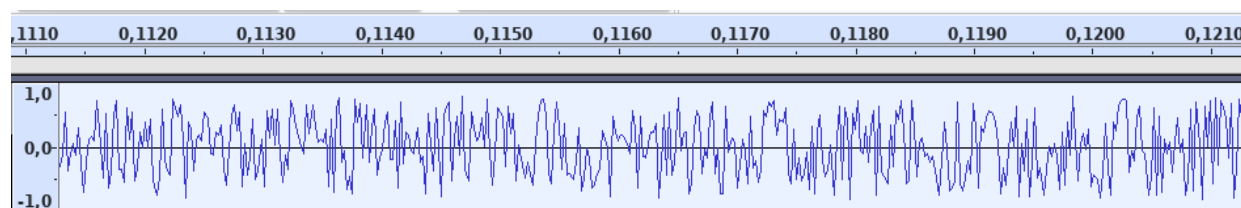


Figure 16: Soroll artificial general per l'Audacity [37]. És un senyal format per totes les freqüències i completament aleatori.[1]

3.4.1. Visualització de cascada

La visualització de cascada és una altra forma de representar els senyals que rebem. Ens permet veure mitjançant colors quines freqüències estan sonant en aquell mateix instant i en els instants passats. El nom de visualització de cascada prové de la forma en què veiem el resultat: la part superior de la imatge és l'instant actual, que va baixant per l'eix vertical com si fos l'aigua d'una cascada (veure Figure 17).

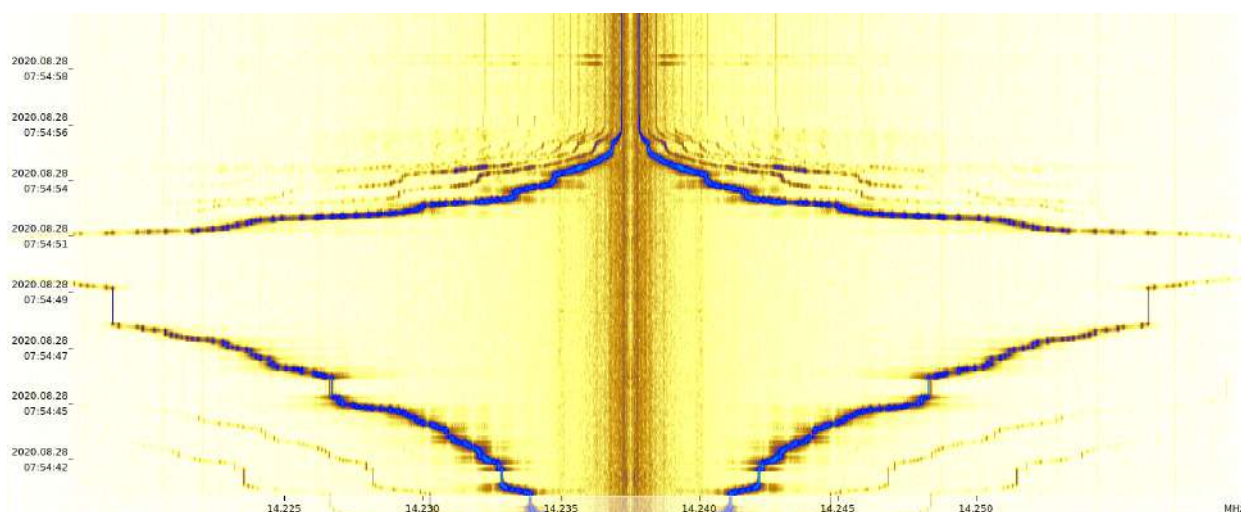


Figure 17: Exemple de visualització de cascada. El color blau representen una pujada i baixada de ones sinusoidals des de 100Hz fins a 15000Hz (el senyal). Les línies grogues i taronges del fons són soroll del propi micròfon. En el centre hi trobem les freqüències més baixes, com més allunat del centre són més altes (la línia central fa de mirall). Aquesta cascada està generada amb el GQRX [38], un software per visualitzar senyals de ràdio i de micròfons.[1]

3.4.2. FFTW

L'FFTW és una eina per fer la Transformada de Fourier discreta (és a dir, per a un nombre finit de mostres) de la forma més ràpida possible en un ordinador [39]. Utilitza moltes tècniques d'optimització a nivell de software i de hardware i molts algorismes per als diferents casos que es pot trobar per aconseguir el mateix resultat que la transformada tradicional en un temps molt més ràpid. Està programada en C i donada una llista de mostres d'àudio retorna una llista de vectors complexos.

3.5. Espai de color

Els espais de colors són formes d'organitzar els colors per representar-los digitalment. L'espai de color que utilitzen la majoria de monitors és **l'RGB** («red-green-blue») [40]. La idea principal de l'RGB és

Treball de recerca • Comparació d'ones

que es pot crear qualsevol color si se sumen tres raigs de llum: un de vermell, un de verd i un de blau. Per aconseguir els diferents colors, es canvia la quantitat de llum de cada raig (veure Figure 18).

		<<< Control 2 - Adjust FROM Colour to White >>>																	
		1			2			3			4			5			6		
		R	G	B	R	G	B	R	G	B	R	G	B	R	G	B	R	G	B
Red	23	255	0	0	255	100	100	255	150	150	255	200	200	255	230	230	255	255	255
	22	255	100	0	255	110	20	255	120	100	255	165	155	255	230	210	255	255	255
Orange	21	255	150	0	255	140	20	255	180	40	255	220	120	255	240	190	255	255	255
	20	255	200	0	255	210	20	255	230	100	255	240	150	255	255	200	255	255	255
Yellow	19	255	255	0	255	255	20	255	255	80	255	255	150	255	255	210	255	255	255
	18	200	255	0	200	255	20	220	255	80	230	255	150	235	255	180	255	255	255
Chartreuse	17	150	255	0	150	255	20	200	255	40	220	255	150	230	255	170	255	255	255
	16	100	255	0	110	255	20	150	255	40	200	255	150	230	255	160	255	255	255
Green	15	0	255	0	100	255	100	150	255	150	200	255	200	225	255	225	255	255	255
	14	0	255	100	20	255	150	100	255	180	180	255	210	235	255	220	255	255	255
Aquamarine	13	0	255	150	20	255	160	100	255	200	160	255	210	200	255	220	255	255	255
	12	0	255	200	20	255	200	40	255	220	140	255	230	200	255	255	255	255	255
Cyan	11	0	255	255	20	255	255	40	255	255	150	255	255	180	255	255	255	255	255
	10	0	200	255	20	210	255	40	230	255	60	240	255	170	255	255	255	255	255
Azure	9	0	150	255	20	180	255	40	200	255	60	230	255	150	255	255	255	255	255
	8	0	100	255	20	150	255	40	180	255	60	200	255	140	255	255	255	255	255
Blue	7	0	0	255	20	100	255	40	150	255	80	180	255	160	220	255	255	255	255
	6	100	0	255	120	20	255	140	100	255	150	150	255	200	200	255	255	255	255
Violet	5	150	0	255	150	50	255	180	100	255	220	140	255	230	180	255	255	255	255
	4	200	0	255	200	20	255	220	60	255	240	90	255	255	160	255	255	255	255
Magenta	3	255	0	255	255	60	255	255	100	240	255	150	250	255	180	255	255	255	255
	2	255	0	200	255	20	220	255	40	230	255	140	240	255	170	255	255	255	255
Rose	1	255	0	150	255	20	200	255	40	200	255	100	200	255	160	240	255	255	255
	0	255	0	100	255	20	120	255	80	140	255	100	160	255	140	180	255	255	255




Figure 18: Taula de colors RGB. Canviant la quantitat de llum de cada raig es poden aconseguir tots els colors. La quantitat màxima de cada raig és de 255 i la mínima (apagat) és 0. [42]

Tot i que l'RGB és l'espai de color més utilitzat, en la part pràctica s'utilitzarà un espai de color diferent per simplificar la programació: **l'HSV** («hue-saturacion-value») [41]. L'HSV, també conegut com a HSB (**brightness**), funciona d'una manera més semblant a la percepció dels humans. Consta de tres valors:

- » El matís (hue), que correspon al color. En el programa s'expressa amb un percentatge sobre 1, però també es pot expressar amb graus (veure Figure 19). Com que l'RGB té 3 colors bàsics, es divideix aquest 1 en tres parts, de tal forma que queda $\frac{1}{3}$ per a cada color (o 120°). Per les barreges de colors s'utilitzen els graus intermedis, de tal forma que $\frac{1}{6}$ (o 60°) és el color groc (suma de vermell i verd, veure 3).

Taula 3: Diferents valors del matís, el seu valor en RGB i el color resultant. [41]

Valor en graus	Valor en percentatge	Valor en RGB	Color
0°	0	(1,0,0)	
60°	$\frac{1}{6}$	(1,1,0)	
120°	$\frac{1}{3}$	(0,1,0)	
180°	$\frac{1}{2}$	(0,1,1)	

Valor en graus	Valor en percentatge	Valor en RGB	Color
240°	$\frac{2}{3}$	(0,0,1)	
300°	$\frac{5}{6}$	(1,0,1)	
360° = 0°	0	(1,0,0)	

- » La saturació (saturation) que es pot definir com la «puresa» del color. També s'expressa amb un percentatge sobre 1, de forma que com més pròxim és el valor de la saturació al 0 més gris serà la tonalitat del color (si la saturació és 0 el resultat serà una escala de grisos), i com més proper a 1 més viu serà el color.
- » El valor o la brillantor (value, brightness) que és la quantitat de llum que té el color. També és un percentatge sobre 1. Si el valor és 0, el color sempre serà negre. No obstant, si el valor és 1 pot ser blanc o un color més o menys saturat.

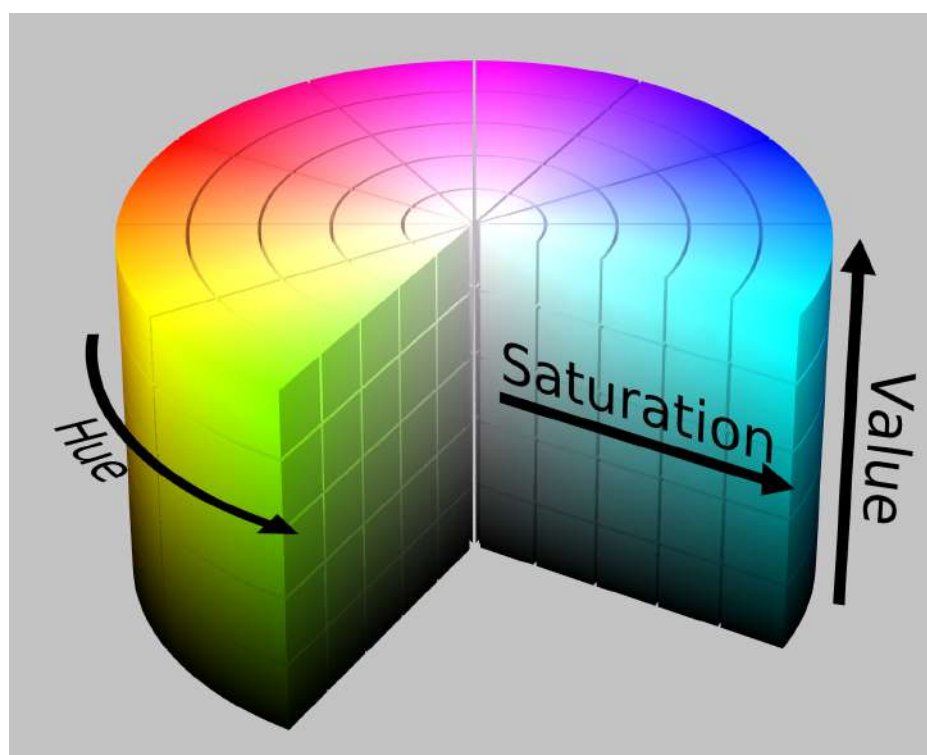


Figure 19: Representació gràfica de l'espai HSV. El Hue (color) és l'angle, la saturació és la profunditat del cilindre, que equival a la tonalitat, i el valor és la brillantor (la part inferior és negra, la part superior és el color més viu i brillant).[43]

3.6. Resta d'imatges

La resta d'imatges és una tècnica que s'utilitzarà en la part pràctica. Per restar dues imatges el que es fa és agafar el **valor numèric** de cada píxel (la **brillantor** si la imatge està en blanc i negre) de la primera imatge i restar-lo amb la segona imatge. Llavors s'agafa el valor absolut del resultat, de tal forma que tan és restar «a» amb «b» com «b» amb «a». Per aconseguir resultats precisos cal que les dues

imatges a restar tinguin les mateixes dimensions (el mateix nombre de píxels), ja que si els nombres són diferents no es poden comparar i s'han de redimensionar⁵ perquè siguin iguals i poder-les restar. El resultat de la resta és una imatge on queda de color el que és diferent de les dues imatges (blanc si la imatge és en blanc i negre) i el que és idèntic queda negre (veure exemples a sota). Al final d'aquest apartat s'exposa un exemple il·lustratiu de l'aplicació de la resta d'imatges en les Figure 20, Figure 21, Figure 22 i Figure 23.

Cal destacar que el cost de processament de la resta és **quadràtic** (segueix una funció de grau 2), és a dir, si el nombre d'imatges a restar és n , el nombre d'imatges resultants després de la comparació serà n^2 (si hi ha 5 imatges a comparar i cada imatge es compara amb les 5 imatges, ens trobem amb $5 \cdot 5$, és a dir 5^2 , que és 25 imatges resultants). Un exemple visual d'aquest cost seria quan es vol enrajolar una superfície; si es vol enrajolar la mateixa superfície amb unes rajoles grans es tardarà poc perquè se n'hauran de col·locar poques, però si es volen utilitzar rajoles molt petites (més imatges a comparar), es tardarà molt més perquè n'hi haurà moltes més. En termes de computació, això significa que com més imatges es vulguin comparar més estona es tardarà, però aquest creixement del temps també serà quadràtic.

⁵ Convertir el nombre de píxels d'una imatge en un altre, de tal forma que la imatge resultant sigui el menys diferent a la original. Si la relació d'aspecte canvia (el nombre de píxels de l'eix y dividit per el nombre de píxels de l'eix x), la imatge quedarà deformada, però és necessari per realitzar la resta.



Figure 20: Exemple d'imatge 1. [1]



Figure 21: Exemple d'imatge 2. [1]



Figure 22: Resta de la imatge 1 amb la imatge 2. En blanc el telèfon mòbil que és la principal diferència. També hi ha petites diferències de llum en la part superior esquerra. [1]

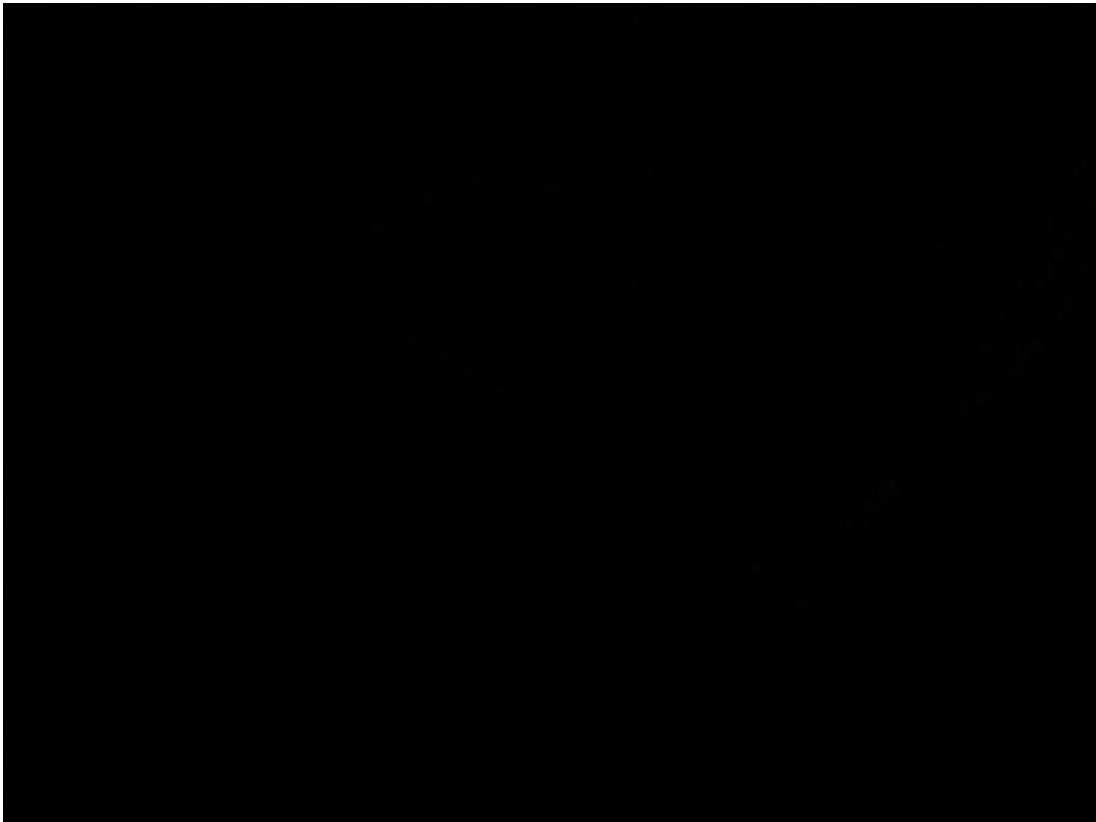


Figure 23: Resta de la imatge 1 amb la imatge 1. El resultat és totalment negre perquè són la mateixa imatge. [1]

3.7. Matrius de correlació

Per veure els resultats de la part pràctica d'una forma entenedora i visual s'utilitzaran el que s'anomena matrius de correlació. En el cas de fer una comparació de dos arxius «a» i «b», que és el cas del treball, es col·loca a l'eix y un dels arxius que es compara (a) i a l'eix x l'altre arxiu (b). El valor de cada cel·la és el percentatge de similitud que té l'arxiu que li correspon a la columna i a la fila (veure Figure 24).

Un fet que sempre es troba és que una de les diagonals sempre quedarà 1. Això és degut a que en aquesta diagonal l'arxiu de la columna i el de la fila és el mateix, per tant com que es compara exactament el mateix el percentatge de similitud serà del 100% (o 1 sobre 1).

	cica- da-01	cica- da-02	cica- da-03	cica- da-04	cica- da-05	cica- da-06	cica- da-07	cica- da-08	cica- da-09	cica- da-10	cica- da-11	cica- da-12	cica- da-13	cica- da-14	grill- 01	grill- 02	grill- 03	grill- 04	grill- 05	grill- 06	grill- 07	grill- 08	grill- 09	grill- 10
Comparing																								
cicada-01	1	0,969	0,967	0,964	0,967	0,964	0,937	0,935	0,962	0,963	0,962	0,91	0,921	0,914	0,924	0,916	0,885	0,885	0,909	0,907	0,875	0,878	0,906	0,906
cicada-02	0,969	1	0,967	0,964	0,966	0,963	0,938	0,937	0,962	0,963	0,961	0,913	0,924	0,917	0,926	0,917	0,888	0,888	0,909	0,909	0,877	0,88	0,906	0,908
cicada-03	0,967	0,967	1	0,968	0,969	0,963	0,944	0,943	0,958	0,959	0,957	0,923	0,934	0,927	0,934	0,925	0,899	0,899	0,917	0,913	0,887	0,89	0,911	0,912
cicada-04	0,964	0,964	0,968	1	0,968	0,962	0,947	0,946	0,955	0,956	0,954	0,926	0,937	0,931	0,933	0,927	0,903	0,904	0,919	0,915	0,891	0,893	0,911	0,911
cicada-05	0,967	0,966	0,969	0,968	1	0,963	0,944	0,942	0,959	0,959	0,958	0,92	0,931	0,924	0,932	0,922	0,896	0,896	0,915	0,911	0,885	0,888	0,91	0,911
cicada-06	0,964	0,963	0,963	0,962	0,963	1	0,939	0,938	0,956	0,957	0,956	0,912	0,924	0,917	0,922	0,914	0,888	0,888	0,912	0,907	0,877	0,88	0,905	0,905
cicada-07	0,937	0,938	0,944	0,947	0,944	0,939	1	0,953	0,927	0,929	0,925	0,945	0,951	0,947	0,94	0,936	0,928	0,929	0,924	0,91	0,912	0,915	0,907	0,906
cicada-08	0,935	0,937	0,943	0,946	0,942	0,938	0,953	1	0,925	0,927	0,924	0,947	0,953	0,949	0,937	0,939	0,93	0,931	0,924	0,911	0,913	0,915	0,907	0,907
cicada-09	0,962	0,962	0,958	0,955	0,959	0,956	0,927	0,925	1	0,97	0,97	0,899	0,91	0,903	0,916	0,905	0,874	0,874	0,902	0,901	0,864	0,867	0,903	0,903
cicada-10	0,963	0,963	0,959	0,956	0,959	0,957	0,929	0,927	0,97	1	0,969	0,902	0,913	0,906	0,918	0,907	0,877	0,877	0,905	0,902	0,867	0,87	0,904	0,905
cicada-11	0,962	0,961	0,957	0,954	0,958	0,956	0,925	0,924	0,97	0,969	1	0,898	0,91	0,903	0,915	0,904	0,873	0,873	0,902	0,9	0,863	0,866	0,902	0,903
cicada-12	0,91	0,913	0,923	0,926	0,92	0,912	0,945	0,947	0,899	0,902	0,898	1	0,97	0,972	0,949	0,95	0,96	0,96	0,93	0,91	0,938	0,939	0,909	0,907
cicada-13	0,921	0,924	0,934	0,937	0,931	0,924	0,951	0,953	0,91	0,913	0,91	0,97	1	0,972	0,951	0,95	0,953	0,953	0,932	0,915	0,932	0,933	0,914	0,912
cicada-14	0,914	0,917	0,927	0,931	0,924	0,917	0,947	0,949	0,903	0,906	0,903	0,972	0,972	1	0,95	0,95	0,957	0,957	0,931	0,912	0,935	0,936	0,912	0,91
grill-01	0,924	0,926	0,934	0,933	0,932	0,922	0,94	0,937	0,916	0,918	0,915	0,949	0,951	0,95	1	0,947	0,939	0,939	0,93	0,91	0,924	0,928	0,913	0,91
grill-02	0,916	0,917	0,925	0,927	0,922	0,914	0,936	0,939	0,905	0,907	0,904	0,95	0,95	0,95	0,947	1	0,945	0,945	0,929	0,909	0,929	0,932	0,908	0,906
grill-03	0,885	0,888	0,899	0,903	0,896	0,888	0,928	0,93	0,874	0,877	0,873	0,96	0,953	0,957	0,939	0,945	1	0,97	0,927	0,895	0,944	0,945	0,896	0,894
grill-04	0,885	0,888	0,899	0,904	0,896	0,888	0,929	0,931	0,874	0,877	0,873	0,96	0,953	0,957	0,939	0,945	0,97	1	0,928	0,896	0,944	0,944	0,897	0,894
grill-05	0,909	0,909	0,917	0,919	0,915	0,912	0,924	0,924	0,902	0,905	0,902	0,93	0,932	0,931	0,93	0,929	0,927	0,928	1	0,898	0,909	0,914	0,899	0,899
grill-06	0,913	0,914	0,917	0,918	0,915	0,913	0,913	0,913	0,906	0,908	0,906	0,91	0,916	0,912	0,911	0,908	0,892	0,892	0,896	1	0,879	0,882	0,944	0,944
grill-07	0,875	0,877	0,887	0,891	0,885	0,877	0,912	0,913	0,864	0,867	0,863	0,938	0,932	0,935	0,924	0,929	0,944	0,944	0,909	0,882	1	0,936	0,88	0,878
grill-08	0,878	0,88	0,89	0,893	0,888	0,88	0,915	0,915	0,867	0,87	0,866	0,939	0,933	0,936	0,928	0,932	0,945	0,944	0,914	0,884	0,936	1	0,882	0,881
grill-09	0,913	0,913	0,916	0,916	0,916	0,912	0,911	0,911	0,91	0,911	0,909	0,911	0,916	0,913	0,915	0,909	0,894	0,894	0,898	0,944	0,879	0,881	1	0,957
grill-10	0,913	0,914	0,916	0,915	0,916	0,912	0,909	0,909	0,91	0,911	0,91	0,908	0,913	0,91	0,911	0,906	0,891	0,891	0,897	0,944	0,876	0,879	0,957	1

Figure 24: Matriu de correlació del cant d'uns grills amb el cant d'unes cigales. Destacar com els valors de la diagonal són 1, ja que es comparen els mateixos arxius i són idèntics. [1]

A més a més, les matrius solen anar acompanyades d'escalas de color per remarcar quins arxius són més semblants i quins menys de forma visual i entenedora.

Un altre fet que sempre es trobarà és la **simetria** de la matriu. L'eix de la simetria sempre és la diagonal formada per els valors 1, i és una simetria desviada 90°: si ens fixem en els valors propers a la diagonal, veiem que els valors que veiem a sota de l'1 també es troben a la dreta d'aquest, ja que s'estan comparant els mateixos arxius. Aquesta simetria també és aplicable amb el valor superior i el de l'esquerra, i també es manté pels valors llunyans a la diagonal, sempre que es moguin el mateix nombre de caselles cap amunt (o avall) i cap a la dreta (o l'esquerra). És cert que, si ens fixem en la matriu de dalt, hi ha alguns valors que no compleixen aquesta simetria: la causa és desconeguda i no hauria de ser així, probablement és degut a aproximacions que l'ordinador fa a l'hora de calcular-la.

3.7.1. Desviació estàndard

Una de les eines que ajuda més quan es treballa amb moltes dades és la desviació estàndard [44]. Aquesta ens permet veure quant s'allunyen els resultats d'un conjunt de números amb la mitjana d'aquests. Per tant, com més petita sigui la desviació estàndard més a prop de la mitjana estaran els resultats. Es calcula amb la següent fórmula, on x és el valor de cada dada, μ és la mitjana del conjunt de dades i N és el nombre total de dades:

$$DE = \sqrt{\frac{\sum |x - \mu|^2}{N}}$$

La desviació estàndard ens serveix per veure, per exemple en les comparacions de la part pràctica, quins grups tenen una desviació més alta i més baixa. Si el valor és molt petit vol dir que l'impacte de la modificació és petit ja que tots els valors són propers a la mitjana, mentre que si el valor és bastant alt significa que hi ha resultats molt allunyats de la mitjana i per tant que la modificació aplicada té un impacte considerable.

4. Part pràctica

4.1. Introducció

L'objectiu de la part pràctica era aplicar tots aquests coneixements sobre les ones i trobar una manera de comparar senyals i ones. La primera idea era utilitzar un transmissor per enviar diferents senyals de ràdio des de varis llocs, capturar-los amb una antena i veure quina semblança tenia el senyal enviat amb el senyal rebut mitjançant software. Llavors, es podrien valorar quines característiques del clima, de la zona i del relleu afecten més la qualitat de les ones. No obstant, el fet de l'existència de la pandèmia ha dificultat molt el treball exterior i es va decidir canviar lleugerament el treball per uns resultats més digitals. Per aconseguir-ho, es va crear un programa que compara dos arxius de so qualssevol i retorna la semblança dels dos arxius. Les proves han estat realitzades amb arxius sintètics per experimentar en casos específics i també s'han utilitzat exemples del món real per corroborar-ne el funcionament.

4.2. Objectiu

L'objectiu principal és crear una nova manera d'utilitzar els resultats de la Transformada de Fourier per poder comparar dos arxius de so.

Resumidament, s'utilitza la FFTW per aconseguir els diferents vectors de la transformada de cada arxiu. A continuació es representen els vectors en forma d'imatge, de forma que s'aconsegueix l'espectre de cada àudio. Tot seguit, mitjançant tècniques de comparació d'imatges i eines fotogràfiques, es realitza la resta i se n'obté el percentatge de similitud segons el valor de la brillantor de la imatge resultant. Per a mostrar els diferents valors i analitzar-los s'utilitzen matrius de correlació per tal de veure els resultats de forma clara i poder extreure'n conclusions.

4.3. Experimentació

4.3.1. Recerca d'eines existents

En primer lloc es va fer recerca d'eines existents que aconseguissin de forma semblant el resultat proposat als objectius. Un dels resultats destacats va estar un software lliure anomenat `scape-xcorr-sound` [45] [46] que processa fitxers d'àudio i permet fer-los-hi varies operacions. Entre elles, la que destacava era el «`waveform-compare`», que mitjançant una sèrie d'algoritmes retornava el percentatge de semblança entre dos arxius de so. El problema que aquesta eina tenia era que no era suficientment consistent i no semblava precisa del tot. Gràcies a que el software era lliure i obert es va poder descarregar el codi font per veure què feia realment el programa i veure què es podia aprofitar per adaptar-lo a les necessitats del treball.

4.3.2. Creació d'arxius sintètics

Per saber si el programa d'Open Preservation funcionava correctament o no, es va crear un arxíu sintètic (veure Figure 25 i Figure 26). Aquest senyal consistia en un arxíu de 15 segons format per 4 ones sinusoidals, la primera d'una freqüència més baixa que l'última però amb la mateixa amplitud (veure Figure 26), separades per mig segon de silenci entre elles. Per acabar l'arxíu, hi havia uns 5 segons de soroll artificial, per veure com afectava a la comparació dels arxius quan el so era una mescla de totes les ones possibles.

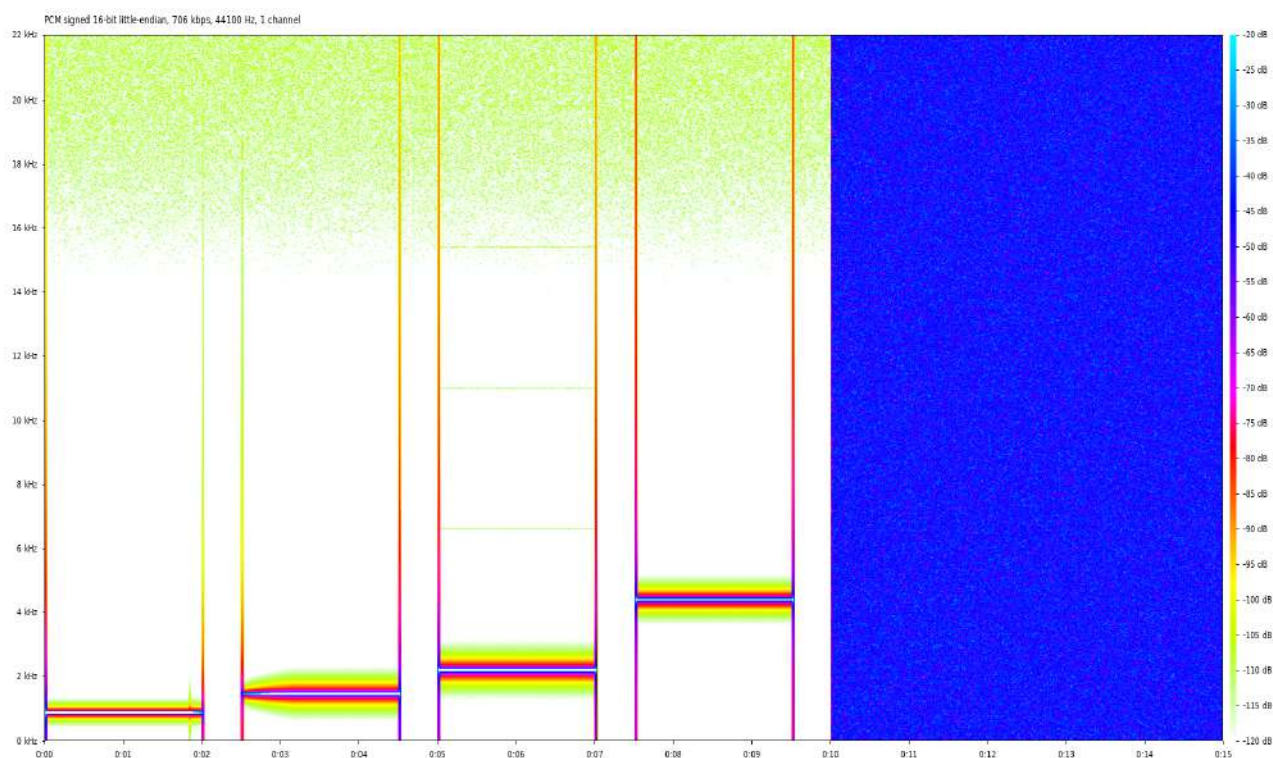


Figure 26: L'arxíu sintètic representat d'una altra forma, amb el software «spek» [0]. A l'eix vertical hi tenim les freqüències, a l'eix horitzontal el temps i els colors representen l'amplitud amb decibels (veure Decibels, pàgina 11). Veiem les quatre ones separades pel silenci (destacar els pics amb totes les freqüències al principi i final de cada ona, deguts al canvi bruscat de silenci a so i viceversa). A la dreta de tot hi ha el soroll, que està format per totes les freqüències.[1]

4.3.3. Modificació i variacions dels arxius

Un cop l'arxíu original va estar creat (l'anomenarem referència), s'hi van crear modificacions vàries per veure què afectava més en la diferència dels diferents arxius. Aquestes són les proves que es van fer als diferents arxius (comparar-les amb la Figure 25 per veure la diferència visual). De cada tipus de modificació se'n van fer diferents variacions amb canvis més i menys agressius per veure l'impacte que tenien amb més claredat.

Atenuació

Primer de tot vaig es va reduir l'amplitud general de l'arxiu referència. El volum era més baix, però les freqüències i la fase estaven totalment intactes.

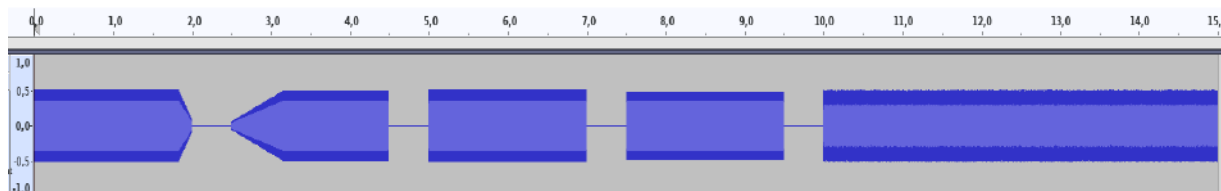


Figure 27: La duració i freqüència de l'arxiu és la mateixa, però l'amplitud és més petita.[1]

Contracció

La contracció de l'arxiu de so consistia en «premsar» l'arxiu, és a dir, reproduir l'arxiu en menys temps i afegir silenci al final per arribar als 15 segons. Aquesta compressió preserva l'amplitud però modifica la freqüència ja que es reproduïx el mateix senyal en menys temps, per tant la freqüència augmenta. Com que són freqüències diferents, la fase també es veu modificada tot i que els arxius comencen al mateix punt.

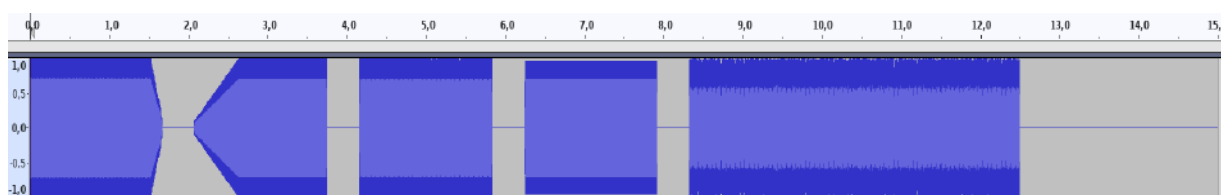


Figure 28: La duració dels blocs es veu considerablement reduïda. L'amplitud, però, és la mateixa.[1]

Expansió

L'expansió és el contrari de la contracció, és a dir, estirar l'arxiu de so perquè ocupi més espai, resultant en una freqüència més baixa. Per mantenir la duració de 15 segons, es va retallar el soroll perquè quedés exactament la mateixa duració d'arxiu.

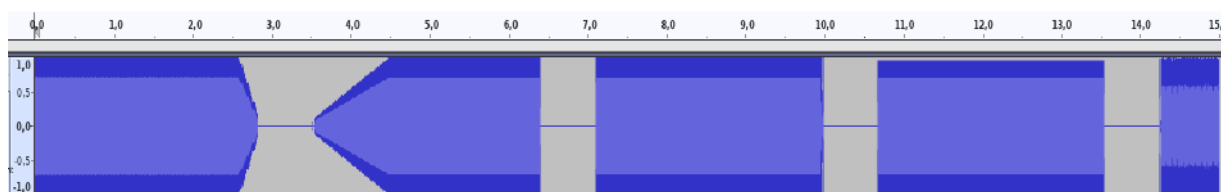


Figure 29: La duració dels blocs es veu considerablement ampliada. L'amplitud, però, és la mateixa. El silenci està retallat perquè tingui la duració exacta.[1]

Creació de soroll de fons

Un problema molt comú a l'hora de gravar qualsevol senyal és l'aparència de soroll de fons. Per simular aquesta situació es van afegir diferents nivells de soroll de fons a l'arxiu, de tal manera que les parts entre blocs on hi hauria d'haver silenci van ser reemplaçades per soroll. Les proves estan fetes amb varis nivells de soroll per veure la modificació (des de un 5% a un 100%).

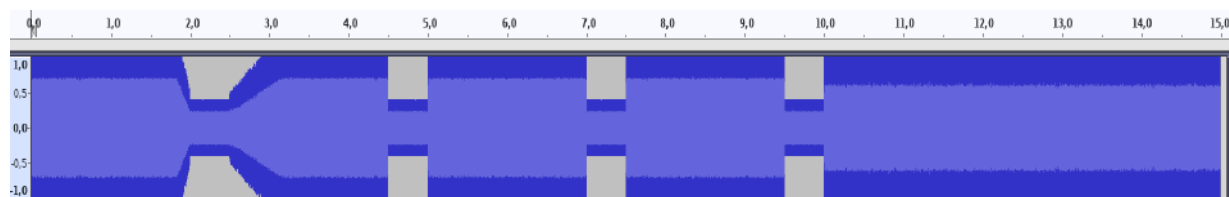


Figure 30: L'arxiu de referència però amb soroll artificial de fons, concretament un 40% de l'amplitud de l'arxiu original. [1]

Canviant la fase

Una altra prova va ser modificar la fase dels dos arxius idèntics per veure com afectava aquest desplaçament.

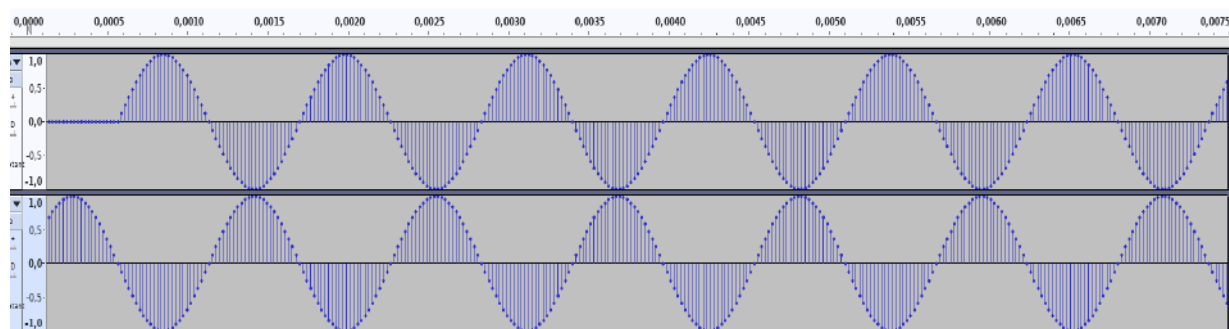


Figure 31: Els dos arxius que no estan en fase, sinó que estan separats per mig cicle. En aquest cas el programa *scape-xcorr-sound* (veure 4.3.1) va detectar aquest «phase-shift» i va donar com a resultat que els dos arxius eren completament idèntics. [1]

Gravant mitjançant el jack de l'ordinador.

Per afegir un nivell de realisme a les proves, es va decidir enviar l'arxiu de referència a un smartphone. Llavors, utilitzant un cable jack de 3.5mm es va connectar un extrem al connector dels auriculars del dispositiu i l'altre a l'entrada de micròfon de l'ordinador. D'aquesta manera, quan es reproduïa qualsevol so pel telèfon es podia gravar des de l'ordinador com si fos un micròfon. Es va repetir la gravació de l'arxiu referència 5 vegades per veure si hi havia alguna diferència amb el so sintètic original.

Gravant mitjançant un altaveu (20cm, 50cm, 100cm)

La última prova que es va fer va ser aprofitar l'arxiu de so que tenia a l'smartphone per reproduir-lo des dels altaveus del mòbil. Llavors, amb un micròfon de sobretaula es va gravar varies vegades el so reproduït des del telèfon. Per controlar les variables, es va gravar des de 20, 50 i 100 centímetres per veure si es notava la diferència en amplitud que suposava la distància de l'altaveu al micròfon.

4.3.4. Gravació de senyals reals

Per complementar els arxius sintètics es va decidir afegir un exemple del món real al conjunt de proves. Dins el món dels animals, un exemple senzill i simple de representar són els insectes, concretament els grills i les cigales. Aquests animals tenen un cant molt diferenciat que facilita l'anàlisi i la comparació. Es van recollir catorze mostres de cigales i deu mostres de grills, per a un total de vint-i-quatre

arxius més de comparació. Les gravacions estan fetes a diferents hores del dia pels boscos propers de Besalú. El micròfon utilitzat va ser el del telèfon mòbil, intentant que el so es sentís el més clar possible i el més aïllat de la resta d'animals que habiten per aquella zona.

Un cop totes les mostres estaven importades a l'ordinador, es van retallar cautelosament perquè tinguessin una duració exacta de 15 segons per poder-les comparar d'una manera justa amb els diferents arxius sintètics creats prèviament.

La meua hipòtesis era que el resultat de la comparació entre un grill i una cigala havia de ser lleugerament semblant, ja que són uns cants relativament semblants. El resultat que havia de ser diferent seria, per exemple, en comparar qualsevol dels insectes amb l'arxiu de referència o qualsevol de les seves modificacions.

4.3.5. Modificació del software per exportar una imatge del FFTW

Un cop vistos que els resultats no eren exactament els esperats, es va decidir reprogramar el software a les necessitats del treball, amb una idea completament diferent per a comparar els sons. Es va aprofitar tot el codi que calculava la Transformada de Fourier de l'arxiu de so, ja que era exactament el que es necessitava. Un cop s'havien recopilat tots els vectors de l'arxiu, se n'aprofitava el mòdul (veure Transformada de Fourier, pàgina 17) a cada instant per crear una imatge en forma de cascada (veure Visualització de cascada, pàgina 19). La imatge, en blanc i negre, estava formada segons el mòdul del vector (el mòdul més llarg de tots vectors serà blanc, el més petit serà negre) de tal forma que les freqüències pures queden marcades en blanc (sempre que no hi hagi soroll de fons).

Script que pinta la imatge. Utilitza totes les dades del resultat del FFTW (la variable `m_listOfaFFTResult`, consulta el valor de les posicions `i` i `j` (imaginaries i reals)) i dibuixa l'arxiu de so amb la visualització de cascada. Per a cada valor canvia el color del bolígraf (veure línia 12) segons el valor del mòdul. Els colors s'expressen amb l'espai de color HSV (veure a sota i veure 3.5). [1]

```

1 a void MainWindow::pintar(){
2 b     double config_module_window_width = 40*mouseXpress;
3 b     double config_module_window_level = -0.5+mouseYpress;
4 b
5 b     QPainter painter(&m_pixmap);
6 b     QPen myPen(Qt::black, 1, Qt::SolidLine);
7 b     for (size_t i = 0; i < m_listOfaFFTResult->size(); i++) {
8 c         for (size_t j = 0; j < m_listOfaFFTResult->at(i).size(); j++) {
9 d             std::complex<double> number = m_listOfaFFTResult->at(i)[j];
10 d             double module = std::log(std::abs(number)) /
11 d                 config_module_window_width + config_module_window_level;
12 d             double angle = (std::arg(number)+M_PI) / (2 * M_PI);
13 d             myPen.setColor(QColor::fromHsvF(0,0,module));
14 d             painter.setPen(myPen);
15 c             painter.drawLine(j,i,j,i);
16 b         }
17 b     }
18 a     m_pixmap.save(m_filename.append(".png"));

```

En la funció de pintar (veure línia 12d, pàgina 31) s'estableix que el valor del matís i de la saturació és 0, ja que els colors en blanc i negre no tenen cap d'aquests valors, i llavors s'estableix que el valor o brillantor és el **logaritme** del mòdul. S'utilitzen logaritmes per fer que els resultats tinguin uns valors més propers i visualment entenedors, ja que si no s'haguessin utilitzat la diferència dels mòduls és tan gran que seria inapreciable. D'aquesta manera, el valor que es pintarà sempre serà una escala de grisos.

Així doncs, la modificació principal del codi va ser la següent:

1. Crear dues variables per poder canviar lleugerament el valor del mòdul segons on cliques amb el ratolí (línia 10, veure pàgina 31). Veure les Figure 33 i Figure 34 per entendre el canvi.
2. Llegir tots els resultats de la transformada i dibuixar l'arxiu en forma de cascada. S'utilitzen logaritmes perquè sinó la diferència entre el valor més gran i el més petit fa que no es pugui veure bé la diferència de colors.
3. Mostrar el resultat de la imatge en pantalla perquè es pugui veure i es pugui passar a color o blanc i negre segons es desitgi. També hi ha la possibilitat d'ajustar la brillantor amb les tecles numèriques per aclarir el resultat (totes les comparacions han estat fetes amb la mateixa brillantor).
4. Guardar el resultat de la transformada de l'arxiu en format «png» per després analitzar-lo.

La imatge resultant queda estructurada de la següent forma (veure Figure 32):

- » L'eix vertical està invertit: la primera fila de píxels és el segon 0, i l'última fila és l'últim segon de l'arxiu (en el cas dels arxius sintètics, 15 segons). L'alçada de la imatge, doncs, depèn de la duració de l'arxiu⁶ que es vol comparar (un arxiu de 15 segons té una alçada de 150 píxels, un de 20 segons de 200, etc.).
- » A l'eix horitzontal hi veiem representades les freqüències. Com més properes al límit lateral de la imatge més baixa és aquesta freqüència (com més al centre més alta és).
- » Les zones fosques representen el silenci ja que el mòdul del vector és molt petit, les zones blanques representen les freqüències altes que generalment formen les ones pures.
 - » Les línies verticals grises que s'intueixen al costat de les freqüències pures són harmònics (veure Harmònics, pàgina 13).

Per facilitar la visualització es va afegir aquest petit fragment de codi per ajudar a veure millor el resultat:

⁶ El programa divideix l'arxiu en diferents blocs. Per defecte aquesta duració és de 100 mil·lisegons, per això la mida de la imatge resultant és de 150 píxels (15000 mil·lisegons / 100 mil·lisegons = 150 píxels).

Script que detecta quan el ratolí es prem i canvia el valor de les variables de posició del ratolí. A continuació crida l'ordre pintar i es torna a dibuixar el mapa de píxels amb una lleugera modificació al valor del mòdul (veure `config_module_window_width` i `config_module_window_level`). [1]

```

19 a void MainWindow::mousePressEvent(QMouseEvent *event) {
20 b     event->setAccepted(true);
21 b     mousePress = this->mapFromGlobal(QCursor::pos());
22 b     mouseXpress = double(mousePress.x()/(this->width()));
23 b     mouseYpress = double(mousePress.y()/(this->height()));
24 b     qDebug() << "es prem el ratoli a les coordenades: " << mouseXpress << ', '
        << mouseYpress;
25 b     MainWindow::pintar();
26 b     this->repaint();
27 a }

```

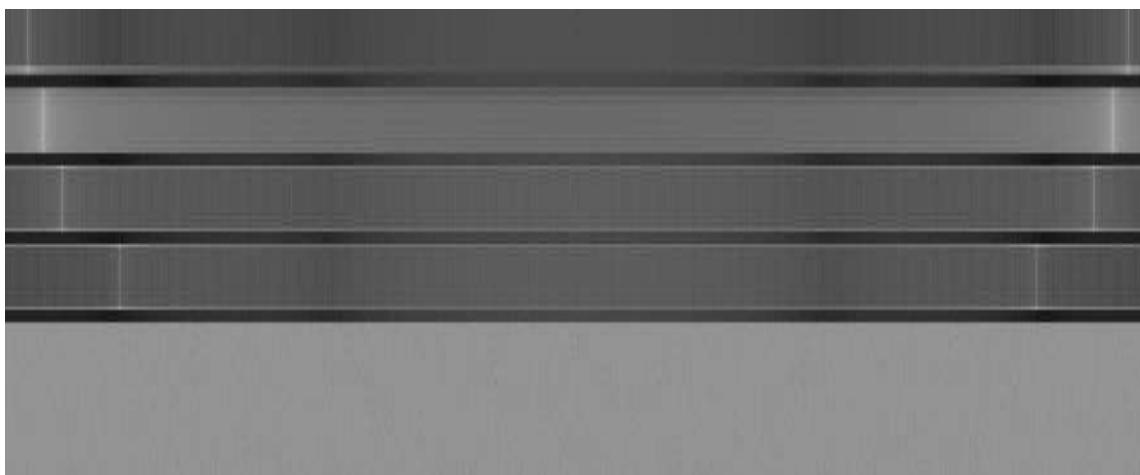


Figure 32: Representació de l'arxiu referència amb el nou software. La part superior és el segon 0, l'eix vertical queda invertit (la última fila de píxels correspon al segon 15). Les vuit ratlles blanques representen els pics de les freqüències pures que formen el senyal (es fa mirall des del centre, en realitat hi ha quatre ratlles). La part inferior grisa és el soroll, la barreja de totes les freqüències.[1]

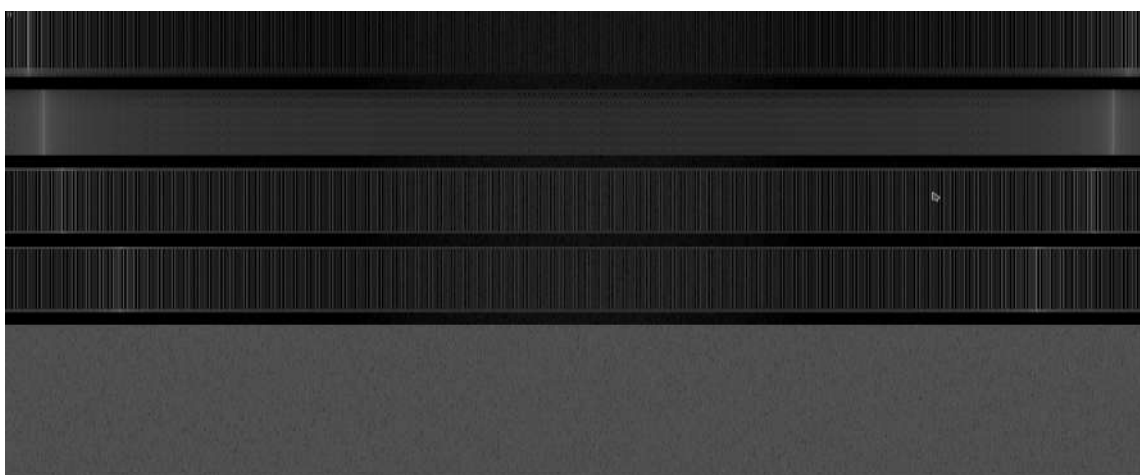


Figure 33: Representació de l'arxiu referència havent fet clic a la posició on es troba el ratolí (imatge fosca).[1]

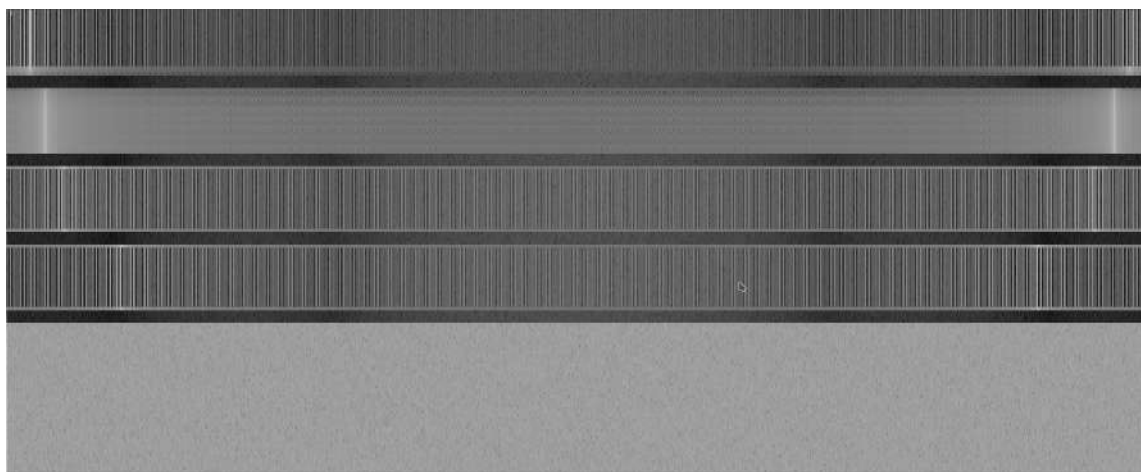


Figure 34: Representació de l'arxiu referència havent fet clic a la posició on es troba el ratolí (imatge clara).[1]

Aquest fragment de codi registra les coordenades del ratolí cada cop que es clica. Segons on es cliqui es canvia lleugerament el valor dels mòduls perquè quedin valors més alts (i per tant una imatge amb més llum, veure Figure 34) o valors més petits (la imatge resultant més fosca, veure Figure 33).

Exemple d'script de quan es prem la tecla 1. El programa col·loca el cursor a la posició (17/25, 8/25) de l'amplada de la finestra on s'executa i a continuació simula un clic per ajustar la brillantor de la imatge. [1]

```

28 a void MainWindow::keyPressEvent(QKeyEvent *keyEvent){
29 a
30 b     bool pressed = false;
31 a
32 b     if (keyEvent->key() == Qt::Key_1){
33 c         cursor().setPos((17*this->width())/25, (8*this->width())/25);
34 c         QTest::mouseClick(this, Qt::LeftButton, 0, cursor().pos());
35 b     }
36 a }

```

Aquest és un exemple de les tecles numèriques que es van programar. El que fan és simular un clic a la pantalla de forma que en tots els ordinadors en què s'executi el programa s'obté el mateix resultat (sempre que l'aspecte de la pantalla⁷ sigui 16:9). La resta de tecles numèriques estan programades amb multiplicadors diferents però essencialment és el mateix concepte.

Per acabar, només cal executar el programa amb tots els arxius d'àudio que es volen comparar per obtenir-ne totes les imatges. Amb aquest petit script de consola de Linux⁸ s'aconsegueix:

Script que executa el programa de comparació amb tots els arxius «.wav» de la carpeta.[1]

```

37 a #!/bin/sh
38 a
39 a for i in *.wav
40 a do

```

⁷ L'aspecte de la pantalla és la relació entre els píxels d'amplada i els d'alçada. La més comuna és 16:9, tot i que per exemple existeixen monitors de 4:3 o 21:9.

⁸ Sistema operatiu basat en Unix, lliure i obert. En aquest cas s'utilitza Debian, una de les moltes distribucions que hi ha disponibles gratuïtament.

```

41 b   waveform-compare $i reference.wav
42 b   rm $i.png
43 a   done

```

4.3.6. Comparació dels resultats

Un cop el software havia exportat totes les imatges dels arxius, es va utilitzar la tècnica de la resta d'imatges (veure 3.6) per aconseguir el percentatge de similitud entre les diferents imatges aconseguides, utilitzant el valor de la brillantor:

L'script general que fa la comparació de totes les imatges amb totes les imatges. [1]

```

44 a   #!/bin/bash
45 a
46 a   # In order to ignore the output of the pushd command.
47 a   pushd () {
48 b       command pushd "$@" > /dev/null
49 a   }
50 a
51 a   # In order to ignore the output of the popd command.
52 a   popd () {
53 b       command popd "$@" > /dev/null
54 a   }
55 a
56 a   echo "Comparing;Compared;Raw result"
57 a   for i in *.png
58 a   do
59 b       ./difference.sh $i #It substracts all the images in the folder for the
60 b       cp ./identify.sh ./difference/$i/identify.sh #It copies the identifying
61 b       pushd ./difference/$i/
62 b       ./identify.sh
63 b       rm identify.sh
64 b       popd
65 a   done

```

El que l'script «difference.sh» fa és restar cada una de les imatges del programa entre elles, de tal forma que si hi ha 5 imatges es resten entre elles i produeixen 25 imatges resultants. Per fer la comparació utilitzo un software lliure anomenat ImageMagick [47] que té l'opció de fer la resta entre moltes d'altres:

Script per fer la resta de totes les imatges entre totes les imatges. El que fa és crear una carpeta amb el nom de l'arxiu que es compara i llavors fa la comparació de les dues imatges i les guarda dins la carpeta recentment creada. [1]

```

66 a   #!/bin/sh
67 a   for i in *.png
68 a   do
69 b       mkdir -p ./difference/$1
70 b
71 b       # Difference of two images using imagemagick.
72 b       composite -compose difference $i $1 difference/$1/$i
73 a   done

```

Una falta d'optimització que té el programa és que en comparar totes les imatges entre elles es produeixen resultats idèntics (és el mateix restar la primera imatge amb la segona que la segona amb la primera). Per tant, la meitat dels resultats que l'script obté són correctes però duplicats.

Un cop el programa ha fet la resta d'imatges, cal obtenir-ne resultats. Una altra eina del software ImageMagick [47] que s'utilitza és la de «identify». El que fa és analitzar cada imatge i retornar-ne diferents valors. El que és interessant és el valor que s'anomena «mean», és a dir la mitjana de lluminositat de la imatge resultant de cada resta que s'ha fet. Si les dues imatges són completament iguals el resultat és una imatge completament negra (per tant el valor mitjà de la lluminositat és 0). Com més alt és el valor de la brillantor més diferents són les imatges. El valor màxim que el programa pot retornar és 255, i el mínim és 0.

Script d'identificació de dades (identify.sh). La comanda identify recull el valor de la mitjana de lluminositat. [1]

```
74 a #!/bin/sh
75 a
76 a for i in *.png
77 a do
78 b     echo -n "$(printf '%s\n' "${PWD##*/}")"
79 b     echo -n ";\n$i\";"
80 a
81 b     # Image statistics using imagemagick. We only take the mean.
82 b     identify -verbose $i | grep "mean:" | cut -d ':' -f 2 | cut -d ' ' -f 2
83 a done
```

Aquest script està programat per poder guardar els resultats en format csv (comma-separated values) [48]. Aquest format és pràctic perquè permet importar les dades en un full de càlcul i faciliten la creació de taules per veure i entendre els resultats obtinguts.

4.3.7. Organització i anàlisis dels resultats

Un cop el programa s'ha executat es crea l'arxiu de format «csv» amb totes les dades. Aquest arxiu segueix la següent estructura:

Taula 4: Exemple del format en el que es reben les dades per col·locar al full de càlcul (no són les dades reals). [1]

Comparing	Compared	Raw result
Grill-1.wav	Grill-1.wav	0
Grill-1.wav	Cigala-1.wav	15
Grill-1.wav	Cigala-2.wav	23
Cigala-1.wav	Grill-1.wav	15
Cigala-1.wav	Cigala-1.wav	0
Cigala-1.wav	Cigala-2.wav	6

A la primera columna s'hi col·loca l'arxiu que es compara primer mentre que a la segona s'hi col·loca l'arxiu amb què es compara aquest primer arxiu. La columna de la dreta és el resultat de similitud que

retorna l'script de «identify.sh» (veure a sobre). Cal destacar la segona i la quarta fila de la taula, ja que el resultat de la comparació és el mateix perquè és la mateixa resta deguda a la falta d'optimització ja mencionada.

» Tractament de les dades

El resultat que ens dona el programa no és suficientment clar com per treballar-hi. Per acabar d'aprofitar les dades i crear representacions visuals es converteixen els resultats obtinguts en un percentatge sobre 1. D'aquesta manera els arxius que són idèntics tenen un 1 i els que són diferents tenen valors més propers al 0. Per aconseguir aquest resultat s'aplica la següent fórmula a la columna de «Raw Result» (veure 4):

$$f(x) = \frac{256 - x}{256}$$

El resultat d'aquesta fórmula llavors es col·loca a una columna al costat per poder utilitzar les dades de les dues columnes en cas que es necessités.

Per poder preparar la creació de taules dinàmiques amb el full de càlcul el que es fa és crear agrupacions dels diferents arxius sintètics i reals. Quan es van crear els arxius es va utilitzar una nomenclatura específica («grup»-«variació»). Mitjançant una altra fórmula del full del càlcul es retalla el nom del grup de la primera columna i el grup de la segona, de tal manera que el resultat abans de comparar els números és el següent (veure la diferència entre la 4 i la 5):

Taula 5: Exemple del format en el que queden les dades per estudiar-les.[1]

Grup 1	Grup 2	Comparing	Compared	Raw result	Dades
Grill	Grill	Grill-1.wav	Grill-1.wav	0	1
Grill	Cigala	Grill-1.wav	Cigala-1.wav	15	≈0,94
Grill	Cigala	Grill-1.wav	Cigala-2.wav	23	≈0,91
Cigala	Grill	Cigala-1.wav	Grill-1.wav	15	≈0,94
Cigala	Cigala	Cigala-1.wav	Cigala-1.wav	0	1
Cigala	Cigala	Cigala-1.wav	Cigala-2.wav	6	≈0,98

Un cop s'ha fet aquest tractament s'obté una llista de dades amb 6 columnes. Llavors es crea una matriu de correlació (veure 3.7) automàticament amb les taules dinàmiques. En aquestes taules es col·loca a l'eix y el «Grup 1» i el «Comparing» i a l'eix x el «Grup 2» i el «Compared». D'aquesta forma es pot controlar quins grups d'arxius es vol que es comparin amb aquella taula amb pocs clics. D'altra banda, en el camp de les dades de la taula se'n col·loquen dos: el valor de la comparació sobre 1 i, en el cas que es comparin els grups directament, també s'hi afegeix la desviació estàndard per veure si les dades de dins de cada grup són molt disperses o no (veure els resultats al següent apartat).

Treball de recerca • Comparació d'ones

Per veure els resultats de forma clara s'utilitzen les escales de colors. Aquesta escala de colors es forma amb el valor mínim de 0,75 i el valor màxim de 1. Això significa que qualsevol valor inferior a 0,75 estarà pintat del mateix color, però com que hi ha pocs casos i són de comparacions que tenen poc sentit (per exemple comparar soroll amb silenci), ajuda a que els altres resultats que realment són interessants es vegin més clars.

4.4. Resultats de les diferents proves

Les diferents matrius de correlació a continuació són els resultats dels diferents arxius sintètics de les proves. S'han utilitzat uns colors diferents per remarcar encara més la diferència, amb l'escala de color on el gris representa un 80% de similitud i el blau representa un 100%.

Si ens fixem més detalladament en els resultats es poden observar varis fets interessants. Una de

	atenuation		cicada		contracted		expanded		grill		linein		noise		reference		shift		silence		speakerin100cm		speakerin20cm		speakerin50cm	
	Mitjana - Result	Desviació - Result	Mitjana - Result	Desviació - Result	Mitjana - Result	Desviació - Result	Mitjana - Result	Desviació - Result	Mitjana - Result	Desviació - Result	Mitjana - Result	Desviació - Result	Mitjana - Result	Desviació - Result	Mitjana - Result	Desviació - Result	Mitjana - Result	Desviació - Result	Mitjana - Result	Desviació - Result	Mitjana - Result	Desviació - Result	Mitjana - Result	Desviació - Result	Mitjana - Result	Desviació - Result
Group 1																										
atenuation	0,973	0,016	0,839	0,007	0,83	0,035	0,874	0,018	0,844	0,022	0,916	0,016	0,832	0,03	0,96	0,018	0,959	0,017	0,802	0,017	0,912	0,009	0,924	0,008	0,922	0,005
cicada	0,847	0,006	0,947	0,025	0,836	0,011	0,858	0,013	0,911	0,022	0,876	0,014	0,823	0,034	0,836	0,005	0,839	0,005	0,775	0,034	0,885	0,015	0,884	0,01	0,891	0,01
contracted	0,83	0,035	0,854	0,011	0,899	0,049	0,842	0,021	0,842	0,013	0,854	0,038	0,825	0,03	0,834	0,04	0,835	0,04	0,792	0,022	0,858	0,032	0,855	0,033	0,857	0,03
expanded	0,874	0,018	0,856	0,019	0,842	0,021	0,898	0,043	0,855	0,026	0,886	0,033	0,831	0,04	0,871	0,028	0,873	0,028	0,801	0,036	0,89	0,028	0,89	0,026	0,893	0,022
grill	0,854	0,016	0,912	0,022	0,833	0,017	0,86	0,019	0,926	0,034	0,862	0,029	0,787	0,042	0,834	0,015	0,836	0,016	0,831	0,025	0,869	0,031	0,874	0,027	0,881	0,027
linein	0,916	0,016	0,866	0,016	0,854	0,038	0,886	0,033	0,85	0,036	0,982	0,012	0,91	0,024	0,934	0,008	0,936	0,007	0,729	0,008	0,966	0,003	0,96	0,009	0,958	0,002
noise	0,832	0,03	0,811	0,036	0,825	0,03	0,831	0,04	0,774	0,049	0,91	0,024	0,965	0,019	0,864	0,024	0,865	0,024	0,641	0,023	0,903	0,022	0,887	0,025	0,882	0,023
reference	0,96	0,018	0,824	0,007	0,834	0,04	0,871	0,028	0,821	0,023	0,934	0,008	0,864	0,024	1	0	0,992	0,004	0,769	0	0,917	0,003	0,928	0,002	0,921	0,002
shift	0,959	0,017	0,827	0,007	0,835	0,04	0,873	0,028	0,824	0,023	0,936	0,007	0,865	0,024	0,992	0,004	0,992	0,007	0,768	0,001	0,919	0,002	0,931	0,002	0,923	0,002
silence	0,802	0,017	0,794	0,032	0,792	0,022	0,801	0,036	0,84	0,03	0,729	0,008	0,641	0,023	0,769	0	0,768	0,001	1	0	0,736	0,006	0,751	0,011	0,757	0,002
speakerin100cm	0,912	0,009	0,877	0,017	0,858	0,032	0,89	0,028	0,859	0,037	0,966	0,003	0,903	0,022	0,917	0,003	0,919	0,002	0,736	0,006	0,978	0,009	0,964	0,006	0,965	0,003
speakerin20cm	0,924	0,008	0,876	0,012	0,855	0,033	0,89	0,026	0,865	0,033	0,96	0,009	0,887	0,025	0,928	0,002	0,931	0,002	0,751	0,011	0,964	0,006	0,974	0,012	0,967	0,003
speakerin50cm	0,922	0,005	0,885	0,012	0,857	0,03	0,893	0,022	0,873	0,033	0,958	0,002	0,882	0,023	0,921	0,002	0,923	0,002	0,757	0,002	0,965	0,003	0,967	0,003	0,982	0,008

Figure 35: Matriu de correlació dels diferents grups de les proves. Hi ha dos resultats per cada comparació: la mitjana de similitud entre les diferents proves del grup (l'escala amb color de l'esquerra) i la desviació estàndard (el valor de la dreta). [1]

les preguntes que sorgien mentre es comparaven els diferents arxius era què tindria un efecte més gran sobre l'arxiu referència d'entre l'atenuació, la compressió i l'expansió. Si creem una matriu amb els resultats d'aquestes tres proves (veure Figure 35) podem deduir varies coses:

Si ens fixem amb la columna del grup «reference», que conté només l'arxiu original, podem veure quin efecte han tingut les diferents modificacions amb l'arxiu inicial. El valor de la casella de la similitud, però, correspon a la mitjana de totes les comparacions, per agrupar tant els exemples menys exagerats

Treball de recerca • Comparació d'ones

i els més exagerats. Tot i així, si ens fixem amb el valor de l'atenuació (0,96), és considerablement més gran que el de l'expansió (0,871), que també és més gran que el de la contracció (0,834). Per tant, l'atenuació és, d'entre els tres exemples a comparar, el que causa menys impacte amb el resultat. Això és degut al logaritme que hi ha en el codi, que redueix considerablement el valor del mòdul dels vectors i per tant redueix l'impacte del volum de les mostres.

Si mirem la Figure 35, veiem que el resultat de la comparació entre els mateixos grups (atenuació amb atenuació, contracció amb contracció, etc.) no és 1. Això és degut a que en aquesta matriu es contemplen tots els resultats d'una forma agrupada: la casella d'atenuació inclou la mitjana de tots els resultats d'atenuació. Per tant, el valor de la casella de l'atenuació amb l'atenuació, que és 0,973, equival a la mitjana dels resultats en fer la matriu d'atenuació amb atenuació:

Comparing	ate-nua-tion-02db	ate-nua-tion-04db	ate-nua-tion-06db	ate-nua-tion-08db	ate-nua-tion-10db	ate-nua-tion-12db	ate-nua-tion-14db	ate-nua-tion-16db	ate-nua-tion-18db	ate-nua-tion-20db
atenuation-02db	1	0,987	0,981	0,975	0,968	0,962	0,956	0,95	0,944	0,938
atenuation-04db	0,987	1	0,987	0,981	0,975	0,968	0,962	0,956	0,95	0,944
atenuation-06db	0,981	0,987	1	0,987	0,981	0,975	0,969	0,963	0,957	0,951
atenuation-08db	0,975	0,981	0,987	1	0,987	0,981	0,975	0,969	0,963	0,957
atenuation-10db	0,968	0,975	0,981	0,987	1	0,987	0,981	0,975	0,969	0,963
atenuation-12db	0,962	0,968	0,975	0,981	0,987	1	0,986	0,981	0,975	0,969
atenuation-14db	0,956	0,962	0,969	0,975	0,981	0,986	1	0,986	0,98	0,975
atenuation-16db	0,95	0,956	0,963	0,969	0,975	0,981	0,986	1	0,986	0,98
atenuation-18db	0,944	0,95	0,957	0,963	0,969	0,975	0,98	0,986	1	0,986
atenuation-20db	0,938	0,944	0,951	0,957	0,963	0,969	0,975	0,98	0,986	1

Figure 36: Matriu de correlació dels arxius d'atenuació comparats entre ells. Si s'agafen tots els valors d'aquesta matriu i se'n fa la mitjana, s'obté el valor de la casella esmentat a dalt. [1]

D'aquesta manera s'obté en una taula molt més reduïda molta més informació. Per ajudar a entendre tota la informació de la taula també s'hi ha afegit la desviació estàndard (veure 3.7.1), que ens permet veure la dispersió dels resultats, és a dir, la diferència que hi ha entre els diferents valors de la matriu de cada grup. Si observem la Figure 36, veiem que el valor més petit és 0,938 i el més gran és 1, resultant en una desviació estàndard molt petita (0,016) ja que els valors s'allunyen poc de la mitjana (0,937).

Si ens fixem amb el soroll (veure Figure 37) veiem que al principi la diferència entre un 1% de soroll i un 5% és molt més gran que entre un 95% i un 100%. Això és degut a que en un punt el soroll «tapa» completament el senyal original i és gairebé irreconeixible. D'altra banda, la comparació amb el silenci és el resultat més baix de totes les dades recollides perquè és compara la suma de totes les freqüències amb l'absència de freqüències, retornant la diferència més gran de totes les comparacions.

	noi-se-0.01	noi-se-0.05	noi-se-0.10	noi-se-0.15	noi-se-0.20	noi-se-0.25	noi-se-0.30	noi-se-0.35	noi-se-0.40	noi-se-0.45	noi-se-0.50	noi-se-0.55	noi-se-0.60	noi-se-0.65	noi-se-0.70	noi-se-0.75	noi-se-0.80	noi-se-0.85	noi-se-0.90	noi-se-0.95	noi-se-1.00	
Comparing																						
reference-	0,935	0,908	0,894	0,885	0,878	0,873	0,869	0,865	0,862	0,859	0,856	0,854	0,852	0,85	0,848	0,846	0,845	0,843	0,842	0,84	0,839	
silence-	0,71	0,683	0,669	0,66	0,654	0,649	0,645	0,641	0,638	0,636	0,633	0,631	0,629	0,627	0,625	0,624	0,622	0,621	0,62	0,618	0,617	

Figure 37: Matriu de correlació dels arxius referència i silenci amb els del soroll. [1]

	reference-0.25Wave	shift-0.50Wave	shift-0.75Wave	shift-1.00Wave	silence-
Comparing					
reference-	1	0,995	0,984	0,995	0,769
shift-0.25Wave	0,995	1	0,984	0,994	0,768
shift-0.50Wave	0,984	0,984	1	0,984	0,771
shift-0.75Wave	0,995	0,994	0,984	1	0,768
shift-1.00Wave	0,995	0,994	0,984	0,994	1
silence-	0,769	0,768	0,771	0,768	1

Figure 38: Matriu de correlació dels arxius de canvi de fase. [1]

Si ens fixem en les diferents proves de la fase (veure Figure 38), veiem que el resultat de la comparació és molt gran. Això és degut a que la durada del blocs amb els quals el programa fa les comparacions és massa gran com perquè la fase tingui un efecte (per defecte és de 100 mil·lisesegons). Com que les freqüències són altes, aquests mil·lisesegons no permeten que la fase tingui un impacte més gran.

Un últim dubte que queda per resoldre és perquè el valor més baix és un 61,7%, considerablement alt. La resposta és que perquè els dos arxius fossin completament diferents, és a dir un 0% de correspondència, haurien de ser inversos. Llavors la imatge resultant seria completament blanca i s'aconseguiria aquest resultat, que és gairebé impossible a menys que estigui preparat.

En últim lloc, es vol destacar la imatge resultant de l'arxiu de silenci (sense fer la comparació, és a dir l'espectre de la imatge que retorna el software abans de fer la resta). Tot i que l'arxiu és silenci pur, l'ordinador necessita donar un valor mínim a l'amplitud de l'arxiu per poder-lo exportar, ja que sinó no tindria duració. El que es fa a continuació és atribuir-li a diferents freqüències molt baixes (que nosaltres els humans no podem escoltar) una amplitud molt i molt baixa però que tot i així té valor. Aquest valor tan ínfim és el que veiem reflectit en les diferents taques que es formen a la imatge de l'arxiu silenci, que semblen inexplicables a primera vista però el resultat del programa és correcte perquè és cert que l'amplitud en aquelles ones té un valor (veure Figure 39).



Figure 39: Representació de l'arxiu silenci abans de fer qualsevol resta. Destacar les diferents «taques» que es poden apreciar al llarg de la imatge, resultat de les aproximacions que l'ordinador fa. [1]

	cica- da-01	cica- da-02	cica- da-03	cica- da-04	cica- da-05	cica- da-06	cica- da-07	cica- da-08	cica- da-09	cica- da-10	cica- da-11	cica- da-12	cica- da-13	cica- da-14	grill- 01	grill- 02	grill- 03	grill- 04	grill- 05	grill- 06	grill- 07	grill- 08	grill- 09	grill- 10
Comparing																								
cicada-01	1	0,969	0,967	0,964	0,967	0,964	0,937	0,935	0,962	0,963	0,962	0,91	0,921	0,914	0,924	0,916	0,885	0,885	0,909	0,907	0,875	0,878	0,906	0,906
cicada-02	0,969	1	0,967	0,964	0,966	0,963	0,938	0,937	0,962	0,963	0,961	0,913	0,924	0,917	0,926	0,917	0,888	0,888	0,909	0,909	0,877	0,88	0,906	0,908
cicada-03	0,967	0,967	1	0,968	0,969	0,963	0,944	0,943	0,958	0,959	0,957	0,923	0,934	0,927	0,934	0,925	0,899	0,899	0,917	0,913	0,887	0,89	0,911	0,912
cicada-04	0,964	0,964	0,968	1	0,968	0,962	0,947	0,946	0,955	0,956	0,954	0,926	0,937	0,931	0,933	0,927	0,903	0,904	0,919	0,915	0,891	0,893	0,911	0,911
cicada-05	0,967	0,966	0,969	0,968	1	0,963	0,944	0,942	0,959	0,959	0,958	0,92	0,931	0,924	0,932	0,922	0,896	0,896	0,915	0,911	0,885	0,888	0,91	0,911
cicada-06	0,964	0,963	0,963	0,962	0,963	1	0,939	0,938	0,956	0,957	0,956	0,912	0,924	0,917	0,922	0,914	0,888	0,888	0,912	0,907	0,877	0,88	0,905	0,905
cicada-07	0,937	0,938	0,944	0,947	0,944	0,939	1	0,953	0,927	0,929	0,925	0,945	0,951	0,947	0,94	0,936	0,928	0,929	0,924	0,91	0,912	0,915	0,907	0,906
cicada-08	0,935	0,937	0,943	0,946	0,942	0,938	0,953	1	0,925	0,927	0,924	0,947	0,953	0,949	0,937	0,939	0,93	0,931	0,924	0,911	0,913	0,915	0,907	0,907
cicada-09	0,962	0,962	0,958	0,955	0,959	0,956	0,927	0,925	1	0,97	0,97	0,899	0,91	0,903	0,916	0,905	0,874	0,874	0,902	0,901	0,864	0,867	0,903	0,903
cicada-10	0,963	0,963	0,959	0,956	0,959	0,957	0,929	0,927	0,97	1	0,969	0,902	0,913	0,906	0,918	0,907	0,877	0,877	0,905	0,902	0,867	0,87	0,904	0,905
cicada-11	0,962	0,961	0,957	0,954	0,958	0,956	0,925	0,924	0,97	0,969	1	0,898	0,91	0,903	0,915	0,904	0,873	0,873	0,902	0,9	0,863	0,866	0,902	0,903
cicada-12	0,91	0,913	0,923	0,926	0,92	0,912	0,945	0,947	0,899	0,902	0,898	1	0,97	0,972	0,949	0,95	0,96	0,96	0,93	0,91	0,938	0,939	0,909	0,907
cicada-13	0,921	0,924	0,934	0,937	0,931	0,924	0,951	0,953	0,91	0,913	0,91	0,97	1	0,972	0,951	0,95	0,953	0,953	0,932	0,915	0,932	0,933	0,914	0,912
cicada-14	0,914	0,917	0,927	0,931	0,924	0,917	0,947	0,949	0,903	0,906	0,903	0,972	0,972	1	0,95	0,95	0,957	0,957	0,931	0,912	0,935	0,936	0,912	0,91
grill-01	0,924	0,926	0,934	0,933	0,932	0,922	0,94	0,937	0,916	0,918	0,915	0,949	0,951	0,95	1	0,947	0,939	0,939	0,93	0,91	0,924	0,928	0,913	0,91
grill-02	0,916	0,917	0,925	0,927	0,922	0,914	0,936	0,939	0,905	0,907	0,904	0,95	0,95	0,95	0,947	1	0,945	0,945	0,929	0,909	0,929	0,932	0,908	0,906
grill-03	0,885	0,888	0,899	0,903	0,896	0,888	0,928	0,93	0,874	0,877	0,873	0,96	0,953	0,957	0,939	0,945	1	0,97	0,927	0,895	0,944	0,945	0,896	0,894
grill-04	0,885	0,888	0,899	0,904	0,896	0,888	0,929	0,931	0,874	0,877	0,873	0,96	0,953	0,957	0,939	0,945	0,97	1	0,928	0,896	0,944	0,944	0,897	0,894
grill-05	0,909	0,909	0,917	0,919	0,915	0,912	0,924	0,924	0,902	0,905	0,902	0,93	0,932	0,931	0,93	0,929	0,927	0,928	1	0,898	0,909	0,914	0,899	0,899
grill-06	0,913	0,914	0,917	0,918	0,915	0,913	0,913	0,913	0,906	0,908	0,906	0,91	0,916	0,912	0,911	0,908	0,892	0,892	0,896	1	0,879	0,882	0,944	0,944
grill-07	0,875	0,877	0,887	0,891	0,885	0,877	0,912	0,913	0,864	0,867	0,863	0,938	0,932	0,935	0,924	0,929	0,944	0,944	0,909	0,882	1	0,936	0,88	0,878
grill-08	0,878	0,88	0,89	0,893	0,888	0,88	0,915	0,915	0,867	0,87	0,866	0,939	0,933	0,936	0,928	0,932	0,945	0,944	0,914	0,884	0,936	1	0,882	0,881
grill-09	0,913	0,913	0,916	0,916	0,916	0,912	0,911	0,911	0,91	0,911	0,909	0,911	0,916	0,913	0,915	0,909	0,894	0,894	0,898	0,944	0,879	0,881	1	0,957
grill-10	0,913	0,914	0,916	0,915	0,916	0,912	0,909	0,909	0,91	0,911	0,91	0,908	0,913	0,91	0,911	0,906	0,891	0,891	0,897	0,944	0,876	0,879	0,957	1

Figure 40: Matriu de correlació dels grills amb les cigales. [1]

4.5. Aplicacions a la vida real

Aquest software permet comparar qualsevol tipus d'arxiu de so, ja sigui el cant d'un insecte o tota una cançó. Pot servir per a comparar si dues cançons aparentment iguals ho són o no, o per saber la qualitat que es perd en comprimir un arxiu de so.

L'exemple a continuació és el resultat de la diferència entre una cançó i la mateixa cançó però en versió instrumental, és a dir sense veu. Per aquest exemple, s'han agafat dues cançons sense copyright que tenen tant una versió amb lletra com una versió idèntica sense («Linked» de Jim Yosef i Anna Yvette, del 2017 [49] [50]). Després es van col·locar a l'Audacity [37] i es van retallar perquè tinguessin exactament la mateixa duració. A continuació es va executar tot el codi del programa amb els arxius i el resultat va ser el següent:

Taula 6: Dades recollides de les dues cançons.[1]

Comparing	Compared	Raw result
Linked_instrumental-.png	Linked-.png	9,0387
Linked_instrumental-.png	Linked_instrumental-.png	0
Linked-.png	Linked-.png	0
Linked-.png	Linked_instrumental-.png	9,0387

Taula 7: Dades tractades i apunt per fer la matriu de correlació.[1]

Group 1	Group 2	Comparing	Compared	Raw result	Result
Linked_instrumental	Linked	Linked_instrumental-.png	Linked-.png	9,0387	≈0,965
Linked_instrumental	Linked_instrumental	Linked_instrumental-.png	Linked_instrumental-.png	0	1
Linked	Linked	Linked-.png	Linked-.png	0	1
Linked	Linked_instrumental	Linked-.png	Linked_instrumental-.png	9,0387	≈0,965

Després de tot el procés i la interpretació de les dades es pot veure que les dues cançons són iguals en un 96,5 % aproximadament. Degut a que no tota la cançó té lletra i moltes parts són compartides s'obté un resultat molt elevat.

Comparing	Linked-.png	Linked_instrumental-.png
Linked-.png	1	0,964692578125
Linked_instrumental-.png	0,964692578125	1

Figure 41: Matriu de correlació de les dues cançons. Només hi ha 4 valors, que estan repetits entre ells. El que és interessant és el de la comparació entre les dues cançons, que és d'un 96,5 % de similitud aproximadament. [1]

No s'han adjuntat les imatges de la Transformada de Fourier perquè no es poden interpretar ja que hi ha massa freqüències sonant al mateix moment i no es pot veure clarament en la imatge. No obstant, aquestes imatges es podran trobar als arxius de l'annexe.

5. Conclusions

Després de diverses proves i tests s'ha comprovat que l'eina que s'ha programat funciona correctament i com s'esperava, tant en un espai preparat amb unes proves sintètiques adaptades com en la vida real amb un exemple que qualsevol aficionat del so es podria plantejar. El software, però, està limitat per una sèrie de condicions de programació que no afecten als resultats però sí a la forma en què se li han d'entregar els arxius a comparar (veure a sota).

En primer lloc, el percentatge de comparació de tots els arxius mai ha estat inferior al 50%. Això és degut a la relativitat de la mesura: tots els valors que el programa executarà en condicions normals seran entre 0,5 i 1 perquè l'única forma en què el valor pugui ser inferior a 0,5 és quan es compara el mateix arxiu (o semblant) però **invertit**⁹. Aquest fet es deu a que l'única forma d'aconseguir una imatge resultant blanca és si els valors de la resta estan el més allunyats possible (255-0 en el cas de restar el blanc pur amb el negre pur), donant un valor molt alt a la imatge resultant i, per tant, que aquesta tingui molta lluminositat. Per tant, pel que als humans seria una diferència del 100%, per a l'ordinador és del 50%, cosa que pot generar confusió.

Un altre fet que cal remarcar és la naturalesa del soroll. Tot i que per a l'oïda humana el soroll sempre se sent de la mateixa manera (i quan pensem en soroll sempre sentim el mateix so), per a l'ordinador, cada soroll generat és diferent perquè, tot i que hi haurà les mateixes freqüències (totes), estaran recol·locades de diferents formes i tindran amplituds diferents i, per tant, si comparem dos arxius de soroll totalment **diferents** però de mateixa duració, el resultat **no** podrà ser 1. Si, per exemple, es resten dos sorolls aleatoris, el resultat esperat ha de ser del voltant del 50%, per tant la imatge de la resta quedaria de color «gris».

En l'apartat de resultats es posa èmfasi en les troballes més interessants que es produeixen a l'hora de comparar els diferents sons sintètics generats. La més destacable és que si ens fixem amb la matriu dels grups (veure Figure 35), veiem que el resultat a l'hora de comparar els grups amb ells mateixos són els més alts de tots (es pot dir que hi segueix havent el valor més alt de cada columna a la diagonal). Això és degut a que com que s'aplica la mateixa modificació a tots els arxius, el resultat de la correlació és molt elevat.

Tanmateix, quan ens fixem amb els grills i les cigales veiem que obtenim els valors de correlació més alts en comparar les cigales entre elles i els grills entre ells. No obstant, també es veuen uns valors molt alts en comparar, per exemple, els grills 1,2,3 i 4 amb la cigala 14 (escoltar els àudios de l'annexe i veure Figure 40). És curiós perquè, segons la oïda humana, els cants de les cigales i el dels grills també són molt semblants, per tant, tot i que el programa és objectiu s'assembla a la nostra percepció dels di-

⁹ La inversió de color és quan es resta el valor de color (de lluminositat en el cas del treball) del valor màxim que pot arribar a tenir, normalment 255. Per tant, si una imatge té un valor de 255 de lluminositat, el seu invers serà 0 (255-255=0). També es pot aplicar a imatges amb color. [51]

ferents sons que podem trobar. Un altre fet que observem relacionat amb els insectes és que no trobem la mateixa correlació entre les diferents gravacions de les cigales. Veiem, per exemple, que les cigales 1, 2, 3, 4 i 5 tenen molta semblança perquè van estar gravades a la mateixa zona. Si comparem qualsevol dels animals anteriors amb una altra cigala diferent (per exemple la 12, 13 i 14), veiem que el resultat és més baix, tot i ser el mateix animal. D'aquest fet podem deduir que això és degut a que les gravacions han estat fetes a llocs diferents i a hores diferents; per tant, tot i que l'animal és el mateix, la gravació és diferent i la distància entre l'insecte i el micròfon també ho era, modificant la correlació.

En últim lloc, cal remarcar les limitacions que el programa encara té però que són solucionables:

- » En primer lloc, els arxius han de tenir exactament la mateixa duració, ja que la resolució de la imatge del programa depèn de la llargada i per a fer una resta acurada cal que tinguin la mateixa dimensió.
- » Un altre inconvenient és que tots els arxius que es volen comparar han d'estar en format «wav» (que és una llista de mostres sense comprimir), però hi ha software disponible tant en línia com en local que permet fer la conversió senzillament.
- » D'altra banda, també cal remarcar que aquest programa ignora la fase (deliberadament fet així), com s'ha vist en els resultats. Tot i que la fase es podria emprar en la comparació (és el valor de la part imaginària del vector de la transformada), els humans no tenim el costum de comparar a nivell de fase dos sons monofònics, sempre quan pensem en una comparació, els sons estan alineats. Els únics moments on sembla que el cervell pot tenir en compte la fase és si s'escoltessin els dos sons fora de fase, un amb cada orella. És gràcies a això, i al volum amb què sentim, que podem determinar la direcció dels sons que sentim (en freqüències baixes).

En conclusió, el software que s'ha creat permet fer la comparació amb uns arxius que tenen preparació prèvia. Cal remarcar que el programari està pensat per comparar sons i no soroll, i que tampoc té en compte la fase perquè la mínima diferència hauria provocat un resultat molt diferent. El projecte queda obert a disposició de qualsevol usuari que desitgi seguir millorant aquest software i convertir-lo en una eina més útil i multiplataforma, ja que està disponible sota GPLv2¹⁰.

¹⁰ Llicència que estableix que el programari és codi lliure. És GPLv2 i no 3 perquè el projecte del qual aquest projecte està derivat no permet usar la v3 de forma opcional. [52]

Bibliografia

- 1: JUANOLA, Aniol, Font pròpia, 2020, Besalú
- 2: Et al. *Waveform* [en línia]. , . < <https://en.wikipedia.org/wiki/Waveform> >.
- 3: TechTerms. *Waveform definition* [en línia]. , . < <https://techterms.com/definition/waveform> >.
- 4: Soundbridge. *What are waveforms and how do they work?* [en línia]. , . < <https://soundbridge.io/what-are-waveforms-how-they-work/> >.
- 5: COMEAU, Josh. *Let's learn about waveforms* [en línia]. , 2018. < <https://pudding.cool/2018/02/waveforms/> >.
- 6: MERCADÉ, Joan; SERRA, Salvador; ARMENGOL, Montserrat. (2009). *Física Batxillerat 2. (1a)*. Aravaca, Madrid: McGraw-Hill. (sèrie Fluvià).
- 7: TIPLER, Paul Allen; MOSCA, Gene. (2010). *Física para la ciencia y la tecnología. (6a)*. Alabama: Editorial Reverté. (Volumen 2: Electricidad y magnetismo/Luz).
- 8: Et al. *Frequency* [en línia]. , . < <https://en.wikipedia.org/wiki/Frequency> >.
- 9: FEYNMAN, Richard P. (1964). *The Feynman Lectures on Physics. (11a)*. Estats Units d'Amèrica: Addison-Wesley. (Volum 1).
- 10: Et al. *Hertz* [en línia]. , . < <https://en.wikipedia.org/wiki/Hertz> >.
- 11: Fluke. *What is frequency?* [en línia]. , . < <https://www.fluke.com/en-us/learn/best-practices/measurement-basics/electricity/what-is-frequency> >.
- 12: All about circuits. *AC Phase* [en línia]. , . < <https://www.allaboutcircuits.com/textbook/alternating-current/chpt-1/ac-phase/> >.
- 13: Et al. *Fase* [en línia]. , . < [https://ca.wikipedia.org/wiki/Fase_\(ona\)](https://ca.wikipedia.org/wiki/Fase_(ona)) >.
- 14: Researchgate. *Alternation* [en línia]. , . < https://www.researchgate.net/figure/Wavelength-amplitude-and-frequency-of-an-electromagnetic_fig21_303377557 >.
- 15: Techplayon. *Wavelength, Frequency, Amplitude and phase - defining Waves!* [en línia]. , . < <http://www.techplayon.com/wavelength-frequency-amplitude-phase-defining-waves/> >.
- 16: H, Will. *Speeds in different mediums* [en línia]. , . < <https://alexawillsound.weebly.com/speeds-in-different-mediums.html#> >.
- 17: Light Colour Vision. *Wavelength and Speed of Light in Different Media* [en línia]. , . < https://lightcolourvision.org/resource-library/speed-and-wavelength-of-light-in-different-materials/?attribute_pa_file-format=jpg >.
- 18: SZIELENSKI, Jeff. *dB vs. dBW vs. dBm Tutorial* [en línia]. , 2015. < <http://www.szienlenski.com/2015/12/db-vs-dbw-vs-dbm-tutorial.html> >.
- 19: Et al. *Harmònic* [en línia]. , . < <https://ca.wikipedia.org/wiki/Harm%C3%B2nic> >.
- 20: Electronics Tutorials. *Harmonics* [en línia]. , . < <https://www.electronics-tutorials.ws/accircuits/harmonics.html> >.
- 21: Et al. *Còclea* [en línia]. , . < <https://ca.wikipedia.org/wiki/C%C3%B2clea> >.
- 22: Centro auditivo Valencia. *Tonotoía, la percepción de las frecuencias* [en línia]. , 2015. < <https://www.centroauditivo-valencia.es/2015/09/23/tonotop%C3%ADa-la-percepci%C3%B3n-de-las-frecuencias/> >.
- 23: Et al. *Frecuencia de muestreo* [en línia]. , . < https://es.wikipedia.org/wiki/Frecuencia_de_muestreo >.
- 24: HyperPhysics. *Microphones* [en línia]. , . < <http://hyperphysics.phy-astr.gsu.edu/hbase/Audio/mic2.html> >.
- 25: Et al. *Micròfon* [en línia]. , . < <https://ca.wikipedia.org/wiki/Micr%C3%B2fon> >.
- 26: Et al. *Audio bit depth* [en línia]. , . < https://en.wikipedia.org/wiki/Audio_bit_depth >.
- 27: Mastering The Mix. *Sample rates and bit depth... In a nutshell* [en línia]. , . < <https://www.masteringthemix.com/blogs/learn/113159685-sample-rates-and-bit-depth-in-a-nutshell> >.
- 28: DaleMultimedia. *Digital Sampling* [en línia]. , . < <https://m.themusictelegraph.com/247> >.

- 29: Et al. *Muestreo digital* [en línia]. , . < https://es.wikipedia.org/wiki/Muestreo_digital >.
- 30: Et al. *Frecuencia de muestreo* [en línia]. , . < https://es.wikipedia.org/wiki/Frecuencia_de_muestreo >.
- 31: Toms Rays Audio Mastering. *Sample Rate and Bit Depth* [en línia]. , 2017. < <https://tomsrayaudiomastering.com/sample-rate-and-bit-depth/> >.
- 32: Et al. *Nyquist–Shannon sampling theorem* [en línia]. , . < [https://en.wikipedia.org/wiki/Nyquist %E2%80%93Shannon_sampling_theorem](https://en.wikipedia.org/wiki/Nyquist_%E2%80%93Shannon_sampling_theorem) >.
- 33: Pro Audio Files. *Digital Audio 101: What is Aliasing?* [en línia]. , 2014. < <https://theproaudiofiles.com/digital-audio-aliasing/> >.
- 34: Et al. *Transformada de Fourier* [en línia]. , . < https://ca.wikipedia.org/wiki/Transformada_de_Fourier >.
- 35: Et al. *Fast Fourier transform* [en línia]. , . < <https://aavos.eu/glossary/fourier-transform/> >.
- 36: Et al. *Radio noise* [en línia]. , . < https://en.wikipedia.org/wiki/Radio_noise >.
- 37: MAZZONI, Dominic; DANNENBERG, Roger. *Audacity* [en línia]. , 1999. < <https://www.audacityteam.org/> >.
- 38: CSETE, Alexandru. *Gqrx SDR* [en línia]. , . < <https://gqrx.dk/> >.
- 39: FRIGO, Matteo; JOHNSON, Steven G.. *The Fastest Fourier Transform in the West* [en línia]. , 1997-2018. < <http://fftw.org/> >.
- 40: Et al. *RGB color space* [en línia]. , . < https://en.wikipedia.org/wiki/RGB_color_space >.
- 41: Et al. *Modelo de color HSV* [en línia]. , . < https://es.wikipedia.org/wiki/Modelo_de_color_HSV >.
- 42: _Richard. *Drive RGB LED - Lookup Table?* [en línia]. , 2010. < <https://www.microchip.com/forums/m494386.aspx> >.
- 43: SharkD. *HSV color solid cylinder* [en línia]. , 2010. < https://es.wikipedia.org/wiki/Archivo:HSV_color_solid_cylinder_saturation_gray.png >.
- 44: Khan Academy. *Calcular la desviación estándar paso a paso* [en línia]. , . < <https://es.khanacademy.org/math/statistics-probability/summarizing-quantitative-data/variance-standard-deviation-population/a/calculating-standard-deviation-step-by-step> >.
- 45: Open Preservation. *Xcorr-sound* [en línia]. , 2012. < <http://xcorr-sound.openpreservation.org/> >.
- 46: Open Preservation Foundation. *Scape-xcorr-sound* [en línia]. , 2014. < <https://github.com/openpreserve/scape-xcorr-sound> >.
- 0: KOJEVNIKOV, Alexander. *Spek - Acoustic Spectrum Analyser* [en línia]. , 2016. < <http://spek.cc/> >.
- 47: ImageMagick Studio LLC. *ImageMagick* [en línia]. , . < <https://imagemagick.org/index.php> >.
- 48: Varis autors. *Comma-separated values* [en línia]. , . < https://en.wikipedia.org/wiki/Comma-separated_values >.
- 49: YVETTE, Anna; YOSEF, Jim. *Linked (Instrumental Version)* [en línia]. , 2017. < <https://www.youtube.com/watch?v=Tj7mbQ7o6dk> >.
- 50: YVETTE Anna; YOSEF, Jim. *Linked* [en línia]. , 2017. < <https://www.youtube.com/watch?v=yHLtE1wFeRQ> >.
- 51: Home and learn. *Invert image colors* [en línia]. , . < <https://www.homeandlearn.co.uk/extras/image/image-invert-colors.html> >.
- 52: The Free Software Foundation. *GNU General Public License, version 2* [en línia]. , 1991. < <https://www.gnu.org/licenses/old-licenses/gpl-2.0.html> >.

6. Annex

Al final del treball s'hi trobaran varis DVDs amb totes les dades recollides del treball:

- » Totes les imatges i àudios que s'han utilitzat per a fer la part pràctica del projecte.
- » El codi modificat que s'ha utilitzat al llarg del treball juntament amb els scripts¹¹.
- » Les imatges resultants de les restes i les comparacions.
- » Els arxius «csv» amb totes les dades recollides.
- » Els diferents fulls de càlcul amb les taules dinàmiques, que es poden modificar per veure els resultats de formes diferents.

¹¹ Per a poder executar el codi es necessitarà compilar-lo (tot i que la carpeta «build» amb el programa compilat està inclosa) i és probable que hi hagi mancances de dependències. Visitar la pàgina de GitHub de [scapecxcorrsond](#) per resoldre-les i per a la guia de compilació. Els scripts s'executen en una consola «bash».