Examensarbete

HÖGSKOLAN I SKÖVDE

Användning av olika samplingsfrekvenser inom dataspelsmusik

Hur påverkas spelupplevelsen av olika samplingsfrekvenser i musiken, 44.1 kHz respektive 96 kHz?

Using different sample rates in video game music

How does the gaming experience get affected by different sample rates in music, 44.1 kHz compared to 96 kHz?

Examensarbete inom huvudområdet Medier, estetik och berättande
Grundnivå 30 högskolepoäng
Vårtermin 2020

Sebastian Winbladh, Erik Lövgren

Handledare: Lars Bröndum
Examinator: Markus Berntsson

Sammanfattning

I denna studie undersöktes skillnader mellan 44.1 kHz och 96 kHz som samplingsfrekvens inom musikproduktion. Samplingsfrekvensen 44.1 kHz är sedan slutet av sjuttiotalet industristandard, dock säljs många av dagens ljudkort med en kapacitet att spela in frekvenser upp till 96 kHz eller 192 kHz. Detta har i sin tur skapat en debatt inom musikproduktionsvärlden om för och nackdelar mellan högre och lägre samplingsfrekvenser. I denna studie undersöktes kvantitativ data från tidigare studier och artiklar inom forskningsfältet gällande samplingsfrekvenser. Med tidigare forskning som utgångspunkt genomfördes en kvantitativ fältstudie. Resultatet av fältstudien indikerar att majoriteten av testdeltagarna upplevde en skillnad i kvalité mellan 44.1 kHz respektive 96 kHz som samplingsfrekvens. Fältstudien kombinerade dataspel och lyssningstest, hur olika dataspel samt lyssningstest splittrar testdeltagarnas fokus mellan dem bör undersökas vidare i framtiden.

Nyckelord: Samplingsfrekvens, Ljudkvalité, Dataspelsmusik, Musikproduktion

Innehållsförteckning

1 Introduktion **1**

2 Bakgrund **2**

2.1 Utvecklingen av dataspelsmusik 2

2.2 Samplingsfrekvens 3

2.2.1 Hertz 4

2.2.2 Skillnad mellan analoga och digitala ljud 4

2.2.3 Nyquist Teorin 5

2.3 Analog-Digitalomvandlare 5

2.3.1 Exempel av ADC med samplingsfrekvens 100 Hz 5

2.3.2 Aliasing/harmonisk förvrängning 6

2.3.3 Anti-aliasing-filter och ADC 6

2.3.4 Aliasing Intermodulation Distortion (AID) 7

2.4 Infra- och ultraljud 7

2.4.1 Människans hörselomfång under specifikt reglerade omständigheter 8

2.5 44.1 kHz och högre samplingsfrekvenser 8

2.5.1 Val av samplingsfrekvens 8

2.5.2 Hårdvara - mer än bara ljudkort 8

2.5.3 Future Proofing 9

2.5.4 Placebo 10

2.6 Relaterad forskning 10

2.6.1 44.1 kHz jämfört med 96 khz och 192 kHz som samplingsfrekvens 11

3 Problemformulering **12**

3.1 Metodbeskrivning 12

3.1.1 Försöksgrupp 12

3.1.2 Upplägg av studie 13

3.1.3 Problematisering kring studien och etik 13
3.1.4 Kvantitativ studie med enkät 13
3.6.5 Metoddiskussion 14
4 Genomförande 15
4.1 Spelet Temple Defender 15
4.2 Skapandet av kompositionen “Djungle Delay” 16
4.3 Inspelning av slagverk 19
4.4 Inspelning av kompositionen 21
4.5 Mixning av inspelningarna 22
4.5.1 Begrepp som används under mixningen 22
4.5.2 Sammanfattning av mixningsprocess 24
4.5.3 Tre plugins 24
4.5.4 Redigering av ljudfiler från inspelningarna 25
4.5.5 Slutet av mixningen samt exportering 26
4.6 Pilotstudie 27
5 Utvärdering 30
5.1 Presentation av undersökning 30
5.2 Analys 30
5.3 Slutsatser 32
6 Avslutande diskussion 33
6.1 Sammanfattning 33
6.1.1 Samhällsetiska aspekter 33
6.2 Diskussion 33
6.3 Framtida arbete 34
6.3.1 Ett par dagar extra 34
6.3.2 En större studie 35
6.3.3 Generella reflektioner över teknologisk utveckling & samplingsfrekvens 35
Referenser 37

1 Introduktion

Denna rapport innehåller fem kapitel. I bakgrundsdelen redovisas information om hur musik började användas i datorspel. Den definierar begreppet samplingsfrekvens och beskriver varför 44.1 kHz blev industristandard inom musikproduktion. Bakgrunden innefattar även information om infraljud, ultraljud och uppkomsten av högre samplingsfrekvenser.

I bakgrunden beskrivs också resonemang från forskare och musikproducenter gällande för- och nackdelar med dagens inspelningsutrustning. Eftersom den moderna utrustningen kan spela in högre frekvenser en 44.1 kHz upplever producenter i vissa fall att det kan vara ett problem. Kapitlet presenterar således fältstudier där det genomförts lyssningstester med varierande samplingsfrekvenser och nedsampling.

Därpå följande kapitel beskriver problemformuleringen. Där diskuteras valet av frågeställning för studien, samt metodbeskrivning, studiens upplägg, enkätfrågorna och urval av försöksgrupp. Studiens frågeställning är följande: *Hur påverkas spelupplevelsen av olika samplingsfrekvenser i musiken, 44.1 kontra 96 kHz?*

Tidigare fältstudier har fokuserat på ett mestadels audiellt perspektiv. Författarna av denna studie ville utforska vad som händer när visuella och audiella medier används i kombination.

En kvantitativ fältstudie genomfördes med åtta deltagare.

I genomförandedelen beskrivs hur författarna med utgångspunkt i tidigare forskning utformade sin egna studie. Vidare förklaras hur författarna skapade kompositionen som utgjorde artefakten, samt hur inspelning, mixning och pilotstudie genomfördes.

Utvärderingskapitlet presenterar fakta från författarnas fältstudie och jämför den med tidigare studier från bakgrunden. Dessutom problematiseras det kring varför högre samplingsfrekvenser verkar vara av större vikt beroende på genre och specifika användningsområden.

I den avslutande diskussionsdelen resoneras kring val av metod, utförande och på vilka sätt det kan ha påverkat resultatet. Vidare redovisas samhällsetiska aspekter samt förslag på vidare forskning.

2 Bakgrund

I en studie genomförd av Sarlins (2011, s. 3) kontrasterades att spela upp musik med ultraljud i 96 kHz jämfört med 44.1 kHz. Testdeltagarnas resultat indikerar att musik uppspelad i 96 kHz i samplingsfrekvens upplevs annorlunda mot 44.1 kHz.

Andra studier, exempelvis Biasuttis (1997, s. 1) antyder att det redan vid 3 kHz blir svårt att höra nyanser i musiken.

Harris studie blev en inspiration till att undersöka om skillnad kunde höras mellan 44.1 kHz och 96 kHz i samplingsfrekvens (Harris et al. 1998, s. 2).

Pras och Guastavinos (2010, s. 3) undersökning var något som efterliknades i denna studie då författarna använde en mikrofon som kunde ta upp frekvenser mellan 20 Hz till 50 kHz.

Grafiken i *Geometry Dash* (Topala u.å.) använde mycket neoninspirerad grafik, små trianglar som lyser i en neonfärg samt blinkande runda cirklar. Detta efterliknades i spelet för denna studie med inkräktare som är ritade med neonfärger.

Det hämtades inspiration till musiken från spelet *Geometry Dash*. Musiken som spelats med gitarr *delay* skulle låta som syntarna i *Geometry Dash*. Delay är en effekt som lägger till eko till ljudsignalen (TeachMeAudio 2020b). Låten *Stereo Madness* spelas i durtonarten E. Studiens komposition spelas också i en durtonart, och har även ett trumkomp som liknar *Stereo Madness* trumkomp. Anledningen till varför inspirationen har tagits från spelet *Geometry Dash* är för att grafiken i det spelet liknar den grafik som användes i denna studie.

2.1 Utvecklingen av dataspelsmusik

KyuSik Chang, GyuBeom Kim och TaeYong Kim studerade tillsammans på Chung-Ang Universitetet i Sydkorea där de genomförde fältstudien "Video Game Console Audio: Evolution and Future Trends". Studien fokuserar på hur ljud och musik inom dataspel har utvecklats samt hypoteser om framtida utveckling. När de första spelen lanserades fanns det inte ljud i spelkonsolerna. Exempelvis var Magnavox: Odyssey, som kom ut 1972, helt ljudlös (Chang, Kim, G & Kim, T 2007, s. 1). Det första system som kunde spela tre toner monofoniskt (Nationalencyklopedin u.åm) var konsolen Channel F: Fairchild som kom ut 1976. Dessa toner modulerades mellan varandra och spelades upp via konsolen. (Chang et al. 2007, s. 1). År 1977 lanserades den första spelkonsolen som kunde spela upp flera toner polyfoniskt, kallad Atari 2600 (Chang et al. 2007, s. 1).

År 1990 lanserades konsolen Super Nintendo Entertainment System. Vad som var speciellt för sin tid avseende denna spelkonsol var att den hade ett lågpassfilter på varje ljudkanal. Ett lågpassfilter är ett filter som tar bort frekvenser över en bestämd gränsfrekvens (Nationalencyklopedin u.åg). Denna teknik medförde att låga frekvenser upplevdes låta bättre (Chang et al. 2007, s. 3). År 1994 släpptes den första Playstationkonsolen, vilken spelade upp ljud med filformatet PSF (Portable Sound Format). En låt som var inspelad i detta filformat tog mindre lagringsutrymme, jämfört med inspelningar gjorda i de andra filformaten för ljud och musik som var tillgängliga vid denna tid (Chang et al. 2007, s. 3).

En spelkonsol är en typ av dator, vilken läser information som ettor och nollor. Programmen i datorn måste överensstämma med varandra gällande hur dessa ettor och nollor ska läsas. Ett datorprogram kan inte läsa information från filer som programmet inte kan förstå. Om ett program försöker läsa ett filformat som inte stöds, kan det medföra att programmet kraschar (Nationalencyklopedin u.å̊d). Vissa filformat är skapade för ett specifikt program, och kan därmed inte läsas av andra program, medan andra filformat kan läsas av ett flertal olika program (Nationalencyklopedin u.å̊d). Filformat kan således beskrivas som en uppsättning regler för hur en viss typ av data ska sparas (Nationalencyklopedin u.å̊d). Ett av de vanligaste och mest använda filformaten för musik i dagsläget är Mp3-filer. Mp3 är en metod för att komprimera stora filer, och därigenom ta upp mindre lagringsutrymme. Beteckningen på dessa filer är MPEG 1, Audio Layer-3. Filerna kan komprimeras till en tiondel av originalstorleken, vilket gör att filerna lättare kan skickas via internet eller lagras på en hårddisk (Nationalencyklopedin u.å̊j).

När Nintendo 64 kom ut 1996 kunde den spela upp både Mp3-filer samt MIDI. MIDI står för *Musical Instrument Digital Interface* och är en datakommunikationsstandard för musikinstrument (Nationalencyklopedin u.åi). Nintendo 64 kunde spela upp Mp3-filer samt MIDI med en samplingsfrekvens på 48 kHz och ett bitdjup på 16 bit. Dock var lagringsutrymmet för ljud och musik med denna kvalité begränsat (Chang et al. 2007, s. 4). År 2001 byggde Microsoft, som är ett amerikanskt programvaruföretag (Nationalencyklopedin u.å̊h), för första gången in ljudstandarden Dolby Digital i en konsol. Denna konsol hette Xbox (Chang et al. 2007, s. 4). Dolby Digital är ett digitalt ljudsystem för bio, HDTV (High Definition Television), DVD (digital video disc) och andra digitala medier med ljud. Ljudsystemet kan omfatta allt ifrån en till sex ljudkanaler, varav en kanal är till enbart för basljud (Nationalencyklopedin u.å̊c). När efterträdaren till Xbox kom 2006 (Xbox 360) var det industristandard att alla spel måste stödja 5.1 Dolby Digital surroundljud (Chang et al. 2007, s. 5).

Idag används formatet 22 kHz med 16 eller 8 bitars bitdjup fortfarande inom dataspel, men format som 44.1 kHz med ett bitdjup av 16 bitar (CD kvalitet) börjar långsamt bli standard (Marks 2012, s. 4). Formatet Voc (Creative Voice File) gjorde det möjligt för kompositörer att komponera musik till spelen i musikprogram, istället för att som tidigare programmera in musiken i ljudkort (Marks 2012, s. 4).

2.2 Samplingsfrekvens

Begreppet samplingsfrekvens kan användas i sammanhang gällande konvertering av analogt till digitalt material (Zawistowski & Shah 2018). Begreppet är en sammansättning av två separata ord - sampling och frekvens. Sampling definieras i Nationalencyklopedin som:

> "sampling [saˈmpliŋ], inom kommunikations-, ljud- och processteknik metod att med regelbundna tidsintervall, givna av *samplingsfrekvensen*, ta prov på en kontinuerlig signal och representera den som en diskretiserad tidssekvens av sampel."
>
> (Nationalencyklopedin u.å̊k)

Frekvens beskrivs som hur ofta något händer (Nationalencyklopedin u.åe). Inom musikproduktion beskriver således begreppet samplingsfrekvens hur ofta (frekvens) ljud (sampling) spelas in. I de sammanhang begreppet samplingsfrekvens används i denna rapport åsyftas endast ljudmaterial.

2.2.1 Hertz

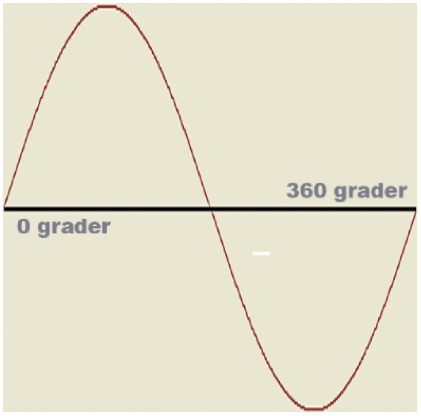
Samplingsfrekvensen anges i Hertz, Hz, vilket är enheten för hur ofta en våg upprepas per sekund (Nationalencyklopedin u.åf). Ofta används kilohertz, kHz, inom musikproduktion (1 kHz är lika med 1000 Hz). Industristandarden för samplingsfrekvens är 44.1 kHz (Meseu Das Comunicacoes u.å).

2.2.2 Skillnad mellan analoga och digitala ljud

Alla ljud som skapas i den riktiga världen är analoga ljud (Meseu Das Comunicacoes u.å). Exempel på analoga ljud kan vara en människoröst, en studsande boll eller spöregn. Det som alla analoga ljud har gemensamt är att de alltid är kontinuerliga. Ingen paus uppstår mellan början och slutet av ett analogt ljud.

Digitala ljud existerar enbart i den digitala världen. Till skillnad från analoga ljud är digitala ljud inte kontinuerliga (Meseu Das Comunicacoes u.å). När analoga ljud spelas in sparas dessa som digitala ljudsamplingar i ettor och nollor, även kallat för binär kod. När digitala ljudsamplingar spelas upp av digitala enheter måste samtliga ettor och nollor beräknas efter varandra vilket resulterar i ett icke kontinuerligt ljud. Dock upplever vi människor dessa digitala samplingar som ett kontinuerligt ljud om tillräckligt många samplingar spelas upp per sekund (Meseu Das Comunicacoes u.å).

Enligt nationalencyklopedin är en sinuston en kontinuerlig ren ton utan övertoner, se figur 1. Exempel på ljud som liknar sinustonen är visslande toner, eller toner från stämgafflar (Nationalencyklopedin u.ån).



Figur 1: Vid noll grader börjar tonen och vid 360 grader så har tonen gått hela varvet runt i en cykel.

2.2.3 Nyquist Teorin

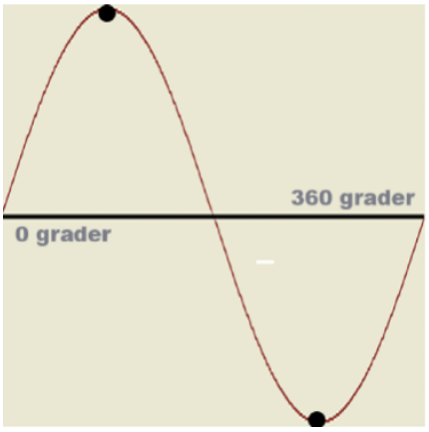
Bogdan är lektor på University of Lucian Blaga Sibiu i Rumänien, och beskriver Nyquist Teorin med följande citat: *"The minimum sampling frequency required to represent the signal should at least be twice the maximum frequency of the analog signal"* (Bogdan 2009, s. 1). Citatet betyder att man, i samband med inspelning, behöver minst dubbla samplingsfrekvensen för den högsta frekvens man vill spela in. Om en frekvens på 20 kHz korrekt skall återskapas behövs således en samplingsfrekvens på 40 kHz. Nyquistfrekvensen är i detta fall 20 kHz.

2.3 Analog-Digitalomvandlare

ADC är en förkortning av Analog to Digital Converter och DAC är förkortningen av Digital to Analog Converter (Meseu Das Comunicacoes u.å). Dessa begrepp beskriver analog-digitalomvandlare (Nationalencyklopedin u.åa). ADC omvandlar analoga ljudvågor till binärkod, som sedan datorer, och därmed musikprogram kan läsa. DAC omvandlar binärkod till elektroniska impulser vilka får membranen i högtalare och hörlurar att svänga, vilket i sin tur skapar ljudvågor som människor kan höra (Meseu Das Comunicacoes u.å).

2.3.1 Exempel av ADC med samplingsfrekvens 100 Hz

Om en frekvens på 50 Hz ska spelas in, behövs som minst samplingsfrekvensen 100 Hz eftersom det behövs 100 samplingar per sekund för att återskapa det analoga ljudet som digitalt ljud (Bogdan 2009, s. 1). Analoga ljud är vanligtvis mer komplexa och innefattar större frekvensomfång. För att illustrera hur en ADC fungerar används dock endast en sinusvåg av frekvensen 50 Hz i figur 2.



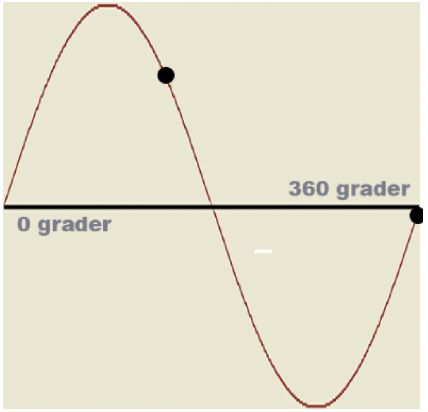
Figur 2: ADC med samplingsfrekvens 100 Hz, samplar frekvensen 50 Hz. Punkterna indikerar var frekvensen 50 Hz samplas.

ADC kommer med 100 samplingar per sekund konvertera den analoga ljudvågen av frekvensen 50 Hz till binärkod. På så vis skapas digitala ljud som datorer kan förstå. Eftersom samplingsfrekvensen är dubbla värdet av frekvensen som spelas in, kommer frekvensen att samplas korrekt. Nyquistfrekvensen i detta exempel är 50 Hz.

Om samplingsfrekvensen hade varit mindre än dubbla värdet av frekvensen som spelades in, skulle detta leda till att frekvensen ej samplas korrekt (Bogdan 2009, s. 1).

2.3.2 Aliasing/harmonisk förvrängning

Aliasing, även kallat harmonisk förvrängning, betyder att det uppstår ljud på frekvenser som inte finns i originalversionen. Om man i exemplet ovan istället försökt spela in en frekvens på 70 Hz med en samplingsfrekvens på 100 Hz, så skulle en harmonisk förvrängning uppstått (Bogdan 2009, ss. 1-2). Den förvrängningen betyder att inspelningen tar bort sinuskurvans toppar och åstadkommer en inspelad kurva med lägre frekvens. Denna harmoniska förvrängning skapar då undertoner på frekvensen 30 Hz på grund av att tonen 70 Hz speglas och blir 30 Hz. Sampling med en för låg samplingsfrekvens medför att inte hela vågen kommer med i samplingen. Samplingen görs inte på topparna av sinusvågen vilket innebär att en ny sinusvåg med lägre frekvens skapas och det skapas då en ny sinusvåg (frekvens) som inte fanns med i själva inspelningen (Bogdan 2009, ss. 1-2). Se figur 3.



Figur 3: Visar en sinusvåg på 70 Hz som ej korrekt samplas med en samplingsfrekvens av 100 Hz. Punkterna visar hur samplingen uppstår vid fel tillfälle.

2.3.3 Anti-aliasing-filter och ADC

Enligt Sarlin (2011, s. 3) kan människor normalt höra frekvenser mellan 20 Hz till 20 kHz. Ju äldre vi blir desto svårare blir det för oss att uppfatta de högre frekvenserna inom människans hörselomfång (Sarlin 11, s.3). Sarlin är mjukvaruutvecklare på Dizparc i Jönköping. Han har studerat till ljudtekniker på Luleås Tekniska universitet och till civilingenjör på Umeås Universitet. För att vara säker på att frekvenser över Nyquistfrekvensen inte skapar harmonisk förvrängning använder man ett lågpassfilter som tar bort frekvenser över Nyquistfrekvens. Lågpassfiltret i detta sammanhang kallas för ett anti-aliasing filter (Bogdan 2009, s. 2).

Ett anti-aliasing-filter fungerar på samma sätt som ett lågpassfilter. Dock har lågpassfilter fler användningsområden eftersom de finns i både analoga och digitala varianter. Anti-aliasing-filter är analoga och används enbart vid inspelning.

Bandbredden (bandwidth) är i detta sammanhang kopplat till det frekvensomfång anti-aliasing-filtret har när det skär bort frekvenser mellan till exempel 19.85 kHz till 24.25

kHz. Bandbredden är skillnaden i Hertz mellan den höga och den låga frekvensen (Nationalencyklopedin u.åo).

När lågpassfilter används som anti-aliasing-filter är det vanligt att dessa börjar vid 45% av samplingsfrekvensen och skär av helt vid 55% enligt Vest och Scheelke (u.åa, s. 1).
Om samplingsfrekvensen är 44.1 kHz, börjar alltså lågpassfiltret ta bort frekvenser vid 19,85 kHz, och tar bort alla frekvenser över 24,26 kHz. Mikael Vest och Peter Scheelke arbetar båda med försäljning på DAD (Digital Audio Denmark).

2.3.4 Aliasing Intermodulation Distortion (AID)

Om användning av anti-aliasing-filter, med samplingsfrekvensen 44.1 kHz används, kommer frekvenserna mellan 22,05 kHz (halva samplingsfrekvensen) och 24,6 kHz att skapa harmonisk förvrängning mellan frekvenserna 19,98 kHz och 22, 05 kHz.
Om dessa frekvenser över 19,98 kHz spelas upp separat från resten av en inspelningens hörbara frekvenser menar Vest och Scheelke (u.åa, s. 1) att människor inte kommer uppfatta dem.

Vest och Scheelke (u.åb, s. 1) förklarar hur människor oftast ej kan urskilja frekvensinformation över 20 kHz. Efter 12 kHz blir det betydligt svårare att höra skillnad mellan 12 kHz eller högre frekvenser. Därför borde en harmoniska förvrängning mellan 19, 98 kHz och 22,05 kHz vara irrelevant för människor. Så är dock inte fallet. När DAC ska konvertera de digitala ljuden till analoga ljud som sedan spelas upp via högtalare eller hörlurar, kommer dessa frekvenser mellan 19, 98 kHz och 22,05 kHz modulera med varandra och skapa nya undertoner i frekvensomfånget 1 kHz till 5 kHz. Detta fenomen kallas Aliasing Intermodulation Distortion (AID). Författarna av artikeln menar att AID kan motverkas genom att samplingsfrekvensen ökas till exempel till 96 kHz eller 192 kHz. (Vest & Scheelke, u.åb s. 1).

2.4 Infra- och ultraljud

Frekvenser under 20 Hz kallas för infraljud och frekvenser över 20 kHz ultraljud. Infra och-ultraljud är utanför människans generella hörselomfång (Sarlin 2011, s. 3). Vad gäller ultraljud kan dessa dock uppfattas på ett annat sätt av människor, beroende på hur begreppet ljuduppfattning definieras. Sarlin beskriver människans ljuduppfattning som följande:

> *"Research has been made and we do not only hear sounds through the tympanic membrane, when exposed to sound. It is also perceived through bone conduction which means that the sound wave skips both the middle ear and the tympanic membrane and instead the vibrations that sound waves generate in the bones goes straight for the cochlea which is the part of the ear that translates vibrations into brain impulses."*
>
> (Sarlin 2011, s. 6)

Sarlins forskning visar på att människor kan uppfatta infra- och ultraljud utanför sitt upplevda frekvensomfång (20 Hz till 20 kHz). Dessa frekvenser fångas inte upp av trumhinnan (tympanic membrane), utan går istället direkt genom skelettet till

hörselsnäckan (cochlea) vilken i sin tur omvandlar dessa frekvenser till hjärnimpulser. Det innebär att även om människor inte kan höra dessa frekvenser på samma sätt som frekvenserna 20 Hz till 20 kHz skapar ändå infra- och ultraljud hjärnimpulser, vilket gör att människor registrerar dessa frekvenser. Ultraljud påverkar dock inte hjärnan om enbart ultraljud spelas, men om vanlig musik inom 20 till 20 000 Hz spelas i kombination med ultraljud kan en skillnad i hjärnimpulser registreras. Det kan därav fastställas att dessa högre frekvenser noteras hos människor, men om de har någon påverkan på ljudupplevelsen är svårt att säga i dagsläget, menar Sarlin (2011, s. 10).

> *"Ultrasonic hearing is an auditory effect that allows humans to react to frequencies far above the limitations of the Tympanic membrane, in the right circumstances the human might be able to react on frequencies up to 100kHz and probably higher. Therefore it is of great relevance to the field to know whether we perceive ultrasound or not when concluding if ultrasound is necessary in audio equipment."*
>
> (Sarlin 2011, s. 9)

2.4.1 Människans hörselomfång under specifikt reglerade omständigheter

I sin forskning hänvisar Sarlin till en fältstudie som gjorts på Tokyo City University i Japan, om huruvida människor normalt sett kan höra frekvenser ovanför 20 kHz. Fältstudien genomfördes av studenterna Motoi Koubori, Kaoru Ashihara, Mizuki Omata, Masaki Kyoso och Shogo Kiryu (2010, s. 4-5). Deras forskning visar på att det är möjligt att höra frekvenser upp till 24 kHz, dock under specifika reglerade omständigheter. Deltagarna kunde höra högre frekvenser när de lyssnade via högtalare, jämfört med när de lyssnade via hörlurar. För att deltagarna skulle höra frekvenser upp till 24 kHz behövde dessa frekvenser spelas upp i minst 80 decibel, via högtalare i ett ljudisolerat rum (Koubori et al. 2010, s. 4-5). Decibel är en enhet för att mäta ljusstyrka på ljud (Nationalencyklopedin u.a).

2.5 44.1 kHz och högre samplingsfrekvenser

2.5.1 Val av samplingsfrekvens

Enligt Sarlin kan skillnad tydligt ses gällande olika ljudkorts möjlighet att spela in i olika frekvenser, eftersom informationen finns tillgänglig i exakta siffror. Vad gäller mikrofoner och högtalare är informationen dock ej lika lättillgänglig. Majoriteten av mikrofon- och högtalartillverkare redovisar ej specifik information gällande produkternas användning av ultra- eller infraljud, utan fokuserar på frekvenserna 20 till 20 kHz (Sarlin 2011, s.12).

2.5.2 Hårdvara - mer än bara ljudkort

Sarlin skriver att den upplevda skillnaden av samplingsfrekvens 192 kHz jämfört med 44.1 kHz är relativt liten. Den blir också svår att mäta eftersom inte enbart ljudkort utan även mikrofoner och högtalare måste var kapabla att kunna spela in och spela upp ultraljud. Högtalares frekvensomfång fokuserar ofta på människans vanliga frekvensomfång, det vill säga 20 Hz till 20 kHz (Sarlin 2011, s. 12).

Idag finns högtalare som kan spela upp frekvenser långt över 20 kHz. Till skillnad från när inspelningar genomförs med ljudkort och olika samplingsfrekvenser, används generellt inget lågpassfilter i högtalare. Hur mycket volym ultraljuden genererar, jämfört med 20 Hz till 20 kHz, är dock ofta svårt att få reda på (Sarlin 2011, s. 12). Det samma gäller för mikrofoner, vilka heller inte har något lågpassfilter. Dock är det vanligt att mikrofoner "naturligt" (utan lågpassfilter) reducerar ljudmottagningen från cirka 20 kHz.

När en inspelning väl skall genomföras är det, enligt Sarlin, (2011, s. 17) många olika faktorer som påverkar huruvida det faktiskt är möjligt att spela in, samt spela upp, ultraljud. De mikrofoner som används måste kunna registrera frekvenser över 20 kHz. Ljudkortet behöver ha förmåga att spela in dessa frekvenser. Det innebär att om en mikrofon tar upp frekvenser upp till 30 kHz, måste ljudkortet kunna använda en samplingsfrekvens av minst 60 kHz. I annat fall kommer frekvenser förloras under inspelningen. När ljudet är inspelat och skall spelas upp, bör även högtalarna kunna producera (spela upp) frekvenser upp till 30 kHz för att dessa skall vara möjliga att höra. Även om alla krav uppfylls, kan dock komprimerade filformat radera stora delar av frekvenserna över 20 kHz. Ett exempel på detta är det mest använda filformatet i dagsläget; Mp3 (Sarlin 2011, s. 10).

2.5.3 Future Proofing

För att försäkra sig om att musiken skall klara av framtidens krav, har vissa musikproducenter börjat använda 96 kHz eller högre samplingsfrekvens som en sorts "future proofing". Det vill säga att musik som spelas in i dagsläget skall kunna användas långt framöver, även om industristandarden skulle höjas. Warren Huart, professionell producent och grundare av webbsidan Produce Like a Pro (producelikeapro.com), hjälper genom sin hemsida många nyare producenter med informativa videor om mixning i sin helhet samt hårdvara. Huart uppger följande gällande 192 kHz och höga samplingsfrekvenser generellt:

> *"The reason why I said to do this in order to future-proof recordings is because film and TV often require that recordings be at a much higher sample rate, and before you know it, everything will be 192Khz. So when I say "future-proof", I mean record at the highest rate that you can, so if somebody likes your song and they want to use it in a film or TV show, you have recorded it at the sample rate and bitrate that they want it to be at. This is a really smart business reason to record at a higher sample rate, because you will be making sure that you can always be competitive and your music can be used in every different format!"*

(Huart 2019)

2.5.4 Placebo

Den upplevda ökade kvalitén av högre samplingsfrekvenser anses vara relativt liten, vilket gör det svårt att avgöra om en betydelsefull skillnad uppstår (Sarlin 2011), eller om detta endast är en placeboeffekt. Arthus Shapiro var en psykiatriker som är känd för sin forskning inom bland annat tourettes syndrom, han beskrev placeboeffekten år 1968 som följande.

"A placebo is defined as any therapeutic procedure (or a component of any therapeutic procedure) which is given (1) deliberately to have effect, or (2) unknowingly and has an effect on a symptom, syndrome, disease, or patient but which is objectively without specific activity for the condition being treated. The placebo effect is defined as the changes produced by placebos."

(Shapiro 1968, s. 136)

2.6 Relaterad forskning

Pras och Guastavino (2010, s. 2) undersöker i en fältstudie huruvida det är möjligt att höra skillnad mellan ljud inspelat i 44.1 kHz, 88.2 kHz eller 88.2 kHz nedsamplat till 44.1 kHz. Vid nedsampling av en frekvens tas samplingar bort i själva cykeln av frekvensen. Det medför att ljudet kommer ta mindre plats i datorn då det är färre samplingar av frekvensen (Bland 2002, s. 7). Amandine Pras är Universitetslektor inom området Digital Audio Arts på University of Lethbridge i Alberta och forskare på School for Advanced Studies in the Social Sciences i Paris. Catherine Guastavino är docent på School of Information Studies i Kanada. Dennis Bland har arbetat som programvaruutvecklare i drygt 25 år.

Alla deltagare i studien, förutom en, var studenter från Université du Québec à Montréal och McGill University och samtliga hade erfarenhet av att jobba i musikstudio. 16 deltagare, 15 män och en kvinna, med en medelålder av 30 år, genomförde testet. Pras och Guastavino (2010, s. 2) spelade in fem musikstycken inför undersökningen. All musik var inspelad med mikrofoner som kan ta upp frekvenser mellan 10 Hz till 60 kHz, vilket gör att ultraljud ej förloras när inspelning med högre samplingsfrekvenser genomförs. Enligt Pras och Guastavino (2010, s. 3) så var inspelningarna fem till åtta sekunder långa.

Testdeltagarna genomförde ett abx-test, vilket innebär att inspelning X var den samma som antingen A eller B. Deltagarna fick i uppgift att lyssna på alla tre versionerna A, B och X, för att sedan försöka avgöra om A eller B var samma inspelning som X. Deltagarna behövde lyssna på alla tre versionerna minst en gång, men fick valet att lyssna fler gånger om de ville. Efter testet fick deltagarna bedöma svårigheten på testet genom att uppge en skattad siffra på en skala ett till tio. Ett innebar lätt och tio innebar svårt. Testdeltagarna skattade i genomsnitt nio av tio i svårighetsgrad. De upplevde att testet var mycket krävande på grund av hur koncentrerade de behövde vara genom hela testet.

Undersökningen visade att deltagarna kunde höra skillnad på musik inspelad i 44.1 kHz jämfört med 88.2 kHz, men det gällde endast vissa instrument. Det orkestrala utdraget var det utdrag där deltagarna kunde höra störst skillnad. Detta betyder, enligt Pras och Gustaviano, att högre samplingsfrekvens för inspelning i större rum med orkestral verksamhet fångar upp mer detaljer inom musik och akustik. När testdeltagarna blev tillfrågade hur de kunde höra dessa skillnader i musiken uppgav de rumsakustik, höga frekvenser, timber och precision som de viktigaste faktorerna. Dock var dessa faktorer olika svåra att identifiera beroende på vilket utdrag som spelades upp (Pras & Guastavino, s. 6). Fältstudien visade också på att deltagarna hörde en tydligare skillnad mellan originalversionen av 88.2 kHz och den nedsamplade versionen av 88.2 kHz. Detta, menar

författarna av artikeln, kan ha att göra med på vilket sätt programmet man använder för nedsampling utför sin konvertering. Författarna menar att det behövs mer forskning inom detta område för att skapa en klarare bild gällande olika konverteringsmetoders betydelse (Pras & Guastavino, s. 7).

2.6.1 44.1 kHz jämfört med 96 khz och 192 kHz som samplingsfrekvens

I en studie av Harris, Kelly, Mcleod och Story (Harris, Kelly, Mcleod & Story 1998 s. 2), jämfördes skillnader mellan 44.1 kHz med 16 bit , 96 kHz med 24 bit, och 192 kHz och 24 bit. Skillnaderna beskrivs som följande.

> *"Recording engineers, and many musicians – particularly in the classical area – are becoming aware that material recorded and edited using these higher sample rates has some attractive qualities. Current theory on how human hearing works has so far been unable to explain the basis for these qualities, but they are nonetheless easy to demonstrate."*
>
> (Harris et al. 1998 s. 2)

Studien visade att deltagarna upplevde en ökad kvalité i samband med högre samplingsfrekvens. Jämfört med 44.1 kHz upplevdes rumsakustiken och stereobilden vara tydligare i 96 kHz och 192 kHz. Vid 192 kHz upplevdes ljudet vara av högre kvalitet i jämförelse med både musik i 96 kHz och 44.1 kHz i samplingsfrekvens. Stereobilden kunde även upplevas som bredare i 192 kHz, dock kunde basen uppfattas som ur takt.

För studien användes ATR-100 bandspelare, en analog-digitalomvandlare med ADC och DAC funktion, en förstärkare och två högtalare. Anordningen kunde snabbt byta mellan analogt material och digital material, med olika samplingsfrekvenser och bitdjup (Harris et al. 1998 s. 1). En förstärkare är en enhet som förstärker den ljudsignal som inmatas genom enheten med hjälp av ström (Agarwal & Lang 2005, s. 331).

3 Problemformulering

År 1982 definierade Sony och Philips samplingsfrekvensen 44.1 kHz som standard för CD-kvalité. Sedan dess har diverse alternativ med högre kvalité lanserats, dock utan någon större kommersiell succé (Pras & Gustavino, s. 1). Även om högre samplingsfrekvenser än 44.1 kHz används vid inspelning i dagsläget, så nedsamplas den mesta musiken till 44.1 kHz i slutändan när den skall konsumeras som en färdig produkt (Marks 2012). En stor del av inspelningsutrustningen som produceras idag är kapabel att spela in med samplingsfrekvenser upp till 192 kHz. Om ett projekt inom en DAW (Digital Audio Workstation) använder 96 kHz är det möjligt att nedsampla projektet till 44.1 kHz eller till 48 kHz vilka båda är vanliga samplingsfrekvenser. Processen är dock inte möjlig åt det omvända hållet. Ett individuellt projekt kan enbart använda en samplingsfrekvens åt gången.

Det kan således ifrågasättas varför inte samtliga musikproducenter använder industristandarden 44.1 kHz. Inom musikproduktion råder delade uppfattningar kring huruvida man bör spela in i högre samplingsfrekvenser eller om detta enbart är onödigt eftersom det kräver större lagringsutrymme och högre datorprestanda. Det faktum att högre samplingsfrekvenser än 44.1 kHz kan anses onödigt har sin grund i Nyquist Teorin.

I debatten gällande för och nackdelar med högre samplingsfrekvenser blandas konkret fakta med subjektiva åsikter. Detta skapar en gråzon gällande vad som objektivt är sant och vad som endast är upplevda effekter av högre samplingsfrekvenser. Därav kan det vara svårt att skapa sig en klar uppfattning gällande varför man borde använda högre samplingsfrekvenser eller ej. Det här ledde fram till studiens frågeställning: *Hur påverkas spelupplevelsen av olika samplingsfrekvenser i musiken, 44.1 kHz respektive 96 kHz?*

3.1 Metodbeskrivning

För att svara på frågeställningen användes en kvantitativ metod med frågeformulär. För denna studie komponerades ett stycke, cirka en minut långt. Stycket varierar i dynamik, intensitet, frekvensomfång samt antal instrument, vilka kommer in gradvis under styckets gång. Denna komposition spelades in två gånger, en gång med 44.1 kHz som samplingsfrekvens, och en gång med 96 kHz.

3.1.1 Försöksgrupp

I studien deltog 8 personer i åldrarna 22 till 27 år. Deltagarna rekryterades med hjälp av sociala nätverk inom vänskapskretsen. Deltagarna var kunniga inom musik och ljudskapande och samtliga deltagare förutom en har gått en ljud- eller musikutbildning på högskolan i Skövde. Även om att bara män deltog så har författarna inte sett någon forskning som visar på att män och kvinnors hörsel skiljer sig åt, utan att ålder är den viktigaste faktorn för att höra högre frekvenser. Ett inklusionskriterie bestämdes därmed vara ålder, från 20 till 30 år. I de tidigare fältstudier som nämns i bakgrundskapitlet av denna studie, har forskarna lagt vikt vid huruvida studiedeltagarna varit erfarna inom musikproduktion eller inte. Sannolikt har studiedeltagare med en bakgrund inom musikproduktion lättare att veta vad de mer specifikt ska lyssna efter i ljudbilden, så kallad “Purposive Sampling”

(Denscombe 2003, ss. 15-16). Därav gjordes valet att i första hand rekrytera deltagare som var utbildade inom musik.

3.1.2 Upplägg av studie

Testdeltagarna lyssnade på de olika inspelningarna samtidigt som de spelade ett tillhörande spel. Spelet som användes var Temple Defender som har konstruerats samt programmerats av en av författarna av denna studie. Inspelningarna var cirka en minut långa, och spelades upp efter varandra med en kort paus emellan. Testdeltagarna visste ej vilken version som spelades upp eller vilken ordning, testen gjordes som blindtest. Grundläggande instruktioner gällande testets funktion gavs till deltagarna före testet. De fick också instruktioner hur de kunde ställa in ljudkortet till en volym som upplevdes bekväm för var och en. Testdeltagarna fick lyssna till två olika inspelningar efter varandra och det ombads lyssna till skillnader mellan dessa. Efter spelsessionen fick testdeltagarna svara på vilken version som var inspelad med lägre respektive högre samplingsfrekvens och fylla i ett frågeformulär.

3.1.3 Problematisering kring studien och etik

Innan själva fältstudien genomfördes en pilotstudie, för att uppmärksamma eventuella brister i utformningen av studien. Författarna ansvarade för att testdeltagarnas medverkan hölls hundra procent konfidentiell. Den data som redovisades kunde ej härledas till enskilda individer. Innan studien började tillfrågades samtliga testdeltagare huruvida de vill delta i studien eller ej. Deltagandet var helt frivilligt och testdeltagarna hade all rätt att avsluta testerna när de själva önskade (Denscombe 2003, ss. 39, 71-79).

3.1.4 Kvantitativ studie med enkät

Testdeltagarnas upplevelser dokumenterades med hjälp av en enkät i Google Formulär. Enkäten, vilken besvarades efter att kompositionen spelats upp med både samplingsfrekvens 44.1 kHz och 96 kHz i kombination med spelet Temple Defender.

Frågorna som användes, och varför de användes, redovisas nedan.

- **Upplevde du någon skillnad i immersion mellan spelomgångarna?**

  Denna ja- eller nejfråga klargjorde huruvida musiken som var inspelad med en samplingsfrekvens på 96 kHz hade någon påverkan på spelupplevelsen kontra versionen på 44.1 kHz. Frågan har en direkt koppling till hela studiens frågeställningen som är *Hur påverkas spelupplevelsen av olika samplingsfrekvenser i musiken, 44.1 kontra 96 kHz*?.

- **Upplevde du någon skillnad i musiken mellan spelomgångarna?**

  Syftet med denna fråga var att se om deltagarna kunde höra skillnad på musikaliskt innehåll mellan spelomgångarna för att kunna analysera om 96 kHz i samplingsfrekvens har någon betydande fördel gentemot 44.1 kHz.

- **Om du hörde skillnad var det något specifikt instrument detta gällde?**

  Denna fråga undersöker om det är något särskilt instrument som framträder mer i den ena versionerna kontra den andra. Vissa instrument har högre frekvenser än

andra, vilket gjorde att det kunde undersökas om deltagarna kunde uppfatta skillnaden i de olika versionerna.

- **Vilken av spelomgångarna upplevde du hade högst ljudkvalitet?**

  Denna frågas syfte var att ta reda på om deltagarna upplevde 96 kHz i samplingsfrekvens som överlägset i förhållande till 44.1 kHz.

3.6.5 Metoddiskussion

Att genomföra en kvantitativ studie hade många fördelar i sammanhanget. Författarna kan i efterhand se ett tydligt samband mellan svaren på enkätfrågorna och vår frågeställning. Eftersom frågeställningen till stor del fokuserar på huruvida testdeltagarna upplevde en skillnad mellan samplingsfrekvenserna, användes konkreta ja- eller nejfrågor. Dock kan nackdelarna med att göra en kvantitativ studie vara att författarna inte kommer att få utförliga svar, till skillnad från vid en kvalitativ studie (Denscombe 2003, s. 6). Men eftersom fokus i denna studie framför allt ligger på om man alls kan höra en skillnad, och inte vilken skillnad man eventuellt hör, anses en kvantitativ studie lämplig. En kvalitativ studie hade medfört att testdeltagarna måste anstränga sig mer för att besvara varje fråga. Författarna vill att testdeltagarna ej ska behöva fundera över hur de formulerar sig och på så vis få mer precisa svar.

4 Genomförande

4.1 Spelet Temple Defender

Spelet fick namnet Temple Defender för att belysa att det går ut på att försvara ett neonfärgat tempel i mitten av en bana mot olika fiender. Templet har liv som fienderna försöker förstöra.

Spelet är också lätt att lära sig, vilket möjliggör ett fokus på studien istället för att lära testdeltagare att spela. Fördelen med att göra ett eget spel är också att det programmeras med egen kod. Det blir då möjligt att skräddarsy spelet till vår studie, något som inte hade varit genomförbart om ett kommersiellt spel använts. Spelet kunde till exempel formas efter våra ideèr, bilder kunde laddas upp från spelet för denna studie utan att behöva ta hänsyn till upphovsrätten.

Svårigheterna med att utveckla spelet var att locka till ett engagemang i spelandet. I början var spelet repetitivt på grund av att det bara fanns en fiende. För att motverka detta problem skapades ytterligare tre fiender med olika egenskaper. En kan till exempel explodera, en annan har extra mycket motståndskraft och den tredje kan flyga snabbt. Bombfienden kan till exempel döda andra inkräktare som äter på templet. Allt eftersom templet får mindre liv måste spelaren trycka på en knapp i mitten av templet för att ge energi till templet.

Tidigt i utvecklandet gjordes valet att bygga spelet till mobiltelefonen. Spelmotorn som användes för att programmera spelet heter Cocos2d (Cocos u.å.). Cocos2d är en spelmotor som kan användas för att göra spel till både Android och iPhone. Spelets grafik ritades i Photoshop (Adobe u.å.). Grafiken kan ses i figur 5.

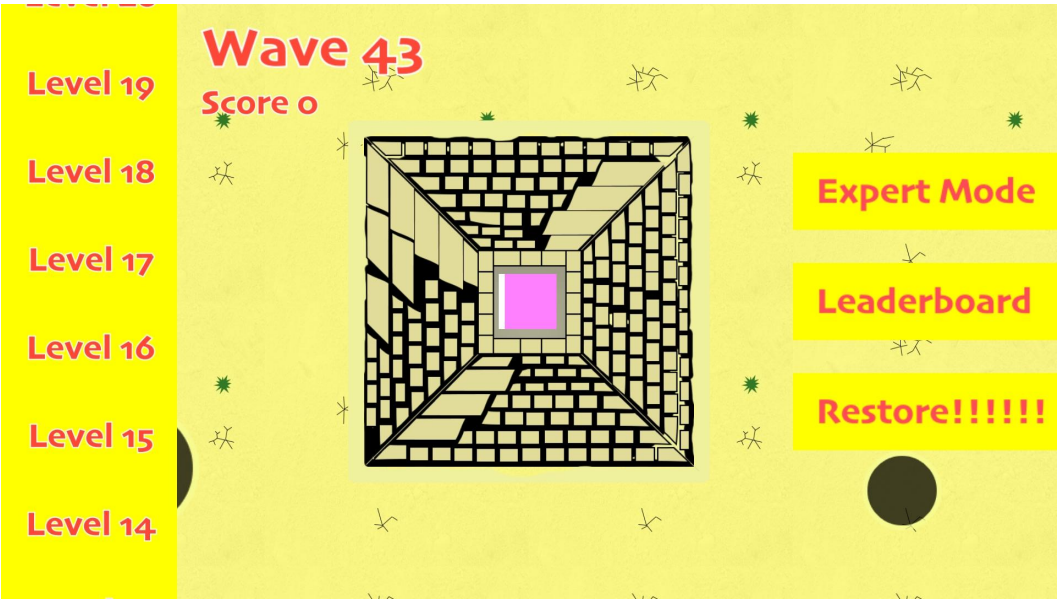
Spelmotorn kunde användas till att utveckla ett spel för en telefon, vilket är positivt då en mobiltelefon är lätt att ta med sig. Om spelet istället hade funnits på en dator hade det genast blivit svårare att ta med utrustningen, med tanke på att det också behövdes ett externt ljudkort som kunde spela 96 kHz. Att använda någon form av spel var också viktigt för vår studie.

Tidigt i utvecklingsprocessen fanns ingen grafik till spelet eftersom logiken skapades först. Logiken är alltså det som lägger grunden för spelet. Tidigt i utvecklingen gjordes valet att spelaren skulle slå på inkräktarna för att döda dem med hjälp av knappen till höger (se figur 5). Alternativet till att slå hade varit att skjuta projektiler på inkräktaren. Anledningen till att slå istället för att skjuta gjordes för att skapa en bättre spelupplevelse.

I figur 5 kan en knapp ses i mitten av templet. Knappen ser ut som en hammare och är till för att laga templet när templet har låg energi. Denna knapp var från början inte tänkt att finnas, men programmerades in eftersom templet behövde kunna lagas.

Spelets första grafik var ritat med ett sandfärgat tempel i mitten av banan. Se figur 4. Grafiken ändrades sedan till ett mörkare tema. Objekten i upplaga nummer två var tänkt att målas tydligare så att föremålen bättre kunde ses av användaren. Se figur 5.

Grafiken i *Geometry Dash* (Topala u.å.) använder mycket neoninspirerat, till exempel små trianglar som lyser i neonfärg samt blinkande runda cirklar. Detta efterliknades i spelets andra upplaga med inkräktare som är ritade med neonfärg.



Figur 4: Tidigare grafik till spelet Temple Defender.

Figur 5: Spelets nuvarande grafik.

4.2 Skapandet av kompositionen "Djungle Delay"

Många av de virtuella instrument som finns tillgängliga idag använder relativt låga samplingsfrekvenser, vilket ej skulle fungera i denna studie som använder 96 kHz i samplingsfrekvens. Därav innefattar stycket och inspelningen endast analoga och akustiska

(riktiga) instrument, vilka spelats in via mikrofon. I kompositionen används akustisk gitarr, elgitarr och mindre slagverk.

Den här studien utgår ifrån att spela in samma komposition två gånger, en första gång med samplingsfrekvens 44.1 kHz och en andra gång med 96 kHz, samt 24 bitars djup i båda inspelningarna. I kompositionen användes elgitarr, akustisk gitarr och mindre slagverk. Ursprungligen var det planerat att använda ett trumset på högskolan i Skövde som slagverk, men på grund av den pågående epidemin av Covid-19 så valde författarna av denna studie att istället spela in mindre slagverk hemifrån.

Kompositionen skrevs med mobilspelet *Geometry Dash* som inspirationskälla, eftersom det spelet upplevdes vara av liknande karaktär som *Temple Defender*. Musiken i *Geometry Dash* använder virtuella instrument, vilket författarna funderade på att använda tidigare under skapandet av kompositionen. Dock skulle dessa virtuella instrument ej fungera i denna studie. Detta beror på att de flesta ljudbibliotek är inspelade i 44.1 kHz. Eftersom denna studie jämför skillnaden mellan 44.1 och 96 kHz som samplingsfrekvens skulle dessa virtuella instrument ej fungera för inspelningen i 96 kHz. Därav valdes det istället att enbart använda analoga instrument.

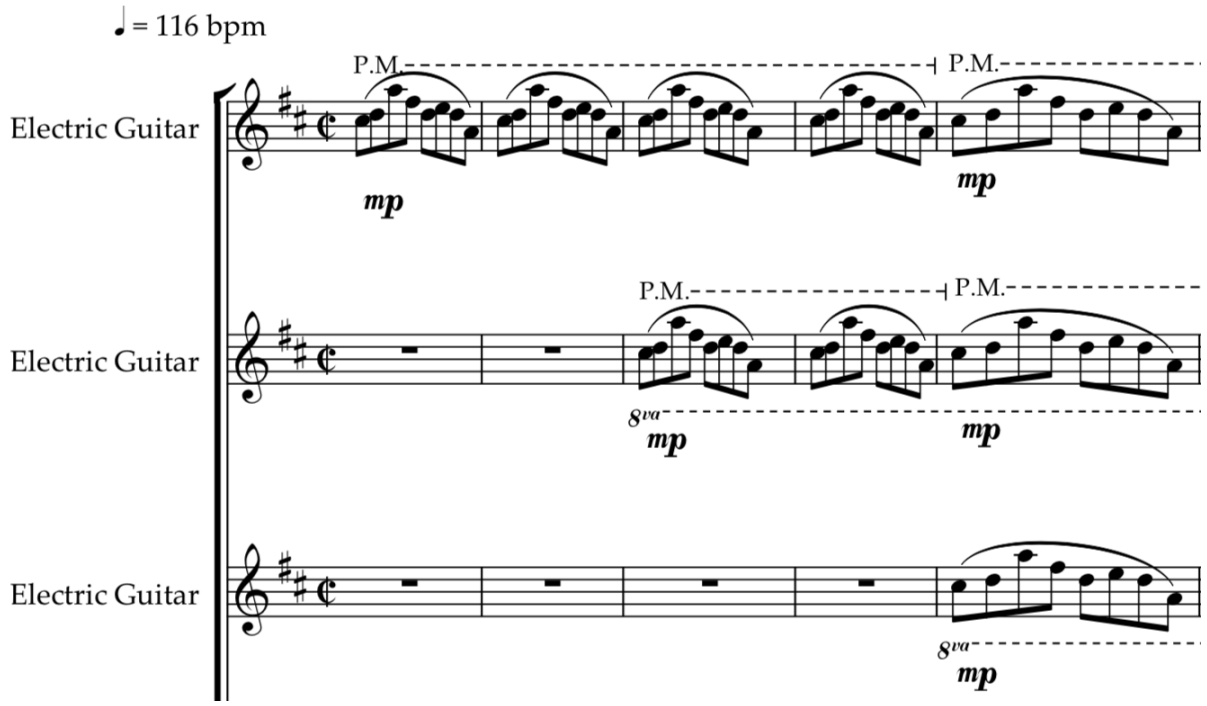
Elgitarr och effektpedaler användes för att efterlikna syntharna i *Geometry Dash* (Topala u.å.). Synthar har ofta ett flertal effekter lagda på varandra samtidigt, vilket skapar en stor ljudbild och kan upplevas ta mycket av uppmärksamheten i musiken. För att försöka efterlikna de stora synthljuden i *Geometry Dash* användes reverb och ett kraftigt delay i tre lager.

Reverbet på gitarrförstärkaren som användes, gav författarna möjligheten att skapa en större upplevd rumsklang, jämfört med hur elgitarren lät från början. Då elgitarr spelades in med delaypedal (se 4.6.1) på tre ljudspår, skapades en betydligt större ljudbild jämfört med när delaypedal inte använts, se figur 6.

Figur 6: Delaypedal Hardwire DL-8, användes under inspelning för att skapa delay på elgitarren.

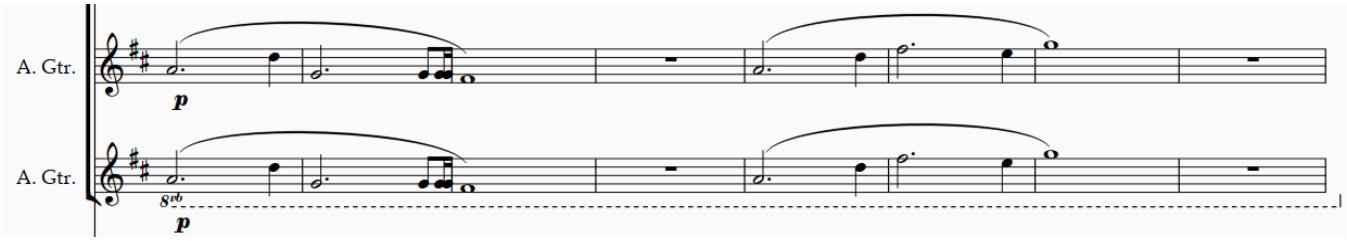
Kompositionens grund bygger på en upprepande melodi, som tillsammans med delaypedalen skapar en konstant rytm genom hela kompositionen. Denna korta melodi är inspelad i två oktaver med palm mute, samt en tredje gång i den högre oktaven utan palm mute. Palm mute är en gitarrteknik där gitarristen lätt dämpar strängarna med sin högra hand samtidigt som hen spelar (Palm Mute u.å). Det mest problematiska under inspelningen var att hitta rätt bpm, i kombination med delaypedalen som användes. På delaypedalen Hardwire DL-8 går det inte att ställa in exakt bpm, istället används en tids-ratt (time) för att ställa in tempo. Eftersom elgitarrerna spelades in i flera lager, var det viktigt att ställa in en bpm i Pro Tools som synkade i tempo med delaypedalen. Annars skulle inspelningarna kunna låta ur takt. Pro Tools och delaypedalen synkade ihop i tempo med 116 bpm inställt i Pro Tools.

Kompositionen går i tonarten D-dur för att efterlikna musiken i *Geometry Dash*, vilken också är skriven i en durtonart. Noter för elgitarrerna, vilka är grunden för kompositionen kan ses i figur 7 nedan.



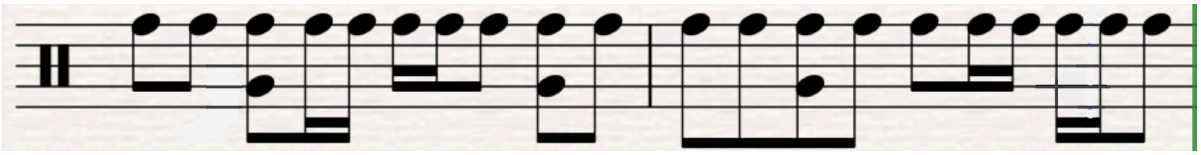
Figur 7: Elgitarr, åttondelar i tre lager.

För att skapa variation i kompositionen skrevs även en melodislinga som spelas två gånger. Denna melodi spelades från början in på på elgitarr, men valdes senare att spelas in på akustisk gitarr. Anledningen var att den akustiska gitarren skapade mer övertoner samt ultraljud. Melodin spelas i två oktaver. Melodin spelades även in i en lägre oktav för att simulera en akustisk bas, se figur 8.



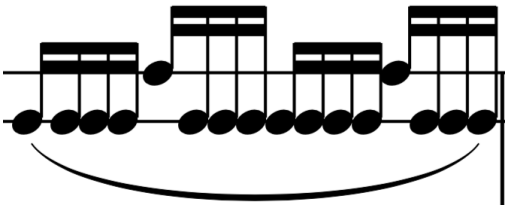
Figur 8: Akustisk Gitarr, melodi i två oktaver.

4.3 Inspelning av slagverk



Figur 9: Slagverk från melodin *Stereo Madness* från spelet *Geometry Dash*.

För att få fram ett bra ljud användes olika slagverksinstrument. Slag på en gitarr användes som ett av slagverken i kompositionen. För att kompositionen skulle likna *Stereo Madness* snabba takter gjordes slagen i 16-delar. Dessa slag spelades in med hjälp av en Zoom H1 mikrofon som placerades 10 centimeter från gitarrens resonanslåda (se Figur 10).



Figur 10: Slag som slås på en gitarrkropp. Det övre slagen i notbilden är betoningar.

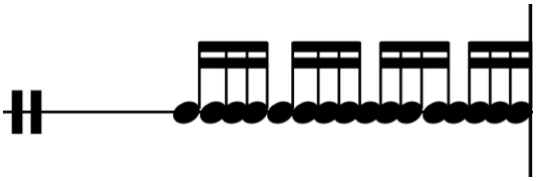
Denna del var viktig att få i takt till en metronom, eftersom samma sak sedan skulle spelas in på nytt i 96 kHz. Det var viktigt att inspelningarna blev så lika varandra som möjligt så att testdeltagare inte skulle reagera på att inspelningarna var olika på grund av att den ena inspelningen var ur takt. Slag mot gitarrkropp användes i kompositionen. Gitarren hölls i halsen så att resonanslådan var fri från objekt som skulle kväva gitarrens resonans. Mikrofonen placeras 10 centimeter från gitarrens resonanslåda. Takten kan ses i figur 11.

Figur 11: Slag på gitarrens resonanslåda.

Med en tamburin kan frekvenser på 47 kHz uppnås. För att uppnå samma frekvens men med en mjukare ton så användes en shaker i stället för en tamburin. I melodin *Stereo Madness* lät slagverket som en shaker vilket underlättade valet av instrument. Vid inspelningen av shakern placerades en Zoom H1 mikrofon med 10 centimeters avstånd från instrumentet. För att ta upp ett kvalitativt bra ljud med hög volym utan distribution så behövde mikrofonen vara rätt inställd, respektive ställas in noggrant. Inspelningen togs om ett par gånger för att försäkra att shakern var i takt med musiken. Shakern spelades in med två olika rytmer. Den första rytmen kan ses i figur 12.

Figur 12: En enkel rytm som spelas med en shaker.

För att efterlikna *Stereo Madness* ytterligare användes shakern för att markera takten i 16-delar. Det kan ses i figur 13. Detta var en rytm som var svår att få inspelad i takt, därför spelades denna rytm in några gånger innan ett lyckat resultat uppstod.



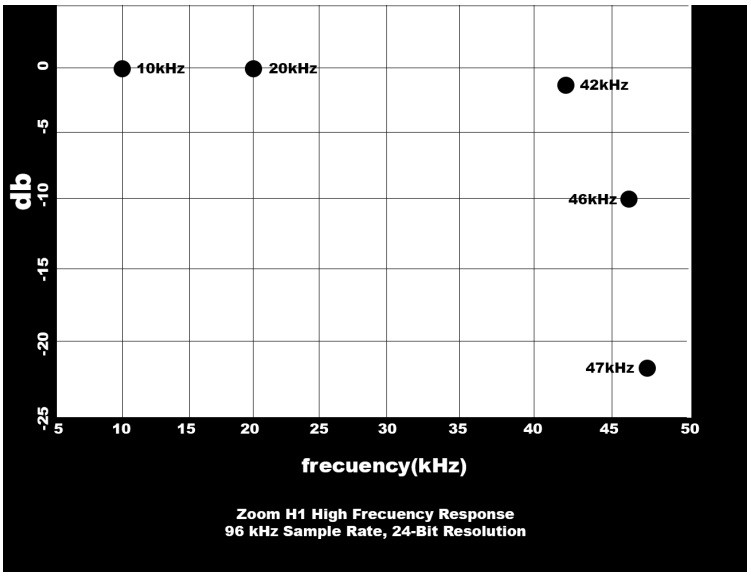
Figur 13: Takt i 16-delar.

Det sista slagverket som spelades i kompositionen var guiro (även kallad gurka) som är ett latinamerikanskt rytminstrument. För att spela in guiron placerades Zoom H1 mikrofonen på 20 centimeters avstånd. Guiron spelades in en gång per version eftersom rytmen var lättspelad. En del av noterna för guiron kan ses i figur 14.

Figur 14: En enkel rytm som spelas med en guiro.

4.4 Inspelning av kompositionen

Beträffande val av mikrofon fanns tre olika alternativ tillgängliga, vilka var SM 57, Röde NT1-A och Zoom H1. Då kompositionen skulle spelas in med både 44.1 och 96 kHz var det viktigt att använda en mikrofon med möjlighet att spela in ultraljud. Av de tre mikrofonerna var Zoom H1 den enda som kunde spela in ultraljud. Zoom H1 kan spela in upp till ungefär 50 kHz, frekvensomfånget för Zoom H1 kan ses i figur 15.



Figur 15: Frekvensomfång Zoom H1.

Om Zoom H1 ansluts direkt till en dator och inspelningsprogram via dess USB kan den enbart spela in i högst 48 kHz med 16 bitars djup. Lösningen blev att använda line out från Zoom H1 till ett Motu M2 ljudkort, med 96 kHz 24 bitars djup inställt på båda två, och sedan ändra både Zoom H1 och Motu M2 till 44.1 kHz med 24 bitars djup för den andra inspelningen. Detta kan ses i figur 16 nedan.

Figur 16: Zoom H1 stereo line out, till två mono tele in i Motu M2.

4.5 Mixning av inspelningarna

4.5.1 Begrepp som används under mixningen

DAW(Digital Audio Workstation):
Ett datorprogram som kan bearbeta ljud eller musik.

Delay:
Delay är en effekt som lägger till eko till ljudsignalen (TeachMeAudio 2020b)

Plugin:
Plugins är mjukvara som installeras som tillbehör för ett program och utökar detta programs förmågor (computerhope 2020).

Reverb plugin:
Reverb beskriver det rumseko som bildas när ljud studsar mot väggar i ett stort rum (Brooks 2018). Reverb plugins ger möjligheten att digitalt skapa en större upplevd rumsklang, jämfört med hur originalinspelningen lät.

Spectrum analyser plugin:
En spectrum analyser kan grafiskt visa en inspelnings frekvensomfång och de olika frekvensernas volym (Tektronix u.å).

Mastering plugin:
En mastering plugin används oftast på master fader som en sista detalj när en inspelning är färdigt mixad. Används för att justera volym värdet innan musik exporteras till olika filformat. *(*Producelikeapro 2019).

Master fader:
Master fader kontrollerar ljudstyrkan allt ljud i ett DAW projekt (sweetwater 2005).

Kompressor plugin:
En kompressor ökar de svaga delarna i ett ljud och minskar de starka delarna för att jämna ut ljudet. *(*Practical Music Production u.å).

Volym:
Volym är en vanlig term för ljudnivå (Nationalencyklopedin u.å1).

Ljudspår:
När analoga och digitala ljud spelas in används DAW. Varje ljud läggs på ett eget ljud spår för att kunna redigeras separat. Till exempel kan en elgitarr spelas in till ljudspår 1, och sång spelas in till ljudspår 2 (Apple 2019a).

Auxiliary spår:
Ett auxiliary spår är ett extra spår som det inte går att spela in ljud på, men ljud kan kan skickas till spåret och sedan ändras med hjälp av pluginer. Man kan inte lagra inspelat ljud på ett auxiliary, men man kan skapa effekter med ett auxiliary spår(sweetwater 2013).

Buss:
En buss är en koppling som transporterar ljud från ett ljudspår till ett auxiliary spår (HomeBrewAudio 2020).

Grid:
Grid innebär att man kan flytta inspelade ljud på ett ljudspår till ett bestämt rutnät i en DAW. Funktionen grid finns i många olika DAW, som Logic och Pro Tools *(*Apple 2019b).

Clip Gain:
Clip Gain används för att justera specifika delar av ljudet i ett ljudspår. Det går att höja eller sänka volymen av olika delar i ljudspåret utan att förändra volymen i ljudspåret som helhet *(*ProToolsProduction 2016).

Crossfade:
Crossfade används för att skapa skarvlösa övergångar från ett ljud till ett annat *(*Margaret 2005).

Fade in - Fade out:
Fade in och Fade out används för att tona in och tona ut ljud (audacityteam 2020).

Panorering:
Panorering betyder att ett ljud distribueras ut till antingen höger eller vänster högtalare. Ljudet behöver inte panoreras i sin helhet åt höger eller vänster högtalare, utan det kan varieras (TeachMeAudio 2020a).

Session Data:
Session data betyder att inställningarna från ett projekt i en DAW kan kopieras till ett annat projekt (Gilliland 2017).

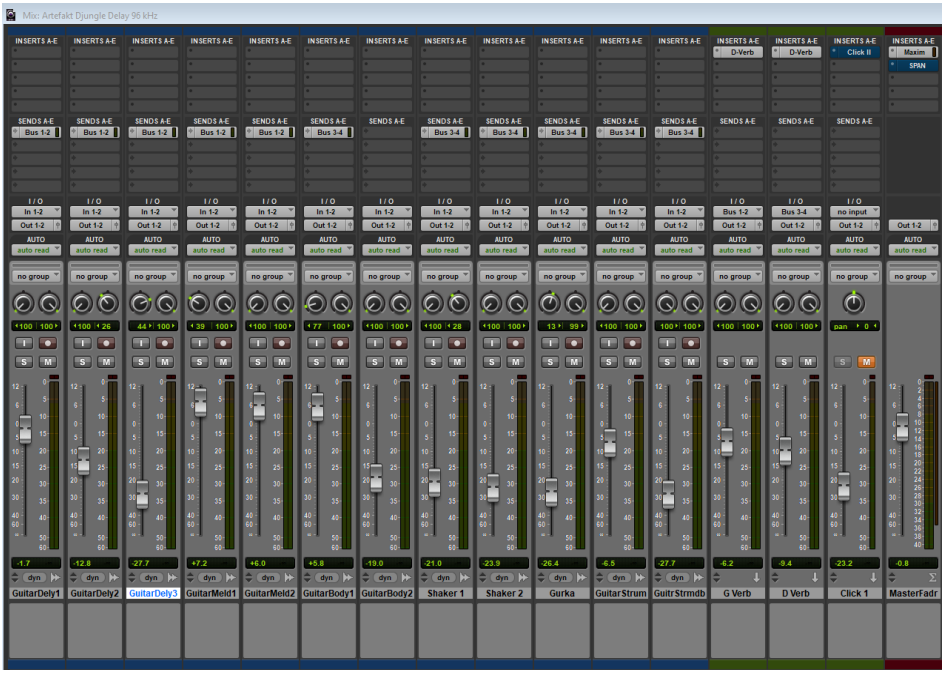
4.5.2 Sammanfattning av mixningsprocess

Kompositionen spelades in två gånger i programmet Pro Tools (Avid 2020). Inspelningen som gjordes i 96 kHz redigerades och exporterades vi *session data* till 44.1 kHz. Med hjälp av *session data* importerades inställningar på *ljudspår, auxiliary spår, bussar, plugins, panorering och master fader* till 44.1 kHz. C*lip gain, crossfades, fade in och fade out* inställningarna kunde inte importeras med hjälp av *session data*. Istället användes skärmdumpar av dem i 96 kHz, för att sedan efterlikna dessa så mycket som möjligt i 44.1 kHz. All *clip gain, crossfades, fade in och fade out* gjordes med *grid* påslaget. Detta minimerade risken att redigering i det båda projekten, 96 kHz respektive 44.1 kHz skulle bli olika. Slutligen, när inspelningarna exporterades från Pro Tools, användes en *spectrum analyser* för att bekräfta skillnaden av frekvensomfång inom infraljud mellan 44.1 kHz och 96 kHz.

4.5.3 Tre plugins

Under mixningen användes 3 *plugins*, vilka var *reverb* (d-verb), *mastering plugin* (Maxim) samt *spectrum analyser* (Span). *Reverb* användes på två *auxiliary spår*, ett *auxiliary spår* för elgitarr och akustisk gitarr, det andra för slagverk. *Auxiliary spåren* anslöts via bussar till ljudspåren med instrument inspelningar. *Mastering plugin* användes på *master fader* för att justera volymen på inspelningarna innan de exporterades till Wav-filer.

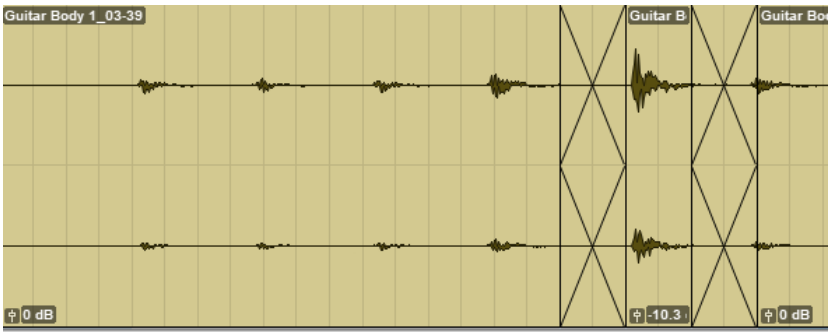
Ju fler *plugins* som används, desto större var risken att förändra, ta bort eller lägga till nya frekvenser från de ursprungliga inspelningarna. Eftersom studien grundar sig i att höra eventuella skillnader mellan inspelningarna i 44.1 kHz respektive 96 kHz, användes därför så få plugins som möjligt. Mixer och plugins kan ses i figur 17 nedan.



Figur 17: Mixer och plugins.

4.5.4 Redigering av ljudfiler från inspelningarna

För redigering av själva ljudfilerna användes *clip gain, crossfades, fade in* och *fade out*. Även om båda inspelningarna gjordes med samma inställningar på Zoom H1 och Motu M2, så kommer olika inspelningar aldrig låta exakt likadana på grund av den mänskliga faktorn. Dock var det tacksamt att kompositionen var relativt repetitiv i sin helhet, vilket innebar att de delar av inspelningarna som lät mest likt varandra kunde kopieras genom kompositionens gång. All redigering gjordes med *grid* påslagen i Pro Tools, vilket resulterade i att redigeringen av 96 kHz och 44.1 kHz blev så lika varandra som möjligt. Exempel på *clip gain* och *crossfades* som gjordes med *grid* påslaget i 16:delar kan ses nedan i figur 18.

i. 

Figur 18: *Clip gain* samt *crossfades* med *grid* påslaget i Pro Tools.

Kryssen som ses i figur 18 är *crossfades* med *grid* påslaget, *clip gain* kan ses längst ner i vänstra hörnet samt två gånger längst ner i det högra hörnet. *Clip gain* var användbart då mixen gjordes utan kompressor. Ljudfilen i figur 18 härrör från den akustiska gitarr som användes som trumma. Det näst sista slaget sänktes med 10.3 decibel med hjälp av clip gain.

4.5.5 Slutet av mixningen samt exportering

Efter att båda inspelningarna var färdigredigerade finjusterades volymen på de olika ljudspåren, *master fader* och på *mastering plugin*. Som tidigare nämnts gällande redigering kommer olika inspelningar aldrig att låta exakt likadant på grund av den mänskliga faktorn. Detta gäller inte bara för hur korrekta inspelningarna är, med hänsyn till kompositionens noter, utan även för volymen i det olika inspelningarna. De exporterade Wav-filerna jämfördes därför och små ändringar i mixen gjordes i både 96 kHz och 44.1 kHz. Ändringarna gjordes tills det båda inspelningarna upplevdes så lika som möjligt, både i avseende till volym och beträffande skillnader mellan det olika instrumenten. Små ändringar i mixen gjordes i både 96 och 44.1 kHz tills de båda inspelningarna upplevdes vara så lika varandra som möjligt, både avseende volym generellt och även beträffande eventuella skillnader mellan olika instrument. När båda projekten var färdigmixade, användes *spectrum analyser* på *master fader* i båda inspelningarna. En tydlig skillnad i ultraljud kunde ses mellan inspelningen med 44.1 kHz respektive 96 kHz i samplingsfrekvens, se figur 19 och 20.

Figur 19: 44.1 kHz sample rate, master fader, spectrum analyser.

Figur 20: 96 kHz sample rate, master fader, spectrum analyser.

4.6 Pilotstudie

Pilotstudien utfördes på två testdeltagare, båda 28 år. Innan pilotstudien började informerades testdeltagarna om vad som skulle studeras. Det klargjordes också för dem att de fick avsluta testet när de själva ville. I pilotstudien användes wav-filer som uppspelningsformat. För att undvika användning av två olika projekt till en DAW under studien testades om en exporterad wav-fil innehåller frekvenser över 20 000 Hz genom att kopiera wav-filen med musiken in i ett 96 kHz projekt i Pro Tools. Därefter användes en frequency analyser som visade att frekvenser över 20 000 Hz är närvarande i wav-filen. Anledningen till varför Wav-filer användes var att det skulle spara tid under testet jämfört med att stänga av och öppna projekt i Pro Tools med olika samplingsfrekvenser. Pro Tools kan inte ha två projekt öppna med olika samplingsfrekvenser samtidigt. Samplingsfrekvensen behövde dock fortfarande ändras på ljudkortet från 44.1 kHz till 96 kHz, men detta upplevdes av författarna gå tillräckligt fort för att studien smidigt skulle kunna genomföras. Eftersom frekvensomfånget mellan olika hörlurar samt högtalare kan variera, valdes inför studien ett par hörlurar som alla deltagare använde under testet, vilket var ett par AKG K701. Modellen kan spela frekvenser så högt som 39800 Hz (AKG u.å). Figur 21 visar att wav-filen kan spela det högre frekvenserna som är nödvändig för att utföra testet.

Figur 21: Illustrerar att wav-filen som användes under pilotstudien innehåller höga frekvenser.

För att de höga frekvenserna skulle kunna höras användes ett par AKG K701 hörlurar som kan spela frekvenser mellan 10 Hz till 39800 Hz (AKG u.å). Studiens testdeltagare fick använda en mobiltelefon som hade spelet *Temple Defender* installerat. Testdeltagarna spelade *Temple Defender* samtidigt som de lyssnade på 96 kHz inspelningen (symboliseras som spelomgång 1 i figur 22) av kompositionen. Efter det spelades 44.1 kHz inspelningen (symboliseras som spelomgång 2 i figur 22) upp. Resultatet visade att testdeltagarna kunde höra vilken version som hade högst samplingsfrekvens. Detta kan ses i figur 22.

Figur 22: Resultatet av pilotstudien.

Efter pilotstudien upptäcktes det att svarsalternativet *Inget (instrument)* behövde läggas till på frågan *om du hörde skillnad, var det något specifikt instrument detta gällde?* i enkäten. Det upptäcktes också att deltagare inte kunde välja flera alternativ på den frågan, vilket åtgärdades. Testdeltagarna klagade på brus utifrån när de lyssnade på inspelningarna. Det var viktigt att den slutliga studien sedan skulle göras i ett tyst rum för att undvika annat inkommande ljud. En viktig upptäckt var att samlingsnamnet på wav-filerna måste tas bort. Det behövde göras för att testdeltagarna inte i förväg skulle veta vilken version de lyssnade på. Om de fått veta det skulle det kunnat påverka resultatet. Risk förelåg för placebo-effekt (Sarlin 2011).

5 Utvärdering

5.1 Presentation av undersökning

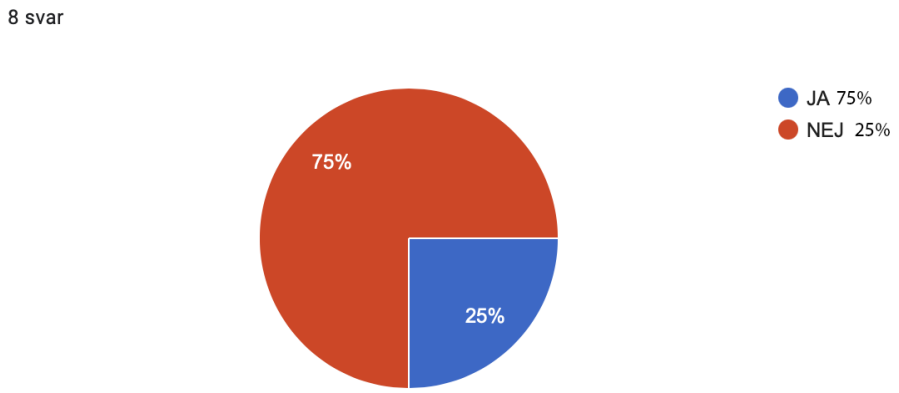
Studien gjordes på 8 deltagare. Alla testdeltagare gjorde testet individuellt. Samtliga testdeltagare var manliga studenter vid Högskolan i Skövde i åldrarna 22 till 27 år. Testdeltagarna fick först information om försöket, om enkätens upplägg och om deras rätt om frivillighet till att delta. Därefter fick de spela med tillhörande musik i det två olika versionerna, 44.1 kHz respektive 96 kHz. Därefter ombads det att svara på fyra frågor.

Viktigt att poängtera är att testdeltagarna inte visste vilken samplingsfrekvens de lyssnade på. Den första musikinspelningen de lyssnade på hade 96 kHz i samplingsfrekvens och den andra 44.1 kHz. Fyra frågor ställdes till deltagarna efter studien.

1. *Upplevde du någon skillnad i immersion mellan spelomgångarna?*
2. *Upplevde du någon skillnad i musiken mellan spelomgångarna?*
3. *Om du hörde skillnad, var det något eller några specifika instrument detta gällde?*
4. *Vilken av spelomgångarna upplevde du hade högst ljudkvalite?*

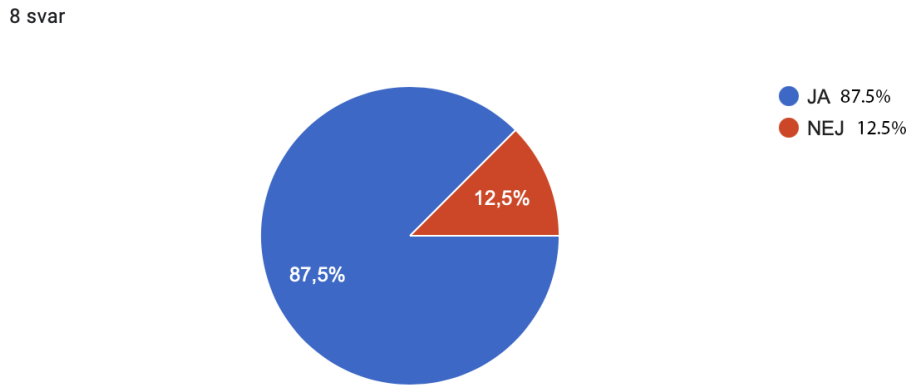
5.2 Analys

Författarna använde Sarlins studie (2011) som ett underlag som visar att människors hjärnor uppfattar ultraljud när de förekommer samtidigt som hörbara frekvenser. Resultatet på fråga 1 visar att 6 av 8 tyckte att de inte upplevde en större immersion mellan spelomgångarna. 2 av 8 upplevde att det kände en större immersion. Se figur 23.



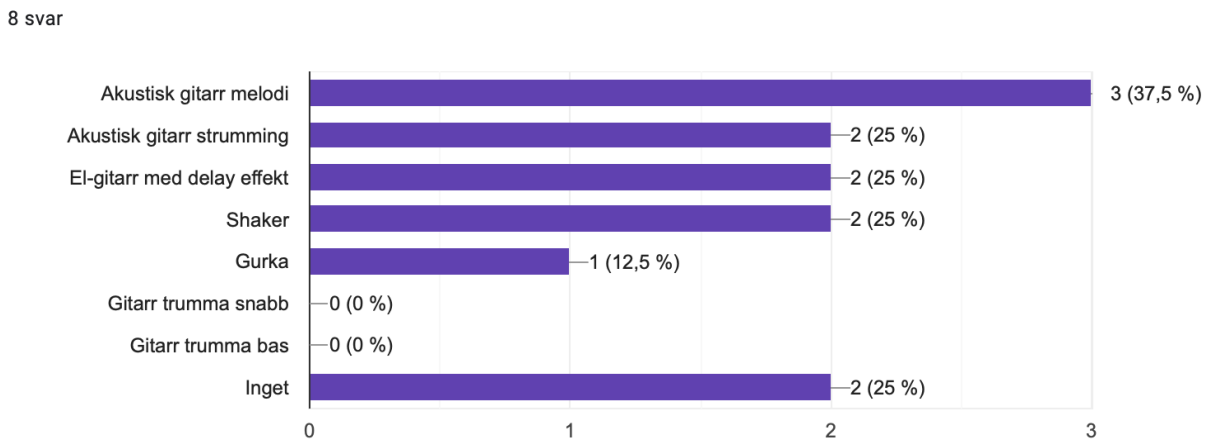
Figur 23: 6 av 8 personer svarade *Ja*.

Beträffande skillnad i musik mellan spelomgångarna uppgav sju deltagare att de hörde skillnad. En av deltagarna uppgav ingen upplevd skillnad. Se figur 24.



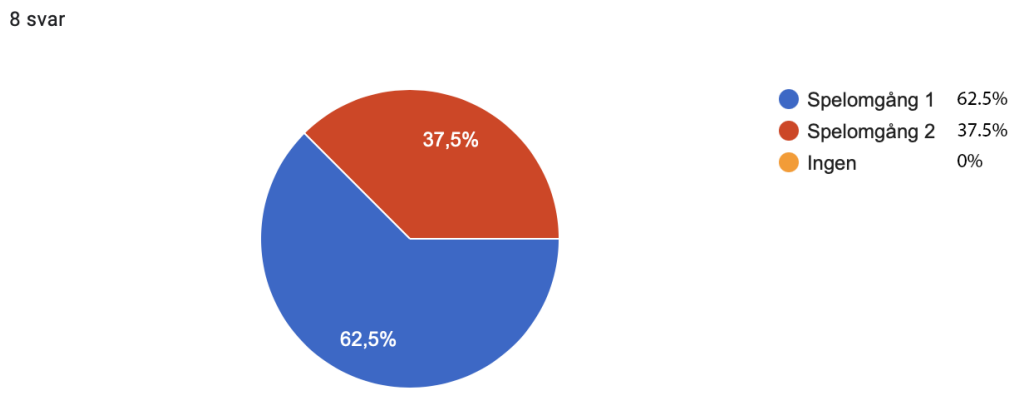
Figur 24: 7 av 8 svarade *Ja*.

På fråga tre varierade svaren. Tre hörde skillnad på instrumentet *Akustisk gitarr melodi*. Två personer hörde skillnad på *Akustisk gitarr strumming, Elgitarr med delay effekt* och *Shaker,* en svarade att hen hörde skillnad på *Shakern*. En testdeltagare angav svarsalternativ inget.



Figur 25: Upplevd skillnad gällande specifika instrument.

Beträffande frågan om spelomgångarnas ljudkvalitet så ansåg majoriteten av testdeltagarna att 96 kHz uppvisade högst ljudkvalitet medan 37,5 procent ansåg att 44.1 kHz i samplingsfrekvens var bästa kvalitet. Denna fråga står i relation till ABX testet som deltagarna i Pras och Guastavinos (2010) undersökning genomförde, eftersom det i båda undersökningarna ställs frågor om ljudkvalité och samplingsfrekvens. Frågan handlar om samma sak som Harris et al. (1998) studie, vilken visade att personer upplever en ljudmässigt ökad kvalité i samband med högre samplingsfrekvens.



Figur 26: Upplevda skillnader mellan olika spelomgångar.

5.3 Slutsatser

Med hjälp av enkäten kunde författarna besvara frågeställningen, *Hur påverkas spelupplevelsen av olika samplingsfrekvenser i musiken, 44.1 respektive 96 kHz?*

Resultatet visade att spelupplevelsen påverkades generellt inte av musik uppspelad i olika samplingsfrekvens.

62.5% av testdeltagarna tyckte ljudkvalitén var bättre med 96 kHz i samplingsfrekvens. 75 % upplevde att de inte kände någon skillnad gällande immersion, alltså kunde en tendens ses att ljudupplevelse och immersion inte har någon koppling.

Studien visade att sju av åtta upplevde en skillnad på musiken. Detta behöver dock inte betyda att deltagarna hörde divergens på stereobilden eller ljudkvalitén i helhet, utan kan eventuellt betyda att de olika inspelningarna var inspelade på skiftande sätt. Mixningen eller instrumentens klang kan ha skilt sig åt.

Eftersom tre personer hörde skillnad på *Akustisk gitarr melodi* fanns det en tendens att deltagarna trodde sig höra höga frekvenser på detta instrument trots att det inte fanns. Detta kan bero på placeboeffekten. Vad gäller de andra instrumenten, det vill säga *Akustisk gitarr strumming, Elgitarr med delay effekt* och *Shaker,* var det enbart två testdeltagare på vardera instrumentet som hörde skillnad.

Instrumentet guiro (även kallad gurka) hade många höga frekvenser men fick bara en röst på frågan *om du hörde skillnad, var det något eller några specifika instrument detta gällde?*.

Vad gäller instrumentet *Akustisk gitarr melodi* fick det 3 röster trots att det inte hade höga frekvenser i sig.

6 Avslutande diskussion

6.1 Sammanfattning

Frågeställningen som undersöktes löd *Hur påverkas spelupplevelsen av olika samplingsfrekvenser i musiken, 44.1 kHz respektive 96 kHz?*

En kvantitativ studie genomfördes. Författarna skapade samt spelade in en artefakt i 44.1 kHz samt 96 kHz. Testdeltagarna lyssnade på båda variationerna av artefakten direkt efter varandra samtidigt som de spelade mobilspelet Temple Defender.

Deltagarna fick skatta huruvida de upplevde skillnad mellan kompositionerna som helhet eller ej. Vid upplevd skillnad fick deltagarna ange inom vilka specifika delar av kompositionerna som skillnaderna ansågs finnas. Resultatet visade att det förelåg en skillnad i upplevd ljudkvalite mellan 44.1 kHz respektive 96 kHz beträffande spelupplevelsen som helhet. Bland de deltagare som uppgett att de hört en skillnad, uppgav flest att de hört skillnad vid jämförande av akustisk gitarrmelodi inspelad i respektive frekvens. Inom två aspekter, gitarr trumma snabbt samt gitarr trumma bas, uppgav deltagarna att de ej hört någon skillnad avseende just dessa parametrar, även om de tyckte sig uppleva en skillnad mellan inspelningarna som helhet.

6.1.1 Samhällsetiska aspekter

Högre samplingsfrekvenser bidrar enligt denna studies resultat till en upplevd ökad kvalité generellt. Det bedöms också medföra en tydligare stereobild. När spel skapas för människor med till exempel nedsatt syn blir ljudbilden särskilt viktig för att spelet ska fungera för dem. Utifrån ett tillgänglighetsperspektiv kan således ljudkvalitén anses vara en viktig samhällsetisk aspekt.

Avseende konsumenter av musikinstrument samt ljudinspelningsutrustning föreligger en etiskt faktor beträffande ekonomi. Människor som exempelvis skall köpa ljudkort, mikrofon eller andra dylika produkter kan ha nytta av att veta mer om hur olika samplingsfrekvenser påverkar kvaliteten. Om en faktisk skillnad i kvalité upplevs föreligga, kan det utifrån en konsuments perspektiv sannolikt upplevas befogat att betala en större summa pengar för mer avancerad utrustning. Hade däremot ingen skillnad noterats hade det kunnat anses oetiskt av försäljare gentemot konsumenter att marknadsföra utrustningen som bidragande till en ökad upplevd kvalité.

6.2 Diskussion

Resultatet visade på att det flesta personer i undersökningen kunde höra vilken version som var inspelad med 96 kHz i samplingsfrekvens. En av anledningarna till att många svarat spelomgång 1 på frågan: *Vilken av spelomgångarna upplevde du hade högst ljudkvalite?* kan vara att de har erfarenhet av att jobba med musik eller ljudproduktion. En annan betydande faktor skulle kunna vara att testdeltagarnas ålder är mellan 22 till 27 år. Det betyder att deras hörsel inte har hunnit försämras avsevärt. Pras och Guastavino (2010, s. 2) använde personer med en medelålder runt 30 i deras undersökning. Deltagarna i deras

studie kunde höra skillnad på ett musikstycke inspelat i 44.1 och 88.2 kHz. Detta betyder således att testdeltagarna i denna studie, troligtvis hade lika stora chanser att höra ultraljud som de i Pras och Guastavinos undersökning. Vidare hade undersökningar med olika åldersgrupper varit viktig för att jämföra ifall äldre personer ej hade kunnat höra dessa skillnader i lika stor utsträckning som yngre.

Pras och Guastavinos (2010, s. 2) nämner att deltagarna i deras undersökning kunde höra skillnad på olika instrument. Ett liknande resultat framkom i vår studie då resultatet visade på att deltagarna hörde skillnad gällande ett antal instrument. Pras och Guastavino använde även ett flertal olika stycken, vilka varierade i genre, instrument samt intensitet. Det orkestrala stycket var det stycke som flest deltagare i deras undersökning hörde skillnad på, gällande 44.1 kHz kontra 88 kHz som samplingsfrekvens. I studien som genomfördes av Harris et al. (1998), visades ytterligare underlag för att inspelning av orkestrala stycken upplevs vara de stycken där tydligast skillnad hörs gällande 44.1 kHz kontra högre samplingsfrekvenser. En teori är att eftersom stereobilden blir tydligare i högre samplingsfrekvenser, är detta en skillnad som upplevs starkare när större ljudbilder spelas in. Orkesteruppsättningar är bland de större uppsättningar som finns i dagsläget.

Pras och Guastavino genomförde även ett ABX-test där deltagarna hade möjligheten att lyssna på de olika inspelningarna flera gånger om de så önskade. I vår studie spelades de två olika inspelningarna upp endast en gång. Detta skedde samtidigt som testdeltagarna var upptagna med att spela Temple Defender. Spelarnas fokus blev på detta sätt splittrade mellan spelande och lyssnande. Det finns en möjlighet att resultatet blivit ett annat ifall deltagarna kunnat fokusera endast på musiken, eller kunnat lyssna flera gånger i rad. Studiens syfte var dock just att undersöka upplevelsen av musik i samband med spel och ej de två som skilda företeelser.

De hittills viktigaste faktorerna gällande att höra skillnad mellan olika samplingsfrekvenser är ålder, erfarenhet inom musik och ljudproduktion samt hur stor ljudbild och antal instrument som spelas in när man jämför de olika samplingsfrekvenserna.

Högre samplingsfrekvenser kan minska risken för harminiskörvrängning eller uppkomst av AID. Vest och Scheelke menar att en samplingsfrekvens över cirka 60 kHz troligtvis skulle leda till att all aliasing problem försvinner. De menar också att fenomenen inte innebär någon större skillnad vid 196 kHz eller 96 kHz (u.åa, s. 1). Därav kan det argumenteras för att högre samplingsfrekvenser bör användas oavsett upplevd skillnad i ljudkvalité eller ej.

6.3 Framtida arbete

6.3.1 Ett par dagar extra

Om den här studien kunde fortgått några dagar till skulle inspelningen av den akustiska gitarren gjorts om både vad gäller 44.1 kHz och 96 kHz, eftersom inspelningen medförde icke önskvärda och störande ljud. Eventuellt medförde dessa störande ljud att skillnaderna i detaljer mellan inspelningen i 44.1 kHz kontra 96 kHz ej hördes lika tydligt. Ett annat alternativ kunde vara att minska det störande ljudet med hjälp av EQ-plugin. Dock var det en

teknik som inte var aktuell i det här fallet eftersom det kunde orsakat att frekvenser adderats eller subtraherats från orginalinspelningen.

6.3.2 En större studie

Förutom ovan nämnda förslag på åtgärd avseende att reducera brus vid inspelning skulle flera olika mikrofoner kunnat användas vid inspelning. Mikrofonen Zoom H1 fungerade väl när det gällde att spela in höga frekvenser. Däremot var den inte specialiserad för specifika typer av instrument. Ifall fler olika mikrofoner använts är det möjligt att större hörbara skillnader för specifika instrument mellan samplingsfrekvenser skulle uppstå.
Det hade varit intressant att spela in flera olika stycken med varierande musikinstrument. Som tidigare nämnts hade ett orkestralt stycke varit intressant att spela in då andra studier indikerade att störst upplevda skillnad mellan samplingsfrekvenser hördes i orkestrala inspelningar.

I en tidigare nämnd studie av Koubori (et al- 2010, s. 4-5) testades om deltagarna kunde höra frekvenser ovanför 20 kHz. Resultatet visade att deltagarna kunde höra upp till 24 kHz. Dock gällde det enbart vid användning av högtalare i ett helt ljudisolerat rum. Att jämföra vår studie där studiedeltagarna använde hörlurar med ett annat test med enbart högtalare skulle kunna ge svar på om resultaten skilde sig åt eller inte mellan hörlurar eller högtalare. En ytterligare fråga som kunde varit aktuell att undersöka är om resultaten hade blivit detsamma ifall testpersonerna spelat ett för dem ett välbekant spel istället för spelet Temple Defender. Eftersom deltagarna ej sett detta spel tidigare krävde det sannolikt större koncentration av dem vilket kan ha påverkat upplevelsen som helhet.

6.3.3 Generella reflektioner över teknologisk utveckling & samplingsfrekvens

Genom sin tid i digitala medier som exempelvis, film, tv, radio och dataspel har ljud och musik inneburit många begränsningar, inte minst avseende lagringsutrymme. Kompromisser har ofta varit tvungna att göras för att det över huvud taget skall vara möjligt att använda ljud och musik inom dataspel. Att spela in i högre samplingsfrekvenser kräver mer utrymme samt prestanda av datorer. Samplingsfrekvensen 44.1 har nu varit industristandard i ungefär fyrtio år. När denna standard bestämdes var det inte teknologiskt möjligt att använda 88, 96 eller 192 kHz vid inspelning. Dock minskar begränsningarna och musikens kvalité ökar i takt med den teknologiska utvecklingen.

Således vore det en rimligt slutsats att musikens kvalité alltid kommer att öka i takt med den tekniska utvecklingen. Inom världen av musikproduktion är många producenter fortfarande tveksamma till huruvida det är nödvändigt med högre samplingsfrekvenser än 44.1 kHz. Dock finns det tydliga underlag för att en samplingsfrekvens över cirka 60 kHz skulle lösa många av de problemen som kan uppstå via harmonisk förvrängning, AID och liknande problem. Om till exempel VR-spel inom en snar framtid blir mer vanligt skulle antagligen en högre samplingsfrekvens vara att föredra eftersom stereobilden blir tydligare i en högre frekvens. Ifall en högre samplingsfrekvens ej användes i VR-spel skulle eventuellt en dissonans mellan ljud och bild kunna uppstå. Människor vill ofta ha bättre och snabbare teknologi, högre bildkvalité och bättre internetanslutning. Ibland skiljer sig inte den nya tekniken nämnvärt från den gamla. Trots det vill många människor ändå ha den senaste tekniken.

Referenser

Adobe. (u.å.). *Photoshop*. https://www.adobe.com/se/products/photoshop.html?sdid=8JD95K3M&mv=search&ef_id=Cj0KCQjw6ar4BRDnARIsAITGzlBmC9OmuGC-5c4KFL2e7yfqMBtjzCmwvQyKXqA8gzPzO77jcLe1sfwaAs0iEALw_wcB:G:s&s_kwcid=AL!3085!3!393367673656!e!!&gclid=Cj0KCQjw6ar4BRDnARIsAITGzlBmC9OmuGC-5c4KFL2e7yfqMBtjzCmwvQyKXqA8gzPzO77jcLe1sfwaAs0iEALw_wcB [2020-07-12].

Agarwal, A,. & Lang, J. (2005). Foundations of Analog and Digital Electronic Circuits. *Elsevier Inc.*

AKG. (u.å.). *AKG K701*. https://www.akg.com/Headphones/Professional%20Headphones/K701.html [2020-05-23].

Apple (2019a). *Logic Pro X: Get started with audio tracks*. https://support.apple.com/kb/PH13418?locale=sv_SE [2020-08-02].

Apple (2019b). *Logic Pro X: Snap items to the grid*. https://support.apple.com/kb/PH13026?locale=en_US [2020-08-02].

audacityteam (2020). *Fade and Crossfade*. https://manual.audacityteam.org/man/fade_and_crossfade.html [2020-08-02].

Avid (2020). *Pro Tools Overview*. https://www.avid.com/pro-tools [2020-05-24].

Biasutti, M. (1997). Sharp low- and high-frequency limits on musical chord recognition. *Department of Educational Science, PhD program, University of Padua, Piazza Capitaniato 3, Italy*.

Bland, D. (2002). Method and system for sampling rate conversion in digital audio applications. *WIND RIVER SYSTEMS, INC., CALIFORNIA*.

Bogdan, M. (2009). Sampling rate and aliasing on a virtual laboratory. *University of Lucian Blaga Sibiu, Faculty of Engineering, Str. Emil Cioran, no.4, 550025 Sibiu, Romania*

Cocos (u.å.). *Cocos2d-x*. https://www.cocos.com/en/ [2020-07-12].

Computerhope (2020). *Plugin*. https://www.computerhope.com/jargon/p/plugin.htm [2020-08-03].

Chang, K., Kim, G. & Kim, T. (2007). Video Game Console Audio: Evolution and Future Trends. *Computer Graphics, Imaging and Visualisation, Bangkok, Thailand*

Denscombe, M. (2003). The Good Research Guide, for small-scale social research projects: *Second Edition, Open University Press, Maidenhead, Berkshire, England.*

Gilliland, S. (2017). *Importing Session Data in Pro Tools*.
https://www.pro-tools-pc.com/importing-session-data-pro-tools/ [2020-07-28].

Harris, M, N., Kelly, R., McLeod, D, A. & Story, M, J. (1998). Effects in High sample Rate Audio Material. *Presented at 20th Tonmeistertagung, Karlsruhe, Germany*.

HomeBrewAudio (2020). *What is a Bus in Audio Recording?*.
https://www.homebrewaudio.com/what-is-a-buss-in-audio-recording/ [2020-08-02].

Koubori, M., Ashihara, K., Kiryu, S., Kyoso, M. & Omata, M. (2010). Psycho-Acoustic Measurement and Auditory Brainstem Response in the Frequency range between 10 kHz and 30 kHz. *AES Convention paper 8294, Presented at the 129th Convention 2010 November 4-7*.

Margaret, R. (2005). *crossfade*.
https://whatis.techtarget.com/definition/crossfade [2020-08-02].

Marks, A. (2012). The complete guide to Game Audio: For Composers, Musicians, Sound Designers, and Game Developers. *CMP Books, Lawrence, Kansas, USA*.

Meseu Das Comunicacoes (u.å.). *ADC and DAC*
http://www.cmm.gov.mo/eng/exhibition/secondfloor/MoreInfo/ADConverter.html[2020-08-02].

Nationalencyklopedin. (u.å[a]). *Analog–digitalomvandlare*.
http://www.ne.se/uppslagsverk/encyklopedi/lång/analog-digitalomvandlare [2020-06-21].

Nationalencyklopedin. (u.å[b]). *Decibel*.
https://www.ne.se/uppslagsverk/encyklopedi/lång/decibel [2020-03-10].

Nationalencyklopedin. (u.å[c]). *Dolby Digital*.
https://www.ne.se/uppslagsverk/encyklopedi/lång/dolby-digital [2020-03-17].

Nationalencyklopedin. (u.å[d]). Filformat.
https://www.ne.se/uppslagsverk/encyklopedi/lång/filformat [2020-03-10].

Nationalencyklopedin. (u.å[e]). *Frekvens*.
https://www.ne.se/uppslagsverk/encyklopedi/lång/frekvens [2020-03-09].

Nationalencyklopedin. (u.å[f]). *Hertz*.
https://www.ne.se/uppslagsverk/encyklopedi/lång/hertz [2020-03-10].

Nationalencyklopedin. (u.å[g]). *Lågpassfilter*.
https://www.ne.se/uppslagsverk/encyklopedi/lång/lågpassfilter [2020-03-10].

Nationalencyklopedin. (u.å[h]). *Microsoft Corporation*.
https://www.ne.se/uppslagsverk/encyklopedi/lång/microsoft-corporation [2020-06-30].

Nationalencyklopedin. (u.å). *MIDI*.
https://www.ne.se/uppslagsverk/encyklopedi/lång/midi [2020-03-10].

Nationalencyklopedin. (u.å). *Mp3*. https://www.ne.se/uppslagsverk/encyklopedi/lång/mp3 [2020-03-10].

Nationalencyklopedin. (u.å). *Sampling*.
https://www.ne.se/uppslagsverk/encyklopedi/lång/sampling [2020-06-10].

Nationalencyklopedin. (u.å). *Volym*.
https://www.ne.se/uppslagsverk/encyklopedi/lång/volym-(ljud) [2020-06-10].

Nationalencyklopedin. (u.å). *Enstämmighet*.
https://www.ne.se/uppslagsverk/encyklopedi/lång/enstämmighet [2020-08-05].

Nationalencyklopedin. (u.å). Sinuston.
https://www.ne.se/uppslagsverk/encyklopedi/lång/sinuston [2020-08-05].

Nationalencyklopedin. (u.å). andbred.
https://www.ne.se/uppslagsverk/encyklopedi/lång/bandbredd [2020-08-05].

Palm Mute. (u.å). *Palm Mute Guitar in Heavy Metal*.
http://www.fretjam.com/palm-mute-guitar.html [2020-08-6].

Brooks, D. (2018). *What does a reverb effect do? | The Basics*
. https://reverb.com/news/what-does-a-reverb-effect-do-the-basics [2021-06-09].

Pras, A. & Gustavino, C. (2010). Sampling Rate Discrimination: 44.1 kHz vs 88.2 kHz. *Audio Engineering Society, Convention Paper, Presented at the 128th Convention, London, UK*.

Practical Music Production (u.å). *What is an Audio Compressor?*.
https://www.practical-music-production.com/audio-compressor/ [2020-08-02].

Producelikeapro (2019). *The Best Mastering Plugins*.
https://producelikeapro.com/blog/best-mastering-plugins/ [2020-08-04].

ProToolsProduction (2016). *Pro Tools 10: Clip Gain*.
https://www.protoolsproduction.com/clipgain/ [2020-07-28].

Sarlin, T. (2011). Do we perceive ultrasound well enough to use it while recording and reproducing music? *Luleå University of Technology, Department of Arts, Communication and Learning*.
https://www.diva-portal.org/smash/get/diva2:1020605/FULLTEXT02

Shapiro, K., Wilensky, H. & Elmer, L. (1968). Comprehensive Psychiatry: Study of the

placebo effect with a placebo test. *Elsevier Inc.*

Sweetwater (2013). *Aux Track*. https://www.sweetwater.com/insync/aux-track/ [2020-07-28].

Sweetwater (2005). *Master Fader*. https://www.sweetwater.com/insync/master-fader/ [2020-08-02].

TeachMeAudio (2020a). *Panning*. https://www.teachmeaudio.com/mixing/techniques/panning [2020-07-28].

TeachMeAudio (2020b). *Delay*. https://www.teachmeaudio.com/mixing/equipment/effects/delay [2020-08-04].

Tektronix (u.å). *What is a spectrum analyzer?*. https://www.tek.com/spectrum-analyzer [2020-08-03].

Topala, R. (u.å.). *RobTop*. http://www.robtopgames.com [2020-05-23].

Vest, M, Sheelke, P. (u.åa). The tradeoff of 192 kHz sampling. *Digital Audio Denmark*.

Vest, M, Sheelke, P. (u.åb). How to avoid Aliasing Intermodulation Distortion (AID) The invisible distortion in digital audio. *Digital Audio Denmark*.

Waren, H. (2019). Should you be using a higher sample rate? FAQ Friday. https://producelikeapro.com/blog/should-you-record-at-a-higher-sample-rate-faq-friday/ [2020-03-12].

Winters, L. (2014). *Gain and Levels*. https://www.practical-music-production.com/audio-compressor/ [2020-08-02].

Zawistowski, T. & Shah, P. (2018). *An Introduction to Sampling Theory*. http://dspace.bmsit.ac.in/xmlui/bitstream/handle/123456789/297/DC_Casestudy_parallel course_IA1.pdf?sequence=1&isAllowed=y [2020-08-03].

ZOOM. (u.å.). *ZOOM H1*. https://www.zoom-na.com/products/field-video-recording/field-recording/zoom-h1-handy-recorder#specs [2020-05-23].