

Ingeniørhøjskolen i Århus
IKT
Dalgas Avenue 2
8000 Århus C

20. december 2004

4. Semesterprojekt
Projektrapport

MyP3000

I4PRJ4 E2004

Gruppe 4:

Benjamin Sørensen, 02284

Tomas Stæhr Berg, 03539

Nikki Ashton, 01087

Jonas Livbjerg, 02797

Vejleder:

Sten Hansen

1 Versionshistorie

Ver.	Data	Initialer	Beskrivelse
0.1	28.10.2004	BS	Opsætning af dokumentet i LaTeX
1.0	03.12.2004	TB	Påbegyndelse af projektrapport
1.1	17.12.2004	TB	Review rettelser
1.2	17.12.2004	NA	smårettelser samt engelsk resumé

2 Godkendelsesformular

Forfatter(e):	Benjamin Sørensen(BS), Tomas Stæhr Berg(TB), Nikki Ashton(NA), Jonas Livbjerg(JL)
Projektnummer:	4. semesterprojekt
Antal sider:	26
Godkendes af:	Sten Hansen

Ved underskrivelse af dette dokument anses semesterprojektet som afsluttet og rettidigt afleveret. Underskrivelsen er udtryk for ønske om individuel evaluering af projektet, herunder eksaminering.

Dato og underskrift:

Benjamin Sørensen

Tomas Stæhr Berg

Nikki Ashton

Jonas Livbjerg

Sten Hansen
Vejleder

Indhold

1	Versionshistorie	1
2	Godkendelsesformular	1
3	Resume	4
3.1	Dansk - Resume	4
3.1.1	Problemstilling	4
3.1.2	Formål	4
3.1.3	Materialer og metoder	4
3.1.4	Resultater	5
3.1.5	Konklusion	5
3.1.6	Ordliste	5
3.2	English - Abstract	5
3.2.1	Problem	6
3.2.2	Rationale	6
3.2.3	Resources and methods	6
3.2.4	Results	7
3.2.5	Conclusion	7
4	Forord	8
5	Indledning	9
6	Projektbeskrivelse	10
6.1	Afgrænsning	10
6.2	Gennemførelse	10
6.3	Metoder	12
6.3.1	Iterativ udvikling	12
6.3.2	Komprimeringsmetoder	13
6.3.3	Testmetoder	16
6.3.4	Metodevalg	17
6.4	Analyse og design	18
6.4.1	Funktionsanalyse	18
6.4.2	Arkitektur design	19
6.4.3	Dokumentation	22
6.5	Udviklingsværktøjer	22
6.6	Eksterne projektkrav	23
6.7	Resultater	23
6.7.1	Diskussion af opnåede resultater	23
6.7.2	Projektets fortræffeligheder	23
6.8	Forbedringer	24
6.8.1	Funktionelt	24

<i>INDHOLD</i>	3
6.8.2 Projektarbejde	24
7 Konklusion	25
8 Referencer	26

3 Resume

3.1 Dansk - Resume

I det følgende afsnit er et kort resumé af projektarbejdet. Afsnit 3.2 indeholder resuméet oversat til engelsk.

3.1.1 Problemstilling

I forbindelse med semesterprojektet i kurset I4PRJ4 skal der udvikles en MyP3 encoder og en decoder. Projektet skal udføres som et udviklingsprojekt, hvor et nyt problemområde skal udforskes. Der skal ikke laves et færdigudviklet system, kun en prototype. Erfaringerne fra projektet skal danne grundlag for en vurdering af det nye problemområde. Der ligges derfor mere vægt på udforskningen af området, end på udfærdigelsen af produktet.

3.1.2 Formål

Formålet med projektet er at anvende viden fra tidligere signalbehandlingskurser, til at foretage en komprimering af lyd. Med komprimering menes, kontrolleret tab af præcision, uden hørbar tab af kvalitet.

Projektet foretages i grupper. Det er derfor vigtigt at kunne samarbejde, planlægge og udvikle et projekt som en gruppe.

3.1.3 Materialer og metoder

Dette afsnit beskriver kort hvilke materialer og metoder vi har brugt.
Materialer:

- Oplæg fra Henrik Karstoft - Specifikation til MyP3 standarden.
- Lydopfattelse [Torben Poulsen 1998]
- Perceptual Coding of Digital Audio
- Matlab
- MiKTeX
- CVS

Disse metoder er anvendt:

Udviklingsmæssig:

Iterativ udvikling.

Psykoakustiske modeller:

Simultan Maskering.

Undersøgelse af høretærskel.

3.1.4 Resultater

Foruden denne projektrapport skal der udarbejdes en projektdokumentation indeholdende:

- Målsætning (papir).
- Systemarkitektur (papir).
- Dokumenteret kildekode (digital).
- Lyttetest (papir).
- Eksempelfiler (digital).

Som digitalt medie er der valgt cd-rom.

3.1.5 Konklusion

Vi har udviklet et funktionelt produkt. Der er gennemgået dobbelt blindtest på det ikke-komprimerede og det komprimerede, der viser at der i nogen grad kan høres forskel. Komprimeringsgraden ligger mellem 0 og 50 procent i forhold til det oprindelige.

Vi er tilfredse med produktet, og projektarbejdet der ligger bag. Vi har fungeret godt som gruppe, og udnyttet hinandens individuelle forcer optimalt.

3.1.6 Ordliste

- MyP3: Lydformatet vi benytter.
- M fil: Fil med Matlab kode.
- Psykoakustisk model: Metode til lydkomprimering.
- LaTeX: Program til tekstbehandling.
- CVS: Concurrent Versions System - Program til versionsstyring.
- IDE: Integrated Development Environment - Integreret udviklingsmiljø.

3.2 English - Abstract

The following section is an abstract of our work on the 4th semester IKT project.

3.2.1 Problem

In relation to our semester project in the course I4PRJ4, we are to develop a MyP3 encoder and decoder.

The project will be developed as a proof of concept project, where a new problem-domain is to be researched. Experience learned from this project will be the basis for a re-assessment of the problemdomain. Because of this, we will concentrate more on researching the domain, than making a final product.

3.2.2 Rationale

The rationale of the project is to use knowledge from earlier signalprocessing courses to compress sound. Defining compression as a controlled loss of precision, without a hearable loss of quality.

The project is developed in teams. Therefore it is important to be able to work together, plan, and complete a project as a team.

3.2.3 Resources and methods

This section briefly describes the resources and methods we have used.

Resources:

- Introduction from Henrik Karstoft - Specifications of the MyP3 standard.
- Lydopfattelse [Torben Poulsen 1998]
- Perceptual Coding of Digital Audio
- Matlab
- MiKTeX
- CVS

Methods we have used:

In development:

Iterationel development.

Psychoacoustic models:

Simultaneous masking.

Research of hearing threshold.

3.2.4 Results

In addition to this project document, elaborate project documentation has to be made, containing:

- Ambitions (paper).
- System architecture (paper and digitally).
- Documented sourcecode (digitally).
- Listening test (paper).
- Example files (digitally).

We have chosen CD-ROM as digital medium.

3.2.5 Conclusion

We have developed a functional product. Der er gennemgået dobbelt blindtest på det ikke-komprimerede og det komprimerede, der viser at der i nogen grad kan høres forskel. Komprimeringsgraden ligger mellem 0 og 50 procent i forhold til det oprindelige.

We are satisfied with the product and the research, and the work put into it. We have come together as a team, and benefited from individual skills.

4 Forord

Denne rapport omhandler et semesterprojekt afviklet på Ingeniørhøjskolen i Århus. Projektet er tilknyttet undervisningen på 4. semester, især kurset I4PRJ4. Kurset har bl.a. til formål at binde de øvrige fag sammen ved at gennemføre et udviklingsforløb.

Denne projektrapport er skrevet som afslutning på projektet. Formålet med rapporten er at adskille projektdokumentation fra produktokumentation. Rapporten indeholder således problemstilling, metodeovervejelser, projektbeskrivelse og efterfølgende resonementer.

Inspiration og skrivevejledning til rapporten er hentet fra studie metroen [Metro www].

Tidligt på semestret udmeldte Jeppe Hasager sig fra gruppen. Jeppe er derfor ikke krediteret på dokumenterne, og der indgår ikke materiale udarbejdet af Jeppe alene.

Projektet vurderes udelukkende på denne projektrapport.

5 Indledning

Emnet for dette projekt er lydkomprimering. Vi har fået til opgave at fremstille et produkt bestående af en encoder og en decoder, der fungerer efter MyP3 standarden.

Vi ønsker med dette projekt at opnå følgende:

- Tilegne os ny viden og praktisk erfaring igennem udforskning af signalbehandling i praksis. Dette gøres gennem forsøg med lydkomprimering vha. frekvensanalyse.
- Gennemføre et udviklingsforløb, der adskiller sig fra tidligere projekter.
- Anvende Matlab til implementering samt LaTeX og CVS.

De opstillede punkter skal opnås igennem en udvikling af dette produkt. Vi vil herefter anvende det interne navn MyP3000 for produktet. Produktet skal fungere som prototype.

Projektet løses i tre faser. Første fase vil udgøre en 80% løsning, anden fase en 100% løsning og tredje fase en 120% løsning. Dette betyder, at vi forventer at nå til og med anden fase, men forsøger at nå tredje fase.

Indholdet af de tre faser beskrives under afsnit 6.2 Gennemførelse.

Denne projektrapport består af en gennemgang af udviklingsforløbet. Den giver et indblik i de beslutninger vi har taget og hvorfor. Derudover giver den et billede af hvilke teknikker og metoder vi har tilegnet os igennem arbejdet med projektet.

6 Projektbeskrivelse

6.1 Afgrænsning

Dette afsnit beskriver projektets begrænsninger.

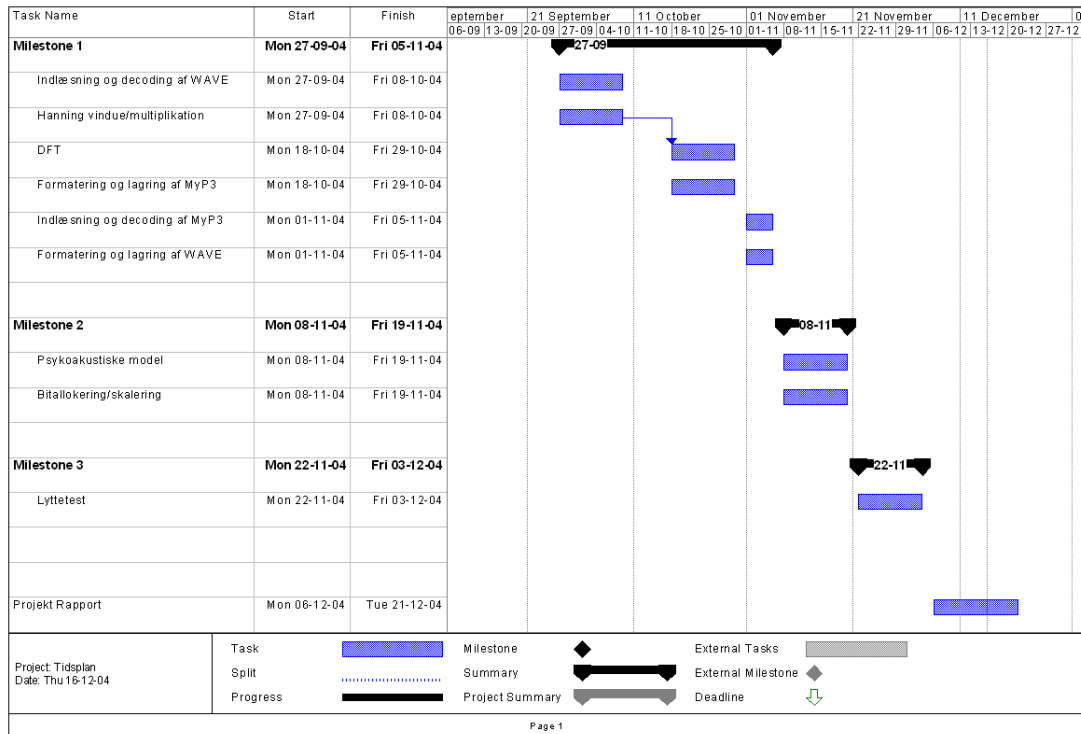
- Der skal laves en simpel MyP3 encoder/decoder med henblik på komprimering af lyd.
- Der skal benyttes en psykoakustisk model indeholdende forskellige udvalgte principper.
- Vi vil kun kode et mono signal.
- Vi vil kun arbejde med én samplingsfrekvens på 44.1kHz.
- Der skal ikke udvikles en decideret afspiller. MyP3 filen decodes tilbage til WAVE formatet og kan derefter afspilles.
- Der er tale om et proof of concept projekt. Det vil sige, at der ikke skal leveres et brugervenligt, gennemtestet produkt til en slutbruger.

6.2 Gennemførelse

Det første punkt i gennemførelsen af dette projekt, var at analysere opgaven og nedbryde den til mere håndterbare delopgaver. Vi fandt frem til 9 delmål, som vi fordelte til 3 milestones der repræsenterer henholdsvis vores 80%, 100% og 120% løsning omtalt i målsætningsdokumentet.

- Milestone 1: Konvertering til og fra MyP3.
 - Indlæsning og decoding af WAVE
 - Hanningvindue/multiplikation
 - DFT
 - Formatering og lagring af MyP3
 - Indlæsning og decoding af MyP3
 - Formatering og lagring af MyP3
- Milestone 2: Komprimering.
 - Psykoakustisk model
 - Bitallokering/skalering
- Milestone 3: Lyttetest.
 - Lyttetest

Der er herudfra lavet en tidsplan, se fig. 1. På tidsplanen ses de 3 milestones, med tilhørende delmål. For hver milestone angives hvilke delmål der indgår, samt hvor mange dage det vurderes at vare.



Figur 1: Tidsplan

Der er hver uge afholdt statusmøder med deltagelse af vejleder, hvor der bl.a. er diskuteret metoder og problemer.

I afsnittet 6.7 Resultater vil vi samle op på overholdelse af tidsplanen.

6.3 Metoder

Dette afsnit omhandler udvælgelsen af metoder og begrundelse for valgene. Vi vil ikke beskrive de forskellige udviklingsmetoder fra sidste semester, men blot beskrive iterativ udvikling, da vores udvikling er baseret herpå.

Derudover vil vi beskrive forskellige psykoakustiske- og testmetoder. Til sidst i afsnittet vil vi forklare valget af metoder imod en række kriterier.

6.3.1 Iterativ udvikling

Den iterative udviklingsproces er en videreudvikling af vandfaldsmodellen¹. Ideen er at gennemgå de fire faser (analyse, design, implementering og test) flere gange (heraf navnet). Efter hver iteration er der evaluering og evt. kundeleverance.

Fordele:

- Reducering af spildtid da Code and Fix minimeres.
- Forbedre tidsplanlægning og budgettering. Specielt vil man kunne korrigere tidsplanen efter 1. iteration.

Ulemper:

- Svært at begrænse funktionaliteten og beholde fokus på den enkelte iteration.

Vi vil kort beskrive to måder at iterere på.

Minimal-baseret

Ide: Opretholde et minimalt system, og forbedre de enkelte dele efterfølgende. Man laver en kørende implementation i 1. iteration, og udvikler så denne i de efterfølgende iterationer.

Fordele:

- Man får hurtigt et overblik over problemområder i implementeringen.
- Ressourcetunge dele af projektet afdækkes.

Ulemper:

- Grænseflader i systemet kan være svære at ændre, da der vil være kode på begge sider af fladen, der skal opdateres.

¹Se [vandfald www].

Stub-baseret

Ide: Opdele systemet i små separate dele. I første iteration laves de fleste dele som stubbe. Efterhånden som man itererer implementeres stubbene.

Fordele:

- Grænseflader kan nemt ændres, da stubsiden af fladen er trivielt at opdatere.

Ulemper:

- Man har ikke altid føling med, hvor lang tid implementeringen af hver stub tager.

Kilde(r) [Bruce Douglass 1999]

6.3.2 Komprimeringsmetoder

Dette afsnit beskriver forskellige metoder til komprimering af lyd, baseret på virkemåden af det menneskelige øre. Vi har valgt af koncentrere os om psykoakustiske modeller som bygger på destruktiv komprimering. Vi kunne derudover også have valgt at arbejde med lossless komprimering (som f.eks. benyttes i pakkeprogrammer som WinZip).

For at forstå disse principper er det nødvendigt først at kigge på teori bag ørets opbygning. Emperiske målinger foreslår at øret inddeles i 25 kritiske bånd, som hver især er følsomme overfor et fast frekvensområde [PCDA www]. Disse bånd ligger til grund for den opdeling, som er præsenteret i projektoplægget [Karstoft 2003].

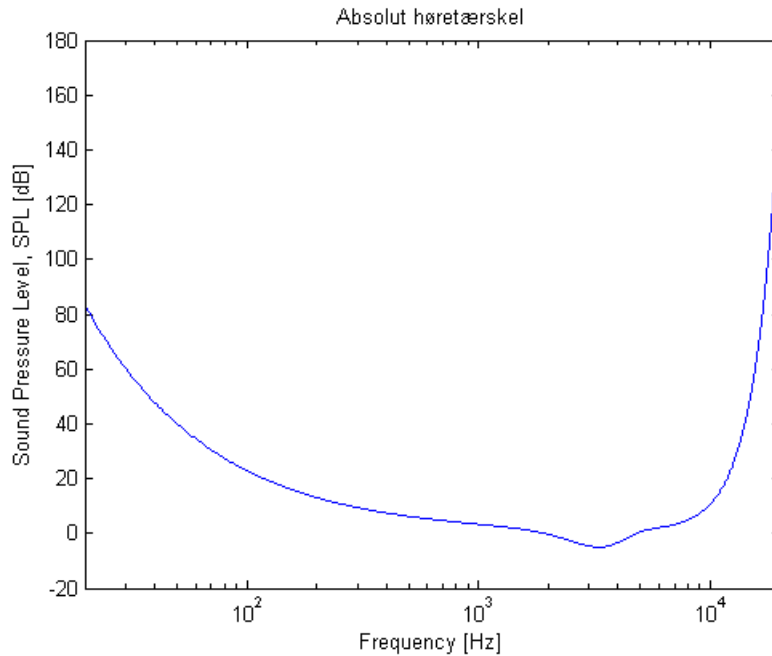
For at udføre analyser i frekvensområdet, er det nødvendigt at transformere signalet vha. diskret Fourier transformation (DFT). Hver DFT-pind i et bånd er fælles om én skalering og én allokeringsværdi, og det er derfor ikke muligt at ændre f.eks. allokeringen af en enkelt pind.

Absolut høretærskel:

Den absolutte høretærskel er baseret på undersøgelser af det menneskelige øres følsomhed overfor forskellige frekvensområder. Nedenstående figur er baseret på en approksimation lavet af det amerikanske National Institutes of Health (NIH) [PCDA www].

Figuren viser tydeligt, at ørets følsomhed for frekvenser over 10 kHz falder kraftigt (der skal et stort lydtryk til for at øret bliver påvirket). Det stemmer overens med den almindelige opfattelse af at den maksimale hørbare frekvens ligger mellem 10 kHz og 20 kHz afhængig af personen.

Vi kan drage nytte af denne viden ved at lægge høretærsklen ind over vores blokke af samples. Hvis signalindholdet i et bånd ligger fuldstændig under høretærsklen kan vi sætte allokeringen af dette bånd til 0. Med denne metode vil vi få en kraftig



Figur 2: Absolute høretærskel

komprimering, der dog er afhængig af signalet.

Fordele:

- Mulighed for store besparelser af data.
- Relativ simpel implementering.

Ulemper:

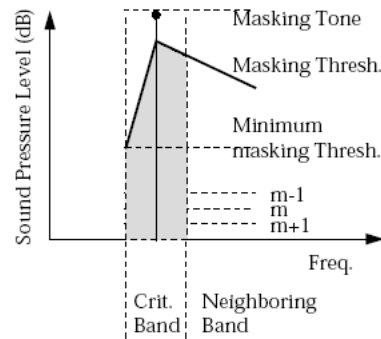
- Præcis tilpasning af høretærsklen kræver forholdsvis mange forsøg.

Simultan maskering:

Maskering opstår, når øret påvirkes af et kraftigt lydtryk i et snævert frekvensområde. F.eks. en enkelt kraftig tone. Denne tone opfanges af ét bestemt bånd i øret, og dette bånd er pga. den fysiske påvirkning mindre modtagelig overfor andre toner.

Ovenstående figur illustrerer dette fænomen. Signaler indeholdt i det skraverede område maskeres af den enkelte kraftige tone (masking tone). I teorien spredde maskeringen sig ud over båndet til de nærliggende bånd. Dette kan beskrives vha. en spredningsfunktion, som på figuren er kurven der afgrænser det skraverede område:

$$SF_{dB}(x) = 15.81 + 7.5(x + 0.474) - 17.5\sqrt{1 + (x + 0.474)^2}dB$$



Figur 3: Simultan maskering

Disse teorier kan benyttes til at komprimere vores signal. Hvis der i et enkelt bånd opstår maskering som følge af et kraftigt signal, kan vi introducere en vis mængde kvantiseringsstøj uden en hørbar forringelse af signalet. Højden mellem maskeringstonen og det maskerede areal ligger typisk på ca. 14,5 [dB][PCDA www]. Udfra denne værdi og spredningsfunktionen, kan vi udregne det maksimale niveau vores kvantiseringsstøj må antage, uden at der introduceres hørbare fejl.

Fordele:

- Mulighed for kraftigt øget kvantisering uden hørbar effekt.

Ulemper:

- Fuldstændig implementering kræver stor indsigt og er tidskrævende.
- Simple implementering kræver flere forsøg med tilpasning.

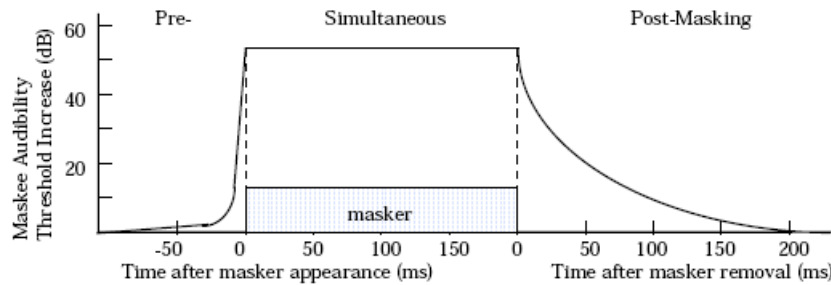
Tidsmæssig maskering:

Maskering optræder også som en tidsmæssig effekt. Som det ses på figur 4, findes der både præ- og postmaskering. Præmaskeringen er kortvarig og varer kun et par millisekunder. Postmaskeringen varer derimod mellem 50 og 200 ms, afhængig af maskeringstonens styrke og varighed. Kurven på figur 4 viser hvor mange dB høretærsklen hæves før, under og efter maskeringstonen optræder.

Denne form for maskering kan bruges til at analysere signalet på tværs af blokkene. Hvis en maskeringstone opstår i signalet, kan vi efterfølgende (og evt. i blokken før) enten skære ikke hørbare dele af signalet væk, eller introducere kraftigere kvantiseringsstøj.

Fordele:

- Mulighed for besparelser af hele blokke.



Figur 4: Tidsmæssig maskering

Ulemper:

- Kræver avanceret system for at kunne arbejde på tværs af blokke.
- Tidskrævende at implementere.

6.3.3 Testmetoder

Softwaretest:

Da det er et proof of concept projekt, er det ikke lagt vægt på test af koden. (unittest og integrationstest). Projektet skal bruges til at bevise det er muligt at komprimere lyd, med de kendskaber vi har fra signalbehandling. I et slutprodukt vil det kun være metoderne fra projektet der genbruges. Implementationen skal genskriveres fra bunden, og i et andet programmeringsprog.

Undervejs i projektet har vi testet koden, men dette har ikke været systematisk og er derfor ikke dokumenteret.

Lyttetest:

Der er lagt vægt på udførelsen af en lyttetest, for at validere vores komprimering. Lyttetestens primære formål er at afsløre afvigelser i den komprimerede lyd i forhold til den originale. Det sekundære formål er at bedømme kvaliteten. Det primæremål varetages af en kvantitativ test, og det sekundære af en kvalitetstest. Der er blevet lagt vægt på lyttetesten, da destruktiv komprimering nemt kan ødelægge et stykke musik. Det er derfor vigtigt at påvise, at det ikke har en negativ effekt på den hørbare kvalitet.

Kvantitativ test:

Testpersonen hører to musik stykker A og B samt lyden X, som vælges tilfældig blandt A og B. Test personen skal nu afgøre om lyden X, er A eller B.

I medicinalbranchen kendes denne metode som en dobbelt blindtest, da hverken testpersonen eller personen, der udfører testen, på forhånd er klar over, hvilket musik stykke X svarer til. I HIFI-kredse kaldes metoden for en ABX test.

Da der i hver test er 50% chance for at gætte rigtigt, er det vigtigt at gentage

testen med samme musikstykker og testperson. Sandsynligheden for gætte vil være uforholdsmæssig stor i de første tests, men den vil med tiden flade ud. Denne form for test viser udelukkende om test personerne er i stand til at høre forskel.

Fordele:

- Sikkerhed for at testpersonen og personen der udfører testen, ikke er påvirket af at kende det rigtige svar på forhånd.
- Testpersonens subjektive præferencer har ingen påvirkning på resultatet.

Ulemper:

- Testen siger intet om, hvorvidt en eventuel hørbar forskel er til stor eller lille gene, og dermed intet om kvalitetsforringelse.

Kvalitativ test:

Testpersonen hører de to musikstykker A og B, men ved ikke hvilken af dem der er den originale. Testpersonen kan give fra 1 til 5 point, hvor 5 point er lig med en ikke hørbar ændring. Der gives mindre end 5 point, hvis testpersonen kan høre en forskel.

Point betydning:

5. Ikke hørbar forskel
4. Hørbar forskel, men ikke generende
3. En smule generende
2. Generende
1. Meget generende

Fordele:

- Der gives en kvalitativ vurdering af en eventuel kvalitetsforringelse.

Ulemper:

- Testpersonens subjektive præferencer har indvirkning på resultatet, og det kan aldrig blive en objektiv vurdering.

6.3.4 Metodevalg

Udvikling

Vi valgt en stub-baseret iteration. I første iteration implementeres pakkerne WA-VE IO, MyP3 IO og Bitallokering/skalering.

I anden iteration implementeres en psykoakustisk model. Indholdet af pakkerne

er specificeret i afsnit 6.4.2

Valg af psykoakustiske metoder:

Vi valgte at benytte os af en kombination af 2 af de 3 beskrevne metoder. Vi kombinerede metoden for absolut høretærskel med en simpel udgave af simultan maskering. Maskeringen er gjort simplere ved kun at regne indenfor det samme bånd og ikke tage højde for den forskellige spredning i hver retning. Derved opnåes en simplere implementering, som kræver væsentlig mindre tid.

Vi fravalgte tidsmæssig maskering da vi vurderede at det ville stille for store krav til systemet.

Valg af testmetode:

På grund af tidspres har vi kun udført den kvantitative test. Afsnittet 6.7 Resultater indeholder en resonering over testresultaterne.

6.4 Analyse og design

Dette afsnit omhandler analyse- og designprocessen. Analyseprocessens formål er at identificere de essentielle karakteristika ved systemet, der kan bruges til de mulige løsningsforslag.

Da dette er et udviklingsprojekt består analysen af et målsætningsdokument. Målsætningsdokumentet beskriver den ønskede funktionalitet af systemet.

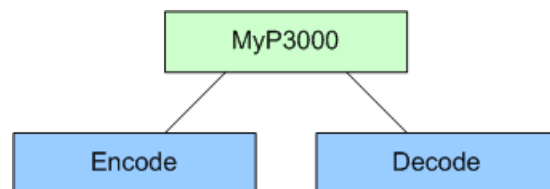
Efter de funktionelle krav er afdækket, udarbejdes en arkitektur for produktet. Denne arkitektur er dokumenteret i systemarkitektur dokumentet.

Afsnittet indeholder designet af vores løsning. Der er ikke opstillet alternative designs eller løsninger på konkrete problemer. Dette er undladt, da hensigten med projektet ikke er analyse og design, men derimed udforskning af psykoakustiske metoder med henblik på komprimering.

6.4.1 Funktionsanalyse

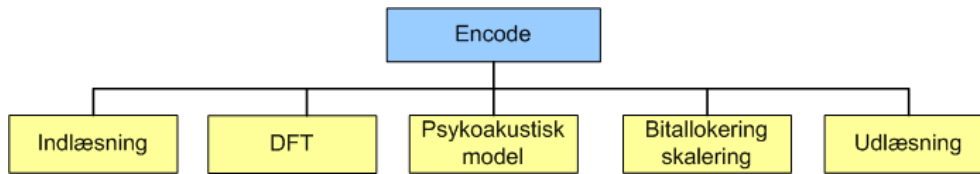
Formålet med funktionsanalysen er at bearbejde de funktionelle krav.

Det er oplagt at opdele systemets funktionalitet i 2 dele, encoder og decoder, som vist på figur 5.



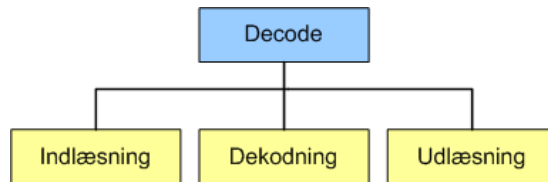
Figur 5: Opdeling af grundfunktionalitet

Funktionaliteten specificeret i MyP3 standarden [Karstoft 2003] giver en naturlig opdeling af encoderen.



Figur 6: Funktionalitet i encode

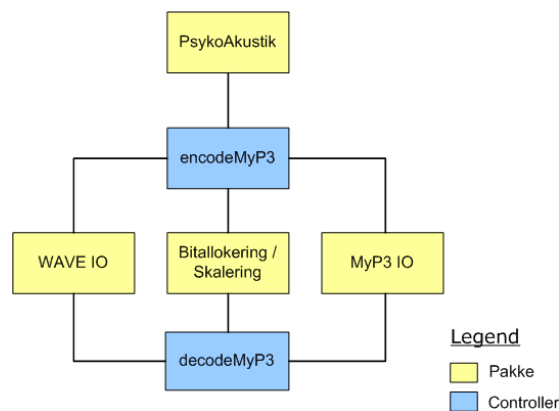
Funktionaliteten for decoderen danner grund for denne opdeling.



Figur 7: Funktionalitet i decode

6.4.2 Arkitektur design

Da Matlabs objektorienterede egenskaber er mangelfulde, har vi valgt at designet skal være funktionsbaseret. Vi kan umiddelbart oversætte ovenstående funktionelle krav til en række tilsvarende funktioner. Da funktionerne i decoderen vil være de inverse af dem i encoderen, har vi valgt denne pakkeopdeling.



Figur 8: Sammenkobling af encode og decode funktioner i pakker

WAVE IO

- **Ansvar:**
Håndtering af WAVE filer. Både indlæsning og udlæsning af WAVE fra Matlab.
- **Funktioner:**
readWAVE - Indlæsning af en WAVE fil. Returnerer samples, samplingsfrekvens samt bit opløsning.

writeWAVE - Udlæsning af samples til WAVE fil. Er hardcoded til en samplingsfrekvens på 44100 [Hz] og en bitopløsning på 16.

hanning - Indbygget Matlab funktion.

MyP3 IO

- **Ansvar:**
Håndtering af MyP3 filer. Både indlæsning og udlæsning af MyP3 fra Matlab.
- **Funktioner:**
myp3writer - Udlæsning af MyP3 fil. Skriver bearbejdet data til disken samt header, allokering og skalering.

myp3reader - Indlæser 1 frame af en MyP3 fil fra et givet offset. Offset er i bytes og leveres som parameter.

Bitallokering og skalering

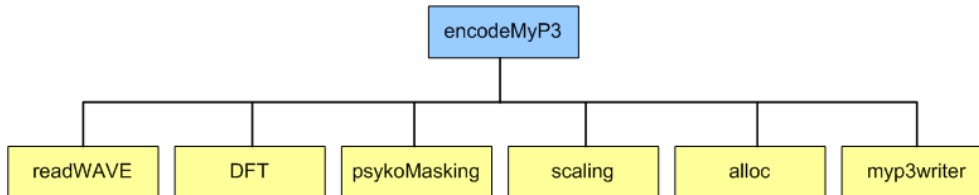
- **Ansvar:**
Håndtering af allokeringen og skalering af samples.
- **Funktioner:**
scaling - Foretager maksimal skalering indenfor hvert bånd.

alloc - Foretager kvantisering iht. den givne allokering.

Psykoakustik

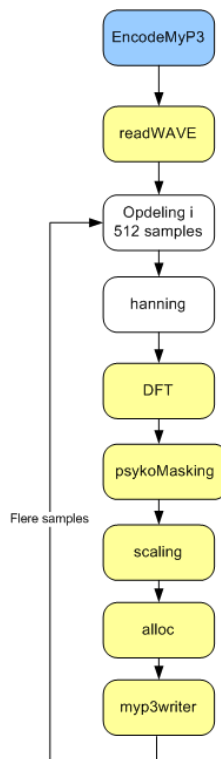
- **Ansvar:**
Håndtering af psykoakustikken der er grundlag for bitallokeringen.
- **Funktioner:**
psykoMasking - Foretager udvælgelsen af bitopløsligheden for hvert bånd.

Vi vil nøjes med at se på encoderen, da decoderen følger den analogt. Med udgangspunkt i funktionaliteten for encoderen kan vi opstille dette funktionsdiagram



Figur 9: Funktioner der indgår i encode

Ser vi på den sekventielle udførelse af programmet, kan vi beskrive den således.



Figur 10: Flowchart over encode

Hele inputfilen læses ind. Vi opsplitter filen fortløbende i blokke bestående af 512 samples med 50 % overlap. Hver blok færdigbehandles og gemmes, hvorpå den næste blok udvælges. Dette fortsættes indtil vi er igennem hele filen.

6.4.3 Dokumentation

Yderligere dokumentation ligger i Matlab filerne. Hver funktion er beskrevet med en header, som dækker over den hjælpetekst Matlab viser for funktionen, ved kald af `help`. Derudover vil vitale dele af koden være kommenteret.

6.5 Udviklingsværktøjer

Det har været vores hensigt at tilegne os ny viden om udviklingsværktøjer. Vi har derfor valgt at benytte adskillige værktøjer, som vi ikke har beskæftiget os med før.

Matlab:

Efter samråd med vejleder, valgte vi at benytte Matlab til programmeringen. Vi havde tidligere kun arbejdet kort med Matlab, så det var ret nyt for os.

Vores erfaring med Matlab er, at det er et rigtig godt valg til udvikling af prototyper. Der er indbygget en masse matematik rutiner, som man i andre sprog ville have brugt unødigt tid på at lave. Desuden skal man ikke holde styr på, hvordan man får oprettet og nedlagt sine data.

Vi kunne have valgt at benytte C++, som vi har væsentlig mere erfaring med. C++ har den fordel, at det bliver compileret til maskinkode, og derved kan udvikles langt mere effektivt end Matlab, der er et fortolkningsprog.

Hvis produktet var henvendt til slutbruger, havde der været større krav til eksekveringshastigheden, og vi ville have valgt at kode det i C++, eventuelt med de mest krævende dele i assembler.

LaTeX:

Alle vores dokumenter er skrevet i LaTeX. Grunden til vores valg af LaTeX var hovedsageligt for at prøve noget nyt. En anden grund er at det er meget stringent, så når man først har sat sig ind i hvordan det fungerer, skal man ikke tænke så meget på dokumentopsætning, hvilket medfører en mere koncentreret skriveproces. Vi har desuden benyttet TeXnicCenter, der er et IDE til LaTeX.

Alternativet til LaTeX var MS Word, som vi har benyttet på de øvrige semestre.

CVS:

Vi har valgt at benytte CVS til versionsstyring. CVS består af 2 dele, en server og en klient. Idéen bag det er, at flere personer kan arbejde på de samme dokumenter samtidig. Når man har lavet et dokument, som andre i projektgruppen skal se, eller skrive videre på, bruger man sin CVS klient til at checke det ind på serveren. Når et dokument er sat under CVS styring, kan de andre fra gruppen hente det ned. Hvis flere personer redigerer i det samtidig, for eksempel skriver et afsnit hver, finder CVS serveren selv ud af at sammensætte dokumentet, når det bliver checket ind. Desuden er der mulighed for at hente en tidligere version

af dokumentet.

6.6 Eksterne projektkrav

De gældende eksterne krav for projektet er beskrevet i MyP3 standarden, som er specificeret i projektoplægget side 2-4,6 [Karstoft 2003].

6.7 Resultater

Vi har udviklet et fuldt funktionelt produkt, der er afleveret til vejleder, sammen med dokumentationen, d. 22. december 2004.

Produktet består af en encoder, samt en decoder, der begge overholder version 1.3 af MyP3 standarden. Encodereren kan komprimere mellem 5-50% i forhold til originalen. Dette er afhængigt af indholdet i musikstykket. Dette svarer til 100% løsningen specificeret i afsnit 6.2. Desuden har vi gennemført en kvantitativ lyttetest, og dermed startet på 120% løsningen.

Udfra resultaterne af lyttetesten kan vi slutte, at i størstedelen af testene er der overvejende sandsynlighed for, at der kan høres en forskel. De rå testdata siger intet om den hørbare forskel består i en tydelig kvalitetsforringelse i form af f.eks. støj eller direkte fejl i musikstykket. Efter et interview med hver testperson viser det sig, at det var en lille smule periodisk støj der gjorde at der kunne høres forskel.

Med hensyn til tidsplanen, er det ikke lykkedes overholde den helt. Første milestone blev afsluttet præcis til tiden. Anden milestone blev afsluttet d. 3. december, og skød dermed tidsplanen med en uge. Vi nåede dermed ikke alle tre milestones inden eksamen, der lå i uge 50. Trejde milestone varede fra d. 6. til d. 15. december. Dette tog altså også længere tid end beregnet.

6.7.1 Diskussion af opnåede resultater

Vi har haft mulighed for at afprøve teori fra signalbehandling, og har endda tilegnet os en del viden inden for lydkomprimering. Også erfaringen fra DTM4, specielt mht. Matlab, har vi draget nytte af.

6.7.2 Projektets fortræffeligheder

Afsnittet omhandler nogle af de punkter i projektet vi er specielt stolte over. Vi har opdelt punkterne i tre grupper: Psykoakustik, Lyttetest og arbejdsmæssigt.

Psykoakustik:

Vi har med succes implementeret nogle grundlæggende psykoakustiske teorier, og dermed kraftigt reduceret størrelsen af MyP3 filen.

Lyttetest:

Vi har udført en lyttetest og lavet en simpel statistisk analyse af resultatet.

Arbejds-mæssigt:

Vi har fungeret godt internt i gruppen. Der har været en positiv og konstruktiv tone under hele semesteret. Det har bl.a. udmøntet sig i fast kageordning og motionsordning.

6.8 Forbedringer

Dette afsnit omhandler de umiddelbare funktionelle forbedringer vi har til produktet, samt forbedringer af projektarbejdet.

6.8.1 Funktionelt

Forbedringer af struktur:

Strukturen af produktet kan med fordel ensrettes på de to sider. Derved ville der blive en mere overskuelig opdeling.

Forbedringer af psykoakustiske metoder:

Her kan der arbejdes med at implementere den fulde udgave af simultan maskering samt den tidsmæssige maskering. Derved kan de nuværende fejl i lyden elimineres og kompressionen øges yderligere.

Forbedringer af testmetoder:

De udførte tests kan forbedres ved at udvide testen med flere testpersoner og evt. flere musikeksempler.

Derudover kan der med fordel udføres en kvalitativ test for at bestemme yderligere detaljer.

Forbedringer af hastighed:

De opnåede erfaringer angående psykoakustiske metoder m.m. kan benyttes til at udvikle programmet i C/C++. Eventuelle krævende dele kan implementeres i Assembler.

6.8.2 Projektarbejde

Der har til tider været en tendens til at fokus blev flyttet fra projektarbejdet til de øvrige kurser. Dette har bl.a. været årsag til en forskydning af tidsplanen.

7 Konklusion

Opnåede resultater:

Vi har udviklet et fuldt funktionelt produkt, og vi er godt tilfredse med komprimeringsgraden sammenholdt med tabet i kvalitet.

Tidsplan:

Vi har ikke overholdt tidsplanen til fulde, men er overordnet tilfreds med projektstyringen.

Metoder lært:

Vi har lært at arbejde med Matlab.

Vi har lært at håndtere praktiske signalbehandlingsproblemer.

Vi har lært selv at tilegne os viden om et ukendt emne. I tidligere projekter er teorien blevet præsenteret i undervisningen.

Vi har lært at benytte CVS til versionsstyring og organisering af filer.

Vi har lært at skrive dokumenter i LaTeX.

Kompromiser:

Der opstår små "ticks" i de komprimerede musikstykker - men det er ikke generende og fremkommer få steder. Lyttetesten afslørede disse fejl, og derfor er vi godt tilfredse med testen.

Samlet konklusion:

Vi er tilfredse med produktet MyP3000, og projektarbejdet der ligger bag. Vi har fungeret godt som gruppe, og udnyttet hinanden individuelle forcer optimalt.

8 Referencer

[Metro www]

<http://studiemetro.au.dk/html/index.html>

[Vandfald www]

I3PRJ3 Gruppe 1 - Projektrapport

IHA 2004

<http://www.benjagroup.dk/ccsystem/documents/projektrapport/projektrapport.pdf>

[Bruce Douglass 1999]

Doing Hard Time:

Developing Real-Time Systems using UML, Objects, Frameworks, and Patterns

Reading, MA: Addison-Wesley, 1999

[PCDA www]

Perceptual Coding of Digital Audio

Painter & Spanias

Arizona State University 85287-7206

<http://www.eas.asu.edu/~spanias/audiopaper1.pdf>

[Karstoft 2003]

MyP3-optager/-afspiller

Audio-kompression og -dekompression

Henrik Karstoft

IHA 2003

[Torben Poulsen 1998]

Lydopfattelse, 2. udgave

Torben Poulsen

DTU 1998