

Ingeniørhøjskolen i Århus
IKT
Dalgas Avenue 2
8000 Århus C

20. december 2004

4. Semesterprojekt
Målsætning

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	15.09.2004	BS	Opsætning af dokumentet i LaTeX
0.2	17.09.2004	BS	Indledning og generel beskrivelse
1.0	27.09.2004	TB	Dokument færdiggjort

2 Godkendelsesformular

Forfatter(e):	Benjamin Sørensen(BS), Tomas Berg(TB), Nikki Ashton(NA), Jonas Livbjerg(JL)
Projektnummer:	4. semesterprojekt
Antal sider:	15
Godkendes af:	Tomas Berg

Ved underskrivelse af dette dokument accepteres det som værende målsætningen for udviklingen af det ønskede system.

Dato og underskrift:

Tomas Berg

Indhold

1	Versionshistorie	1
2	Godkendelsesformular	1
3	Indledning	4
3.1	Formål	4
3.2	Referencer	4
3.3	Læsevejledning	4
3.4	Ordliste	4
4	System oversigt	5
4.1	Systemets funktioner	5
4.1.1	Encoder	5
4.1.2	Decoder	5
4.1.3	Komprimering	6
4.2	Systemets begrænsninger	6
4.3	Systemets fremtid	6
5	Målsætning	7
5.1	Delmål 1 - Indlæsning og decoding af WAVE	7
5.1.1	Analysering	7
5.1.2	Antagelser	7
5.1.3	Succeskriterier	7
5.2	Delmål 2 - Hanningvindue/multiplikation	8
5.2.1	Analysering	8
5.2.2	Antagelser	8
5.2.3	Succeskriterier	8
5.3	Delmål 3 - DFT	8
5.3.1	Analysering	8
5.3.2	Antagelser	8
5.3.3	Succeskriterier	9
5.4	Delmål 4 - Formatering og lagring af MyP3	9
5.4.1	Analysering	9
5.4.2	Antagelser	9
5.4.3	Succeskriterier	9
5.5	Delmål 5 - Indlæsning og decoding af MyP3	9
5.5.1	Analysering	9
5.5.2	Antagelser	9
5.5.3	Succeskriterier	9
5.6	Delmål 6 - Formatering og lagring af WAVE	10
5.6.1	Analysering	10

5.6.2	Antagelser	10
5.6.3	Succeskriterier	10
5.7	Delmål 7 - Psykoakustisk model	10
5.7.1	Analyse	10
5.7.2	Antagelser	10
5.7.3	Succeskriterier	10
5.8	Delmål 8 - Bitallokering/skalering	10
5.8.1	Analyse	10
5.8.2	Antagelser	10
5.8.3	Succeskriterier	10
5.9	Delmål 9 - Lyttetest	11
5.9.1	Analyse	11
5.9.2	Antagelser	11
5.9.3	Succeskriterier	11
6	Udførelse	12
6.1	Milestone 1	12
6.2	Milestone 2	12
6.3	Milestone 3	13
7	Afslutningskriterier	14
7.1	Procentopdeling	14
8	Oversigter	15
8.1	Bilagsoversigt	15
8.2	Figuroversigt	15

3 Indledning

3.1 Formål

I forbindelse med semesterprojektet i kurset I4PRJ4 skal der udvikles en prototype for encoding og decoding af lyd.

Der lægges vægt på udforskningen af lydkomprimering. Projektet er et udviklingsprojekt, som udføres for at grundlægge firmaets vurdering af rentabiliteten. Dette gøres ved udvikling og lancering af en ny lydkomprimeringsstandard.

3.2 Referencer

Projektoplægget: myp3 ver. 1.3.pdf.

Forklaring af WAVE formatet: WaveFormat.htm

3.3 Læsevejledning

Dette dokument indeholder en generel beskrivelse, målsætning og udførelse af projektet. Derudover indeholder det en afklaring af, hvad der anses som afslutningskriterier for projektet. Disse afdækker hvad der forventes implementeret ved projektets afslutning.

3.4 Ordliste

MyP3: Lydkomprimeringsstandard.

WAVE: Specificering af lagring af multimedia filer.

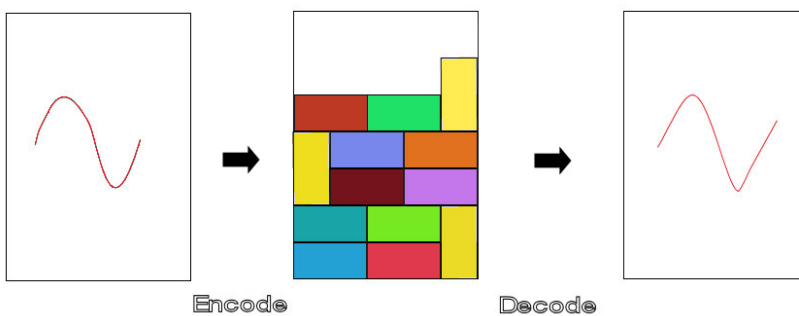
Encoder: Software der, i vores tilfælde, håndterer komprimering af filer fra WAVE til MyP3.

Decoder: Software der, i vores tilfælde, håndterer tilbagekonvertering fra MyP3 til WAVE.

4 System oversigt

Dette afsnit giver et overordnet billede af systemet, samt de omgivelser systemet skal fungere i.

Systemets navn er MyP3000. MyP3000 opfylder to formål. Det primære mål er komprimering af wavefiler ved brug af MyP3 standarden. Det sekundære mål er at dekomprimere MyP3 filer tilbage til WAVE filer.



Figur 1: Sammenhæng mellem primær og sekundær mål.

Det primære mål varetages af en encoder med indbygget komprimering. Det sekundære mål varetages af en decoder. Vi skal lave både encoder og decoder.

4.1 Systemets funktioner

Herunder beskrives henholdsvis encoder, decoder og komprimering.

4.1.1 Encoder

Encoderen skal indlæse lydsignalet fra en WAVE fil og opdele det i blokke. Disse blokke bliver behandlet med 50% overlap. Der ganges et Hanningvindue på og efterfølgende udføres en DFT, så vi transformerer signalet over i frekvensområdet. Herefter foretages den egentlige komprimering af data, dette gøres vha. en psykoakustisk model, se under "Komprimering".

4.1.2 Decoder

Decoderen skal udføre en invers DFT på blokkene for at transformere signalet over i tidsområdet. Til sidst samles blokkene, og de gemmes i WAVE lydformatet.

4.1.3 Komprimering

Komprimeringen finder sted efter at lydsignalet er transformeret over i frekvensområdet.

Komprimeringen er destruktiv. Det betyder, at vi smider dele af dataene i lydsignalet væk for at gøre komprimeringen bedre. Dette kan vi tillade os, da det menneskelige øres opbygning gør, at det ikke er i stand til at opfatte visse dele i et lydsignal. Der benyttes 2 principper:

1. Alt efter det menneskelige øres følsomhed i forskellige frekvensområder, kan vi allokere et større antal bits til dataene i et følsomt frekvensområde, og et mindre antal bits i et mindre følsomt frekvensområde.
2. Ørets opbygning gør, at hvis det præsenteres for en kraftig puls i et snævert område, vil det være mindre følsomt overfor fejl i nærliggende områder. Derved kan vi udfra en frekvensanalyse afgøre, hvor stor en kvantiseringsfejl vi kan introducere, og stadig have et transparent signal.

Ovenstående principper udgør vores psykoakustiske model. Derudover kan vi også skalere de enkelte frekvenser, og derved opnå bedst mulig udnyttelse af den valgte præcision.

4.2 Systemets begrænsninger

Vi vil kun arbejde med mono signaler med en samplingsfrekvens på 44.1 kHz. Decoderen skal lave MyP3 filer om til WAVE filer. Der skal altså ikke implementeres en afspilning af hverken MyP3 eller WAVE filer.

4.3 Systemets fremtid

Da MyP3000 udvikles som en prototype, lever produktet ikke videre efter endt udvikling. Vitale dele af systemet vil kunne indgå i fremtidige færdige systemer.

5 Målsætning

Dette afsnit indeholder en nærmere beskrivelse af de mål vi har sat os for dette projekt. De er opdelt i 9 delmål fordelt på 3 grupper, der hver leder hen til en milestone. Projektet består som beskrevet i forrige afsnit af flere opgaver, og vi vil inddele dem i følgende delmål.

5.1 Delmål 1 - Indlæsning og decoding af WAVE

5.1.1 Analysering

Første delmål består af indlæsning af en WAVE fil. Vi skal have lavet en programstump, der åbner en WAVE fil og gør den klar til læsning.

Det drejer sig om at få analyseret WAVE formatet, og få splittet filen op. En WAVE fil består af forskellige blokke, som vi kan udtrække vores samples fra.

5.1.2 Antagelser

Som input lægger vi os fast på WAVE formatet. Programmet skal ikke kunne læse andre formater i dette delmål. Vi antager at inputfilen ligger i WAVE PCM standard formatet.

5.1.3 Succeskriterier

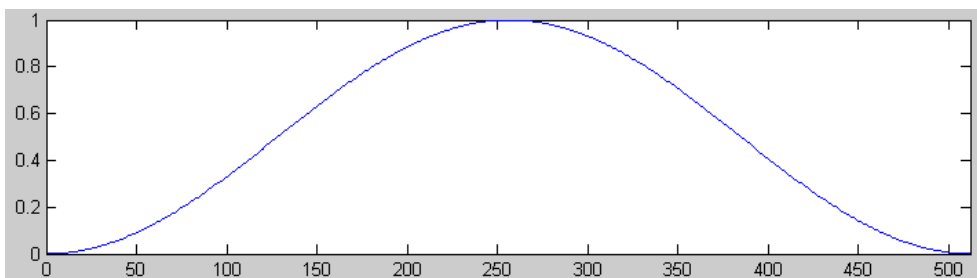
Programstumpen skal producere rene samples klar til behandling.

5.2 Delmål 2 - Hanningvindue/multiplikation

5.2.1 Analysering

Andet delmål indeholder databehandling. Vi skal have læst blokke af 512 samples ind med 50% overlap. Dvs. at for hver blok vi behandler, er der data der går igen. Herefter skal der ganges et Hanningvindue på hver blok.

Et Hanningvindue dæmper starten og slutningen af signalet. Vinduet genereres ud fra en cosinus og kan skiteres således.



Figur 2: Hanning vindue længde 512.

5.2.2 Antagelser

Der skal være adgang til de rene samples.

5.2.3 Succeskriterier

Signalet er blevet inddelt i blokke og multipliceret med en Hanningvindue.

5.3 Delmål 3 - DFT

5.3.1 Analysering

Tredje delmål skal udføre en DFT (Discrete Fourier Transform) på signalet og derved overføre det fra tidsområdet til frekvensområdet. Denne disciplin giver os mulighed for senere at gå ind og komprimere signalet i forskellige frekvensbånd.

5.3.2 Antagelser

Vi vil benytte en skaleret DFT vha. følgende formel:

$$X(k) = \frac{1}{512} \sum_{n=0}^{511} x[n] e^{-j(\frac{2\pi}{512}k)n}$$

5.3.3 Succeskriterier

Signalet skal være konverteret til frekvensområdet så det er muligt at arbejde med individuelle frekvensbånd.

5.4 Delmål 4 - Formatering og lagring af MyP3

5.4.1 Analysering

De konverterede sampleblokke skal lægges i MyP3 formatet, som er defineret nedenfor. Derefter skal blokkene skrives til en fil til senere brug.

5.4.2 Antagelser

Vi lægger os fast på følgende format:

Header 12 bits	Allokationsbit ($25 \cdot 4 = 100$ bits)	Skaleringsfaktorer ($25 \cdot 4 = 100$ bits)	Kvantiserede frekvensdata
-------------------	--	--	------------------------------

5.4.3 Succeskriterier

En fil er blevet skrevet i det fastlagte MyP3 format.

5.5 Delmål 5 - Indlæsning og decoding af MyP3

5.5.1 Analysering

Det femte delmål består af indlæsning af en fil i MyP3 formatet. Denne fil skal pakkes ud af MyP3 formatet ved at lave en invers skalering af frekvensbåndene. Derefter skal de enkelte samples lægges sammen for at kompensere for vores 50 % overlap. Til sidst skal der laves en invers DFT på signalet for at konvertere det tilbage til tidsområdet.

5.5.2 Antagelser

Det antages at filen er i MyP3 formatet og dermed er konstrueret vha. 50 % overlap og et Hanningvindue.

5.5.3 Succeskriterier

MyP3-filen er blevet konverteret og splittet op i rene samples.

5.6 Delmål 6 - Formatering og lagring af WAVE

5.6.1 Analysering

Her skal sammensættes blokke i WAVE PCM formatet udfra rene samples. Derefter skal signalet gemmes i en fil.

5.6.2 Antagelser

Der skal være adgang til de rene samples.

5.6.3 Succeskriterier

Signalet er blevet formateret og gemt i en WAVE fil.

5.7 Delmål 7 - Psykoakustisk model

5.7.1 Analyse

Undersøgelse af forskellige modeller og teorier omkring komprimering af forskellige bånd. En simpel implementering af en psykoakustisk model.

5.7.2 Antagelser

Det vil være muligt at fjerne information fra WAVE filen uden hørbar tab af kvalitet.

5.7.3 Succeskriterier

Opnå komprimering i forhold til original filen.

5.8 Delmål 8 - Bitallokering/skalering

5.8.1 Analyse

Foretage en bitallokering og skalering udfra den psykoakustiske model. Den optimale bitallokering vil være 16 bit i alle bånd. Når denne skrues ned, opnås komprimering.

5.8.2 Antagelser

Det er muligt at kvantisere WAVE filen, uden tab af kvalitet.

5.8.3 Succeskriterier

Opnå den forventede komprimering, som den psykoakustiske model skulle give.

5.9 Delmål 9 - Lyttetest

5.9.1 Analyse

Udføre en lyttetest for at undersøge tabet af kvalitet.

5.9.2 Antagelser

Ved en dobbelt blindtest vil eventuelle afvigelser fra originalen komme frem.

5.9.3 Succeskriterier

Testen udføres.

6 Udførelse

Udviklingen deles op i milestones, der baseres på delmålene beskrevet i det foregående afsnit. For hver milestone angives hvilke delmål der indgår, samt hvor mange dage det vurderes at vare. Efter gennemførelsen af en milestone reflekteres over problemer og resultater.

6.1 Milestone 1

Første milestone indeholder disse delproblemer:

- Indlæsning og decoding af WAVE
Tidshorisont: 10 dage.
Ansvarlig: NA, JL
- Hanningvindue/multiplikation
Tidshorisont: 10 dage.
Ansvarlig: BS, TS
- DFT
Tidshorisont: 10 dage.
Ansvarlig: BS, TS
- Formatering og lagring af MyP3
Tidshorisont: 10 dage.
Ansvarlig: NA, JL
- Indlæsning og decoding af MyP3
Tidshorisont: 5 dage.
Ansvarlig: NA, JL
- Formatering og lagring af WAVE
Tidshorisont: 5 dage.
Ansvarlig: NA, JL

6.2 Milestone 2

Anden milestone indeholder disse delproblemer:

- Psykoakustisk model
Tidshorisont: 10 dage.
Ansvarlig: BS, TS
- Bitallokering/skalering
Tidshorisont: 10 dage.
Ansvarlig: NA, JL

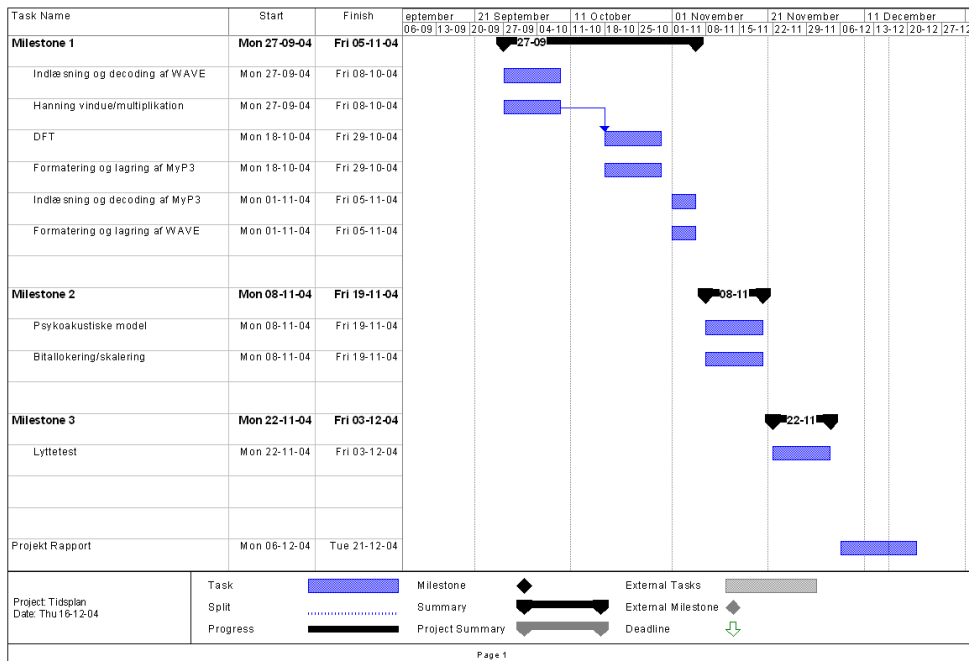
6.3 Milestone 3

Tredje milestone indeholder dette delproblem:

- Lyttetest
Tidshorisont: 10 dage
Ansvarlig: NA

Milestone 1 og 2 indeholder sideløbende processer. Disse udføres parallelt.

For at give et overblik over udførelsen, samles denne i en tidplan.



Figur 3: Tidsplan

7 Afslutningskriterier

I dette afsnit vil vi definere hvad vi forstår ved 80 %, 100 % og 120 % løsninger. Vi planlægger at nå 120 % løsningen, men vi vil være tilfredse med at nå en 100 % løsning. 80 % løsningen beskriver projektet som værende godt på vej, men ikke fuldendt. Denne løsning vil vi ikke acceptere som det endelige produkt. Procentopdelingen kan ses om afslutningskriterier for projektet.

7.1 Procentopdeling

- 80 % løsning
Milestone 1 anses som en 80 % løsning. Med denne løsning kan vi indlæse en WAVE fil, konvertere den til MyP3 og lagre denne. Der vil ikke være nogen komprimering.
- 100 % løsning
Milestone 2 anses som en 100 % løsning. Med denne løsning kan vi desuden foretage en vis komprimering af filen. Dette vil vi være tilfreds med, såfremt der ikke er nogen hørbar forskel på WAVE filen og MyP3 filen.
- 120 % løsning
Milestone 3 anses som en 120% løsning. Denne milestone består af en omfattende lyttetest, der skal dokumentere brugbarheden af produktet.

8 Oversigter

8.1 Bilagsoversigt

Bilag 1: Tidsplan

8.2 Figuroversigt

Figurer

1	Sammenhæng mellem primær og sekundær mål.	5
2	Hanning vindue længde 512.	8
3	Tidsplan	13