

Ingeniørhøjskolen i Århus
IKT
Dalgas Avenue 2
8000 Århus C

20. december 2004

4. Semesterprojekt
Testdokumentation

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 | 13.12.2004 | BS | Opsætning af dokumentet i LaTeX |
| 0.2 | 13.2.2004 | NA | Første udkast |
| 0.4 | 20.12.2004 | NA | Diverse rettelser og tilføjelser |
| 1.0 | 20.12.2004 | TB | Afsluttende rettelser |

2 Godkendelsesformular

| | |
|----------------|--|
| Forfatter(e): | Benjamin Sørensen(BS), Tomas Berg(TB), Nikki Ashton(NA), Jonas Livbjerg(JL) |
| Projektnummer: | 4. semesterprojekt |
| Antal sider: | 9 |
| Godkendes af: | Nikki Ashton |

Dato og underskrift:

Nikki Ashton

| | |
|----------------|---|
| <i>INDHOLD</i> | 2 |
|----------------|---|

Indhold

| | |
|--|----------|
| 1 Versionshistorie | 1 |
| 2 Godkendelsesformular | 1 |
| 3 Introduktion | 3 |
| 3.1 Formål og omfang | 3 |
| 3.2 Referencer | 3 |
| 3.3 Definitioner og forkortelser | 3 |
| 3.4 Læsevejledning | 3 |
| 4 Organisering af test | 4 |
| 4.1 Softwaretest | 4 |
| 4.2 Lyttetest | 4 |
| 4.2.1 Automatisering | 4 |
| 4.2.2 Analyse af resultatet | 4 |
| 4.2.3 Fremgangsmåde ved test | 5 |
| 5 Konklusion | 7 |
| 6 Figuroversigt | 8 |
| 7 Bilag | 9 |
| 7.1 Bilag 1 | 9 |

3 Introduktion

3.1 Formål og omfang

Dette dokument beskriver testforløb og resultater. Tests er dokumenteret for at andre end projektdeltagerne kan danne sig et overblik over hvilken testform, og hvilke tests der er udført.

3.2 Referencer

WinABX: <http://www.kikeg.arrakis.es/winabx/>

3.3 Definitioner og forkortelser

Lyttetest: En subjektiv test af lyden/musikken.

Freeware: Et stykke software der er gratis at bruge.

Unittest: Test af en enkelt funktion.

Integrationstest: Test af funktioner i sammenhæng.

3.4 Læsevejledning

Organisering af test indeholder en beskrivelse af hvad og hvordan vi har testet.

Konklusion indeholder en analyse og konklusion på baggrund af testresultaterne.

4 Organisering af test

Softwaretesten er af begrænset omfang, da dette er et udviklingsprojekt, og softwaren derfor ikke skal ud til en almindelig bruger. Derimod vil der blive lagt vægt på lyttetesten, da det er vigtigt at musik, der har været igennem MyP3 formatet, ikke er væsentlig forringet.

Lyttetestens formål er at evaluere om man kan høre forskel på musik før og efter, at det har været igennem MyP3 encoderen/decoderen.

4.1 Softwaretest

Da det er et udviklingsprojekt, er der ikke lagt vægt på test af softwaren (unittest og integrationstest). Dette er fordi softwaren kun skal bruges til at bevise at det er muligt at komprimere lyd, ved bl.a. at behandle det i frekvensområdet. I en videreudvikling af konceptet vil det kun være vitale dele der genbruges fra dette udviklingsprojekt. Disse dele vil skulle genskrives fra bunden, da det sandsynligvis ikke vil være i Matlab produktet programmeres, men i et andet programmeringsprog.

Vi har løbende lavet unittest og integrationstest, men det er ikke dokumenteret yderligere.

4.2 Lyttetest

Metoden vi bruger til at teste kaldes i 'HIFI kredse' for en ABX test.

En testperson kan høre de to musikstykker A og B samt stykket X, som er en tilfældig af de to. Testpersonen skal bestemme om lyden X er den samme som A eller B. Denne metode svarer til hvad man i medicinalbranchen kalder en dobbelt blindtest, da hverken testpersonen eller personen der udfører testen, på forhånd er klar over hvilket musikstykke X svarer til.

4.2.1 Automatisering

Til automatisering har vi fundet et program der opfylder vores krav til testen. Programmet hedder WinABX og er freeware.

4.2.2 Analyse af resultatet

For at udregne sandsynligheden for at testresultatet er fremkommet ved gæt-værk, bruges en kumulativ binomial fordeling. Dvs. sandsynligheden for at der i N forsøg er svaret korrekt mindst Y gange ved en tilfældighed.

Summen af $P(x)$ hvor x løber fra Y til N :

$$P_k = \sum_{x=Y}^N P(x)$$

$$P(x) = \frac{N!}{x!(N-x)!} \cdot 0.5^x \cdot 0.5^{N-x}$$

Her er det naturligt, at jo større N er, jo mere præcis vil analysen være. For ikke at overbebyrde testpersonen, sætter vi N antal tests til 16. Dette er ikke et tal vi har regnet os frem til, men et vi har fastsat ved en afvejning af hvor mange tests vi vil udsætte en person for, samt hvor mange tests der er nødvendige for at få et nogenlunde pålideligt resultat.

Hvis $P_k < 5\%$ vil vi gå ud fra at resultatet ikke er fremkommet ved gætværk, og at testpersonen kunne høre en forskel.

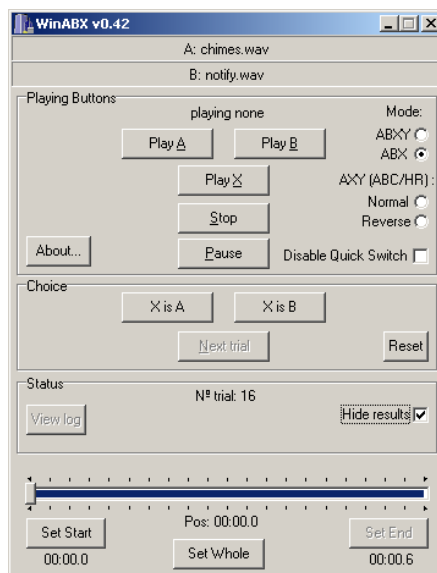
4.2.3 Fremgangsmåde ved test

Der testes på følgende musikstykker:

- Klassisk 1: High Fidelity, Reference CD nr. 65, DeBussy - Clair de Lune, offset 90sek, 13sek.
- Klassisk 2: High Fidelity, Reference CD nr. 65, DeBussy - Clair de Lune, offset 165sek, 13sek.
- Klassisk 3: Canon 'a 3 Violinis con suo Basso' - Johann Pachelbel, 10sek.
- Enkelt instrument: TestCD udleveret til PRJ - sqam-castanets-left-1st, 7sek.
- Rock: Dire Straits - On Every Street: My Parties, 10sek.
- Elektronisk: Astral Projection - In the mix: Burning Up, 10sek.

Hver testperson følger disse punkter for at gennemføre en test:

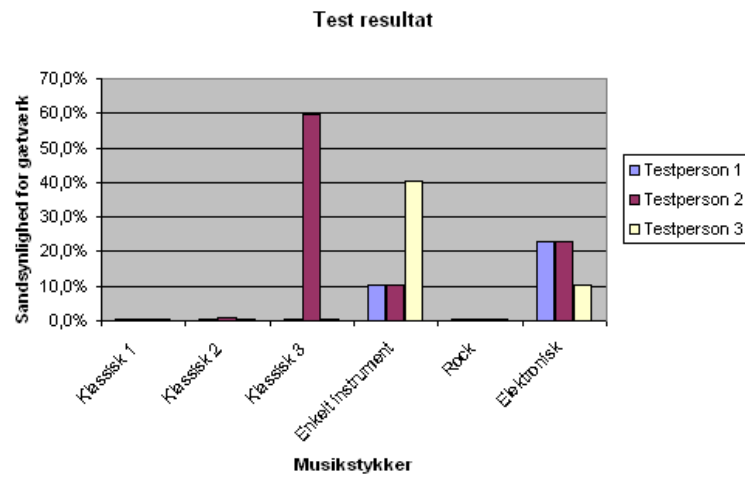
1. WinABX startes og de to WAVE filer læses ind. Ved 'Mode' vælges 'ABX'.
Se evt. figur 1
2. Testpersonen har mulighed for at høre lyd A, B og X så mange gange det er nødvendigt ved at trykke på henholdsvis 'Play A', 'Play B' og 'Play X'.
3. Testpersonen trykker på enten 'X is A' eller 'X is B' knappen, alt efter hvad testpersonen mener X lyder mest som.
4. Testpersonen trykker på 'Next Trial' knappen.
5. Testpersonen gentager testen fra punkt 2, indtil N.trial når 17.
6. Der gentages fra punkt 1, indtil alle musikstykkerne er testet.



Figur 1: Opsætning af WinABX.

5 Konklusion

På baggrund af resultaterne i bilag 1, ses det at i størstedelen af testene er der overvejende sandsynlighed for, at der kan høres en forskel. Se evt. figur 2. De rå testdata siger intet om hvorvidt 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.



Figur 2: Graf over resultatet.

6 Figuroversigt

Figurer

| | | |
|---|-------------------------------|---|
| 1 | Opsætning af WinABX. | 6 |
| 2 | Graf over resultatet. | 7 |

7 Bilag

7.1 Bilag 1

- Klassisk: High Fidelity, Reference CD nr. 65, DeBussy - Clair de Lune, of-fset 90sek, 13sek.
Nikki: 15/16 P<0.1%
Benjamin: 15/16 P<0.1%
Jonas: 16/16 P<0.1%
- Klassisk: High Fidelity, Reference CD nr. 65, DeBussy - Clair de Lune, of-fset 165sek, 13sek.
Nikki: 15/16 P<0.1%
Benjamin: 14/16 P=0.2%
Jonas: 16/16 P<0.1%
- Klassisk: Canon 'a 3 Violinis con suo Basso' - Johann Pachelbel, 10sek.
Nikki: 16/16 P<0.1%
Benjamin: 8/16 P=59.8%
Jonas: 16/16 P<0.1%
- Enkelt instrument: TestCD udleveret til PRJ - sqam-castanets-left-1st, 7sek.
Nikki: 11/16 P=10.5%
Benjamin: 11/16 P=10.5%
Jonas: 9/16 P=40.2%
- Rock: Dire Straits - On Every Street: My Parties, 10sek. Nikki: 16/16 P<0.1%
Benjamin: 15/16 P<0.1%
Jonas: 16/16 P<0.1%
- Elektronisk: Astral Projection - In the mix: Burning Up, 10sek.
Nikki: 10/16 P=22.7%
Benjamin: 10/16 P=22.7%
Jonas: 11/16 P=10.5%