

Digital Kino og HDTV

Øystein Hansen

Master i elektronikk
Oppgaven levert: Juni 2007
Hovedveileder: Andrew Perkis, IET

Oppgavetekst

Film blir i dag mer og mer tatt opp og editert i digitalt format for så og overføres til 35mm film for visning. Framtidens kino vil digitalisere alle ledd i verdikjeden og distribuere en digital versjon av filmen i en såkalt "Digital Cinema Package - DCP". Denne består av komprimerte stillbilder og tilhørende metadata. Kompresjonsformatet som er valgt er JPEG2000 Part 1. En DCP vil typisk være en fil på ca. 5-800 GByte.

Filmoriginalen- representert ved DCPen -vil det være interessant å senere kunne distribuere til kringkastere for visning på TV. Kringkastere bruker et annet format og vil derfor måtte trekke ut en annen oppløsning på bildet enn det som brukes i Digital kino. Oppgaven går ut på å sette seg inn i JPEG2000 og studere hvordan et HDTV signaler kan trekkes ut fra en JPEG2000 kodet 2K original. I tillegg skal det gjøres klart for og muligens gjennomføres subjektive tester på JPEG2000 kodede sekvenser.

Oppgaven gitt: 15. januar 2007
Hovedveileder: Andrew Perkis, IET

Digital kino og HDTV

Øystein Hansen

19. juni 2007



Norges teknisk-naturvitenskapelige universitet
Fakultet for informasjonsteknologi, matematikk og elektroteknikk
Institutt for elektronikk og telekommunikasjon

Oppgavetekst

For vårsemesteret 2007 var jeg tildelt Master oppgaven “Digital kino og HDTV”. Denne Masteroppgaven er en videreføring av prosjektoppgaven, “Digital Kino 2 - HDTV/JPEG2000” [3], gjennomført høsten 2006. Oppgaven var satt opp av Andrew Perkis og hadde følgende oppgavetekst. [4]

“Film blir i dag mer og mer tatt opp og editert i digitalt format for så og overføres til 35mm film for visning. Framtidens kino vil digitalisere alle ledd i verdikjeden og distribuere en digital versjon av filmen i en såkalt “Digital Cinema Package - DCP”. Denne består av komprimerte stillbilder og tilhørende metadata. Kompresjonsformatet som er valgt er JPEG2000 Part 1. En DCP vil typisk være en fil på ca. 5-800 GByte.

Filmoriginalen- representert ved DCPen -vil det være interessant å senere kunne distribuere til kringkastere for visning på TV. Kringkastere bruker et annet format og vil derfor måtte trekke ut en annen oppløsning på bildet enn det som brukes i Digital kino. Oppgaven går ut på å sette seg inn i JPEG2000 og studere hvordan et HDTV signaler kan trekkes ut fra en JPEG2000 kodet 2K original. I tillegg skal det gjøres klart for og muligens gjennomføres subjektive tester på JPEG2000 kodede sekvenser.“

I samtale med faglærer Andrew Perkis og veileder Marlon Th M Nielsen, kom vi frem til at det i Master oppgaven skal hovedsaklig legges vekt på gjennomføring av en subjektiv test. Den subjektive testen skal utføres på HDTV utstyr og sammenligne H.264 og JPEG2000.

Sammendrag

I Masteroppgaven ser vi på den subjektive kvaliteten til JPEG2000 og H.264. Vi ser på den subjektive kvaliteten i forbindelse med konverteringen fra en digital kino pakke (DCP) til video klar for kringkasting på HDTV. En subjektiv undersøkelse har blitt gjennomført, som prøver å gi svar på hva riktig kvalitet for HDTV er og hvordan JPEG2000 og H.264 står i forhold til hverandre med bittrate og kvalitet. I Masteroppgaven har vi sett videre på konverteringsprosessen fra en DCP til en H.264 kodet video for kringkasting på HDTV, og kommet frem til en mer automatisert metode med flere utbyggingsmuligheter. Fra den subjektive undersøkelsen kan man trekke ut følgende.

- For middels og lave bittrater er H.264 overlegen i forhold til JPEG2000 for videoenkoding, med tanke på kvalitet.
- H.264 kodet video for HDTV er av god kvalitet ved bittrater på 2Mbps.

Forord

Denne rapporten oppsummerer og forklarer hva som ble gjort under Masteroppgaven. Den tar også med deler av prosjekt arbeidet, dette for å hjelpe rapporten til å bli mer fullstendig, og gi en så god beskrivelse av emnet som mulig. Under arbeidet med Masteroppgaven har jeg benyttet KafeMedia (B133) på Gløshaugen som medialabb. Her har jeg fått god hjelp av Krzysztof Orleanski til å tilpasse KafeMedia slik at den ble klar til den subjektiv undersøkelsen. Det faglige har hovedsaklig Marlon Th M Nielsen bistått meg med. Faglærer Andrew Perkis har dette semesteret hatt, mer enn for prosjekt oppgaven, den overordnede styringen, og dermed ikke hvert med på alle smådeler prosjektet består av. Per Børler har via epost hvert svært behjelpelig med informasjon om subjektiv testing og HDTV innstillinger, som har gitt gode retningslinjer for gjennomføringen av den subjektive undersøkelsen.

Dessverre ble det en del kluss med tilgjengelighet på en HDTV under Masteroppgaven. Dette førte til at testing mot HDTV og den subjektive undersøkelsen kom svært sent igang. Det oppstod da også en del døtid i midten av Masteroppgave-tiden som ikke ble benyttet 'i linje' med den subjektive undersøkelsen. Noe av det som kom ut av denne døtiden kan du finne under kapittel 8.

Halvveis gjennom den subjektive undersøkelsen ble løsningen for 'bleke farger i DCP materialet funnet. Siden undersøkelsen var godt igang ble denne løsningen ikke innført i den subjektive undersøkelsen. Denne løsningen blir isteden beskrevet separat i kapittel 7.1.

Innhold

Oppgavetekst	I
Sammendrag	III
Forord	V
Innhold	VIII
Figurer	IX
Forkortelser	XI
1 Introduksjon og motivasjon	1
2 Tidligere arbeid	3
3 Subjektiv undersøkelse av kvalitet for HDTV	5
3.1 Gjennomføringen av undersøkelsen	5
3.1.1 Miljø	5
3.1.2 Kilde signal	6
3.1.3 Utvelgelse av test materiale	6
4 Beskrivelse av rådata	9
4.1 EBU test sekvenser	9
4.1.1 Test sekvens - Dance	9
4.1.2 Konvertering fra YUV8 til AVI	9
4.2 SVT test sekvenser	10
4.2.1 Test sekvens - Crowd Run	10
4.2.2 Test sekvens - Ducks Take Off	10
4.2.3 Test sekvens - In The Tree	10
4.2.4 Test sekvens - Old Town Cross	10
4.2.5 Test sekvens - Park Joy	13

4.2.6	Konvertering fra SGI til AVI	13
4.3	DCP sekvenser	13
4.3.1	Sekvens - The Breakup	15
4.3.2	Sekvens - John Tucker Must Die	15
4.3.3	Sekvens - My Super Ex-Girlfriend	16
4.3.4	Sekvens - The Devil Wears Prada	18
4.3.5	Sekvens - Slither	19
4.3.6	Konvertering fra DCP til AVI	21
4.4	NRK sekvenser	21
4.4.1	Konvertering fra YUV10 til AVI	21
5	Gjennomføring av den subjektive undersøkelsen	25
5.1	Utvalg av matriale	25
5.2	Presentasjon av materialet	25
5.2.1	Tekniske bak presentasjon av matrialet	25
5.2.2	Presentasjon av matrialet	27
5.3	Miljøet hvor den subjektive undersøkelsen ble gjennomført	27
6	Resultater	31
7	Dekoding av JPEG2000 og koding av H.264	35
7.1	Korreksjon av bleke farver i DCP matrialet	35
7.1.1	Bakgrunn	35
7.1.2	Konverterings metode	35
7.2	Koding av H.264	37
8	Refleksjoner om HDTV, Digital kino og HD materiale	39
8.1	Konverteringstid	39
8.2	Bilderate for HDTV	39
9	Konklusjon	41
	Referanser	42

Figurer

2	Konvertering fra YUV8 til AVI	10
1	Test sekvens - Dance	11
3	Test sekvens - Crowd Run	11
4	Test sekvens - Ducks Take Off	12
5	Test sekvens - In The Tree	12
8	Konvertering fra YUV8 til AVI	13
6	Test sekvens - Old Town Cross	14
7	Test sekvens - Park Joy	14
9	DCP sekvens - The Breakup	17
10	DCP sekvens - John Tucker Must Die	17
11	DCP sekvens - My Super Ex-Girlfriend	20
12	DCP sekvens - The Devil Wears Prada	20
14	Konvertering fra YUV8 til AVI	21
13	DCP sekvens - Slithers	22
15	NRK sekvens - nrk004	22
17	Konvertering fra YUV10 til AVI	23
16	NRK sekvens - nrk013	24
18	Representasjon av sammensetningen av spillelisten for del 1 . .	27
19	Representasjon av sammensetningen av spillelisten for del 2 . .	27
20	Grafisk representasjon av skjæring mellom H.264 og JPEG2000 sin bitrate	32
21	Grafisk representasjon av H.264 satt opp mot JPEG2000 ved like bitrater.	32
22	Grafisk representasjon av den generelle kvaliteten til to klipp kodet med H.264 ved forskjellige bitrater.	33
23	Bilde fra DCP i XYZ fargerom, representert som RGB	36
24	Bilde fra DCP i RGB fargerom, representert som RGB	36
25	Konvertering fra JPEG2000 til TIFF med XYZ til RGB kon- vertering	37

Forkortelser

bps Bits per second

DCP Digital Cinema Package

DVD Digital Versatile Disc

FTP File Transfer Protocol

GPL GNU General Public License

GUI Graphical User Interface

H.264 Digital video codec standard (MPEG4 Part 10, AVC)

HD High Definition

HDTV High-definition television

JPEG2000 Graphics File Format

MXF Material eXchange Format

NRK Norsk rikskringkasting AS

TIFF Tagged Image File Format

XML Extensible Markup Language

1 Introduksjon og motivasjon

HDTV begynner nå å bli populært blant befolkningen, det nye digitale bakkenettet er ved å åpne og mange kinoer investerer i digitalt kino utstyr. Digital multimedia, som dette er basert på, bringer med seg mange fordeler. En fordel er at man kan få hvilket som helst format og kvalitet, ved hjelp av en høykvalitets original og en konvertering. Man slipper da lagring og distribusjon av et stort sett med formater. Innenfor HDTV og Digital Kino, som denne Masteroppgaven omhandler, kan vi gjøre akkurat dette. Vi tar for oss multimedia laget for Digital Kino og gjør dette tilgjengelig for kringkasting på HDTV. Det finnes også mange andre muligheter for en slik konvertering, man kan f.eks. tenke seg følgende scenario. På vei til jobb eller skole en morgen mottar du en film trailer på mobiltelefonen, utpå ettermiddagen ser du filmen på kino, etter filmen får du muligheten til å kjøpe DVD'en og senere på natta blir filmen vist på HDTV. Alle disse forskjellige formatene som er nødvendig kan lages ut fra en høykvalitets original. I denne Masteroppgaven legger vi vekt på den subjektive kvaliteten til multimedia i formatene JPEG2000 for Digital Kino og H.264 for HDTV, og ser på hva som er 'riktig' kvalitet.

2 Tidligere arbeid

Måling av kvalitetsnivå er et stort og komplisert tema, og det er skrevet mye om dette de siste årene. Flere standarder og dokumenterte tester er benyttet som grunnlag til denne rapporten. Både for veiledning under selve gjennomføringen og som grunnlag for rapport skrivingen. Følgende artikler, rapporter og dokumenter er benyttet som grunnlag:

- Digital Cinema System Specification v1.0 [1]
- ITU-R BT.500-11 Methodology for the subjective assessment of the quality of television pictures [8]
- Masteroppgave, Digital Kino av Marlon Nielsen [6]
- Masteroppgave, Subjective quality evaluation of the effect of packet loss in High-Definition Video, av Sander Sunde Vorren [9]
- Prosjektoppgave, Digital Kino 2 - HDTV/JPEG2000 av Øystein Hansen [3]

3 Subjektiv undersøkelse av kvalitet for HDTV

I denne Master oppgaven er det gjennomført en undersøkelse hvor kvaliteten til HD materialet ble undersøkt, ved hjelp av en subjektiv test. Grunnen til at det er benyttet en subjektiv test er at på denne måten får man et inntrykk av hva folk flest trenger og ønsker av kvalitet for HDTV sendingene sine, og ikke bare hva ekspertene i fagfeltet mener. Samtidig som vi ser på den HDTV nære standarden H.264 skal vi også se på standarden JPEG2000 som er en typisk bærer for høykvalitets video.

3.1 Gjennomføringen av undersøkelsen

For å gi et troverdig resultat fra undersøkelsen må man gjennomføre undersøkelsen etter bestemte metoder og under bestemte forhold. Retningslinjer for gjennomføring av slike undersøkelser er gitt av ITU [12]. Spesifikke retningslinjer for TV og HDTV kan bli funnet i spesifikasjonen ITU-R BT.500-11 [8], et sammendrag som gjenspeiler de metoder som er benyttet i undersøkelsen er som følger:

3.1.1 Miljø

Miljøet for test personen skal ha følgende egenskaper. Ser her på et hjemme miljø.

- Forholdet mellom en inaktiv skjerm og full lumminans ≤ 0.02
- Lysstyrke og kontrast settes opp etter PLUGE (ITU-R BT.818 og ITU-R BT.815)
- Maks observasjons vinkel 30°
- Ingen bildeprosessering i skjermen
- Full luminans $200cd/m^2$
- Ekstern belysning av skjermen $200lux$
- Personens avstand til skjermen er definert etter PVD tabellen

Skjerm diagonal		Skjerm høyde	PVD
4/3 (in)	16/9 (in)	(m)	
12	15	0.18	9
15	18	0.23	8
20	24	0.30	7
29	36	0.45	6
60	73	0.91	5
>100	>120	>1.53	3-4

Tabell 1: PVD (forholdet mellom skjerm høyde og observatør sin avstand fra skjermen)

3.1.2 Kilde signal

Video kilden som benyttes bør være av ypperste kvalitet, dette for å oppnå stabile resultater. Typisk benyttes digitalt lagrede bilder og videoer som kilde, dette gir utmerket repeterbarhet, både innad i en undersøkelse og mellom uavhengige undersøkelser.

3.1.3 Utvelgelse av test materiale

Det skal velges materiale som kategoriseres som materiale som kan benyttes for TV distribusjon. Altså de må kunne tenkes at test materialet skal sendes på TV. At materialet er av en slik art at de kunne tenkes sendt på TV. For generell evaluering stille det ikke spesielle krav til materialet sin oppbygning, men derimot ved evaluering av kodek standarder må man passe spesielt godt på bitrater samt at materialet vises i riktig bilde rate.

Når man benytter koding og koderen er av en slik art at koderen sin ytelse avhenger mye av senens sammensetning og sammensetning av sener i sekvensen, som er typisk for nye kodeker, er valg av test materiale ennå viktigere.

Generelt er det viktig å inkludere sekvenser i testmaterialet sitt som er av en vanskelig art, dette for å teste koderen på disse områdene. Tanken er den at man kan ikke ekstrapulere resultater fra et generelt sett med sekvenser inn i et vanskelig område, men ved å ta med både generelle sekvenser og vanskelige sekvenser får man et fullværdig resultat for hele området. Man må dog huske på og kun inkludere vanskelige sekvenser som kan forekomme i en virkelig situasjon innen TV kringkastning. Når vi her nevner en såkalt vanskelig sekvens mener vi en sekvens som setter koderen på prøve, f.eks.

en sekvens med høy dynamikk, en sekvens med raske bevegelser, en sekvens med høye kontraster.

4 Beskrivelse av rådata

For Masteroppgaven er det samlet inn video for bruk under tester og presentasjoner. Dette kapitlet tar for seg disse rådataene og beskriver hvor det kommer fra, hvilke formater det har vært i og hvordan det er videre behandlet. Grunnen til denne nøye gjennomgangen av rådataene er for å gi leseren mulighet til å etterkontrollere fremgangsmåten fra rådata til visning på skjerm, og dermed ha mulighet til å avdekke feilkilder som kan ha oppstått underveis.

4.1 EBU test sekvenser

European Broadcasting Union (EBU) har tilgjengelig for sine medlemmer test sekvenser i HD kvalitet [2]. Test sekvenser fra EBU er tilgjengelig som enkeltrammer i YUV8 filformatet. YUV8 filformatet kan også beskrives som YUV422 [10]. Kvaliteten til test sekvensen er av særdeles god kvalitet.

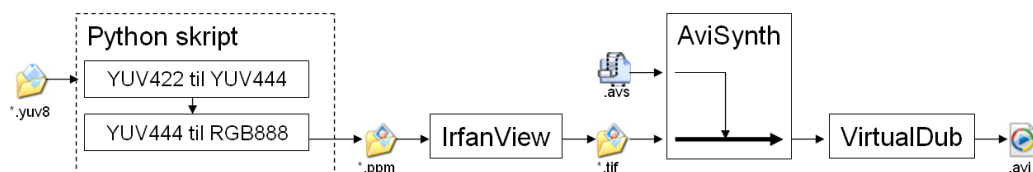
4.1.1 Test sekvens - Dance

Denne sekvensen er av to dansere fra en parade på en fotball arena. Danserne kommer ut av den Brasilianske seksjonen av paraden og viser et dansenummer forran noen fotografer. Figur 1 viser en ramme som er representatibel for innholdet i sekvensen.

Denne sekvensen er levert i formatene 720p50 (1280x720 ved 50fps progressif), 1080p50 (1920x1080 ved 50fps progressif) og 1080i25 (1920x1080 ved 25fps interlaced).

4.1.2 Konvertering fra YUV8 til AVI

For presentasjon og videre behandling av testsekvensen har bilderammene blitt konvertert og samlet til ukomprimerte AVI filer. For konverteringen har det vært benyttet Python skriptet 'yuv8.py' , programmet IrfanView, AviSynth script og VirtualDub.



Figur 2: Konvertering fra YUV8 til AVI

4.2 SVT test sekvenser

Sveriges Televisjon (SVT) har flere test sekvenser som er fritt tilgjengelig via en FTP server på internett [2]. Disse sekvensene er i HD oppløsning og kan lastes ned som enkeltrammer i SGI filformatet [7]. SVT test sekvensene er levert i formatene 720p50 (1280x720 ved 50fps progressift), 1080p50 (1920x1080 ved 50fps progressift) og 1080i25 (1920x1080 ved 25fps interlaced). Alle SVT test sekvensene er klipp fra et seks minutters program kaldt 'Fairytale'. Kvaliteten til test sekvensen er av særdeles god kvalitet.

4.2.1 Test sekvens - Crowd Run

Denne sekvensen består av et hundretalls menesker som løper. Denne sekvensen ses på som vanskelig å kode. Figur 3 viser en ramme som er representatibel for innhodet i sekvensen.

4.2.2 Test sekvens - Ducks Take Off

Denne sekvensen består av enner som er ved å ta av fra et vann. Denne sekvensen ses på som vanskelig å kode. Figur 4 viser en ramme som er representatibel for innhodet i sekvensen.

4.2.3 Test sekvens - In The Tree

Denne sekvensen viser kameraets ferd over en eiendom og inn mot et tre. Denne sekvensen ses på som lett å kode. Figur 5 viser en ramme som er representatibel for innhodet i sekvensen.

4.2.4 Test sekvens - Old Town Cross

Denne sekvensen ses på som lett å kode. Figur 6 viser en ramme som er representatibel for innhodet i sekvensen.



Figur 1: Test sekvens - Dance



Figur 3: Test sekvens - Crowd Run



Figur 4: Test sekvens - Ducks Take Off



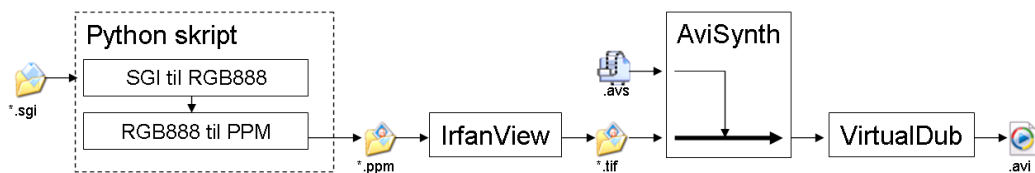
Figur 5: Test sekvens - In The Tree

4.2.5 Test sekvens - Park Joy

Denne sekvensen består av pent kledde menesker løpende langs en elvebredd. Denne sekvensen ses på som vanskelig å kode. Figur 7 viser en ramme som er representatibel for innholdet i sekvensen.

4.2.6 Konvertering fra SGI til AVI

For presentasjon og videre behandling av testsekvensen har bilderammene blitt konvertert og samlet til ukomprimerte AVI filer. For konverteringen har det vært benyttet Python skriptet 'sgi.py', programmet IrfanView, AviSynth script og VirtualDub. En grafisk representasjon av konverteringsmetoden finnes som figur 14.



Figur 8: Konvertering fra YUV8 til AVI

4.3 DCP sekvenser

DCP sekvensene er video ekstrahert fra digitale kino pakker. En digital kino pakke er en samling med XML og MXF filer, satt sammen etter et forhonds-definert system (alla innholdet på en Digital Video Disc (DVD)). Det at dette systemet er definert gjør det mulig å ekstrahere videoen fra de digitale kino pakkene. Videoen er lagret i MXF filene som enkeltrammer komprimert med JPEG2000 [5]. Rammene er kodet med konstant kvalitet og her med høy kvalitet. Under studering av enkeltrammene kan man, med det blotte øyet, ikke oppdage none kompresjons artefakter og kan dermed se på rammene som originale, ennå det er et lite avvik fra produksjonsoriginalen. Video rammene kommer i formatet 2K (2048x1080 i 24fps progressift). Benyttet areal for det faktiske bilde internt i rammen kan dog variere, og hvis det faktiske bilde arealet er mindre en 2K formatet vil bilde være tilført sorte kanter for å fylle 2K formatet. Ved nøye undersøkelse finner man at noen av DCP sekvensene inneholder artefakter. Artefaktene kan ligne på en avart av blokkeffekter, sannsynligvis introdusert under produksjon av traileren sin produksjons original, og ikke på grunn av kompresjon til DCP / JPEG2000



Figur 6: Test sekvens - Old Town Cross



Figur 7: Test sekvens - Park Joy

formatet. Dette betyr at påstanden om at disse DCP'ene er tilnærmet originale fortsatt stemmer, dette siden avviket fra produksjonsoriginalen fortsatt er liten.

4.3.1 Sekvens - The Breakup

Denne DCP sekvensen er fra en trailer på filmen The Breakup med hovedrolleene Vince Vaughn og Jennifer Aniston. Video sekvensen er av god kvalitet, man ser noe film støy generert under opptak. Det kan også tyde på at det er påført filtre eller kompresjon i produksjonskjeden siden man kan se blokkeffekter med en blokkstørrelse på 55x55 punkter. Video sekvensen består av flere sener med varierende mengde bevegelse for objekter og kamera. Figur 9 viser en ramme som er representatibel for innholdet i video sekvensen. Tabell 2 oppsummerer egenskapene til videosekvensen.

Tittel	The Breakup
Bilderate	24
Oppløsning	2048x1080
Antall rammer	2424
Varighet	1 minutt og 41 sekunder
Enkoding	Progresiv
Kompresjon	JPEG2000
Studio	Orca
Produksjons år	2006
Bildeformat	1999x1080
Bildeoffset	25,0
DCP Produsent	Doremi Labs Inc.

Tabell 2: Egenskaper for DCP sekvensen The Breakup

4.3.2 Sekvens - John Tucker Must Die

Denne DCP sekvensen er fra en trailer på filmen John Tucker Must Die. Video sekvensen er av god kvalitet, men man kan se noe film støy sannsynligvis generert under opptak. Videoen består av flere sener, siden filmen sin arena er et amerikansk collage, består mange av senene av mange mennesker og mye bevegelse. Sener av en mer rolig art er også representert i video sekvensen. Figur 10 viser en ramme som er representatibel for innholdet i video sekvensen. Tabell 3 oppsummerer egenskapene til videosekvensen.

Tittel	John Tucker Must Die
Bilderate	24
Oppløsning	2048x1080
Antall rammer	3408
Varighet	2 minutt og 22 sekunder
Enkoding	Progresiv
Kompresjon	JPEG2000
Studio	Orca
Produksjons år	2006
Bildeformat	2048x858
Bildeoffset	0,112
DCP Produsent	Doremi Labs Inc.

Tabell 3: Egenskaper for DCP sekvensen John Tucker Must Die

4.3.3 Sekvens - My Super Ex-Girlfriend

Denne DCP sekvensen er fra en trailer på filmen My Super Ex-Girlfriend. Video kvaliteten er god for denne sekvensen, en smule bedre enn de andre DCP sekvensene, men fortsatt har denne sekvensen også noe film støy. Ser man nærmere på videorammene ser man også her noe blokkeffekt, ikke så fremtredend men kan tydes som blokker på 16x16 punkter. Det er et godt utvalg av sener med forskjellig mengde bevegelse og kompleksitet i video sekvensen. Noen av senene har også en del animerte spesial effekter, dette for å lage sener med unaturlige begivenheter. Disse senene er godt integrert i videoen, og på den måten har de ikke innvirkning på en eventuell videoenkoders ytelse. Figur 11 viser en ramme som er representativ for innhodet i videosekvensen. Tabell 4 oppsummerer egenskapene til videosekvensen.



Figur 9: DCP sekvens - The Breakup



Figur 10: DCP sekvens - John Tucker Must Die

Tittel	My Super Ex-Girlfriend
Bilderate	24
Oppløsning	2048x1080
Antall rammer	3606
Varighet	≈2 minutt og 30 sekunder
Enkoding	Progresiv
Kompresjon	JPEG2000
Studio	Orca
Produksjons år	2006
Bildeformat	2048x858
Bildeoffset	0,111
DCP Produsent	Doremi Labs Inc.

Tabell 4: Egenskaper for DCP sekvensen My Super Ex-Girlfriend

4.3.4 Sekvens - The Devil Wears Prada

Denne DCP sekvensen er fra en trailer på filmen The Devil Wears Prada. Video sekvensen er av god kvalitet, men man kan se noe film støy også i denne videosekvensen. Videoen består av flere sener, hovedsaklig er disse senene fra et lyst kontorlandskap som vil gi en eventuell video koder fine glatte og hvite flater å jobbe med. Dette kan igjen redusere bittrater eller øke kvaliteten på det enkodede resultatet. Figur 12 viser en ramme som er representatibel for innhodet i video sekvensen. Tabell 5 oppsummerer egenskapene til video sekvensen.

Tittel	The Devil Wears Prada
Bilderate	24
Oppløsning	2048x1080
Antall rammer	4403
Varighet	≈3 minutt og 4 sekunder
Enkoding	Progresiv
Kompresjon	JPEG2000
Studio	Orca
Produksjons år	2006
Bildeformat	2048x858
Bildeoffset	0,111
DCP Produsent	Doremi Labs Inc.

Tabell 5: Egenskaper for DCP sekvensen The Devil Wears Prada

4.3.5 Sekvens - Slither

Denne DCP sekvensen er fra en trailer på filmen Slither. En grøsser med alle aspekter av en grøsser film, mørke sener, skumle monstre og vettskremte skuespillere. Video sekvensen er av god kvalitet, man ser noe film støy og noe blokkeffekt med blokkstørrelse på 8x8 punkter. Senene er hovedsaklig mørke og inneholder mye detaljert natur. Figur 13 viser en ramme som er representabel for innhodet i video sekvensen. Tabell 6 oppsummerer egenskapene til video sekvensen.

Tittel	Slither
Bilderate	24
Oppløsning	2048x1080
Antall rammer	2118
Varighet	≈1 minutt og 28 sekunder
Enkoding	Progresiv
Kompresjon	JPEG2000
Studio	Orca
Produksjons år	2006
Bildeformat	1994x1080
Bildeoffset	27,0
DCP Produsent	Doremi Labs Inc.

Tabell 6: Egenskaper for DCP sekvensen Slither



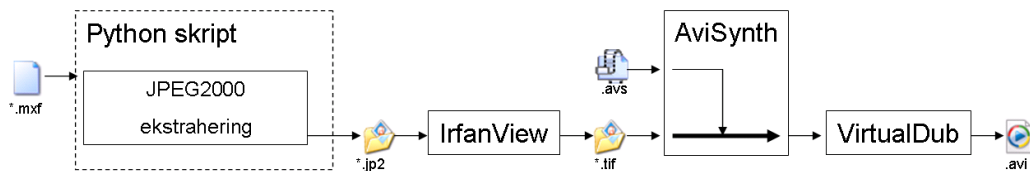
Figur 11: DCP sekvens - My Super Ex-Girlfriend



Figur 12: DCP sekvens - The Devil Wears Prada

4.3.6 Konvertering fra DCP til AVI

For presentasjon og videre behandling av sekvensene har innholdet i MXF filen blitt konvertert til ukomprimerte AVI filer. For konverteringen har det vært benyttet Python skriptet 'mxf_extractt.py', programmet IrfanView, AviSynth script og VirtualDub. Ved å benytte flere AviSynth script får man også andre visningsformater. For disse DCP'ene benytter vi tre visningsformater, 2K (2048x1080 i 24fps progressift), 1080p24 (1920x1080 i 24fps progressift) og 720p24 (1280x720 i 24fps progressift). En grafisk representasjon av konverteringsmetoden finnes som figur 14.



Figur 14: Konvertering fra YUV8 til AVI

4.4 NRK sekvenser

NRK sekvensene er HD video mottatt av NRK for bruk som test sekvenser. Materialet består av to klipp på flere minutter. Klippene er av høy kvalitet, i oppløsningen 1080p50 (1920x1080 ved 50fps progressift). De er tatt-opp med et Sony HDC-1500 kamera og lagret på disk uten komprimering. Figur 15 og 16 viser en ramme fra hver av sekvensene.

4.4.1 Konvertering fra YUV10 til AVI

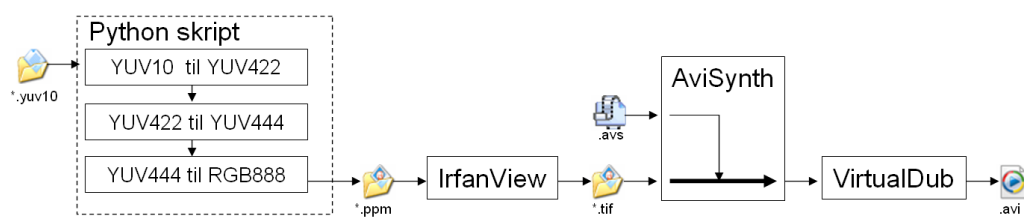
For presentasjon og videre behandling av testsekvensen har bilderammene blitt konvertert og samlet til ukomprimerte AVI filer. For konverteringen har det vært benyttet Python skriptet 'yuv10.py', programmet IrfanView, AviSynth script og VirtualDub.



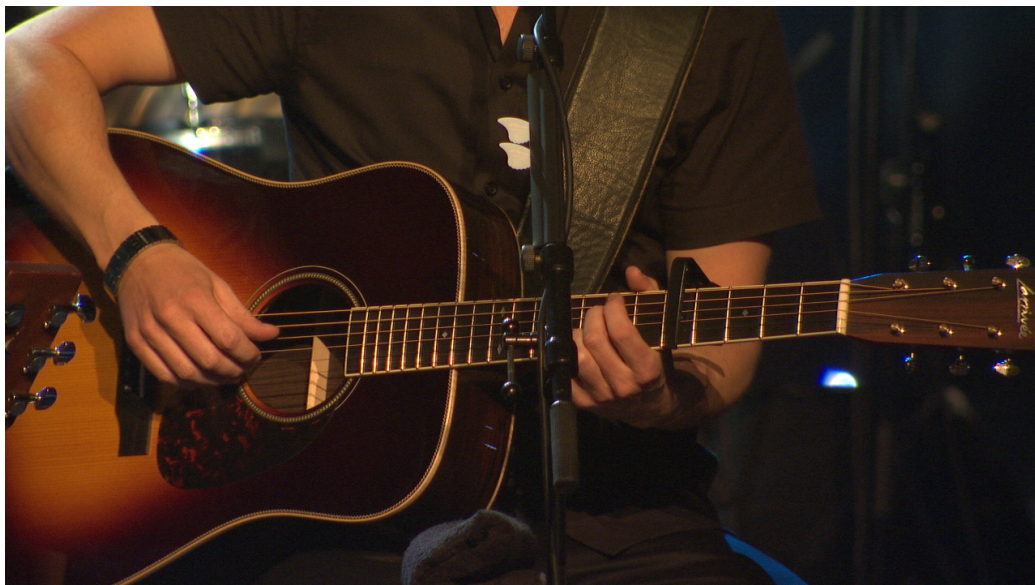
Figur 13: DCP sekvens - Slithers



Figur 15: NRK sekvens - nrk004



Figur 17: Konvertering fra YUV10 til AVI



Figur 16: NRK sekvens - nrk013

5 Gjennomføring av den subjektive undersøkelsen

Gjennomføring av den subjektive undersøkelsen kom igang i begynnelsen av mai, og ble avsluttet i begynnelsen av juni. Jeg vil her beskrive hva som ble gjort og under hvilke forhold, slik at det i etterkant kan vurderes sammen med resultatene.

5.1 Utvalg av materiale

Video sekvensen “In to Tree” ble benyttet for å introdusere deltakeren til koderene og kvalitene som ble benyttet i testen. Sekvensen ble vist i 10 sekunder i de kodede formatene JPEG2000 og H.264, og i to bitrater hver.

Video sekvensen “Crowd Run” ble benyttet for å definere skjæringspunktet mellom JPEG2000 og H.264 i bitrate. Denne sekvensen ble vist i flere bitrater i kode formatene JPEG2000 og H.264, hvor en stigende H.264 bitrate ble satt opp mot en synkende JPEG2000 bitrate. For denne video sekvensen lå bitratene i området 750Kbps til 10Mbps.

Video sekvensen “Park Joy” ble benyttet for å se på JPEG2000 og H.264 ved ekstremt høye og ekstremt lave bitrater. Her ble sekvensen vist i bitrater fra 10Mbps og oppover, og fra 1Mbps og nedover.

Et 10 sekunders sekvens fra NRK materialet (NrK01), og et 10 sekunders sekvens fra DCP materialet Slither (Slither02), ble benyttet i del to av den subjektive testen her ble video sekvensene vist i bitrater fra 250kbps til 15Mbps i formatet H.264. Her skulle deltakeren gi sin mening på den generelle kvaliteten på video sekvensene, og på den måten gi en veiledning til hva som er riktig kvalitet for kringkasting av slik video.

5.2 Presentasjon av materialet

5.2.1 Tekniske bak presentasjon av materialet

Video materialet ble vist på en, Sony KDL-46X2000, 46 tommer HDTV i 720p25 (1280x720 punkter med 25 bilder per sekund). HDTV'en var koblet til en PC via HDMI og oppløsningen 1080p50 (1920x1080 punkter med 50 bilder per sekund). Dette førte da til at man kunne benytte “pixel to pixel mapping” på HDTV'en. Skaleringen fra 720p25 til 1080p50 ble utført på PC'en. For å oppnå mest mulig korrekt bilde på HDTV'en ble alle interne filtre og kopansasjoner skrudd av på HDTV'en. Tabell 7 viser innstillingene

Picture:	
↦PictureMode	= Custom (Vivid, Standard , Custom)
↦Backlight	= 5 (0 - 10)
↦Contrast	= 10 (0 - 10)
↦Brightness	= 50 (0 - 100)
↦Colour	= 50 (0 - 100)
↦Colour Tempterature	= Neutral (Cool, Neutral, Warm 1, Warm 2)
↦Sharpness	= 50 (0 - 100)
↦Noise Reduction	= Off (High, Medium, Low, Off, BNR)
↦Advanced Settings:	
↦↦Black Corrector	= Off (Max, High, Mediu n , Low , Off)
↦↦Adv. Contrast Enhancer	= Off (High, Medium , Low, Off)
↦↦Gamma	= Off (Max, High, Medium, Low , Off)
↦↦Clear White	= Off (High, Low, Off)
↦↦Colour Space	= Normal (Wide, Normal)
↦↦White Balance:	
↦↦↦R Gain	= 0 (-20 - 0)
↦↦↦G Gain	= 0 (-20 - 0)
↦↦↦B Gain	= 0 (-20 - 0)
↦↦↦R Bias	= 0 (-10 - 10)
↦↦↦G Bias	= 0 (-10 - 10)
↦↦↦B Bias	= 0 (-10 - 10)
↦↦Detail Enhancer	= Off (Max, High, Medium , Low, Off)
↦↦Edge Enhancer	= Off (High, Medium, Low , Off)

Tabell 7: Innstillingene benyttet for Sony KDL-46X2000 under undersøkelsen

som ble benyttet på HDTV'en. I tabellen er standard verdier uthevet, og mulige verdier er i parentes.

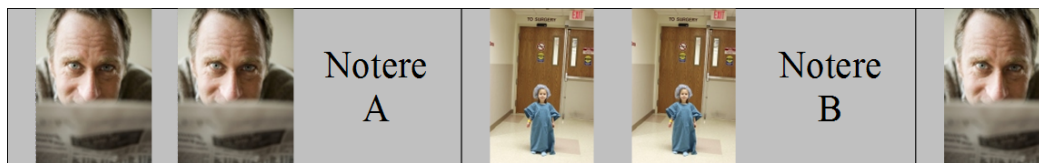
En kombinasjon av programmet “Windows Media Player 11” [13] og “Event-Ghost” [11] ble benyttet for å gi deltakeren en følelse av å bruke en TV i motsetning til en PC som det faktisk ble spilt av fra. “Windows Media Player 11” holdt styr på fremvisning og spillelisten, mens “EventGhost” muliggjorde det å bruke fjernkontroll for styring av mediaspilleren og spillelisten. Deltakeren kunne da enkelt forflytte seg frem i spillelisten når deltakeren var ferdig med å noterer ned sine resultater.

5.2.2 Presentasjon av materialet

Den subjektive testen var delt opp i to deler. Del 1 som inneholdt “In to Tree”, “Crowd Run” og “Park Joy” viste en sekvens i to formater i like eller ulike bitrater to ganger slik som indikert i figur 18. Ut fra det deltakeren kunne se skulle deltakeren notere ned i hvilken kvalitet video sekvens A hadde i forhold til video sekvens B. Del 2 inneholdt “NrK01” og “Slither02”. Disse sekvensene ble vist hver for seg og hver 10 sekunders sekvens ble vist to ganger, så fikk deltakeren tid til å notere ned hvilken kvalitet deltakeren syntes video sekvensen generelt hadde. Figur 19 illustrerer slik videosekvensene fremstod.



Figur 18: Representasjon av sammensetningen av spillelisten for del 1



Figur 19: Representasjon av sammensetningen av spillelisten for del 2

Alle klippene ble vist i en semitilfeldig rekkefølge. (en rekkefølge som vil se tilfeldig ut for deltakeren, men er fast fra deltaker til deltaker, og kan ha overordnet struktur.) Tabell 8 viser den faktiske rekkefølgen som ble benyttet.

5.3 Miljøet hvor den subjektive undersøkelsen ble gjennomført

Den subjektive undersøkelsen ble gjennomført på KafeMedia (B133) i elektrobygget, NTNU Gløshaugen. Dette var en lab som kunne dedikeres til undersøkelsen, slik at forstyrrelser fra medelever eller ansatte ikke skulle skje under undersøkelsene. Rommet ble tilpasset så godt som mulig til ITU spesifikasjonen, men følgende punkter ble avvirket for å kunne gjennomføre testen.

- Lysforhold ble ikke testet i rommet, pga. manglede måle utstyr. Generelt var lysstyrken i rommet noe høyt siden det ble benyttet den floriserende takbelysningen, som var den eneste tilgjengelige belysningen i dette rommet.
- HDTV skjermen ble ikke testet via PLUGE, da måleutstyr for test og test bilde ikke var tilgjengelig.

InToTree	H.264	250kbps
InToTree	JPEG2000	750kbps
InToTree	H.264	8Mbps
InToTree	JPEG2000	80Mbps
CrowdRun	JPEG2000	10Mbps
CrowdRun	H.264	750Kbps
CrowdRun	H.264	4Mbps
CrowdRun	JPEG2000	8Mbps
CrowdRun	JPEG2000	6Mbps
CrowdRun	H.264	2Mbps
CrowdRun	JPEG2000	4Mbps
CrowdRun	H.264	4Mbps
ParkJoy	H.264	10Mbps
ParkJoy	JPEG2000	10Mbps
ParkJoy	JPEG2000	20Mbps
ParkJoy	H.264	20Mbps
ParkJoy	H.264	40Mbps
ParkJoy	JPEG2000	40Mbps
ParkJoy	JPEG2000	1Mbps
ParkJoy	H.264	1Mbps
ParkJoy	JPEG2000	500Kbps
ParkJoy	H.264	500Kbps
ParkJoy	H.264	250Kbps
ParkJoy	JPEG2000	250Kbps
Nrk01	H.264	250Kbps
Nrk01	H.264	2Mbps
Nrk01	H.264	15Mbps
Nrk01	H.264	4Mbps
Nrk01	H.264	8Mbps
Nrk01	H.264	750Kbps
Slither02	H.264	250Kbps
Slither02	H.264	2Mbps
Slither02	H.264	15Mbps
Slither02	H.264	4Mbps
Slither02	H.264	8Mbps
Slither02	H.264	750Kbps

Tabell 8: Spilleliste for subjektiv test

6 Resultater

Fra den subjektive undersøkelsen ble det hentet inn data fra deltakerene. Disse dataene ble behandlet og vi kommer så ut med følgende resultater. Figur 20 viser i hvilken grad H.264 er bedre eller dårligere enn JPEG2000 ved bestemte bitrater. Skalaen er som følger:

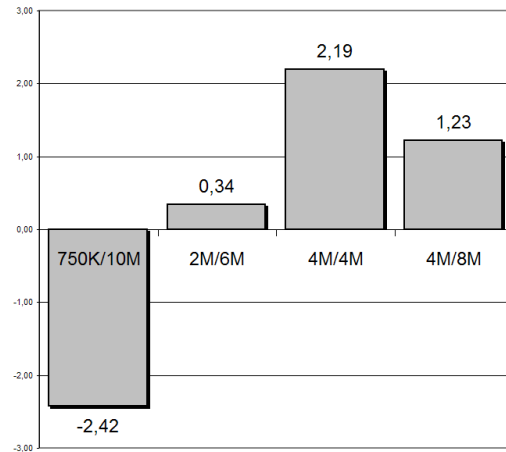
- 3 H.264 er *mye værre* enn JPEG2000
- 2 H.264 er *værre* enn JPEG2000
- 1 H.264 er *litt værre* enn JPEG2000
- 0 H.264 er *lik* med JPEG2000
- 1 H.264 er *litt bedre* enn JPEG2000
- 2 H.264 er *bedre* enn JPEG2000
- 3 H.264 er *mye bedre* enn JPEG2000

Bitratene er satt opp slik at man skal kunne se et skjæringspunkt mellom JPEG2000 og H.264 sine bitrater, dette skjæringspunktet vil være når kvaliteten er lik for koderne. Fra figur 20 ser vi at dette punktet ligger noe før 2Mbps / 6Mbps punktet, hvor H.264 er kodet med 2Mbps og hvor JPEG2000 er kodet med 6Mbps. Jeg estimerer at skjæringspunktet vil ligger rundt 1.7Mbps / 6Mbps, eventuelt også representert som 2Mbps / 7Mbps. Det vi ser ut av dette er at, for det spesifiserte området, H.264 gir mye mindre bitrate enn JPEG2000 for samme kvalitet.

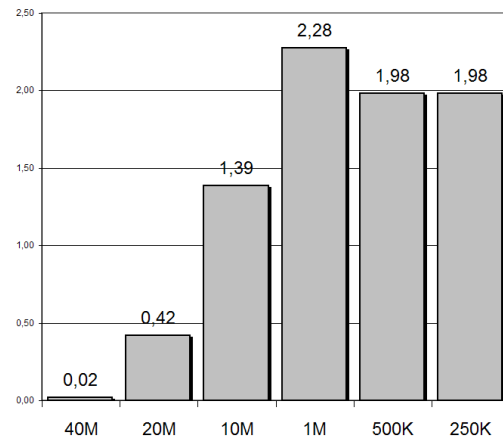
Figur 21 tar for seg H.264 og JPEG2000 ved lik bitrate, vi ser her på hvilken av koderne som gjør det best i områder med høy bitrate og områder med lav bitrate. I området fra 1Mbps ned til 250Kbps ser vi at H.264 gjør det bedre enn JPEG2000. I området 10Mbps opp til 40Mbps ser vi at H.264 også her gjør det bedre enn JPEG2000, men vi ser også tendensen om at de blir svært like ved høye bitrater.

Figur 22 viser den generelle kvaliteten til to video sekvenser. Dette plottet er skalert med følgende intervall:

- 4-5 Videoen er *utmerket*
- 3-4 Videoen er *bra*



Figur 20: Grafisk representasjon av skjæring mellom H.264 og JPEG2000 sin bitrate



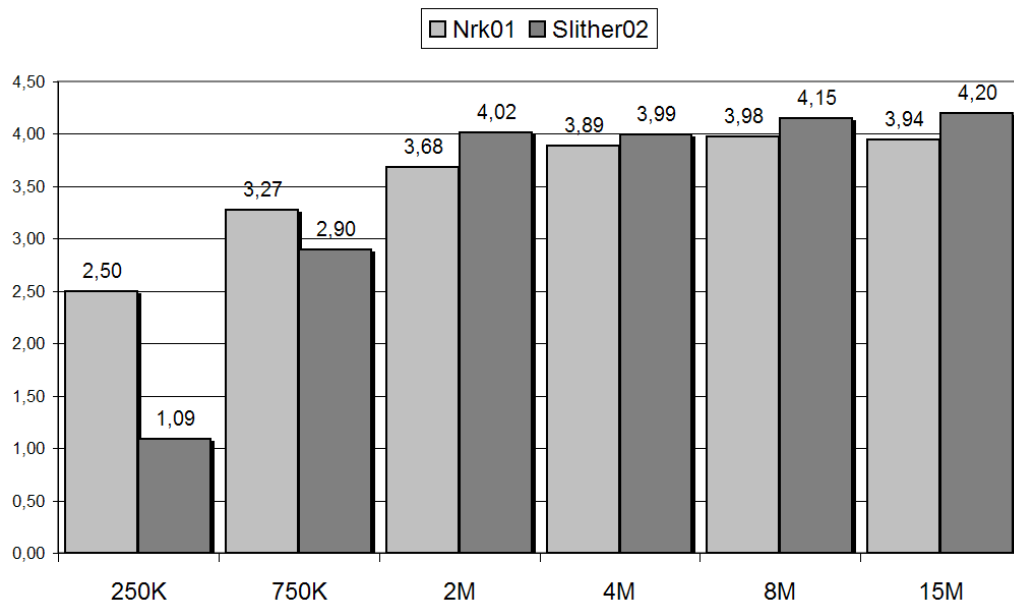
Figur 21: Grafisk representasjon av H.264 satt opp mot JPEG2000 ved like bitrater.

2-3 Videoen er *tilfredsstillende*

1-2 Videoen er *dårlig*

0-1 Videoen er *veldig dårlig*

Hver av video sekvensene er representert i bittrater fra 250Kbps til 15Mbps. For “Slither02” video sekvensen, som representerer en typisk film, ser vi at ved høye bitrater, fra 2Mbps og oppover mener deltakeren at video sekvensen ikke oppnår noe særlig kvalitets økning. Under 2Mbps ser vi at deltakeren ikke synes video sekvensen er mer en tilfredsstillende, og når vi passerer 250Kbps synes deltakeren at videoen blir veldig dårlig. For “NrK01” video sekvensen ser vi dermed at videoen er tilfredsstillende allerede ved 250Kbps. “NrK01” video sekvensen er av en typisk NRK produksjon med mye statiske sener og rolige kameraføringer, som nok gir H.264 koderen økt kompresjon under koding. For området over 2Mbps ser vi at, også for “NrK01” video sekvensen, deltakeren mener at video sekvensen ikke oppnår noe særlig kvalitets økning.



Figur 22: Grafisk representasjon av den generelle kvaliteten til to klipp kodet med H.264 ved forskjellige bitrater.

7 Dekoding av JPEG2000 og koding av H.264

7.1 Korreksjon av bleke farver i DCP matrialet

I slutfasen av Masteroppgaven, ble problemet med bleke farger for DCP matrialet løst. Beskrivelsen av konvertering fra JPEG2000 til avi som beskrevet i kapittel 4.3.6 stemmer for de video sekvensene som ble benyttet under den subjektive undersøkelsen. Men siden en ny metode nå foreligger for denne konverteringen skal jeg her gå inn på denne.

7.1.1 Bakgrunn

Videoen til en DCP kommer pakket som enkeltstående JPEG2000 rammer. Disse rammene er i forkant av kompresjon med JPEG2000, kodet i fargerommet XYZ, dette forde det utstyret som benyttes for fremvisning av matrialet på en digital kino, arbeider i XYZ fargerummet. Når en slik JPEG2000 ramme åpnes på en PC som jobber i RGB fargerummet, burde programvaren som benyttes gjøre denne konverteringen uoppfordret, JPEG2000 rammene er tross alt merket med profilen “CINEMA2K” som tilsier dette fargerummet. Programvaren Kakadu som er benyttet i Masteroppgaven, sies å ha støtte for “CINEMA2K” profilen, men mangler tydeligvis denne XYZ til RGB konverteringen. En representasjon for denne fargerom-feilen vises med figur 23 og 24. Her viser figur 23 en ramme fra en DCP uten konvertering fra XYZ til RGB. Rammen vil da være representert feil siden den blir tolket her som RGB. Figur 24 har gått gjennom konverteringen fra XYZ til RGB og fremkommer korrekt.

7.1.2 Konverterings metode

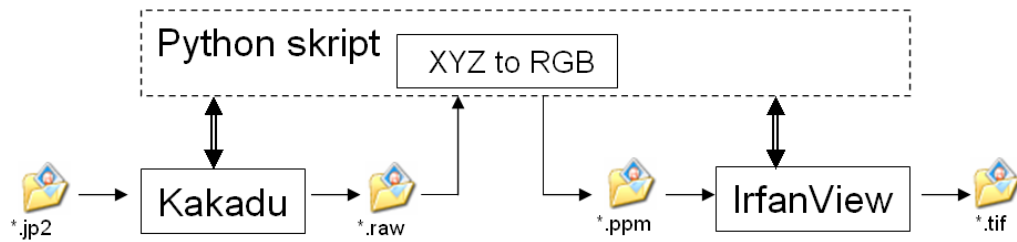
Vi lar Kakadu programvaren eksportere rå-pixeldata istedet for å nedsample og lagre i et TIFF format. Så benytter vi Python til å samle rå-pixeldataene, gjøre XYZ til RGB transformen, og til slutt lagre det hele til TIFF filer. Python skriptet “DciRaw.py” gjør denne jobben. Skriptet er et overordnet skript som man starter med parametre om ‘fra katalog’ og ‘til katalog’ og ordner selv med å starte og stoppe både Kakadu programvaren og IrfanView programvaren. Figur 14 viser en grafisk representasjon av denne konverteringen.



Figur 23: Bilde fra DCP i XYZ fargerom, representert som RGB



Figur 24: Bilde fra DCP i RGB fargerom, representert som RGB



Figur 25: Konvertering fra JPEG2000 til TIFF med XYZ til RGB konvertering

7.2 Koding av H.264

Som i prosjekt oppgaven “Digital Kino 2 - HDTV/JPEG2000” [3], benytter vi x264 enkoderen for H.264 enkoding. De tre store fordelene med denne enkoderen sett fra mitt ståsted er at

1. x264 enkoderen er GPL, altså gratis å benytte så lenge man følger lisensbetingelsene.
2. x264 enkoderen er styrt via et shell. Slik kan den benyttes i BAT filer og i skript for full kontroll og automasjon.
3. x264 har støtte for flere kvalitetsbringende egenskaper ut over grunnprofilen.

I Master oppgaven, i motsetning til under prosjekt oppgaven, benytter vi ikke det grafiske grensesnittet til x264 men istedet utnytter vi shell egenskapen og automatiserte prosessen som genererer H.264 enkodete videosekvenser. Dette medførte at man kunne generere de H.264 enkodete videosekvensene, som skulle representeres i et titals forskjellige bittrater, for den subjektive testen med enkelhet. Med enkelhet mener vi at man kunne i tekstfiler spesifisere hvilke bitrater, hvilke videosekvenser og hvilke H.264 enkodings profiler som skulle benyttes, og over tid ble de ønskede video sekvensene, med hver sine parametre, generert automatisk.

8 Refleksjoner om HDTV, Digital kino og HD materiale

8.1 Konverteringstid

Konverteringstiden mellom en DCP og HDTV er lang. Typisk konverteres 1 til 4 bilder per sekund i snitt for hele konverteringsalgoritmen. Raskere PC'er i fremtiden vil redusere konverteringstiden, men en metode for å oppnå nær sanntidskonvertering har vi i dag ved å benytte flere PC'er sammen for å konvertere en video. JPEG2000 og H.264 muliggjør dette ved distribuert prosesering. Man setter opp en maskinpark og lar hver maskin konvertere en liten del av hele videoen. I dag finnes ikke en effektiv metode for dette for HDTV, men dette kan enkelt bli implementert via Python skript.

8.2 Bilderate for HDTV

Man sier at 25 bilder per sekund er nok til å lure hjernen til å tro at man ser på et levende bilde. Med de testene som ble gjennomført på Sony HDTV'en så man derimot at dette ikke lenger stemmer. For en HDTV som viser med høy oppløsning og ikke smører bildet utover, ser man tydelig hakking i scener med raske bevegelser. Dette kommer sannsynligvis av at med den høyere oppløsningen oppfatter øyet skjæringslinjen mellom objekter og bakgrunn bedre og når disse beveger seg, forflytter skjæringslinjen seg en oppdagbar avstand for hver ramme, og dermed ser hakkete ut. Ved tester, med tilsvarende video, ved 50 bilder per sekund kan man ikke se denne hakkingen. Etter min bedømming vil man få et bedre resultat ved å øke bilderaten til 50fps på en 720p video kontra å øke video størrelsen til 1080p. Noen kompresjonsalgoritmer er kjent for å smøre bildet, så kompresjon og / eller filtrering kan eventuelt fjerne noe av hakkingen. Det kan også være andre grunner til at videoen ble oppfattet som hakkete ved 25fps, så videre undersøkelser på dette er nødvendig for å kunne dra en konklusjon.

9 Konklusjon

Fra den subjektive undersøkelsen ser vi at H.264 gir subjektivt bedre kvalitet enn JPEG2000 i det testede området 250Kbps til 40Mbps, vi ser at H.264 har et stort forsprang i de lave og middels bitratene, men ved høye bitrater tar JPEG2000 igjen H.264 og de er tilsynelatende like gode. For H.264 ble det for visse videomaterial-typer indikert at en bitrate på 250Kbps var tilfredsstillende, og generelt lå en bra video kvalitet på rundt 2Mbps.

Det å benytte en DCP som grunnlag for kringkasting til HDTV er nå gjennomførbart, verktøyene er nå tilstede og teknikken er beskrevet. Det vi derimot ser er at kvaliteten ikke er tilstrekkelig for visse DCP'er, så man må være oppmerksom på kvaliteten på materialet man baserer seg på. Avhengig av kompresjon og bruk av filtrering, kan også bilderaten ha noe å si på den subjektive opplevelsen. Seeren kan typisk oppleve hakking i scener med mye bevegelse grunnet DCP'en sin lave bilderate på 24 bilder per sekund.

Referanser

- [1] LLC Digital Cinema Initiatives. Digital cinema system specification v1.0, July 2005.
- [2] www.ebu.ch/en/technical/hdtv/test-sequences.php European Broadcasting Union. Hd test sequences.
- [3] Øystein Hansen. Digital kino hdtv / jpeg2000, desember 2006.
- [4] NTNU Institutt for elektronikk og telekommunikasjon. Forslag til masteroppgaver innen signalbehandling, 2007.
- [5] Jim Wilkinson Matt Beard Nick Wells, Bruce Devlin and Phil Tudor. The mxf book. Focal Press, 2006.
- [6] Marlon Thomas Montejo Nielsen. Digital kino, produksjon av en digital cinema package". Master's thesis, Norges teknisk-naturvitenskapelige universitet, Juni 2006.
- [7] local.wasp.uwa.edu.au/~pbourke/dataformats/sgirgb/sgiversion.html Paul Haeberli, Silicon Graphics Computer Systems. The sgi image file format.
- [8] International Telecommunication Union. Methodology for the subjective assessment of the quality of television pictures.
- [9] Sander Sunde Vorren. Subjective quality evaluation og the effect of packet loss in high-definiton video. Master's thesis, Norwegian University of Science and Technology, June 2006.
- [10] wikipedia.org/wiki/YUV. Wikipedia - yuv.
- [11] www.eventghost.org. Eventghost.
- [12] www.itu.int. International telecommunication union.
- [13] www.microsoft.com/windows/windowsmedia/. Windows media player.