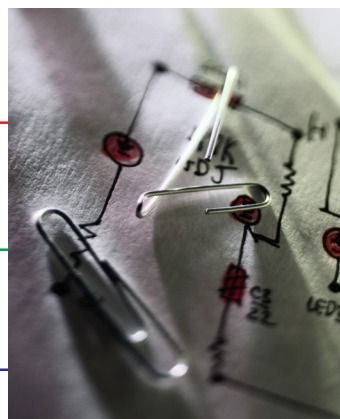


Digital Video Audio



Udarbejdet af:
Flemming Rathsach
Videoinform ApS.
© 2004
Uredigeret førsteudgave

 **Videoinform**
VIDEOWORLD

Kongevejen 49, 2840 Holte
Telefon: 4454 0999; Fax: 4454 0998
E-Mail: av@videoinform.dk
Home page: www.videoinform.dk

Indhold

Historie	02
Rank 9000	02
Standarden siger:	02
Det digitale signal	03
Digitale grundbegreber	03
Eksempel:	03
Det Binære Tal system	03
Eksempel:	03
Eksempel:	04
Sampling	05
CCIR-601	05
4:2:2	05
4:1:1	06
4:2:0	06
4:2:2:4	06
Sammensætning af udstyr med forskellig signalbehandling	07
Hvorfor anvende digital teknik	08
Parallel Digital Video	08
Seriell Digital Video	08
Digital Audio	08
SDI Signalet	09
Tekniske udtryk	010
Online redigering	010
Offline redigering	010
Lineær redigering	010
Nonlineær redigering	011
Hybrid løsninger	011
PC'en som produktions værktøj	011
Den rene PC løsning	011
Hybrid løsningen	011
Den rene harddisk løsning	011
Komprimering	012
Intra-frame komprimering (JPEG)	012
Sekvens 1	013
Sekvens 2	013
Inter-frame komprimering (MPEG2 4:2:2P@ML)	013
JPEG komprimering	014
MPEG komprimering	014
Fast eller Dynamisk	014
Effekter	014
JPEG OG MPEG komprimering	015
MPEG Del I	015
JPEG	016
MPEG Del II	017
MPEG Kodning	017
Groups of Pictures	018
Sony "SX" komprimering	019
Tekniske specifikationer for SX komprimering:	019
De Digitale formater	020
Sampling	020
DVCAM	021
DVCPRO	021
Audio Lock Mode	021
Konklusion	022

Historie

Allerede i starten af 70'erne begyndte man at eksperimentere med at indspille digitale videosignaler på en videobåndmaskine.

Rank 9000

De første forsøg blev udført i England på en Rank 9000 2" videobåndmaskine. Samme maskine blev præsenteret på Montreux udstillingen i 1977, hvor den indspillede et digitalt videosignal med en hastighed på 8" pr. sekund.

På IBC i 1978 blev der vist en BCN 1" maskine med samme tekniske data som Rank 9000, men med et båndforbrug på 10" pr. sekund.

Det digitale signal var i begge maskiner baseret på det composite videosignal, der blev digitaliseret med en sampling frekvens på $2 \times f_{sc}$ (subcarrier frekvens).

Allerede året efter i Montreux viste Ampex en 2" maskine med en $4 \times f_{sc}$ sampling teknik.

Sony viste en modificeret 1" maskine der arbejdede med $2 \times f_{sc}$.

Umiddelbart efter denne udstilling stod det klart, at det var nødvendigt at få fastsat en digital videostandard, alle kunne leve med. De to organisationer EBU og SMPTE nedsatte i 1980 specialistgrupper, der skulle udarbejde foreslag til en digital videostandard. På trods af at det var englænderne, der først konstruerede en digital video recorder, var det Sony der fik den største indflydelse på den standard, der blev vedtaget i 1981.

Den norm der blev vedtaget hedder CCIR601 for videorecording og playback. Standarden kaldes 4:2:2, hvilket refererer til de tre component signaler Y, R-Y og B-Y.

Standarden siger:

Antallet af samples pr. aktive linie skal være :
864 for Y-signalet og
432 for hhv. R-Y og B-Y signalerne.

Med en billed frekvens på 50 fields pr. sekund og 625 linier, bliver samplingfrekvensen for Y-signalet: $312,5 \times 864 \times 50 = 13,5 \text{ MHz}$.

For R-Y og B-Y signalerne bliver de tilsvarende sampling frekvenser:
 $312,5 \times 432 \times 50 = 6,75 \text{ MHz}$ for hhv. R-Y og B-Y signalerne.

Hver sample skal have en opløsning på 8 bit svarende til 256 mulige kombinationer. Der skal være plads til 4 digitale lydspor samt lineære spor for time code og audio cue spor.

Det var et ønske fra EBU at der skulle anvendes en videokassette som lageringsmedie. Med Sony's erfaring fra U-Matic kassetten blev der valgt en kassette med et 19 mm bredt bånd og en spilletid på 90 minutter.

Man blev enige om, at der skulle kunne laves op til 10 generationer uden tab af kvalitet i video- og audiosignalet.

Det digitale signal

Den nyeste udvikling indenfor kamerateknik er det digitale kamera. I stedet for at behandle signalet der kommer fra CCD cellerne som et analogt signal, digitaliseres signalet umiddelbart efter det er kommet ud af CCD cellen.

Hvad betyder det så at signalet digitaliseres? I en A/D-konverter (analog til digital konverter) laves det analoge signal om til et digitalt signal, et signal der består af en række tal der kun kan have to tilstande „1“ eller „0“. Antallet af et-taller og nuller, samt deres indbyrdes placering giver en bestemt talværdi.

Digitale grundbegreber

Digitale signaler er bygget op omkring det binære talsystem.

I det Binære tal system eksistere der kun to værdier:

„1“ (HØJ) for åben eller „0“ (LAV) for lukket. Tallene 0 og 1 kan have forskellig fysisk mening. Eksempelvis kan 1 betyde at der er et signal tilstede og 0 at der ikke er et signal, eller 1 kan betyde en positiv spænding og 0 en negativ. Med de to tal er der beskrevet en tilstand, men ingen værdi.

De to tilstande 1 og 0 kaldes Bit.

Det talsystem vi normalt arbejder med kaldes decimaltalsystemet og går fra 0 til 9. En talværdi kan bestå af flere tal, der er ordnet i en rækkefølge med en faldende eksponent af roden 10.

Eksempel:

$$259 = 2 \times 10^2 + 5 \times 10^1 + 9 \times 10^0$$

$$200 + 50 + 9 = 259$$

Det Binære Tal system

I det Binære tal system eksisterer der kun to tal 0 og 1.

Her anvendes roden 2 med en faldende eksponent.

Eksempel:

$$1 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 1 \quad 1$$

$$1x2^8 + 0x2^7 + 0x2^6 + 0x2^5 + 0x2^4 + 0x2^3 + 0x2^2 + 1x2^1 + 1x2^0 =$$

$$256 + 0 + 0 + 0 + 0 + 0 + 0 + 2 + 1 = 259$$

Omsætning af et decimaltal til et Binært tal foregår efter følgende formel. Tallet divideres med to, hvis resultatet er et helt tal noteres et 0. Resultatet divideres med 2 igen. Bliver det nye resultat et decimaltal noteres et 1, herefter nedrundes til nærmeste hele tal der igen divideres med 2. Processen fortsættes til resultatet bliver 0.

Eksempel:

22:2	0	<i>Mindst betydende ciffer (LSD)</i>
11:2	1	
5:2	1	
2:2	0	
1:	1	<i>Mest betydende ciffer (MSD)</i>

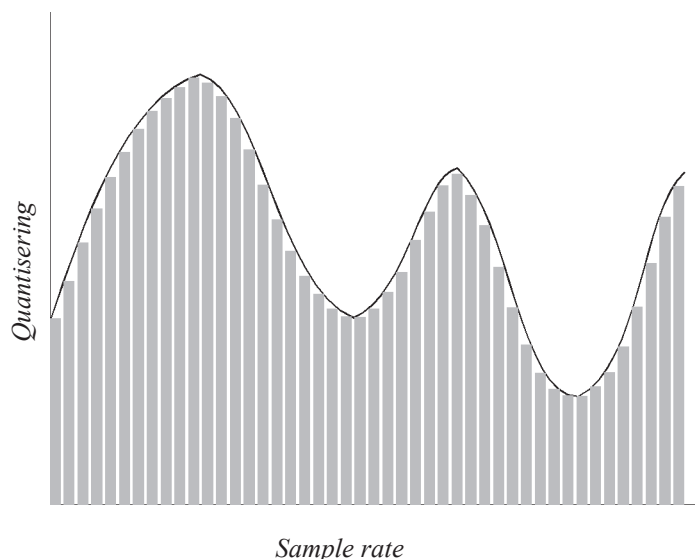
22 omskrevet til et binært tal ser herefter således ud:

$$\begin{array}{cccccc}
 1 & 0 & 1 & 1 & 0 & \\
 1x2^4 & + & 0x2^3 & + & 1x2^2 & + & 1x2^1 & + & 0x2^0 & \\
 16 & + & 0 & + & 4 & + & 2 & + & 0 & = & 22
 \end{array}$$

Når der her anvendes udtrykket mest og mindst betydende ciffer er det fordi det er vigtigt fra hvilken side man læser det binære ord. Som vi så før, har det mest betydende ciffer i det binære tal 100000011, værdien 256 medens det mindst betydende har værdien 0 eller 1.

Digitalisering af videosignaler er ikke noget nyt, teknikken har været anvendt i digitale time base correctore i mange år. Standarden for det digitale video signal blev fastsat i 1981 af de to organisationer EBU og SMPTE.

Sampling



Sampling betyder det antal gange man måler på et analogt signal. En af de vigtigste grundregler indenfor samplingteknikken er udarbejdet af matematikeren Nyquist, der sagde at sampling- frekvensen mindst skal være det dobbelte af den højeste frekvens, der ønskes gengivet.

Quantiseringen er den nøjagtighed, man måler med, dvs. det antal bit der bruges i hver sample. I CCIR 601 normen er det fastsat til 8-bit.

I de sidste nye digitale kameraer er man begyndt at anvende en Quantisering på 10-bit, det giver en opløsning på 1024 mulige kombinationer, eller sagt på en anden måde 1024 gråtoner i videosignalet.

CCIR-601

Selvom CCIR-601 standarden er blevet erstattet af ITU-R standarden, anvendes CCIR stadig som synonym for 4:2:2 sampling i digitalt videoudstyr.

4:2:2

Digitalt videoudstyr der arbejder efter CCIR-601 standarden, altså 4:2:2, samples Y-signalet med 13,5 MHz og de to difference signaler (R-Y) og (B-Y) med hver 6,75 MHz. Det giver en maksimal farvebåndbredde på 3,37 MHz, hvilket er nok til lave chroma key i en høj kvalitet.

Med CCIR-601 normen som udgangspunkt er der opstået en række nye formater baseret på de samme grundfrekvenser som i 4:2:2.

4:1:1

Her anvendes en samplingfrekvens på 13,5 MHz for Y-signalet og for hver af de to differenssignaler en samplingfrekvens på 3,37 MHz. Her har man bevaret luminans båndbredden, men halveret chrominance båndbredden.

4:2:0

Y-signalet samples med 13,5 MHz i hver linie mens (R-Y) og (B-Y) samples med 6,75 MHz i hver anden linie. Eksempelvis en linie samples med 4:0:0, den næste med 4:2:2. I en normal 4:2:2 sampling er chrominance opløsningen i en vertikal linie det halve af opløsningen i en horisontal linie. I en 4:2:0 sampling er chrominance opløsningen den samme vertikalt og horisontalt.

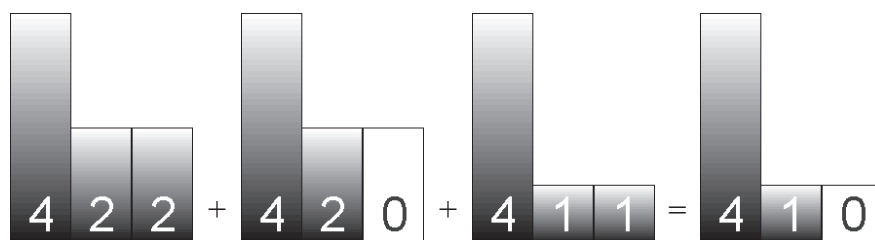
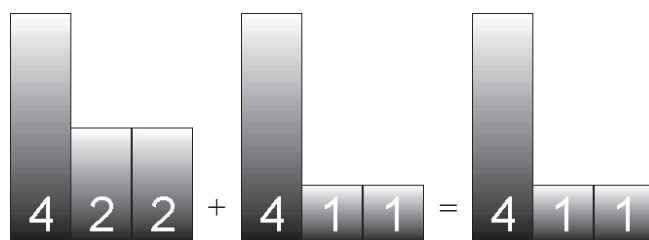
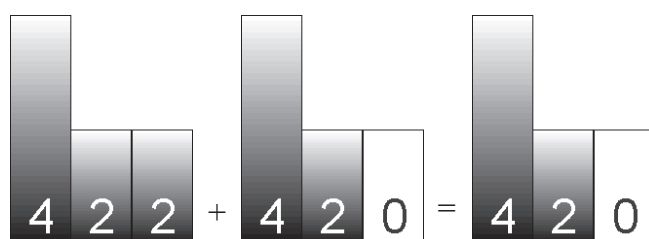
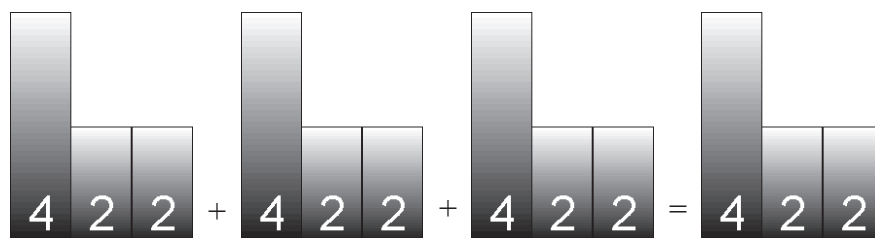
4:2:2:4

Er det samme som 4:2:2 men med en key kanal som den fjerde komponent.

Når man i digitalt udstyr har valgt at reducere sampling frekvenserne for de to farve component signaler, er det for at mindske antallet af data der skal behandles og dermed gøre signalbehandlingen lettere. Det er dog ikke uden omkostninger. Skal et 4:2:0 eller et 4:1:1 signal signalbehandles mange gange sker der en forringelse af farveopløsninge. En forringelse der gør det svært at opnå en god chroma key effekt.

En anden ting der skal tages hensyn til er sammensætning af udstyr. Hvis et 4:2:2 samplet signal skal igennem en enhed med 4:2:0 eller 4:1:1 sampling vil der ske en forringelse af farveopløsningen og dermed en forringet mulighed for at opnå en god chroma key. Samtidig begrænses muligheden for multigenerations bearbejdning af materialet.

Sammensætning af udstyr med forskellig signalbehandling



Hvorfor anvende digital teknik

Med indførelsen af digital video og audio gik drømmen om den tabsfrie kopiering i opfyldelse. Uanset hvor mange generationer man ønsker vil kopien altid være som originalen. Kameraets justeringer bliver mere præcise og mere stabile. Signal behandlingen er som i en computer - rent regnearbejde.

Digitale videosignaler er generelt mindre følsomme over for brum-, gain- og fasefejl, overhøring og alle de andre ting man er plaget af i den analoge verden.

Parallel Digital Video

I CCIR-601 standarden for component digitale videosignaler der blev fastsat i midten af 1980'erne, er der kun specificeret sample og kvantiserings frekvenser, men ingen metoder til at overføre signalerne mellem forskellige apparater. SMPTE 125M standarden blev fastsat udfra CCIR-601 normen. Standarden specificerer en måde at overføre det digitale videosignal som et parallelsignal, hvor der anvendes et separat sæt af kabler for de enkelte komponenter i det digitale signal. For at kunne overføre alle signalerne skulle der anvendes et 25-bens stik, også kaldet 25-pin D sub connector.

Det blev dog hurtigt klart at brugen af 25-leder kabler og tilhørende stik ikke var en farbar vej, når der skulle opbygges store komplekse TV-studier.

Der var brug for en anden metode til overførelse af de digitale signaler. Ønsket var at anvende et almindeligt coax kabel med BNC stik.

Seriel Digital Video

Seriel Digital Video var svaret. Selvom det er store datamængder der skal overføres, ligger det indenfor mulighederne i et standard coax kabel.

Overførelshastigheden for et component digital video signal er: 270 Mbit/sek. (Megabit pr. sekund). Tallet fremkommer som summen af de tre digitaliserede component signaler: $(13,5 \text{ MHz} + 6,75 \text{ MHz} + 6,75 \text{ MHz}) \times 10$ (kvantiseringen) = 270Mbit/s.

Digital Audio

Digital lyd er blevet en del af vores hverdag, den anvendes i CD og DAT afspillere. Når man er gået over til at anvende digital lyd er det udfra de samme argumenter som for digital video.

De to organisationer AES (Audio Engineering Society i USA) og EBU (European Broadcast Union) har udarbejdet fælles specifikationer for digital audio.

Som hovedregel anvendes der to sampling frekvenser 44,1 kHz der primært anvendes i CD afspiller og 48 kHz der hovedsagelig anvendes til professionelt brug.

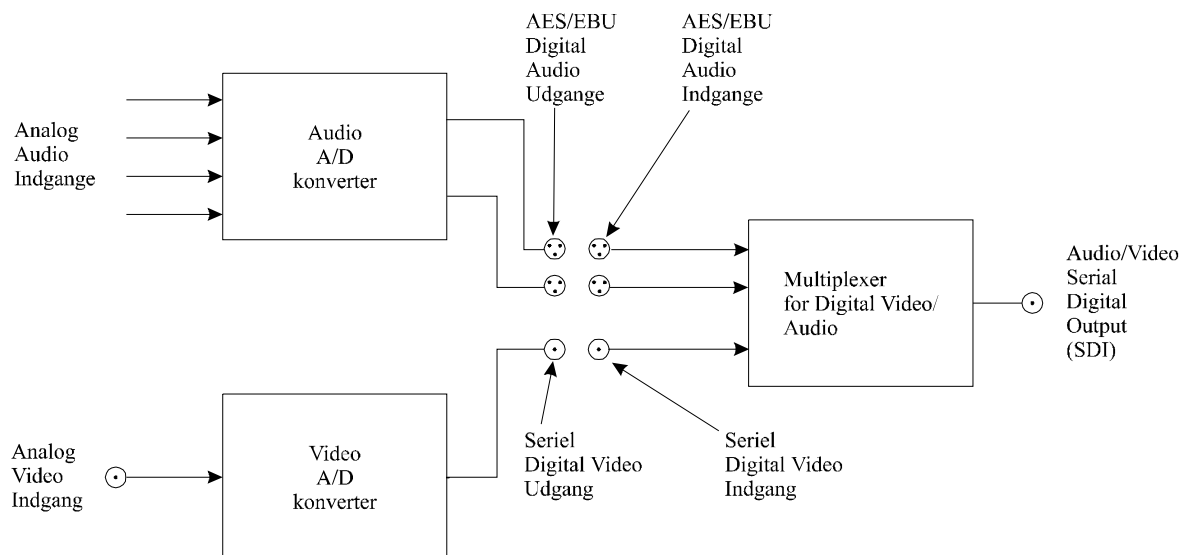
Ligesom i video er opløsningen vigtig. I digital audio anvendes der kvantiseringer der spænder fra 16-bit til 24-bit. I CD afspiller bruges der 16-bit, og til professionelt brug anvendes der enten 18-bit eller 20-bit. I Sony's Digitale Betacam maskiner anvendes der 48kHz sampling og 20-bit kvantisering.

AES/EBU digital lyd kan ligesom digital video overføres som én bit strøm via et enkelt kabel. Det er derfor muligt via et multiplexer kredsløb at kombinere digital lyd og digital video. I et standard SDI signal (Serial Digital Interface) er der foruden det digitale video-signal 4 digitale audiokanaler.

SDI Signalet

Det kombinerede digitale video/audiosignal - SDI - kan fordeles på samme måde som et almindeligt analogt composite videosignal, der skal dog anvendes specielle krydsfelter og fordelere til et SDI signal.

Den store fordel ved SDI signalet er at signalet både indeholder video og fire kanaler

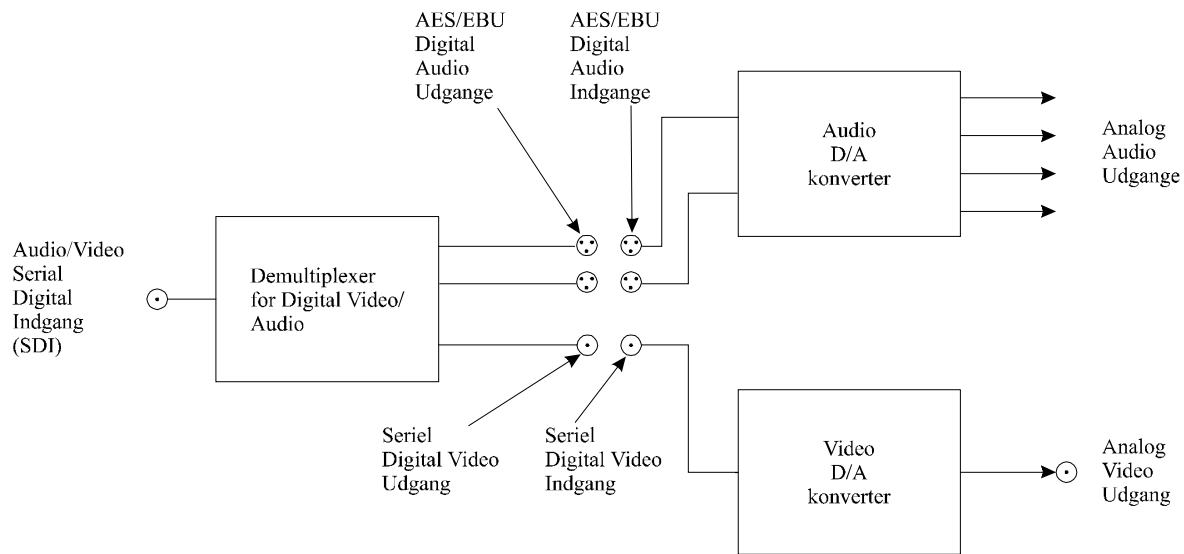


Typisk multiplexer system

(embedded) lyd. Der hvor man har brug for at behandle lyd hver for sig, indsættes en demultiplexer der skiller lyden fra videoen.

SDI signalet kan distribueres via et standard 75 ohm videokabel. Afhængig af kabeltypen kan signalet transporteres ca. 200 m.

Princippet er det samme for et component video system. SDI signalet fra Sony's digitale Betacam maskiner indeholder også tidskode signalet.



Typisk demultiplexer system

Tekniske udtryk

Der findes en lang række måder at redigere video på. Hver måde har fået sin egen terminologi, der nøjagtigt beskriver, hvad det er der sker. For at bringe orden i forvirringen vil vi her gennemgå de enkelte udtryk og deres indhold.

Online redigering

Denne redigeringsform er den mest kostbare og samtidig den mest ufleksible måde at redigere på. I online redigering redigeres masterbåndene uden nogen mellemliggende forberedelser, altså en direkte færdiggørelse af den endelige master. Det er en tidskrævende proces, der stiller store krav til både redigeringsteknikeren og til hardwaren .

Offline redigering

Offline redigering betragtes som et forstadium til online redigeringen. I denne redigeringsform bliver kildematerialet overspillet til et semi-professionelt format, som f.eks. VHS, og forredigeret her.

Ud fra forredigeringen bliver der udarbejdet en "Edit Decision List" (EDL) der senere danner grundlag for redigeringen af den endelige master i online redigeringen.

Lineær redigering

I en lineær redigering hvor der både anvendes bånd og videobåndmaskiner har redigeringsteknikeren ikke direkte adgang til de enkelte scener på båndet. For at finde en bestemt scene skal der spoles frem og tilbage på båndet. Da optagelserne ligger på båndet i den rækkefølge de blev optaget, giver denne form for redigering et stort slid på redigeringsudstyret.

Nonlineær redigering

Nonlineær redigering på en PC eliminerer alle disse problemer. Med kildematerialet på digital form kan dette behandles som alle andre typer datafiler. For at få indspillet kildematerialet på en hard disk er det nødvendigt at bringe video/audio-signalet på digital form. Men ikke nok med det, man er også nødt til at komprimere signalet, så det fylder mindre. Når først materialet er lagt ned på harddisken, er nonlinear redigering langt den hurtigste redigeringsform. En af ulemperne ved harddisk redigering er, at der er begrænset lagerplads på disken, og at denne lagerplads er forholdsvis dyr. I mange tilfælde opvejes prisen på harddisk plads dog af redigeringshastigheden.

Hybrid løsninger

Her udnytter man fordelene fra den lineære og den nonlinear redigering, idet der anvendes en kombination af videobåndmaskiner og harddisk, begge dele styret fra den samme PC.

PC'en som produktions værktøj

Her anvendes PC som en fællesbetegnelse for både PC og Mac platformen. PC'en kan anvendes på to måder. Som en ren styreenhed for forskelligt periferiudstyr, videobåndmaskiner og mixere, eller som et integreret system, hvor der indbygges et kort i computeren med alt hvad der skal anvendes af video- og lydmixer, DVE osv. I begge tilfælde udnyttes PC'ens formidable evne til at udføre kommandoer hurtigt og præcist.

Den rene PC løsning

Den enkleste form for redigering der kan udføres med en computer, er den hvor PC'en udelukkende anvendes til at styre eksterne enheder med et program via en kommunikationsport. Editman er et godt eksempel. Her har man lavet et program - DOS baseret - der kan styre og kontrollere en ekstern boks, hvortil de enkelte enheder der skal styres, tilsluttes. Denne type styring har den fordel, at der ikke skal gøres indgreb i computeren. En hvilken som helst computer kan i princippet anvendes.

Hybrid løsningen

En anden løsning er den, hvor der indbygges et kort i PC'en. Her får man dels en maskinstyring og dels en video/audiomixer med alle de effekter det nu er muligt at lægge på et enkelt kort. Her er Fast Video Machine et godt eksempel. Fast Video Machine er modulopbygget, således at man kan starte med et enkelt kort der indeholder video/audiomixer og styring til tre maskiner. Anlægget kan senere udbygges til en hybrid løsning med harddisk recording og videomaskiner.

Den rene harddisk løsning

Endelig er der de rene non-lineære redigeringsløsninger. Non-lineær redigering eller harddisk redigering blev oprindeligt udviklet til filmindustrien i Hollywood. Hele ideen gik ud på at overspille filmmaterialet til en harddisk i stærk komprimeret form. Her var det ikke billedkvaliteten der var det væsentlige, men derimod hastigheden.

I stedet for at sidde og klippe filmstrimler sammen kunne man nu meget hurtigere flytte rundt på klip. Når produceren var færdig med harddisk redigeringen af filmen, kunne han blot overlade en klippeliste med tidskode til filmklipperen, der herefter udførte det fysiske klippearbejde.

Denne teknik er senere blevet udviklet til også at omfatte forredigering af videofilm. Det der i dag går under betegnelsen non-lineær videoredigering.

I stort set alle de anlæg der er på markedet, kan brugeren selv vælge graden af komprimering. Det betyder, at hvis der kun er brug for VHS kvalitet vælger man en høj komprimeringsgrad, eksempelvis 1:20. Skal programmet bruges til Broadcast anvendes en komprimeringsgrad på 1:3. Graden af komprimering der kan anvendes til de forskellige formål, er helt afhængig af de enkelte fabrikanters komprimerings algoritmer - den matematiske formel der anvendes til at regne komprimeringen ud.

Til TV-produktion anvendes der to former for komprimering. JPEG der har fået sit navn efter "The joint photographic experts group" der udviklede denne komprimeringsform. Den anden form, MPEG, har fået sit navn efter "The moving pictures experts group".

Komprimering

Et fuldt digitaliseret videosignal fylder 270 Mb/s. Da der er et udbredt ønske om at kunne arbejde med component digitale signaler, lige fra optagesiden til det færdig redigerede materiale, er det af økonomiske og tekniske årsager nødvendigt, at reducerer mængden af data til et niveau, hvor signalet kan behandles og optages, uden at det går udover billedkvaliteten.

Der findes to typer af komprimering.

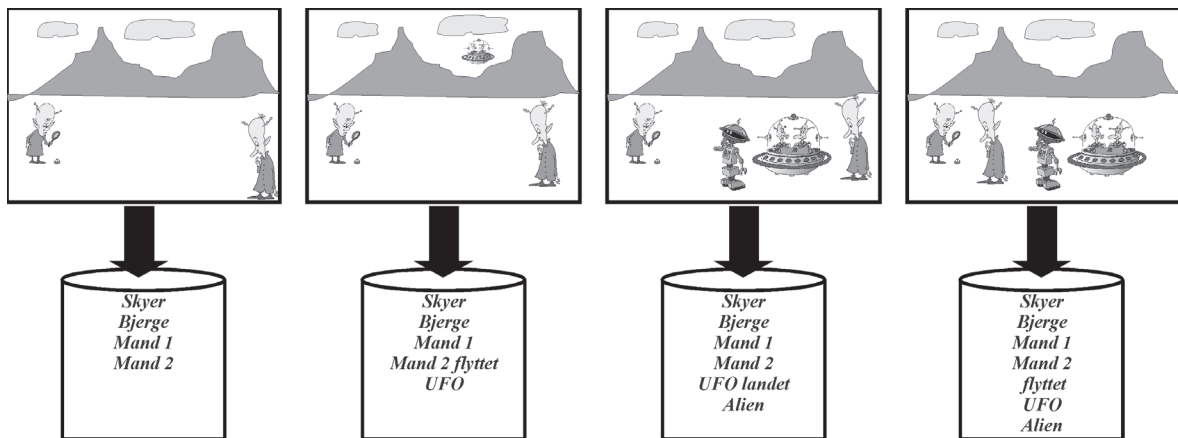
Intra-frame komprimering

Inter-frame komprimering

Intra-frame komprimering (JPEG)

I denne komprimeringsform behandles hvert billede for sig, uden hensyn til de billeder der ligger før og efter. I Sekvens 1 kan vi se, hvad der sker når hvert billede bliver komprimeret for sig. Spandene under hvert billede symboliserer datamængden for hvert billede. Som det fremgår af tegningen, bliver spandene mere og mere fyldte, datamængden der skal behandles stiger. Datamængden kan blive så stor, at der går data tabt. Nu kan man løse det problem på flere måder, øge systemets kapacitet - gøre spandene større - eller ændre datamængden i billedet. En af måderne at gøre det på er at bruge 4:2:0 eller 4:1:1 sampling på kildematerialet, inden komprimeringen.

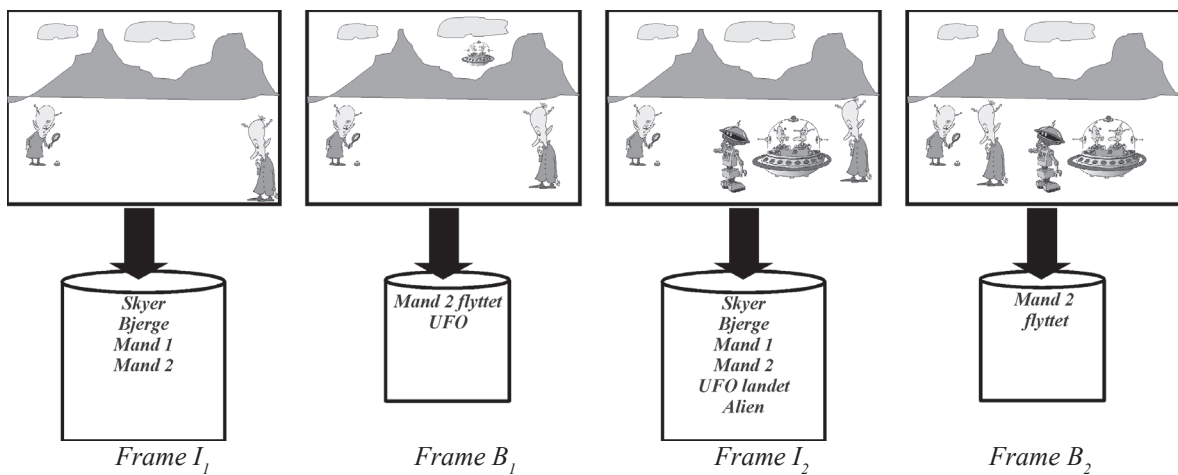
Sekvens 1



Inter-frame komprimering (MPEG2 4:2:2P@ML)

Inter-frame komprimering gør brug af det faktum at gennemsnits indholdet i en video-sekvens, er det samme. En type inter-frame kompression der gør brug af dette forhold kaldes I/B-frame coding.

Sekvens 2



Frame I₁ komprimeres på samme måde som en Intra-frame - med en fuld beskrivelse af indholdet. Det næste billede i sekvensen bygger på indholdet af denne I-frame, men beskriver kun hvad der er forandret i billedet. Denne frame kaldes en B-frame. Den næste I₂ frame beskriver igen hele indholdet af billedet, mens B₂ kun beskriver hvad der er ændret i billedet. Med denne form for komprimering har man reduceret den komprimerede data-mængde betragteligt. Det betyder at man kan arbejde med et fuldt 4:2:2 samplet video-signal og samtidig opnå en høj komprimeringsgrad.

JPEG komprimering

JPEG komprimerer det enkelte billede. Denne komprimeringsform tager udgangspunkt i det faktum, at et TV billede består af en række punkter hvoraf nogle indeholder samme information. Men lad os først se på hvordan et TV billede er opbygget.

En fuld PAL frame består af ca. $575 \times 767 = 441.025$ aktive pixels. Hver pixel bliver kodet i en 4:2:2 YUV teknik med 8 bit til Y-signalet og hver 4 bit til (R-Y) og (B-Y) signalerne, tilsammen 16 bit = 2 byte.

Der skal derfor anvendes ca. $441.025 \times 2 = 0.882050$ Mbyte pr. frame eller 22.05125 MB/sek. ved 25 billeder i sekundet. Det betyder at der på en 1 Gigabyte harddisk kan være ca. 49 sekunder ukomprimeret video. Omregner man prisen for 1 Gigabyte harddisk til videobånd, kan man få adskillige timer bånd for de samme penge. For at få en bedre udnyttelse af harddisken er det derfor nødvendigt at komprimere videosignalet.

Som vi så før består et TV billede af en lang række pixels. Ser man på de enkelte linier, vil der være en stor del af disse, der indeholder samme information. Den eneste forskel er deres indbyrdes placering. Hele øvelsen går nu ud på, kun at gemme informationen for en enkelt pixel. For de pixels der har samme indhold, gemmer man kun deres adresse. På den måde kan informationen i et videobillede reduceres betydeligt uden nævneværdig kvalitetsforringelse.

MPEG komprimering

Denne komprimeringsform tager både udgangspunkt i det enkelte billede og i de efterfølgende billeder. Her udnyttes det faktum at forskellen mellem to på hinanden følgende billeder ikke er ret stor. I princippet gemmer man kun den information, der fortæller, hvor billedet har ændret sig. Med denne komprimeringsform kan der opnås meget store komprimeringsgrader. MPEG anvendes ikke til nonlinear redigering, da klipning i materialet vil kunne forårsage en afbrydelse af forbindelsen til det oprindelige billede. MPEG vil derfor kun blive anvendt til Broadcast, satellit transmission og kabel TV. For at give en idé om hvor meget TV signalet kan komprimeres, regner man med at kunne overføre 16 Digitale TV-kanaler på det samme som én analog TV-kanal fylder i dag.

Fast eller Dynamisk

I de fleste anlæg vælges et fast komprimeringsforhold, der passer til det kvalitetsniveau der ønskes. Der er naturligvis ingen grund til at vælge et komprimerings forhold på 1:3, hvis det er et VHS signal der skal redigeres. Her vil man typisk vælge et forhold på 1:20 eller mere. I nogle anlæg skal der slet ikke vælges komprimeringsforhold, her vælger maskinen selv det bedste komprimerings forhold ud fra, hvor detaljeret billedet er.

Effekter

Alle non-lineære redigeringsanlæg indeholder video- og audiomixere samt mere eller mindre avancerede Digitale Video Effekt generatorer. Man skal være opmærksom på, at ikke alle anlæg udfører effekter i real time, men kun simuleret. Det betyder at maskinen skal have tid til efterfølgende at beregne de valgte effekter.

JPEG OG MPEG komprimering

Fremtidens professionelle videoproducenter må se i øjnene at 99% af de billeder, de så møjesommeligt har komponeret på det bedst tænkelige udstyr, bliver smidt væk. Kompression af video og lyd giver så mange muligheder for både kommercielle TV stationer og producenter af CD plader, at det langt opvejer de ulemper, der er forbundet med kompressionen. Det bliver nødvendigt for fremtidens TV-producenter at sætte sig ind i den nye tekniks - fordele og ulemper.

Enhver professionel videoproducent ved, at det ikke hensigtsmæssigt at bruge en mættet blå tekst på en rød baggrund når signalet skal udsende over en almindelig PAL TV sendestation. Ligeledes ved producenterne at underbelyste, mørke scener vil forsvinde i støj inden de når forbrugeren. Vi ved af erfaring, hvordan vi får det bedste ud af de 625 linier vi har fået i PAL systemet.

Med indførelsen af komprimeret video opstår der en række nye problemer, producenterne skal tage højde for, inden de går i gang med en produktion. Producenterne skal være klar over, hvilke mærkelige og uventede effekter der kan opstå når, den endelig produktion skal afspilles på en CD eller udsendes i komprimeret form. Med andre ord, der skal opbygges et helt nyt sæt af erfaringer, afhængig af hvordan produktet skal bruges - CD eller Broadcast, måske et helt tredje medie.

Nu kan man med rette spørge, hvorfor så anvende komprimering. Jo, for TV stationerne er det indlysende. I stedet for én kanal får de med komprimering plads til fire kanaler. CD producenterne får plads til en hel spillefilm på en enkelt CD, så det økonomiske aspekt vejer tungt.

Fordelene ved at arbejde med digitale formater er mange. En af de mest indlysende er at video/audio på digitalt format er meget hurtigere at arbejde med i forhold til traditionelle tapes. Der er stor set ingen søgetider, filer kan hentes, flyttes og kopieres fuldkommen som i et tekstbehandlingssystem. Tidsforbruget i en redigerings-suite kan reduceres med op til 40%, når der arbejdes med non-lineær redigering.

En af de helt store vanskeligheder ved digital video/audio er at det ukomprimeret fylder enormt. Et sekund ukomprimeret levende video fylder 27Mbyte på en harddisk. Et klip på et minut vil kræve 1.6Gbyte harddisk plads. Filer af denne størrelse vil sætte computerne på en hård prøve.

Derfor komprimering. Der hersker en del misforståelser omkring de to komprimeringsformer JPEG og MPEG. Misforståelser der bør ryddes af vejen.

MPEG Del I

MPEG er den internationale standard for digital videokomprimering fastsat af "The Moving Picture Expert Group".

Under MPEG paraplyen findes der to komprimeringsformer MPEG1 og MPEG2.

MPEG1 dækker kodningen af signaler op til 4 mega bit pr. sekund, anvendes hovedsaglig til

CD indspilninger i forbindelse med multimedia programmer.

MPEG2 dækker kodningen mellem 4 og 8 mega bit pr. sekund, anvendes til Broadcast via satellit.

For at få en forståelse af MPEG kodningen er det lettere at se på JPEG kodningen først, da MPEG bruger en stor del af de samme metoder som i JPEG komprimering.

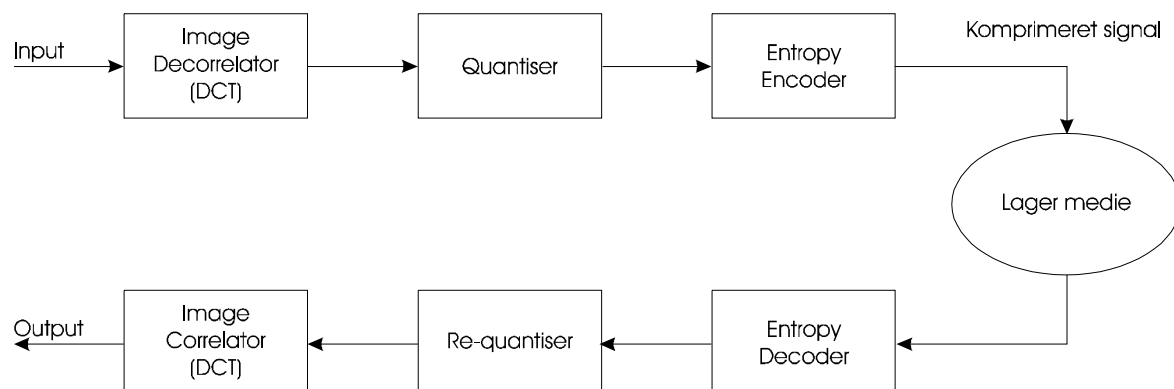
JPEG

JPEG komprimeringen arbejder med component signaler. JPEG og MPEG arbejder begge ud fra den betragtning at en stor del af de data der er indeholdt i et billede er gentagelser. Komprimeringsalgoritmerne er konstrueret så de minimerer gentagelserne, uden at det går for meget udover billedkvaliteten.

I første del af komprimeringsprocessen fjernes halvdelen af farveopløsningen både vertikalt og horisontalt. Her udnytter man det faktum, at menneskets farveopfattelse er forholdsvis ringe. Herefter opdeles de tre component signaler i blokke på 8x8 pixels, gennemsnitsværdien af de 64 pixels i hver 8x8 blok, bliver omsat til lavfrekvente energiniveauer. Resultatet af denne konvertering bliver nu kodet ved hjælp af en bestemt algoritme, således at gentagne mønstre bliver overført med få bits medens sjældne mønstre bliver overført med mange bits.

De fleste komprimeringsalgoritmer er koncentreret omkring den lavfrekvente del af video-signalet, netop fordi det er her den meste energi er samlet. Det er også i dette område vores øjne er mest følsomme.

JPEG blev konstrueret til at behandle stillbilleder. Afspilning af sekvenser af stillbilleder giver bevægelse, deraf navnet Motion JPEG. JPEG anvendes til alle former for redigering og bearbejdning af videosignaler, netop på grund af at det er komprimering af enkelt billeder.

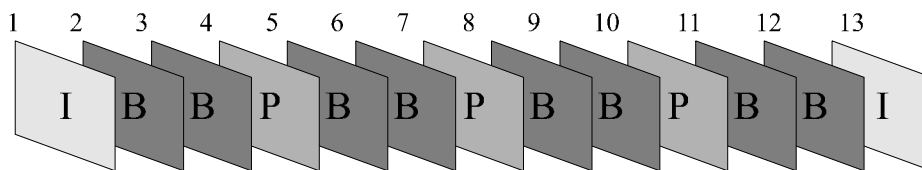


Typisk JPEG koder- og dekoderkredsløb.

MPEG Del II

MPEG arbejder ud fra at et videosignal består af en bevægelig forgrund og en næsten statisk baggrund. Det er derfor kun nødvendigt at overføre ændringerne af forgrunds-billedet. MPEG er mest en algoritme til mønster genkendelse. MPEG video består af tre typer billeder:

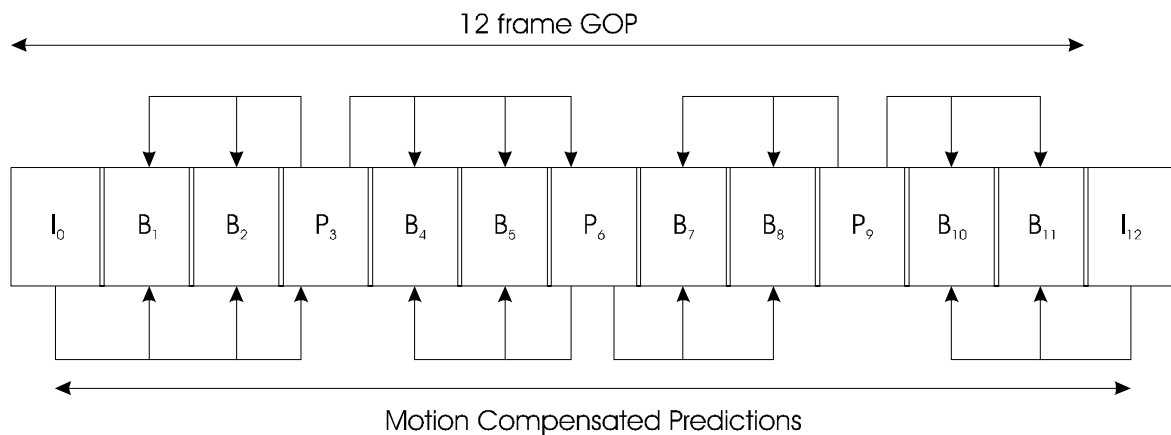
MPEG Kodning.



I-billeder (Intraframes) der som JPEG er komprimeret i sig selv.

P-billeder (Predicted Frames) billeder der er bevægelseskomprimeret i forhold til det foregående billed.

B-billeder (Bi-directional frames) der er komprimeret både i forhold til det foregående og til det efterfølgende forventede billede.



Groups of Pictures

Med de tre billedtyper I, P og B kan der dannes en sekvens af billeder, der passer til det formål de skal anvendes. Det er vigtigt i opbygningen af sekvenser, at der med passende mellemrum bliver indsat et "ankerbillede" en I-frame. En billedsekvens der går fra en I-frame til det billede der kommer umiddelbart før den næste I-frame, kaldes en "GOP" - Group of Pictures.

En 12 frame sekvens som den viste, har en længde på ca. 0,5 sekunder, det giver en tilpas kort koder/dekoder forsinkelse, uden at det går udover komprimeringen.

Når man i MPEG kan kassere 99% af det originale billedmateriale, er det fordi, der er fuld kontrol med, hvad der kasseres og hvad der bibeholdes. En af de vigtigste grundregler ved komprimering er, at slut resultatet er afhængig af master materialet. Kommer man "skidt ind" kommer der "skidt ud".

Et af de helt store problemer ved MPEG er den begrænsede båndbredde. Scener med bevægelse og mange detaljer kan få MPEG enkoderen til at bryde sammen, fordi der ikke er bits nok til rådighed.

Sony "SX" komprimering

Den nye generation af Sony Betacam hedder Betacam SX. Her anvendes der en komprimeringsform, der ligger under MPEG-2 paraplyen, men med en ny standard der kaldes 422P@ML. I den nye standard anvendes der kun I- og B-frames i en 2-frame sekvens.

Tekniske specifikationer for SX komprimering:

- komprimeringsforhold 1:10
- en bitrate på 18Mbit/sekund for video, 20 - 21 Mbit/s med lyd.
- fuld broadcast kvalitet der er væsentligt bedre end Betacam SP.
- nøjagtig fuld frame editing
- multi generationer med effekter.
- mulighed for chroma key.

En MPEG komprimering med fuld båndbredde, editerings faciliteter og multi - generations muligheder.

De Digitale formater

De nye digitale videoformater DV, DVCAM og DVCPRO er kommet eller er lige på trapperne. Hvordan vælges det rigtige format? Hvilke parametre skal der lægges vægt på? Hvilken betydning får valg af format for de fremtidige muligheder? Det er nogle af de problemstillinger vi vil prøve at behandle her.

Mere end 50 producenter af videoudstyr er blevet enige om et fælles digitalt videobånd-format DV, beregnet til den avancerede forbruger af hjemmevideoudstyr. Formatet er baseret på 5:1 digital komprimering af et 4:2:0 digitalt videosignal. Men hvad er et 4:2:0 digitalt videosignal?

Allerede tilbage i 1981 vedtog EBU normen for et broadcast digitalt videosignal, den vi idag kender som CCIR-601 standarden. Normen siger at den sort/hvide komponent af billedet skal samples med 13,5MHz og de to farve komponenter (R-Y) og (B-Y) hver med 6,75MHz. I daglig tale kaldes denne norm CCIR-601 4:2:2.

Sampling

Datastrømmen i et 4:2:2 digitaliseret videosignal er ca. 220 Mbit/sek. hvilket er en høj bitstrøm, der kræver stor processor kapacitet og stor lagerplads. Hvordan kan man nu reducere denne store bitstrøm uden at det går nævneværdigt udover billedkvaliteten?

Da det menneskelige øje ikke er særligt følsomt over for ændringer i farveopløsning, var der muligheder i de to farvekomponenter (R-Y) og (B-Y). I DV formatet har man valgt at reducerer den vertikale farveopløsning til det halve, altså kun sample hver anden linie med 4:2:2, med 4:2:0 som resultat.

Denne opløsning kendes fra PAL systemet - de daglige TV udsendelser. En anden mulighed var at reducere den horisontale farveopløsning til det halve, altså 4:1:1 - en teknik der anvendes i en del digitale effektenheder.

Den reducerede farveopløsning har især betydning ved Chroma-key effekter. Men det får nogle konsekvenser, hvis man blander de to digitale samplingsmetoder.

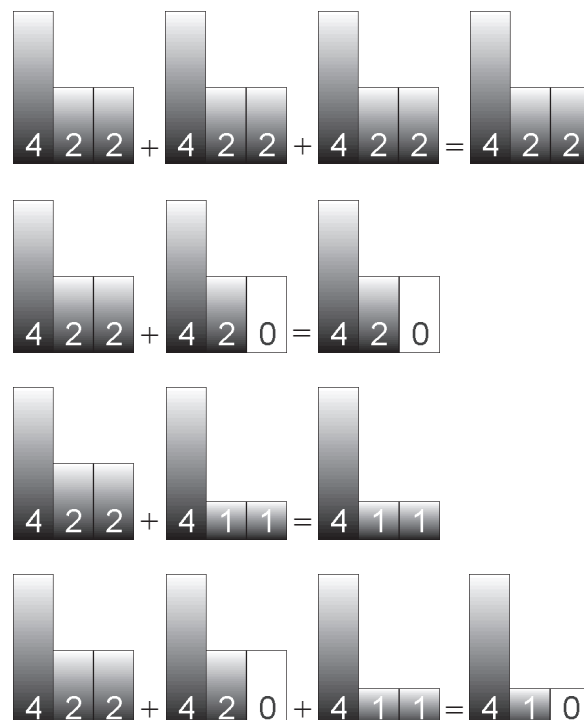


Fig. 1

I figur 1 kan man se resultatet af en sammenblanding af udstyr med forskellig signalbehandlingsmetoder. Som det fremgår af figuren, er det den laveste fællesnævner, der danner grundlag for det endelige resultat. Af hensyn til 16:9 formatet har man i consumer DV formatet valgt at bevarer den horisontale opløsning - 4:2:0.

Med udgangspunkt i DV formatet har Panasonic og Sony udviklet hver deres professionelle DV format - DVCAM og DVCPRO - det er signalbehandlingen i de to formater vi vil se lidt nærmere på.

DVCAM

Sony har i deres professionelle DVCAM system valgt den samme sampling metode som i DV formatet - 4:2:0 - halv vertikal farveopløsning. De har valgt samme båndtype, dog med en større sporbredde - 15 mikron. Det betyder at der er fuld kompatibilitet med consumerformatet DV, der som sagt også arbejder med 4:2:0 sampling.

Kompatibiliteten gælder begge veje dvs. at et DVCAM bånd kan afspilles i en consumer båndmaskine, dog ikke i en consumer camcorder og et DV bånd kan umiddelbart afspilles i en DVCAM maskine. Der kan optages på et DV bånd i en DVCAM maskine.

For at få spoletallerkenerne til at passe til de små DV kassetter har man valgt at bruge samme metode som i Betacam maskinerne, hvor spoletallerkenerne automatisk tilpasser sig kassettebåndstørrelsen.

DVCPRO

Panasonic har i deres professionelle DVCPRO system valgt 4:1:1 - halv horisontal farveopløsning - som samplingsmetode. De har valgt at anvende et Metal-partikelbånd og en sporbredde på 18 mikron.

Som det fremgår af figur 2, vil afspilning af et DV consumer bånd i en DVCPRO maskine resultere i en

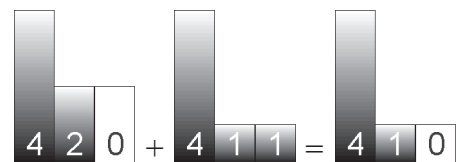


Fig. 2

reduceret farveopløsning - 4:1:0 - altså halv vertikal og halv horisontal farveopløsning.

For at kunne afspille et DV bånd i en DVCPRO maskine kræves en kassetteadaptor, som det kendes fra VHS-C formatet. Der kan ikke optages på et DV bånd i en DVCPRO maskine, heller ikke med en kassetteadaptor.

Panasonic betragter DVCPRO som et selvstændigt broadcast format.

Audio Lock Mode

Consumer DV bruger PCM (Pulse Code modulation) til lydoptagelse i CD kvalitet. Der kan vælges mellem to optage metoder: 16-bit, 48kHz sampling, der giver 2 lydkanaler i CD kvalitet, eller 12-bit, 32kHz sampling, der giver 4 lydkanaler. Alle kanalerne kan bearbejdes separat. I "Lock mode" er lyden låst til videoen hvilket betyder at der ved 48kHz er 1920 sampling punkter pr. frame (48000/25) og ved 32kHz 1280 samplingpunkter pr. frame (32000/25).

I "Unlock mode" er lyden ikke låst til videoen, antallet af samplingpunkter varierer i 48kHz/1920 med +/- 24 samplingpunkter pr. frame og i 32kHz/1280 med +/- 16 samplingpunkter pr. frame.

Kopiering af et bånd med unlocked audio i 4 x hastighed er derfor ikke mulig. Skal et DV bånd med unlocked audio redigeres på et professionelt non-lineært redigeringsanlæg, skal materialet overføres første gang i 1:1 hastighed.

Ved overspilningen bliver audioen låst. Herefter kan signalet overspilles frem og tilbage mellem harddisk og båndmaskine i 4 x hastighed uden problemer.

DVCAM har 4 lydkanaler som i DV standarden, men DVCAM arbejder altid med lyden låst til videoen - "locked audio".

I DVCPRO er der mulighed for 2 lydkanaler med 16-bit, 48kHz sampling. DVCPRO har udover de to digitale lydspor et analogt CUE spor.

Konklusion

Videokomprimering er nok en af væsentligste udviklinger indenfor videoindustrien, siden indførelsen af farve-TV i begyndelsen af 1950'erne. Der nok ingen tvivl om at vi også i fremtiden vil se nye og forbedrede former for videokomprimering - en udvikling - der vil smitte af på hjemmeelektronikken.

Der findes andre komprimeringsformer end dem jeg har beskæftiget mig med her. Wavelet og Fractal komprimering. Wavelet komprimering anvendes i ImMix TurboCube non-lineære - on-line redigeringsystem. Jeg har ikke på nuværende tidspunkt kendskab til broadcast anlæg der anvender Fractal komprimering, men jeg ved at der arbejdes ihærdigt på udvikling og forbedring af denne komprimeringsform.

JPEG anvendes i non-lineære redigeringsanlæg, fordi det er enkeltbilled komprimering. MPEG anvendes til CD produktion og Broadcast, fordi det fylder mindre og der skal ikke klippes i materialet.