

## AV

### Hoofdstuk 1: Achtergrond

#### Digitale voorstelling:

Een bit is een data-eenheid die 1 van de 2 waarden kan aannemen.

Groepen van bits kunnen geïnterpreteerd worden als getallen met basis 2, maar kunnen ook behandeld worden als karakters, kleur, etc...

#### **1. Digitalisatie:**

Analoge data moet geconverteerd worden in digitale vorm vooraleer het gemanipuleerd kan worden door een computerprogramma

Digitaal: geschikt voor precieze berekeningen

Analoog: gedetailleerder

Digitalisatie: proces waarbij een analoog signaal geconverteerd wordt naar een digitaal signaal.

Dit proces bestaat uit 2 stappen : Bemonstering (=sampling)

Kwantificering

Bemonstering: waarde van het signaal meten op discrete tijdsintervallen

Kwantificering: waarde beperken tot een vaste verzameling van niveau's

Deze processen worden meestal uitgevoerd door speciale hardware apparaten= ADC's= analoog-naar-digitaal-converters.

Bemonsteringssnelheid: Het aantal samples in een vaste tijdshoeveelheid of ruimtehoeveelheid.

Kwantificeringsniveau's: de niveaus waarop een signaal gekwantificeerd wordt

Voordelen digitale voorstellingen:

- bepaalde signaalwaarden zijn geldig
- elke fluctuatie die veroorzaakt werd door ruis zal getransformeerd worden van een geldige in een ongeldige waarde.
- Robuuster dan analoog, want geen verlies wanneer ze gekopieerd worden of overgebracht worden naar media onderhevig aan ruis.

Undersampling: wanneer er niet genoeg samples worden genomen, spreken we van undersampling. Het signaal zal vervormen, details zullen verloren gaan (bemonsteringssnelheid is te laag)

Signalen in tijd en ruimte bestaan uit pure sinusgolfcomponenten van verschillende frequenties. Daarbij kan elke golfvorm opgedeeld worden in frequentiecomponenten.

Hoe meer frequentiecomponenten we toevoegen, hoe beter de benadering.

De Fouriertransformatie kan gebruikt worden om de voorstelling van het signaal in zijn frequentiedomein te berekenen.

Hogere frequentiecomponenten worden geassocieerd met abrupte overgangen. De stijgende en vallende bogen van de golfvorm worden bijna verticaal.

\*bij functies zoals 'sharpen' en 'smooth' worden bepaalde frequenties verwijderd.

Bestaat er een bemonsteringssnelheid die een accurate reconstructie van het signaal garandeert ?

⇒ Samplingstelling

= Deze stelling stelt als de hoogste frequentiecomponent van een signaal  $F_h$  is, kan het signaal juist gereconstrueerd worden als het bemonsterd wordt op een frequentie groter dan de Nyquist-snelheid  $2 F_h$ .

⇒ Nyquistsnelheid: limietwaarde

Als we bemonsteren met een snelheid dat lager is dan de Nyquistsnelheid dan zullen bepaalde frequentiecomponenten in het originele signaal omgezet worden in andere frequenties wanneer het signaal gereconstrueerd wordt. (=Undersampling)

Undersampling leidt tot ALIASING.

⇒ bij geluid: distortion

⇒ bij beelden: waarnemen in de vorm van gekartelde lijnen

⇒ bij animatie: schokkende beelden

Het effect van onvoldoende aantal kwantificeringsniveaus:

⇒ posterisatie / brightness contouring

⇒ Moiré-effecten in beeld

⇒ Kwantificeringsruis in geluid (distortion), duidelijk waarneembaar bij 8 bits

## 2. Compressie

Mediadata heeft enorm veel opslagruimte nodig. Ze hebben veel bandbreedte nodig om ze te verzenden over netwerken.

Oplossing: compressie = reduceren van opslaghoeveelheid

Als data gecomprimeerd werd, zal een inverse decompressie-operatie nodig zijn om data terug te brengen naar een vorm waarin het kan afgebeeld worden.

Codec: software die compressie en decompressie uitvoert.

Er zijn twee soorten compressiealgoritmen:

- **Lossless algoritme:**

Het is altijd mogelijk om data te decomprimeren die gecomprimeerd was en om een exacte kopie te verkrijgen van de originele data.

Exact omkeerbaar, zonder (kwaliteits)verlies

- **Lossy algoritme:**

Bepaalde data werden verwijderd in het compressieproces en kunnen niet hersteld worden, zodat de gedecomprimeerde data enkel een benadering is van het origineel.

Niet-exact omkeerbaar, met verlies

Lossy algoritmen worden veel gebruikt, want ze zijn goed in het bewaren van de kwaliteit van stilstaand beeld, video en geluid, ook al werd een grote hoeveelheid data verwijderd. => wordt gebruikt voor multimediatoepassingen

Lossless algoritmen zijn minder effectief dan lossy algoritmen. => wordt gebruikt voor tekst.

De meeste lossy algoritmen maken gebruik van bepaalde lossless technieken. In het ideale geval wordt lossy compressie enkel toegepast in de laatste mogelijke stap in het klaarmaken van media voor aflevering. Elke bewerking die ervoor moet gebeuren, zal door lossless compressie gebeuren. Waarom?

- als data lossy gecomprimeerd is, dan kan de verloren informatie nooit meer teruggehaald worden, waardoor de data heeltijd gecomprimeerd en gedecomprimeerd moet worden en zo zal de kwaliteit enorm verslechteren.

- Bepaalde bewerkingen kunnen het verlies aan kwaliteit verergeren.

⇒ Best werken met ongecomprimeerde data en enkel comprimeren om uiteindelijk af te leveren.

Verschillende compressiealgoritmen zijn toepasbaar op verschillende types mediadata. Hun effectiviteit hangt af van de karakteristieken van de data zelf.

Bv. Lossy : JPG voor beeldcompressie, MP3 voor audiocompressie...

Lossless: computer gegenereerde beelden met grote vlakken van 1 kleur.

Algemene eigenschap van elk compressieschema:

Er bestaan bepaalde data waarvoor de gecomprimeerde versie groter zal zijn dan de ongecomprimeerde.

### **3. Digitale voorstelling van media:**

Datavoorstellingen in detail bespreken van stilstaand beeld, video, animatie, geluid en tekst. Er zijn vaste manieren om deze voor te stellen in bits.

Mediadata kan voorgesteld worden als een tekstuele beschrijving in een bepaalde taal, of als binaire data met een specifieke structurele vorm.

#### ***Stilstaand beeld:***

Beeld wordt weergegeven als een array van pixels en voorgesteld met een intern model. Om een afbeelding weer te geven op een monitor moet het programma elke pixel zetten

op een gepaste kleur of grijsschakering en er zo voor zorgen dat het patroon van het gewenste beeld produceert.

Resolutie = aantal pixels dat een apparaat kan afbeelden

Rendering= proces om een patroon pixels te genereren uit een model

Twee verschillende benaderingen van grafische modellen

- vectorbeelden
- bitmapbeelden

### 1. Bitmapbeelden

Beeld wordt gemodelleerd door een array van pixelwaarden, die rechtstreeks gekoppeld kunnen worden aan de fysische pixels op het scherm. (logische pixels: opgeslagen waarden)

### 2. Vectorbeelden

Beeld wordt bewaard als wiskundige beschrijving van een verzameling lijnen, curven en vormen die het beeld samen voorstellen. Dit vereist berekening om het te renderen.

Verschillen

<b>Bitmapbeelden</b>	<b>Vectorbeelden</b>
Moet waarde van elke pixel bijhouden	Beschrijving van alle objecten van het beeld, hoe complexer de afb. hoe meer objecten, hoe groter de beschrijving, maar omdat we geen pixels bewaren, is de grootte onafhankelijk van de resolutie.
	Kleiner dan Bitmap
Schaling met kwaliteitsverlies	Schaling zonder kwaliteitsverlies
Distortion of blurring is eenvoudig op een bitmapbeeld	Vectorbeeld moet eerst omgezet worden in bitmapformaat
Bewerken en retoucheren van foto's	Grafieken en diagrammen

### ***Video en animatie:***

Alle methoden om bewegende beelden weer te geven zijn gebaseerd op het fenomeen dat gekend staat als 'nawerking van het oog'. (=nabeelden)

Als een reeks beelden weergegeven wordt aan onze ogen aan een voldoende hoge snelheid(fusiefrequentie), ervaren we een continue visuele gewaarwording in de plaats van losse, individuele beelden.

De fusiefrequentie hangt af van de helderheid van het beeld in vergelijking met de achtergrond, maar is rond 40 beelden per seconde. (=flikkerend effect)

Twee verschillende manieren om beelden te genereren:

- reeks frames worden geproduceerd door de eigenlijke beweging op te nemen wanneer ze zich in de reële wereld voordoet
- illusie van beweging synthetiseren door elk frame individueel aan te maken; bv afzonderlijke beelden na mekaar.

Dus we kunnen bewegende beelden maken door live-actie of door animatie.

Live-actie kan bewaard worden als video

Videoframes zijn bitmapbeelden en vereisen daarom ook veel opslagruimte, dus video moet constant gecomprimeerd worden.

PAL-standaard: frame moet 576 pixels hoog zijn en wordt afgespeeld aan snelheid van 25 frames per seconde.

geen compressie ? clip 10 minuten : 15 GB

⇒ Daarom wordt er altijd compressie toegepast

Animatie kan voorgesteld worden op meer flexibele en efficiënte manieren.

Animatieframes kunnen vectorbeelden zijn. Dit kan de bestanden kleiner maken en is goed geschikt voor cartoonachtige tekeningen.

Interpolatie: objecten tekenen in bepaalde keyframes (animatie 1.21)

### ***Geluid***

Wordt veroorzaakt door snelle veranderingen in de luchtdruk

Kan voorgesteld worden als een reeks samples na digitalisatie, als de bemonsteringssnelheid hoog genoeg was.

CD audio : 44.100 samples per seconde (=44,1 kHz)

Audio, die verspreid is op internet, is gecomprimeerd (MP3)

### ***Tekst***

Nood aan karakterset

= koppeling van karakters aan karaktercodes

ASCII-karakterset

UNICODE-karkterset: standaard voor alle digitale tekst

➔ in staat om tekst voor te stellen in alle gekende talen

Lettertype= reeks karaktersvormen, glyphs genaamd.

Plain-tekst bestanden: bestanden die geen lay-out informatie bevatten

wordt gebruikt voor de broncode van programma's

Tekst kan een dubbele rol spelen in mediadata. Het is niet enkel mediadata zelf maar kan eveneens functioneren als een taal voor het beschrijven van de structuur en het voorkomen van de data (vb html, css)

PDF(Portable Document Format)=

- gebruikt voor tekstdocumenten
- exacte positie van glyphs wordt gespecificeerd
- weinig structuur behouden

- ideaal formaat voor boeken en documenten waarvan lay-out bewaard moet blijven
- afleverformaat
- meest geschikt om documenten weer te geven

### ***Interactiviteit***

Interactiviteit wordt geproduceerd door een programma uit te voeren als antwoord op de gebruikersinvoer. In multimedia worden programma's dikwijls geschreven in een scripttaal zoals Javascript of Actionscript.

De cpu in een computer kan geen programma's uitvoeren in hun hoog-niveau vorm. Eerst moet een **compilatie** worden uitgevoerd (=het programma, geschreven door een persoon, moet vertaald worden in een reeks machine-instructies.

Een programma kan ook uitgevoerd worden door **interpretatie** (= een programma leest de bronversie van een ander programma en voert de operaties uit die het specificeert). Specificatie is trager dan compilatie, maar het is wel eenvoudig te implementeren en is bovendien machine-onafhankelijk.

Interpreters worden vaak gebruikt om scripttalen te implementeren=>java en action

Scripts:        tekstbestanden, waarvan de inhoud de syntactische regels volgt die gedefinieerd zijn door de taal waarin het script geschreven is.

### ***Metadata***

= gestructureerde data over data, die gekoppeld kan zijn aan mediabestanden om het zoeken en classificeren te vereenvoudigen.

Bv. Metadata van een foto: datum, tijd, gps-coördinaten, naam fotograaf, belichtingstijd..

Elk van deze info kan opgeslagen worden als een apart metadata-item of veld.

Belangrijk voor mediadata => maakt organisatie en zoeken uitvoerbaar

Verschillende soorten metadataformaten:

- **Exif:** Exchangeable Image File Format)  
Geschikt voor digitale foto's  
Kan uitgebreid metadata opnemen (automatisch)
- **IPTC:** International Press Telecommunication Council)  
Uitwisseling van data tussen nieuwsorganisaties  
Meer begaan met bijdrage en rechten  
Metadata moet manueel toegevoegd worden
- **Dublin Core:**  
Algemene specificatie die 15 soorten metadata-items definieert, die het elementen noemt.  
Elk element is optioneel ( lijst p 50)  
Kan gebruikt worden om elk soort data te beschrijven

## Hoofdstuk 2: Bitmapbeelden

Bitmapbeelden zijn veel eenvoudiger dan vectorbeelden en bovendien bitmapvoorstellingen gebruikt worden om foto's voor te stellen.

### 1.Resolutie

= maat van hoe fijn een apparaat continue beelden benadert met eindige pixels, het is dus nauw gerelateerd aan bemonstering.

#### ***Apparaatresolutie***

Resolutie van scanners en printers: uitgedrukt in pixeldichtheid(dots per inch)

Bij scanners gaat men soms hele hoge scannerresoluties in dots per inch aanhalen door geïnterpoleerde waarden te tellen die niet rechtstreeks gescand zijn.

Daardoor zal je soms de fysieke resolutie van een scanner uitgedrukt zien in eenheden van samples per inch.

Resolutie van videoframes en computermonitoren: uitgedrukt in pixeldimensies(br x h)

De pixeldichtheid zal op alle monitoren hetzelfde zijn, d.w.z dat een beeld ook altijd dezelfde proporties van een scherm zal innemen met dezelfde pixeldimensies.

Resolutie van digitale camera's: uitgedrukt als het aantal pixels in het grootste beeld dat het kan maken.

Hiervoor moet je de verhoudingen van het beeld kennen, nl de breedte en de hoogte. De standaardverhouding is 4:3 voor tv-schermen en monitoren (low-mid-range camera's)

Bij high-range camera's spreken we van een 3:2 verhouding.

#### ***Beeldresolutie***

Bij bitmapbeelden is het beeld een array van pixelwaarden, dus heeft het noodzakelijk pixeldimensies. Het heeft geen afmetingen

De pixeldichtheid van monitoren en camera's is gemiddeld 72dpi(schermresolutie).

#### Algemeen:

Afmeting= pixeldimensie / pixeldichtheid

Pixeldichtheid: de apparaatresolutie (pixels per lengte-eenheid)

⇒ Als 2 beelden dezelfde afmetingen hebben op apparaten met verschillende resoluties, moeten de beelden verschillende pixeldimensies hebben. Door die

verschillende groottes zullen die beelden verschillende hoeveelheden details weergeven.

⇒ Vb. als beeld 72 dpi is en afbeelding is 198px breed en 148px hoog en we willen de afmeting berekenen:

Bij 

$(198 / 72) * 25,4$ 70 mm	en	$(148 / 72) * 25,4$ 52,5 mm
------------------------------	----	--------------------------------

 de lage resolutie is het beeld grof, dit noemt

men een pixellation. Dit is een vorm van aliasing.

Dus

- ⇒ een bitmapbeeld heeft pixeldimensies, maar geen intrinsieke pixeldichtheid.
- ⇒ De fysische grootte van een beeld wanneer het afgebeeld wordt hangt af van de pixeldichtheid van het apparaat waarop afgebeeld

De meeste beeldformaten bewaren een resolutie(pixeldichtheid) samen met de beelddata, dit is gewoonlijk de resolutie van het apparaat van waar het afkomstig is.

Deze resolutie wordt uitgedrukt in eenheden van pixels per inch(ppi)

Als een beeld nu wordt ingescand bij 600 dpi, zal het opgeslagen beeldresolutie 600ppi zijn, omdat de pixels in het beeld waren gegenereerd op deze resolutie en zo worden de afmetingen van het beeld berekend uit de pixeldimensies en de beeldresolutie.

Algemeen: apparaatresolutie/ beeldresolutie

Bv. Foto met afmetingen 6 inch op 4 inch en wordt ingescand op 600 dpo => bitmapbeeld zal bestaan uit 3600 op 2400 pixels.

## 2. Hersampling

Schalen van een bitmapbeeld:

Schalen verandert de pixeldimensies van een beeld, wat betekent dat er bepaalde pixels moeten weggegooid worden als het beeld verkleind wordt, of toegevoegd als het beeld vergroot wordt.

Downsampling:      pixeldimensie reduceren  
Upsampling:        pixeldimensie uitbreiden      }      Hersampling

⇒ Beide kunnen tot kwaliteitsverlies leiden

Schalen kan gebeuren door zowel een transformatie uit te voeren op elke originele pixel of door de inverse transformatie op elke pixel in het geschaalde beeld uit te voeren.

Hiervoor is interpolatie nodig wegens de eindige grootte van pixels. (zie p62 / 63)

Het pixelrooster voor een geschaald beeld zal niet uitlijnen met gehele pixels wanneer het teruggekoppeld is met het origineel.

Verschillende interpolatieschema's:

- Dichtste-buur
- Bilineaire interpolatie



- Bicubische interpolatie

Zie illustratie p 65: We beelden ons in dat we een pixel tekenen, de doelpixel, gecentreerd op  $x, y$  wat het sample zal zijn dat we nodig hebben. Deze pixel gaat 4 pixels van het originele beeld overlappen.

Dichtste buur: fig. 2.5

- Eenvoudigste interpolatieschema
- Gebruikt de waarde van de dichtste pixel(buur)
- Werkt het snelst
- Produceert resultaten met slechte kwaliteit (undersampling)
- Beeld ziet eruit alsof de originele pixels uitvergroot werden door een vergrootglas

Bilineaire interpolatie: fig. 2.6

- Deze techniek gebruikt de waarden van alle 4 de naburige pixels
- Eenvoudige berekening, gebaseerd op veronderstelling dat de waarden lineair veranderen in beide richtingen (**bilineair**) doorheen de regio van de 4 pixels
- Mogelijk dat de waarden discontinuïteiten vertonen

Bicubische interpolatie

- Interpolatie is gebaseerd op cubische splines
- Deze techniek gebruikt 16 pixels uit het originele beeld
- Tussenliggende waarden liggen langs een Bézier-curve die de bewaarde pixels verbindt
- Traag

nadeel interpolatiemethodes:

- ⇒ veroorzaakt afzwakking van scherpe randen
- ⇒ verlies aan informatie die niet opnieuw kan opgehaald worden

### **3. Beeldcompressie**

Beelden kunnen lossless gecomprimeerd worden met verschillende methoden:

RLE: run-length encoding

- Eenvoudigste lossles compressiealgoritme om te begrijpen
- Niet het meest effectieve

Huffman-codering:

Woordenboekgebaseerde compressietechnieken:

- Werkt door aanmaken tabel of woordenboek waarin strings van bytes worden gestoken
- Ruimtebesparing

- LZ77: deflate compressie :
- LZ78
- LZW: 1 van de meest gebruikte technieken, vormt basis van Unix compress utility en van Gif bestanden

### ***JPEG-compressie:***

Joint Photographic Experts Group

Jpeg wordt het meest gebruikt als lossy compressiemethode voor silstaand beeld.

Elke pixel geeft een z-waarde aan zijn x en y coördinaten, zodat het beeld een driedimensionale vorm definieert.

= we krijgen een complexe 3D-vorm

- ⇒ elke golfvorm kan getransformeerd worden in het frequentiedomein met behulp van de Fouriertransformatie
- ⇒ de hoogste frequentiedomeinen worden geassocieerd met abrupte veranderingen in intensiteit
- ⇒ mensen kunnen hoge frequenties niet nauwkeurig ontvangen, zeker niet in gekleurde beelden

Jpeg-compressie werkt echter niet met de Fouriertransformatie maar met de DCT-transformatie

=Discrete Cosinus Transformatie

=men neemt een array van pixels en produceert een array van coëfficiënten die de amplitude van de frequentiecomponenten in het beeld voorstellen.

De discrete cosinustransformatie wordt toegepast op 8 x 8 blokken van pixels.

De DCT toepassen reduceert de grootte van de data niet, omdat de array van frequentiecoëfficiënten van dezelfde grootte is als de originele pixelarray.

We kunnen zeggen dat de verschillende frequenties gekwantificeerd worden op een verschillend aantal niveaus. Bij JPEG-compressie kan het aantal kwantificeringsniveaus die gebruikt worden voor elke frequentie coëfficiënt gespecificeerd worden in een kwantificeringsmatrix die een waarde bevat voor elke coëfficiënt. De kwantificering verwijdert bepaalde informatie.

Kwantificeringsproces :beeld bewaren op 2 manieren:

- Er zullen veel componenten na kwantificering eindigen met nulcoëfficiënten.
- Er zullen minder bits nodig zijn om coëfficiënten die niet nul zijn te bewaren.

De niet-nulcoëfficiënten worden gecomprimeerd met de Huffman-codering. (zig-zag sequentie)

Decompressie van JPEG-data gebeurt door het omkeren van het compressieproces, met de inverse DCT om het beeld terug te halen uit zijn frequentiedomeinvoorstelling. Zo kunnen de waarden opnieuw behandeld worden als pixels in een beeld.

Merk op ! Er is geen 'omgekeerde kwantificerings' stap; eens de informatie verloren is, is het voor altijd verloren.

## Eigenschappen JPEG-compressie

Het is mogelijk om de graad van de compressie te controleren, dus ook de kwaliteit van het gecomprimeerd beeld, door de waarden in de kwantificeringsmatrix te wijzigen.

Het gedecomprimeerde beeld kan compressieartefacten tonen, inclusief blurring en zichtbare randen tussen de 8 x 8 pixelblokken, vooral bij lagere kwaliteitsinstellingen. (zie p 76)

## **4. Bestandsformaten:**

### ***GIF en PNG***

Ad hoc standaarden

GIF:

- Graphics Interchange Format
- Gebruiken LZW-compressie, zodat ze lossless zijn
- Compressie is het efficiëntst bij eenvoudige beelden
- Techniek: geïndexeerde kleur, beperkt tot 256 kleuren => ongeschikt voor gekleurde foto's
- 1 kleur kan aangegeven worden als transparant
- Gif's worden veel gebruikt op World Wide Web, voor grafische elementen zoals logo's
- Er werd een alternatief ontwikkeld omdat het gebruik van LZW-compressie een betaling van een licentie vereiste (patent op) => PNG

PNG:

- Portable Network Graphics
- Bedoeld om GIF te vervangen
- Gebruikt deflate compressiemethode (geen patent op)
- Niet beperkt tot 256 kleuren
- Ondersteunt alpha-kanalen, die gedeeltelijk transparantie toelaat
- Dit formaat werd ontwikkeld onder toezicht van W3C
- PNG-bestanden worden gebruikt in andere contexten (bv iconen mac OS X)
- PNG is standaard formaat voor webbeeldeditor Fireworks

### ***JPEG-bestanden***

JPEG-data kan bewaard worden in verschillende formaten

JFIF: JPEG File Interchange Format

Wordt universeel gebruikt voor het bewaren van JPEG-data in webbeelden.

SPIFF: Still Picture Interchange File Format

ISO-standaard

JFIF en SPIFF gebruiken dezelfde conventies voor bestandsextensies(. jpg . jpeg) en beide hebben hetzelfde Internet Media Type (image/ jpeg)  
Exif kan zowel JPEG- als TIFF-data bevatten, samen met uitgebreide data.

### ***Andere formaten***

Belangrijkste bitmapbeeldformaat: PSD ( photoshop)

Grootste rivaal: TIFF ( Tag Image File Format)

= is een uitbreidbaar formaat dat full-colour bitmapbeelden kan bewaren door gebruik te maken van verschillende compressieschema's, inclusief LZW en JPEG.

= TIFF bestanden worden vaak gebruikt om ongecomprimeerde digitale foto's te bewaren en om beeld uit te wisselen

BMP-formaat: Microsoft Windows Bitmap Formaat

= platformafhankelijk, maar kan gelezen worden door de meeste programma's op andere systemen.

= ondersteunt een eenvoudige vorm van losless compressie en deze bestanden worden gewoonlijk zonder compressie bewaard.

PDF-formaat: PDF documenten kunnen bitmapdata bevatten, die kunnen gecomprimeerd worden met JPEG, JPEG2000, LZW,...

Camera-raw-data: wordt gebruikt wanneer complete controle over beeldwerking vereist is, maar er is geen standaardformaat voor

Adobe's DNG: Digital Negative

Dient als archiveringsformaat voor camera-raw-beelden

Gebaseerd op TIFF-formaat

### ***Metadata***

Beeldbestandformaten verschillen in hun ondersteuning in metadata

De meest voor de hand liggende plaats om beeld meta data bij te houden is in het beeldbestand zelf.

## Hoofdstuk 3: Kleur

Toevoeging van kleur is niet noodzakelijk een verbetering (Technicolor MGM musical 1950). Er zijn immers voordelen aan het gebruiken van beelden zonder kleur.

- bitmapbeelden in zwart-wit zijn veel kleiner
- zwart-wit beelden zijn immuun aan de variaties in kleurreproductie van verschillende monitoren
- Sommige mensen zien geen kleur en sommigen kunnen geen onderscheid maken tussen bepaalde kleuren
- Door zwart-wit beelden te gebruiken vermijd je het produceren van beelden die niet correct gezien kunnen worden omwille van deze redenen.

MAAR:

KLEUR is essentieel voor het doel waarvoor het gebruikt wordt, bijvoorbeeld aankoopgedrag beïnvloeden bij tijdschriften.

### ***Kleur en wetenschap***

Kleur is een subjectieve waarneming die geproduceerd wordt in de hersenen. We hebben een kleurmodel nodig dat de subjectieve waarneming relateert aan meetbare en reproduceerbare fysische fenomenen.

Licht is een vorm van elektromagnetische straling en zo kunnen we zijn golflengte en intensiteit meten. (golflengte tss 400 en 700 nm)

Deze metingen combineren we in een SPD (spectral power distribution)

=beschrijving van hoe de intensiteit van licht uit een particuliere bron varieert met zijn golflengte

SPD wordt opgebouwd door licht te splitsen in zijn componentgolflengtes en zijn intensiteit van elke component te meten.

SPD is te moeilijk om mee te werken voor kleur te specificeren => andere benadering.

Menselijk oog:

- Staafjes: voorzien nachtzicht en kunnen geen kleur onderscheiden
- Kegeltjes: komen overeen met verschillende golflengten van licht, komt op 3 manieren voor

Kleur kan kan gespecificeerd worden door juist 3 waarden: **tristimulustheorie**

= impliceert dat elke kleur kan geproduceerd worden door het mixen van de juiste hoeveelheden van de 3 additieve primaire kleuren

Elk type kegeltje antwoordt ofwel op rood, groen of blauw licht. Elk kleur kan geproduceerd worden door geschikte hoeveelheden van rood, groen en blauw.

Additieve primaire kleuren: rood, groen , blauw.

Pixels die samengesteld zijn uit 3 fosforpunten die rood, groen en blauw licht uitzenden. Deze fosforpunten worden geactiveerd door middel van 3 elektronenstralen, 1 voor elke kleur.

Gewenste kleur produceren ?

Intensiteit van elke elektronenstraak regelen

### ***RGB-kleur***

3 primaire kleuren: Rood, groen en blauw

Enkel kleuren in het RGB-gamma kunnen voorgesteld worden op deze manier.

Het is niet mogelijk elke zichtbare kleur voor te stellen (vinvormige oppervlakte is een ruimtelijke voorstelling van al de mogelijke kleuren); bepaalde kleuren kunnen niet samengesteld worden door groen, rood en blauw samen te voegen.

Kleur wordt vaak geschreven als een hexadecimaal getal van 6 cijfers met R, G en B tussen 0 en 255.

### ***Kleurdiepte***

Elke kleurwaarde neemt 24 bits in of 3 bytes.

Kleurdiepte

= aantal bits om een kleurwaarde te bevatten

Verschillende kleurdiepten(24-bit, 1-bit, 16-bit,...)

1-bit: zwart-wit

8-bit: 256 kleuren en wordt gebruikt voor monitoren en World Wide Web.

16-bit: meest gebruikt, maar is niet deelbaar door 3, dus ofwel wordt 1 bit niet gebruikt of er wordt een verschillend aantal bits toegekend aan de 3 componenten.

Menselijk oog is gevoeliger voor groen licht dan rood en blauw.

30, 36, 48-bit: voor scanners

Grote kleurdieptes hebben 2 doelen:

- De bijkomende info die bewaard wordt in de extra bits maakt het mogelijk om accuratere benaderingen te gebruiken wanneer het beeld gereduceerd wordt naar een lagere kleurdiepte voor weergave.
- Het is mogelijk om extreem fijn onderscheid te maken tussen kleuren, zodat effecten zoals chroma-key heel accuraat kunnen toegepast worden.

24-bit:	miljoenen kleuren:	True Colour
16-bit:	duizenden kleuren (65536 waarden)	Hi Colour
8-bit:	256 kleuren	

Kleurdiepte is een cruciale factor in het bepalen van de grootte van een bitmapbeeld

Elke logische pixel vereist 24 bits voor 24 bitkleur enz....

Vandaar mogelijk om de kleurdiepte van een beeld te reduceren van 24 naar 8 bit en bestandsgrootte zal afnemen met factor 3.

### ***Geïndexeerde kleur***

Directe kleur ⇔ Geïndexeerde kleur

Bij geïndexeerde kleur gebruikt men een 8-bit i.p.v. een 24-bit kleurwaarde die dient als een index in een kleurtabel. Het voorziet ook een manier om een palet van 256 specifieke kleuren te associëren met elk beeld.

Wanneer het beeld wordt weergegeven, zoekt het grafische systeem de kleur op uit het palet overeenkomstig met elke 1-byte waarde die bewaard is bij elke pixel. De kleurtabel bevat het palet van kleuren gebruikt in het beeld.

Zo'n geïndexeerde kleurtabel = CLUT = Colour Lookup Table of palet

= voert koppeling uit van logische kleuren naar fysische kleuren

= wanneer geen palet wordt toegevoegd: standaard systeempalet

Bepaalde kleuren uit het originele beeld kunnen ontbreken in het palet.

Hoe oplossen ?

De ontbrekende kleur moet vervangen worden door 1 die zich in het palet bevindt. Er zijn 2 populaire manieren hiervoor:

- De kleurwaarde van elke individuele pixel vervangen met de CLUT-index van het dichtstbijzijnde kleur=> ongewenste effecten: detail kan verloren gaan.  
Posterisatie= verenigen van kleuren en andere visuele artefacten kunnen opduiken waarbij gradaties van kleur vervangen worden door scherpe randen.
- Vlakken van 1 kleur kunnen vervangen worden door een patroon punten van verschillende kleuren.

Dithering = Gebruikt om posterisatie te verminderen.

= Uitbreiding van het gebruik van half-toning

Een blok van  $n \times n$  pixels kan  $n^2 + 1$  verschillende grijsniveaus simuleren.

Een palet zal je vullen met de meest belangrijke kleuren in jouw afbeelding.

Webveilig palet: 216 kleuren, kleuren zijn goed gereproduceerd door webbrowsers op elk systeem dat gebruik maakt van een 8-bit kleur.

### ***Andere kleurmodellen***

CMYK: cyaan, magenta en geel zijn de subtractieve hoofdkleuren. Ze zijn de complementaire kleuren van rood, groen en blauw.

Thomas Young 1801: stralen van de drie additieve hoofdkleuren worden geprojecteerd op een wit oppervlak.

Bij combinatie van deze drie wordt er wit licht gevormd.

+ additief mengen van licht

-aftrekken van licht

$$C = G + B = W - R$$

$$M = R + B = W - G$$

$$Y = R + G = W - B$$

De kleur links wordt de complementaire kleur genoemd van 1 van de kleuren uiterst rechts.

Relevantie van dit experiment is 2-voudig:

- Het vormt de basis voor een theorie rond kleurenleer die een grote invloed heeft gehad op het gebruik van kleur in kunst en design

- De idee van het vormen van kleuren voorziet door het aftrekken van licht, ipv het toevoegen ervan, een kleurenmodel dat geschikt is voor drukken en schilderen, omdat deze substanties zijn die hun gekleurde voorkomen toeschrijven aan de manier waarop ze licht absorberen

Dunne lagen inkt absorberen bepaalde componenten van het invallend licht.

Overlappende inkt mixt kleur subtractief.

Kleurengama van CMYK is niet hetzelfde als van RGB.

Als C, M en Y gecombineerd worden, krijgen we geen uitstekend zwart. Daarom worden deze subtractieve primaire kleuren uitgebreid met zwart. Deze 4 kleuren worden proceskleuren genoemd wanneer ze gebruikt worden bij het drukken.

Het CMYK-gamma is kleiner dan RGB, maar het bevat wel kleuren dat RGB niet bevat. CMYK essentieel voor kleurreproducties van hoge kwaliteit (inkt en papier), maar niet voor multimedia.

### ***HSB***

Elke kleur kan geïdentificeerd worden door zijn Hue, Saturation en Brightness.

Isaac Newton identificeerde 7 kleurschakeringen (regenboog)

Een pure kleurschakering kan meer of minder verbleekt worden door het te mengen met wit. Verzadigde kleuren zijn pure kleurschakeringen.

Bij verf:

- verminderen van verzadiging door wit toe te voegen : produceert tint
- Verminderen van helderheid door zwart toe te voegen: produceert toon

“Kleurenwiel”:

subtractieve primaire kleuren worden gelijk verdeeld langs de omtrek en additieve kleuren ertussen zodat elke primaire kleur tegengesteld is aan zijn complement.

Hues kunnen geplaatst worden langs de rand van het kleurenwiel met de complementaire kleuren tegenover elkaar. Een hue-waarde is de hoek tussen zijn positie op het wiel en de positie van rood.

Kleurenwiel kan uitgebreid worden om alternatief kleurenmodel te bekomen :

- De onderscheiden primaire, secundaire en tertiaire kleurschakeringen kunnen uitgebreid worden zodat ze een continuüm vormen langs de volledige cirkel.
- Elke kleurschakering kan gespecificeerd worden als een hoek rond de cirkel, gemeten tegen wijzerszin.
- Verzadiging kan toegevoegd worden door wit toe te voegen in het centrum van de cirkel.
- Om helderheid toe te voegen: derde dimensie nodig

HSB kan voorgesteld worden in een cilindrische vorm, zodat het voorgesteld kan worden als kleurenkiezers( colour-pickers)

### ***Kleurruimten gebaseerd op kleurverschillen***



Het is handig om de helderheidsinformatie van een beeld te scheiden van zijn kleur. Door die scheiding is het mogelijk om een beeld te verzenden op zo'n manier dat de kleurinformatie niet gedetecteerd werd door een zwart/wit ontvanger. Kleurruimten bestaande uit een brightness-component en 2 kleurverschilwaarden worden gebruikt bij video. Ze laten toe chrominance sub-sampling te gebruiken.

Y= Luminantie / helderheid

YUV-kleur= wordt gebruikt om te verwijzen naar elke kleurruimte bestaande uit een component van helderheid samen met 2 componenten voor kleurverschil

Digitale tv's gebruiken een variant genaamd Y1 Cb Cr.

## Hoofdstuk 4: Video

Ontwikkelingen video , hoe ?

- snelle stijging in processorsnelheden en geheugen, schijfcapaciteit en netwerkbandbreedte
- ontwikkeling van standaarden voor digitale videosignalen en interfaces
- verschuiving naar digitale video werd gedreven door zijn gemak en robustheid en de flexibiliteit en relatieve inhoud van digitale videobewerking werd in vergelijking met zijn analoog equivalent.
- 

VTR: video tape recorder, gebruikt 1 van de varianten van het DV-formaat. (mini DV ) verbonden met een Fire-Wire interface.

Er zijn 3 varianten van het DV-formaat: Mini-DV, DVCAM en DVCPRO

Deze gebruiken verschillende tapeformaten en zijn voorzien van verschillende graden van foutcorrectie en compatibiliteit met analoge studiouitrusting.

Mini : gebruikersformaat

DVCAM en DVCPRO: professioneel gebruik

Ze zijn alle 3 ondersteunt door een Device Control ( stoppen, starten, verplaatsen positie)

DV apparatuur gebruikt meer en meer High Definition.

DV is zeer goed van kwaliteit maar het is een gecomprimeerd formaat, d.w.z dat het compressieartefacten veroorzaakt en storingen bij opeenvolgende verwerking. De gebruiker heeft geen controle over de kwaliteit van DV.

Goedkope videocamera's in mobiele telefoons of laptops zijn geen DV apparaten, maar worden gedefinieerd door het MPEG-4 standaard of 3GP/ USB 2.0.

### ***Videostandaarden***

We moten in multimediacproducties werken met signalen die overeenkomen met de standaarden die gebruikt worden voor tv.

Videostandaarden variëren van land tot land

### ***Analoge videostandaarden***

Er zijn 3 reeksen standaarden

- NTSC: heeft het ontworpen  
525 lijnen: 480 beeld, weergegeven aan 59,95 velden (29,97 frames) per seconde (525 / 59.94)
- PAL-standaard:  
625 lijnen: 576 beeld, weergegeven aan 50 velden (25 frames) per seconde (625/50)
- SECAM- standaard

Deze drie standaarden handelen over de technische details waarop kleur-televisiebeelden gecodeerd worden als signalen..

Tv's worden gebaseerd op CRT's, cathode ray tubes. Ze werken volgens een raster scanning principe.

Schermd wordt opgedeeld in horizontale lijnen. Daarnaast worden 3 elektronenbundels, 1 voor elke additieve primaire kleur, gezonden en afgebogen door een magnetisch veld zodat ze zweven over het scherm, een lijn trekken, vervolgens naar beneden bewegen om de volgende lijn te trekken, enz.

Beeld moet ongeveer 40 keer per seconde ververs worden om flikkering te voorkomen.

Daarom moet elke frame worden opgedeeld in 2 velden (even en oneven lijnen)

Het weergegeven beeld wordt dus van boven naar onder opgebouwd als een sequentie lijnen. (1 per 1)

Interlacing: lijnen worden 1 per 1 verzonden zodat elke frame opgebouwd is door het samenvoegen van velden.

= progressive scanning, er treedt geen flikkering op omdat scherm ververs wordt uit de framebuffer op een hoge snelheid.

### ***Digitale videostandaarden***

Onvermijdelijk wegens de noodzaak aan compatibiliteit met de bestaande uitrusting.

Ze moeten in staat zijn om zowel 625/50 en 525/59.94 voor te stellen. Video moet eerst bemonsterd worden om geconverteerd te worden in een digitale vorm.

- ⇒ CCIR 601 : definieert standard definition digital video sampling met 720 luminantiesamples en 2 x 360 kleurverschilsamples per lijn.
- ⇒ Y' Cb Cr met **4:2:2 chrominance subsampling** = Om de grootte van digitale video te reduceren worden minder samples genomen voor elk van de kleurverschilwaarden als voor de luminantie.
- ⇒ CCIR 601 bestaat uit een luminantiecomponent en 2 kleurverschilcomponenten.

Hoogte-breedteverhouding bij PAL en NTSC is 4:3

Pal-frames zijn 720 x 576

NTSC-frames zijn 720 x 480

De pixels zijn niet vierkant.

Waarom niet 4:1:1 ?

Dit komt omdat er andere schikkingen voor de bemonstering mogelijk zijn. Bij 4:1:1 wordt elke 4<sup>e</sup> pixel op elke lijn bemonsterd voor de kleur. Bij 4:2:2 worden de kleurwaarden gesubsampled met een factor 2 in zowel de horizontale als de verticale richtingen.

### ***DV en MPEG***

Deze 2 standaarden zijn beide gebaseerd op Yt Cb Cr- componenten.

Digitale video is gebaseerd op de DV-standaarden. DVCAM EN DVPRO gebruiken allebei hetzelfde compressiealgoritme en dezelfde datastroom. Bemonstering: 4:1:1 subsampling. En bij het comprimeren hebben ze altijd een constante datasnelheid van 25 Mbits en compressieverhouding van 5:1.

MPEG heeft verschillende ISO-standaarden geproduceerd door ISO/IEC (Motion Picture Experts Group)

1<sup>e</sup> standaard : MPEG-1: hoofdzakelijk bedoeld voor video cd formaat

Opvolger : MPEG-2: voor digitale studiouitrusting (TV en DVD)

Opvolger : MPEG-4: ontworpen voor een groot bereik aan multimediate data te ondersteunen aan bitsnelheden van 10 Kbps per seconden.  
=>voor gebruik van mobiele telefoons tot HDTV

MPEG-4: onderverdeeld in verschillende delen: Multi-part standaard

- audiocompressie
- afleveren van data over netwerken
- bestandsformaten
- ..... => 23 delen in totaal

Deel 2 & deel 10: videocompressie

MPEG-4 part 2: MPEG-4 video

MPEG-4 part 10: verdere verfijning, en er wordt verwezen naar Advanced Video  
=H.264/AVC

Elke standaard van MPEG definieert verschillende profielen en levels.

Elk level definieert bepaalde parameters(max framegrootte), datasnelheid en chrominance subsampling.

Elk profiel kan geïmplementeerd worden op 1 of meerdere levels.

MPEG-2: meest voorkomende combinatie: Main Profile at Main level

- ⇒ MP@ML gebruikt CCIR 601 met 4:2:2 chrominance subsampling en ondersteunt een datasnelheid van 15 Mbits per seconde
- ⇒ Formaat wordt gebruikt voor digitale tv-uitzendingen en DVD-video

## H.264 / AVC

- Baseline Profile: geschikt voor video-conferencing
- Extended Profile: streamen van video
- Main Profile: algemeen gebruik
- High Profile: HDTV en Blu-Ray

Elk van deze profielen specificeren 16 verschillende waarden van parameters zoals de framegrootte en bitsnelheid (Levels)

BP@L1 (baseline profile level 1):

- Bitsnelheid: 64 kbps
- Framegrootte: 176 x 144 pixels
- Framesnelheid: 15 fps

HP@L5(High Profile level 5):

- Bitsnelheid: 300 Mbps
- Framegrootte: 4096 x 2048 pixels
- Framesnelheid: 30 fps

Elk level heeft 2 of meer sub-levels (L.s)

MPEG-4 definieert ook een **QuickTime-formaat** (gebruikt om gecomprimeerde videodata, audio en metadata op te slaan)

En ook het **3GP-formaat**: eenvoudigere versie, gebruikt bij mobiele telefoons

### ***High Definition formaten***

1<sup>e</sup> generatie digitale video: niet-viekkante pixels en vast aantal horizontale samples.

PAL & NTSC: SD ( standard definition)

HD-video: alles met grotere frames dan SD, gebruikt hogere resoluties en kan progressively gescand zijn.

Frames moeten 720 x 1080 pixels zijn en een verhouding van 16:9.

P: progressive

I: interlaced

HD-video gebruikt meer bits en eist meer schijfruimte, bandbreedte en verwerkingskracht.

### ***Videocompressie***

Elk videocompressiealgoritme bestaat uit een sequentie bitmapbeelden.

Er zijn 2 manieren waarop deze sequentie gecomprimeerd kan worden:

- elk individueel beeld apart comprimeren m.b.v. technieken: **Spatial** compressie : intra -frame
- deelsequenties van frames comprimeren door enkel de verschillen ertussen te bewaren: **Temporal** compressie: inter-frame

Spatial : compressie

Lossless: produceert geen voldoende hoge compressieverhoudingen

Lossy: leidt tot achteruitgang in beeldkwaliteit

Als er genoeg schijfruimte is, best werken met ongecomprimeerde video tijdens de postproductiefase. (tijdens het bewerken en toepassen van effecten)

Temporal compressie:

Key-frames (op constante intervallen)

De frames tussen de key-frames worden vervangen door een verschilframe, dat enkel de verschillen bewaart tussen het frame dat zich origineel in die positie bevond.

Waar pixels identiek zijn bij het aftrekken zal het resultaat 0 zijn en wordt weergegeven als zwart.

Veel types videomateriaal bestaan uit relatief statische vlakken, waarvan slechts een klein deel van het frame in beweging is.

Indien compressie en decompressie dezelfde tijd innemen spreekt men van een symmetrische codec.

### ***Spatial compressie***

Gebaseerd op DCT: Discrete Cosinus Transformatie

Gemakkelijkste benadering: JPEG-compressie

= wordt toegepast op 3 componenten van een kleurbeeld en werkt op dezelfde manier ongeacht de kleurruimte die gebruikt wordt om beelddata te bewaren.

Videodata wordt bewaard met Yt Cb Cr kleur met chrominance subsampling.

Op elk frame wordt motion JPEG of MJPEG-compressie toegepast.

DV-compressie begint met chrominance subsampling, deze hangt af van de videostandaard, PAL of NTSC. DV-compressie is puur Spatial.

Dus er wordt DCT toegepast: op 2 manieren:

1)

- frame is statisch: transformatie op volledige 8x8 blok
- frame beweegt : blok wordt opgesplitst in twee 8x4 blokken, elke blok wordt dan onafhankelijk getransformeerd
- De compressor moet bepalen of er beweging is adhv motion compensation

2)

- Uitvoeren uitgebreid proces van herschikking op de blokken die samen het volledige frame vormen
  - 14 bytes voor elk 8x8 blok
  - shuffling toepassen: videosegment construeren door systematisch 8x8 blokken te nemen uit 5 verschillende vlakken van het frame.
- ⇒ gevolg: hoeveelheid detail wordt evenredig verdeeld in elk videosegment.
- ⇒ het detail wordt verspreid over de videosegmenten wat een efficiënte compressie over het gehele beeld eenvoudiger maakt

### ***Temporal compressie***

Berekent verschil tussen de frames i.p.v. elk volledig frame te bewaren

### MPEG-1 standaard:

Combineert temporal compressie gebaseerd op motion compensation met spatial compressie gebaseerd op kwantificering en het coderen van frequentie coëfficiënten afkomstig uit een discrete cosinustransformatie van de data.

Hoe ?

Aftrekken van de waarde van elke pixel in een frame van dezelfde pixel in de vorige frame => verschilframe

MPEG-1 compressors verdelen elke frame in 16x16 blokken (=Macroblokken).

Hier moeten we de motion vectoren bewaren die de voorspelde verplaatsing van macroblokken tussen frames beschrijven. Deze motion vectoren kunnen op zichzelf worden gecomprimeerd.

MPEG key frames:

I-beelden (I: intra) => puur spatial compressie

P-beelden: verschromen die vorige frames gebruiken, kunnen gebaseerd zijn op een vorig I- beeld of P-beeld.

B-beelden: frames die voorspeld worden uit volgende frames en kunnen motion compensation gebruiken uit de volgende I of P beelden.

GOP: Group Of Pictures (een herhalende sequentie en begint altijd met een I-beeld)

Als B-pictures worden gebruikt moet een GOP herordend worden in bitstroomvolgorde voor de decoding.

Dus bvb IBBPBB IBBPBBI = MPEG-sequentie in weergavevolgorde



GOP

SIF: Source Input Format : klassiek formaat voor MPEG-1 video

MPEG-1 kan niet werken met interlacing en HDTV-formaten

⇒ NOOD aan MPEG-2 voor uitzendingen en studiowerk

Maar bestanden zijn te groot: 650 CD-ROM : 40 minuten video en 8,75 GB heeft ruimte voor meer dan 9u.

### MPEG-4 en H.264/AVC

Hogere profielen van deze standaard ontwikkelen een methode voor het onderverderen van een scène in willekeurig gevormde video-objecten die apart kan gecomprimeerd worden.

Lagere profielen worden beperkt tot rechthoekige objecten (complete frames)

MPEG-4 gebruikt global motion compensation en sub-pixel motion compensation om de kwaliteit van MPEG-1 en MPEG-2 te verbeteren.

Global motion compensation: comprimeert statische scènes met conventionele camerabewegingen. De bewegingen kunnen gemodelleerd worden als een vectortransformatie van de originele scène en voorgesteld worden door waarden en enkele parameters.

Sub pixel motion compensation: de verplaatsingsvectoren nemen de beweging op op een accuraatheid dat fijner is dan 1 pixel.

Bij Simple Profile :  $\frac{1}{2}$  pixel

Bij advanced Profile :  $\frac{1}{4}$  pixel

H624/AVC:

Sterk geoptimaliseerde versie van MPEG-4 part 2, deze codec moet op alle Blu-ray players implementeren.

- ⇒ kan de best mogelijke MPEG-2 kwaliteit geven aan de helft van de datasnelheid.
- ⇒ Laat toe data te gebruiken uit een stapel frames om het even waar in een film
- ⇒ Vergelijkbaar met woordenboekgebaseerde benadering van compressie bij LZ-algoritmes
- ⇒ Gebruikt dezelfde benadering als JPEG en andere MPEG-videocodes (B, I en P frames)
- ⇒ Gebruikt betere transformatie dan DCT, met keuze van 8x8 of 4x4 blokken
- ⇒ Werkt met filters voor verwijderen van bepaalde compressie-artefacten; resultaat: betere beeldkwaliteit
- ⇒ De-blocking filter: gebruikt om de karakteristieke discontinuïteiten tussen de pixelblokken te verzachten, die apart getransformeerd zijn.
- ⇒ Niet praktisch voor live video
- ⇒ Te traag om ruwe voorvertoningen te maken
- ⇒ Bevat één-stap modus waarin de video zo snel mogelijk moet gecomprimeerd worden

### ***Andere videocodes***

-Windows Media 9 (VC-1): WMV9

gelijkenissen met H624/AVC:

- bi-directionele voorspelling
- motion compensation
- de-blocking filter

Verschillen

- Ondersteunt differentiële kwantificering = verschillende kwantificeringsmatrices worden gebruikt voor verschillende delen van een frame
- WMV kan zijn DCT toepassen op elk 8x8 blok van pixels of dit opbreken in twee 8x4 blokken, twee 4x8 blokken of vier 4x4 blokken.

- Deze kleinere transformatieblokken kunnen de visuele artefacten reduceren aan de randen van de blokken.
- In staat om extra compressie te bekomen bij fade transitities, door de fades te detecteren

WM9 codec gestandaardiseerd door SMPTE, onder de naam VC-1=> verplicht voor Blu-Ray players

VC-1 heeft verschillende profielen en levels

On2 VP6 codec: gebruikt voor Flash Video (geen officiële standaard)

Het is een andere DCT-gebaseerde techniek, met inter-frame compressie en motion compensation. Het ondersteunt geen bi-directionele voorspelling. P-beelden kunnen enkel afhangen van P- en I-beelden die voorafgaan.

Voordeel= eenvoudig om video te decomprimeren dat hiermee gecompriemd werd.

Ogg Theora:

- Gebruikt JPEG-achtig lossy compressiealgoritme gebaseerd op DCT gevolgd door kwantificering
- Het is gekoppeld met inter-frame compressie met motion compensation
- Enkel I en P beelden worden ondersteund, geen B

On2 VP6 en Ogg Theora zijn minder krachtig, maar gratis en veel gebruikt.

Alle moderne codecs produceren excellente kwaliteit aan 2 Mbps en hoger.

### ***Aflevering***

Video kan bezorgd worden over netwerken, als een gedownload bestand en kan gestreamd worden of afgeleverd worden door progressieve download.

### ***Streaming***

- Afspelen zonder dat het bewaard moet worden op de harde schijf
- Biedt de mogelijkheid om live-video af te leveren
- Het is niet beperkt tot 1 enkele transmitter die uitzendt naar veel consumenten.
- Grootste obstakel: bandbreedte
- Eenvoudigste methode: ingebedde video= filmbestand overbrengen van server naar de machine van een gebruiker en start het afspelen van zodra het volledige bestand is aangekomen.
- True streaming
- Verfijning van deze methode: progressive download / http-streaming

Progressive download

=de film start met spelen wanneer de tijd nodig om de overblijvende frames te downloaden kleiner is dan de tijd om de hele film te spelen (zie p156 fig)



Vaak vertraging voor het afspelen begint omdat progressief gedownloadde films meestal gemaakt zijn met een datasnelheid die de bandbreedte van de netwerkconnectie overstijgt

- ⇒ filmbestand blijft op harde schijf, in de cache, dus er moet genoeg schijfruimte voorzien zijn (niet gebruikt voor speelfilms)
- ⇒ geen live-video

True streaming

- Video wordt nooit bewaard op de harde schijf
- Kleine buffer kan gebruikt worden om jitter weg te werken (p155 groen)
- Elke frame wordt gespeeld van zodra het aankomt over het netwerk
- De stream kan open-ended zijn
- Kan gebruikt worden voor live-video
- Willekeurige toegang tot specifieke punten is mogelijk
- Geschikt voor video-on-demand toepassingen en video-conferencing
- Raadzaam om kleinere frames te gebruiken

Nadeel : speciale server vereist

Progressive download kan opgezet worden met een gewone webbrowser

### ***Architecturen en formaten***

Multimedia-architectuur bevat verschillende eigenschappen:

- API: Application Programming Interface die mogelijkheden voorziet voor media, capture, compressie en het afspelen.
- 1 of meerdere codecs
- containerformaat voor bewaren van mediadata
- streaming-server
- software-tools, zoals een player

een goede multimedia-architectuur maakt het mogelijk om bijna elke movie af te spelen dat het aangeboden krijgt zodat gebruikers zich geen zorgen hoeven te maken over codecs en formaten.

Eerste architectuur: **QuickTime** , geïntroduceerd door Apple in 1991.

Containerformaat is het moviebestand: .mov

Beschikbaar op Macintosh en windowssystemen

Multimediasysteem windows: **DirectShow**, met containerformaat ASF(advanced systems format). Deze hebben meestal VMW bestanden. = .vmw

QuickTime en DirectShow zijn uitbreidbaar met componenten waardoor ze veel codecs kunnen gebruiken, inclusief H.264/AVC en WMV 9 en bijkomende bestandsformaten kunnen lezen en schrijven, zoals MP4 en AVI.

Flashvideo is vaak gebruikt voor webvideo. FLV-bestanden worden afgespeeld in de Flash Player met een SWF die het afspelen controleert.

Ogg is een open formaat, gebruikt met de Theora-codec om filmpjes te produceren die niet onderhevig zijn aan beperkingen of licentiebijdragen.

Bij webvideo wordt framegrootte en framesnelheid gereduceerd voor compressie en exporteren naar een geschikt formaat dat afgespeeld kan worden op de meeste systemen.

MP4 en FLV zijn de twee meest vanzelfsprekende keuzes om video te verspreiden. MP4 is een standaard en afspeelbaar door DirectShow, QuickTime en OpenSource speler. FLV is afspeelbaar door de Flash Player, die beschikbaar is op alle grote platformen en geïnstalleerd is op de meeste machines.

## Hoofdstuk 5: Audio

Geluid heb je niet altijd nodig, ze kunnen irriterend zijn.

Er zijn twee types geluid: muziek & spraak

### ***Aard van geluid***

Geluiden worden geproduceerd door conversie van energie in trillingen in de lucht of een ander elastisch medium, die gedetecteerd worden door het oor en geconverteerd worden in zenuwimpulsen die we ervaren als geluid.

Vorken van een goede stemvork vibreren zuiver op 1 enkele frequentie. De meeste andere geluidsbronnen vibreren op een meer gecompliceerde manier.

Frequentiespectrum

= beschrijving van de relatieve amplitudes van zijn frequentiecomponenten

Menselijk oor is in staat frequenties te detecteren tussen 20 Hz en 20 Khz.

De mogelijkheid om hogere frequenties te horen hangt af de leeftijd.

### ***Golfvormen***

Golfvorm van een geluid afbeelden door zijn amplitude uit te zetten in functie en tijd.

Een geluidsvorm toont hoe de amplitude varieert over de tijd.

De afbeelding van een golfvorm kan een bepaalde hoeveelheid van het karakter van een geluid tonen, maar geen details.

Grootste voordeel van visuele weergaven is hun statisch karakter. Dit maakt de analyse van de temporele structuur van geluid relatief eenvoudig.

### ***Perceptie***

1 van de meest gebruikbare illusies in geluidspceptie is stereofonie

Proces van kwantificering en bemonstering.

De geluids informatie moet geconverteerd worden naar een elektrisch signaal vooraleer het gedigitaliseerd kan worden. (kan gebeuren door microfoon of andere transducer)

### ***Bemonstering***

Als de grens van het horen genomen wordt op 20 kHz, dan is een minimum van 40 kHz vereist omwille van de bemonsteringssnelheid.

Die snelheid die gebruikt wordt voor audio-cd's is 44,1 kHz.

Subveelvouden kunnen gebruikt worden voor lage-kwaliteit digitale audio. (22,05 voor audio & 11,025 voor spraak)

Sommige professionele en semi-professionele opname-apparatuur gebruiken bemonsteringssnelheden die veelvouden zijn van 48 kHz.

CD-spelers en solid-state recorders hebben het voordeel dat ze een digitale uitvoer kunnen genereren.

Bij hoge bemonsteringsfrequenties is er een kleine tolerantie voor jitter.

### ***Kwantificeren***

Meest voorkomende keuze van bemonsteringsgrootte is 16 bits en geeft 65536 kwantificeringsniveaus.

Kwantificeringsruis kan verzacht worden door dithering

= kleine hoeveelheid random ruis toevoegen die de scherpe transitie van kwantificeringsruis verzacht.

De aanwezigheid van de ruis heeft ervoor gezorgd dat de samples snel wisselen tussen de kwantificeringsniveau's, in plaats van zuiver en abrupt te springen van het ene naar het andere. (figuur p 178-179) De scherpe transitie wordt afgezwakt.

Hoger betrouwbare reproductie: 24 bits

### ***Formaten***

AIFF voor MacOS

WAV voor Windows

AU voor Unix

Op het internet is het MP3-formaat dominant.

Podcasts gebruiken MP3.

MP3-data kan bewaard worden in QuickTime en Flash movies.

MP3 is hoofdzakelijk een codering, geen bestandsformaat.

Streaming is succesvoller voor geluid dan voor video, wegens de lagere bandbreedte die vereist is voor audio.

### ***Compressie***

Geluid is moeilijk te comprimeren met lossless methoden uitgezonderd speciale gevallen.

Huffman-codering kan effectief zijn als de amplitude van het geluid hoofdzakelijk valt onder het maximum niveau dat kan voorgesteld worden in de gebruikte bemonsteringsgrootte.

Een goede compressie techniek: verwijderen van stilte. In de plaats van 44100 samples te gebruiken met een nulwaarde voor elke second stilte, bewaren we de lengte van de stilte. = speciaal geval van run-length codering die lossless is. Stilte is echter zelden absoluut => weinig compressie

### ***spraakcompressie***

companding gebruikt niet-lineaire kwantificering om spraak te comprimeren.

De hogere niveau's worden verder uit elkaar geplaatst dan de lagere zodat rustige geluiden in groter detail worden voorgesteld dan de luidere.

Compressie ! => er zijn minder bits nodig om het volledige bereik van mogelijke invoerwaarden voor te stellen dan een lineair kwantificeringsschema.

2 ITU-recommendations voor gebruik in de telecommunicatie (telefonie):

- u-law: gebruikt in Noord-Amerika
- A-law: gebruikt in de rest van de wereld

Telefoonsignalen worden bemonsterd op slechts 8 kHz

Een andere techniek die gebruikt wordt:

ADPCM: Adaptive Differential Pulse Code Modulation

= werkt door het bewaren van informatie over het verschil tussen een sample en een waarde voorspeld uit het voorgaande sample

ADPCM wordt ook gebruikt in de telefonie met datasnelheden van 16 kbps en 32 kbps.

Lineair Predictive coding gebruikt een wiskundig model van de toestand van het vocale tractaat als zijn voorstelling van spraak. Het zendt parameters die de overeenkomstige toestand van het model beschrijven.

### ***Perceptuele compressie (waarnemingsgebaseerde compressie)***

Deze compressie verwijdert onhoorbare geluiden