

H1: Achtergrond

Digitale voorstelling

Digitalisatie

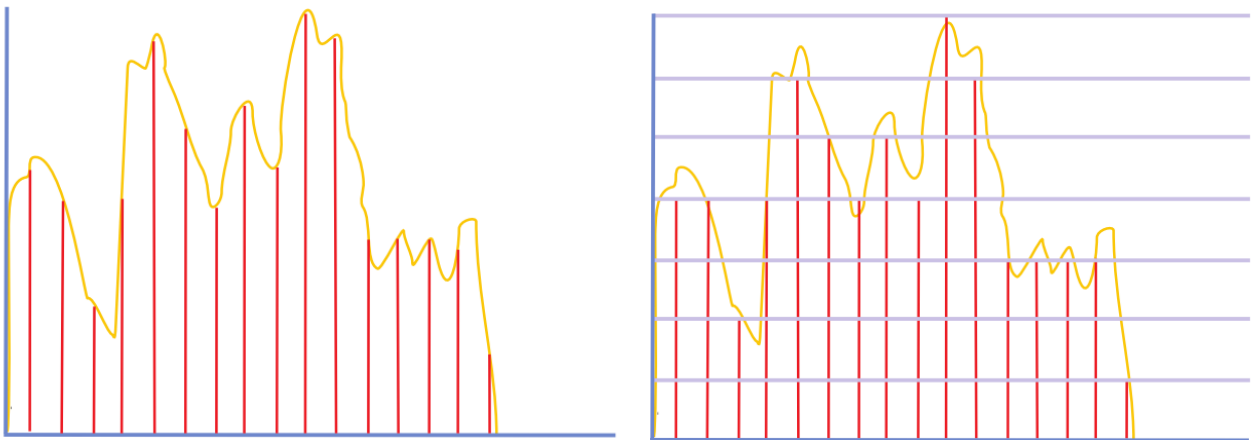
oorsprong van digitale media:

- soms in digitale vorm (beelden worden gemaakt met een tekenprogramma)
- meestal omzetting van een analoge voorstelling (continu fysisch fenomeen) naar een digitale voorstelling
 - ↳ *foto nemen met digitale camera*: intensiteit/kleur van het licht varieert continu, wordt omgezet in een continu variërende elektrische lading in de sensor van de camera
 - ↳ *geluid opnemen met microfoon*: de amplitude van een geluidsgolf varieert continu, wordt omgezet in een continu variërende elektrische lading in de microfoon
 - ↳ *analoge/digitale klok*: een analoge klok is gedetailleerder (seconden kunnen geschat worden), maar een digitale klok is eenvoudiger en makkelijker om berekeningen te doen

signaal: een waarde die we meten en die kan variëren in tijd en/of ruimte

digitalisatie: het proces waarbij we een analoog signaal converteren naar een digitaal signaal

1. bemonstering/sampling (het meten van de waarde op discrete tijdsintervallen)
 2. kwantificering (het beperken van de waarde tot een vaste verzameling van niveaus)
- ↳ kunnen uitgevoerd worden in willekeurige volgorde
 - ↳ meestal uitgevoerd door speciale hardware: ADC's (*analoog-naar-digitaal-converters*)



bemonsteringssnelheid/sample rate: het aantal monsters in een vaste tijds- of ruimteteveelheid

kwantificeringsniveau's: niveaus waarop een signaal gekwantificeerd wordt

- ↳ gelijk verspreid bij een vaste sample rate

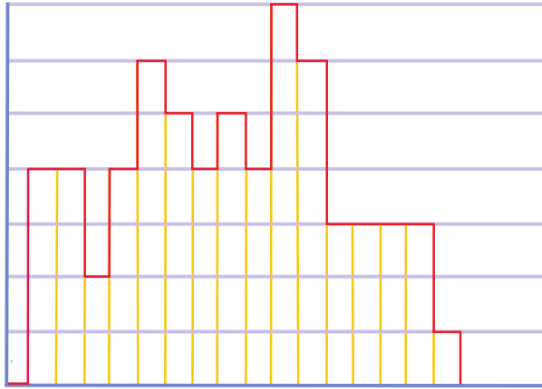
digitale signalen ondergaan **geen verlies** wanneer ze overgebracht worden naar media onderhevig aan ruis

- ↳ analoog → ruis is min of meer ondetecteerbaar (elke analoge waarde is aannemelijk, vervuild signaal kan niet onderscheiden worden van een proper signaal)
- ↳ digitaal → elke kleine fluctuatie wordt getransformeerd naar een ongeldige waarde (tussen de kwantificeringsniveaus), het originele signaal kan hersteld worden door opnieuw te kwantificeren
 - enkel als de ruis voldoende groot is om het signaal te wijzigen naar een ander niveau zal een fout in de transmissie opduiken

Reconstructie van analoge uit digitale signalen

= beslissen wat er in de gaten tussen de monsters geplaatst zal worden

- **bemonster en houd aan:** de waarde van een sample wordt gebruikt voor de volledige reikwijdte tussen de huidige en de volgende sample
 - ↳ abrupte overgangen → niet echt goede benadering van het origineel
 - ↳ vertragingen/onvolkomenheden van het uitvoerapparaat zorgen ervoor dat deze discontinuïteiten verminderen, met een vrij goed resultaat tot gevolg



undersampling: details van het analoge signaal vallen weg omdat ze tussen de monsters liggen

- ↳ vervormingen/artefacten die altijd ontoelaatbaar zijn
- ↳ = bemonsteren op een snelheid lager dan de Nyquistsnelheid

muzieknoot: pure tonen (sinusgolven van één enkele frequentie) die samengevoegd zijn

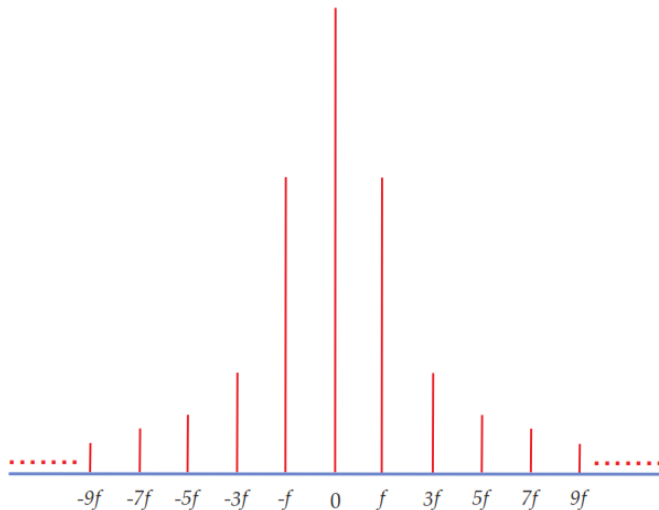
- grondtoon (toon die geassocieerd is met de noot)
- harmonische tonen/bijtonen (geven de noot zijn onderscheidend timbre/klankkleur)

hoe meer frequentiegolven, hoe beter de benadering

Fouriertransformatie: wiskundige operatie die een signaal weergeeft als een frequentiespectrum

frequentiespectrum: een grafiek waarbij de horizontale as de frequentie voorstelt en de verticale as de amplitude

- ↳ bestaat typisch uit een reeks pieken op verschillende frequenties overeenkomstig de verschillende componenten
- ↳ geen faseverschuivingen = symmetrische voorstelling (neg. frequenties = pos. frequenties)
- ↳ DC-component = integraal van het signaal = gemiddelde waarde van het signaal = nul-piek



inverse Fouriertransformatie: een signaal uit het frequentiedomein omzetten in het tijdsdomein

hogere frequentiecomponenten worden geassocieerd met abrupte overgangen

samplingstelling: als de hoogste frequentiecomponent van een signaal f_h is, kan het signaal correct gereconstrueerd worden indien het bemonsterd werd op een frequentie die groter is dan $2f_h$ (= *Nyquistsnelheid*).

aliasing: bepaalde frequentiecomponenten in het originele signaal worden omgezet in andere frequenties wanneer het signaal gereconstrueerd wordt (door *undersampling*)

- schokkende bewegingen bij bewegende beelden (bv. een roterende schijf lijkt in tegenwijzerzin te draaien terwijl het eigenlijk in wijzerzin draait)
- distortion bij geluid
- gekartelde lijnen of Moirépatronen bij beelden

posterisatie/brightness contouring: doordat er een onvoldoende aantal kwantificeringsniveaus aanwezig zijn, zijn er maar een beperkt aantal kleuren mogelijk en smelten gekleurde vlakken samen.

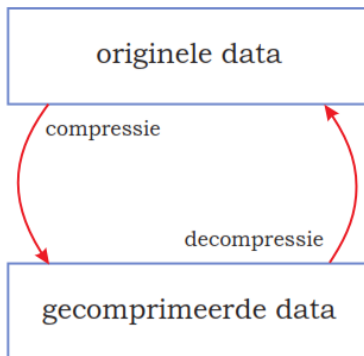
- reduceert het aantal bits die gebruikt wordt om elke monster te bewaren
- reduceert de bestandsgrootte

kwantificeringsruis: doordat er een onvoldoende aantal kwantificeringsniveaus aanwezig zijn, ontstaat er een vorm van distortion op geluid (sissend geluid, verlies aan rustige passages, algemene vaagheid)

- ↳ duidelijk waarneembaar bij bemonstering op 8 bits (256 niveaus)
- ↳ niet waarneembaar bij bemonstering op 16 bits (65536 niveaus, bij audio-cd's)

Compressie

= elke operatie die uitgevoerd kan worden op de data om de opslaghoeveelheid te reduceren



codec: software die compressie en decompressie uitvoert (vooral in de context van audio/video)

lossless compressie: na decompressie wordt een exacte kopie verkregen van de originele data

- ↳ werkt door het herorganiseren van data op verschillende manieren
- ↳ niet handig om mee te werken
- ↳ hoeveelheid compressie hangt af van de data en het type/de structuur van de data
- ↳ bv. tekstcompressie: een verschillend aantal bits gebruiken voor elk karakter en minder bits toekennen aan de meest voorkomende karakters

lossy compressie: bepaalde data werden onherstelbaar verwijderd in het compressieproces

- ↳ werkt door het verwijderen van niet betekenisvolle data
- ↳ worden vaak gebruikt en zijn goed in het bewaren van de kwaliteit van beeld, video en geluid
- ↳ toegepast voor de meeste multimediatoepassingen
- ↳ gebruikt ook bepaalde lossless technieken als onderdeel van het totale compressieproces
- ↳ wordt in het ideale geval enkel toegepast op het eindproduct
 1. bij herhaaldelijke compressie/decompressie zal de kwaliteit gradueel verslechteren
 2. bepaalde bewerkingen kunnen het kwaliteitsverlies verergeren
- niet altijd mogelijk: video/beeld wordt vaak gecompriemerd in de camera
- ↳ o.a. JPEG voor beeldcompressie, MP3 voor audiocompressie

op een bepaald moment zal de gecompriemede versie van de data groter zijn dan de ongecompriemede versie

Digitale voorstelling van media

tekstuele voorstelling

- ↳ moet voldoen aan de syntaxregels van een bepaalde taal om verwerkt te kunnen worden
- ↳ bv. SVG-bestand

binaire voorstelling: directe voorstelling van het beeld door bitpatronen die niet geïnterpreteerd kunnen worden als karakters

- ↳ moet voldoen aan een bepaald bestandsformaat om verwerkt te kunnen worden door software
- ↳ hebben een bepaalde structuur (bv. eerste vier bytes = lengte van de mediadata)
- ↳ bv. TIFF-bestand

Stilstaand beeld

= rechthoekige array van pixels die optisch samenvloeien wanneer deze op voldoende afstand bekeken worden

resolutie: het aantal pixels dat een apparaat kan afbeelden binnen een lengte-eenheid

- ↳ anders bij verschillende apparaten en monitors
- ↳ dikwijls bewaard met de beelddata

rendering: het proces om een patroon pixels te genereren uit een model

- ↳ gewoonlijk in de vorm van een expliciete datastructuur die een beschrijving van het beeld bijhoudt
- ↳ model moet bijgehouden worden in het bestand om het beeld later opnieuw af te beelden
- ↳ begint met een programma dat de data inleest, decomprimeert en d.m.v. functies uit een grafische bibliotheek een interne datastructuur bouwt, overeenkomstig met de beeldbeschrijving van het bestand

2 verschillende benaderingen van grafische modellen:

- **bitmapbeelden** (het beeld wordt gemodelleerd door een array van pixelwaarden)
 - ↳ verschil tussen logische pixels (opgeslagen waarden) en fysische pixels (punten op scherm)
 - ↳ array van pixelwaarden = de bitmap (≠ de afbeelding zelf)
 - ↳ de grootte is afhankelijk van de grootte en resolutie van het beeld (de waarde van elke logische pixel wordt bewaard)
- **vectorbeelden** (het beeld wordt bewaard als een wiskundige beschrijving van een verzameling lijnen, curven en vormen die het beeld samen voorstellen)
 - ↳ vergt bepaalde berekeningen die moeten uitgevoerd worden om zo het model te interpreteren en een array van pixels te genereren
 - ↳ bv. lijn voorstellen door zijn eindpunten te bewaren, bij renderen worden de coördinaten van de pixels die op de lijn liggen berekend
 - ↳ de bestandsgrootte is afhankelijk van de complexiteit van de vormen (≠ resolutie)

Verschillen		
	Bitmapbeelden	Vectorbeelden
Bestandsgrootte	Afhankelijk van de grootte en resolutie van het beeld	Afhankelijk van de complexiteit van de vormen
Visuele karakteristieken	Realistische schakeringen en texturen	Gestileerde kwaliteit (duidelijk uitgelijnde vormen, gemaakt uit een collectie vloeiende curven, opgevuld met kleur)
Individuele elementen	Het beeld is een array van pixels, zonder onderliggend model dat individuele elementen bepaalt	Elk element van de afbeelding behoudt zijn identiteit en kan bewerkt worden als een object
Effecten	Eenvoudig, omdat elke pixel getransformeerd wordt, mogelijk door gebruik te maken van de waarde van zijn burens	Beeld eerst omzetten naar bitmap-formaat
Schalen	Brengt een duidelijk verlies aan kwaliteit met zich mee (gekarteld/wazig)	Eenvoudige wiskundige operatie voordat de pixelwaarden berekend worden (kwaliteit verandert niet)
Voorbeelden	Gescande beelden, screenshots, foto's en videoframes	Grafieken, diagrammen en andere visualisaties van computer-gegenereerde data

Video en animatie

nawerking v.h. oog: vertraging in het antwoord van het oog op visuele stimuli dat resulteert in 'nabeelden'.

- als een reeks beelden weergegeven wordt aan onze ogen aan een snelheid boven de fusiefrequentie, zien we bewegende beelden in plaats van losse, individuele beelden.
- de fusiefrequentie hangt af v.d. helderheid v.h. beeld in vergelijking met de achtergrond, maar is rond 40 beelden per seconden

het begrip video:

- vroeger: systemen gebruikt voor het opnemen van televisiesignalen naar tape (zowel voor als na transmissie) en bewegende beelden opgenomen door deze systemen.
- nu: digitale video
 - ↳ kan gemanipuleerd worden door computers
 - ↳ veel flexibeler (geen dure apparatuur, niet verbonden met tv-signalen of standaarden)
 - ↳ standaard middel om live-actie te bewaren
 - ↳ bestaat in feite uit een reeks bitmapbeelden, maar toch complexer:
 - moet bewaard, verzonden en efficiënt afgespeeld kunnen worden
 - moet gecomprimeerd kunnen worden (o.a. H.264/AVC, On VP6, WMV, Ogg Theora)

standaarden:

- elke frame SD-video gegenereerd door een DV-camera is 720 pixels breed
- PAL-standaard (grootste deel van Europa, Australië): 576px hoog aan 25 frames per seconde
- NTSC-standaard (o.a. Noord-Amerika): zelfde grootte (kleinere frames, hogere fps)

Geluid

= veroorzaakt door snelle veranderingen in de luchtdruk in de tijd, die gedetecteerd worden door de oren

- digitale effecten kunnen eenvoudig toegepast worden
 - worden geïmplementeerd door algoritmes in programma's
 - volledige muziekstukken kunnen gemaakt worden door samples en loops te combineren
 - bv. toonhoogte veranderen zonder de duur te veranderen
- analoge effecten zijn afhankelijk van de beschikbaarheid van circuits die de gewenste veranderingen aan het signaal maken

H2: Bitmapbeelden

Inleiding

voordelen:

- conceptueel veel eenvoudiger dan vectorbeelden (geen nood aan wiskundige modellering)
- veel meer visuele mogelijkheden (niet beperkt door het tekenen van bepaalde vormen)
- kunnen gebruikt worden om foto's voor te stellen

nadelen:

- potentieel uitgebreide bestandsgrootte
- objecten binnen een beeldbestand kunnen niet geïdentificeerd worden in het geval van bewerking/transformatie

kunnen gemaakt worden door:

- externe hardware die kleuren meet en zo een array van waarden aanmaken (bv. scanners/camera's)
- zelf digitaal tekenen door visuele artiesten, met programma's zoals *Paint*
- vectorbeelden renderen

Resolutie

= een maat van hoe fijn een apparaat continue beelden benadert m.b.v. eindige pixels

→ eenvoudig concept, maar verschillende betekenissen kunnen tot verwarring leiden

Apparaatresolutie

oorspronkelijk: een maat van de mogelijkheid van een apparaat om fijn detail te onderscheiden

- ↳ kan gemeten worden door zowel analoge/digitale apparaten
- ↳ camera/scanner: hoe ver uit elkaar de lijnen op een doelrooster moeten zijn voor het apparaat om elke lijn te zien
- ↳ printer/scherm: hoe ver uit elkaar het apparaat de lijnen op zo'n doelrooster kan produceren
- ↳ digitale apparaten → aantal lijnen/inch = aantal punten/inch (een lijn is één punt breed)

nu: de pixeldichtheid dat een apparaat kan weergeven of bemonsteren

- ↳ uitgedrukt in *dots per lengte-eenheid* (Engelstalige landen → *dots per inch/dpi*)

scanners: resolutie ≠ pixeldichtheid (effect van optische componenten)

- ↳ fabrikanten halen hoge scannerresoluties aan door geïnterpoleerde waarden te tellen die niet rechtstreeks gescand zijn
- ↳ maat voor fysieke resolutie → samples per inch

video: resolutie is de pixeldimensie (grootte van een frame in pixels) i.p.v. pixeldichtheid

- ↳ elke monitor (bep. videostandaard) gebruikt hetzelfde pixelraster, onafhankelijk van de grootte
- ↳ bv. PAL-frame = 768x576, NTSC-frame = 640x480

computermonitors: resolutie betekent

- de beeldgrootte (zelfde technologie als videomonitoren)
- de pixeldichtheid (beeld voorbereiden of software schrijven om het weer te geven)
 - ↳ altijd vaste waarde, nl. 72 dpi
 - ↳ zorgt ervoor dat een beeld altijd dezelfde proporties van een scherm zal innemen met dezelfde pixeldimensies
 - ↳ zal veranderen wanneer monitoren met veel hogere resoluties wijdverspreid geraken

digitale camera's: resolutie betekent het totaal aantal pixels in beeld

- ↳ werken met een bepaalde beeldverhouding
 - low- tot mid-range camera's → 4:3 (= standaard)
 - high-range camera's → 3:2
- ↳ bv. 5 megapixel-camera (+/- 5 miljoen pixels) = 2584x1936 (4:3) of 2736x1824 (3:2)

Beeldresolutiepixeldimensie (array van pixelwaarden): zegt hoeveel detail vervat zit in het beeld

- ↳ bv. 1080 pixels x 720 pixels

pixeldichtheid (dpi): zegt

- ✓ hoe groot het beeld zal zijn op het uitvoerapparaat
- ✓ hoe eenvoudig het zal zijn om individuele pixels te zien
- ↳ bv. 72 dpi

$$\text{afmeting} = \frac{\text{pixeldimensie}}{\text{pixeldichtheid}}$$

- ↳ indien twee beelden dezelfde afmetingen blijken te hebben op apparaten met verschillende resoluties, hebben de beelden verschillende pixeldimensies

pixellation: vorm van aliasing, waarbij individuele pixels duidelijk zichtbaar worden bij een vergroting

- ↳ anti-aliasing kan niet worden toegepast op bitmapbeelden

natuurlijke grootte van beelden:

- grootte van een origineel voor het gescand werd
- grootte van het canvas in een tekenprogramma
- standaard printgrootte voor afbeeldingen
- ↳ meeste beeldformaten bewaren een *beeldresolutie (ppi)* samen met de data
- ↳ is gewoonlijk gelijk aan de *apparaatresolutie (dpi)*
- ↳ beelden afbeelden op natuurlijke grootte = schalen met de factor $\frac{\text{apparaatresolutie (dpi)}}{\text{beeldresolutie (ppi)}}$

foto- en videocamera's:

- ↳ vaste pixeldichtheid (72 dpi) om overeen te komen met de resolutie van monitoren
- ↳ beeld van een camera wordt afgebeeld op een monitor met de eigenlijke pixels
 - video → hedendaagse monitoren zijn groot genoeg
 - foto → hedendaagse monitoren zijn te klein (5+ Megapixel) → schalen

Hersampling

hersampling: pixeldimensies veranderen; *downsampling* en *upsampling*

- ↳ elke pixel moet getransformeerd worden om te kunnen schalen
- ↳ de nieuwe pixels komen meestal niet overeen met het oorspronkelijke pixelraster
- ↳ een bepaalde interpolatiemethode moet gebruikt worden om de waarde van bep. pixels te berekenen (= reconstructie, dus verlies aan zichtbare kwaliteit)

downsampling: de pixeldimensies van een beeld reduceren (de resolutie verminderen)

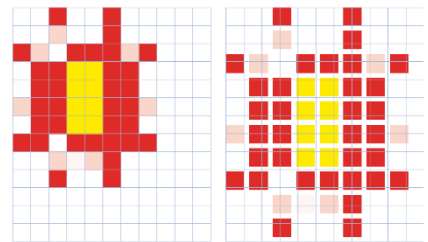
upsampling: de pixeldimensies van een beeld uitbreiden (de resolutie vergroten)

Pixels schalen

transformatie met schaafactor s en de pixel op coördinaten (x, y) :

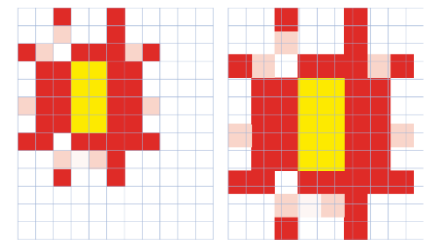
$$(x', y') = (sx, sy)$$

- ↳ de meeste pixels worden afgebeeld op punten die niet liggen op het pixelrooster van het nieuwe beeld
- ↳ sx en sy zullen meestal geen gehele waarden zijn → gaten tussen de geschaalde pixels



inverse koppeling waarbij we de pixel op coördinaten (x', y') in het uitvergrote beeld gelijk stellen aan de waarde op $(x, y) = (x'/s, y'/s)$ in het origineel

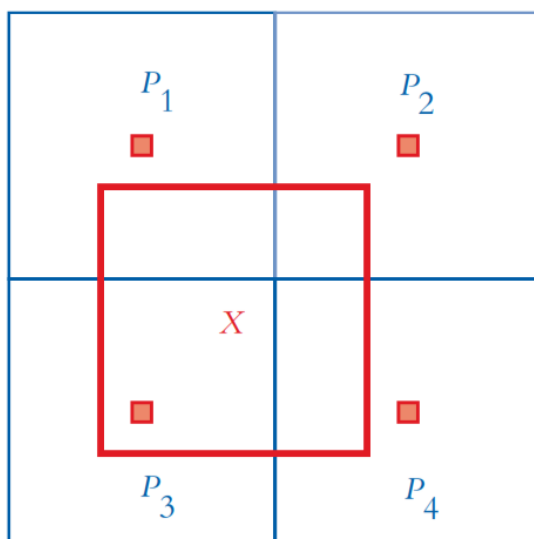
- ↳ voor elke pixel in het nieuwe beeld een pixel vinden in het originele beeld
- ↳ we berekenen enkel de pixelwaarden die we nodig hebben
- ↳ x'/s en y'/s zullen meestal geen gehele waarden zijn → slechte uitlijning van de pixelroosters
- ↳ waarden van verschillende naburige pixels moeten gecombineerd worden



Interpolatiemethoden

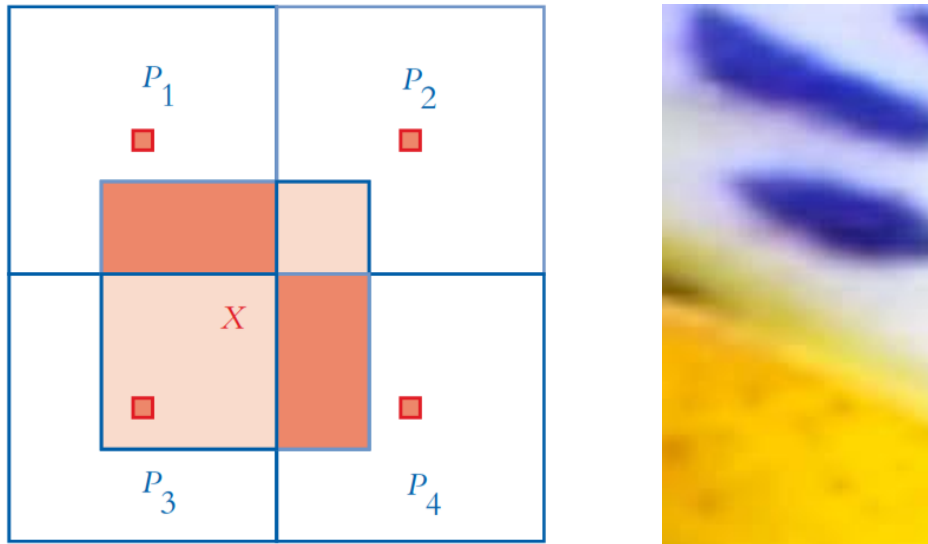
dichtste buur: de waarde van de pixel waarvan het midden het dichtst ligt bij (x, y) wordt gebruikt (hier: P_3)

- ↳ gebruikt 1 pixel uit het originele beeld voor elke pixel van het resultaat
- ↳ resultaat ziet eruit alsof de originele pixels uitvergroot werden door een vergrootglas



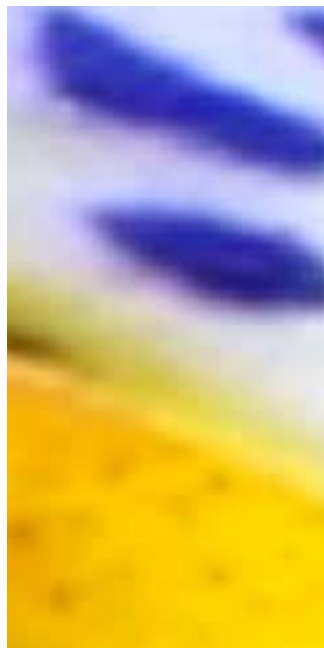
bilineaire interpolatie: de waarden van alle vier naburige pixels worden gecombineerd in verhouding met de oppervlakte van het intersectie met de doelpixel

- ↳ hier: P_1 , P_2 , P_3 en P_4 worden vermenigvuldigd met de opp. van de gekleurde vlakken in de overeenkomstige kwadranten
- ↳ gebruikt 4 pixels uit het originele beeld voor elke pixel van het resultaat



bicubische interpolatie: men gaat er vanuit dat de tussenliggende waarden liggen langs een Bézier-curve die bewaarde pixels verbindt, i.p.v. een rechte lijn (gebaseerd op bicubische splines)

- ↳ worden gebruikt omdat ze (net zoals curven) vloeiend samenvoegen
- ↳ gebruikt 16 pixels uit het originele beeld voor elke pixel van het resultaat
- ↳ berekeningen duren langer dan de andere twee methoden
- ↳ het resultaat is een vloeiend beeld, maar scherpe randen kunnen afgezwakt worden



Beeldcompressie

RLE (run-length encoding): het vervangen van herhalende patronen in data door het aantal herhalingen en wat herhaald moest worden

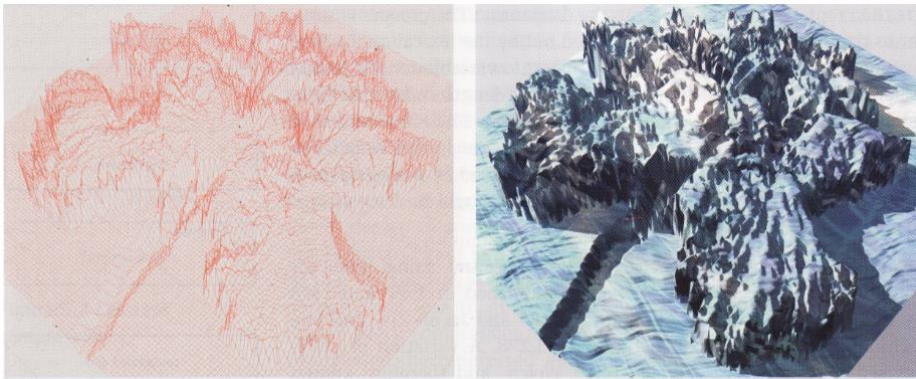
↳ lossless compressietechniek (er gaat geen informatie verloren tijdens de (de)compressie)

twee klassen lossless algoritmen:

- hercoderen van data met variabele-lengte codes
 - ↳ de meest frequente waarden bezetten de minste bits
 - ↳ dateert uit de begintijd van datacompressie en informatietechnologie
 - ↳ meest bekende algoritme: de Huffman-codering
- woordenboek-gebaseerde compressie
 - ↳ werkt door het aanmaken van een tabel/woordenboek waarin strings van bytes worden gestoken die men tegenkwam in de data
 - ↳ alle voorkomens van een string in het resultaat worden vervangen door een verwijzing
 - ↳ gebruikt codes met een vaste lengte die verwijzen naar variabele-lengte strings in het woordenboek
 - ↳ origineel twee technieken voor het aanmaken van woordenboeken:
 - LZ77 (Abraham Lempel)
 - LZ78 (Jacob Ziv)
 - ↳ LZW: compressietechniek uitgevonden door onderzoeker Terry Welch
 - variant op LZ78, met een verbeterde implementatie
 - één van de meest gebruikte compressiemethodes
 - basis van de Unix compress utility en GIF-bestanden
 - ↳ deflate compressie:
 - ontwikkeld als reactie op onpopulaire licentiepraktijken van LZW (patenten zijn ondertussen vervallen → originele motivatie is niet meer relevant)
 - combineert:
 - de techniek van het bouwen van het woordenboek uit LZ77
 - een Huffman-codering van de woordenboekingen
 - basis van de zlib-bibliotheek en gebruikt bij PNG-bestanden

JPEG-compressie

= Joint Photographic Experts Group



- lossy compressietechniek (bepaalde hoeveelheid aan dataverlies kan getolereerd worden)
- beelden kunnen gecomprimeerd worden tot 5% van hun originele grootte
- meest geschikt voor foto's en beelden met fijn detail en continue schakeringen
- werkt door een beeld te transformeren in zijn frequentiecomponenten door een discrete cosinustransformatie (DCT)
 - ↳ zet een array van pixels (*bitmap*) om in een tweedimensionale array met getallen die de amplitude van de frequentiecomponenten (*helderheid/kleurwaarden*) voorstellen
 - ↳ tweedimensionale array omdat de intensiteit kan variëren in zowel de x- als y-richting
 - ↳ maakt van het beeld een driedimensionale vorm, waarbij de frequentiecomponenten de hoogte bepalen (elke pixel geeft een z-waarde aan zijn x- en y-coördinaten)
 - ↳ is rekenkundig duur
 - onpraktisch om ineens op een volledig beeld toe te passen
 - beelden worden opgedeeld in 8x8 pixelvierkanten en apart getransformeerd
 - ↳ verandert de data in een vorm die gecomprimeerd kan worden
- vervolgens wordt de data gecomprimeerd zonder dat er waarneembare informatie verdwijnt
 - ↳ informatie over de hoogste frequenties wordt verwijderd
 - ↳ de verschillende frequenties worden gekwantificeerd in een verschillend aantal niveaus (worden gespecificeerd in een kwantificeringsmatrix)
 - ↳ reduceert de ruimte die nodig is om het beeld te bewaren op twee (lossless) manieren:
 1. veel componenten zullen na kwantificering eindigen met nulcoëfficiënten
 - worden gecodeerd volgens de run-length encoding
 - worden verwerkt d.m.v. een zig-zag sequentie om de lengte van de reeks nullen te maximaliseren: waarneembare informatie is geconcentreerd in de linkerbovenhoek, het deel rechtsonder zal vol nullen zijn
 2. minder bits zullen nodig zijn om coëfficiënten die niet nul zijn te bewaren
 - worden gecodeerd volgens de Huffman-codering
- decomprimeren gebeurt door het omkeren van het compressieproces
 1. de reeksen worden opnieuw ontplooid
 2. de Huffman-gecodeerde coëfficiënten worden gedecomprimeerd
 3. inverse discrete cosinustransformatie (IDCT) wordt toegepast
 - ↳ data terug van frequentiedomein naar het ruimtelijk domein
 - ↳ vereist evenveel tijd als de DCT
 4. geen omgekeerde kwantificering-stap (informatie is voor altijd verloren)

- graad v. compressie kan gecontroleerd worden door waarden i.d. kwantificeringsmatrix te wijzigen
 - ↳ geschikt evenwicht zoeken tussen beeldkwaliteit en compressie
- compressieartefacten: ongewenste eigenschappen in een gecomprimeerd beeld
 - ↳ de grenzen van de 8x8 vierkanten worden zichtbaar in het beeld
 - ↳ scherpe bogen (bv. tekst) worden vervaagd

Bestandsformaten

GIF (Graphics Interchange Format)

- ontwikkeld door CompuServe
- gebruikt LZW-compressie (lossless)
- geschikt voor eenvoudige, met de computer getekende beelden (bv. logo's)
- minder succesvol voor gescande en fotografische beelden (beperkt tot 256 kleuren)
- één kleur kan aangegeven worden als transparant

PNG (Portable Network Graphics)

- alternatief voor de gepatenteerde (en dus betalende) LZW-compressie
- gebruikt de deflate compressiemethode (combinatie van LZ77 en de Huffman-codering)
- niet beperkt tot 256 kleuren
- ondersteunt alpha-kanalen (gedeeltelijke transparantie)
- ontwikkeld onder toezicht van het W3C, is een ISO-standaard
- ondersteuning is slechts langzaam gekomen (geen nood aan nieuwe alternatieven)

JPEG

- geen bestandsformaat, enkel (de)compressiealgoritmen en een uitwisselingsformaat
- kan opgenomen worden in verschillende soorten bestanden (o.a. PDF en Exif)
- enigste lossy gecomprimeerde bitmapbestanden die ondersteund worden door elke browser

JFIF (JPEG File Interchange Format)

geen standaard, maar wordt bijna universeel gebruikt voor het bewaren van JPEG-data in webbeelden

SPIFF (Still Picture Interchange File Format)

- ontwikkeld door de JPEG-groep (5 jaar na de originele specificatie), ISO-standaard
- kunnen bevatten:
 - ✓ gecomprimeerde JPEG-data
 - ✓ zwart/wit-beelden (geen grijs)
- gecomprimeerd met JBIG (een gerelateerde standaard)
- onderscheid met JFIF is niet relevant (beiden zijn volledig compatibel)

EXIF (Exchangeable Image File Format)

- ontwikkeld door de Japan Electronic Industry Development Association (JEIDA)
- kan zowel JPEG- als TIFF-data opnemen, audio kan toegevoegd worden
- meestal gebruikt door digitale camera's om foto's op te slaan

JPEG2000-standaard

- JP2-formaat (mediatype image/jp2)
- JPX-formaat (extensie van JP2; mediatype image/jpx)
 - ↳ heel beperkte browserondersteuning
 - ↳ kunnen ingebed worden in andere formaten (o.a. PDF)

Andere formaten

PSD (Adobe Photoshop)

in veel gevallen het belangrijkste bitmapformaat (industriestandaard voor beeldmanipulatie)

TIFF (Tag Image File Format)

- gedetailleerd uitbreidbaar bestandsformaat
- kan full-colour bitmapbeelden opslaan
- kan gecomprimeerd zijn d.m.v. o.a. LZW en JPEG
- kan ongecomprimeerd zijn (bv. digitale camera's, + Exif-metadata)
- heel brede, maar ongelijke ondersteuning (TIFF-bestanden aangemaakt in een bep. programma kunnen soms niet gelezen worden door een ander programma)

BMP (Microsoft Windows Bitmap)

- platformafhankelijk, maar kan gelezen worden door de meeste programma's
- kan gecomprimeerd zijn d.m.v. een eenvoudige vorm van lossless compressie
- worden gewoonlijk ongecomprimeerd bewaard

PDF (Portable Document Format)

- is in staat om alles voor te stellen wat kan opduiken in een document (tekst + bitmapbeelden)
- bewaart vectorbeelden, maar bitmapbeelden worden behandeld als ondeelbare objecten
- kan gecomprimeerd worden door o.a. LZW, deflate, RLE, JBIG, JPEG en JPEG2000

RAW

- bevat de onbewerkte gegevens die uit de beeldsensor van de camera worden uitgelezen
- geen algemene standaard (elke camerafabrikant heeft eigen implementatie)

DNG (Adobe Digital Negative Specification)

- standaard voor RAW-beelden, gebaseerd op TIFF
- weinig camera's produceren rechtstreeks DNG-bestanden, maar er is een gratis conversieprogramma voorzien

H3: Kleur

Inleiding

- kleur is meestal niet noodzakelijk → zwart-wit:
 - ✓ succes van zwart/wit-fotografie en film
 - ✓ veel kleiner dan gekleurde bitmapbeelden
 - ✓ zien er grotendeels hetzelfde uit op ≠ monitoren
 - ✓ zijn ook duidelijk voor mensen die geen kleuren kunnen zien
- kleur is soms essentieel (bv. kleuren in modetijdschriften die aankoopgedrag beïnvloeden)

Kleur en wetenschap

kleur is een **subjectieve waarneming die geproduceerd wordt in de hersenen**

- ↳ digitaal manipuleren → kleurmodel nodig dat deze waarneming koppelt aan meetbare fysische fenomenen
- ↳ licht is een vorm van elektromagnetische straling → golflengte en intensiteit zijn meetbaar

spectral power distribution (SPD): beschrijving van hoe de intensiteit van licht uit een particuliere bron varieert met zijn golflengte

- ↳ opgebouwd door licht te splitsen in zijn componentgolflengtes en de intensiteit van elke component te meten
- ↳ komt nauw overeen met wat we bedoelen met kleur, maar is moeilijk om digitaal mee te werken

tristimulustheorie: elke kleur kan samengesteld worden door bep. hoeveelheden van de additieve primaire kleuren te mengen (rood/groen/blauw)

- ↳ televisieschermen/computermonitoren:
 - elke pixel bestaat uit drie fosforpunten die rood, groen en blauw licht uitzenden
 - deze worden geactiveerd d.m.v. drie elektronenstralen, één voor elke kleur
 - optisch mixen van het uitgezonden licht door de drie fosforpunten van elke pixel zal resulteren in een pixel van een bepaalde kleur
- ↳ in het oog zitten:
 - ✓ *staafjes* (nachtlicht, kunnen geen kleur onderscheiden)
 - ✓ *kegeltjes* (drie ≠ soorten, komen overeen met verschillende golflengten van licht)
- ↳ het is een vereenvoudigde theorie (problemen kunnen optreden bij o.a. printen)

RGB-kleur

RGB kleurmodel: drie waarden die de hoeveelheden rood (R), groen (G) en blauw (B) licht aangeven die moeten gecombineerd worden om licht van de gewenste kleur samen te stellen

- ↳ RGB = drie primaire SPD's
 - ✓ televisie/video: verschillende dergelijke standaarden
 - ✓ computerwereld: geen echte standaard (meer en meer kleurenspecificatie voor HDTV)
- ↳ het is niet mogelijk elke zichtbare kleur voor te stellen als combinatie van RGB-componenten
 - ✓ wel het overgrote deel van de kleuren die waargenomen kunnen worden in de wereld
 - ✓ maakt van het RGB-model een eenvoudige en efficiënte manier om kleuren voor te stellen
- ↳ RGB-waarden zijn relatief en subjectief (iedereen ziet kleuren anders)

additieve kleurenmengingen: licht van een bepaalde kleur wordt toegevoegd aan het donker

- ↳ RGB (rood, groen, blauw)
- ↳ bv. beeldschermen, lampen, scanners

subtractieve kleurenmengingen: licht wordt weggenomen van de drager/ondergrond

- ↳ CMYK (cyaan, magenta, geel)
- ↳ bv. schilderen, drukprocessen

Kleurdiepte

= het aantal bits dat gebruikt wordt om een kleurwaarde te bevatten

- 1-bit kleurdiepte: 2 kleuren (bi-level kleur, bv. zwart/wit, oranje/zwart)
- 4-bit kleurdiepte: 16 kleuren (meestal ongeschikt)
- 8-bit kleurdiepte: 1 byte → 256 kleuren (bv. monitoren voor PC's)
- 16-bit kleurdiepte (*duizenden kleuren*, gebruikt op bep. computersystemen)
 - oneven aantal bits
 - rood/blauw: 5 bits, groen: 6 bits (menselijk oog is gevoeliger voor groen licht)
- 24-bit kleurdiepte (*miljoenen kleuren*): kan meer kleuren voorstellen dan het oog kan onderscheiden
 - bv. RGB
- 30/36/48-bit kleurdiepte: meer en meer gebruikt (bv. scanners)
 - ✓ accurate benaderingen wanneer het beeld weergegeven wordt met een lagere kleurdiepte
 - ✓ extreem fijn onderscheid tussen kleuren, bv. voor effecten zoals chroma-key

= cruciale factor voor de bestandsgrootte:

- kleurdiepte reduceren van 24 naar 8 bits → bestandsgrootte afnemen met factor 3
- beeld in grijstinten zal één derde van de grootte zijn van hetzelfde beeld in miljoenen kleuren

Geïndexeerde kleur

directe kleur: de opgeslagen RGB-waarden worden rechtstreeks gebruikt om de intensiteit van de drie elektronenstralen van de monitor te sturen

geïndexeerde kleur: een palet van 256 specifieke kleuren wordt geassocieerd met een beeld

- ↳ VRAM is niet sterk genoeg om 3 bytes/pixel van een full-screen beeld bij te houden
- ↳ pixels worden bewaard als een klein getal dat een 24-bit kleur identificeert uit het palet
- ↳ het palet bevat enkel de 24-bit RGB-waarden die worden gebruikt in het beeld
- ↳ gevolg: kleinere bestandsgrootte (tot 3x)

Color Lookup Table (CLUT): kleurenpalet; geïndexeerde tabel van kleurwaarden

- ↳ koppelt logische kleuren (logische pixels) naar fysische kleuren (fysische pixels)
- ↳ opzoekingsoperatie wordt meestal uitgevoerd door de videohardware van de monitor
- ↳ weergaveprogramma moet enkel het correcte palet inladen
- ↳ wordt bewaard samen met de eigenlijke beelddata

bestandsformaten:

- PNG, BMP, TGA, TIFF → ondersteund
- GIF → enkel 8-bit geïndexeerde kleurbeelden
- SGI → enkel in een apart bestand
- JFIF, SPIFF → niet ondersteund (24-bit JPEG moet worden gereduceerd naar 8-bit bij afbeelden)
- PostScript, EPS, PDF → gesofisticeerde ondersteuning (beperkt tot ingesloten bitmapbeelden)

Wat als een bepaalde kleur van het beeld niet aanwezig is in het palet?

→ vervangen door één die zich wel op het palet bevindt

- vervangen door de CLUT-index van de dichtstbijzijnde kleur
 - ↳ kleuren kunnen uit balans gebracht worden
 - ↳ detail kan verloren gaan
 - ↳ posterisatie (gradaties van kleur worden vervangen door scherpe grenzen)
- dithering: vlakken van één kleur worden vervangen door een patroon punten van verschillende kleuren
 - ↳ optisch mengen in het oog
 - ↳ uitbreiding van het gebruik van *half-toning* (beelden in grijstinten printen op drukpersen die alleen zwart/wit kunnen produceren)
 - ↳ kan slechte resultaten produceren bij lage resoluties
 - ↳ verlies aan detail (effectieve resolutie wordt gehalveerd)

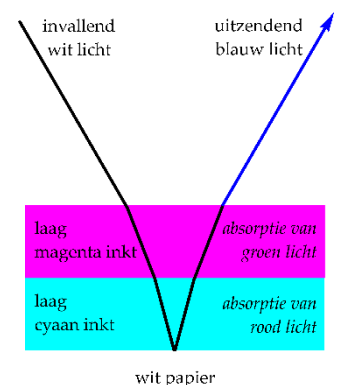
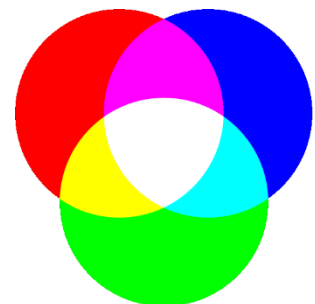
Welke kleuren moeten worden gebruikt in een palet?

- ideaal gezien: de meest belangrijke (meest voorkomende) kleuren
 - ↳ meestal automatisch bepaald door software zoals Photoshop
 - ↳ soms handmatig bepalen om te verzekeren dat alle vitale kleuren aanwezig zijn
- webveilige palet: systeempalet van 216 kleuren, ondersteund door alle browsers op elk systeem dat gebruik maakt van 8-bit kleur

Andere kleurmodellen**CMYK**

= proceskleuren bij het drukken (cyaan, magenta en geel + zwart)

- **experiment door Thomas Young:**
 - wit w gevormd door de drie additieve hoofdkleuren samen te voegen
 - de subtractieve primaire kleuren worden gevormd door twee additieve primaire kleuren samen te voegen
- $$\begin{aligned} \text{cyaan} &= G + B = W - R \\ \text{magenta} &= R + B = W - G \\ \text{geel} &= R + G = W - B \end{aligned}$$
- ↳ **voordelen:**
 - ✓ vormt de basis voor een invloedrijke theorie rond kleurenleer
 - ✓ subtractieve kleurmenging is geschikt voor drukken en schilderen (licht wordt weggenomen/geabsorbeerd)
 - ↳ **nadelen:**
 - ✓ het kleurengamma is niet hetzelfde als het RGB-gamma (kleiner, maar wel bepaalde kleuren buiten het RGB-gamma)
 - ✓ het combineren van de drie kleuren geeft geen uitstekend zwart, is niet goed voor het papier en geeft lange droogtijden
 - complementaire kleuren:
 - rood (RGB) en cyaan (CMYK)
 - geel (RGB) en magenta (CMYK)
 - blauw (RGB) en geel (CMYK)



HSB

komt beter overeen met de manier waarop we kleuren ervaren in de wereld:

- kleurschakering (hue): de mate waarin een kleur lijkt op één of een mengeling van twee van de kleuren rood, geel, groen en blauw
 - de dominante golflengte (waarop het meeste energie van licht geconcentreerd is)
 - Isaac Newton: zeven kleurschakeringen in het spectrum (rood/oranje/geel/groen/blauw/indigo/violet)
- verzadiging (saturation): maat van de puurheid van een kleur
 - mengen met wit = lagere verzadiging = productie van een *tint*
- helderheid/intensiteit (brightness): de intensiteit van het licht
 - minder licht = donkerder = productie van een *toon*

kleurenwiel: de verschillende schakeringen in een cirkel

- ↳ subtractieve primaire kleuren langs de omtrek
- ↳ additieve primaire kleuren ertussen (tegengesteld aan zijn complement)

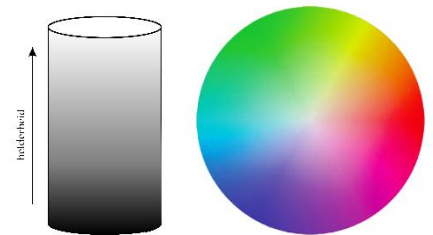


in de kunstwereld:

- ↳ primaire kleuren = subtractieve primaire kleuren
- ↳ secundaire kleuren = complementen van subtractieve primaire kleuren
- ↳ tertiaire kleuren = mengelingen van primaire en secundaire kleuren

HSB-model: uitgebreide versie van het kleurenwiel

- ↳ kleurschakeringen vormen een continuüm langs de volledige cirkel (elke schakering is een hoek tussen eigen positie en positie van rood)
- ↳ verzadiging wordt toegevoegd door een gradatie van het witte midden van de cirkel tot de pure omtrek
- ↳ helderheid wordt toegevoegd door een derde dimensie met een gradatie van de zwarte bodem tot de pure top van de cilinder



Alle kleurmodellen voorzien een manier om

- kleur te specificeren als een reeks getallen (noodzakelijk voor computerprogramma's)
- hun relaties te visualiseren in een driedimensionaal model.

Kleurruimten gebaseerd op kleurverschillen

bepaalde doeleinden → helderheidsinformatie scheiden van kleurinformatie in een beeld

- ↳ manier voor het verzenden van signalen voor kleurentelevisie en zwart/wit-ontvangers tegelijk (kleurinformatie wordt niet gedetecteerd door de zwart/wit-ontvanger)
- ↳ minder bandbreedte gebruiken voor kleur dan voor helderheid (oog is gevoeliger aan helderheid)

helderheid is gerelateerd aan RGB-waarden: Y (*luminantie/helderheid*) = $0.2125R + 0.7154G + 0.0721B$

Y'UV-kleur: elke kleurruimte bestaande uit een helderheidscomponent (Y') samen met twee componenten voor kleurverschil (U en V).

- ↳ bepaald door omvorming van hierboven vermelde formule
- ↳ gebruikt bij analoge televisie

Y'C_BC_R: variant van Y'UV die andere gewichten gebruikt voor de berekening van de kleurverschillen

H4: Video

Inleiding

late jaren '90: analoge videocamera's

- ↳ digitaliseren met speciale capture cards, heel hardware-intensief

nu: digitale video is de norm

- ↳ betaalbare digitale camcorders voor consumenten
- ↳ higher-end digitale uitrusting voor professionals
- ↳ kleine ingebouwde videocamera's in mobiele telefoons en computers
- ↳ screencasting
- ↳ videobewerkingssoftware op bescheiden krachtige systemen voor zowel amateurs als professionals

resultaat: video is een prominente eigenschap van het internet geworden

- ↳ videowebsites
- ↳ nieuwssites met ingebedde videoclips
- ↳ screencasts op websites ter ondersteuning van bepaalde software
- ↳ videochat

oorzaken:

- ↳ snelle stijging in processorsnelheden en geheugen, schijfcapaciteit en netwerkbandbreedte
- ↳ ontwikkeling van standaarden voor digitale videosignalen en interfaces
- ↳ gemak, robuustheid en flexibiliteit van digitale videobewerking t.o.v. analoge equivalent

goede kwaliteit: digitale camcorder/VTR (videotape recorder) + FireWire-interface

- ↳ 3 varianten van het DV-formaat:
 - *mini-DV* (formaat voor gewone en semiprofessionele gebruikers)
 - *DVCAM* of *DV-CPRO* (formaat voor professionele gebruikers, bv. nieuwsverslaggeving)
- ↳ varianten gebruiken hetzelfde formaat, maar verschillende graden van foutcorrectie en compatibiliteit met analoge studio-uitrusting
- ↳ ondersteunt device control (mogelijkheid om de tape te starten/stoppen/verplaatsen via en computer)
- ↳ soms interne harde schijf of rechtstreeks schrijven op DVD's
 - gebruiken DV-formaat via FireWire of MPEG-2-formaat via USB
- ↳ DV-apparatuur gebruikt meer en meer High Definition (HD)-standaarden
- ↳ DV-formaat is gecomprimeerd (dus compressieartefacten en storingen bij recompressie)
- ↳ kwaliteit is niet controleerbaar (beeld dient te voldoen aan de geschikte standaard)
 - betere kwaliteit gewenst = duurdere professionele apparatuur gebruiken

lage kost en goed gemak: MPEG-4- of 3GP-beelden + USB-interface

- ↳ o.a. goedkope videocamera's ingebouwd in mobiele telefoons en laptops, externe webcams
- ↳ kleine bestandsgrootte, maar lage kwaliteit

Videostandaarden

nieuwere digitale apparaten moeten compatibel blijven met de oude, analoge televisiestandaarden

- camera's die enkel gebruikt worden met computers zouden economisch niet praktisch zijn
- veel gebieden hebben nog geen *standard definition*-digitale televisie

Analoge videostandaarden

NTSC (US National Television Systems Committee)

- ↳ gebruikt in Noord-Amerika, Japan, Taiwan en delen van de Caraïben en van Zuid-Afrika

PAL (Phase Alternating Line)

- ↳ gebruikt in het grootste deel van West-Europa, Australië, Nieuw-Zeeland en China
- ↳ naam verwijst naar de manier waarop het signaal gecodeerd is

SECAM (Séquential Couleur avec Mémoire)

- ↳ gebruikt in Frankrijk, voormalige Sovjet-Unie en Oost-Europa
- ↳ naam verwijst naar de manier waarop het signaal gecodeerd is

raster scanning-principe: het beeld wordt van boven naar onder opgebouwd als een reeks horizontale lijnen

- ↳ In een CRT (*cathode ray tube*) worden drie elektronenbundels (elke additieve primaire kleur) gezonden/afgebogen door een magnetisch veld
- ↳ zweven over het scherm, trekken een lijn en bewegen naar beneden om de volgende lijn te trekken
- ↳ hun intensiteit wordt gewijzigd volgens het inkomende signaal
- ↳ meer dan een halve eeuw gebruikt
- ↳ door de nawerking van het oog zien we het als één vloeiend beeld

scherm moest 40 keer/seconde ververs worden → toen nog niet genoeg bandbreedte

- ↳ elk frame werd opgedeeld in twee *velden* (even en oneven genummerde lijnen van elk frame)
- ↳ worden 1 na 1 verzonden, elk frame is opgebouwd door het samenvoegen van velden (*interlacing*)
- ↳ een snel bewegend object zal veranderen van positie tussen de twee velden (*de-interlacing* nodig)
- ↳ snelheid waarmee de velden werden verzonden = lokale AC-lijnfrequentie
 - ✓ PAL: 50 velden/seconde → 25 fps
 - ✓ NTSC: 60 velden/seconde bij zwart-wit, 59.94 velden/seconde bij kleur (interferentie met geluid) → 29,97 fps
- ↳ afspelen op computermonitor = *progressive scanning* (lijnen w. geschreven naar framebuffer)
 - ✓ geen interlacing of flikkering
 - ✓ lagere framesnelheden kunnen gebruikt worden

naast beelddata worden ook synchronisatie- en andere informatie meegezonden in elk frame

- ↳ NTSC: 525 lijnen, 480 ervan zijn beeld
- ↳ PAL/SECAM: 625 lijnen, 576 ervan zijn beeld

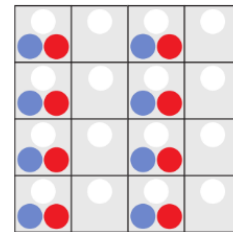
Digitale videostandaarden

minder complexe situatie dan bij analoge video

- compatibiliteit met bestaande uitrusting:
 - 625/50 en 525/59.94-standaarden
 - HDTV-standaarden
- pogingen om de twee huidige formaten te verenigen
- mislukt, ≠ standaarden voor consumenten/professionals

analoge video w. bij omzetting naar digitale vorm bemonsterd volgens officiële standaard **CCIR 601**:

- bemonsterd in beide richtingen (video = 2D)
- scanlijnen in een verticale schikking
- horizontaal bemonsteringsbeeldformaat
- bestaat uit:
 - 720 luminantiesamples
 - 2 reeksen van 360 kleurenverschilsamples per lijn (volgens $Y'C_B C_R$)
- PAL → 720x576, NTSC → 720x480
 - ↳ PAL/NTSC = analoge standaarden → lijnen in een frame worden gegenereerd door een continu signaal i.p.v. opbreking in pixels
 - ↳ aantal pixels in een lijn = aantal beeldlijnen (576 of 480) * de hoogte-breedteverhouding (4:3)
- pixels zijn niet vierkant
 - ↳ 625 lijnsystemen → lichtjes breder dan hoog
 - ↳ 525 lijnsystemen → lichtjes hoger dan breed
- chrominance subsampling: minder samples voor kleurverschilwaarden als voor luminantie
 - ↳ om de bestandsgrootte te reduceren
 - ↳ oog is minder gevoelig voor variaties in kleur dan voor variaties in helderheid
- schikking van samples wordt 4:2:2-bemonstering genoemd
 - ↳ in elke lijn:
 - 4 Y-samples
 - 2 B - Y-samples
 - 2 R - Y-samples
 - ↳ kleurbemonstering is co-sited (alle 3 de waarden worden bemonsterd op hetzelfde punt)
- andere schikkingen voor bemonstering:
 - 4:1:1-bemonstering: elke vierde pixel op elke lijn wordt bemonsterd voor de kleur
 - 4:2:0-bemonstering: kleurwaarden w. gesubsampled met factor 2 in beide richtingen



DV en MPEG

- **bemonstering** produceert een digitale voorstelling van een videosignaal
- andere standaarden zijn nodig voor **compressie/formaat**: DV en MPEG-familie
 - ↳ gebaseerd op $Y'C_B C_R$
 - ↳ gescand overeenkomstig CCIR 601
 - ↳ verdere chrominance subsampling (hangt af van videostandaard: PAL = 4:2:0, NTSC = 4:1:1)
- niet enkel standaarden: ook onofficiële/open codecs en bestandsformaten

DV-standaarden: digitale videouitrusting voor consumenten, semi-prof. gebruik en nieuwsgaring

- DV en hoofdvarianten (DVCAM/DVPRO) gebruiken:
 - zelfde compressiealgoritme
 - zelfde datastroom (datastroom van 25 Mbits/sec.)
 - zelfde compressieverhouding van 5:1
- DVPRO/Digital-S → 4:2:2-bemonstering
 - betere kwaliteit
 - hogere bitsnelheden
 - professioneel gebruik
- DV → 4:1:1-bemonstering
- HDDV: HD-versie van DV, geschikt voor low-budget film-making

MPEG-standaarden: versch. ISO-standaarden geproduceerd door het ISO/IEC Motion Picture Experts Group

- MPEG-1: hoofdzakelijk bedoeld voor het Video CD-formaat (basis voor volgende standaarden)
- MPEG-2: eerste generatie digitale studiouitrusting, TV en DVD
- MPEG-4: ambitieuze standaard ontworpen om een groot bereik aan mediadata te ondersteunen
 - ↳ bitsnelheden van 10 kbits/sec. tot 300+ Mbits/sec.
 - ↳ toepassingen van mobiele telefoons tot HDTV
 - ↳ onderverdeeld in 23 delen:
 - ✓ audiocompressie
 - ✓ videocompressie (delen 2 en 10)
 - ✓ afleveren van data over netwerken
 - ✓ bestandsformaten
 - ✓ etc.
 - ↳ MPEG-4 Part 2: een verfijning van MPEG-2
 - gewoonlijk naar verwezen als MPEG-4 video
 - extra compressietechnieken → kleinere bestanden aan dezelfde kwaliteit
 - ↳ MPEG-4 Part 10: gekend als H.264/AVC (Advanced Video Coding)
 - verdere verfijning van compressietechnieken
 - één van de leidende compressietechnieken voor webvideo
 - gebruikt voor de tweede generatie HD (Bluray) DVD's
 - ↳ MPEG-4 bestandsformaat:
 - gebaseerd op het QuickTime-formaat
 - gebruikt om gecomprimeerde video + audio + metadata op te slaan
 - ↳ 3GP-bestandsformaat: vereenvoudigde versie van het MP4-formaat
 - ondersteunt compressie volgens MPEG-4 Part 2 en H.264/AVC

elk van de MPEG-standaarden definieert **verschillende profielen en levels**

- profiel: deelverzameling eigenschappen van de datastroom
 - ↳ kan geïmplementeerd worden op één of meerdere levels
 - ↳ niet elke combinatie van level en profiel is gedefinieerd
- level: definieert bepaalde parameters
 - ↳ bepalen o.a. maximum framegrootte, datasnelheid en chrominance subsampling
 - ↳ 16 verschillende levels (bv. BP@L1: 64 kbps, 176 x 144 px, 15 fps)
 - ↳ elk level heeft 2 of meer sublevels (bv. HP@L5.1: 300 Mbps, 4096 x 2048, 30 fps)
 - ↳ bv.
- meest bekende in MPEG-2: Main Profile at Main Level (MP@ML)
 - ↳ CCIR 601
 - ↳ 4:2:0 chrominance subsampling
 - ↳ datasnelheid van 15 Mbits/sec.
 - ↳ gebruikt de meest uitgewerkte voorstelling van gecomprimeerde data door MPEG-2
 - ↳ gebruikt voor digitale televisieuitzendingen en DVD-video
- meest bekende in H.264/AVC:
 - Baseline Profile (BP): geschikt voor video-conferencing en mobiele apparaten met beperkte rekenmogelijkheden
 - Extended Profile (XP): geschikt voor streamen van video
 - Main Profile (MP): geschikt voor algemeen gebruik
 - High Profile (HIP): gebruikt voor HDTV en Bluray (vervangt MP voor TV-uitzendingen)

High Definition-formaten

Standard Definition (SD): dezelfde beeldgrootte als analoog, even weinig detail, maar minder ruis

- ↳ resoluties van PAL en NTSC
- ↳ Enhanced Definition: 480p, progressief gescand SD-frame

High Definition (HD): alles met grotere frames dan SD

- ↳ verschillende globale HD-standaarden voor uitzendingen
- ↳ één globale beeldverhouding, nl. 16:9
- ↳ acht mogelijke combinaties van de verschillende variabelen:
 - ✓ framehoogte (genoeg om de resolutie te specificeren)
 - ✓ framesnelheid (25 of 30 fps, overeenkomstig de SD-standaarden) of veldsnelheid
 - ✓ een paar interlacede velden of één enkel progressief gescand frame
- notatie: framehoogte + frame-/veldsnelheid + i (interlaced) / p (progressief)
- bv. 720 25 p → 1280 x 720 px, 25fps, progressief gescand
- bv. 1080 60i → 1920 x 720 px, interlaced op 60 velden/sec. (realiteit: 59.94 zoals SD NTSC)
- ↳ vereist passende apparatuur voor opname/transmissie/ontvangst/weergave
- ↳ heeft zijn eigen tapeformaten (o.a. HDCAM, DVCPRO-HD)
- ↳ heeft zijn eigen optische media (Blu-ray DVD)
- ↳ vereist meer schijfruimte, bandbreedte en verwerkingskracht dan SD-video

Videocompressie

twee soorten (kunnen apart/samen gebruikt worden):

- spatial (intra-frame) compression: elke individuele frame wordt apart gecomprimeerd m.b.v. bitmap-compressietechnieken
 - ↳ onderscheid tussen lossless/lossy technieken:
 - lossless → niet effectief (behalve computergegenereerd materiaal)
 - lossy
 1. film wordt opgenomen/geselecteerd
 2. video wordt gedecomprimeerd voor postproductie
 3. video wordt opnieuw gecomprimeerd voor aflevering
- temporal (inter-frame) compression: deelsequenties w. gecomprimeerd door enkel de verschillen ertussen te bewaren
 1. keyframes worden geplaatst op constante intervallen in de sequentie
 - ↳ interval kan op voorhand bepaald worden
 - ↳ blijven ongecomprimeerd of worden enkel spatial gecomprimeerd
 2. tussenliggende frames worden vervangen door verschilframes die enkel de verschillen met de vorige frame/keyframe bewaren
 - ↳ verschilframe zal veel minder informatie bevatten dan normale frame (veel types videomateriaal bestaan uit grote relatief statische vakken)

symmetrische codec: compressie en decompressie nemen evenveel tijd in

asymmetrische codec: compressie en decompressie nemen niet evenveel tijd in

- ↳ meestal → compressie duurt veel langer dan decompressie
- ↳ omgekeerd → codec is nutteloos (afspelen gebeurt op redelijk hoge framesnelheid)

Spatial compressie

MJPEG-compressie (motion JPEG): het comprimeren van video door JPEG-compressie toe te passen

- ↳ videoframes worden bewaard met Y'C_BC_R kleur met chrominance subsampling
- ↳ JPEG-compressie kan rechtstreeks toegepast worden op deze data om voordeel te halen uit de bestaande compressie van de subsampling
- ↳ was populair voor analoge video/digitale fotocamera's met eenvoudige videomogelijkheden

nu: belangrijkste technologie die alleen spatial compressie gebruikt is **DV**

- ✓ gebruikt zelfde technieken als JPEG-compressie
 1. blokken van 8x8 px uit elk frame worden getransformeerd m.b.v. de DCT
 2. kwantificering
 3. run-length- en Huffman-codering langs een zigzag-sequentie
- ✓ gebruikt enkele slimme trucs om hogere beeldkwaliteit te bekomen
 - DCT wordt toegepast op de 64 pixels in elk blok
 - (bijna) statisch frame, geen verschil tussen elk veld
 - toegepast op het volledige 8x8-blok (opgebouwd uit afwisselende lijnen van de even en oneven velden)
 - veel beweging, velden verschillen
 - blok w. opgesplitst in twee 8x4-blokken en onafhankelijk getransformeerd
 - leidt tot efficiëntere compressie bij frames met beweging
 - ↳ de compressor bepaalt of er al dan niet beweging is tussen de frames:
 - ✓ d.m.v. motion compensation (beschreven onder MPEG)
 - ✓ door beide versies v.d. DCT te berekenen en het kleinste resultaat te kiezen
 - shuffling: uitgebreid proces van herschikking, toegepast op de blokken, om beschikbare ruimte optimaal te gebruiken
 - ↳ per seconde moeten alle beschikbare bytes (25 Mbits) verdeeld worden onder de verschillende blokken
 - ↳ blokken met veel nulcoëfficiënten (door transformatie) hebben minder bytes nodig
 - ↳ elk videosegment wordt geconstrueerd door 8x8-blokken te nemen uit 5 verschillende vlakken van het frame
 - ↳ zo wordt de hoeveelheid detail evenredig verdeeld in elk videosegment

Temporal compressie

- gebruikt door alle moderne videocodecs (Windows Media 9, Flash Video, delen van MPEG-4)
- oorspronkelijk gebruikt door MPEG-1

MPEG-1 definieert geen compressiealgoritme, enkel een syntax voor datastroom en een decompressor

- ↳ mogelijkheid tot competitie op de markt openlaten
- ↳ in praktijk (impliciet) redelijk grondig gedefinieerd: ze combineert
 - ✓ temporal compressie (gebaseerd op motion compensation)
 - ✓ spatial compressie (gebaseerd op kwantificering en het coderen van frequentiecoëfficiënten uit de DCT)

de waarde van elke pixel in een frame wordt afgetrokken van de overeenkomstige pixel in het vorige frame

- ↳ resultaat: verschilframe met veel nulpixels op de plaatsen waar geen verschil is tussen de frames

motion compensation: de techniek van het opnemen van de relatieve verplaatsing van objecten in de verschilframes (bv. motion vectors)

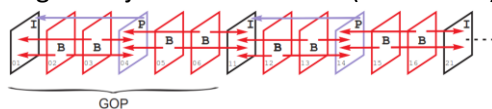
1. elke frame wordt verdeeld in macroblokken van 16x16px
2. de plaats van een macroblok in het volgende frame wordt voorspeld
3. verschilframes worden geconstrueerd door elke macroblok af te trekken van zijn voorspelde tegenhanger
 - ↳ voordeel: heel veel nulpixels en een kleiner verschilframe na spatial compressie
 - ↳ nadeel: motionvectoren moeten bewaard worden (beschrijven de voorspelde verplaatsing)
 - worden gecompriemd door verschillen tss. nabije motionvectoren te bewaren

een video kan gecodeerd worden in gecompriemde vorm als een **sequentie van I-, P- en B-beelden**

- I-beelden (intra-): MPEG keyframes; worden puur spatial gecompriemd
- P-beelden (predictive-): verschilframes; kunnen gebaseerd zijn op een vorig I- of P-beeld
- B-beelden (bi-directionally predictive-): frames die voorspeld worden uit volgende frames
 - ↳ achterwaartse voorspelling
 - ↳ kunnen motion compensation gebruiken uit de volgende I- of P-beelden

encoders gebruiken typisch een herhalende sequentie, gekend als een **GOP (Group Of Pictures)**

- ↳ begint altijd met een I-beeld (bv. IBBPBB)



- ↳ alle drie types beelden worden gecompriemd m.b.v. de MPEG-1-versie van JPEG-compressie
- ↳ bij de keuze van een patroon van een GOP moet een evenwicht gevonden worden tussen compressie en rekenkundige complexiteit
- ↳ voldoende I-beelden moeten opgenomen worden om willekeurige toegang toe te laten op verschillende frames per seconde
- ↳ moet van de MPEG-1 specificatie geen regulier patroon vormen → frequentie van I-beelden worden aangepast naargelang de aard van de video

volgorde van weergave: sequentie beelden in de eigenlijke framevolgorde

- ↳ probleem → bepaalde informatie die nodig is om het overeenkomstige frame te reconstrueren, zit in een I- of P-beeld dat later in de sequentie voorkomt

bitstroomvolgorde: sequentie beelden in geschikte ordening voor transmissie

- ↳ voorspelde frames komen later in de sequentie dan de beelden die er afhankelijk van zijn

SIF (Source Input Format): het klassieke formaat voor MPEG-1 video, maar ook voor andere framegroottes en -snelheden

besluit: MPEG-(de)compressie zijn rekenkundig dure taken, met grote bestanden tot gevolg

- ↳ kon initieel enkel afgespeeld worden met speciale hardware
- ↳ kon later afgespeeld worden door software alleen door de vooruitgang in processorsnelheden

H5: Geluid

Inleiding

- complexe mengeling van fysische/psychologische factoren, is moeilijk nauwkeurig te modelleren
- je hebt het niet altijd nodig (kan soms zelfs storend zijn)
- gebruikers nemen het niet altijd waar (bv. oude computers zonder geluidskaart)
- het is bedachtzaam om altijd een bep. alternatief te voorzien
- twee speciale types geluid: muziek en spraak
 - ↳ meest gebruikte types in multimediateproducties
 - ↳ vervullen een speciale (culturele/linguïstische) rol

De aard van geluid

Werking van een stemvork:

1. Stemvork strijken over hard oppervlak → vibreren op een precieze frequentie
 2. Lucht wordt samengedrukt en dunner in de tijd met de vibraties
 3. Door interactie met naburige luchtmoleculen planten deze vibraties zich voor als golven
 4. De geluidsgolf bereikt het oor → trommelvlies vibreert op dezelfde frequentie
 5. De vibratie wordt doorgezonden en omgezet in zenuwimpulsen
 6. Deze zenuwimpulsen interpreteren we als het geluid van een stemvork
- alle geluiden worden geproduceerd door de conversie van energie naar trillingen in de lucht of een ander elastisch medium
 - de meeste geluidsbronnen vibreren op een meer gecompliceerde manier dan stemvorken
 - we beschrijven geluid in functie van zijn frequentiespectrum (de amplitudes van zijn frequentiecomponenten)
 - het menselijk oor is in staat frequenties te detecteren tussen 20Hz en 20kHz
 - ↳ verschilt van individu tot individu
 - ↳ vermindert vrij snel met toenemende leeftijd

Compressie

moeilijk lossless te comprimeren door de complexe en onvoorspelbare aard van geluidsgolven

- ↳ Huffman-codering kan effectief zijn wanneer de amplitude valt onder het hoogste niveau dat voorgesteld kan worden in de gebruikte bemonsteringsgrootte
- ↳ signaal kan dan voorgesteld worden in een kleinere bemonsteringsgrootte

in het algemeen → lossy comprimeren

- ↳ verwijderen van stilte (alle samples onder een bep. waarde) d.m.v. run-length encoding
- ↳ hoge frequenties kunnen niet verwijderd worden zoals bij beeldcompressie (zijn belangrijk bij audio)

Spraakcompressie

telefoonbedrijven sinds de jaren '60 → gedwongen om effectieve compressietechnieken te ontwikkelen

companding: *compressing/expanding*; niet-lineaire kwantificeringsniveau's worden gebruikt

- ↳ hogere niveaus worden verder uit elkaar geplaatst dan de lagere
- ↳ rustige geluiden worden in groter detail voorgesteld dan de luidere
- ↳ komt overeen met de manier waarop we verschillen in volume ervaren
- ↳ expanding → invers proces van expansie is vereist wanneer het signaal gereconstrueerd wordt

ITU Recommendations: definieert de belangrijkste companding-functies voor gebruik in telecommunicatie

- ↳ μ-law (*Recommendation G.711*): gebruikt in Noord-Amerika en Japan voor
 - compressie van audio op Sun- en NeXT-systemen
 - kan een dynamisch bereik van 12 bits terugbrengen tot 8 bits (bij 8kHz bemonstering)
- ↳ a-law: gebruikt in de rest van de wereld

Adaptive Differential Pulse Code Modulation (ADPCM): bewaart het verschil tussen opeenvolgende monsters in plaats van een absolute waarde

- ↳ gebaseerd op inter-frame videocompressie
- ↳ DPCM berekent een voorspelde waarde voor een sample, gebaseerd op vorige samples, en bewaart het verschil tussen de voorspelling en de eigenlijke waarde (goede voorspelling = klein verschil)
- ↳ Adaptive DPCM bekommt extra compressie door dynamisch de stapgrootte te variëren gebruikt om de gekwantificeerde verschillen voor te stellen
 - hoeveelheid detail schaalte mee met de grootte van het verschil
- ↳ ITU Recommendation G.271: vorm van ADPCM voor gebruik in telefonie (16/32 kbps)

Linear Predictive Coding: gebruikt een wiskundig model als voorstelling van spraak

- ↳ zendt parameters die de toestand van het model beschrijven (i.p.v. audiosamples)
- ↳ kan verzonden worden op snelheden lager dan 2.4 kbps
- ↳ geluid heeft een machineachtige kwaliteit

Perceptuele (waarnemingsgebaseerde) compressie

geheim van effectieve lossy compressie = **onhoorbare geluiden verwijderen**

- ↳ drempel van het gehoor: het minimum niveau waarbij een geluid gehoord kan worden
 - geluid met lage/hoge frequentie moet luider zijn dan een middentoon om gehoord te worden
- ↳ het algoritme moet een psycho-akoestisch model gebruiken
 - een wiskundige beschrijving van de manier waarop het oor en de hersenen geluiden ervaren

masking: luide tonen verbergen zachtere tonen die voorkomen op hetzelfde ogenblik

- ↳ = een wijziging van de curve van de drempel van het gehoor in de buurt van de maskerende toon
- ↳ *masking curve*: de gestegen portie (is niet-lineair en asymmetrisch)
- ↳ kan gebruikt worden om kwantificeringsruis te verbergen

implementatie:

1. het signaal wordt in (vaak 32) banden van frequenties gesplitst
2. het gemiddelde signaalniveau in elke band wordt berekend
3. masking niveau w. berekend voor elke band (m.b.v. deze waarden en een psycho-akoestisch model)
4. signaal in een band valt volledig onder zijn masking niveau → de band wordt verwijderd
5. anders → signaal wordt gekwantificeerd m.b.v. het minst aantal bits die nodig zijn om kwantificeringsruis te maskeren

best werkende algoritmes: MPEG-1 en MPEG-2 (gebruikt op zichzelf/in combinatie met video)

- ↳ audiodeel van MPEG-1: drie lagen van audiocompressie
 - ✓ gebaseerd op bovenstaande principes
 - ✓ coderingsproces wordt complexer van Laag 1 naar 3, datasnelheid neemt af
- ↳ MPEG-1 Laag 3 (MP3): compressieverhoudingen van ong. 10:1, hoge kwaliteit behouden
 - ✓ hogere bitsnelheden kunnen gebruikt worden, geven een betere kwaliteit
 - ✓ lagere bitsnelheden kunnen gebruikt worden, bv. streaming (64kbps stereo = FM radio)
 - ✓ codering met VBR (variabele bitrate) is mogelijk
 - verandert de bitrate zodat passages die niet eenvoudig comprimeren, gemakkelijk kunnen gecodeerd worden op een hogere snelheid dan deze die dat wel doen
- ↳ audiodeel van MPEG-2:
 - ✓ codering is identiek met MPEG-1 audio + bepaalde uitbreidingen i.v.m. surround sound
 - ✓ definieert een nieuwe audiocodec: AAC (Advanced Audio Coding)
 - niet achterwaarts compatibel met vroegere MPEG-standaarden of lagere lagen
 - zo in staat om hogere compressieverhoudingen te bekomen op lagere bitsnelheden
 - AAC is superieur aan MP3 op dezelfde/lagere bitsnelheden