BVA4 – Bildverarbeitung in Medizin und Biologie

GRUNDLAGEN DIGITALER BILDER Gerald Zwettler

Software Engineering, Bildverarbeitung und Computergrafik School of Informatics, Communications and Media gerald.zwettler@fh-hagenberg.at

> University of Applied Sciences Upper Austria FH OÖ Studienbetriebs GmbH

> > Softwarepark 21

4232 Hagenberg/Austria

Tel.: +43 (0)5 0804-22038 Fax: +43 (0)5 0804-22099

auf Basis der BVA5 Vorlesung von Prof. Werner Backfrieder

Digitale Signale

- Digitale 1D Signale i.d.R. Output von Messinstrumenten, z.B. Temperatur, EKG, EEG, Geschwindigkeit.
- 1D Signale oft als Funktion der Zeit interpretiert I (t)
- Durch Abtastung sind die Zeitpunkte, an denen Messwerte vorliegen, diskret: $t_{i} = T_{o} + i * \Delta t$, wobei im Falle einer Repräsentierung als Array damit i als Index zu verwenden ist.

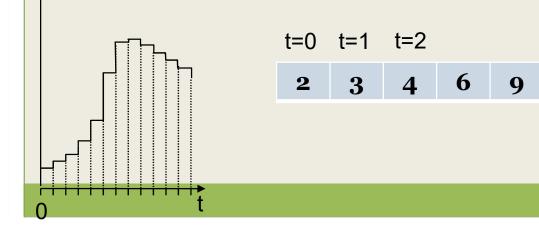
20

21

19

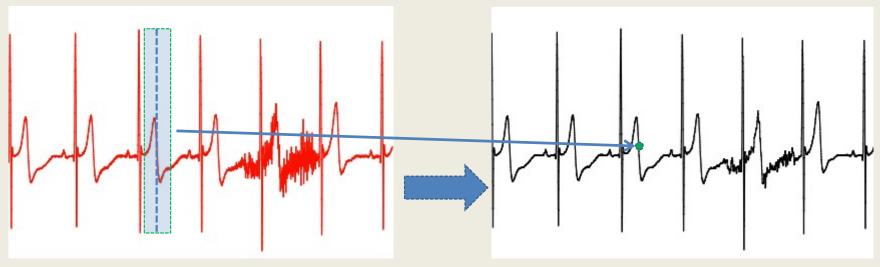
18

• Gängige Repräsentierung: 2D Chart



Digitale Signale

- Digitale Signale i.d.R. mit Rauschen behaftet → Anwendung von digitalen Filtern bei schlechtem Signal-Rauschverhältnis notwendig
 - o Symmetrischer Filter-Kernel nur anwendbar, wenn keine live Analyse zu erfolgen hat. Ansonsten kann Bearbeitung nur anhand der Historie stabilisiert werden im Sinne eines *weighted moving average*

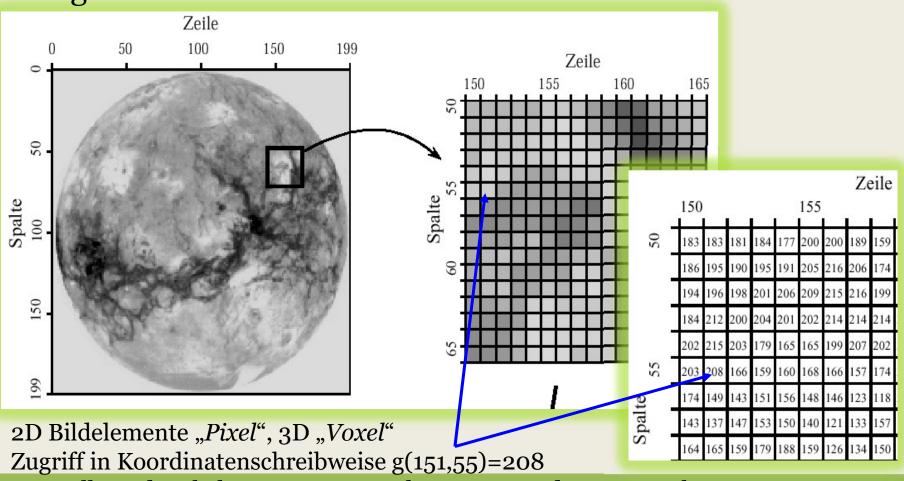


Median Filter, r=15

Digitale Bilder

4

• Digitale Bilder als 2D Masken von skalaren Messwerten



Darstellung der skalaren Werte z.B. als Grauwert oder RGB Farbwert

Mathematische Definition 2D Bilddaten

5

Digitale 2D Bilddaten repräsentieren eine zweidimensionale Matrix skalarer Messwerte, die als Graustufen oder Farben (z.B. RGB oder HSV) interpretiert werden. Ein ganzzahliges 2D Bild I stellt dabei eine Abbildung $f: \mathbb{N}^2 \to \mathbb{N}$ dar. Das Bild I ist folglich eine Funktion $I(x,y) \in \mathbb{N}$ mit $x, y \in \mathbb{N}$, die vom Koordinatenraum $[0,\ldots,N_x] \times [0,\ldots,N_y] \to \{0,1,\ldots,g-1\}$ auf den skalaren Wertebereich $g \in \mathbb{N}^+$ abbildet. Die maximale Ausdehnung N_x und N_y der beiden Dimensionen wird dabei als Breite (width) und Höhe (depth) bezeichnet.

Die Koordinaten eines Bildes I definieren sich als Menge an zweidimensionalen Tupel $\{c_i\}$ bezeichnet als \mathcal{C} , mit $c_i = \{c_x, c_y\}$. Das Bild I kann an allen diskreten Positionen c_i , mit $c_i \in \mathcal{C}$, ausgewertet werden, sodass der skalare Werte mit $I(c_i) = I(c_{i_x}, c_{i_y})$ retourniert wird. Ein Teilbild bzw. eine Teilregion \mathcal{R} definiert sich als Teilmenge $\mathcal{R} \subseteq \mathcal{C}$ in Bezug auf die Koordinaten.

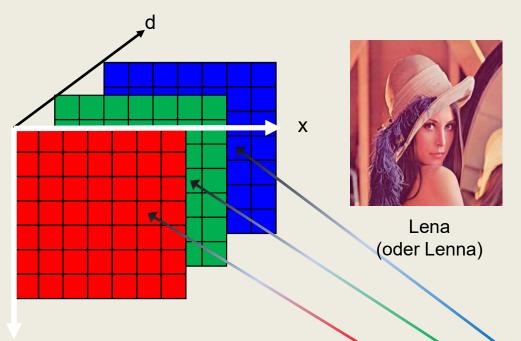
- gängige Dimensionen der Bilddaten z.B. 512 oder 1024 (Medizin) bzw. 4096+ (digitale Fotografie)
- skalarer Wertebereich 8bit [0;255] oder 12bit [-1024;3071] (Medizin) bzw.
 32bit+ (digitale Fotografie), negativer Wertebereich wird via Offset realisiert.
- Bilder als Funktionen und Koordinaten-Mengen

Farbbilder

6

• mehrere Kanäle (z.B. RGB) für jede Pixel-Position anstatt

eines einzelnen skalaren Wertes



- Zugriff Position g(5,4) = (255, 100, 0)
- Zugriff Farbkanal g(5,4,1)=100

У



Umrechnung Farbbilder in Graustufen



Mittelwert-Methode (average):

$$\circ$$
 g = (R+G+B) / 3

Helligkeits-Methode (lightness):

$$\circ$$
 g = (MAX(R, G, B) + MIN(R, G, B)) / 2

• Gemäß Luminosität (luminosity):

$$\circ$$
 g = 0.21R + 0.72G + 0.07B

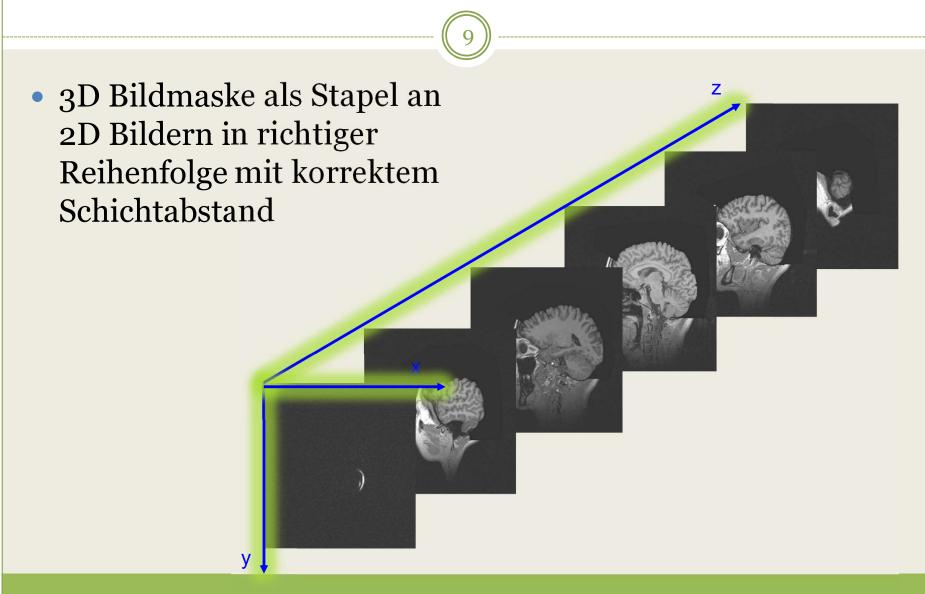
Mathematische Definition 3D Bilddaten

8

Dreidimensionale Bilddaten sind eine Erweiterung von 2D Bilddaten um die Dimension Tiefe (depth) z. Das 3D Volumen ist dabei ein Mapping $f: \mathbb{N}^3 \to \mathbb{N}$ des Koordinatenraums auf den skalaren Wertebereich als $[0, \ldots, N_x] \times [0, \ldots, N_y] \times [0, \ldots, N_z] \to \{0, 1, \ldots, g-1\}$. 3D Volumen definieren folglich die Zugriffsfunktion $I(x, y, z) \in \mathbb{N}$ mit $x, y, z \in \mathbb{N}^+$. Die Koordinaten ergeben sich als Tupel $\{c_i\}$ mit $c_i = \{c_x, c_y, c_z\}$.

- 3D Bilddaten werden entweder als komplette Volumen in einer einzigen Datei repräsentiert (z.B. DICOM Format), oder setzen sich aus einer Sequenz von einzelnen 2D Dateien zusammen (z.B. JPEG, PNG, DICOM). Dabei muss die Reihenfolge bzw. der Schichtabstand bekannt sein.
- Für Filme, Animationen und medizinische Anwendungen gibt es eine Erweiterung auf 4D Daten (3D + t), die eine Zeitserie an bewegten 3D Volumina repräsentieren. Gleichsam, ist etwa auch bei der 2D Cine-Angiografie die dritte Dimension als Zeitachse zu verstehen.

3D Bilddaten



Repräsentierung von 2D und 3D Bilddaten

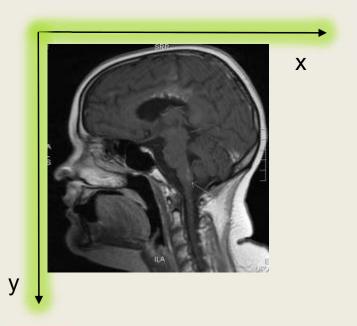


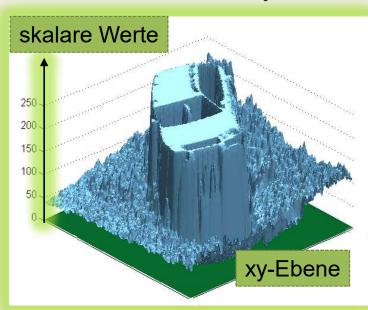
- Neben den eigentlichen skalaren Daten, die meist im RAW-Format, d.h. als Byte-Sequenz abgespeichert werden, sind Metainformationen zur korrekten Darstellung erforderlich:
 - Länge Header → Beginn des Datenblocks
 - **Skalare Wertdarstellung**: Datentyp (int, unsigned char) und Byte-Order (Little / Big Endian) bzw. optional Offset für negative Werte
 - \circ **Abmessungen** (in x/y/z)
 - o **Spacing** (Maskengröße) für proportionsgetreue Wiedergabe
 - Ursprung und Orientierung des Koordinatensystems (vgl. row/column major order bzw. links/rechts-händiges Koordinatensystem)
- Bei Bildformaten wie DICOM, png, gif, tiff, jpeg, bmp findet sich diese Information am Beginn in einem Header-Block, bzw. werden sie bei klassischen RAW-Formaten in einem eigenen Header-File bereitgestellt (Analyze mit *.hdr und *.img)

Skalarwerte als Topografie



- Die Repräsentierung der skalaren Werte als Graustufen oder RGB Farbwerte stellt nur eine Möglichkeit zur Visualisierung dar.
- Die skalaren Werte aus 2D Bilddaten können ferner etwa auch als Topografie interpretiert werden, wobei die skalare Bandbreite als Höheninformation interpretiert wird.
- Das 2D Bild wir folglich als diskrete Funktion auf einer xy-Ebene aufgefasst.

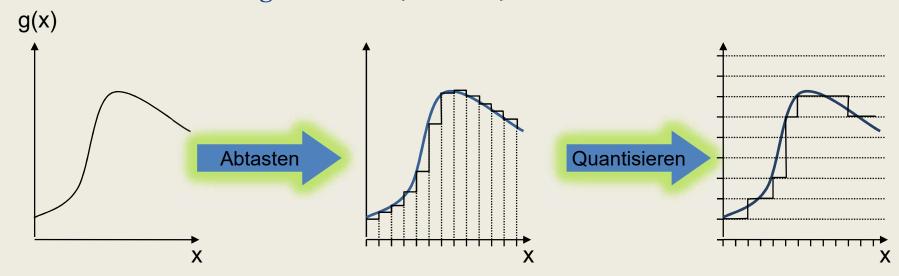




Diskretisierung bei digitalen Signalen



- gegeben: kontinuierliche Funktion bzw. analoges Signal g
- gesucht: diskrete Funktion bzw. digitales Signal.
 - **Abtasten**: die Funktion wird an diskreten und i.d.R äquidistanten Punkten abgetastet
 - **Quantisierung**: der kontinuierliche Wertebereich wird auf endlich viele Werte eingeschränkt (Runden)



Diskretisierung bei digitalen Signalen



- Diskretisierung und Quantisierung wird im AD-Wandler vollzogen.
- Bei der Quantisierung wird der kontinuierliche Wertebereich auf einen diskreten Wert zur Basis 2 konvertiert, z.B.
 - o 2^8 für 256 Werte, [0;255] bei unsigned byte
 - o 2^16 für 65536 Werte, [0;65535] bei unsigned short (2 bytes)
- Zur Quantisierung muss die mögliche Bandbreite der analogen Signale bekannt sein (min,max) → dann Form der Normierung
- Auch negative skalare Werte denkbar (vgl. HU bei CT) oftmals jedoch via Offset modelliert

Signaltheorie I



- Wellenlänge = Größe einer (periodischen) Struktur
 - o bei Bilddaten: Einheit "Pixel" (2D) bzw. "Voxel" (3D)
- Ortsfrequenz = Anzahl der Wiederholungen einer (periodischen) Struktur pro Einheitslänge
 - o bei Bilddaten: Wiederholungen pro "Pixel" bzw. "Voxel"

Wellenlänge
$$\lambda$$
Ortsfrequenz $f = \frac{1}{\lambda}$

Signaltheorie II



• die kleinste periodische darstellbare Struktur bei Bildern hat die Wellenlänge == Größe $\lambda_{min} = 2$ Pixel

$$f_{\text{max}} = \frac{1}{\lambda_{\text{min}}} = \frac{1}{2}$$
 Wellen pro Pixel

 die größte im Bild darstellbare Struktur hat die Größe der jeweiligen Dimension im Bild, z.B. Bildkantenlänge 256

$$f_{\min} = \frac{1}{\lambda_{\max}} = \frac{1}{256}$$
 Wellen pro Pixel

Signaltheorie III



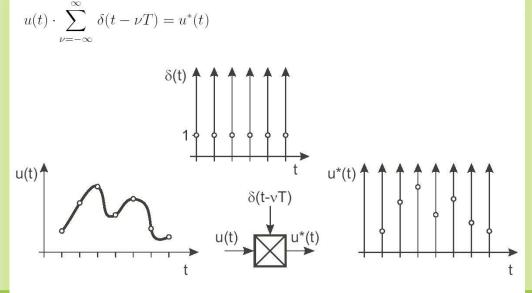
- die Anzahl der Wiederholungen einer periodischen Struktur wird als **Wellenzahlindex u** bezeichnet
 - Verwendung bei Bildern statt Frequenz, da Werte >=1
- $oldsymbol{u}$ die maximale Wellenzahl bei gegebener Abtastrate $oldsymbol{u}_{max}$ heißt $oldsymbol{Grenzwellenzahl}$
- in einem 2D Bild der Größe 512x512 kann eine periodische Struktur höchstens 256 mal wiederholt werden
 - o es ergeben sich folglich u_{max}=256, u_{min}=1

Signaltheorie IV



- Abtastung von Signalen zu äquidistanten Zeitpunkten
 - O Abtastzeit T, diskrete Positionen v
 - o Abtastzeitpunkte **t=vT**, mit **v**=0,1,...,n
 - \circ es ergibt sich die Abtastfrequenz $\omega_{\rm T} = 2\pi/{\rm T}$
 - Abtastung beim A/D-Wandler mittels Dirac-Kamm-Funktion

δ(t):



Graphisches Modell des A/D-Wandlers

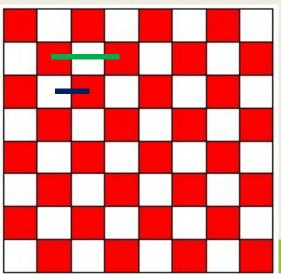
Signaltheorie V

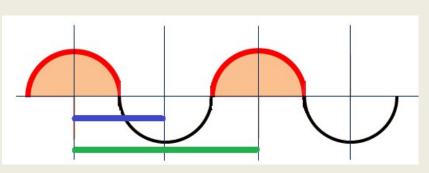


$$f_{sample} \ge 2 \cdot f_{\max}$$

- die Abtastfrequenz muss mind. doppelt so groß sein, wie die größte im Bild vorkommende Frequenz.
- die Verwendung der doppelten Abtastfrequenz ist notwendig, jedoch keine hinreichende Bedingung

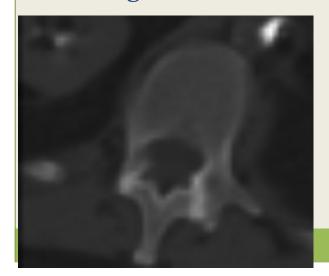
kleinste Periode Abtastintervall

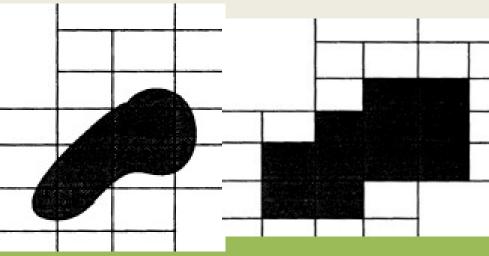




Signaltheorie VI

- Bei unzureichender räumlicher Abtastung fallen Kanten und Details, i.e. rapide lokale Änderungen der skalaren Signalwerte, auf einen Signalwert zusammen
 - o Partial Volume Effekt bei Computertomografie
 - o feine Details können am Raster nicht mehr abgebildet werden
 - o führt zu generellen Unschärfe; Vermischung der skalaren Werte
 - o die Extremwerte (Maximum / Minimum) werden massiv abgeschwächt

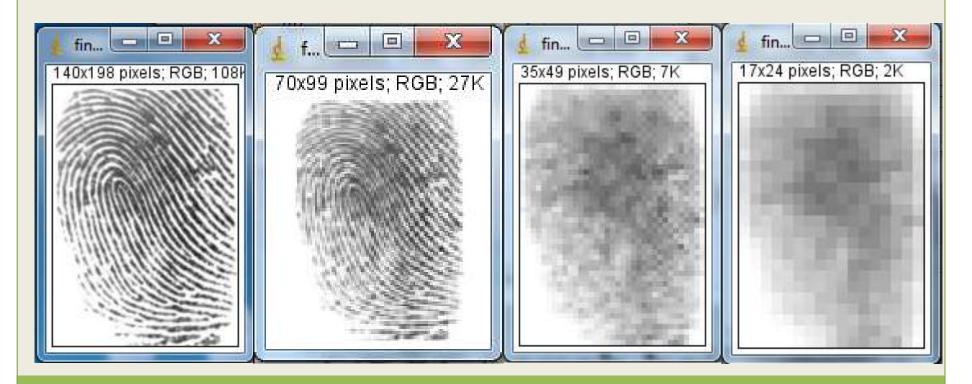




Signaltheorie VII

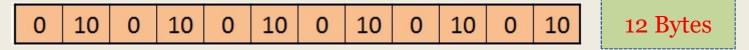
20

• Räumliche Auflösung auch bei Reskalierung von Bilddaten zu beachten. Je kleiner das Bild wird, desto eher verschwimmen feine Strukturen.



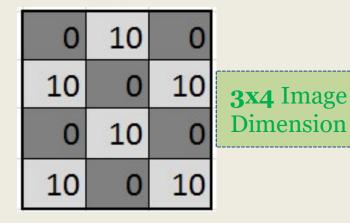
Speichermodell

- Viele medizinische Dateiformate (*Analyze, DICOM*) separieren Metadaten und die eigentlichen (*unkomprimierte*) Bilddaten, die als Byte-Sequenz persistiert werden.
- Nur durch Kenntnis der Abmessungen ist eine korrekte Bildwiedergabe möglich



0	10	0	10
0	10	0	10
0	10	0	10

4x3 Image Dimension

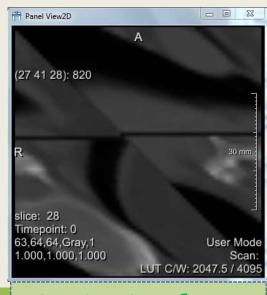


Speichermodell

- Bei bildverarbeitenden Anwendungen (Vergrößern, Verkleinern, Komprimieren,...) kann es zu Problemen bei der Bildrepräsentierung kommen, wenn die Metadaten nicht korrekt nachgezogen werden
- Bild dann entweder völlig verrauscht (*Indiz für falschen Datentyp*) oder periodisch versetzt (*Fehler bei der Dimension*)



64x64x64, 2bytes



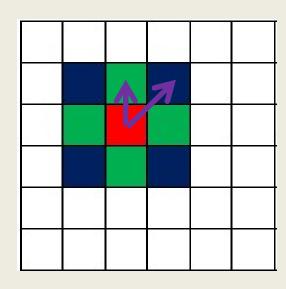


63x64x64, **2**bytes

64x64x64, 1byte



• Für Pixel in der diskreten Bildmatrix können Nachbarschaftsbeziehungen definiert werden:



- Für ein Pixel q (rot) gibt es bei 2D
 Bildern direkte Nachbarn in N₄
 (grün), d.h. links, rechts, oben und unten
- \circ Die diagonalen Nachbarn N_D (**blau**) weisen Pixelunterschiede in x- und y-Richtung auf
- \circ Die maximale Nachbarschaftsadjazenz N_8 ergibt sich als:

$$N_8 = N_4 \cup N_D$$



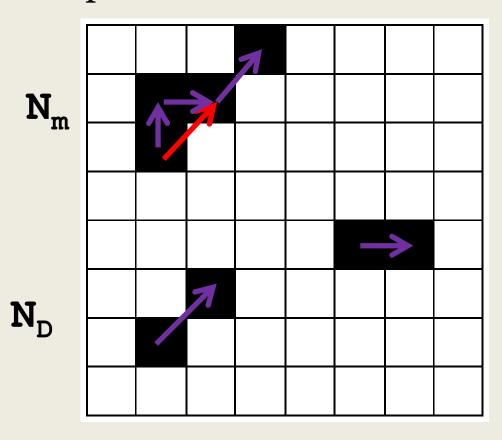
ullet Für Pixel $oldsymbol{p}$, $oldsymbol{q}$ im Bild $oldsymbol{V}$ kann nun definiert werden:

$$egin{array}{lll} egin{array}{lll} egin{array}{lll} egin{array}{lll} N_4 & (p,q) := & |p,q| & = & 1 \\ ON_D & (p,q) := & |p,q| & = & \sqrt{2} \\ ON_8 & (p,q) & := & N_4 & (p,q) & V & N_D & (p,q) \\ ON_m & (p,q) := & N_4 & (p,q) & V & (N_D & (p,q) & \Lambda \\ & & & ! & N_4 & (p,q) &) \end{array}$$

• N_m , i.e. "mixed adjacency", garantiert eindeutige Pfade

25

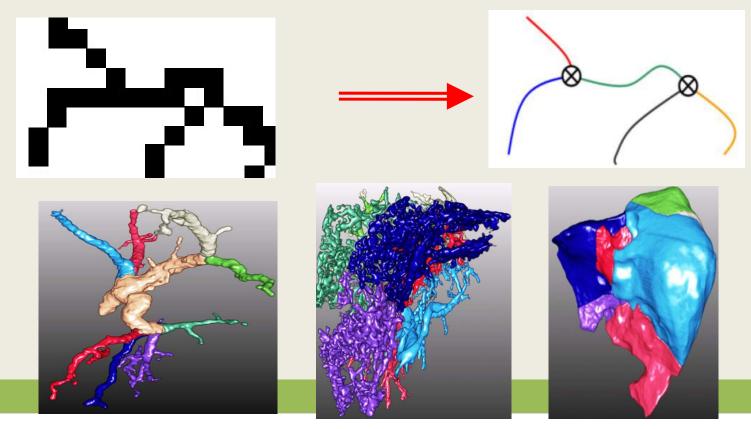
• Beispiele:



Bei der mixed adjacency ist die rote diagonale Verbindung nicht erlaubt, um
 eindeutige Pfade sicherstellen zu können

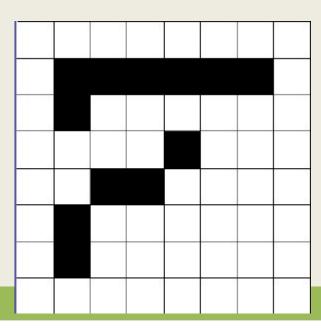
M-Adjazenz - Anwendungsgebiet

- O Überführung von Binärpfaden in geordnete Punktzüge (Vectorization). Erkennung von Verzweigungspunkten (bifurcation points) zum Aufbau von Pfad-Graphen.
 - x BSP: Gefäßhierarchien:



• **Pfade**: Zwischen 2 Pixeln **p**, **q** gibt es einen Pfad unter Verwendung einer spezifischen Nachbarschaftsadjazenz, wenn:

```
\circ \text{connN}_8(p,q) := N8(p,q) \lor (\exists z \in V : N8(p,z) \land N8(q,z))
```



Pfade in N_8

Distanzmetriken



- Abstandsberechnung für benachbarte Pixel p, q um die "Kosten" des Pfades bestimmen zu können.
- Allgemeine Bedingungen für die Pfadlänge:
 - odist(p,q) ≥ 0, d.h. es sind nur positive Pfadlängen zulässig
 - odist(p,q) ≤ dist(p,z) + dist(z,q), d.h. es gilt die Dreiecksungleichung
 - o dist(p,q) == dist(q,p), d.h. die
 Distanzberechnung ist kommutativ → Reihenfolge bei der Bearbeitung der Pfadelemente muss Ergebnisinvariant sein

Distanzmetriken



 Der Abstand zwischen zwei benachbarten Pixeln p, q kann wie folgt bestimmt werden:

• Euklid: dist(p,q) =
$$\sqrt{(x_p - x_q)^2 + (y_p - y_q)^2}$$

- O Manhattan: dist(p,q)= $|x_p-x_q|+|y_p-y_q|$
- O Schachbrett:

$$dist(p,q) = max(|x_p - x_q|, |y_p - y_q|)$$

 Die geamte Pfadlänge ergibt sich aus den Einzeldistanzen:

Oppath =
$$\sum_{i=1}^{length-1} dist(p_i, p_{i+1})$$

Distanzmetriken

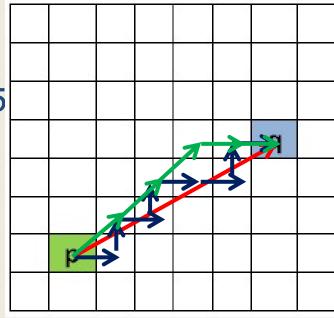


BSP

o *Euklid*: dist(p,q) = $\sqrt{(1-6)^2+(6-3)^2}$ = $\sqrt{34} \approx 5.831$

o Manhattan: dist(p,q) = 8

o Schachbrett: dist(p,q) = 5



Farben und Licht

31

RGB

CMY

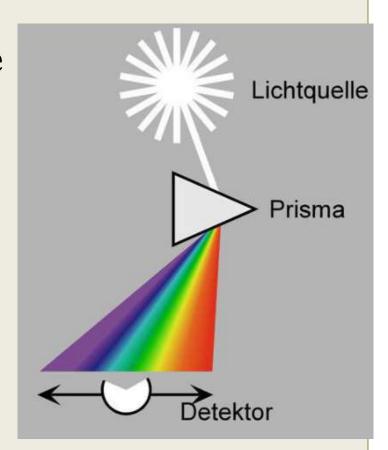
YUV

HSV

CIE LAB

Spektrale Zerlegung des Lichts

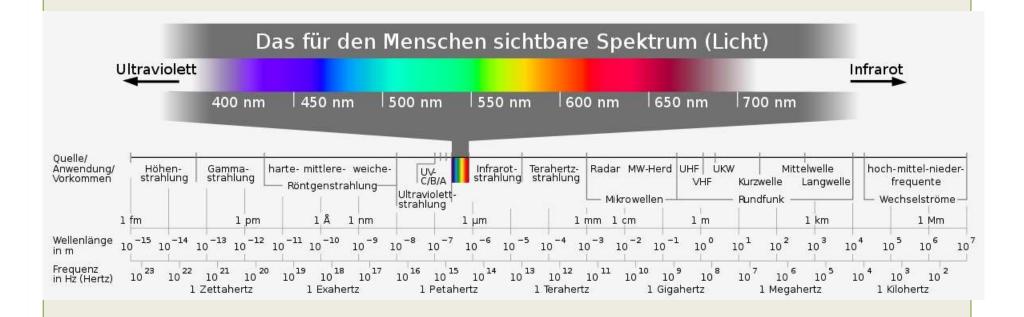
- Mittels optischem Prisma wird weißes Licht in seine Farbanteile zerteilt.
- Für Menschen sichtbares Licht: Bandbreite von 400nm (lila → grenzt an Ultraviolett) bis 700nm Wellenlänge (rot → grenzt an Infrarot)



Spektrale Zerlegung des Lichts

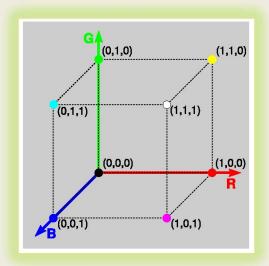


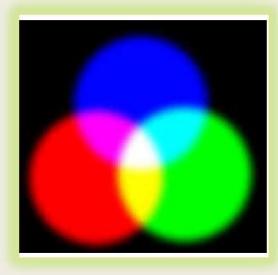
- Optisch sichtbares Licht nur sehr kleiner Ausschnitt des möglichen Wellen-Spektrums
 - Nicht sichtbar etwa Röntgenstrahlung, UV-Licht, Infrarot oder Rundfunk-Wellen





- **RGB** (red, green, blue)
 - o additives, hardwareorientiertes Farbmodell
 - Farben als Linearkombination ihres Rot-, Grün- und Blau-Anteils
 - Verwendet z.B. bei Monitoren / Displays

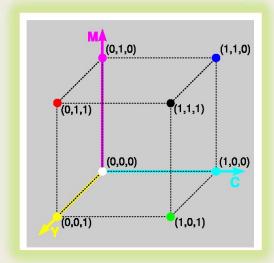


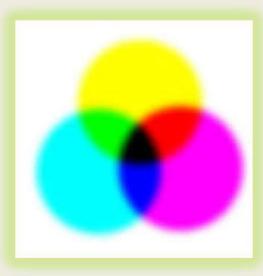


- RGB Speicher-Formate:
 - 24-bit true color: 1 Byte pro Farbkanal → 256^3
 == 16,777,216 darstellbare Farben
 - o 16-bit high-definition color → 5bit für Rot- und Blau-Kanal sowie 6bit für den Grün-Kanal obwegen physiologischer Eigenheit des menschlichen Auges
 - o 32-bit-Format:
 - ▼Im Falle von Opazität ein weiter Kanal notwendig (*RGBA*)
 - × Verarbeitung von 32 bit i.d.R. performanter − d.h. auch ohne Bedarf für Alpha-Kanal zu bevorzugen



- CMY (cyan, magenta, yellow), auch CMYK
 - o subtraktives, hardwareorientiertes Farbmodell
 - Farben werden als Linearkombination ihres Cyan-, Magenta- und Yellow-Anteils beschrieben.
 - o verwendet etwa bei Druckern



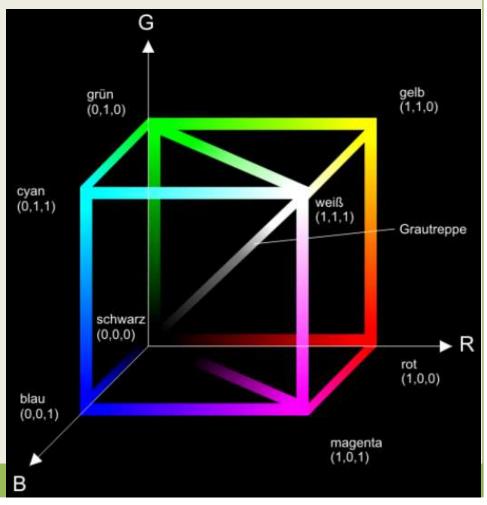




Umrechnung RGB ←→ CMY

$$\begin{pmatrix} R \\ G \\ B \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} - \begin{pmatrix} C \\ M \\ Y \end{pmatrix}$$

$$\begin{pmatrix} C \\ M \\ Y \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} - \begin{pmatrix} R \\ G \\ B \end{pmatrix}$$





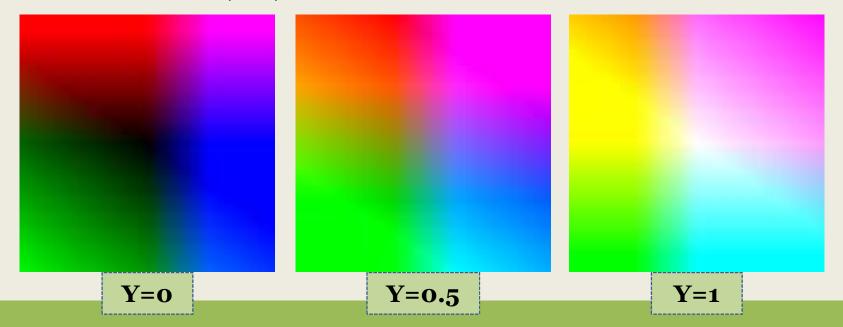
• YUV-Modell:

- Farbmodell für Farbfernsehen (*PAL-Standard-Europa*)
- Ein Kanal überträgt das Luminianz-Signal Y
 (Helligkeit == SW-Signal → Kompatibilität zu älteren SW-Geräten)
- *Chrominanz*-Signal UV enthält Farbe mit *U*=Blau-Luminanz und *V*=Rot-Luminanz
- o RGB-Signal lässt sich aus YUV rekonstruieren
- SW-Geräte ignorieren UV-Komponente → Abwärts-Kompatibilität



• YUV-Modell:

- Chromazitäts-Diagramm abhängig von spezifischem Luminanz-Wert Y
- o Bei **Y=o →** R,G,B dominant und etwas Magenta
- o Bei **Y**=**1** → C, M, Y dominant und etwas Grün





• YUV-Modell:

o Umrechnung RGB → YUV möglich

$$\begin{pmatrix} Y \\ U \\ V \end{pmatrix} = \begin{pmatrix} 0.299 & 0.587 & 0.114 \\ -0.147 & -0.289 & 0.436 \\ 0.615 & 0.515 & 0.100 \end{pmatrix} \cdot \begin{pmatrix} R \\ G \\ B \end{pmatrix}$$

$$U = 0.492 \cdot (B - Y)$$
$$V = 0.877 \cdot (R - Y)$$

$$Y \in [0,1]$$
 $U \in [-0.436,0.436]$ $V \in [-0.615,0.615]$

Technische Umsetzung:



Composite Video

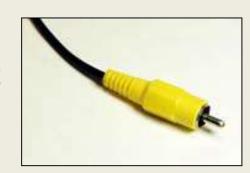
- Alle Signale (+Steuersignale) auf einer Leitung
- ▼ Demodulation der einzelnen Signale
- ▼ Gelber RCA-Jack, Ton (weiß,rot)
- × TV, VHS

o S-Video

- ★ Luminanz und Chrominanz auf eigenen Leitungen
- 4 Pin mini-DIN Stecker mit gepaarten Ground-Level

Anlog Component Video

- ➤ RGB Signal auf getrennten Coaxial-Kabeln
- Sync-Signale extra oder auf grün







Technische Umsetzung:



- VGA (video graphics array)
 - X Analoges Signal heutzutage nicht mehr benötigt, da Monitore ohnehin digitale Bilder verarbeiten → bei TFT transformation
 - 🗷 R, G, B als 3 der 15 Kanäle
- DVI-D (digital visual interface)
 - × Digitale Übertragung von Videodaten
 - **×** Seit TFT-Monitore
 - ➤ Serielle Übertragung der 8bits von R,G,B über den Daten-Kanal
- HDMI (high definition multimedia interface)
 - ➤ Verbesserte Unterstützung von YCbCr
 - ➤ Hohe Datenübertragung (4Gb/sec), Audio
 - × Auch 48bit Farbtiefe möglich







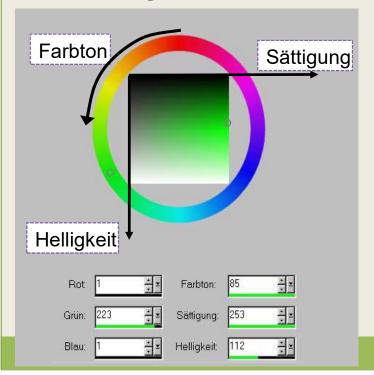


• **HSV** (hue, saturation, value) Farbmodell, auch HSB

o entspricht eher der menschlichen Farbwahrnehmung

o Farben ergeben sich durch Farbton, Sättigung und

Helligkeit

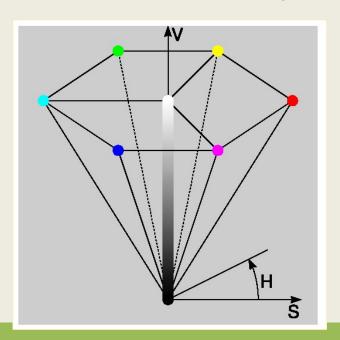


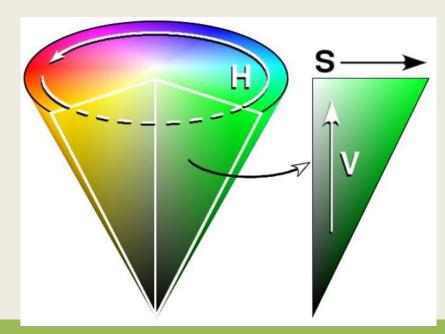


Lena (oder Lenna)



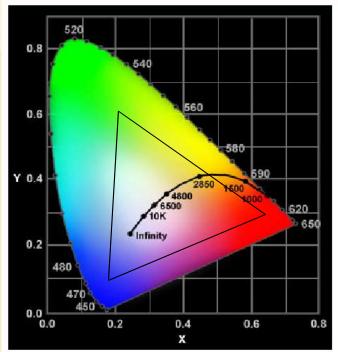
- HSV Modell als perzeptionsorientiertes Farbmodell
 - × hue (H): Winkel um vertikale Achse, definiert Basisfarbe (Rot o°)
 - x saturation (S): Farbanteil in Bezug zur aktuellen Helligkeit
 - ▼ value (V): Helligkeit zwischen Weiß und Schwarz
 - ▼ Wertebereich H [0;360], S und V [0.0;1.0]







- Chromazitäts-Diagramm CIE (Commission Internationale de *I'Eclairage*, L*a*b)
 - o der menschlichen Wahrnehmung nachempfunden / angepasst → bester Farbraum für Farbempfinden
 - Spezifikation eines 3D Koordinatensystems mit einer Untermenge, in der die sichtbaren Farben eines bestimmten Farbbereiches • Y = g/(r+g+b)(Gammut) liegen
 - Spektralfarben außen (in nm)
 - o gleiche Distanzen == gleiche Wahrnehmungsunterschiede



•
$$X = r/(r+g+b)$$

•
$$Y = g/(r+g+b)$$

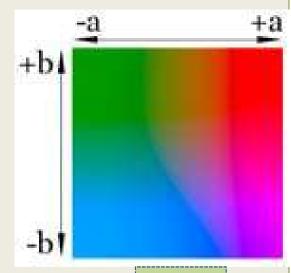
•
$$Z = b/(r+g+b)$$

Daraus folgt:

$$\circ$$
 X+Y+Z = 1



- CIE L*a*b cont'd
 - o heute z.B. auch für Qualitätsanalysen von Kameras (vgl. SmartPhones) verwendet
 - o enthält Weisspunkt W (alle Farben zu 1/3), sowie Black-Body-Kurve (Farben als Temperatur in Kelvin in Bezug auf einen B.B.)
 - Farbwiedergabe, z.B. RGB als Teilmenge (vgl. Dreieck)
 - o L: Luminanz Werte zwischen o und 100
 - x L=0 ... schwarz, L=100 ... weiß
 - × a: Grün (-) und Rot(+)
 - **★** b: Blau (-) und Gelb (+)



L=50

CIE-Diagramm cont.



• Neuerdings auch zum Benchmark Vergleich bei Displays (Fernseher, Monitore, SmartPhones)

