

# Einführung in die Medieninformatik

Für EMI im 1. Semester beim Bart

## Allgemeines

### Aufgabenstellungen

Berechenbarkeit (Lösen mathematischer Probleme), Unberechenbarkeit, Curchsche These (Gleichwertigkeit aller Maschinenkonzepte), Chomsky-Hierarchie (Zusammenhang Automaten  $\dot{\jmath}$ - $\dot{\zeta}$  Sprachgrammatik), Programmkomplexität, Programmverifikation

### Medienkompetenz

Medialitätsbewusstsein + Medienwissen, Medienspezifisches Rezeptionsmuster, Medienbezogene Genussfähigkeit (Bedürfnis nach Identifikation und Unterhaltung), Medienbezogene Kritikfähigkeit. Medien: Mittel zur Speicherung von Information (Physisch oder Format, Text, Video, Audio, )

## Text und HTML

### Zeichensätze

Zeichensatz definiert Zuordnung von Zahlen zu Zeichen.

1. **ASCII**: 7-bit-Zeichensatz (128 Zuordnungen), 0-31 Steuerzeichen. Neuere Zeichensätze sind meist abwärtskompatibel (Erste 128 Zuordnungen entsprechen ASCII)
2. **ISO-8859-1, ISO-8859-2 ... ISO-8859-16**: 8-bit-Zeichensätze (256 Zuordnungen), 0-127 ist ASCII, darüber Sonderzeichen wie deutsche Umlaute (-1: Latin1 westeuropäisch, -2: Latin2 osteuropäisch, -5: Kyrillisch, -6: Arabisch, -7: Griechisch [...], -15: Westeuropäisch mit Eurozeichen)
3. **Unicode**: Internationaler Zeichensatz. Verschiedene Standards: UCS-2 mit 2 hoch 16 Zeichen, UCS-4 mit doppelt so vielen Ebenen. Unicode 5 kennt 99k Zeichen. Unicode-Formate:

### Braille-Schrift

Von Louis Braille 1820 für Blinde entwickelt, für viele natürliche Sprachen verfügbar. Gibt Zeichen als fühlbare Punkte auf dem Papier wieder (6 - 8 Punkte kodieren ein Zeichen). Spezielle Schreibweisen für Formeln, Musik o. ä.

- LittleEndian: Höchstes Byte steht ganz links, BigEndian: Höchstes Byte steht ganz rechts. Standard bei Intel / x86 ist LittleEndian.
- **UTF-32**: Immer 4 Bytes für ein Zeichen
- **UTF-16**: 2 Bytes pro Zeichen mit Möglichkeit für mehr, wenn in ersten 2 Bytes spezielle Surrogate-Kodierung benutzt wird
- **UTF-8**: Wichtigstes Unicode-Format, weit verbreitet z. B. für XML. Besteht aus einem Byte pro Zeichen, das nur ASCII ist, wenn das oberste Bit gesetzt ist, ansonsten gehören die nächsten Bytes auch mit zum Zeichen, dann 2-4 Bytes. Aufbauschema: Höchstes Bit ist nicht gesetzt, wenn es nur ein ASCII-Zeichen ist und somit nur aus einem Byte besteht. Ist das höchste Bit gesetzt, gibt die Anzahl der dann folgenden Bits die 1 sind an, aus wie vielen Bytes das Zeichen besteht. Ist das erste Byte z. B. 11100000, folgen noch zwei weitere Bytes für das gleiche Zeichen.

## Schrift

### Eigenschaften



(Bild: Brian Ammon, CC-BY)

1. Höhe wird in Punkt angegeben. Ein Punkt ist meist 1/72 Zoll (TeX, DTP / Postscript), 1 Pica sind 12 Punkte.
2. Schriftstärke (mager, normal, fett), Schriftbreite (schmal, normal, breit), Schriftlage (normal, kursiv)
3. Vektorschriften: Definition durch Vektoren, Rasterschriften: Definition durch Pixelgrafik für je ein Zeichen
4. Serifen: Helfen bei Lesbarkeit, sollte nur bei Fließtext benutzt werden.
5. Kerning: Anpassung von Zeichenabständen bei parallelen Diagonalen ("BRAVO": Zusammenrücken von A und V), Definition pro Buchstabenpaar. Kann doof sein bei geringen Abständen.

### Absätze

Absatzformat gibt Abstände und Einrückungen an. Möglich sind z. B. Blocksatz (Automatische Anpassung des Wortabstandes, sodass stets ganze Zeile gefüllt wird), Flattersatz / Mittelachsensatz (optimale Wortabstände). Seitenlayout: Kopf, Rumpf, Fuß.

### Formate

Eigenschaften: Editierbar? WYSIWYG? Lesbar? Medien einbindbar? Funktionsumfang? Programmierbar?

- **Plain**
- **Rich Text** (Microsoft), eigene Definition für Medien, nicht programmierbar, alt, WYSIWYG, MS Word, L<sup>A</sup>T<sub>E</sub>X, PostScript, PDF, XHTML -  

```
{\rtf Hallo \par {\i Dies} war kursiv. \par Noch ne Zeile.}
```
- L<sup>A</sup>T<sub>E</sub>X(Frei), voll programmierbar, kein WYSIWYG (wirft PDF o. ä. aus), Medien einbindbar, gute Formelsetzungsbibliotheken, vielseitig und toll, aber kompliziert
- **PDF** (Adobe), baut auf Postscript auf, Layout-Treu, Rechtemanagment, Einbetten von Schriften problematisch, kaum editierbar, wird aus anderen Formate generiert
- **HTML**: Trennung von Struktur und Style mit CSS, Scripting (Document Object Model: Abbildung des XML-Baums auf JavaScript), programmierbar, kein WYSIWYG, Plugins

## Internetz

- **HTTP**: Protokoll zwischen Web-Server und Web-Client (Browser)
- **URI**: Bezeichnung und Adressierung beliebiger Daten im Netz (z. B. <http://www.wh2.tu-dresden.de>)
- **Hypertext**: "Nicht-lineares" Textmedium, verbindet Inhalte durch Hyperlinks (URI, entfernte URIs wie "http://google.de", "ftp://..." oder lokale URIs wie "mailto:", die an andere Programme auf dem Rechner gehen)
- **Mark-Up-Sprachen**: Beschreibung von Text durch Text, bspw. HTML mit öffnenden Tags und schließenden Tags und Inhalt dazwischen

## HTML

```
<html><head><title>Hallo Welt</title>
</head><body><h1>Grosse Ueberschrift
</h1><a href="uri">Link</a></body></html>
```

- Attribute bspw. href mit Ziel-URI, align mit Ausrichtung o. ä.
- <a>-Elemente sind Anker, verlinkt wird per "#ziel" und gesetzt per name-Attribut
- Bilder mit 
- Listen mit ul (unordered list) und ol (ordered list)-Elementen, Einträge mit li (list item)-Elementen
- Tabellen mit table, tr für Zeilen und td für Spalten.  

```
<table><tr><th>1</th><th>2</th></tr>
<td>A</td><td>b</td></tr></table>
```
- div für allgemeine Blöcke

## Barrierefreiheit

Barrieren für Behinderte auf Webseiten (Sehbehinderte, Farbenblinde, Gehörbehinderte, Motorisch behinderte, geistig behinderte)

## CSS

Strikte Trennung zwischen Präsentation und Design, definierte Formate für XHTML-Tags. Text-Formatierung, Positionierung von Elementen, Rahmen und Leerräume, Hintergrund, Formatierung von Tabellen, CSS einbinden mit <style type="text/css"></style> für direktes reinschreiben von CSS, CSS-Datei im Head des HTML-Dokuments verlinken mittels <link rel="stylesheet" type="text/css" href="datei.css">. HTML-Elemente koennen auch mit style-Attribut und CSS-Angaben gestyled werden.

CSS besteht aus Selektoren für Elemente (mittels ID oder Klasse) und Style-Attributen. Bspw.: #foo { color: red; } setzt die Schriftfarbe im Element mit der ID foo auf rot. Mögliche Styleattribute sind z. B. "font-family" (Arial, sans-serif), "font-weight" (bold).

## Positionierung

Mit CSS werden Elemente nach dem Boxmodel layoutet, d. h. jedes Element bildet eine Box. CSS-Attribute für jedes Element sind z. B. "margin" (Außenrand um die Box des Elements, z. B. "5px"), "padding" (Innenabstand um den Boxrand), "border-width" (Randbreite), "border-color" (Randfarbe).

## Navigationstechniken

Websites sollten sets einwandfrei navigierbar sein (vor / zurück, Nutzer muss sich zurechtfinden können). Bspw. durch Setzen von Navigationsleisten, Suchmaschinen etc.

## Karten in HTML

Map-Element definiert Karte. Bsp.:

```
<map name="schland">
<area shape="rect" coords="11,10,59,29"
href="uri" alt="Hier"></map>

```

Beispielsweise rect mit coords="x1,y1,x2,y2" mit 1=Obere Linke Ecke,2=Untere rechte Ecke oder "circle" mit "x,y,r" oder "poly" "x1,y1,x2,y2..."

## Dynamisches Web

Dynamik z. B. durch auf dem Web-Server dynamisch generierte Inhalte durch PHP, ASP.net o. ä. Im Browser wird JavaScript benutzt.

## JavaScript

1. **Objekt**: Ein JavaScript-Objekt besteht aus Schlüsseln, denen Werte zugeordnet werden (bspw. das document-Objekt hat den Schlüssel bgcolor mit dem Wert white).
2. **Funktionen / Methoden / Aktionen**: Wie Funktionen in C, können auch Variablen zugeordnet werden
3. **Ereignisse**: Das Ereignis "onmouseover" eines HTML-Elements kann durch JavaScript auf eine Funktion gesetzt werden.

4. **Variablen:** Können Zahlen, Strings, Booleans, Objekte, Funktionen sein. Bsp. zur Deklaration: `var month = "Januar";`
5. Ausdrücke ordnen einer Variable einen Wert zu (Beispiel s. o.)
6. Operatoren verändern Variablen (siehe C, bspw. "+", "==")
7. Einbinden einer JavaScript-Datei in HTML:

```
<script src="js.js"
type="text/javascript"></script>
```

Einbetten von JS in HTML:

```
<script type="text/javascript">
```

Zugreifen auf Dokumenteigenschaften in JavaScript über das globale window-Objekt, bspw. `window.location.href = 'http://google.de'`. window-Objekt enthält document-Objekt, über das HTML-Elemente mittels `"getElementById"` abgefragt werden können.

## HTML5

HTML5 als zukünftiger Standard führt HTML-Elemente für Strukturierung ein (header, nav) und neue Dinge wie das Video-Tag.

## XML

- Medien in Dokumenten definieren sich durch Inhalt (Semantik), Struktur (Syntax) und Präsentation (lexikalische Information)
- Für Arbeiten mit HTML bedient man sich der Sprache, Werkzeugen und Arbeitsmethoden (Layouten o. ä.). HTML basiert auf XML, das als Metasprache Text beschreibt und strukturiert (Baumstruktur).
- Dokumente bestehen aus Tags  
<person>Der Bart</person> oder für leere Elemente <br/>, Elementname (person) legt das Tag als Ganzes fest. Metazeichen von XML sind entsprechend zunächst <, > und /.

- Elemente können Elemente enthalten (Schachtelung), Elemente können Text enthalten, Groß- und Kleinschreibung spielt eine Rolle und das öffnende Tag muss genauso lauten wie das Schließende
- In XML-Dokumenten immer XML-Deklaration am Dokumentenbeginn: `<?xml version="1.0">`. Folgen kann dann ein Stylesheet und eine DTD sowie die Daten.
- **Wohlgeformtes** XML ist syntaktisch korrekt (s. o., mit korrekten Schachtelungen, richtig verwendeten Metazeichen etc), **Valides** XML erfüllt die angegebene DTD
- Wichtige Begriffe: Elemente, Knoten (Tags), Vorgänger, Nachfolger, Nachbarknoten, Wurzel, Attribute, Text
- XML-Tags müssen mit Buchstaben anfangen
- Entitäten in XML sind z. B. `&lt;` & `&gt;` & `&quot;`; für das `<>`-Zeichen (&#x27E9;) sind "Fluchtsymbole" in XML. Dadurch können alle Metazeichen in XML auch verwendet werden (& wird durch `&amp;` gesetzt)
- `<![CDATA[<svg></svg>]]` ermöglicht Textabschnitt ohne Fluchtsymbole
- In der XML-Deklaration kann Attribut `encoding="iso8859-1"` und `standalone="no"` (gibt an, dass die DTD im Dokument selbst drin ist) verwendet werden
- Kommentare: `<!-- Kommentar -->`

```
<!DOCTYPE html PUBLIC
"-//W3C//DTD XHTML 1.0 Strict//EN"
"http://www.w3.org/TR/xhtml1/DTD/
xhtml1-strict.dtd">
```

Gibt an, dass die XML im Internet zu finden ist.

## DTD

- DTD ist selbst kein XML, gibt Aufbau der XML-Daten mit Regeln an
- Beispieldefinition:

```
<!DOCTYPE person [
<!ELEMENT person (fname, profession*)>
<!ELEMENT profession (#PCDATA)>
<!ELEMENT fname (#PCDATA)>
]>
```

```
<person>
<fname>Tim</fname>
<profession>Troll</profession>
<profession>Toll</profession>
</person>
```

- `<!ELEMENT (...)>` definiert ein Element und in Klammern die erwarteten Kindelemente, in fester Reihenfolge. Man kann ? für ein-oder-keinmal anfügen, \* für beliebig oft (oder gar nicht), + für ein oder mehr. Dabei geht Gruppierung mit ((a, b) | (c, d)). Als vorgegebenes Kind-Element kann auch #PCDATA genommen werden (also beliebig viel normaler Text). Es kann auch EMPTY statt der Kindelementliste geschrieben werden, wenn das Element immer leer sein soll (oder ANY für komplett egal)
- Attributtypen sind z. B. CDATA (Text), ID (eindeutiger XML-Name), NMTOKEN (wohlgeformter XML-Name), IDREF (Name eines anderen im XML vorkommenden Elements), IDREFS (Mehrere Namen von anderen XML-Tags im Dokument, getrennt durch Leerzeichen)
- `<!ATTLIST img source CDATA #REQUIRED w CDATA #REQUIRED` gibt Attributliste für das img-Element an. Möglich sind #REQUIRED (muss vorhanden sein), #FIXED (muss immer einen bestimmten Wert haben, der danach angegeben wird), #IMPLIED (optional). Alternativ kann eine beliebige Zeichenkette geschrieben werden, die den Standardwert angibt.
- Namespaces in XML:

```
<svg xmlns="http://www.w3.org/2000/svg">
<svg:ellipse rx="1" ry="1"/>
</thml>
```

- XML für Dokumente: Text Encoding Initiative mit TDT für Theater o. ä.
- DocBook / OpenEBook: XML für Bücher wie Informatikbücher
- DTDs können keine Zahlen oder Maximallängen vorgeben, keine Modularisierung, DTDs sind keine XML-Dokumente

## XSLT

- Daten von der Präsentation trennen, Transformation von XML-Struktur in neue Struktur, Formatierung der neuen Struktur
- Einbindung eines XSLT-Stylesheets:

```
<?xml-stylesheet type="text/xsl"
href="helloworld.xsl"?>
<?xml version="1.0" encoding
="iso-8859-1"?>
<greeting>Hallo Welt!</greeting>
```

Beispiel des XSLT-Stylesheets:

```
<xsl:stylesheet version="1.0"
xmlns:xsl="..." >
<xsl:template match="/">
<html><head>
<title>Wort zum Tag
</title></head>
<body><p>
<xsl:value-of select="greeting"/>
</p></body></html>
</xsl:template>
</xsl:stylesheet>
```

- Verarbeitung basiert auf Templates, auch ECMA-Script möglich
- `<xsl:apply-templates>` führt die Templates, bei Angabe von `select` nur auf entsprechende Elemente, `<xsl:value-of select="elemname">` holt den Wert eines XML-Tags und fügt ihn an dieser Stelle ins Dokument ein, `<xsl:for-each select="a">` gibt für jedes folgende Element den definierten Text aus,

`<xsl:template match="people">` gibt für alle `people`-Elemente im XML den definierten Text aus

- Mit dem Attribut `mode` für `apply-templates` und `template` kann Status definiert werden, z. B.
- `<xsl:variable name="width" select="50"/>` legt Variable `width=50` fest
- `<xsl:if test="position()=1"></xsl>`
- `<xsl:for-each select="book"></xsl>`
- `<xsl:sort select="author"></xsl>`
- XPATH wird in XSLT benutzt, um Tags anzusprechen (`match`-Attribut). ist die Wurzel, Tagnamen werden als Nachfolger interpretiert, der zugehörige Tag eines Attributes kann mit `@`born angesprochen werden
- In XPATH kann gerechnet werden (`div`, `+`, `-`, `*`, `mod`, `position()`, `first()`, `count()`, `round()`, `substring()`)

## Bild

### SVG

- EncapsulatedPostScript, Bildinformationen in speziellem Code, programmierbar, keine Alphakanäle, nur PostScript-Geräte
- Scalable Vektor Graphics (SVG) in XML für Bilder, Text, Animation, Formen (bspw. mit Inkscape malbar)
- Pfade: `<path d=""></path>`, Möglich anzugeben: M x y (Zeichenstift ohne zu zeichnen bewegen), L x y (Linie zum Ziel malen), Q x1 y1 x2 y2 (Quadratische Bezier-Kurve), C x1 y1 x2 y2 x3 y3 (Kubische Bezier-Kurve), Z (Pfad schließen), H x (Horizontale Linie), V y (Vertikale Linie), A ??? (Ellipsenbogen), S x y (Quadratische Bezierkurve fortsetzen, vorherigen Kontrollpunkt spiegeln) - Groß geschriebenes Zeichen steht für absolute, kleines für relative Koordinaten. Pfade beginnen immer mit `moveto` (M) und können mit Z enden.

- Primitive: circle-Element mit Attributen `cx`, `cy` und `r` (CenterX und CenterY), ellipse mit `cx`, `cy`, `rx` und `ry`, rect mit `x`, `y`, `width`, `height`, line mit `x`, `y`, `width`, `height`, polygon / polyline mit `points` (`x1 y1 x2 y2...`)

- Text: text-Element mit `x`, `y`, `size`

- Image: image-Element mit `x`, `y`, `width`, `height`, `xlink:href`

- Einfärben von Formen mit `fill`-Attribut und einer Farbe in Hex (`#FF0000` z. B.) oder Name.

- Rand malen mit `stroke`-Attribut (enthält Farbcode) und `stroke-width`-Attribut (Randbreite in Pixeln)

- transform-Attribut gibt Rotation, Transformation, Skalierung oder Stauchung an: z. B. `translate(20, 150) scale(0.5) rotate(30,255,230) skewX(25) skewY(25)` - Bei Rotate dienen der zweite und dritte Parameter dabei als Rotationszentrum

- Clipping: Beschränken eines zu zeichnenden Elements auf den Bereich, den ein anderes Element definiert

- Filter in SVG: "Fotorealismus" durch Unschärfe, gezielte Einbringung von Bildstörungen o. ä.

- Blurring: Unschärfefilter

- Events in SVG mit JavaScript definieren (svg onload-Event, onclick-Event etc.)

- `<ellipse id="elli" cx="100" cy="100" rx="48" ry="90" style="fill:green;" /><animate xlink:href="#elli" attributeType="XML" attributeName="rx" begin="0s" dur="6s" from="48" to="198" fill="freeze" />`

- In `<defs>...</defs>` können SVG-Elemente definiert werden, die nicht gezeichnet werden, sondern mit `<use xlink:href="#id">` später wieder benutzt werden

# Licht

- Licht ist eine elektromagnetische, transversale Welle (schwingt parallel zur Ausbreitungsrichtung)
- Lichtgeschwindigkeit im Vakuum  $c = 299792458 \frac{km}{s}$  ist vom Medium, in dem sich das Licht bewegt, abhängig (meist Luft),  $c = \lambda * f$  (f: Frequenz des Lichts,  $\lambda$ : Mediumskonstante)
- Frequenz ist Maß für Energie  $E = hf = \frac{hc}{\lambda}$  (h ist das Plancksche Wirkungsquantum mit ca.  $h = 6.63 * 10^{-34} Js$ ), d. h. je höher die Frequenz, desto energiereicher ist das Licht.
- Frequenz des sichtbaren Lichts beträgt ca. 400 nm - 780 nm, davor liegen die Frequenzbereiche von UV-Licht, Röntgen- und Gammastrahlung, darüber die von infrarotem Licht, UKW-Radio etc
- Je höher die Frequenz, desto geringer die Wellenlänge (antiproportional)
- Lichtemission: Durch Gasentladung (z. B. Xenon Lampen), Photolumineszens (Phosphor "wandelt" einfallendes Licht in andere Wellenlänge um)
- Chemische Reaktion: Verbrennung (Flamme strahlt Licht)
- Phosphor: Elektronenbeschuss bringt Phosphor zum Leuchten
- Absorption: Bspw. beim Durchlaufen einer farbigen Folie wird ein Teil des Lichtspektrums absorbiert
- Reflexion: Ein Teil des Lichts wird beim Auftreffen auf eine Oberfläche reflektiert
- Additive Farbmischung: Überlagerung von Licht mehrerer Emissionsquellen, Spektren werden addiert
- Subtraktive Farbmischung: Einfallendes Licht fällt auf Filter (Absorption / Reflexion) ("Multiplikation mit Filterfunktion")
- Gedruckte Farben erscheinen je nach Beleuchtung unterschiedlich (z. B. Geld wird rötlich bei roter Beleuchtung...)

- Farbtemperatur in Kelvin: "Wärme" des Lichtes, je höher desto weißer - Glühlampe ca. 2800K, 6500K: Tageslicht
- Mensch "sieht" Lichtintensität, visuelle Wirkung, Kontrast, Formen, Bewegung
- Sehen beeinflusst Wachsamkeit, Wohlbefinden, innere Uhr [...]
- Spezielle Rezeptoren, spezieller Nervenpfad im Hirn
- Retina im Auge mit Photorezeptoren besetzt, Fovea im Zentrum des Sehfeldes, "blinder Fleck" des Auges ca. 1,75mm im Durchmesser (Gehirn "vervollständigt" das Bild)
- Farbwahrnehmung in der Retina erfolgt durch Zapfen und Stäbchen (Zapfen: Farbwahrnehmung, Stäbchen: Schwarz/Weiß)

## Farbmodelle

- RGB-Farbraum: 256 Farbstufen pro Rot, Grün und Blau, also ca. 16 Millionen Farben
- CIE 1932: xyz-Kanäle, y: Wahrgenommene Helligkeit, x: rot/grün, z: blau/gelb-Unterschied
- Farbwahrnehmung: Helligkeit (Luminance), Farbton und Sättigung (Saturation) hauptsächlich
- Hirn teilt wahrgenommene Farben in weiß, grau, schwarz und verschiedene Farben
- 128 Farbtöne (Hues), 130 Farbsättigungen (Saturation Levels), 16 (blau) - 26 (gelbe) Helligkeitswerte, also ca. 7 Millionen wahrnehmbare Farbnuancen
- Farben in Text: Hervorhebung, sparsam (auffällige) Farben einsetzen
- Räumliches Sehen durch Augenabstand (Fusion, Disparität) zum Erkennen von relativer Größe, Perspektive, Beleuchtung
- Hirn bildet 3D-Modell (Fusion): Unterdrückung unvereinbarer Regionen, bei Überlagerung wechselt unvereinbares linkes / rechts Bild

- Räumliches Sehen: Fokus (Raumpunkt, auf den die Augen fokussiert sind), Retinale Position (relative Position zur optischen Achse), Mentales Auge: Fusion kombiniert Information
- Disparität: Unterschied von fusionierten Punkten kann zur Berechnung der relativen Position benutzt werden
- Munsell Farbsystem (Farbton hue, Helligkeit value, Sättigung chroma)
- HLS Farbsystem (Hue, Luminance, Saturation)
- CMY - Farbraum (Cyan, Magenta, Yellow)  $(r, g, b) = (1, 1, 1) - (c, m, y)$
- CMYK - CMY mit Schwarz, damit Schwarz nicht aus Farbe gemischt werden muss (für Drucker toll)
- YUV - Y: Helligkeit (s / w), U, V: Farbbalance
- $YC_bC_r$  - Y: Helligkeit, b: Blau, r: Rot

## Digitale Bilder

Digitale Bilder bestehen aus Pixeln ( $w * h$ ), die pro Pixel meist 8 - 32 bit Farinfos beinhalten (Farbtiefe des Bildes) (ab 24 bit: TrueColor)

- Indizierte Farbformate: Zu Beginn des Bildes Farbtabelle, danach wird jede Farbe mit einem Index auf die Farbtabelle angegeben (sinnvoll bei wenigen verwendeten Farben)
- Chroma Subsampling bei YUV und  $YC_bC_r$  (4 : 2 : 2): Halbe Auflösung für g/b (U) und r/gr (V)-Kanäle
- Alphakanal: Transparenzangabe
- Auflösung des Auges:  $\phi$  minimaler Sehwinkel, bei dem 2 Linien im Abstand  $s$  erkannt werden,  $s = 2a * \tan(\frac{\phi}{2})$
- dpi / ppi: dots per inch / pixels per inch (Pixel pro Zoll,  $1in = 2,54cm$ )
- Monitor ca. 95 ppi, Laserdrucker ca. 1400 dpi

## Digitale Bildformate

- Histogramm: Gibt Verteilung der Graustufen von 0-1 des Schwarzwertes an
- Bildfilter: Beispielsweise Weichzeichner (lineare Mittelwertbildung aus Nachbarn um einen Pixel herum)
- GIF, Graphics Interchange Format: Animierbares farbindiziertes Bildformat mit 256 Farbstufen, kein Alphakanal, verlustbehaftete Kompression
- PNG: Portable Network Graphics: verlustfreie Kompression als GIF, Alphakanäle, freies Format, keine Animation, nicht für Fotos geeignet (verlustfreie Kompression zu schwach), Interlacing (Zunächst werden nur einige Zeilen des Bildes übertragen und dadurch eine Vorschau möglich, bevor das ganze Bild übertragen ist)
- JPEG (Joint Photographic Experts Group): Verlustbehaftete Kompression u. a. mit Huffman-Kodierung, Bildung von 8x8-Blöcken - führt zu Blockbildung, JPEG-2000 als Nachfolger mit geringerer Artefaktbildung - Farbraumrechnung in YUV, Tiefpassfilterung & Unterabtastung der Farbdifferenzsignale U, V (verlustbehaftet 4 : 2 : 2, s.o.), Quantisierung, Umsortierung, Huffman-Kodierung. Gut geeignet für Fotos, nicht gut für alles andere (Diagramme etc)
- JPEG2000 basiert auf XML - ermöglicht Rechtemanagement, verschiedene starke Komprimierung in verschiedenen Bildteilen, Interlacing-Support
- TIF - Tagged Image Format: mehrere Farbräume wie RGB, LAB, verlustfreie Komprimierung mit LZW, schwache Komprimierung, können viele Browser nicht
- PSD - Adobe Photoshop, Kanäle, Bearbeitungsschritte, Text wird nicht gerastert, kann von vielen Programmen nicht gelesen werden

## Vektorgrafik

- Vektorgrafik - Beschreibung des Bildes mittels Formen wie Rechtecken und nicht durch Pixel (Bsp.: SVG)

- Transformationen (Rotation, Translation, Skalierung, Scherung) können mittels der Multiplikation des alten Vektors mit einer Matrix und der Addition eines neuen Vektors ausgedrückt werden.  $\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} e \\ f \end{pmatrix}$
- Translation:  $\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} \Delta x \\ \Delta y \end{pmatrix}$
- Rotation:  $\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \end{pmatrix}$
- Scherung:  $\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} 1 & s \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \end{pmatrix}$
- Skalierung:  $\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \end{pmatrix}$
- Bezierkurven dritten Grades / kubische Kurven: Kurven nach dem Format:  $\begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = (1-t)^3 \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} + 3t(1-t)^2 \begin{pmatrix} x_1 \\ y_1 \end{pmatrix} + 3t^2(1-t) \begin{pmatrix} x_2 \\ y_2 \end{pmatrix} + t^3 \begin{pmatrix} x_3 \\ y_3 \end{pmatrix}$  Dabei sind  $\begin{pmatrix} x_0 \\ y_0 \end{pmatrix}$  und  $\begin{pmatrix} x_3 \\ y_3 \end{pmatrix}$  die Start- bzw. Endpunkte der Kurve und  $\begin{pmatrix} x_1 \\ y_1 \end{pmatrix}$  sowie  $\begin{pmatrix} x_2 \\ y_2 \end{pmatrix}$  definieren die Tangenten an diesen Punkten.
- de Casteljau-Schema:  $t = \frac{1}{4}$
- TrueType-Fonts: Die Kontur von Zeichen werden durch Bezierkurven beschrieben, wird dann gefüllt
- Meta Font (L<sup>A</sup>T<sub>E</sub>X): Glyphenkontur wird durch interpolierende Kurve beschrieben
- Gruppierung von einzelnen Formen, damit Transformation auf alle Elemente angewendet werden können
- Tiefe: Definierung einer Koordinate z, damit die Überlagerungsreihenfolge bei überlappenden Formen definiert werden kann
- Rasterisierung: Übersetzung von Vektorgrafik in Pixelgrafik (durch Test jedes Bildpunkts, ob er in

einer Form liegt oder durch zeilenweises Abtasten [Sweepeline], oder durch Abtasten jeder Form und Setzen der ausgefüllten Pixel)

- Dropouts (bei TrueType-Fonts): Nur halb ausgefüllte Pixel werden falsch als nicht gegeben gesetzt, daher setzen TrueType-Abtaster einen Pixel zwischen zwei anderen Pixeln auch dann, wenn zwischen diesen Pixeln zwei Konturlinien der Glyphen liegen (aber nur dann, wenn beide Konturen in beide Richtungen fortgesetzt werden)

## Wahrnehmung

- Motivation: Notwendige Daten an menschliche Wahrnehmung anpassen, um Platz zu sparen (z. B. nicht wahrnehmbare Töne bei Audio)
- Ergonomisch physische Dinge gestalten, Gestaltung von Benutzerschnittstellen, Layout etc. auf Menschen optimieren
- *Reiz* → *Sensor(analog)* → *Filter* → *Quantisierung(Digitalisierung)* → *Filter(digital)* → *Kompression* → *Codierung* → *Speichermedium*
- *Speichermedium* → *Decodierung* → *Dekompression* → *Filter(digital)* → *Rekonstruktion* → *Filter(analog)* → *Display(analog)* → *Reiz(beimMensch)*
- Wie funktioniert Reiz? Wie werden Reize beim Menschen wahrgenommen (Visuell, Akustisch, Haptisch, Gustorisch / Geschmack, Olfaktorisch / Geruch), Multimedial: Sprache und Deixis, Multimodal: Gedächtnisleistungen. Wie stark wird welcher Reiz wahrgenommen? Wie gleicht man produzierten Reiz mit echtem Reiz ab?
- "Wahrnehmung ist der Prozess, durch den Lebewesen Kenntnisse über ihre Umgebung und über sich selbst in Beziehung zur Umgebung gewinnen. Sie ist der Anfang allen Wissens."
- Reize / Wahrnehmung / Neuronale Prozesse
- Funktionale Physiologie: Erkenntnisse durch Patienten mit verletzten Gehirnen, elektrische

Störungen von außen, durch teilweise Einschläferung, mit Elektro-Enzephalographie, Magneto-Enzephalographie

- Wahrnehmung ist relativ: Größe / Helligkeit
- Wahrnehmung der Länge erfolgt logarithmisch (bei längeren Strukturen müssen Unterschiede größer sein)
- Wahrgenommene Skalen für Größe von Kreisen ist abhängig nach vom Radius
- Spitze Winkel werden unter- und stumpfe Winkel überspitzt
- Reizschwelle  $R_0$ : kleinster wahrnehmbarer Reiz (z. B. Helligkeit des dunkelsten, noch erkennbaren Sterns am Nachthimmel), quantitativs Maß (z. B. Pa beim Hören)
- Unterscheidungsschwelle  $\Delta R$ : kleinster Unterschied von einem Grundreiz (z. B. Helligkeitsdifferenz für einen Punkt, den man gerade noch auf einem grauen Hintergrund wahrnehmen kann), qualitatives Maß
- Messverfahren erfordern Wahrnehmung von Reizen: Methode konstanter Reize (Zufällige Wiedergabe von verschiedenen starken Reizen, der zu 50% erkannte Reiz ist die Reizschwelle - sehr genau, langsam), Herstellungsmethode (VP stellt den Reiz selbst ein - ungenau, schnell), Grenzmethode (Reiz erhöhen bis ihn VP erkennt - genau, langsam)
- Empfindungsstärke  $\Psi$ , Intensität des Grundreizes  $R$ , Unterscheidungsschwelle  $\Delta R$ ,  $k$  von Sinnesmodalität abhängige Konstante
- Webersches Gesetz:  $\frac{\Delta R}{R} = k$
- Weber-Fechnersches Gesetz:  $\Psi(R) = c * \ln \frac{R}{R_0}$
- Quadriert man die Stärke des Grundreizes, dann verdoppelt sich die Empfindungsstärke
- Stevensches Potenzgesetz:  $\Psi(R) = c * (R - R_0)^n$ . Entscheidend ist der Parameter  $n$ . Für Helligkeit gilt  $n = 0.3$ , für ein Gewicht  $n = 1.5$ , für die Länge einer Linie  $n = 1.0$

- Empfindungsstärken wie Lautstärke: Verwendung einer logarithmischen Einheit wie Dezibel (Weglassen von k)
- Bel: Logarithmus des Quotienten zweier Leistungen  $P_{egel} = \lg \frac{R}{R_0}$ , Dezibel: Zehntel eines Bels  $P_{egel} = 10 * \lg \frac{R}{R_0}$
- Wahrnehmung: Extraktion elementarer Merkmale (Kanten, Farbe), Präattentive Verarbeitung (Erkennen elementarer Bestandteile), Attentive Verarbeitung (Konstruktion des Gesamten)
- Wörter eines Textes werden nicht Buchstabe für Buchstabe wahrgenommen (Auge liest 6-8 Buchstaben und springt dann weiter), Hören funktioniert ähnlich
- Gesetz der guten Gestalt / Prägnanz: Übergeordnetes Gesetz zur Wahrnehmung von Gegenständen, Zusammenführung von prägnanten Einzeleindrücken zu Gestalt
- Gesetz von Figur und Grund: Figur wird als vor dem Hintergrund stehend wahrgenommen, Figur und Grund kann nicht gleichzeitig wahrgenommen werden
- Gesetz der Gleichheit: Ähnliche Dinge erscheinen zu zusammengehörigen Gruppen geordnet
- Gesetz der Nähe: Dinge, die sich nebeneinander befinden, erscheinen als zusammengehörig
- Gesetz der Geschlossenheit: Von Linien umschlossene Gegenstände werden als zusammengehörig empfunden
- Gesetz des weiterführenden oder gleichen Verlaufs: Bruchstücke reichen zum Bilden von Gestalten aus
- Gesetz der Symetrie: Symmetrische Gebilde werden als Figur wahrgenommen
- Gesetz des gleichen Schicksals: Objekte, die sich mit gleicher Geschwindigkeit bewegen, werden als zusammengehörig empfunden
- Gesetz der Erfahrung: Wiedererkennen von bestimmten Formen

- Gestaltgesetze betreffen die präattentive Wahrnehmungsphase und finden vor allem z. B. bei der Interfacegestaltung Anwendung (Gestaltung von Icons, Bedienoberflächen)
- Bei rivalisierenden präattentiven Prozessen kommt es zu Wahrnehmungstäuschungen, Linien werden gebogen
- Geone: Dreidimensionale Grundformen werden präattentiv anhand der Kantenanordnung erkannt und dann attentiv zu einem Gegenstand wie einem Tisch zusammeninterpretiert

## Kodierung

### Signale

Ein **analoges Signal** ist die Änderung einer physikalischen Größe entsprechend einem Messwert der zu übertragenen Information. Ein **digitales Signal** orientiert sich an einem festen Raster und gibt Werte aus einem "endlichen Vorrat" möglicher Werte an (diskreter Wert). Analoge Signale sind störanfällig gegenüber Rauschen (Schallplatte, Rundfunk, Foto). Digitale Signale sind unempfindlich gegenüber gegen Störsignale, solange das Signal nicht vollständig verfälscht wird oder ausfällt. 1 bit: Bit, 3 bit: Triade (Oktalsystem), 8 bit: Byte, 16 bit: Word, 16 bit: Doubleword, 32 bit: Quadword

### Entropie

Entropie  $H$  ist ein Maß für den durchschnittlichen Informationsgehalt einer Nachrichtenquelle. Es gilt:

$$H = - \sum_i p_i * \log_2(p_i)$$

Wobei  $p_i$  Auftrittswahrscheinlichkeit des Zeichens  $g_i \in A$  ist und  $A$  das Alphabet. Maßeinheit ist Bit.

### Umrechnung

Dezimal zu Hexadezimal: Zahl durch 16 teilen, Rest als Ziffer nehmen. Dann das Ergebnis der Zahl durch 16 für den nächsten Schritt verwenden (also erneut durch 16 teilen). Zahl am Ende rückwärts nehmen.

Hexadezimal zu Dezimal: Zahl von rechts nach links nehmen, erste Ziffer mit  $16^0$  multiplizieren, zweite Ziffer mit  $16^1$  etc.

## Huffman und Fano

- Alphabet: Definiert lineare Ordnung auf Zeichen. Binärcode ordnet jedem Element eines Alphabets einen Binärcode zu (...).
- Fano-Bedingung / Präfixfreiheit: Kein Zeichen ist Beginn eines anderen Zeichens
- Mittlere Länge eines Binärzeichens  $i$  mit der Länge  $N_i$  wird berechnet als:

$$L = \sum_i p_i * N_i$$

- Shannon: Es gilt immer  $H \leq L$ , Differenz kann sehr klein sein, die Differenz  $R = L - H$  wird Redundanz genannt.
- Fano-Kodierung: Zunächst werden die Zeichen so auf 0 und 1 verteilt, dass die Summe der jeweiligen Wahrscheinlichkeiten möglichst gleich groß ist. Dies wird dann so lange wiederholt, bis jedes Zeichen einen eigenen Code hat
- Huffman-Kodierung: Es wird ein binärer Baum angelegt, dessen Blätter zunächst das Alphabet sind. Die beiden Zeichen mit der geringsten Wahrscheinlichkeit werden genommen und mit 0 und 1 beschriftet. Etc.

## Audio

- Typischer Frequenzbereich 20 Hz - 20kHz
- Analoge Audio-Signale: Periodische Signale (Wellen), Amplitude (maximaler Wert): Lautstärke, Phase bezeichnet den Zeitraum der Wiederholung, Obertöne sind Vielfaches einer Grundfrequenz
- Diskretisierung: Festes Raster von Messpunkten wird festgelegt
- Quantisierung: Rundung der Messwerte auf das Raster / Darstellung als Binärzahl

- Auflösung ist die Anzahl Bits pro Sample - Sample: Ein Wert / eine Abtastung
- Bei zu geringer Abtastrate kann sich ganz anderes Muster ergeben (konstanter Ton statt Schwingung)
- Abtasttheorem:  $f_{abtast} > 2 * f_{max}$ , damit alle Signale dargestellt werden können
- Pulse Code Modulation: Analoge Spannung in digitale Daten umwandeln (analoges Signal über analoge Filterstufe über zeitliche Abtastung über Quantisierung über Codierung zum binären Signal)
- Filter: Tiefpass mit Grenzfrequenz  $f_s$
- Analogsignal wird diskretisiert, 2er-Komplementbildung
- Audio-CD: 44100hz, 16 bit / Sample, 2 Kanäle (Stereo), d. h. 10MB/min. 12 cm Durchmesser, Umdrehungsgeschwindigkeit je nach Entfernung vom Mittelpunkt. Oberfläche besteht aus Pits und Lands (Löcher und Hügel). Hoch redundant, durch 16 bit pro Sample.
- BluRay-Disc: Kürzere Wellenlänge als bei DVD und CD, bessere Fokussierung, 36 MB/s, 25 GB pro Layer, MPEG-2 kodiert, beidseitige Laserabtastung
- RIFF (Resource Interchange File Format): WAV und AVI. Aufbau in Chunks (mit Chunk-ID, Chunk-Size, Chunk-Data)
- WAV: 3 Chunks. Header-Chunk: "RIFF" (4 Byte), Dateigröße (4 Byte), "WAVE" (4 byte), Format-Chunk: "fmt" (4 byte), 15 (4 Byte) als Chunklänge, Format (2 Byte, 1 für PCM) Modus (2 Byte, 1 für Mono, 2 für Stereo), Samplefrequenz in Hz (4 Byte, bspw. 44100), Byte pro Sample (2 Byte), Bit pro Sample (2 Byte) Daten-Chunk
- Weitere Formate: MP3 (verlustbehaftete Kompression), MOD (Atari Musik), AU
- MIDI (30k Bit pro Sekunde), Byteweises Festlegen von Instrument, Tonhöhe, Dauer usw., Standardisiert, Erweiterbar

- Digitalisierung mit PCM: störresistente Übertragung möglich, verlustfreie Weiterbearbeitung, kostengünstige Schaltkreise verfügbar, aber Quantisierungsfehler (Rundung der Werte bei Einfügen in diskrete Tabelle) und Hintergrundrauschen bei der Aufzeichnung
- Quantisierungsfehler maximal mit Betrag von  $\frac{1}{2}$  der Quantisierungsstufe, hörbar als Rauschen
- Rauschen: Einfluss der Codierung auf die Signalqualität, z. B. Quantisierungsfehler, ab 70dB unter Maximalpegel nicht hörbar
- DPCM (Differential Pulse Code Modulation): Differenz zwischen benachbarten Abtastwerten werden quantisiert und codiert, Geringere Wortbreite bei gleicher Qualität (max. Kompressionsfaktor 1 : 2)
- ADPCM (Adaptive Differential Pulse Code Modulation): periodisch neu berechnete Skalierung (leise: feine Einteilung, laut: grobe Einteilung), bessere Vorhersage durch Prädiktor. Versucht, den weiteren Signalverlauf vorherzusagen, und speichert dann nur die Differenz zwischen tatsächlichem und vorhergesagtem Signal. Steile Signale teilw. nicht vollständig reproduzierbar. Bei CD-Qualität 1 : 4-fache Kompression, bis zu 32fache Kompression bei starkem Qualitätsverlust

## Kodierung

- Motivation: Akustische Täuschung. Geringe Toleranz erforderlich (Gehör genau, Abspielrate darf nicht länger als 10ms abweichen, Signal darf nicht wiederholt / ausgesetzt werden für 2ms, Gehörempfindlichkeit ist frequenzabhängig, Sprache: Inhalt wichtig [viel Semantik], bei Musik originalgetreue Wiedergabe wichtig)
- Wellenformcodierung (ISDN, PCM, WAV): Einfache Technologie, möglichst exakte Nachbildung
- Frequenzbereichcodierung (MP3, WMA, OGG Vorbis): Komplexe psychoakustische Modelle, Frequenzbänder werden in unterschiedlicher Qualität übertragen, Ziel: Möglichst genaue



Nachbildung beim Weglassen von möglichst viel Information

- Hörempfindlichkeit ist frequenzabhängig - Iso- phone: Stufen gleicher Lautstärke
- Parametrische Codierung (z. B. GSM Sprach- codierung): Speicherung von Mustersequen- zen (wie indexierte Bildformate), geeignet für Sprache (häufige Wiederholungen)
- Spracherkennung (noch unausgereift), dann auch auf Musik übertragbar
- Redundanz: Elemente ohne zusätzliche neue In- formation
- Irrelevanz: Nicht wahrnehmbare Elemente
- Entropiecodierung: Verlustfreie Kodierung, Eli- minierung der Redundanz
- Reduktion: Eliminierung von irrelevanten Daten
- Dekorrelation: Umformung in andere Darstel- lung ohne Datenmenge zu ändern
- Codec (Encoder / Decoder): Kombination ver- schiedener der oberen Techniken zu einem kom- plexen Modul
- Pipeline: Signalaufbereitung (Abtasttheo- rem, Quantisierung), Signalzerlegung (Deko- rrelation, Frequenzzerlegung o. ä.), Quan- tisierung (Perzeptionsmodelle, Reduktion), En- tropiekodierung (Statische Modelle, z. B. ZIP)
- Mithörschwelle: Laute Töne maskieren leise Töne, Frequenzanteile unterhalb der Kurve we- glassen (laute Töne maskieren bis zu 20ms vor dem Ton und bis 200ms nach dem Ton, abhängig von Lautstärke, Frequenzverhältnis, zeitlicher Lage, Struktur / Dauer des Maskierens)
- Irrelevanzkodierung: Herausfinden der nicht wahrnehmbaren Teile, Anpassen der Quan- tisierung, Löschen der irrelevanten Teile. Dadurch mehr Quantisierungsrauschen, also Kompromiss aus Qualität und Bandbreite nötig
- Annäherung eines Rechtecksignals durch unger- ade Vielfache einer Grundfrequenz

- Frequenzspektrum: Graph mit Pegel auf Y- Achse und Frequenz auf X-Achse
- Interferenz: Überlagerung von Wellen (Kon- struktiv: Addierend, Destruktive: So, dass Welle sich ausgleicht)
- Fourier-Transformation (FFT) und Diskrete Kosinus-Transformation (DCT) erlauben die Umrechnung aus dem Signalraum (Frequenz über Zeit) in den Frequenzraum (Vorkommen von einzelnen Frequenzen)
- Filterbank zerlegt Signal in Frequenzverbänder, Gesamtenergie des Signals meist ungleichmäßig verteilt, viele Bänder mit wenig Energie, wenig Bänder mit viel Energie, pro Frequenzband Maskierung bezüglich Irrelevanz (Hörschwelle, Verdeckung), Bitstrommultiplexer erzeugt digi- talen Datenstrom (inverse Filterbank setzt Sig- nal fast perfekt zusammen)
- MP3: Zerlegung in 32 Frequenzbänder, pro Teil- band 18 Subbänder, Rauschen minimieren, En- tropiekodierung, verdeckte Signaleile erkennen, optimale Codierung finden, 1152 Samples pro Frame, 30 - 230 kbit/s
- Audioverarbeitung: Veränderungen der Ampli- tude oder der Frequenz eines Filters, Lautstärke wird durch den Pegel bestimmt (0dB: 1 mW bei 0,775 Volt)
- Effektivwert  $s_{effektiv} = \sqrt{\frac{1}{T} \int_T s^2(t) dt}$
- Gefahren der Pegelanpassung: Übersteuerung (die höchsten Signalwerte liegen außerhalb des quantisierten Bereichs, Clipping, dadurch Störsignale), Untersteuerung (Die höchsten Signalpegel liegen weit unter dem Maxi- malpegel, dadurch mehr Rauschen und gerin- gerer Signal/Rausch-Abstand)
- Normalisierung: Anheben auf 0dB Pegel, Her- stellung einer Symmetrie um den Nullpunkt
- Hüllkurve: Veränderung, um zeitlich beschränkte Pegelanpassungen vorzunehmen, Ein- und Aus- blenden (Fading) durch Multiplikation mit einem Faktor zwischen 0 und 1

- Filter berechnen einen neuen Signalwert auf Ba- sis eines Signalwerts und seiner Nachbarn (Ana- log: Kondensator, Spule, z. B. Hall, Digital: speziell in digitalen Signalprozessoren)
- Hochpass, Tiefpass: hohe/tiefe Frequenzen wer- den durchgelassen
- Bandpass: Nur ein Frequenzband wird durchge- lassen, Bandsperre: Bestimmtes Frequenzband wird *nicht* durchgelassen
- DSPs arbeiten im Frequenzraum
- Dynamik: Verhältnis zwischen dem größten und kleinsten Amplitudenwert in einem Zeitfen- ster, Dynamikkompression: Laute Stellen wer- den abgesenkt, leise Stellen angehoben
- Resampling: Abspielen des Signals mit einer an- deren Samplefrequenz; doppelt so schnelles Sig- nal dauert halb so lang und klingt eine Oktave höher
- Time stretching: Resampling mit periodis- cher Wiederholung, wodurch das Signal länger dauert, aber nicht in der Tonhöhe verändert wird
- Pitch shifting: Zeitgestrecktes Signal wird mit Originalfrequenz abgespielt
- Phasing, Chorus, Flanging: Verändern die Phase des Signals (Chorus: Addition phasenversetzter Kopien)
- Echo und Hall: Basieren auf Reflexion des Sig- nals, bei Hall wird Raumaufbau mit einbezogen

## Video

## Historie

- Laterna magica (1671): Ähnlich zu Diaprojek- tor, mit Glasscheiben, die gegeneinander bewegt werden
- Daumenkino (1868), Lebensrad (1833)
- Wundertrommel: Streifen mit Zeichnungen in einem drehenden Kreis, Bild wird durch Schlitze beobachtet - Praxinoskop: Spiegel im Inneren,

die auf die Innenseite des drehenden Kreises zeigen und so das gerade vorbeilaufende Bild zeigen, Rotoskop: Daumenkino mit Fotos

- Zelluloidfilm ab 1888 (Photoemulsion auf Zelluloid): Verschiedene Filmformate, heute 35mm, 24 Bilder / Sekunde, 1km film ergibt ca. 36,5 min
- Kinemacolor (1906): 32 Bilder / sek, 16 rot, 16 grün mit rotierendem Filter
- Technicolor (1917): Aufspaltung mit Prisma auf 3 Filme
- Agfacolor (1932): als Farbfilter wirkende Körner auf dem Papier
- Edison Phonograph (1877): Nadel schreibt Ton auf Walze mit
- Lichttonverfahren (1922): Lichtdurchlässigkeit proportional zur Amplitude des Tonsignals
- Später: Kombination mit Magnetton (vgl. Tonband, Cassette)
- 4-Kanalton wird als Matrix in die Lichtkanäle moduliert, Dolby 5.1 Surround auf CD oder digital in Filmperforation, 8-Kanal-SDDS komprimiert digital auf Filmaußenrand
- Interaktive Filme: Mehrere Perspektiven werden aufgenommen und durch den Benutzer gewechselt
- Rendering: Berechnung jedes Bildes durch Rechner (Animationsfilm)
- Kamera (Edison: Kinetograph 1891). Unbelichteter Film wird von Spule in Kamera gerollt, ein Bild belichtet, Blende verdeckt das Bild, Film wird auf zweite Spule gerollt und anschließend entwickelt
- Projektoren für 35mm: Kinetoscope (1892 Edison) mit einem Betrachter, ab 1895 auf Leinwand
- Projektor in modernem Kino: Film auf 3 Teillern gewickelt, rein ins Projektor, raus auf freien Teller
- Digitalprojektor: digitales Bild wird gelesen, Lampe wird extra gekühlt

## Formate

- Flüssige Bewegung ab ca. 20 FPS, Film: 24 Hz
- Flimmern: Bei Dunkelpausen zwischen Bildern, dadurch unangenehmes Flackern
- Flimmerverschmelzungsfrequenz ca. 50 Hz, steigt mit Helligkeit
- Bilder zweifach projiziert, Verdopplung der Dunkelpausen, Flimmerfrequenz 48 Hz
- Computermonitore: hohe Helligkeit, geringer Abstand, hohe Aufmerksamkeit. Frequenz muss mindestens 72 Hz sein
- Keine Dunkelpausen bei TFT-Displays, dadurch kein Flimmern
- Helligkeitsempfindung unterstützen, abhängig vom Zustand des Auges (Auge kann sich an Leuchtdichteunterschiede anpassen, von mondloser Nacht zum sonnenbeschiehenden Schneefeld in 30 min). Das angepasste Auge kann ca. 200 Helligkeitswerte unterscheiden
- Fernsehen: Streulicht schränkt Kontrastwahrnehmung ein, Umfeldleuchtdichte sollte ca. 10 % der Spitzenleuchtdichte des Bildschirms betragen
- Aliasing: Treppenstufen und Geisterfrequenzen bei Abweichung von abgetastetem und abtastenden Raster
- Zeitliches Aliasing (Wagenradeneffekt): Wagenrad scheint sich rückwärts zu drehen bei zu geringer Abtastfrequenz, Stillstand falls genau um 90 Grad verdreht. Vermutlich durch neuronale Adaption verursacht
- Videosignal (s/w analog): Zeilenweise Abtastung der Helligkeitswerte, serielle Übertragung der Bildpunkte, synchrone Darstellung auf dem Monitor
- Progressiv: Ganzes Bild für Zeile für Zeile übertragen

- Interlaced: 2 Halbbilder werden nacheinander übertragen und ergeben zusammen ein Vollbild. Das erste Bild beschreibt nur jede zweite Zeile etc. Ursprünglich zur Reduktion von Bildflimmern entwickelt (50 Hz bei Röhrenmonitoren)
- 100Hz-Röhrenfernseher und LCDs sowie Plasmafernseher zeichnen progressiv, dadurch Umrechnung des TV-Signals erforderlich
- Weave: Gleichzeitige Darstellung der Halbbilder. Funktioniert bei TV nicht, da die Halbbilder zeitlich versetzt aufgenommen werden und so Bullshit entsteht.
- Unschärfe: Weave mit Weichzeichnen
- Motion Compensation: Weave, wobei ungerade Zeilen an gerade Zeilen angepasst werden (aktuell)
- Bobbing: Jedes Halbbild wird durch Interpolation zu einem Vollbild erweitert, um 50 Hz zu erreichen. Dadurch wackelt das Bild, da erste und letzte Zeile jeweils ohne Nachbar sind
- Blending: Bobbing, wobei beide Halbbilder zum Vollbild ergänzt werden und dann der Mittelwert zwischen beiden Bildern genommen wird
- BAS: Helligkeit auf 75 % Signalamplitude skaliert, Austastung durch Zeilenrücklauf (100 % Amplitude für 5 Mikrosekunden), Bildsynchronisation durch volle Amplitude für 10 Mikrosekunden
- FBAS: Farbsignale werden überlagert
- Komponentenvideo: Getrennte Übermittlung von drei Farbkanälen (RGB, YUV oder YIQ)
- Y/C-seperiertes Video: Trennung von Helligkeit und Chrominanzsignalen
- Europäisches Standardsignal PAL (Phase Alternate Line System): interlaced FBAS in YUV, Auflösung 833x625 Pixel, 720x576 sichtbar, 50 fields = 25 fps
- Bandbreite Y: 5,5Mhz, U/V: 1,8 Mhz

- NTSC (National Television Systems Committee): US Standard, Interlaced FBAS in VUY, 700x525 Pixels, 29,97 fps, ohne Kontrolldaten 485 Zeilen, wovon 480 nutzbar sind
- Secam (Sequential Couleur avec Memoire): wie PAL mit Bandbreiten 6/2/2 (Y/U/V)
- NTSC in Amerika, PAL in Europa, Australien, Asien, Secam in Frankreich / Russland
- DVB: Digital Video Broadcast, MPEG2-basiert, Senden als Stream, MPEG Digital Storage Command Control (DSM-CC) für Interaktivität. DVB-C (Cable), DVB-T (Terrestisch über Antenne), DVB-S (per Satellit), Container nach Objektkarussell transportiert Audio, Video und Daten. DVB-H für Verteilung über Mobilfunk nicht mehr verbreitet
- HbbTV (Hybrid Broadcast Broadband TV): DVB mit CE-HTML, JavaScript und Medieninformation (JavaScript zur Zeit nicht standardisiert), OIPF Medienformate verschieden. TV-Programm wird mit Signal mitgeliefert, Application Information Table (AIT) wird bei jedem Sendersignal mitgeliefert und enthält URL für Aufruf beim Drücken der roten Taste
- Bildebenen: schwarzer Hintergrund, Video, Untertitel / Gebärden, HbbTv, Geräteoberfläche (Uhrzeit, Lautstärke). Anwendungen werden über HTTP vom Sender ermittelt.
- Bei Video besonders sinnvoll, Differenz zwischen zwei Bildern zu berechnen (vgl. DPCM)
- H.261: Video-Telefonie über ISDN (Bandbreite 64 kbit/s mal Kanal), weniger als 150ms Verzögerung durch Kompression / Dekompression, CIF / QCIF, YCrCb 4 : 2 : 0, Kompression auf 1,5Mbit / s (24x ISDN)
- H.263: Weglassen unwichtiger Teile und Verringerung der Bittiefe, halbe Datenrate
- Intraframes (JPEG codiert): 16x16 große Blöcke, Unterteilung in 8x8-Blöcke für Helligkeit und zwei Blöcke für Chromakanäle, DCT mit anschließender Huffman-Codierung
- Predicted Frames: Unterteilung in 16x16 Blöcke, für jeden Block Suche nach "Best Match" im vorherigen Frame, Differenzcodierung
- MPEG (Moving Pictures Experts Group): Standard für komprimiertes Speichern von Audio und Video. MPEG-1 (1992, nur progressiv für Videos), MPEG-2 (1993, TV, HDTV), MPEG-3 (HDTV, in MPEG-2 integriert), MPEG-4 (laufende Entwicklung), MPEG-7 (DescriptionInterface), MPEG-21 (Digital Multimedia Framework)
- Sequence: Alle Bilder vom  $s_{header}$  bis zum  $s_{end}$
- GOP: Group of Pictures zwischen 2 GOP-Headern
- Macroblock: 16 x 16 Pixels
- Slice: 16 Zeilen
- Block: 8 x 8 Elemente (Luminanz / Chrominanz / DCT-Koeff.)
- Sample: Einzelner Helligkeitswert eines Pixels
- Bei MPEG-1 werden B-Frames eingeführt, die von vorne und hinten interpoliert werden (Macroblockweise mit DCT und Huffman-Kodierung), I-Frames JPEG-kodiert, ansonsten wie H.261
- Audio: Gliederung in Frames (Folge von Abtastwerten), Audio Access Units (kleinste für sich voll dekodierbare Einheit) und Slots (1 - 4 Byte).
- Video aus 6 Schichten: Sequence Layer, Group of Pictures layer (Abfolge der verschiedenen Bildtypen), Picture layer (kodierte Einzelbilder), Slice Layer (Ebenen, Komponenten) eines Bildes, Macroblock Layer (Makroblöcke), Block Layer (Einzelblöcke)
- MPEG-2 zu MPEG-1: Best-Match-Suche auch nach Fields, Makroblöcke Farb-Subsampling von 4 : 2 : 2, 4 : 4 : 4, die Frame-Größe kann bis zu  $16383^2$  Pixel betragen, nichtlineare Quantisierungstabelle
- Verschiedene Profile (High Level: FullHD, Main Level: HD), in höheren Profilen räumliche Skalierbarkeit (Datenstrom kann in verschiedenen Auflösungen gezeigt werden), Zeitliche Skalierbarkeit (Datenstrom adaptiv zu Veränderungen)
- MPEG-1 bis HD (1,5 Mbps), MPEG-2 für DVD oder Broadcast oder HDTV: 4,6 Mbps, MPEG-4 20 Kbps - 6 Mbps für TV, Virtual Games, Streaming [...]
- Neue Interaktivität in MPEG-4 gewünscht
- MPEG-4: Kodierung von Szenen von audiovisuellen Objekten, Kombination natürlicher und synthetischer Medien-Objekte, Prinzip unabhängig von Bitrate (low bitrate bis lossless)
- 3D-Szenengraph: Objektbaum
- X3D, VRML, OpenGL: Rendering-Pipeline (Tessellation, Transformation, Sichtbarkeit, Beleuchtung, Projektion, Clipping, Transformation in affine Koordinaten, Viewport mapping, Scan conversion, Verdeckung, Scissoring, Shading / Texturisierung, AntiAliasing)
- VRML: "SVG in 3D" ohne XML, z. B:

```

#VRML V2.0 utf8
Shape {
  geometry Box{ size 2 2 2 }
}

```

## YUV

- Farbwahrnehmungsbezogenes Chroma-Subsampling  $Y : C_R : C_B$ , Y-Abtastung 3 Mhz,  $C_R$  und  $C_B$ -Abtastungen ev. geringer, Räumliche Position wird angedeutet
- AVI: Containerformat für beliebigen Codec
- DV-Standard bei Camcordern (DCT, D1, D2..., Digital Betacam), CCIR 601 o. ä.
- MJPEG: JPEG-basierte Kompression der Vollbilder, nicht standardisiert
- MPEG: Standard

# Multimodale Systeme

Multimodale Systeme nutzen mehr als einen Sinneskanal zur Wahrnehmung beim Menschen. Dadurch Verbesserung der Intuition und Erhöhung der Robustheit eines interaktiven Systems z. B. durch Einbringung von Lippenbewegungen, Kommunikation kann barrierefreier gestaltet werden.

- Pipeline zur Wahrnehmung: Stimulation, sensorischer Kurzzeitspeicher, Wahrnehmung, Entscheidungsfindung, Reaktion, ggf. Gedächtnisspeicherung. Aufmerksamkeit ab Wahrnehmung nötig.
- Beispielsweise Datenhandschuh, Joystick

## Stereoskopische Ausgabe

- Shutter-Brillen, Head Mounted Display (Problem: Cyber-Sickness weil Bewegungen un-

natürlich)

- Head Mounted Display: Kopfhörer, Positionssensoren liefern Daten über Lage und Blickrichtung
- Autostereoskopische Techniken: z. B. 3D-LCD-Display mit Linsensystem, führt Kopfbewegungen mit Head-Tracking-System nach (Brechwinkel bedingen Abstand), Objekte können vor und hinter dem Display liegen
- Autostereoskopische Displays sind noch nicht ausgereift: Meist nur ein oder zwei Betrachter, nur Raumillusion, feste Betrachtungsposition, "Blick durch Fenster" (nicht sehr realitätsnah), Objekte, die nur von einem Auge gesehen werden, müssen ausgeblendet werden, noch zu reaktionslahm

- Polarisierte Lichtquellen für 3D: Bild wird polarisiert statt durch Farbfilter überlagert, je ein Auge sieht horizontal oder vertikal polarisiertes Licht
- STB: Interaktiver Empfänger für Fernsehen und Internet. Bindeglied zwischen HiFi, TV, PC etc und Internetz. Standards für APIs, Netzprotokolle, Medienformate. Unterschiedliche Hardwarekonzepte, aber meist Tunermodul, Demodulator, Prozessor, MPEG2-Decodierer, Demultiplexer, Backchannel-Controller, Descrambler, Schnittstellen-Controller, Audio DAC, Entschlüsselungssystem (Conditional Access)