

Grundlagen Multimediatechnik Zusammenfassung

WS19/20 Prof. Dr. Enkelejda Kasneci

Einleitung

Was ist ein Medium / Multimedia:

Ein Medium dient zur Verarbeitung und Darstellung von Informationen. Multimedia ist das gleichzeitige Anwenden mehrerer Medien.

Teilaspekte von Multimedia:

- Multimedialität (kann aus mehreren verschiedenen Medien bestehen)
- Multicodalität (kann unterschiedliche Inhalte sein z.B. Text, Bilder...)
- Multimodalität (kann unterschiedliche Sinne des Menschen ansprechen)
- Interaktivität (Nutzer kann mit dem Medium interagieren)

Arten von Medien:

- Perzeptionsmedien (Wie nimmt der Mensch die Informationen auf)
- Präsentationsmedien (Unterscheidung zwischen Ein- und Ausgabemedium z.B. Mikrofon und Lautsprecher)
- Übertragungsmedien (Worüber werden Informationen übertragen)
- Speichermedien (Worauf werden Informationen gespeichert)
- Repräsentationsmedien (Wie werden Informationen kodiert)

Visuelle Wahrnehmung:

Wahrnehmung erfolgt über Licht (Teilchen). Die Frequenz bestimmt die Farbe und die Amplitude bestimmt die Helligkeit.

Farbwahrnehmung:

- unterschiedliche Farbspektren erzeugen selben Farbeindruck (Metamerie)
- menschliches Auge kann keine Farbspektren zerlegen

Farbmischung:

- aktive lichterzeugende Medien:
 - additive Farbmischung: in digitalen Medien RGB(0-255) (spektrale Intensitäten einzelner Lichtkomponenten werden addiert)
- reflektierende Medien:
 - subtraktive Farbmischung: in Printmedien CMY(0-100) (spektrale Intensitäten werden dem Farbwert des weißen Lichts entfernt)

Farbräume digitaler Medien RGB:

- Farbkanäle R (rot), G (grün), B (blau) je 8 Bit
- 0 = kein Farbanteil
- 255 = maximaler Farbanteil

Farbräume Printmedien CMY:

- Farbkanäle C (cyan), M (magenta), Y (yellow)
- 0 = Papierfarbe (keine Farbe)
- 100 = maximaler Farbanteil

Signalverarbeitung

Fourierreihe:

- Jede Schwingung kann als Summe von harmonischen Schwingungen dargestellt werden
- Fourier-Analyse und Fourier-Transformation zerlegt Schwingung in ihre Frequenzbestandteile
- Fourier-Synthese und inverse Fourier-Transformation zur Zusammensetzung

Signalklassen:

- analog:
 - Nimmt im zeitlichen Verlauf kontinuierliche Werte an
 - Störungen beeinflussen Signal
- digital:
 - Nimmt im zeitlichen Verlauf nur diskrete Werte an
 - Störungen beeinflussen Signal nur im Extremfall
 - verlustbehaftet, da sie in Stufen dargestellt werden

Digitalisierung:

- Abtastung (zu grobe Abtastung)
- Zwischenspeicherung
- Quantisierung (zu grob -> schlechte Darstellung der Abstufungen)

Abtasttheorem:

- Ursprüngliches Signal lässt sich nur aus dem abgetasteten Spektrum rekonstruieren, wenn die Abtastrate $\frac{1}{T}$ größer ist als das doppelte der Grenzfrequenz f_{\max}
- Ist die Abtastrate kleiner => komplett andere Funktion wird rekonstruiert (Aliasing)
- Ist die Abtastrate größer bis genau das doppelte der Grenzfrequenz => breitere Spektren bzw alle Nullstellen abgetastet (Aliasing)

Aliaseffekt:

- Die Abtastungen werden als Vielfaches der Grundfrequenz interpretiert
- Signal kann sich innerhalb des Spektrums beliebig oft wiederholen
- Verhindern durch das Einhalten des Abtasttheorem
- Verhindern durch Tiefpassfilter (Anti-Aliasing-Filter)

Bildkomprimierung

Ortsauflösung:

- Anzahl der Pixel pro Zeile/Spalte in digitalem Bild

Kontrastauflösung:

- Anzahl möglicherweise vorhandenen Farbwerte (0-255) innerhalb eines Bildes

Chroma Subsampling:

- Beispiele der Farbunterabtastungen: 4:4:4 , 4:2:2 , 4:2:0 , 4:1:1 ...
- Bildschärfe Eindruck findet überwiegend über die Helligkeitsinformationen statt.
- Farbinformationen können also reduziert werden. (Pixel teilen sich Informationen)

Speicherung in JPEG:

- Bildvorverarbeitung (Chroma Subsampling)
- Transformation in den Frequenzraum (Diskrete Cosinus Transformation)
- Quantisierung (Kodierung höherer Frequenzanteile in weniger Bits)
- Entropiekodierung (Laufängerkodierung, Huffman-Kodierung)

Diskrete Cosinus Transformation DCT:

- Zerlegung in Wellen unterschiedliche Frequenzen (ähnlich wie Fourier Transformation)
- Transformation ist invertierbar wenn Basisfunktionen Orthogonalbasis bilden
- Bildbereiche werden in Koeffizienten für Frequenz-Repräsentation transformiert
- DC-Koeffizient stellt Grundton des Bildbereichs dar
- AC-Koeffizient geben Anteil höherer Frequenzen an
- Rundung der Koeffizienten verursacht viele 0er (ähnliche Informationen trotz Informationsverlust => Kompression)

Bildverbesserung

Anwendungsgebiete:

- Zeichenerkennung
- Qualitätsprüfung in industrieller Produktion
- Medizinische Bildanalyse
- Luftbildauswertung
- Gesichtserkennung
- Fahrzeugsteuerung
- Inhaltsbasierende Bildsuche im Internet

Histogramm:

- beschreibt die Häufigkeitsverteilung der einzelnen Intensitätswerte in einem Bild
- Eigenschaften des Bildes direkt ableitbar:
 - Dynamik: Anzahl unterschiedlich vorkommenden Pixelwerte
 - Dynamik ist unabhängig vom Kontrast
 - Kontrast: Abstand zwischen minimal und maximal vorkommenden Intensitäten
 - globaler Kontrast (größter Grauwertunterschied im Bild)
 - lokaler Kontrast (durchschnittlicher Grauwertunterschied zwischen Nachbarpixel)
 - Größerer globaler Kontrast bedeutet größerer lokaler Kontrast. Da sich der Abstand zwischen den min. und max. Intensitäten vergrößert, vergrößert sich auch der Grauwertdurchschnitt zwischen den Nachbarpixeln

Lokale Operation:

- Funktion der Bildpunkte die in einer wohldefinierten Umgebung innerhalb des Eingabebildes liegen
- Bsp. Mittelwertfilter, Medianfilter, Min-/Maxfilter

Globale Operation:

- Funktion aller Bildpunkte eines Eingabebildes
- Bsp: Diskrete Cosinus Transformation, Fouriertransformation

Punktoperation:

- Bildverbesserung pixelbasierend
- Abbildung der Farb- bzw Grauwerte unabhängig von Ort und ihrer Zuordnung
- Grauwertspreizung, Log-Transformation

Bildanalyse

Filter:

Bildfilterung ist die Faltung eines Bildes mit einer Filtermatrix

- Boxcar Filter (lineare Glättung)
- Min-/Maxfilter (nicht-lineare Glättung)
 - Min Filter (eliminiert weiße Punkte und verbreitert dunkle Stellen)
 - Max Filter (Gegenteil: eliminiert dunkle Punkte und verbreitert weiße Stellen)
- Medianfilter
 - Alle Elemente in der Filterregion (z.B. 3x3) sortieren
 - Mittleren Element der Sortierten Liste als neuen Pixelwert verwenden
 - Wenn Anzahl der Elemente gerade ist so nimmt man den Mittelwert der beiden mittleren Elemente
- Gauß-Filter
 - Rauschunterdrückung, Glättung
 - Grenzen: Salt-Pepper Rauschen kann nicht entfernt werden.

Kantendetektion:

- Prewitt
- Sobel
- Roberts
- Kirsch
- Laplace
- Laplacian of Gauss

Videokompression:

I-Frames:

- Videodarstellung durch Aneinanderreihen von Einzelbildern
- Vorteil:
 - direkter Zugriff auf Bild
 - kann ausgelassen werden wenn der Frame nicht decodiert werden kann
- Nachteil:
 - Kodiereffizienz da keine Abhängigkeiten definiert

P-Frames (Prädiktiv-kodiert):

- Es wird ein I-frame kodiert und die folgenden Frames durch Bewegungsschätzung generiert
- Vorteil:
 - hohe Kodiereffizienz
- Nachteil:
 - muss ein Frame dekodiert werden müssen alle vorherigen auch dekodiert werden
 - kann ein P-Frame nicht dekodiert werden so kann der rest auch nicht dekodiert werden

B-Frames (bidirektional prädiktiv):

- kodiert Unterschied zu vorherigem oder folgendem I- bzw. P-Frame
- Vorteil:
 - höhere Kodiereffizienz
 - schneller Zugriff auf beliebige Bilder
 - können ausgelassen werden
- Nachteile:
 - verschwendete Bits, da sie kein Beitrag zu anderen Bildern leisten
 - höhere Komplexität und mehr Speicherplatz
 - Kodierreihenfolge entspricht nicht der Anzeigefolge

Group of Picture GOP:

- Gruppe von in Abhängigkeit kodierten aufeinanderfolgenden Frames
- GOP hat 1 I-Frame und eine bestimmte Anzahl an P- und B-Frames
- Bsp: **IBBPBBPBBIBBPBBPBBP** (zwei GOP aneinander)

Audiokompression

was muss (nicht) kodiert werden:

- simultane Verdeckung: starkes Signal verdeckt gleichzeitiges schwaches Signal
- temporäre Verdeckung: starkes Signal verdeckt schwaches Signal nicht nur gleichzeitig sondern wirkt vor bzw. nach.
- Datenstrom in Frames zerlegen
- Frequenzbereiche aufteilen auf 32 Subframes
 - Frequenzmaskierung
 - Zeitmaskierung
 - variable Breite
- Lauteste Frequenzanteile verringern benötigte Auflösung
- Differenz zwischen linkem/rechtem Kanal
- nach psychoakustischem Modell quantisieren
- Huffman kodieren