

Bildunterschriften zu Bildkompression

- **Folie 2:** Licht (links oben durch schwarze Linien angedeutet) transportiert visuelle Information über Objekte (durch einen Pfeil angedeutet) in Form ihrer Spiegelbilder ins Auge. Das Licht fällt dabei durch die Hornhaut (engl. *cornea*), die Pupille (engl. *pupil*) und die Linse (engl. *lens*) auf die Netzhaut (engl. *retina*). Die Netzhaut besteht aus lichtempfindlichen Rezeptoren (links unten vergrößert dargestellt), die im Netzhautgrübchen (engl. *fovea*) gehäuft auftreten. Die Rezeptoren wandeln die optischen Signale in elektrische um, die anschließend über den Sehnerv (engl. *optical nerve*) zum Gehirn gelangen.
- **Folie 3:** Zapfen (links) sind hauptsächlich auf kurz- (S, blau), mittel- (M, grün) oder langwelliges (L, rot) Licht empfindlich (rechts) – die Einfärbung im linken Bild dient der Illustration der Verteilung der verschiedenen Typen. Stäbchen (engl. *rods*, R rechts) reagieren nur auf ein Teilintervall des sichtbaren Wellenbereiches.
- **Folie 4:** Aus dem dargestellten sinusförmigen Helligkeitsmuster kann die individuelle Kontrastempfindlichkeitsfunktion abgelesen werden. Von links nach rechts erhöht sich die Ortsfrequenz exponentiell, von unten nach oben verringert sich der Kontrast (die Amplitude, d.h. die Abweichung von mittelgrau in Richtung weiß bzw. schwarz) exponentiell. Am oberen Bildrand ist eine graue Fläche sichtbar, wo das sinusförmige Muster nicht mehr als solches erkannt wird. Die Grenze zwischen diesem Bereich und jenem mit dem erkennbaren Muster wird als Kontrastempfindlichkeitsfunktion bezeichnet.
- **Folie 5:** Ein digitales Bild (links angedeutet) besteht aus (meist) quadratischen Bildelementen. Die Pixel sind in einem durchgehenden Raster innerhalb des Bildes angeordnet (rechts).
- **Folie 6:** Jede sichtbare Farbe lässt sich aus den Grundfarben Rot, Grün und Blau in entsprechenden Anteilen zusammenmischen (links). Ein Fließerton besteht beispielsweise aus hohen Rot- und Blauanteilen, aber weniger Grünanteilen. Die Grundfarbenanteile werden meist als Gewichte zwischen 0 und 255 von 255 möglichen Teilen angegeben. Die Farbmischung erfolgt additiv, z.B. bei den Bildpunkten eines LCDs (rechts), die jeweils aus roten, grünen und Blauen Teilbildpunkten bestehen, deren Licht sich in entsprechenden Anteilen überlagert.
- **Folie 8:** Bei der Chrominanzunterabtastung bleibt der Y-Kanal unverändert, während die Cb- und Cr-Kanäle absichtlich unterabgetastet werden. 4:4:4 Subsampling (links oben) bezeichnet dabei eine Unterabtastung um den Faktor 1, d.h. effektiv keine Unterabtastung. 4:2:2 Subsampling (rechts oben) reduziert die horizontale Auflösung der Farbdifferenzkanäle um den Faktor 2, 4:1:1 Subsampling (links unten) um den Faktor 4. Das in üblichen Bildkompressionsstandards verwendete 4:2:0 Subsampling (rechts unten)

reduziert die Auflösung horizontal und vertikal jeweils um den Faktor 2, nimmt aber eine Abtastposition in der Mitte der vier ursprünglichen Abtastwerte an.

- **Folie 9:** Verschiedene Arten der Chrominanzunterabtastung (links 4:1:1, Mitte links 4:2:0 und Mitte rechts 4:2:2) führen zu teilweise stark vom Original (rechts 4:4:4, d.h. ohne Unterabtastung) abweichenden Farbdifferenzkanälen (unten). Mit korrekter Luminanzinformation kombiniert (oben) ist die Unterabtastung aber nur an kontrastreichen Rändern, z.B. zwischen dem roten Schild und der blauen Lokomotive, erkennbar.
- **Folie 10:** Bei der Kompression nach JPEG wird das Bild in 8·8 Pixel große Blöcke unterteilt, die mit der DCT transformiert und anschließend quantisiert werden. Die quantisierten Werte werden in Zickzack-Reihenfolge angeordnet („gescannt“) und – sofern es sich um AC-Koeffizienten handelt – anschließend lauffängenkodiert (RLC). Der DC-Koeffizient wird gesondert gespeichert. Abschließend werden die vorverarbeiteten Codewörter huffmankodiert und mit Headern in einen Bitstrom eingebettet.
- **Folie 12:** Die Basisfunktionen der zweidimensionalen DCT mit $M = N = 8$. sind grafisch (schwarz repräsentiert -128, weiß 127) dargestellt und optisch durch blaue Linien voneinander getrennt. Die horizontale Ortsfrequenz k erhöht sich in horizontaler Richtung, die vertikale Ortsfrequenz l in vertikaler Richtung. Die Ortsfrequenzkombination $k = l = 0$ repräsentiert den Gleichanteil. Für $l = 0$ ergeben sich die rein horizontalen Basisfunktionen, die jenen der eindimensionalen DCT in dieser Darstellung entsprechen.
- **Folie 15:** Mit niedrigeren Qualitätsstufen q verringert sich die Bildqualität bei der Kompression nach JPEG. Beim Minimum von 1 (rechts unten) sind deutliche Kompressionsartefakte, wohingegen das Bild beim Maximum von 99 (links oben) annähernd dem Original entspricht. Mit dazwischen liegenden Werten von q (links unten und rechts oben) kann – je nach Anwendung – eine passende Qualität zwischen den beiden Extrema erreicht werden.
- **Folie 16:** Blockartefakte (links) bezeichnen das Sichtbarwerden der rasterförmigen Kanten der separat verarbeiteten Blöcke durch Helligkeitssprünge. Unschärfeartefakte (rechts) beschreiben verschwommene Bildbereiche, die vor der Kompression strukturiert waren (z.B. das Gras vor dem Gebäude).
- **Folie 17:** Überschwingartefakte (links) bezeichnen Rekonstruktionsfehler entlang kontrastreicher Kanten. Ein ursprünglich steiler, d.h. hochfrequenter, Helligkeitsverlauf (Sprungfunktion rechts) wird nach der Kompression niederfrequent approximiert, was zum Unter- und Überschwingen (engl. *ringing*) vor bzw. nach der Kante führt, das immer gemeinsam mit Unschärfe entlang der (weniger steilen) rekonstruierten Kante auftritt.

- **Folie 18:** Beim Zickzack-Scan werden die Koeffizienten vom Gleichanteil (DC-Koeffizienten) ausgehend in einem Zickzackmuster umkehrbar von niederfrequent nach hochfrequent umgeordnet. Da hohe Frequenzanteile nach der Kodierung meist null sind, sorgt die Umordnung für längere Nullfolgen, die von der nachfolgenden Lauflängenkodierung effizienter kodiert werden können.
- **Folie 19:** Die AC-Koeffizienten (links oben) werden nach dem Zickzack-Scan zuerst Runlevel-kodiert (rechts oben), d.h. ein Wert wird mit seinen vorangehenden Nullen gruppiert, z.B. zwei Nullen mit dem Wert -1 (zweite Gruppe). Jeder Wert, der nicht 0 ist, wird binär kodiert (hellgraue Codewörter unten). Die Anzahl der vorangehenden Nullen und die (Bit-)Länge des Codewortes (schrägstrichgetrennte Beschriftungen über den Rechtecken unten) werden Huffman-kodiert (schwarze Codewörter). Jede Kombination eines binären und eines Huffman-Codewortes (mit Rechtecken umrahmt) repräsentiert eine laulängenkodierte Gruppe (strichliert angedeutet). Nullen hinter der letzten Gruppe werden mit dem EOB-Codewort signalisiert.