

Klausurfragen vom 26.01.2015

1. Betrachtet wird ein zeitdiskretes aber amplitudenkontinuierliches Audiosignal $x(n)$ mit der Abtastfrequenz $f_A = 44.1\text{kHz}$ und gaußförmiger Wahrscheinlichkeitsdichte der Amplitude. Geben Sie das SNR in dB nach der Quantisierung mit w bit an. [1]

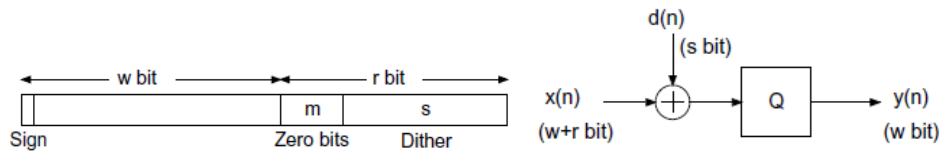
$$\text{SNR} = 6.02 w - 8.5 \quad [\text{dB}]$$

2. Was ist das Ziel von Dithering? [1]

Nichtlineare Verzerrungen, welche bei der Quantisierung von Signalen auftreten sollen linearisiert und gedämpft werden. → Töne werden klarer.

3. Wie wird Dithering durchgeführt? [1]

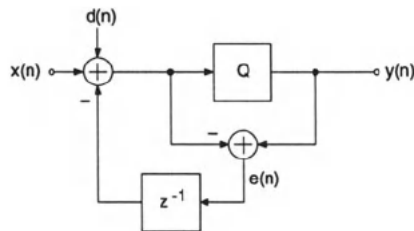
Es wird ein Zufallssignal vor der Quantisierung zum Originalsignal aufaddiert.



4. Welche Art von Dither Signal ist notwendig, um die Rauschmodulation zu unterdrücken? [1]

Durch ein dreiecksförmiges Dither-Signal (TRI-Dither).

5. Zeichnen Sie das Blockdiagramm eines Quantisierers mit Noise-Shaping. [3]



6. Geben Sie das SNR in dB eines überabtastenden AD-Umsetzers mit dem Überabtastfaktor L nach der Quantisierung mit w bit an. Die Amplitude vor der Quantisierung hat eine gaußförmige Wahrscheinlichkeitsdichte. [1]

$$\text{SNR} = 6.02 w - 8.5 + 10\log(L) \quad [\text{dB}] \quad (\text{Zölzer Seite 69})$$

[die (- 8.5) kommen vom gaussverteilten Nutzsignal]

7. Skizzieren Sie das Spektrum des Quantisierungsrauschens nach einem 4-fach überabtastenden AD-Umsetzer mit Noise-Shaping. Schraffieren Sie die Fläche des Quantisierungsrauschens im Nutzband. [3]

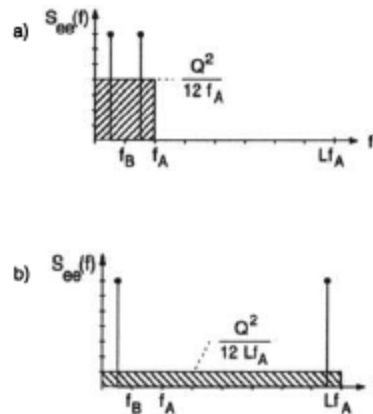
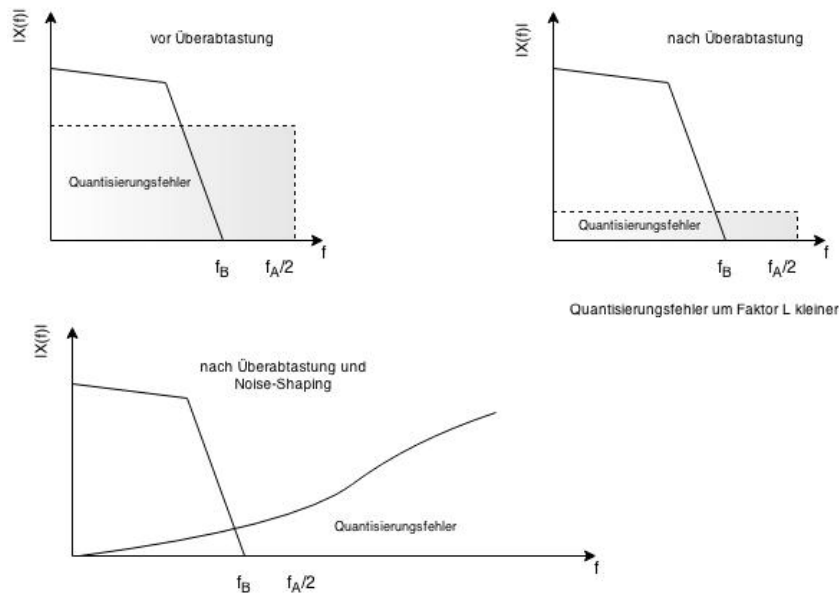


Bild 3.3: Einfluß der Überabtastung auf die spektrale Leistungsdichte des Quantisierungsfehlers

(Bild 3.3 ist ohne Noise-Shaping)



8. Zeichnen Sie das Blockdiagramm eines überabtastenden 1-bit Delta-Sigma AD-Umsetzers, der als digitales Ausgangssignal ein nicht-überabgetastetes w-bit Signal hat. Bezeichnen Sie die Wortbreiten der digitalen Signale. Zeichnen Sie auch das Blockdiagramm des Delta-Sigma Modulators (nicht das zeitdiskrete Modell). [5]

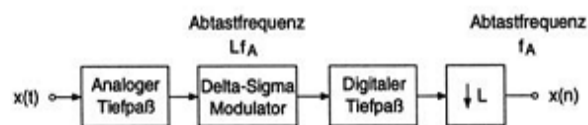


Bild 3.10: Überabtastender Delta-Sigma AD-Umsetzer

Frequenz des Delta-Sigma-Modulators und des Digitalen Tiefpasses ist für beide $L \cdot f_A$
Wortbreite nach Delta-Sigma-Modulator: **1-Bit**, Wortbreite nach Digitaler Tiefpass: **w-Bit**

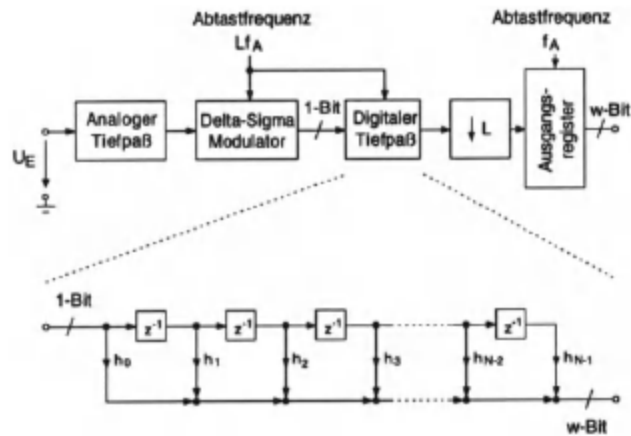
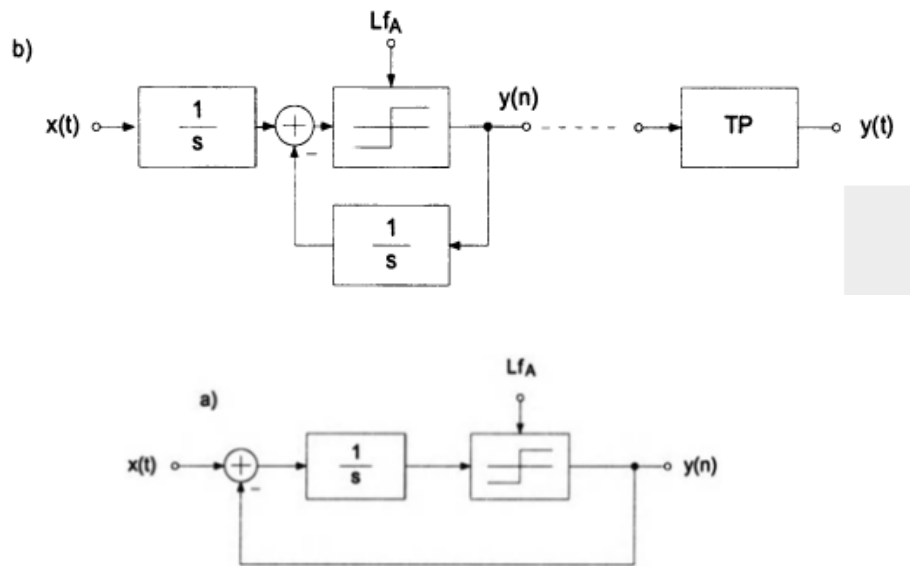


Bild 3.38: Delta-Sigma AD-Umsetzer

(Zölzer, S. 93... sieht in der 3. Auflage etwas anders aus: Der Ausgangsregler ist nicht mehr da und die w-Bit sind bereits nach dem Digitalen Tiefpass eingetragen)



(Sigma delta Modulator nach Zusammenfassung der Integrierer. Zölzer p.71)

9. Erklären Sie die Schritte der Schnellen Faltung zur Berechnung des Ausgangs-Audiosignals $y(n)$ ausgehend vom Mono-Eingangs-Audiosignal $x(n)$ der Länge $N1$ und der Impulsantwort $h(n)$ der Länge $N2$. [2]

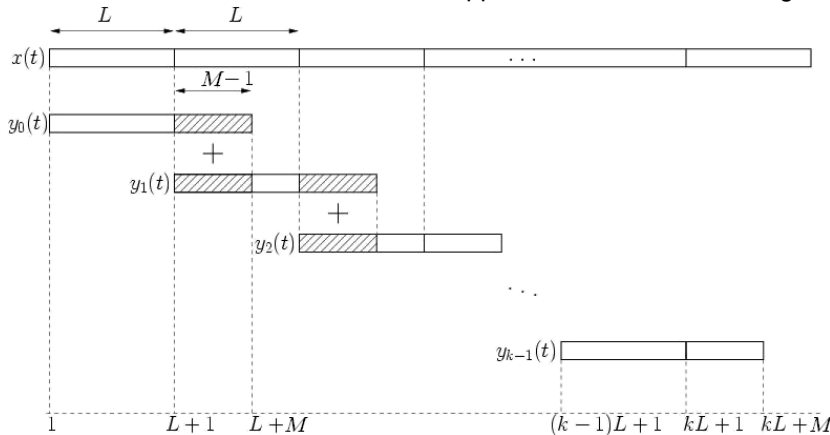
Die Faltung von $x(n)$ der Länge $N1$ mit $h(n)$ der Länge $N2$ führt zu einer Ausgangsfolge der Länge $N1 + N2 - 1$.

L: Bruchteil der Länge $N1$ bei Partitionierung des Eingangssignals (Aufteilung in kleinere Teillängen)
M entspricht $N2$

Die Faltung eines Teils des Eingangssignals $x(n)$ der Teillänge L mit $h(n)$ der Länge M führt zu einem Teilergebn der Ausgangsfolge der Länge $L + M - 1$.

Overlap-Add:

- FFT der Impulsantwort.
- FFT des nächsten Teiles der Sequenz.
- Multiplikation im Frequenzbereich anstatt Faltung im Zeitbereich
- I-FFT des Ergebnisses. Merken des überlappenden Bereichs.
- Hinzuaddieren des zuvor überlappenden Bereiches zum Ergebnis. Weiter bei Punkt 2.



10. Zeichnen Sie das Blockdiagramm einer modifizierten Oktav-Band Filterbank, die eine Aufteilung des Signals in drei Frequenzbänder vornimmt und in der kein Aliasing im Übergangsbereich auftreten kann (ohne Komplementärtechnik). Welche Grenzfrequenz schlagen Sie vor? [5]

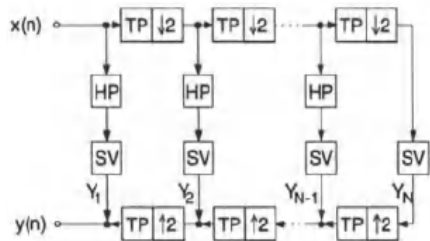


Bild 5.54: Modifizierte Oktav-Band Filterbank

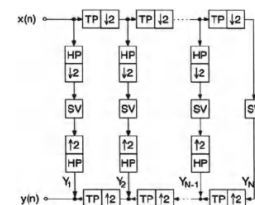


Bild 5.51: Oktav-Band QMF-Filterbank (SV=Signalverarbeitung, TP=Tiefpaß, HP=Hochpaß)

$$\Omega_{Ck} = \frac{\pi}{3} 2^{-k+1} \quad \text{mit} \quad k = 1, 2, \dots, N-1$$

(Zoelezer S. 185)

Die Grenzfrequenzen sollten z.B. wie folgt gewählt werden um Aliasing zu verhindern:

$$\Omega_{Ck} = \pi/3 \cdot 2^{-(k+1)} \quad \text{mit} \quad k=[1,2]$$

$$\Omega_{Ck} = 2 \cdot \pi \cdot f = 44100 \text{ Hz}$$

$$k = 1: \quad 2 \cdot \pi \cdot f = \pi/3 \cdot 2^0 = \pi/(3 \cdot 1) \rightarrow f = 1/6$$

$$44100 \text{ Hz} / 6 = 7350 \text{ Hz}$$

1. Band von 44100 Hz bis 7350 Hz

k = 2: $2\pi f = \pi/3 \cdot 2^{(-1)} = \pi/(3 \cdot 2) \rightarrow f = 1/12$
 $44100 \text{ Hz} / 12 = 3675 \text{ Hz}$

2. Band von 7350 Hz bis 3675 Hz

3. Band von 3675 Hz bis 0 Hz (theoretisch)

k = 3: $2\pi f = \pi/3 \cdot 2^{(-2)} = \pi/(3 \cdot 4) \rightarrow f = 1/24$
 $44100 \text{ Hz} / 24 = 1835,5 \text{ Hz}$

als Zusatzinfo wies weitergehn würde...

Tabelle 5.8: Übergangsfrequenzen f_{Ci} und Übergangsbreiten TB bei einem 8-Band Equalizer

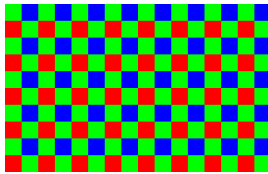
f_A [kHz]	f_{C1} [Hz]	f_{C2} [Hz]	f_{C3} [Hz]	f_{C4} [Hz]	f_{C5} [Hz]	f_{C6} [Hz]	f_{C7} [Hz]
44.1	7350	3675	1837.5	918.75	459.375	≈ 230	≈ 115
TB [Hz]	1280	640	320	160	80	40	20

11. Was versteht man in der Psychoakustik unter Frequenzmaskierung? [1]

Ein Ton verhindert, dass umliegende Frequenzen vom Menschen gehört werden können. Dieser Effekt wird bei der Audio Kodierung genutzt.

12. Beschreiben Sie das Bayer-Muster, welches auf den meisten ... Anwendung findet. Welche Vorteile hat dieses Muster gegenüber einer einfachen Abfolge von roten, grünen und blauen Pixeln? [2]

Das Bayer Pattern enthält doppelt so viele Grün-, wie Rot- und Blau-Punkte, da Grün im menschlichen Auge die meiste Information über die Helligkeit liefert. Somit können wir, physikalisch bedingt, Helligkeits- und Kontrastunterschiede besser wahrnehmen.

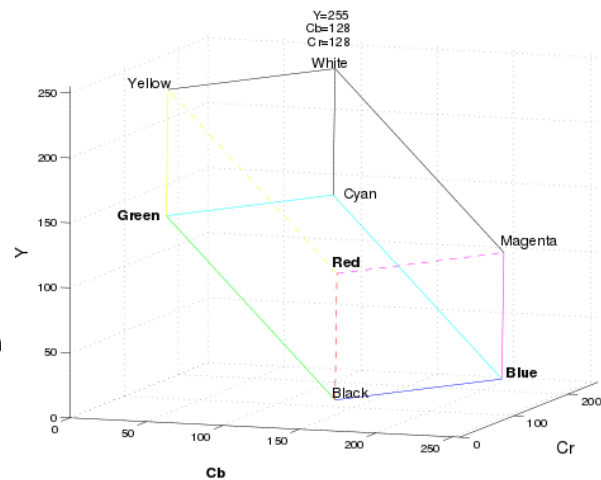
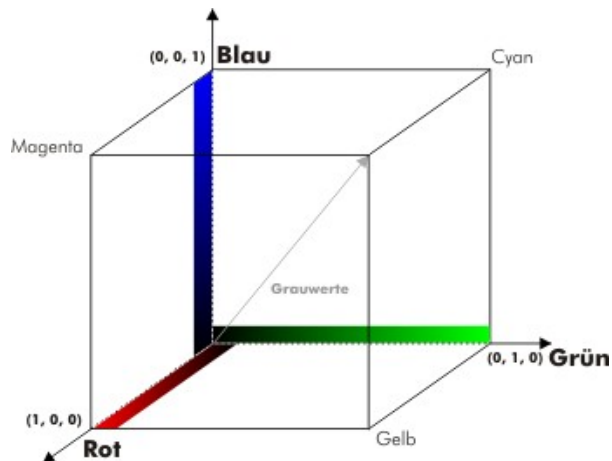


← mögliches Bayer-Pattern

13. Warum ist es zweckmäßig Farbräume als dreidimensional anzunehmen (RGB, YCbCr, ...)? Wie erklärt sich dieser Umstand physiologisch? [2]

So kann jeder darstellbaren Farbe eine entsprechende Koordinate im Raum zugeordnet werden und es braucht nur 3 Werte um alle Farben abzubilden.

14. Skizzieren Sie einen RGB-Farbwürfel und einen YCbCr-Farbwürfel (Ursprung und Koordinatensystem nicht vergessen!) und zeichnen Sie jene Linie ein, die einem Grauverlauf von Schwarz nach Weiß entspricht. [2]



15. Nennen Sie drei Artefakte, die durch das optische System bzw. den Sensor eines Bildaufnahmesystems entstehen können, und wie man diesen durch digitale Bildverarbeitung begegnet. [3]

- *) *Chromatische Aberration*: Farbfehler. Licht unterschiedlicher Wellenlänge wird verschieden stark gebrochen. Gegenmaßnahme: Kombination von Linsen.
- *) *Vignettierung*: Abschattung zum Bildrand hin durch axiale Anordnung von zwei Öffnungen. Gegenmaßnahme: Abblenden.
- *) *Sensor Rauschen*: Spontane Bildung von freien Ladungsträgern durch Wärme. Gegenmaßnahme: Rauschfilter, Kühlen (bei Astronomie)
- *) *Blooming*: Heller Streifen von überbelichteten Bildstellen. Überzählige Ladung wird an Nachbarzellen abgegeben. Gegenmaßnahme: Anti-Blooming Gates am Sensor.
- *) *Rolling Shutter Effekt*: Verzerrung durch zu langsame zeilenweise/spaltenweise Aufnahme. Gegenmaßnahme: Vollständige Aufnahme eines Bildes (z.B. mechanischer Shutter, CCD-Sensoren)

16. Warum wird in Bild- und Videoformaten oftmals der YCbCr-Farbraum gegenüber einem RGB Farbraum bevorzugt? [1]

Um Chrominanz & Luminanz Anteile getrennt voneinander verarbeiten zu koennen. (z.B. Subsampling, Helligkeitsaenderung)

17. Nennen Sie die einzelnen Stufen des JPEG Codecs und beschreiben Sie diese jeweils kurz (ein Satz je Stufe genügt). [5]

- *) *RGB → YUV Konvertierung*: Umwandeln des Informationen in den YUV Farbraum, um Luminanz und Chrominanz getrennt betrachten zu koennen
- *) *Subsampling*: Chrominanz (U-V-Planes) werden mit weniger Bit beschrieben. Das ist moeglich, da in den Farbwerten weniger Information als in der Helligkeit steckt
- *) *DCT*: Blockweise diskrete Kosinustransformation fuer jede Plane. Dadurch wird die Ortsinformation in Frequenzinformation umgewandelt (und nach Frequenz sortiert)
- *) *Quantisierung*: Durch die Quantisierungsmatrix werden die berechneten DCT-Koeffizienten (= Frequenzanteile) der Bloecke unterschiedlich gewichtet. Der DC-Anteil (links oben) ist dabei am

stärksten gewichtet, da hoch-frequente Anteile unwichtiger sind. Die hochfrequenten Anteile fallen somit bei der Quantisierung weg (werden zu Null)

*) *Run Length Encoding*: mit Zick-Zack-Abtastung über jeden separaten Block. Gleiche Werte werden zusammengefasst. Da der rechts-untere Teil aus vielen Nullen besteht, können diese effizient zusammengefasst werden.

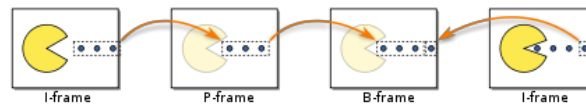
*) *Huffman Encoding*: Aufbauen eines HuffmanBaumes und daraus Erzeugen eines Huffman-Codes. Werte die am häufigsten vorkommen werden mit den wenigsten Bits dargestellt

*) Zum Dekodieren von unten nach oben lesen

18. In (hauptsächlich) welcher Stufe des JPEG Codecs wird die Kompressionsrate des Bildes festgelegt? [1]

Die Elemente der Quantisierungsmatrix legen die Kompressionsrate des Bildes fest. Mit größeren Werten wird eine höhere Kompression erreicht, da dadurch mehr Koeffizient zu Null werden. Die Standardmatrix dient für eine Kompression um 50%. Bei der Quantisierung und dem anschließenden Runden entsteht ein Informationsverlust -> Fehler.

19. Wodurch unterscheiden sich die I-, P- und B-Frames im Kontext des MPEG4 Video Codecs?



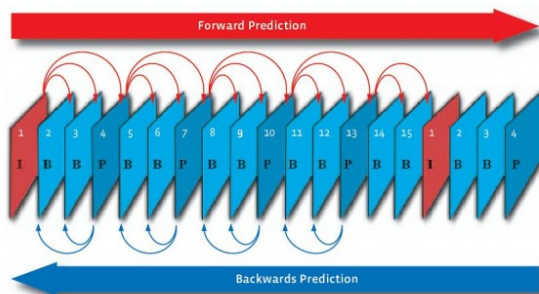
[2]

*) I (Intra)-Frames: Werden als Synchronisations Frames verwendet. Diese werden nicht aus vorhergesagten Werten erstellt.

*) B-Frames: Beschreiben ein Bild, welches aus referenzierten Bildteilen vorhergehender *und* nachfolgender Einzelbilder rekonstruiert wird. Es werden künftige und vergangene Frames referenziert, was die Latenz verlängert.

*) P-Frames: Beschreibt ein Bild, welches aus referenzierten Bildteilen vorhergehender Einzelbilder rekonstruiert wird. Es wird nur jener Teil übertragen, der Änderungen unterlegen war. Wird ein ganzer Bereich verschoben, werden die Differenzvektoren des Bereichs übertragen.

Links mit, rechts ohne B-Frame Kodierung



<http://www.streaminglearningcenter.com/articles/producing-h264-video-for-flash-an-overview.html?page=4>

http://www.axis.com/de/products/video/about_networkvideo/compression.htm

Klausurfragen vom 20.12.2011. Mögliche Punkte in eckigen Klammern pro Bsp. angegeben.

1. Zeitdiskretes und amplitudenkontinuierliches Signal wird mit 44.1kHz abgetastet. Das Signalspektrum ist Gauß-verteilt. Geben sie das SNR an, welches nach einer Quantisierung mit w-Bit vorliegt. [1]

- Sinusförmiges Signal: $\text{SNR} = 6.02 w + 1.76$ [dB]
- Rechteckförmiges Signal: $\text{SNR} = 6.02 w$ [dB]
- Gaußverteiltes Nutzsignal: $\text{SNR} = 6.02 w - 8.5$ [dB]

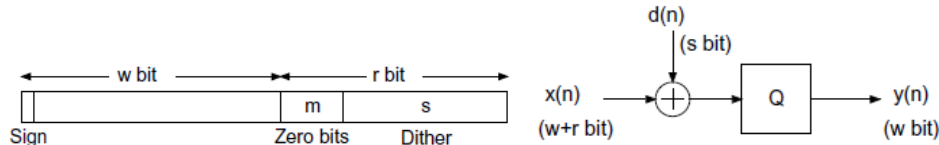
2. Ziel von Dithering [1]

Reduktion von nichtlinearen Verzerrungen, welche bei der Quantisierung von Signalen mit geringer Amplitude auftreten.

Wiederholerspektren werden entfernt/gedämpft. Töne werden dadurch besser/klarer wahrnehmbar.

3. Wie wird Dithering durchgeführt [1]

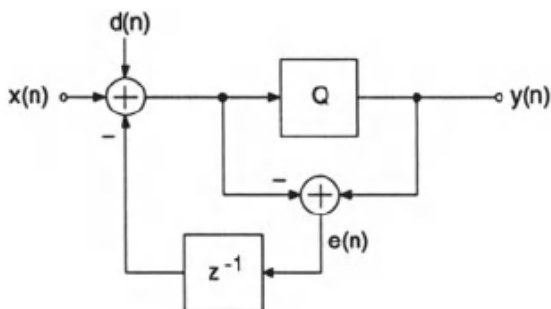
Es wird eine Linearisierung der nichtlinearen charakteristischen Kurve durchgeführt. Hierzu wird dem Signal ein Zufallssignal vor der Quantisierung aufaddiert. r-Bit Dither bedeutet, dass die letzten r-Bits eines Signals mit einer Zufallsfolge ersetzt werden, dies geschieht vor der Quantisierung.



4. Welches Dither-Signal unterdrückt die Rauschmodulation [1]

Die Rauschmodulation wird von einem Dreiecksförmigen Dither-Signal unterdrückt.

5. Blockdiagramm eines Quantisierers mit Noise-Shaping [3]



6. SNR eines überabtastenden ADC mit Überabtastungsfaktor L, Gaus-verteilt [1]

$\text{SNR} = 6.02 \cdot w + 10 \cdot \log(L)$ [dB] ... nicht gaußverteilt? rechteckig oder?

$\text{SNR} = 6.02 \cdot w - 8.5 + 10 \cdot \log(L)$...für gaußverteilte Audiosignale; Seite 23 (Udo Zölzer)

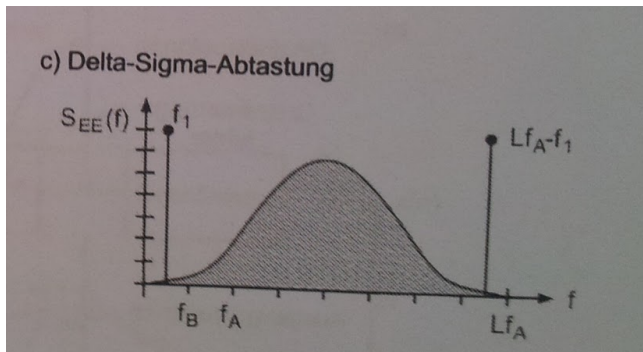
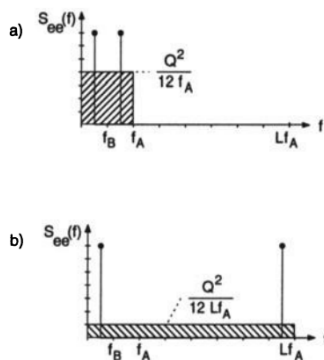
w ... Wortbreite des A/D Umsetzers

L ... Faktor der Überabtastung

Bei Sigma-Delta ADCs ist $w=1$

7. Skizze der Spektren des Quantisierungsrauschens nach Überabtastung mit $L=4$ und Noise-Shaping. Schraffieren Sie das verbleibende Quantisierungsrauschen im Nutzband [3]

Skizze Leistungsdichtespektrum des Quantisierungsfehler



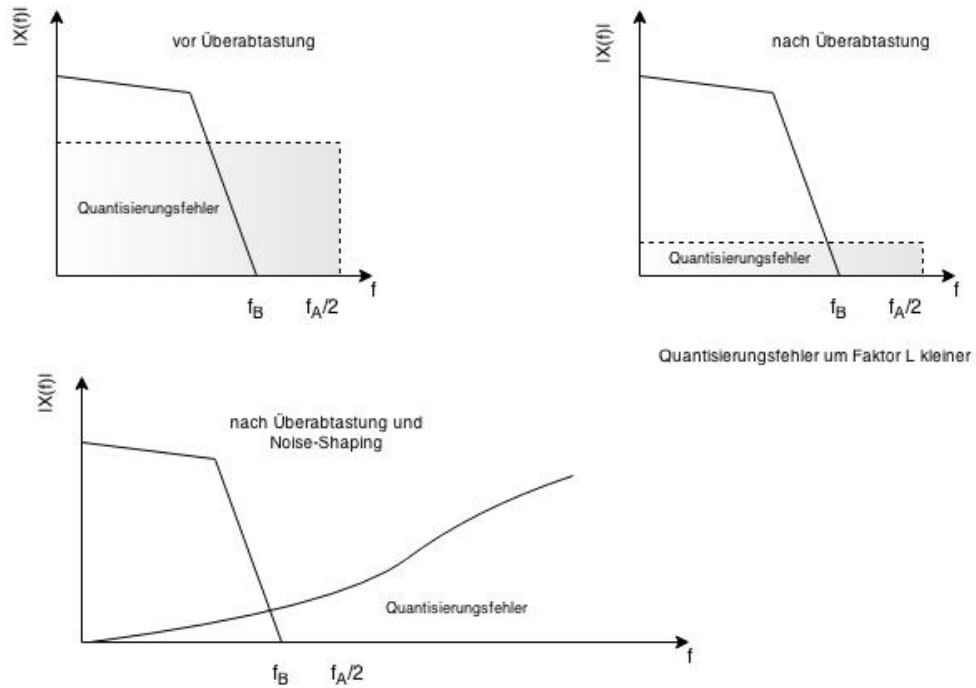
Aufgrund der spektralen Gleichverteilung der Fehlerleistung zwischen Frequenz 0 und f_A kann die spektrale Leistungsdichte um den Faktor L mit der neuen Abtastfrequenz $L \cdot f_A$ verringert werden. Graue Bereich bei Nyquist-Abtastung (a) und er graue Bereich bei der Überabtastung (b) ist gleich. Im Nutzband verringert sich jedoch die Fehlerleistung um den Faktor L (In diesem Fall 4).

- Leistung im Nutzband vor Überabtastung : $Q^2/12$ ($f_A = 2 \cdot f_B$)
- Leistung im Nutzband vor Überabtastung : $Q^2/12 \cdot 1/L$ ($f_A = 2 \cdot f_B$)

Ebenfalls ersichtlich ist die Erhöhung des SNR.

Bei der Delta-Sigma-Modulation wird der entstehende Quantisierungsfehler wie beim Noise-Shaping spektral geformt. Idealerweise wird die Rauschleistung von Frequenzbereichen hoher Gehörempfindlichkeit (wie z.B. 2-4 kHz) in Bereiche geringerer Empfindlichkeit verschoben (zumeist hohe Frequenzbereiche)

Spektren



8. Blockdiagramm eines Sigma-Delta-ADC mit nicht überabgetastetem Ausgang mit w-Bit. [5]

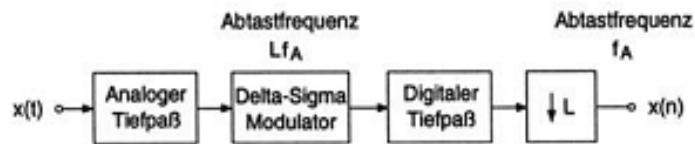
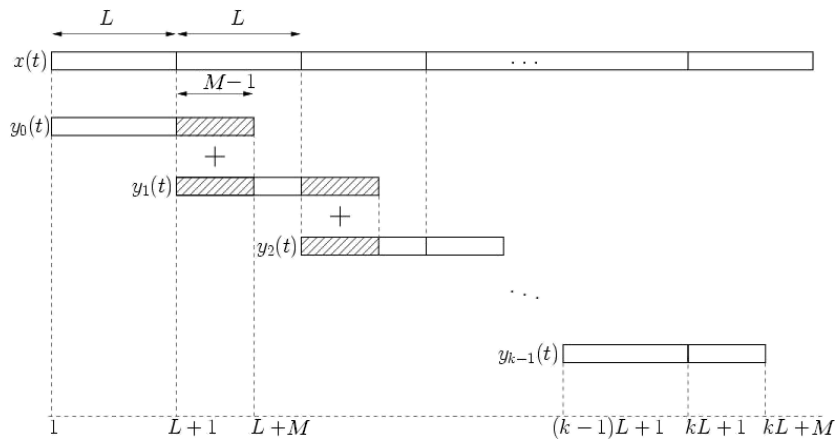


Bild 3.10: Überabtastender Delta-Sigma AD-Umsetzer

9. Schritte der schnellen Faltung [2]

Overlap-Add:

- FFT der Impulsantwort.
- FFT des nächsten Teiles der Sequenz.
- Multiplikation im Frequenzbereich anstatt Faltung im Zeitbereich
- I-FFT des Ergebnisses. Merken des überlappenden Bereichs.
- Hinzuaddieren des zuvor überlappenden Bereiches zum Ergebnis. Weiter bei Punkt 2.



10. Blockdiagramm einer Oktavfilterbank ohne Komplementärtechnik, welche das Eingangssignal in drei Frequenzbänder aufteilt und kein Aliasing auftritt. Machen Sie einen Vorschlag über die Grenzfrequenzen [5].

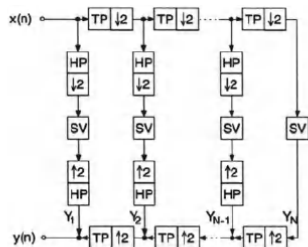


Bild 5.51: Oktav-Band QMF-Filterbank (SV=Signalverarbeitung, TP=Tiefpaß, HP=Hochpaß)

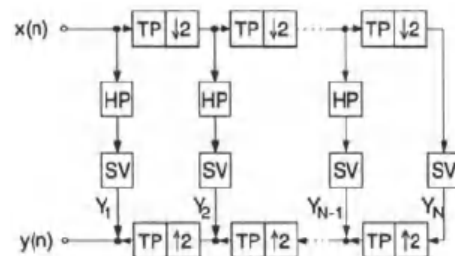


Bild 5.54: Modifizierte Oktav-Band Filterbank

Die Grenzfrequenzen sollten z.B. wie folgt gewählt werden um Aliasing zu verhindern:

$$\Omega_g = \frac{\pi}{3} * 2^{(-k+1)} \quad , \quad k = [1, 2]$$

11. Erläutern sie den Begriff der Frequenzmaskierung [1]

Ein Ton verhindert, dass umliegende Frequenzen vom Menschen gehört werden können. Dieser Effekt wird bei der Audio Kodierung genutzt.

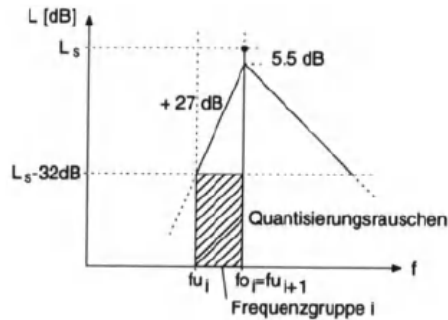


Bild 9.8: Maskierung innerhalb einer Frequenzgruppe

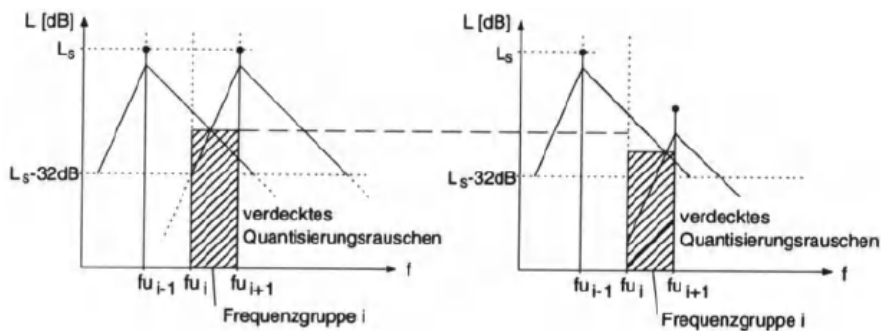


Bild 9.9: Bandübergreifende Maskierung

12. Nennen Sie mindestens zwei Vorteile der DCT hingegen der DFT in der Bildverarbeitung [2]

Ergebnis DCT: Rein reelles Signal, somit weniger Speicherbedarf. Lässt sich mit Hilfe der FFT implementieren.

13. Erklären Sie den Begriff der Separierbarkeit anhand der 2D-DCT [2]

Das Wort Separierbarkeit bezeichnet in der Bildverarbeitung die Eigenschaft, dass sich die Impulsantwort eines zweidimensionalen Filters durch die Multiplikation zweier eindimensionalen Operatoren darstellen lässt. Somit kann die zweidimensionale Faltung zu zwei eindimensionalen Operationen reduziert werden. In der Bildverarbeitung wird das ursprüngliche 2D-Filter in einen x- und y- Kern zerlegt, die dann hintereinander auf das Ursprungsbild angewandt werden. Das Ziel der Separierung ist eine Einsparung von Rechenzeit.

Separierung bei 2D-DCT

DCT:

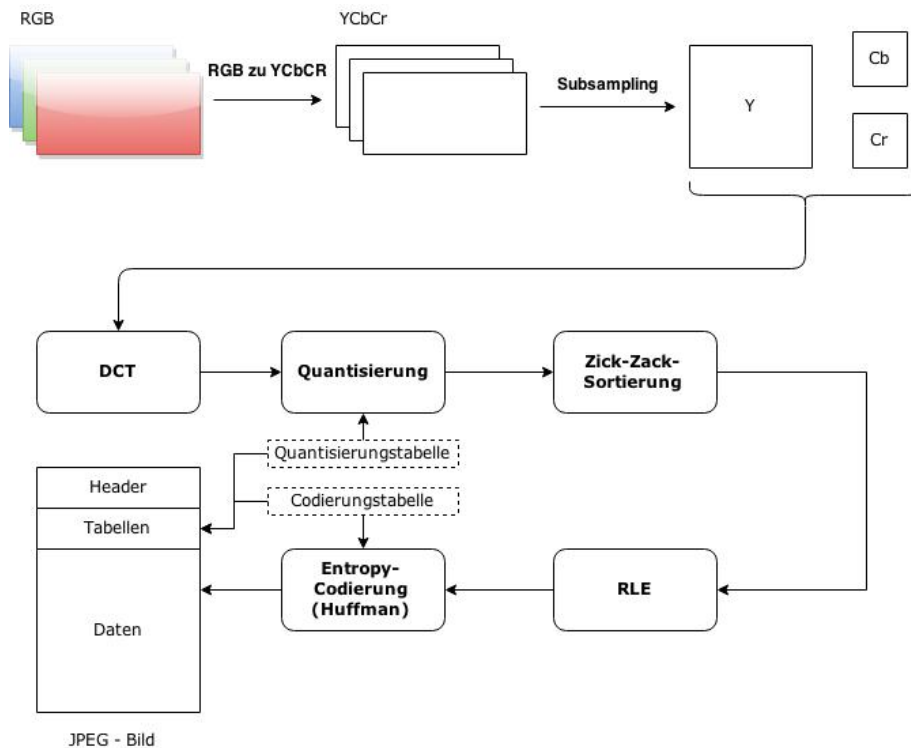
$$\begin{aligned}
 G(m, n) &= \frac{2}{\sqrt{MN}} \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} g(u, v) \cdot c_m \cos\left(\frac{\pi(2u+1)m}{2M}\right) \cdot c_n \cos\left(\frac{\pi(2v+1)n}{2N}\right) \\
 &= \frac{2c_m c_n}{\sqrt{MN}} \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} g(u, v) \cdot D_m^M(u) \cdot D_n^N(v) \quad (15.5)
 \end{aligned}$$

Separierung:

$$G(m, n) = \sqrt{\frac{2}{N}} \sum_{v=0}^{N-1} \underbrace{\left[\sqrt{\frac{2}{M}} \sum_{u=0}^{M-1} g(u, v) \cdot c_m D_m^M(u) \right]}_{\text{one-dimensional DCT}[g(\cdot, v)]} \cdot c_n D_n^N(v) \quad (15.7)$$

Die 1D-DCT kann zuerst auf jede Zeile des Bildes angewandt werden und anschließend auf jede Spalte.

14. Skizzieren sie den JPEG-Encoder [4]



15. Erläutern Sie den Begriff der Run-Length-Encoders anhand des JPEG-Encoders [2]

RLE ist ein verlustfreies Kompressionsverfahren.

RLE (dt. Lauflängenkodierung) nutzt lange Folgen sich wiederholender Zeichen oder Zeichenketten, sog. Läufe (runs), aus. Wird als Wertepaar aus Anzahl des Zeichen und Zeichen dargestellt (4A)

Beispiel: AAAABBBCCCCDEEEEEE RLE: 4A3B4C5D5E

Beispiel: 000111100000111111 RLE: (3,0)(4,1)(5,0)(6,1) oder 3 4 5 6 (falls erster Wert immer eine Null ist).

16. Zeichnen Sie das Blockdiagramm eines Videoencoders mit Motionestimation [4]

17. Erläutern Sie den Vorgang des Adaptive-Rood-Pattern [2]