



Kurzmanuskript zum Praktikumsversuch
Ton- und Bildbearbeitung (TBB)
Elektro- und Informationstechnikpraktikum

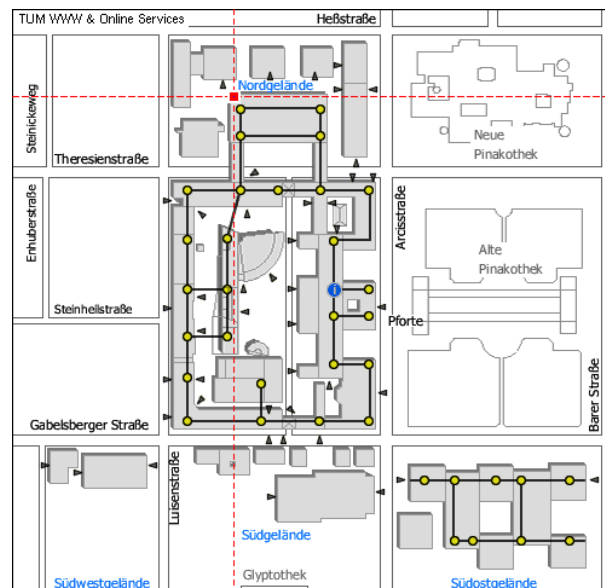
Autor: Dr.- Ing. Dejan Arsić

Betreuer: Patrick Lindemann, M.Sc.

Stand: 01.10.2018

Der Versuch findet im Erdgeschoss des Gebäudes N1 im Nordgelände, Raum N0122 statt.

Lageplan:



Anmerkungen:

- Die Fragen in den mit *** Aufgaben *** gekennzeichneten Kapiteln müssen als Vorbereitung bearbeitet werden. Die schriftlichen Lösungen dazu werden während des Praktikums verwendet und sind anschließend abzugeben.
- Die kursiven Absätze mit dem Titel Zusatzinfo sollen für interessierte Studierende Beispiele aus der Praxis mit der erlernten Theorie in Verbindung bringen. Zur Bearbeitung des Versuchs sind sie nicht dringend notwendig, können aber zum allgemeinen Verständnis beitragen.

Inhaltsverzeichnis

1	Allgemeines	3
2	Abtastung	3
2.1	Tonsignal	3
2.2	Bildsignal	4
2.3	***Aufgaben zu 2***	4
3	Aussteuerung	4
3.1	Quantisierungsrauschen	4
3.2	Unter-/Übersteuerung bei der Digitalisierung eines Tonsignals	4
3.3	Unter-/Übersteuerung bei der Digitalisierung eines Bildsignals	5
3.4	Die Einheit Dezibel	5
3.5	***Aufgaben zu 3***	5
3.6	Versuchsdurchführung	5
4	Aliasing	6
4.1	Abtasttheorem	6
4.2	***Aufgaben zu 4***	7
4.3	Versuchsdurchführung	7
5	Additive Synthese	7
5.1	Fourier-Reihe	8
5.2	***Aufgaben zu 5***	8
5.3	Versuchsdurchführung	8
6	Faltung	9
6.1	Eindimensionale, diskrete Faltung	9
6.2	***Aufgaben zu 6***	9
6.3	Versuchsdurchführung	10
7	Filter	10
7.1	Filtertypen	10
7.2	Zweidimensionale, diskrete Faltung	10
7.3	***Aufgaben zu 7***	10
7.4	Versuchsdurchführung	11

1 Allgemeines

Dieser Praktikumsversuch greift Inhalte der Vorlesung Signaldarstellung auf, und zeigt an Beispielen der Bearbeitung von Bild und Ton deren Anwendung. Dabei wird der Weg der digitalen Signalverarbeitung systematisch nachvollzogen: Ein analoges Signal wird abgetastet und wertdiskretisiert (quantisiert). Die bei diesem Vorgang entstehenden Probleme werden beleuchtet und Lösungsansätze aufgezeigt. Das digitale Signal wird dann mit Filtern bearbeitet. Die einzelnen Versuchsabschnitte erfolgen jeweils für Ton- und Bildsignale getrennt.

2 Abtastung

Da der Versuch am Computer durchgeführt wird, müssen alle Signale in digitaler Form vorliegen. Das heißt *wertdiskret* und *zeitdiskret* für Tonsignale bzw. *wertdiskret*, *ortsdiskret* und *zeitdiskret* für Bildsignale. Die zur Verfügung stehenden Analog/Digital-Wandler in Soundkarte und Webcam erlauben es Ihnen, eigene Dateien in die Versuche einzubinden.

2.1 Tonsignal

Was geschieht, wenn Sie ein akustisches Signal wie z.B. Ihre Stimme digitalisieren? Zunächst wandelt das Mikrofon die Schwingung der Luft in elektrische Größen um. Das Signal ist danach immer noch analog. Der sich auf der Soundkarte befindliche A/D-Wandler tastet dieses Signal mit der Abtastperiode T_s ab. Für jedes Abtastintervall wird ein *Schnappschuss* des momentanen Signalwertes gemacht (z.B. mit Hilfe Sample-and-hold-Schaltung). Dieser Wert wird dann quantisiert, d.h. die unendlich vielen Abstufungen des analogen Signals werden auf endlich viele diskrete Werte abgebildet. Diese Werte werden dann schließlich z.B. binär kodiert:

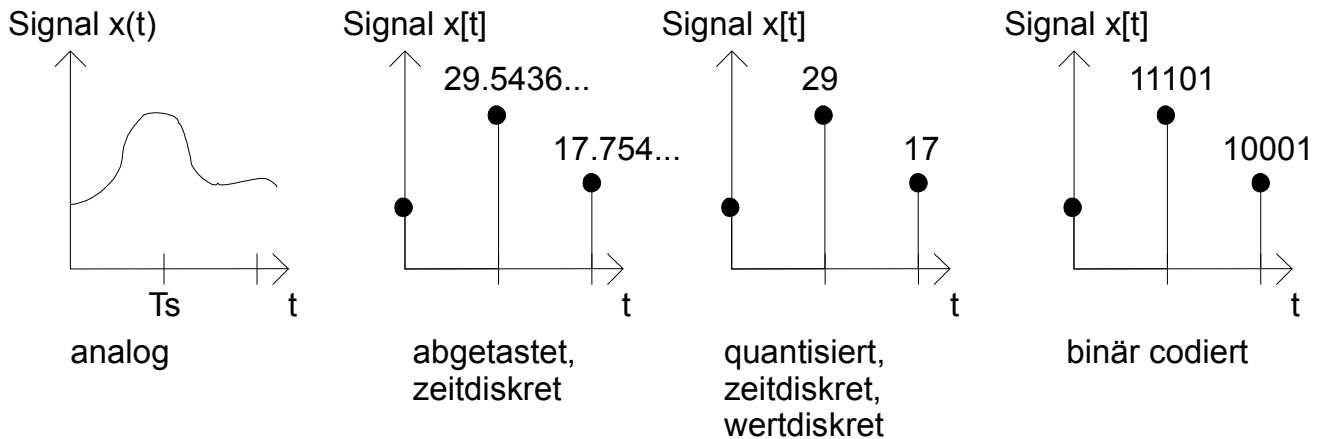


Abbildung 1:

Charakteristische Größen für die Beschreibung eines digitalen Tonsignals sind die Abtastfrequenz F_s ($= 1/T_s = \omega/2\pi$) und die Anzahl der Bit pro Abtastwert. Die Tondateien des Versuchs liegen alle in den für Audio-CDs üblichen Werten vor:

$$F_s = 44,1 \text{ kHz (D.h. das analoge Signal wird 44100 mal pro Sekunde abgetastet)}$$
$$16\text{bit / Abtastwert}$$

Die Amplituden können Werte von 0 bis $2^{\#bit} - 1$ annehmen. Dabei werden die Werte als vorzeichenbehaftete Binärzahlen interpretiert. Die Signale im Versuch werden auf den maximalen Pegel normiert dargestellt. Somit entspricht ± 1 dem maximalen Pegel und der Wert 0.0 Stille bei der Wiedergabe.

Zusatzinfo: Unkomprimierte Tondateien verfügen in der Regel über eine lineare Quantisierung. D.h. dass alle diskreten Werte den gleichen Abstand zueinander haben. Dies ist z.B. bei Dateien im MP3-Format nicht der Fall. Das Microsoft WAVE Format schreibt bei einer Quantisierung von 16bit einen Wertebereich von -32768 bis 32767 vor. Stereodateien benötigen den doppelten Datenspeicherplatz, da hier jeweils für 2 Kanäle separate Abtastwerte gespeichert werden müssen.

2.2 Bildsignal

Die Digitalisierung von zweidimensionalen, flächenhaften Mustern oder Bildsignalen ähnelt der von Zeit-/Tonsignalen. Der Chip der Webcam wandelt das einfallende Licht in elektrische Größen um. Dabei findet schon bei dieser Wandlung die Ortdiskretisierung, also Abtastung, statt. Man kann sich den Chip als Array aus vielen kleinen Kameras vorstellen. Jede dieser Einzelkameras bildet einen Bildpunkt/Pixel (von *picture cell*) ab. Diese einzelnen Bildpunkte werden dann quantisiert und binärkodiert. Je dichter diese Einzelkameras auf dem Chip gepackt sind, desto größer die Auflösung des resultierenden Bildes. Die verwendete Webcam hat eine Auflösung von 352 x 288 Pixel. Übliche Größen für die Beschreibung einer Bilddatei sind die Auflösung und die Anzahl der Bits pro Pixel. Die Bilddateien des Versuchs liegen alle als Grauwert-Bilder mit 8bit pro Pixel vor. Die Auflösung variiert. Die Pixel nehmen Werte zwischen 0 und 255 an. Dabei entspricht 0 dem dunkelsten Wert (schwarz) und 255 dem hellsten Wert (weiß). Die Bilder werden in die Versuchsumgebung durch Strecken/Stauchen eingepasst.

Zusatzinfo: Farbbilder benötigen anstelle eines einzelnen Wertes pro Pixel 3 Werte im RGB-System: Jeweils ein 8bit-Wert für die Farben Rot, Grün und Blau. Die darzustellende Farbe ergibt sich dann durch die Mischung der einzelnen Farbanteile. Die Farbtiefe gibt dabei an, wie viele Bit pro Farbwert verwendet werden. Echtfarben oder Truecolor entsprechen einem Wert von 32bit pro Pixel (RGBA).

2.3 ***Aufgaben zu 2***

Folgende digitalen Ton-/Bildsignale liegen vor:

- Tondatei der Länge 2.2s ($F_s = 44,1$ kHz, 16bit pro Abtastwert)
- Bilddatei der Webcam (Grauwert, 8bit)

Wie viele Bit belegen jeweils die Datensätze? Welcher Datensatz ist größer?

3 Aussteuerung

Die richtige Aussteuerung gewährleistet eine rauscharme, verzerrungsfreie Aufnahme bei Tonsignalen und eine kontrastreiche Abbildung bei Bildsignalen. Eine untersteuerte Aufnahme auf einem analogen Kassettenrekorder kann zwar durch Erhöhen der Abspiellautstärke am Verstärker laut abgespielt werden, allerdings wird dadurch auch das Bandrauschen angehoben und deutlich hörbar. Bei einer übersteuerten Aufnahme hingegen tritt eine Verzerrung auf. Bei einer Digitalisierung treten diese Effekte ebenfalls in ähnlicher Form auf.

3.1 Quantisierungsrauschen

Durch die Quantisierung der Werte bei der Digitalisierung geht Information verloren. Diese *zufällige* Veränderung des ursprünglichen Signals macht sich als Rauschen bemerkbar. Je feiner dabei die diskreten Werte abgestuft sind, desto weniger Rauschen entsteht. *Mehr Bit pro Abtastwert* ist also gleichbedeutend mit *weniger Quantisierungsrauschen*. Pro zusätzlichem Bit wird das SNR (*signal to noise ratio*) um 6dB verbessert.

3.2 Unter-/Übersteuerung bei der Digitalisierung eines Tonsignals

Der A/D-Wandler der Soundkarte weist einer anliegenden Spannung einen Wert zu. Dabei hat er einen begrenzten Wertebereich (vgl. Kap. 2.1) zur Verfügung. Diesen soll er möglichst vollständig ausnutzen, und zugleich nicht über die Grenzen hinaus *übersteuern*. Ein einfaches, konstruiertes Zahlenbeispiel:

Größtmöglicher Wert: 27=11011

Größter erreichter Wert: 13=01101

Das erste Bit (Most Significant Bit oder MSB) wird bei einer derartigen Aussteuerung nicht genutzt, es wird *verschwendet*. Durch nachträgliche Normalisierung, also Multiplikation des digitalisierten Signals mit dem Quotient $\frac{\text{grössterWert}}{\text{grösster,erreichterWert}}$, wird die größtmögliche Zahl zwar erreicht, allerdings wird dadurch auch das Quantisierungs-rauschen mitverstärkt. Bei einer übersteuerten Digitalisierung treten Spannungen auf, die größer sind, als die dem größtmöglichen Wert zugehörige. Diese

Spannungen werden ebenfalls auf den größten Wert abgebildet. Die Spitzen des Signals werden also gekappt (*clipping*). Das resultierende Signal klingt stark verzerrt. Dieser Effekt tritt im Gegensatz zur analogen Übersteuerung viel abrupter ein und wird in der Regel daher als unangenehmer empfunden.

Zusatzinfo: In der Tontechnik wird eine gewisse Toleranz vom Signalmaximum zum maximal möglichen Wert eingehalten. Dadurch wird sichergestellt, dass plötzliche Überhöhungen (z.B. durch einen übermotivierten Drummer) trotzdem noch verzerrungsfrei erfasst werden können. Diese Toleranz nennt man auch Headroom.

3.3 Unter-/Übersteuerung bei der Digitalisierung eines Bildsignals

Bei einer untersteuerten Digitalisierung wird das Bild zu dunkel. Durch Normalisierung kann eine Kontrastverbesserung erreicht werden. Allerdings entarten dann weiche Helligkeitsverläufe zu harten Stufen. Bei einer Übersteuerung wird das Bild zu hell. Eine Kontrastverbesserung durch die oben beschriebene Normalisierung ist nicht möglich.

3.4 Die Einheit Dezibel

Dezibel [dB] ist ein logarithmisches Relativmaß. Es wird häufig in der Akustik und Tontechnik verwendet. In der Versuchsumgebung dient es dazu, das Verhältnis zwischen der größtmöglichen Amplitude (x_o) und der tatsächlichen Signalamplitude ($x[n]$) eines Tonsignals darzustellen [1]:

$$L_{dB} = 20 \log \frac{|x[n]|}{|x_o|}$$

Sollen Intensitäten/Leistungen verglichen werden, gilt folgende Formel:

$$L_{dB} = 10 \log \frac{I[n]}{I_o} = 10 \log \frac{x[n]^2}{x_o^2}$$

3.5 ***Aufgaben zu 3***

Berechnen Sie die fehlenden Größen ($x_o = 1.0$):

L_{dB}	Amplitude $x[n]$
0	
-6	0.25
-20	
$-\infty$ (Grenzwert)	0.0

3.6 Versuchsdurchführung

- Starten Sie den Versuch *Aussteuerung (TON)*.
 - (a) Laden Sie die Datei *Sinus.wav*. Durch Auswahl eines hohen Reskalierungsfaktors und Drücken des Reskalierungsknopfes können Sie diese Datei nachträglich stark übersteuern. Wie hat sich die Sinusschwingung verändert (Zeichnung mit Achsenbeschriftung)? Sie möchten nun die maximale Amplitude um 6dB verringern. Welchen Reskalierungsfaktor müssen Sie verwenden? Gelingt Ihnen durch nochmalige Reskalierung die Rückgewinnung der ursprünglichen Sinusschwingung? Begründung!
 - (b) Laden Sie erneut die Datei *Sinus.wav*. Wählen Sie den Skalierungsfaktor 0.00025. Nach Ausführen der Reskalierung haben Sie die Datei nachträglich stark untersteuert. Normalisieren Sie jetzt und zoomen Sie sich in den Bereich einzelner Schwingungen. Wie hat sich die Sinusschwingung verändert (Zeichnung mit Achsenbeschriftung)? Wie viele Quantisierungsstufen können Sie erkennen? Welche Wortlänge würde also für einen Abtastwert anstelle der ursprünglichen 16bit benötigt?
 - (c) Laden Sie eine Musikdatei (*Musik1.wav* oder *Musik2.wav*). Ist diese Datei optimal ausgesteuert? Falls nicht, normalisieren Sie sie. Senken Sie jetzt die Amplitude um 3dB. Welchen Skalierungsfaktor müssen Sie verwenden? Hören Sie sich Original und Bearbeitung an. Machen sich 3dB Pegelunterschied überhaupt bemerkbar? Laden Sie die Musik erneut, skalieren (Faktor 0.00025) und normalisieren Sie. Deutlich ist das Quantisierungsrauschen wahrzunehmen!

- (d) Laden Sie mit Hilfe des Recorder&Grabber-Tools eine Sprachaufnahme Ihrer Stimme (*Stimme.wav*). Skalieren Sie mit dem Faktor 0.00025 und normalisieren Sie dann. Wieviele Quantisierungsstufen können Sie erkennen? Sind die Worte noch zu verstehen?
- Starten Sie den Versuch *Aussteuerung (BILD)*.
- (e) Laden Sie ein Bild der Webcam (*grabbedImage.tiff*). Reskalieren Sie mit einem hohen Faktor. Wie verändert sich das Bild? Normalisieren Sie. Wie verändert sich das Bild jetzt? Begründung!
- (f) Laden Sie die Datei erneut und skalieren Sie mit Faktor 0.025. Nach Ausführen der Reskalierung haben Sie die Datei nachträglich stark untersteuert. Normalisieren Sie. Wieviele Grauwerte (=Quantisierungsstufen) können Sie erkennen? Welche Wortlänge würde also für einen Bildpunkt anstelle der ursprünglichen 8bit benötigt?
- (g) Wie erklären Sie sich, dass bei Versuch f) aus einem schwarz erscheinendem Bild durch Normalisierung das ursprüngliche Bild annähernd rekonstruiert werden kann?

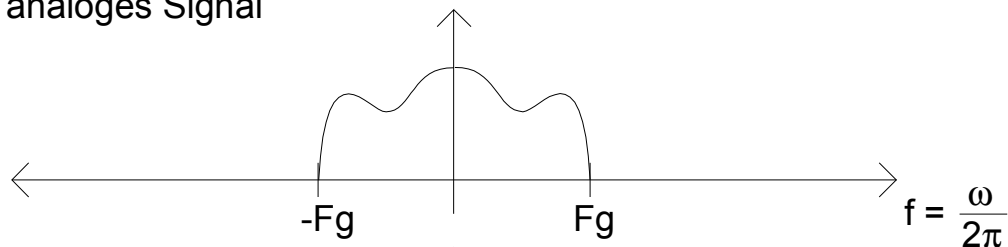
4 Aliasing

Aliasing ist ein zeitbasierter Effekt, der bei der Abtastung auftreten kann und zu Verfälschungen führt.

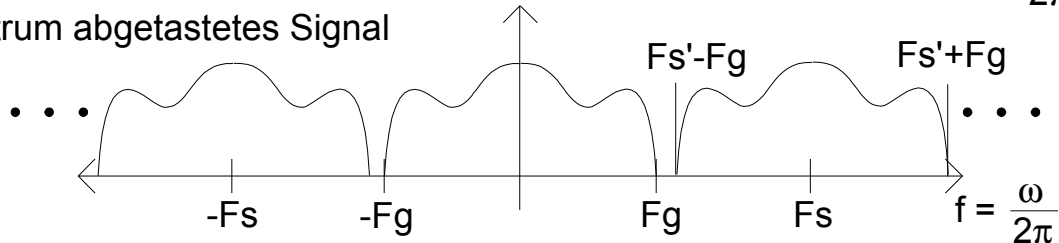
4.1 Abtasttheorem

Das Abtasttheorem besagt, dass ein bandbegrenzttes Signal aus den Abtastwerten ideal rekonstruiert werden kann, wenn die Abtastrate (bzw. Samplingfrequenz) F_s mindestens doppelt so groß ist, wie die Grenzfrequenz F_g des Signals. Angewandt auf die Abtastung eines Tonsignals bedeutet dies: Das menschliche Ohr kann Schalle im Frequenzbereich von 16Hz bis 20kHz hören. Spektralanteile über 20kHz z.B. von Musik müssen also nicht berücksichtigt werden ($\Rightarrow F_g = 20kHz$). F_s muss also größer als 40kHz sein. Durchgesetzt hat sich die Abtastrate von 44.1kHz. Die Abtastung im Zeitbereich entspricht einer periodischen Fortsetzung im Frequenzbereich: Bei einer Überabtastung liegen die periodisierten Spektralanteile weiter

Spektrum analoges Signal



Spektrum abgetastetes Signal



Spektrum bei Unterabtastung

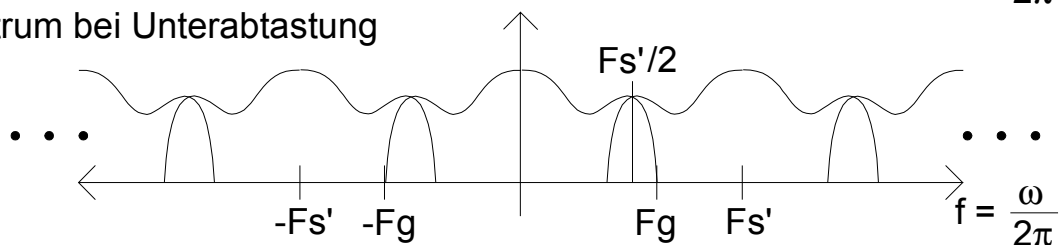


Abbildung 2:

auseinander. Bei einer Unterabtastung überlappen sie sich. Wenn diese Überlappungsfehler auftreten, spricht man von *Aliasing*. Dies kann vermieden werden, wenn das abzutastende Signal zuvor einer Tiefpassfilterung unterzogen wird, die alle Spektralanteile über der halben Abtastfrequenz $F_s'/2$ abschneidet.

4.2 ***Aufgaben zu 4***

Angenommen Sie möchten das analoge Signal aus 2 abtasten, können jedoch nicht die optimale Abtastrate verwenden, sondern nur die Abtastrate F_s' aus der unteren der drei Darstellungen. Um Aliasing zu vermeiden, bearbeiten Sie das analoge Signal mit einem idealen Tiefpassfilter mit der Grenzfrequenz $F_s'/2$. Skizzieren Sie das sich nach der Abtastung ergebende periodische Spektrum.

4.3 Versuchsdurchführung

- Starten Sie den Versuch *Aliasing (TON)*

- (a) Die Grafik zeigt Ihnen das Spektrum eines Sinustons der Frequenz 5kHz, der mit der Abtastrate $F_s = 44100\text{Hz}$ abgetastet wurde. Hören Sie sich diesen Sinuston an. Die periodischen Spektralanteile werden nicht gezeigt (nur Frequenzen $< F_g = F_s/2$).

Hören Sie sich im Vergleich dazu den gleichen Ton bei einer Abtastfrequenz von 22050 Hz und 8000Hz ohne Tiefpassfilterung an.

Sind Unterschiede zu hören oder im Spektrum festzustellen? Wenn ja, warum? Tritt Aliasing auf?

Hören Sie sich nun den Sinuston bei einer Abtastrate von 22050 Hz und 8000Hz mit Tiefpassfilterung an. Sind Unterschiede zu hören oder im Spektrum festzustellen? Wenn ja, warum? Tritt Aliasing auf?

- (b) Wählen Sie nun das Musiksignal zur Analyse aus. Bei welchen Einstellungen tritt Aliasing auf? Ist es eher vorteilhaft oder störend für den Musikgenuss? Wie wirkt sich eine niedrige Abtastrate auf das Klangbild allgemein aus? Würde sich eine Abtastrate von 48kHz positiv auf das Klangbild auswirken?

- Starten Sie den Versuch *Aliasing (BILD)*.

- (c) Setzen Sie die Schwingungszahl auf 1. Es wird eine Sinusschwingung im Ortsbereich dargestellt. Helle Pixel entsprechen positiven Werten, dunkle negativen Werten. Wählen Sie nun die Rechteckschwingung aus. Sie sehen jetzt 1 Schwingung auf 300 Pixel, also 150 schwarze und 150 weiße Pixel. Bei Mausklick auf den rechten Pfeil des Scrollbalkens erhöhen Sie die Anzahl der Schwingungen jeweils um 1, bei Mausklick neben den Scrollbalken um 10. (Direkte Werteingabe ebenfalls möglich!)

Wieviele gleichfarbige Pixel gehören zu einer Halbschwingung bei folgenden Schwingungszahlen:

5, 50, 75, 80, 150, 180, 300?

Lassen sich alle Werte korrekt grafisch darstellen? Begründung!

Welches ist die obere Grenzfrequenz F_g ? Stellen Sie diese nun ein und erhöhen Sie in Zehnerschritten die Frequenz. Was beobachten Sie? (Zoomen!)

- (d) Wählen Sie nun den Versuchsteil *Anti-Aliasing im Bild* aus. Sie sehen viermal eine Darstellung des Buchstaben *K*. Links oben ist das hochauflösende Original zu sehen. In der Abbildung unten links entstand durch Herabsetzung der Auflösung Aliasing.

Diese Treppenbildung kann durch eine Tiefpassfilterung des Originals (oben rechts) vor der Verminderung der Auflösung verringert werden. Das Bild wird dadurch unschärfer (unten rechts), aber die Treppenstufen stechen weniger ins Auge. Erhöhen Sie Ihren Abstand vom Monitor auf ca. 1,5m und vergleichen Sie die beiden unteren Buchstaben!

5 Additive Synthese

Die additive Synthese ist eine Anwendung der Fourier-Reihe. Einfache Sinus- Oszillatoren können benutzt werden, um komplexere Schwingungen zu synthetisieren.

5.1 Fourier-Reihe

Periodische Signale $x(t)$, die beschränkt und im Intervall T_0 bis auf endlich viele Sprungstellen stetig sind, lassen sich in eine Fourier-Reihe entwickeln. Für reelle periodische Signale kann die Fourier-Reihe folgendermaßen geschrieben werden [1]:

$$x(t) = a_0 + \sum_{k=1}^{\infty} a_k \cos(k\omega_0 t); \omega_0 = \frac{2\pi}{T_0}; \omega_0 = 2\pi f_0$$
$$a_0 = \frac{1}{T_0} \int_{T_0} x(t) dt; a_k = \frac{2}{T_0} \int_{T_0} x(t) \cos(k\omega_0 t) dt; b_k = \frac{2}{T_0} \int_{T_0} x(t) \sin(k\omega_0 t) dt$$

Die Fourier-Reihe ist eine Spektraldarstellung des Signals $x(t)$ bestehend aus einem Gleichanteil, der Grundfrequenz ω_0 und den höheren Harmonischen $k\omega_0$. Die Koeffizienten a_k, b_k sind ein Maß für den Anteil, mit dem die Harmonischen zum Gesamtsignal $x(t)$ beitragen. Die zu synthetisierenden Signale Rechteck und Sägezahn sind beide punktsymmetrisch. Damit fällt der gerade Anteil weg: $a_0 = a_k = 0$. Die Fourier-Reihen lauten nun [2]:

$$\text{Rechteck} : \sum_{k=0}^{\infty} \frac{1}{2k+1} \sin((2k+1)2\pi f_0 t); \text{Sägezahn} : \sum_{k=1}^{\infty} \frac{1}{k} \sin(k2\pi f_0 t);$$

Für jeden dieser Summanden stellt der Versuch einen Sinusoszillator mit der $k2\pi f_0 t$ entsprechenden Schwingungsfrequenz zur Verfügung. f_0 entspricht der Grundfrequenz. Im Audioteil des Versuches wird diese Grundfrequenz 200Hz sein. Ihre Aufgabe wird es sein, die Schwingungsamplituden b_k vor dem Sinus-Term einzustellen.

Zusatzinfo: Additive Synthese findet z.B. in Musiksynthesizern Anwendung. Die meisten Analogsynthesizer bieten nur wenige, starre Grundschwingungen. Meist sind dies: Sinus, Rechteck, Sägezahn, Dreieck. Durch die add. Synthese gelingt es, diese Schwingungen nachzubilden und zeitlich zu verändern um einen lebendigeren Klangverlauf zu erreichen. So stellt z.B. der Kawai K5000S Sinus-Oszillatoren für die ersten 128(!) Harmonischen zur Verfügung.

5.2 ***Aufgaben zu 5***

Schreiben Sie die Fourier-Reihen von Rechteck für die ersten 6 Harmonischen und Sägezahn für die ersten 12 Harmonischen. Die Faktoren $1/k$ sind als Dezimalzahlen auf 2 Nachkommastellen zu runden. Ausgehend von der Grundfrequenz 200 Hz: Welche Frequenzen müssen die ersten 6 Sinus-Oszillatoren für die Synthese der Rechteckschwingung haben?

5.3 Versuchsdurchführung

- Starten Sie den Versuch *Additive Synthese (TON)*.
Oben links können Sie Referenzwellenformen auswählen, deren Wellenform und Spektrum angezeigt werden. Durch Bewegen eines Sliders oder direkte Werteingabe unten rechts schaltet die Darstellung auf Wellenform und Spektrum der momentan synthetisierten Schwingung um. Zur Einstellung der Faktoren b_k reichen zwei Nachkommastellen.
 - (a) Machen Sie sich durch Verschieben einzelner Slider die Wirkungsweise der einzelnen Sinus-Oszillatoren klar.
 - (b) Aktivieren Sie nun die Referenzwellenform Rechteck und versuchen Sie diese zu synthetisieren. Fangen Sie dabei zweckmäßigerweise mit der tiefsten Grundfrequenz an.
Wie verändern sich Spektrum und Wellenform nach Hinzufügen einer weiteren Harmonischen? Spielen Sie nach jeder hinzugefügten Harmonischen die neue Wellenform ab. Skizzieren Sie die fertige Wellenform (Achsenbeschriftung)!
 - (c) Setzen Sie die Slider alle auf 0 zurück, aktivieren Sie die Referenzwellenform Sägezahn und versuchen Sie diese zu synthetisieren. Fangen Sie wieder mit der tiefsten Grundfrequenz an. Wie verändern sich Spektrum und Wellenform nach Hinzufügen einer weiteren Harmonischen? Skizzieren Sie die fertige Wellenform (Achsenbeschriftung)!
 - (d) Da nur die ersten 12 Harmonischen verwendet werden, fehlen dem synthetisierten Signal die hohen Frequenzen. Bis zu welcher Harmonischen müsste man gehen, um die Sägezahnschwingung fürs Ohr optimal nachzubilden? Bis zu welcher Harmonischen darf man gehen, wenn man bei einer Abtastung mit 44100Hz und ohne Tiefpassfilter Aliasing vermeiden möchte?

- Starten Sie den Versuch *Additive Synthese (BILD)*.

Die Grundfrequenz ist in diesem Versuchsteil so gewählt, dass genau vier Grund-schwingungen auf der Breite der Grafik dargestellt werden können.

- (e) Synthetisieren Sie zuerst eine Rechteck- und dann eine Sägezahn-schwingung.

6 Faltung

Die Multiplikation zweier Signale im Zeitbereich entspricht einer Faltung im Frequenzbereich und umgekehrt. Wenn ein digitales Signal *gefiltert* werden soll, so geschieht dies durch eine Faltung des Signals mit der Impulsantwort des Filters im Zeitbereich.

6.1 Eindimensionale, diskrete Faltung

Die zeitdiskrete Faltung ist folgendermaßen definiert [1]:

$$y[n] = \sum_{l=-\infty}^{\infty} x[l]h[n-l] = \sum_{l=-\infty}^{\infty} h[l]x[n-l] = x[n] * h[n]$$

$h[n]$ ist die Einheitsimpulsantwort eines zeitdiskreten LTI-Systems (Filter), $x[n]$ das zu bearbeitende Signal. Statt der Integralbildung bei der zeitkontinuierlichen Faltung werden hier nur Addition und Multiplikation benötigt. Man spiegelt die Impulsfolge des Filters an der y-Achse und *schiebt* diese dann von minus unendlich bis plus unendlich über das Signal. Dabei multipliziert man die entsprechenden Werte von x und h (gleiche n -Position), addiert diese Einzelwerte und trägt sie an der Stelle $y[n]$ ein:

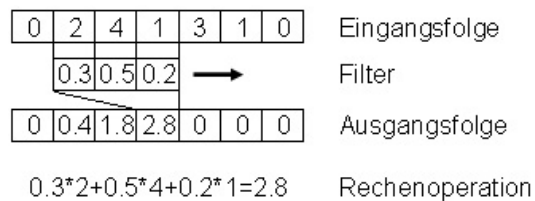


Abbildung 3:

6.2 ***Aufgaben zu 6***

- Was geschieht, wenn ein diskretes Signal mit einem Einheitsimpuls gefaltet wird?
- Wie verhält sich die Sprungantwort zur Impulsantwort?
- Welche Signalform entsteht, wenn zwei gleiche Rechtecke

$$\begin{aligned}
 x[n] = h[n] &= 0 && \text{für } n < 0, \quad n \in \mathbb{Z} \\
 x[n] = h[n] &= 1 && \text{für } 0 \leq n \leq 2 \\
 x[n] = h[n] &= 0 && \text{für } n > 2
 \end{aligned}$$

miteinander gefaltet werden? (Skizze & Rechnung!)

6.3 Versuchsdurchführung

- Starten Sie den Versuch *Faltung*. Der Versuch ermöglicht es Ihnen, eigene Signalfolgen einzugeben, die Faltung schrittweise durchzuführen und das Ergebnis abzulesen.
 - (a) Sie sehen eine zeitbeschränkte Signalfolge (blaue Punkte) und eine kurze Impulsantwort eines LTI-Systems (rote Quadrate). Durch Drücken des *Vorwärts*-Knopfes wird die Faltung schrittweise ausgeführt. Achten Sie darauf, wie die Filterfolge gespiegelt wird! Ab welchen n -Positionen hat die Ausgangsfolge $y[n]$ von 0.0 verschiedene Werte? Ist die Ausgangs- oder die Eingangsfolge länger?
 - (b) Löschen Sie nun alle Werte und überprüfen Sie Ihre Lösungen der Hausaufgaben a) und c) anhand selbstgewählter Einstellungen.
 - (c) Löschen Sie nun alle Werte. Gegeben ist das Filter $[0.333 \ 0.333 \ 0.333]$. Überprüfen Sie an einer selbstgewählten, sinnvollen Eingangsfolge, ob es sich eher um ein Hochpass- oder Tiefpassfilter handelt.
 - (d) Ersetzen Sie bei gleicher Eingangsfolge die Impulsantwort durch folgendes Filter: $[0.75 \ -1.5 \ 0.75]$. Was beobachten Sie?

7 Filter

Beim Durchgang durch ein LTI-System werden die komplexen Amplituden $X(\omega)$ des Eingangssignals $x(t)$ mit dem Frequenzgang $H(\omega)$ moduliert:

$$y(t) = h(t) * x(t); \quad Y(\omega) = H(\omega)X(\omega)$$

Der Frequenzgang (=Übertragungsfunktion) lässt erkennen, in welchen Frequenzbereichen das Signal durchgelassen (=Durchlassbereich) bzw. gesperrt (=Sperrbereich) wird.

7.1 Filtertypen

Ein ideales Filter ist gekennzeichnet durch einen scharfen Übergang zwischen Durchlass- und Sperrbereich (=unendliche Flankensteilheit). Reale Filter weisen abhängig vom Design dagegen eine Welligkeit im Frequenzgang und/oder flachere Übergänge zwischen den Bereichen auf. Die Grenzfrequenz ist die Grenze zwischen Durchlass- und Sperrbereich. Die spektralen Anteile im Sperrbereich werden beim realen Filter nicht vollständig ausgelöscht, sondern nur sehr stark gedämpft. Die häufigsten Filtercharakteristiken sind:

- Tiefpassfilter (TP, engl. lowpass): Lässt Frequenzen unterhalb der Grenzfrequenz durch. Höhere Frequenzen werden gedämpft.
- Hochpassfilter (HP, engl. highpass): Lässt Frequenzen oberhalb der Grenzfrequenz durch. Tiefere Frequenzen werden gedämpft.
- Bandpassfilter (BP, engl. bandpass): Kombination aus TP und HP. Grenzfrequenz des TP liegt über der des HP.
- Bandsperre (BS, engl. notch): Kombination aus TP und HP. Grenzfrequenz des TP liegt unter der des HP.

7.2 Zweidimensionale, diskrete Faltung

Die Impulsantwort des Filters ist bei der zweidimensionalen Faltung nicht mehr ein Vektor, sondern eine quadratische Matrix. Diese *schiebt* man pixelweise über das Bild und führt die schon von der eindimensionalen Faltung bekannten Arbeitsschritte durch: Multiplikation der Einzelwerte von Eingangs- und Filterfolge und anschließende Addition.

7.3 ***Aufgaben zu 7***

- (a) Skizzieren Sie schematisch die Frequenzgänge idealer BP, BS und markieren Sie Sperr- und Durchlassbereich.
- (b) Welche Veränderungen gegenüber dem Original erwarten Sie von einem TP bzw. HP gefilterten Musik- bzw. Bildsignal?

7.4 Versuchsdurchführung

- Starten Sie den Versuch *Filter (TON)*.

- (a) Wählen Sie als Tonsignal die Rechteckschwingung (Grundfrequenz 200Hz) aus. Es wird die Schwingung und das Spektrum angezeigt. Wählen Sie den TP aus. Durch Verändern der Grenzfrequenz können Sie jetzt Schwingungsform und Spektrum beeinflussen.

Welcher Schwingung ähnelt die Rechteckschwingung bei der kleinsten möglichen Grenzfrequenz? Warum? Achten Sie darauf, wie das Filter die höheren Harmonischen zwar dämpft, aber nicht ganz auslöscht!

Ab welcher Grenzfrequenz hören Sie keinen Unterschied mehr zwischen dem ursprünglichen und dem bearbeiteten Signal? Welcher Harmonischen entspricht das?

- (b) Wählen Sie nun den HP aus. Gelingt es Ihnen, die 1. Harmonische (=Grundfrequenz) auszulöschen? Ab welcher Grenzfrequenz ist das Signal nicht mehr zu hören? Stimmt diese mit der von Ihnen festgestellten Grenzfrequenz bei 7.4 a) überein? Verfügen die verwendeten Filter also eher über eine hohe oder geringe Flankensteilheit?

- (c) Wählen Sie nun als Tonsignal das Rauschen aus. Durch Einsatz eines BP können Sie ein bandbegrenztes Rauschen erzeugen. Hören Sie sich dieses bei verschiedenen Grenzfrequenzen an.

Zusatzinfo: Das Spektrum des Rauschsignals sollte theoretisch ein horizontaler Strich sein, da bei einem echten weißen Rauschen alle Frequenzen mit der gleichen Häufigkeit und Intensität auftreten. Da das Spektrum aber nur über einen kurzen Zeitraum analysiert wird und das Rauschen künstlich erzeugt wurde, treten deutliche Schwankungen auf.

BP und BS verfügen über eine obere UND untere Grenzfrequenz. Diese liegen symmetrisch über und unter der jeweils eingestellten Mittenfrequenz und ergeben die Bandbreite des Durchlass- bzw. Sperrbereichs.

- (d) Wählen Sie nun ein Musiksignal Ihrer Wahl aus. Wie macht sich eine TP-/HP-/BP-/BS-Filterung bei einer Grenzfrequenz von um die 3000Hz bemerkbar?

- (e) Wählen Sie Ihre Stimme als Signal aus. Mit welchen Filtereinstellungen/-typen können Sie die Stimme maximal auslöschen?

Welche Einstellungen eignen sich, um das Spektrum der Stimme zu begrenzen, aber trotzdem eine gute Sprachverständlichkeit zu erhalten?

Stichwort Mobilfunk: Schmales Spektrum = geringe Bandbreite = wenige Daten. Stimme im Bereich von 80Hz - 3500kHz, im restlichen Frequenzbereich könnten Daten übertragen werden

- (f) Wählen Sie nun bei Filtertyp *Hallfahne* und bei Tonsignal *Trommelschlag* aus. Hier wird jetzt anstelle eines computerberechneten Filters eine Impulsantwort eines realen Raumes (z.B. Turnhalle) zur Faltung verwendet. Die Faltung wird beim Drücken des Knopfes *PLAY (nach Filter)* ausgeführt. Beschreiben Sie das Resultat!

Zusatzinfo: Wenn das Resultat über die Lautsprecher abgespielt wird, klingt es sehr hallig. Da neben der internen Hallfahne der Turnhalle auch noch die Impulsantwort des TBB-Labors zwischen Lautsprecher und Ohr das Signal moduliert. Um diese zweite Raumimpulsantwort zu eliminieren, kann man Kopfhörer verwenden. So sollte man bei einer guten Hallfahne den Eindruck erhalten, in jener Turnhalle zu sitzen, anstatt im Labor.

- Starten Sie den Versuch *Filter (BILD)*.

- (g) Wählen Sie die Datei *Lena.bmp* und notieren Sie, wie sich eine TP-/HP-Filterung bei einer niedrigen und hohen Grenzfrequenz auswirkt!

Durch eine Bandsperrenfilterung mittlerer Frequenz (ca. 20) können Sie sowohl hohe als auch tiefe Frequenzen im Bild sichtbar machen.

Wie machen sich hohe bzw. tiefe Frequenzen in einem Bild bemerkbar?

- (h) Laden Sie die Bilddatei der Webcam (*grabbedImage.tiff*). Durch Wahl des Filtertyps *Benutzerentwurf* wird das Bild mit der angegebenen Matrix bearbeitet. Welcher Filtertyp ist voreingestellt?

Sie können eigene zweidimensionale Impulsantworten in die Matrix eingeben. Nach jeder Werteingabe wird das Bild automatisch gefiltert.

Durch welche Werte kann das Bild heller bzw. dunkler gemacht werden?

Wie wirken sich folgende Filter aus:

(i)

-1	0	1
-1	0	1
-1	0	1

(j)

1	1	1
0	0	0
-1	-1	-1

(k)

-0.3	-0.3	-0.3
-0.3	3.4	-0.3
-0.3	-0.3	-0.3

Literatur

- [1] Lang: Kurzmanuskript zur Vorlesung Signaldarstellung, 2. Auflage, TUM-Lehrstuhl f. Mensch-Maschine-Kommunikation, München 1998.
- [2] Bronstein, Semendjajew: Taschenbuch der Mathematik, 20. Auflage, Verlag Harri Deutsch, Thun und Frankfurt am Main 1981.
- [3] Lang: Kurzmanuskript zur Vorlesung Mensch-Maschine-Kommunikation 1, 3. Auflage, TUM-Lehrstuhl f. Mensch-Maschine-Kommunikation, München 1999.

Audio Dictionary

Deutsch	English
basslastig	bassy, boomy
blechern	tinny, brassy
dröhnend	boomy
dumpf	muffled, dead
dünn	thin
klar, luftig	clear, airy, crisp
kräftig	strong, powerful
nasal	nasal (pronounced with the voice issuing through the nose)
scharf, aufdringlich	sharp, strident, twangy, harsh
schrill, aggressiv	sharp, piercing, aggressive, edgy
schwammig	spongy, woolly
spitz	topy, trebly
surrend	buzzing, zooming
telefon/megafon	tinny, telephone-like
undefiniert	undefined
verrauscht	noisy
verwaschen	washed out, blurry, wet
voll	full
zischend	hissy
zweistimmig	duophonic, two-part (2 voices)