



UNIVERSITÄT
KOBLENZ · LANDAU

Fachbereich 4: Informatik



Vergleich gebräuchlicher linearer Bildtransformationen

Bachelorarbeit
zur Erlangung des Grades
BACHELOR OF SCIENCE
im Studiengang Informatik

vorgelegt von

André-Gilbert Thomas

Betreuer: Prof. Dr.-Ing. Dietrich Paulus, Institut für Computervisualistik,
Fachbereich Informatik, Universität Koblenz-Landau

Erstgutachter: Prof. Dr.-Ing. Dietrich Paulus, Institut für Computervisualistik,
Fachbereich Informatik, Universität Koblenz-Landau

Zweitgutachter: Dipl.-Inform. Detlev Droege, Institut für Computervisualistik,
Fachbereich Informatik, Universität Koblenz-Landau

Koblenz, im März 2018

Kurzfassung

Das Ziel dieser Bachelorarbeit ist es, die diskrete Fouriertransformation, die diskrete Kosinustransformation und die Hadamard-Walsh-Transformation im Kontext der Bildverarbeitung zu vermitteln und diese unter ausgewählten Gesichtspunkten zu vergleichen. Hierfür soll allgemein das Wissen für den aus der linearen Algebra stammenden Begriff der Transformation gefördert werden und auf die Bildverarbeitung übertragen werden. Anschließend wird das Verständnis für die Fouriertransformation sukzessive aufgebaut und mit den beiden weiteren Transformationen verknüpft. Abschließend werden die Transformationen verglichen und ihr Nutzen innerhalb der Bildverarbeitung erläutert.

Abstract

This bachelor thesis's objective is to offer the reader insight into the discrete Fourier transform, the discrete cosine transform and the discrete Hadamard-Walsh transform in the context of image processing, and also to compare these transformations under various aspects. For this purpose the term of transformation, originated in linear algebra, will be explained and applied to image processing. Subsequently, the understanding of the Fourier transform will successively be built up and connected to the two remaining transforms. Finally, the transformations will be compared and their usefulness in relation to image processing will be explained.

Erklärung

Ich versichere, dass ich die vorliegende Arbeit selbständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe und dass die Arbeit in gleicher oder ähnlicher Form noch keiner anderen Prüfungsbehörde vorgelegen hat und von dieser als Teil einer Prüfungsleistung angenommen wurde. Alle Ausführungen, die wörtlich oder sinngemäß übernommen wurden, sind als solche gekennzeichnet.

Die Vereinbarung der Arbeitsgruppe für Studien- und Abschlussarbeiten habe ich gelesen und anerkannt, insbesondere die Regelung des Nutzungsrechts.

Mit der Einstellung dieser Arbeit in die Bibliothek bin ich einverstanden. ja nein

Der Veröffentlichung dieser Arbeit im Internet stimme ich zu. ja nein

Koblenz, den 1. März 2018

Inhaltsverzeichnis

1	Einleitung	9
2	Hintergrundwissen	13
2.1	Analysis	13
2.1.1	Konvergenz	13
2.1.2	Reihen	14
2.2	Komplexe Zahlen	15
2.2.1	Komplexe Zahlenebene	15
2.2.2	Zusammenhang: Sinus, Kosinus und Exponentialfunktion	17
2.3	Lineare Algebra	18
2.3.1	Vektorraum	18
2.3.2	Linearkombination	20
2.3.3	Basis	21
2.3.4	Lineare Transformation	22
2.3.5	Skalarprodukt	25
3	Lineare Transformationen	27
3.1	Fouriertransformation	27
3.1.1	Mathematischer Hintergrund	28
3.1.2	Eine Idee einer geeigneten Transformationsmatrix	30
3.1.3	Über die Fourierreihe zur DFT-Matrix	32
3.1.4	Interpretation der Zerlegung einer Bildspalte	36
3.1.5	Beispiel: Fouriertransformation am Spaltenvektor	38
3.1.6	2D-DFT im Zusammenhang mit der Bildverarbeitung	42
3.1.7	Anwendungen der DFT auf Bilder	45
3.1.8	Die Fast Fourier Transform (FFT)	53
3.2	Diskrete Kosinustransformation	57
3.2.1	Ähnlichkeiten zwischen DCT und DFT	57
3.2.2	Wichtige Unterschiede zur DFT	58
3.3	Hadamard-Walsh-Transformation	64
3.3.1	Aufbau der Basisfunktionen	64

3.3.2	Anwendung in der Bildverarbeitung	65
4	Vergleich der Transformationen	71

Kapitel 1

Einleitung

Der französische Mathematiker Jean-Baptiste Joseph Fourier (1768-1830) beschäftigte sich mit der Frage, wie Wärmeleitungsgleichungen gelöst werden können und veröffentlichte 1822 in seinem Werk „Théorie analytique de la chaleur“ seine Ideen dazu [3]. Seine Theorie war es, diese Lösungen mittels geeigneter Exponentialreihen berechnen zu können, was ihn zu der Annahme führte, dass beliebige Funktionen als Summe von Schwingungen darstellbar sind. Diese Idee löste viele Kontroversen aus, konnte aber für periodische Funktionen belegt werden und hat sich 200 Jahre später in der Mathematik und Physik als ein wichtiges Werkzeug etabliert, welches als Fourierreihe beziehungsweise Fouriertransformation bekannt ist.

Schwingungen lassen sich an verschiedenen Stellen finden. Angenommen ein Stein wird in ein stilles Gewässer geworfen, so ergeben sich schwingungsförmige Wellen, die von dem Punkt ausgehen, an dem der Stein die Wasseroberfläche durchbrochen hat. Das Gleiche passiert ebenfalls bei Geräuschen, die erzeugt werden. Ob Melodien, Rauschen, Geraschel oder beim Sprechen, es gibt immer geräuscherzeugende Objekte, deren Informationen über Luftschwingungen getragen, wahrgenommen werden können. Diese Schwingungen sind vielfach überlagert, was sich am Klang von Musikinstrumenten nachvollziehen lässt. Die mechanische Schwingung einer Flöte wird durch das Instrument beim Hineinblasen erzeugt. Die Luftsäule innerhalb der Flöte kann infolge des Öffnens und Schließens der Löcher entweder verlängert oder verkürzt werden, wodurch tiefere oder höhere Töne entstehen. Wird gebläset, ist der entstehende Ton nicht etwa rein, sondern wird durch die Vibration des Klangkörpers, also des vibrierenden Holzes der Flöte, überlagert [15]. Noch deutlicher ist dies bei Instrumenten wahrzunehmen, die einen größeren Korpus besitzen, wie eine Geige oder ein Kontrabass. Dies ist dadurch zu erklären, dass Menschen in der Lage sind Töne von 20 bis 20000 Hz wahrzunehmen und deren Höhenunterschiede in gewissem Maße differenzieren können [4]. Da der Ton einer Geige durch den voluminöseren Klangkörper komplexer aufgebaut ist als der einer Flöte, sind die Überlagerungen der Schwingungen besser herauszuhören [5].

Mittels der Fouriertransformation ist es möglich den Klang eines Sängers oder einer Sängerin zu bearbeiten, indem für das menschliche Gehör unharmonische Schwingungen herausgefiltert werden. Das Endergebnis klingt folglich harmonischer. Dies führt unmittelbar zur Signalverarbeitung, bei der das Eingangssignal, das durch den erzeugten Ton entsteht, über ein Mikrofon aufgenommen und an einen Computer weitergeleitet wird. Dieser komplexe Ton wird vom System durch eine Software in seiner Schwingungsform dargestellt und dann in eine Form umgewandelt, in der er bearbeitet werden kann. Das bearbeitete Ergebnis entspricht dem Ausgangssignal. Es ist üblich, dass Systeme periodische Eingangssignale aufnehmen, die periodisch fortgesetzt werden und diese in eine für sich verwertbare Form umwandeln, bearbeiten und weiterleiten.

Ein weiteres Anwendungsgebiet stellt der Bereich der Optik dar. Der gesamte Zweig der Fourieroptik beschäftigt sich mit Schwingungen, was nicht weiter verwunderlich ist, da Licht ebenfalls aus Schwingungen besteht [8]. Die Fouriertransformation kann also praktisch überall dort genutzt werden, wo Schwingungen auftauchen, diese als eine Art Eingangssignal interpretiert und periodisch fortgesetzt werden. Im Fokus dieser Arbeit steht die Anwendung der Fouriertransformation innerhalb der Bildverarbeitung, einem Anwendungsgebiet, in welchem diese Transformation erst einmal nicht vermutet wird.

Was Bilder mit Schwingungen gemeinsam haben und warum Bilder in Schwingungen zerlegt oder aus ihnen rekonstruiert werden können, wird in dieser Arbeit geklärt werden. Außer der Fouriertransformation gibt es einige andere Transformationen dieser Art, wie die Sinustransformation, die Kosinustransformation und die Hadamard-Walsh-Transformation, die im Grunde alle mit demselben Prinzip arbeiten.

Die Fourierreihe sowie die Fouriertransformation sind stark mit verschiedenen mathematischen Bereichen, wie der linearen Algebra und der Analysis, verzahnt. Diese Arbeit soll den entsprechenden Kontext anwendungsorientiert herstellen und dazu beitragen, den Studierenden einen Zugang zu diesem Teilbereich der Bildverarbeitung zu schaffen. Dabei werden die komplexen Sachverhalte auf das Eindimensionale heruntergebrochen und dann auf die zweidimensionale Anwendung in der Bildverarbeitung übertragen. Mathematische Grundlagen werden hierfür zur Auffrischung angeboten und auf einer für die Thematik angemessenen Ebene behandelt. Weiterhin wurde darauf geachtet, die Unterkapitel in eine Form zu bringen, die den Leser mit einem jeweils angemessenen Informationsgehalt nicht überfordern sollen und getroffene Aussagen zahlreich mit Bildern unterstreichen. Dies soll auf möglichst anschaulichem Wege dazu führen, die anwendungsorientierte Fouriertransformation zu erläutern. Darüber hinaus wurde für das Grundverständnis darauf verzichtet überflüssige Formeln einzubringen, die den Einstieg in die Thematik erschweren könnten. Dies soll dazu führen, dass Studierende das abstrakte, mathematische Konstrukt anwendungsbezogen kennenlernen, die Grundlagen verstehen und in anderen Bereichen sukzessive auf den Kenntnissen aufbauen können. Die Grundlage dieser didaktisch motivierten Arbeit bildet demnach ein anschaulicher Ansatz, der auf mathematisch harte Einstiege verzichtet und den In-

halt über einen nachvollziehbaren Weg vermittelt. Diese Veranschaulichung wird dann mit den nötigen mathematischen Konzepten untermauert.

Nach der Fouriertransformation wird die ihr verwandte Kosinustransformation vorgestellt, deren Kapitel sich die zuvor gelegten Grundlagen zu Nutze macht und einen schnellen Einblick in die Materie gibt. Abschließend wird sich der Hadamard-Walsh-Transformation zugewandt, die sich von den zuvor gezeigten Transformationen etwas abgrenzt, aber durch ihre Einfachheit besticht. Zusammenfassend werden diese drei linearen Transformationen untereinander verglichen und ihre Eigenschaften diskutiert.

Kapitel 2

Hintergrundwissen

In diesem Kapitel werden die mathematischen Grundlagen zum Verständnis der Fouriertransformation gelegt.

2.1 Analysis

Im Abschnitt der Analysis wird sich auf das Konvergenzverhalten sowie die Reihendarstellung von $\cos(x)$, $\sin(x)$ und e^x beschränkt.

2.1.1 Konvergenz

Konvergenz wird oft mit dem Schaubild gemäß Abbildung 2.1 ([9] Quelle: Eigene Darstellung) veranschaulicht. Hier wird das Verhalten einer Funktion über die natürlichen Zahlen a_n mit $n \in \mathbb{N}$ dargestellt. Eine solche Funktion wird auch als Folge bezeichnet. a_n oszilliert um den Grenzwert g , bis bei einer hinreichend großen natürlichen Zahl n fast kein Unterschied mehr zwischen dem Wert von g und a_n existiert. Es gilt, dass a_n für ein unendlich großes n den Wert von g asymptotisch annähert [Pap09].

Konvergenz ist ein essentieller Begriff der Analysis. Im Allgemeinen werden Funktionen auf ihr Verhalten untersucht, wenn bestimmte Zahlen in sie eingesetzt werden. Bei Kurvendiskussionen ist es oft interessant zu wissen, wie sich eine Funktion $f(x)$ für x gegen unendlich (formal: $\lim_{x \rightarrow \infty} f(x)$) oder an Stellen verhält, an denen sie nicht definiert ist, den sogenannten Polstellen. Dazu wird beobachtet wie sich $f(x)$ verhält, wenn x schrittweise gegen einen bestimmten Wert wandert. Nähert sich die Funktion einem festen Wert an, wird dies Konvergenz genannt, andernfalls Divergenz.

Gilt $\lim_{x \rightarrow \infty} f(x) = +\infty$ oder $\lim_{x \rightarrow \infty} f(x) = -\infty$ handelt es sich um eine bestimmte Divergenz [Pap09].

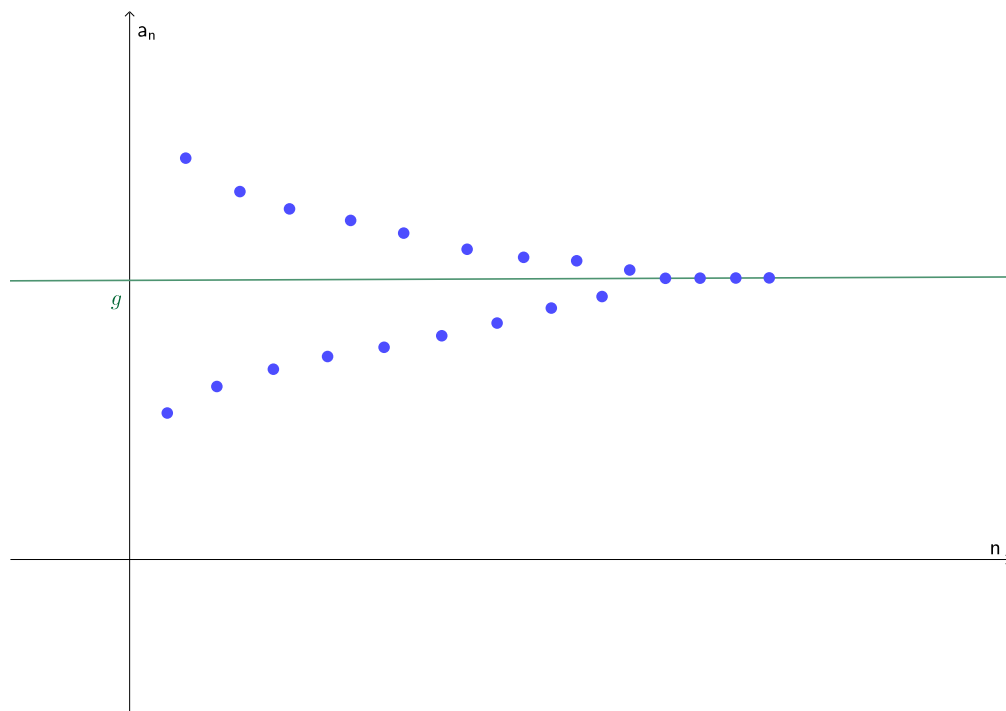


Abbildung 2.1: Konvergenzverhalten einer Folge a_n

2.1.2 Reihen

Das Konvergenzverhalten kann an Summen untersucht werden. Summen, die unendlich viele Summenterme besitzen, werden als „Reihen“ bezeichnet. Bekannte Funktionen wie e^x , $\sin(x)$ und $\cos(x)$ sind als Reihen ausdrückbar:

$$\sin x = \sum_{n=0}^{\infty} \frac{(-1)^{n-1} x^{2n-1}}{(2n-1)!}$$

$$\cos x = \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n}}{(2n)!}$$

$$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!}$$

Das bedeutet, dass unendlich viele aufaddierte Glieder dieser Reihen genau der jeweiligen Funktion entsprechen. Beispielsweise sei e^1 betrachtet. Nach der Reihendarstellung ergibt sich:

$$e^1 = \sum_{n=0}^{\infty} \frac{1^n}{n!} = \sum_{n=0}^{\infty} \frac{1}{n!}$$

Die ersten fünf Glieder ergeben:

$$e^1 = e \approx \sum_{n=0}^4 \frac{1}{n!} = \frac{1}{0!} + \frac{1}{1!} + \frac{1}{2!} + \frac{1}{3!} + \frac{1}{4!}$$
$$e \approx 1 + 1 + \frac{1}{2} + \frac{1}{6} + \frac{1}{24} = 2.708\bar{3}$$

Dies ist eine Näherung der eulerschen Zahl $e \approx 2.718$ und verdeutlicht zusätzlich, dass Reihen ebenfalls konvergent sein können, denn der Grenzwert der Exponentialreihe für $x = 1$ ist e . Im nächsten Kapitel wird ein weiterer nützlicher Zusammenhang untersucht, den diese drei Reihen bieten.

2.2 Komplexe Zahlen

Die bekannten Zahlenmengen beschreiben bestimmte Sachverhalte. Typischerweise sind diese Mengen aus bestimmten Anwendungsfällen hervorgegangen. So werden natürliche Zahlen dazu genutzt, Gegenstände zu zählen, und rationale Zahlen, um Verhältnisse zum Ausdruck zu bringen. Letztendlich verhält es sich mit komplexen Zahlen als Erweiterung der reellen Zahlen ähnlich. Anwendungsgebiete finden sich in der Mathematik und Physik, wo sich Zusammenhänge leichter mit Hilfe der komplexen Zahlen ausdrücken und berechnen lassen. Als Anwendungsbeispiel wird im Kapitel 3.1 die Fouriertransformation vorgestellt, für die die Berechnung und Darstellung durch komplexe Zahlen vereinfacht wird. An dieser Stelle wird auf die komplexe Zahlenebene eingegangen, an der noch einmal verdeutlicht wird, wie sich komplexe Zahlen vorzustellen sind. Des Weiteren wird der Bezug zwischen der komplexen Exponentialreihe, zur Sinusreihe und Kosinusreihe hergestellt und darauf aufbauend eine Interpretation der komplexen Exponentialfunktion gegeben.

2.2.1 Komplexe Zahlenebene

Komplexe Zahlen (\mathbb{C}) in ihrer allgemeinen Form $z = a + bi$ bestehen aus einem Realteil $a \in \mathbb{R}$ und einem Imaginärteil $b \in \mathbb{R}$. i entspricht der imaginären Einheit, die es zulässt,

die Wurzel aus einer negativen Zahl zu ziehen. Der Wert von i ist auf $\sqrt{-1}$ festgelegt. Die Zahl z wird als zweikomponentige Zahl verstanden, die in einer Ebene abgebildet wird, analog zur Menge der Zahlen des \mathbb{R}^2 [MW06]. Die x-Achse bildet den Realteil ab und die y-Achse (Imaginärachse) den Imaginärteil, wie Abbildung 2.2 ([9] Quelle: Eigene Darstellung) zeigt:

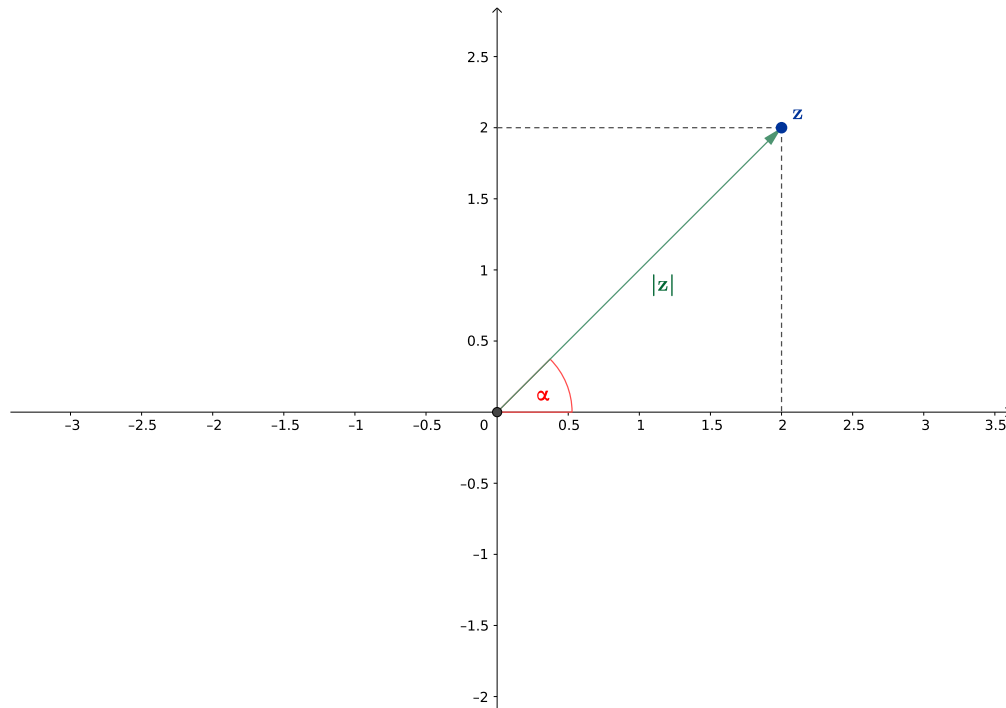


Abbildung 2.2: Komplexe Zahlenebene

Der Betrag einer komplexen Zahl (dargestellt als grüner Vektor) ergibt sich aus dem Abstand vom Ursprung zu den Koordinaten von z . Dieser ergibt sich aus der Wurzel der Summe der Quadrate vom Real- und Imaginärteil. Hier konkret:

$$|z| = \sqrt{2^2 + 2^2} = \sqrt{8}$$

Der Betrag wird auch als Amplitude bezeichnet. Die Phase oder der Winkel, den der Vektor mit der x-Achse einschließt, berechnet sich durch den Arcustangens:

$$\alpha = \arctan\left(\frac{2}{2}\right) = 45.0^\circ$$

Festzuhalten ist, dass komplexe Zahlen eine Zahlenebene bilden und jede komplexe Zahl durch ihre Phase und Amplitude eindeutig darstellbar ist.

2.2.2 Zusammenhang: Sinus, Kosinus und Exponentialfunktion

Zuvor wurde angedeutet, dass zwischen den Reihen ein interessanter Zusammenhang hergestellt werden kann. Durch die Reihendarstellung von Kosinus und Sinus lässt sich zeigen, dass

$$e^{ix} = \cos(x) + i \sin(x)$$

gilt.

Erläuterung:

$$\begin{aligned} e^{ix} &= 1 + ix + \frac{(ix)^2}{2!} + \frac{(ix)^3}{3!} + \frac{(ix)^4}{4!} + \frac{(ix)^5}{5!} + \dots \\ &= \left(1 - \frac{(x)^2}{2!} + \frac{(x)^4}{4!} - \dots\right) + i\left(x - \frac{(x)^3}{3!} + \frac{(x)^5}{5!} - \dots\right) \\ &= \cos(x) + i \sin(x) \end{aligned}$$

Benjamin Peirce (1809-1880) soll einst über die eulersche Identität ($e^{i\pi} + 1 = 0$) gesagt haben [1]:

„Gentlemen, that is surely true, it is absolutely paradoxical; we cannot understand it, and we don't know what it means. But we have proved it, and therefore we know it is the truth.“

Es genügt an dieser Stelle die komplexe Exponentialfunktion aus einem bestimmten Blickwinkel zu verstehen. Da die e -Funktion über $\sin(x)$ und $\cos(x)$ ausgedrückt wird, kann bei der Interpretation stets auf beide Funktionen zurückgegriffen werden. In Abbildung 2.3 ([9] Quelle: Eigene Darstellung) sind die Punkte P_1 und P_2 auf dem Einheitskreis gegeben. P_1 wird durch $\cos(\alpha)$ in x-Richtung beschrieben und durch $\sin(\alpha)$ in y-Richtung. Gezeigt wurde, dass dieser Punkt P_1 genau $e^{i1\alpha}$ entspricht. Das Argument der Exponentialfunktion multipliziert mit 2 ergibt den Punkt P_2 . Ein Punkt, der immer

noch auf dem Einheitskreis liegt, allerdings um 45° gegen den Uhrzeigersinn um den Nullpunkt gedreht ist. Wird der Faktor des Winkels als Zeit t und der Einheitskreis als Uhr interpretiert, so würde $t = 3$ der 22 Uhr-Position entsprechen. Anschaulich kann der Winkel als schrittweise Abtastung der Uhrzeit verstanden werden, was bei einer normalen Uhr entweder Sekunden oder Minuten sind. Entsprechend einer haushaltsüblichen Uhr wäre also die Schrittweite in einen Winkel umgerechnet 5° . t kann nach der Analogie als konkrete Angabe des Zeitpunktes der aktuellen Stunde verstanden werden. Allgemein gilt, dass eine Veränderung des Arguments der Exponentialfunktion als eine Drehung (Rotation) der komplexen Zahl zu verstehen ist.

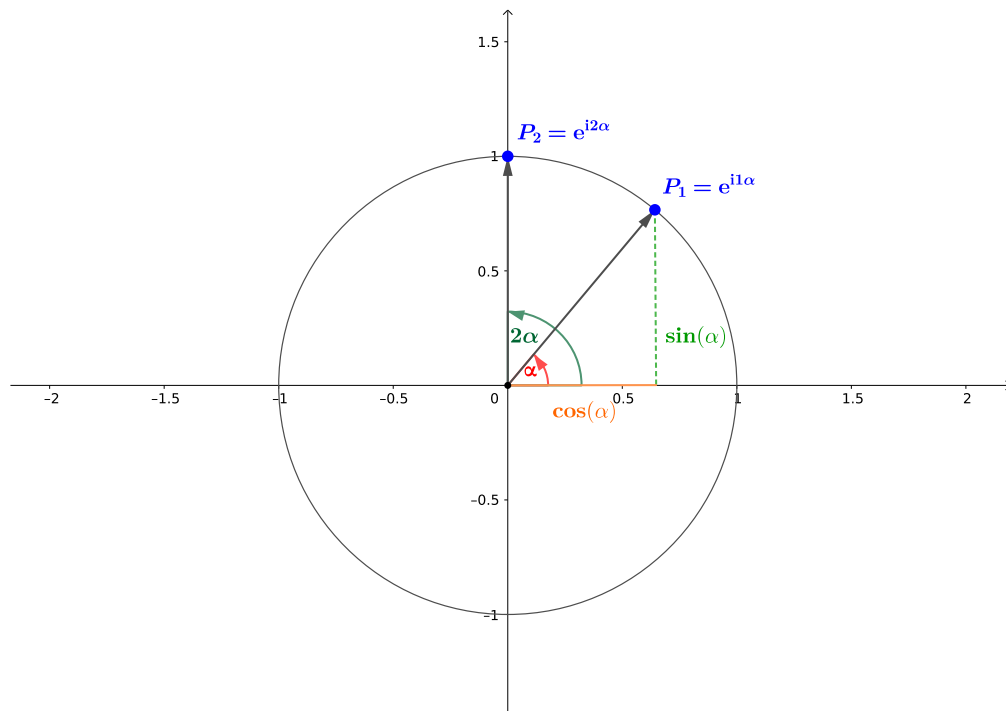


Abbildung 2.3: Komplexer Einheitskreis

2.3 Lineare Algebra

Das Kapitel der linearen Algebra stellt eine Wiederholung der Begriffe Vektorraum, Linearkombination, Basis, Transformation und Skalarprodukt dar.

2.3.1 Vektorraum

Den zentralen Begriff der linearen Algebra bildet der Vektorraum [Beu09]. Vektorräume werden vor allem im \mathbb{R}^n betrachtet, wobei n die Dimensionen des Raumes beschreibt.

Vereinfachend lassen sich die Dimensionen gedanklich auf den \mathbb{R}^2 (ein Blatt Papier) oder auf den \mathbb{R}^3 (einen Raum) reduzieren. In der folgenden Abbildung 2.4 ([10] Quelle: Eigene Darstellung) werden die Komponenten eines Vektorraums ersichtlich:

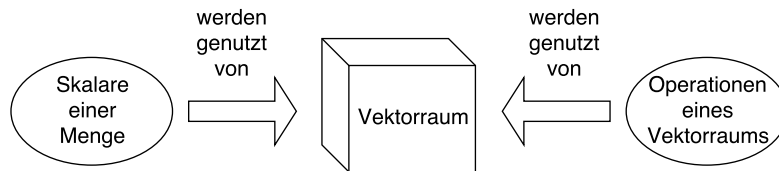


Abbildung 2.4: Graphische Darstellung der Zusammensetzung eines Vektorraums

In der Mitte der Abbildung stellt der Würfel den Vektorraum dar. Dies ist als Veranschaulichung zu verstehen, denn der Vektorraum des \mathbb{R}^3 ist keineswegs endlich, wie das Bild suggeriert. Ein Vektorraum wird hauptsächlich über die Skalare und über die Operationen mit diesen und den Vektoren definiert. Skalare entspringen einer bestimmten Menge. Da diese Menge die Körperaxiome erfüllen muss, wird meistens von Skalaren aus den Mengen \mathbb{R} oder \mathbb{C} , den reellen oder den komplexen Zahlen, gesprochen [Beu09]. Demnach sind Skalare nichts weiter als Variablen, die diesen Mengen angehören, wie etwa 2, da $2 \in \mathbb{R}$.

Wie im Kapitel über komplexe Zahlen 2.2.1 erläutert, können Vektoren als Zeiger auf einen bestimmten Punkt verstanden werden. Diese Vektoren sind fester Bestandteil des Vektorraums. Durch sie kann jeder Stelle des Raumes ein fester Wert zugeordnet werden. Innerhalb dieser Arbeit werden Vektoren wie folgt notiert:

$$\mathbf{v} = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix}$$

Die Einträge v_1 , v_2 und v_3 werden als Komponenten des Vektors bezeichnet und ihre Anzahl spiegelt gleichzeitig die Dimension des Raumes wider. In diesem Fall also den \mathbb{R}^3 . Die Verknüpfung der Vektoren und der Skalare findet über geläufige Operationen statt:

$$\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} + \begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix} = \begin{pmatrix} 5 \\ 7 \\ 9 \end{pmatrix}$$

Durch eine einfache Addition zweier Vektoren kann demnach ein anderer Vektor beschrieben werden.

Festzuhalten ist, dass Vektorräume aus drei Komponenten bestehen:

1. Skalare, eine Menge von reellen oder komplexen Zahlen.
2. Vektoren, die Grundbausteine eines jeden Vektorraumes, die beschreiben, wo etwas innerhalb des Raumes liegt.
3. Operationen, mit deren Hilfe verschiedene Sachverhalte ausgedrückt werden können.

2.3.2 Linearkombination

Wie zuvor angesprochen, können Vektoren durch Operationen verknüpft werden. Aus der Multiplikation eines Vektors mit einem Skalar ergibt sich ein neuer Vektor, wobei jede Komponente des alten Vektors mit dem Skalar multipliziert wurde:

$$3 \begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix} = \begin{pmatrix} 12 \\ 15 \\ 18 \end{pmatrix}$$

Aus dieser Operation und der Addition lässt sich der Zusammenhang der Linearkombination beschreiben:

$$\sum_{n=1}^n s_n \mathbf{v}_n = s_1 \mathbf{v}_1 + s_2 \mathbf{v}_2 + s_3 \mathbf{v}_3 + \dots$$

Wobei s_n aus der Menge der Skalare und \mathbf{v}_n aus der Menge der Vektoren stammt. Wie die Abbildung 2.5 ([9] Quelle: Eigene Darstellung) verdeutlicht, ist es gemäß dieser Formel möglich, mit Hilfe einer gewichteten Kombination anderer Vektoren auf genau denselben Punkt zu zeigen, der schon durch einen Vektor beschrieben ist:

Es gilt:

$$\mathbf{w} = 1\mathbf{u} + 0.5\mathbf{v} = 1 \begin{pmatrix} 3 \\ 3 \end{pmatrix} + 0.5 \begin{pmatrix} 4 \\ 0 \end{pmatrix} = \begin{pmatrix} 5 \\ 3 \end{pmatrix}$$

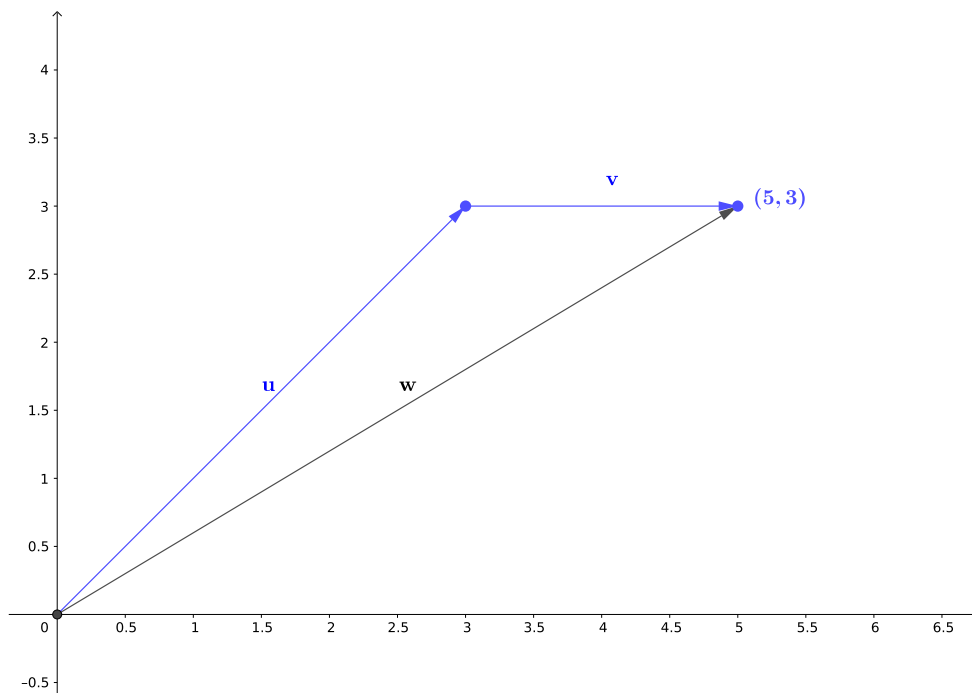


Abbildung 2.5: Beispiel einer Linearkombination

Vereinfacht gesagt, wird die volle Länge des Vektors u genommen und von dem Punkt aus, auf den dieser zeigt, mit Vektor v weitergegangen. Da v allerdings vier Schritte in x -Richtung geht und nur zwei benötigt werden, wird seine x -Richtung durch den Faktor 0.5 gekürzt, sodass der gewünschte Punkt erreicht wird. Jener Punkt, auf den w ebenfalls zeigt. Zuvor wurde von der gewichteten Summe der Vektoren gesprochen. Die Skalare entsprechen bei der Linearkombination den Gewichten, da sie das Verhältnis im Sinne davon angeben, „wie viel“ oder „wie wenig“ von dem jeweiligen Vektor benötigt wird.

2.3.3 Basis

Wie erläutert, können Vektoren durch die gewichtete Summe anderer Vektoren ausgedrückt werden. Auf Grund dessen stellt sich die Frage, ob eine Mindestmenge existiert, aus der alle Vektoren eines Vektorraumes mit Hilfe der Linearkombination dargestellt werden können. Diese Mindestmenge von Vektoren wird als Basis bezeichnet. Ein einfaches Beispiel dafür ist die kanonische Basis des \mathbb{R}^2 :

$$B = \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\}$$

Alle Vektoren des \mathbb{R}^2 können durch die Linearkombination dieser beiden Vektoren ausgedrückt werden. Es sei erwähnt, dass Vektoren einer Basis auch als Basisvektoren bezeichnet werden. Es ist nicht möglich, einen dieser beiden Basisvektoren wegzulassen, da sonst die x-Richtung beziehungsweise y-Richtung nicht mehr beschrieben werden kann. Dies zeigt, dass die beiden Vektoren linear unabhängig voneinander sind, also nicht jeweils durch den anderen Basisvektor ausgedrückt werden können. Demnach bildet diese zweielementige Menge eine Mindestmenge des \mathbb{R}^2 [Sch10].

Festzuhalten ist, dass ein Vektorraum unendlich viele Basen haben kann und dass mit Hilfe einer Linearkombination der Basisvektoren alle Vektoren innerhalb des Vektorraums darstellbar sind.

2.3.4 Lineare Transformation

Knapp formuliert könnte gesagt werden, dass Lineare Transformationen Vektoren zwischen verschiedenen Basen abbilden [AHK⁺09]. Da Transformationen die wichtigste Grundlage sind, wird an dieser Stelle noch einmal genauer auf sie eingegangen.

Wie im vorherigen Abschnitt festgestellt, kann ein Vektorraum durch verschiedene Basen beschrieben werden. Angenommen der \mathbb{R}^2 wird durch zwei verschiedene Basen B_1 und B_2 beschrieben und die beiden Darstellungsmöglichkeiten sollen in Bezug zueinander gesetzt werden. Dazu muss die Überlegung getroffen werden, wie Vektoren, die durch B_1 erstellt werden, mit Hilfe von B_2 ausgedrückt werden können. Selbiges gilt für den umgekehrten Fall. Das folgende Schaubild 2.6 ([10] Quelle: Eigene Darstellung) beschreibt diesen Zusammenhang:

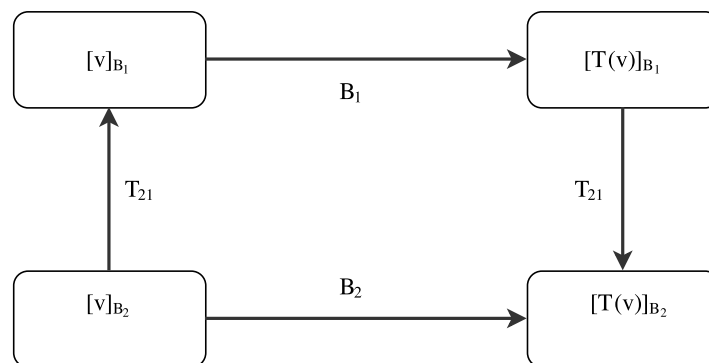


Abbildung 2.6: Transformationszyklus von zwei Basen

Gegeben sind zwei verschiedene Basen B_1 und B_2 . Ein Vektor v wird jeweils durch die Basis B_1 oder B_2 so abgebildet, dass sich das Bild $T[v]_{B_1}$ beziehungsweise das Bild $T[v]_{B_2}$ ergibt. Im Folgenden wird nun untersucht, wie die Transformationsmatrix von $T[v]_{B_1}$ zu $T[v]_{B_2}$, also T_{12} , berechnet werden kann. Hierzu seien folgende Basen gegeben:

$$B_1 = \left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \end{pmatrix} \right\}$$

$$B_2 = \left\{ \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 2 \\ 3 \end{pmatrix} \right\}$$

Es besteht die Möglichkeit, einen beliebigen Vektor mit den Komponenten v_1 und v_2 zu nehmen, diesen mit der Matrix, bestehend aus den Basisvektoren, zu multiplizieren und die Darstellung mit der jeweiligen Basis zu erhalten:

$$\begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} \quad \text{und} \quad \begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}$$

Angenommen der Vektor v besitzt die Komponenten $v_1 = 2$ und $v_2 = 2$, dann berechnet sich die Abbildung $T[v]_{B_1}$ über diese Matrix durch:

$$T[v]_{B_1} = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} 2 \\ 2 \end{pmatrix} = \begin{pmatrix} 4 \\ 6 \end{pmatrix}$$

Das Bild ist in dieser Basis der Vektor mit den Komponenten 4 und 6. Für die Basis B_2 ergibt sich allerdings ein anderes Bild:

$$T[v]_{B_2} = \begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix} \begin{pmatrix} 2 \\ 2 \end{pmatrix} = \begin{pmatrix} 6 \\ 10 \end{pmatrix}$$

Hieraus folgt, dass das gewünschte Ergebnis $T[v]_{B_1} = T[v]_{B_2}$ nicht gegeben ist. Gesucht ist demnach die Lösung dieser Gleichung, um den zugehörigen Vektor für B_2 zu finden:

$$\begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix} \begin{pmatrix} t_1 & t_2 \\ t_3 & t_4 \end{pmatrix} \Rightarrow \begin{pmatrix} t_1 & t_2 \\ t_3 & t_4 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix}^{-1} \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$$

Mit der Berechnung der inversen Matrix, ergibt sich:

$$\begin{pmatrix} t_1 & t_2 \\ t_3 & t_4 \end{pmatrix} = \begin{pmatrix} -3 & 2 \\ 2 & -1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} = \begin{pmatrix} -1 & 1 \\ 1 & 0 \end{pmatrix}$$

Dies entspricht der gesuchten T_{12} -Matrix. Der gesuchte Vektor berechnet sich nun aus:

$$\mathbf{T}[\mathbf{v}]_{B_2} = \begin{pmatrix} -1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 2 \\ 2 \end{pmatrix} = \begin{pmatrix} 0 \\ 2 \end{pmatrix}$$

Zur Probe wird

$$\begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix} \begin{pmatrix} 0 \\ 2 \end{pmatrix} = \begin{pmatrix} 4 \\ 6 \end{pmatrix}$$

berechnet. Dies ergibt das gleiche Ergebnis.

Das Bild, welches entsteht, wenn eine Basis mit einem Vektor multipliziert wird, wird als Transformation des Vektors an der Basis bezeichnet:

$$\mathbf{T}[\mathbf{v}] = \mathbf{T}\mathbf{v}$$

Die Umkehrung entspricht:

$$\mathbf{v}^{-1}\mathbf{T}[\mathbf{v}] = \mathbf{T}$$

Kurzum, es wird eine beliebige Basis und ein Vektor benötigt, um diesen Vektor der Basis spezifisch abzubilden. Transformationen dieser Art sind nicht auf die Dimensionen beschränkt und können daher ebenfalls im \mathbb{R}^n ausgeführt werden.

2.3.5 Skalarprodukt

In diesem Abschnitt wird erläutert, wie das Skalarprodukt als ein Ähnlichkeitsmaß genutzt werden kann. Seien die Vektoren

$$\mathbf{v} = \begin{pmatrix} 1 \\ -3 \end{pmatrix}, \mathbf{u} = \begin{pmatrix} 3 \\ -1 \end{pmatrix}$$

gegeben, so berechnet sich das Skalarprodukt aus der Summe der komponentenweisen Multiplikation:

$$\langle \mathbf{v}, \mathbf{u} \rangle = \begin{pmatrix} 1 \\ -3 \end{pmatrix} \circ \begin{pmatrix} 3 \\ 1 \end{pmatrix} = 1 \cdot 3 + (-3 \cdot (1)) = 0$$

Das Ergebnis 0 zeigt die Orthogonalität der Vektoren. Allgemein ist das Skalarprodukt definiert als:

$$\langle \mathbf{v}, \mathbf{u} \rangle = \mathbf{v} \circ \mathbf{u} = |\mathbf{v}| |\mathbf{u}| \cdot \cos(\alpha)$$

Wobei α den Winkel zwischen den Vektoren bezeichnet und $|\mathbf{v}|$ die Länge von \mathbf{v} beschreibt. Es gilt $\cos(90^\circ) = 0$, daher ist auch das Skalarprodukt 0. Das erklärt, weshalb für eine Prüfung auf Orthogonalität zweier Vektoren $\langle \mathbf{v}, \mathbf{u} \rangle = 0$ getestet wird.

Angenommen \mathbf{u} und \mathbf{v} sind normiert, dann ist die Länge beider Vektoren jeweils 1. In der oben genannten Formel ergibt sich nun:

$$\langle \mathbf{v}_n, \mathbf{u}_n \rangle = \mathbf{v}_n \circ \mathbf{u}_n = |\mathbf{v}_n| |\mathbf{u}_n| \cdot \cos(\alpha) = 1 \cdot 1 \cdot \cos(\alpha) = \cos(\alpha)$$

Somit bewegen sich die Werte, die durch das Skalarprodukt angenommen werden, nur im Wertebereich des Kosinus (im Intervall $[-1, 1]$). Der maximale Wert zweier Vektoren ergibt sich genau dann, wenn der Winkel zwischen ihnen 0° ist und beide Vektoren in dieselbe Richtung zeigen. Das Ergebnis ist dann 1. Analog ergibt sich der niedrigste Wert, wenn beide Vektoren im 180° Winkel zueinander stehen, also in entgegengesetzte Richtungen zeigen. In diesem Fall wäre das Ergebnis -1 .

Das Skalarprodukt liefert demnach Informationen über die Beziehung der untersuchten Vektoren.

Kapitel 3

Lineare Transformationen

3.1 Fouriertransformation

In der Bildverarbeitung ist das Bild von Lenna geläufig [11]. Kantendetektion und Glättung von Bildern durch Faltung, wie in Abbildung 3.1 dargestellt, wird vorausgesetzt.

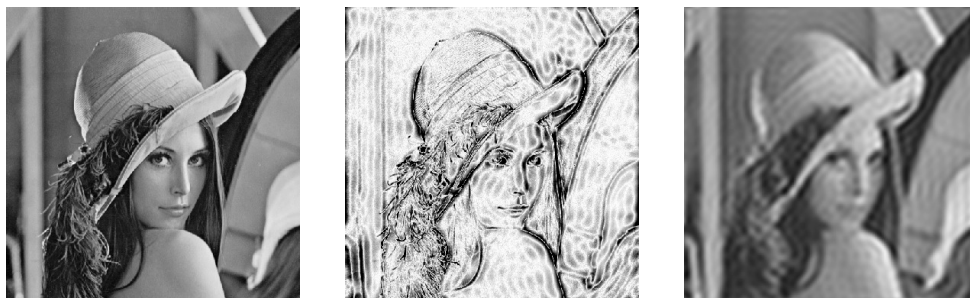


Abbildung 3.1: v.l.n.r. Lenna; Kantendetektion (invertiert); Glättung

Besonders ist, dass diese beiden Operationen hier nicht das Resultat von Faltungen sind, sondern das der diskreten Fouriertransformation (DFT). Die DFT ist eine Basistransformation (siehe Lineare Transformationen 2.3.4), die ein Bild in einen Funktionsraum abbildet, der nur aus Schwingungen besteht. Dieser sogenannte Frequenzraum erleichtert die einfachere Handhabung von Filterungen, lässt sich effizient berechnen und bietet weitere Anwendungsfelder. Die Bilder werden im Frequenzraum so abgeändert, sodass beispielsweise wie in Abbildung 3.1 bei der Rücktransformation eine Kantendetektion oder Glättung entsteht. Faltungsmasken wie der Sobelfilter, werden im Frequenzraum nicht benötigt, sondern durch pixelweise Multiplikationen ersetzt. Alle Bearbeitungen die nicht mit Geogebra oder DrawIO erstellt wurden sind mittels Python und OpenCV erstellt [6].

3.1.1 Mathematischer Hintergrund

Periodische Signale (oder Funktionen) sind überlagerte sinusoidale Schwingungen, die in ihre Bestandteile zerlegt und aus ihren Bestandteilen wieder rekonstruiert werden können. Diese Zerlegung beziehungsweise Zusammensetzung ist eindeutig. Dies ist näherungsweise die Aussage der Fourierreihe [AHK⁺09]. Diese Zerlegung von Signalen entspricht der in ihre Grundschwingung und Oberschwingungen, wobei Oberschwingungen auch als Harmonische bezeichnet werden [AHK⁺09].

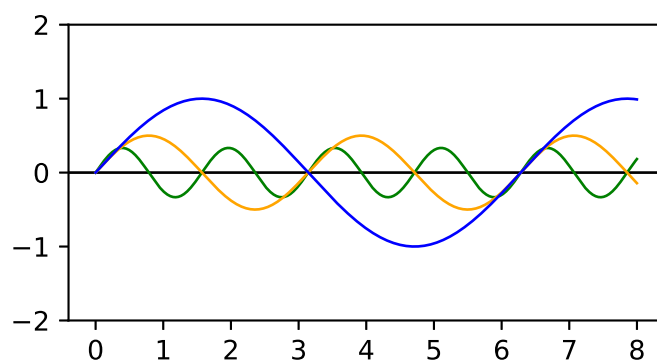


Abbildung 3.2: Grundschwingung und Oberschwingungen einer Sinusfunktion

In Abbildung 3.2 ist zu erkennen, dass die Grundschwingung (blau) der Sinusfunktion einmal im Intervall $[0, 2\pi]$ durchschwingt, wobei die Amplitude 1 beträgt. Bei der orangen Sinusfunktion handelt es sich um die erste Oberschwingung (erste Harmonische). Diese Funktion schwingt mit einer geringeren Amplitude im selben Intervall zweimal komplett durch, wobei der erste und letzte Nulldurchgang mit der Grundschwingung übereinstimmen. Analog nimmt bei der zweiten Harmonischen (grün) die Amplitude weiter ab und die Funktion schwingt doppelt so oft in demselben Intervall wie die erste Harmonische. Im dem Kontext der DFT sind Oberschwingungen und Grundschwingungen mit dem Begriff der Frequenz oft austauschbar, da durch die jeweiligen Oberschwingungen die Frequenz implizit gegeben ist. Die erste Harmonische beschreibt die doppelte Frequenz der Grundschwingung.

Dass es grundsätzlich möglich ist, verschiedenste periodische Signale zu zerlegen, wird an dieser Rechteckfunktion gezeigt (siehe Abbildung 3.3):

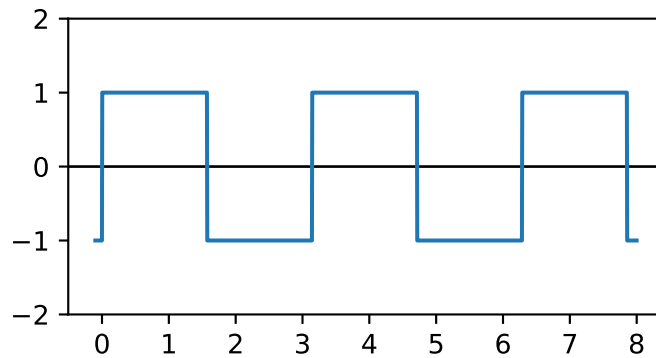


Abbildung 3.3: Rechtecksignal

Charakteristisch für Rechtecksignale ist, dass sie stufenweise Werte annehmen. In diesem Fall die Werte -1 und 1. Werden Schwingungen in einer bestimmten Weise überlagert, fällt auf, dass diese einem Rechtecksignal immer ähnlicher werden. Abbildung 3.4 verdeutlicht dies in einer Abfolge von Bildern:

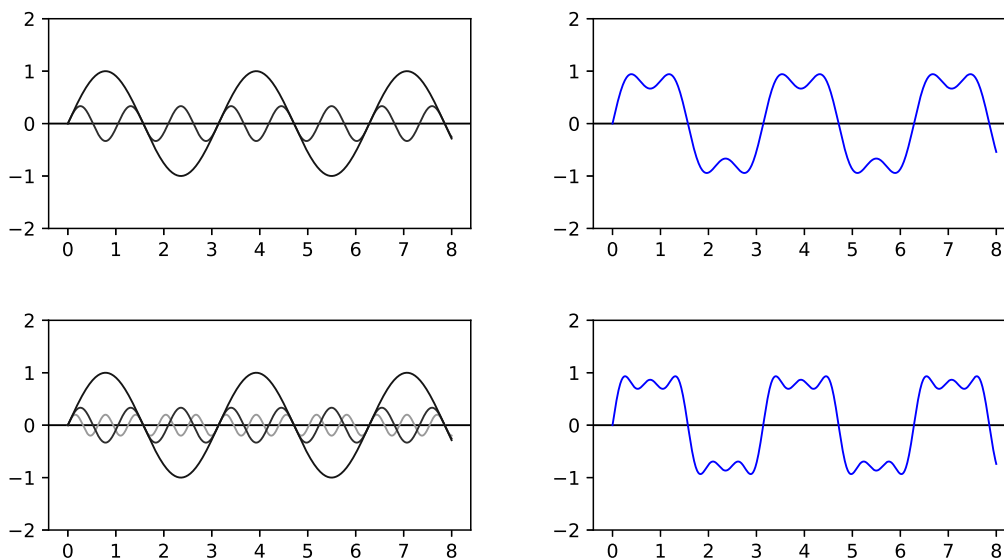


Abbildung 3.4: Approximationen eines Rechtecksignals

Auf der linken Seite kann der Verlauf der verschiedenen Sinuskurven nachvollzogen werden. Auf der rechten Seite sind nun die Funktionen aufaddiert und bilden die jeweils dargestellte blaue Funktion. Wie im Verlauf von oben nach unten zu erkennen ist, wird das Rechtecksignal, welches dargestellt werden soll, angenähert. Die auffällig starken Überschwingungen bei den Sprüngen von -1 auf 1 oder umgekehrt werden als Gibbs-Phänomen bezeichnet und treten immer dann auf, wenn die Ursprungsfunktion eine Sprungstelle hat [7]. Dieses Überschwingen wird für weitere dieser Art hinzugefügte Sinusschwingungen verschwindend gering. Grund dafür ist, dass die Summe dieser Schwingungen sich der Funktion immer weiter annähert, da die Fourierreihe unter gewissen Umständen konvergent ist (siehe Konvergenz 2.1.1) [AHK⁺09].

In diesem Abschnitt wurde eine interessante Betrachtung von Funktionen eingeführt, die auf die Fourierreihe zurückzuführen ist. Grundbausteine verschiedener periodischer Funktionen sind Schwingungen, die in ihrer gewichteten Summe die Funktion wieder darstellen (siehe Kapitel 2.3.2). Umgekehrt kann eine Funktion in ihre einzelnen gewichteten Schwingungen zerlegt werden.

3.1.2 Eine Idee einer geeigneten Transformationsmatrix

In diesem Kapitel wird anschaulich erklärt, wie die Transformationsmatrix der Fouriertransformation aufgebaut ist. Weiterhin wird in den folgenden Kapiteln eine Übersicht 3.5 ([10] Quelle: Eigene Darstellung) angeboten, auf der mit grün farblich markiert ist, was bereits behandelt wurde, beziehungsweise vertieft wird, und mit blau aufzeigt, was im Kapitel thematisch neu hinzukommt:

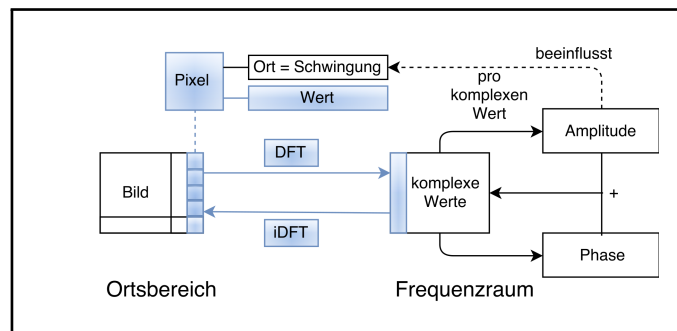


Abbildung 3.5: Übersicht

Bei der Bearbeitung der DFT wird, der Abbildung 3.5 gemäß, das im Verlauf zu behandelnde Bild zunächst zeilenweise untersucht. Die einzelnen Pixel wurden in der Grafik explizit aufgelistet, um zu verdeutlichen, dass sowohl der Pixelwert als auch der Pixelort eine Rolle spielen werden. Das Bild wird vom Ortsbereich über die DFT-Matrix in

den Frequenzbereich abgebildet, wobei komplexe Werte entstehen. Aus diesen werden Amplitude und Phase berechnet (siehe Kapitel 2.2.1) [Bov09]. Die bearbeitete Amplitude hat Auswirkungen auf die Schwingungen. Da es sich um eine lineare Transformation handelt, kann von den veränderten komplexen Werten über die inverse DFT-Matrix zu dem Bild zurücktransformiert werden. Gesucht ist eine Transformationsmatrix

$$\mathbf{T}[\mathbf{v}] = \mathbf{T}\mathbf{v},$$

die diese Transformation vom Ortsbereich in den Frequenzbereich und zurück bewerkstelligen kann. Der Ortsraum wird intuitiv als ein Raum verstanden, der durch Vektoren in kartesischen Koordinaten beschrieben werden kann. Der Begriff Ortsbereich ist ebenfalls geläufig und kann synonym verwendet werden. Der Frequenzraum oder Frequenzbereich ist ein Funktionsraum, der aus Schwingungen besteht [Bov09]. Mathematisch gesehen ist es unproblematisch einen Vektor des \mathbb{R}^2 oder einen Vektor des \mathbb{R}^5 zu transformieren. Neu ist, dass Vektoren \mathbf{v} auf Vektoren $\mathbf{T}[\mathbf{v}]$ abgebildet werden sollen, die Informationen von Schwingungen in ihren Komponenten enthalten. Gibt es eine solche Matrix, so kann diese auch wieder in den Ortsbereich zurücktransformieren:

$$\mathbf{v} = \mathbf{T}^{-1}\mathbf{T}[\mathbf{v}]$$

Dies entspricht der Eigenschaft einer linearen Transformation. Gewünscht ist eine Transformationsmatrix, die zeilenweise als Basisfunktionen Schwingungen enthält. Jede Schwingung wird hierbei mit dem zu transformierenden Vektor \mathbf{v} verknüpft, woraus das Ergebnis dieser Verknüpfung F_k resultiert.

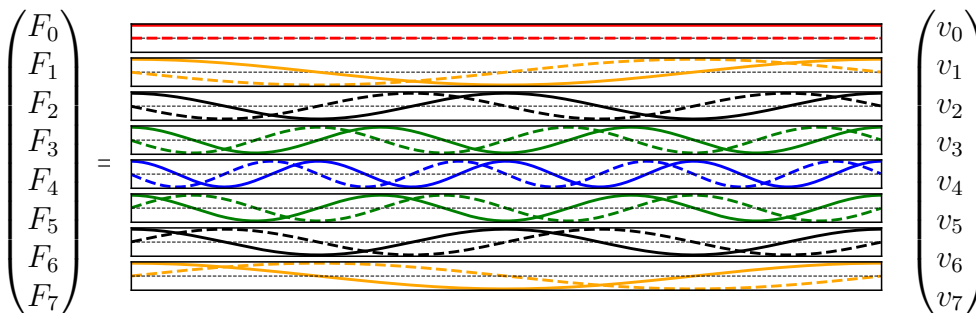


Abbildung 3.6: Transformationsmatrix

Die Matrix besteht beispielsweise aus acht Zeilen (wie in Abbildung 3.6 dargestellt), von denen jede wiederum eine gestrichelte und eine durchgehend gezeichnete Funktion enthält. Die gestrichelte Funktion entspricht hierbei einer Sinusfunktion und die durchgezogene einer Kosinusfunktion. Der Wertebereich jeder dieser Funktionen liegt zwischen -1 und 1 . Die gestrichelte horizontale schwarze Linie kennzeichnet die x -Achse

durch den 0-Punkt. Natürlich würde eine Multiplikation kontinuierlicher Funktionen mit diskreten Werten des Vektors v nicht funktionieren. Die Funktionen innerhalb der Matrix müssten abgetastet werden, sodass sinnvollerweise in diesem Beispiel zeilenweise acht diskrete Werte enthalten wären. Diese diskretisierten Funktionen werden auch als Basisfunktionen bezeichnet und ihre einzelnen abgetasteten Werte werden als Samples bezeichnet. Dass als Basisfunktionen Sinus- und Kosinusfunktionen gefordert sind, ist kein Zufall. Dies wird im nachfolgenden Kapitel deutlich werden. Jede berechnete Komponente F_k des Vektors $T[v]$ trägt Informationen über die multiplizierte Schwingung in sich.

Ziel war es, die Idee der Matrix aufzugreifen, die den Kern der DFT bildet und nichts weiter ist als reihenweise diskretisierte Werte abgetasteter Schwingungen, die mit einem Vektor multipliziert, diesen in den Frequenzbereich abbilden.

3.1.3 Über die Fourierreihe zur DFT-Matrix

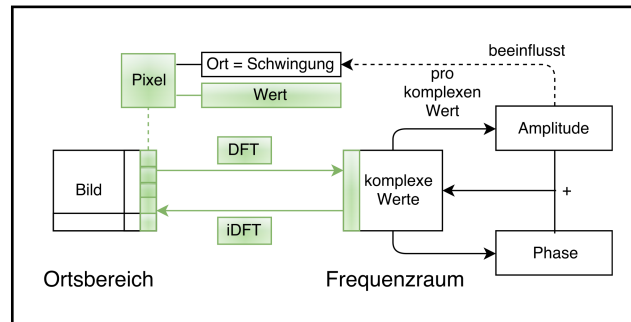


Abbildung 3.7: Übersicht

In diesem Kapitel wird erläutert, wie die Definition der Fouriertransformation in eine kompakte Matrixschreibweise überführt werden kann. Die Fourierreihe, deren funktionsweise am Rechtecksignal verdeutlicht wurde (siehe Kapitel 3.1.1), ist in ihrer komplexen Form gegeben als [AHK⁺09]:

$$f(x) = \sum_{n=0}^{N-1} (e^{\frac{-ikn2\pi}{N}}) p_n$$

Die Formel drückt die Zerlegung einer Funktion $f(x)$ in eine gewichtete Summe ihrer zu Grunde liegenden Schwingungen gemäß der komplexen e -Funktion aus. Die e -Funktion kann entsprechend des Zusammenhangs des Kapitels 2.2.2 in

$$e^{\frac{-ikn2\pi}{N}} = \cos\left(\frac{-ikn2\pi}{N}\right) + i \sin\left(\frac{-ikn2\pi}{N}\right)$$

zerlegt werden, daher besteht $f(x)$ aus sinusoidalen Funktionen. p_n entspricht in der Bildverarbeitung den einzelnen Pixelwerten. Wird der Fokus auf die einzelnen Informationen der vorkommenden Schwingungen, gemäß der Zerlegung, gelegt aus der die Funktion $f(x)$ besteht, ergibt sich die Formel der Fouriertransformation:

$$F_k = \sum_{n=0}^{N-1} (e^{-\frac{ikn2\pi}{N}}) p_n$$

Jedem F_k wird durch die Multiplikation der Pixelwerte mit den gegebenen e -Funktionen ein komplexer Wert zugeordnet. k gibt an, um welche Oberschwingung es sich handelt, die gerade untersucht wird. Demnach würde $k = 1$ bedeuten, dass der komplexe Wert der Grundschwingung ermittelt wird. N entspricht der Anzahl der Komponenten des Vektors. n beschreibt als Laufindex der Summe den Zugriff auf den jeweiligen Pixelwert des Vektors p , der zerlegt werden soll.

Angewandt auf fortlaufendes n und k folgt gemäß der Formel:

$$F_0 = e^{-\frac{i2\pi \cdot 0 \cdot 0}{N}} p_0 + e^{-\frac{i2\pi \cdot 0 \cdot 1}{N}} p_1 + \dots + e^{-\frac{i2\pi \cdot 0 \cdot (n-1)}{N}} p_{n-1}$$

$$F_1 = e^{-\frac{i2\pi \cdot 1 \cdot 0}{N}} p_0 + e^{-\frac{i2\pi \cdot 1 \cdot 1}{N}} p_1 + \dots + e^{-\frac{i2\pi \cdot 1 \cdot (n-1)}{N}} p_{n-1}$$

$$F_2 = e^{-\frac{i2\pi \cdot 2 \cdot 0}{N}} p_0 + e^{-\frac{i2\pi \cdot 2 \cdot 1}{N}} p_1 + \dots + e^{-\frac{i2\pi \cdot 2 \cdot (n-1)}{N}} p_{n-1}$$

$$F_3 = e^{-\frac{i2\pi \cdot 3 \cdot 0}{N}} p_0 + e^{-\frac{i2\pi \cdot 3 \cdot 1}{N}} p_1 + \dots + e^{-\frac{i2\pi \cdot 3 \cdot (n-1)}{N}} p_{n-1}$$

⋮

$$F_{N-1} = e^{-\frac{i2\pi \cdot (N-1) \cdot 0}{N}} p_0 + e^{-\frac{i2\pi \cdot (N-1) \cdot 1}{N}} p_1 + \dots + e^{-\frac{i2\pi \cdot (N-1) \cdot (n-1)}{N}} p_{n-1}$$

Bei genauer Betrachtung fällt auf, dass die Exponentialfunktionen (Basisfunktionen) und die Koeffizienten als Skalarprodukt miteinander verknüpft sind. Das Skalarprodukt entspricht genau einer Multiplikation eines Zeilenvektors mit einem Spaltenvektor (siehe Skalarprodukt 2.3.5). Daher kann dieser Sachverhalt auch ausgedrückt werden als:

$$\begin{pmatrix} F_0 \\ F_1 \\ F_2 \\ F_3 \\ \vdots \\ F_{N-1} \end{pmatrix} = \begin{pmatrix} e^{-\frac{i2\pi \cdot 0 \cdot 0}{N}} & e^{-\frac{i2\pi \cdot 0 \cdot 1}{N}} & \dots & e^{-\frac{i2\pi \cdot 0 \cdot (n-1)}{N}} \\ e^{-\frac{i2\pi \cdot 1 \cdot 0}{N}} & e^{-\frac{i2\pi \cdot 1 \cdot 1}{N}} & \dots & e^{-\frac{i2\pi \cdot 1 \cdot (n-1)}{N}} \\ e^{-\frac{i2\pi \cdot 2 \cdot 0}{N}} & e^{-\frac{i2\pi \cdot 2 \cdot 1}{N}} & \dots & e^{-\frac{i2\pi \cdot 2 \cdot (n-1)}{N}} \\ e^{-\frac{i2\pi \cdot 3 \cdot 0}{N}} & e^{-\frac{i2\pi \cdot 3 \cdot 1}{N}} & \dots & e^{-\frac{i2\pi \cdot 3 \cdot (n-1)}{N}} \\ \vdots & \vdots & \ddots & \vdots \\ e^{-\frac{i2\pi \cdot (N-1) \cdot 0}{N}} & e^{-\frac{i2\pi \cdot (N-1) \cdot 1}{N}} & \dots & e^{-\frac{i2\pi \cdot (N-1) \cdot (n-1)}{N}} \end{pmatrix} \begin{pmatrix} p_0 \\ p_1 \\ p_2 \\ p_3 \\ \vdots \\ p_{N-1} \end{pmatrix}$$

Diese Matrix kann noch etwas kompakter dargestellt werden:

$$\mathbf{T} = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & \dots & 1 \\ 1 & \omega^1 & \omega^2 & \omega^3 & \omega^4 & \dots & \omega^{N-1} \\ 1 & \omega^2 & \omega^4 & \omega^6 & \omega^8 & \dots & \omega^{2(N-1)} \\ 1 & \omega^3 & \omega^6 & \omega^9 & \omega^{12} & \dots & \omega^{3(N-1)} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \omega^{N-1} & \omega^{2(N-1)} & \omega^{3(N-1)} & \omega^{4(N-1)} & \dots & \omega^{(N-1)(N-1)} \end{pmatrix}$$

Hier wurde $\omega = e^{-\frac{i2\pi}{N}}$ gesetzt. Der Exponent über ω entspricht der Multiplikation von $n \cdot k$. Diese Matrix ist die sogenannte Vandermonde Matrix [12]. Für ω^2 ergibt sich zum Beispiel $(e^{-\frac{2\pi i}{N}})^2 = e^{-\frac{4\pi i}{N}}$.

Die im jeweiligen Anwendungsfall berechnete Matrix (hier für $n = 8$) wird als DFT-Matrix bezeichnet:

$$\mathbf{T} = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & \omega & -i & -i\omega & -1 & -\omega & i & i\omega \\ 1 & -i & -1 & i & 1 & -i & -1 & \omega \\ 1 & i\omega & i & \omega & -1 & i\omega & -i & -\omega \\ 1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 \\ 1 & -\omega & -i & i\omega & -1 & \omega & i & -i\omega \\ 1 & i & -1 & -i & 1 & i & -1 & -i \\ 1 & i\omega & i & -\omega & -1 & -i\omega & -i & \omega \end{pmatrix}$$

Es gilt: $\omega = e^{-\frac{i2\pi}{8}} = \frac{1}{\sqrt{2}} - \frac{i}{\sqrt{2}}$.

Da sich das Argument der e -Funktion immerzu anders berechnet, nimmt ω stets andere Werte an.

Eine genaue Betrachtung der Vandermonde Matrix lässt feststellen, dass ihr Inverses gegeben ist als:

$$\mathbf{T}^{-1} = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & \dots & 1 \\ 1 & \omega^{-1} & \omega^{-2} & \omega^{-3} & \omega^{-4} & \dots & \omega^{-(N-1)} \\ 1 & \omega^{-2} & \omega^{-4} & \omega^{-6} & \omega^{-8} & \dots & \omega^{-2(N-1)} \\ 1 & \omega^{-3} & \omega^{-6} & \omega^{-9} & \omega^{-12} & \dots & \omega^{-3(N-1)} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \omega^{-(N-1)} & \omega^{-2(N-1)} & \omega^{-3(N-1)} & \omega^{-4(N-1)} & \dots & \omega^{-(N-1)(N-1)} \end{pmatrix}$$

Daraus ergibt sich analog die Formel der inversen diskreten Fouriertransformation [AHK⁺09]:

$$p_n = \frac{1}{N} \sum_{k=0}^{N-1} (e^{\frac{ikn2\pi}{N}}) F_k$$

Der Faktor $\frac{1}{N}$ dient der Mittelung der Werte und wird an anderer Stelle erläutert (siehe Kapite 3.1.5).

Die komplexe Fourierreihe wurde als ein Werkzeug vorgestellt, das es erlaubt, Funktionen in Schwingungen zu zerlegen. Liegt das Interesse in der Information jeder einzelnen vorkommenden Schwingung, wird dies als Fouriertransformation bezeichnet, die explizit jede Schwingung mit dem zu transformierenden Vektor verknüpft und die Information dieser Multiplikation als komplexe Zahl abbildet. Dieser Sachverhalt kann als lineare Transformation ausgedrückt werden. Die Matrix dieser Transformation wird als Vandermonde Matrix bezeichnet. Das Inverse dieser Matrix ergibt sich im Wesentlichen durch den Vorzeichenwechsel des Exponenten jedes Eintrags. Ist von den konkret berechneten Werten im Anwendungsfall der DFT die Rede, so wird dies als DFT-Matrix bezeichnet.

3.1.4 Interpretation der Zerlegung einer Bildspalte

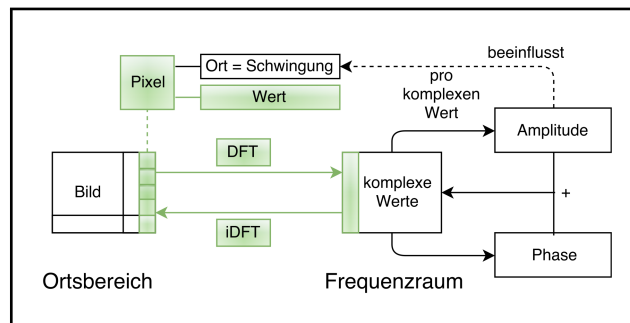


Abbildung 3.8: Übersicht

Durch die Fourierreihe werden periodische Funktionen in Schwingungen zerlegt, oder Funktionen durch ihre enthaltenden Schwingungen approximiert. Im Folgenden wird dies anhand der Bildverarbeitung interpretiert. Gegeben seien die Pixelwerte einer Bildspalte: 12, 36, 196, 220, 22, 64, 132 und 17. Diese Pixelwerte in ein Koordinatensystem übertragen, ergeben folgende Abbildung 3.9:

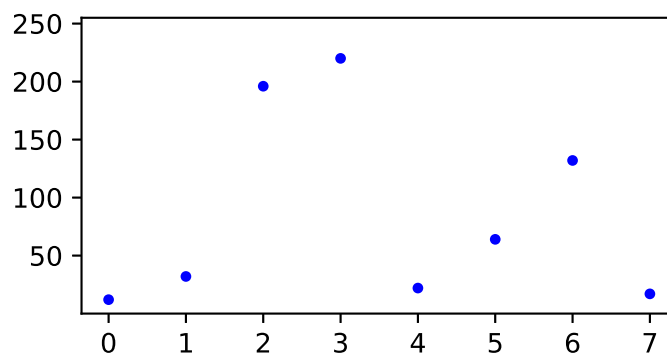


Abbildung 3.9: Eingetragene Pixelwerte

Es existiert eine Funktion $f(x)$, die alle Pixelwerte beinhaltet und die durch die Fourierreihe eindeutig ermittelt werden kann. Es kann also bei mehrfacher Zerlegung nicht passieren, dass unterschiedliche Werte entstehen.

Eine Funktion kann daher dadurch beschrieben werden, dass sie genügend abgetastet wird [13]. Die grüne Funktion der Abbildung 3.9 entspricht der Grundschwingung. Es handelt sich hierbei um eine bekannte Funktion. Im Falle des Pixelvektors der Bildspalte ist es genau umgekehrt: Es handelt sich um eine abgetastete Funktion, aber die dahinterliegende Funktion ist nicht bekannt. Es ist nur klar, dass diese Funktion durch die

Fourierreihe eindeutig zerlegt wird. In ein Koordinatensystem eingetragen, werden die Schwingungsfunktion und die Pixelwerte wie folgt dargestellt (siehe Abbildung 3.10):

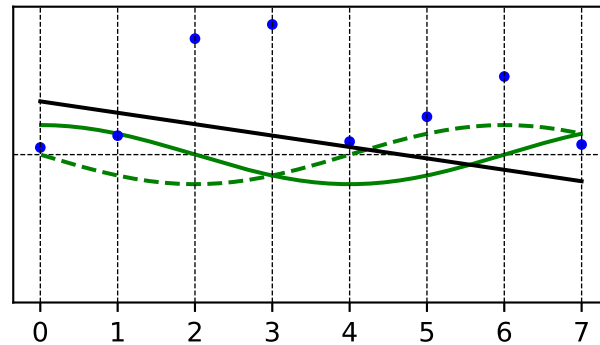


Abbildung 3.10: Korrelation zwischen Basisfunktion und Pixelwerten

Die abgetasteten Werte der e -Funktion, also der additiven Verknüpfung der Sinus- und Kosinusfunktion, die in der Vandermonde Matrix die Zeilenvektoren bilden, sind an den gleichen x -Werten zu finden, wie die Pixelwerte selbst. Wie im Kapitel 2.3.5 erläutert, gibt das Skalarprodukt an, wie ähnlich die Pixelwerte zur gegebenen abgetasteten Schwingung sind. Diese sogenannte Korrelation zwischen den Pixelwerten und der Schwingung entspricht in der Abbildung 3.10 der schwarzen Geraden. Eine negative Steigung, wie hier gegeben, deutet darauf hin, dass diese Schwingung nicht gut mit den Pixelwerten korreliert. Anders formuliert könnte gesagt werden, dass die Schwingung die Pixelwerte nicht gut beschreibt. Das heißt konkret, wenn die Vandermonde Matrix mit einem Vektor multipliziert wird, kann das als Ähnlichkeit des Zeilenvektors zum Spaltenvektor interpretiert werden. Dadurch lässt sich die Matrix-Vektor-Multiplikation als Ähnlichkeitsermittlung zwischen den Zeilenvektoren und dem Spaltenvektor auffassen. Diese Ähnlichkeit wird als komplexe Zahl ausgedrückt.

3.1.5 Beispiel: Fouriertransformation am Spaltenvektor

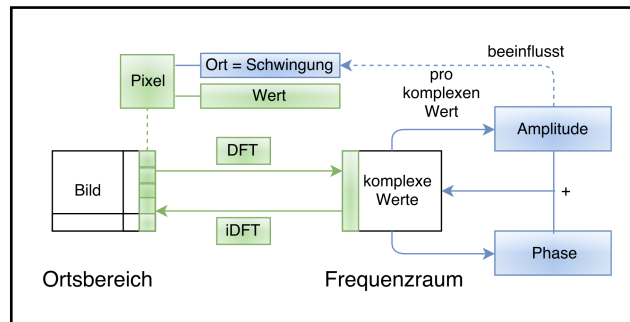


Abbildung 3.11: Übersicht

Da F_0 ein Sonderfall ist, wird dieser separat betrachtet und zuerst auf F_1 eingegangen. Für die Pixelwerte 12, 36, 196, 220, 22, 64, 132 und 17 und die Ähnlichkeit zur Grundschwingung ergibt sich gemäß der Fouriertransformation:

$$\begin{aligned}
 F_1 &= 12 \cdot e^{-\frac{i2\pi \cdot 1 \cdot 0}{8}} + 36 \cdot e^{-\frac{i2\pi \cdot 1 \cdot 1}{8}} + \dots + 17 \cdot e^{-\frac{i2\pi \cdot 1 \cdot 7}{8}} \\
 &= 12 \cdot [\cos(\frac{0}{8}) + i \sin(\frac{0}{8})] + 36 \cdot [\cos(\frac{2\pi}{8}) + i \sin(\frac{2\pi}{8})] + \dots + \\
 &\quad 17 \cdot [\cos(\frac{14\pi}{8}) + i \sin(\frac{14\pi}{8})] \\
 &= 12 \cdot 1 + 36 \cdot (0.707 + 0.707i) + \dots + 17 \cdot (0.707 - 0.707i) \\
 &\approx -173.34 - 187.74i
 \end{aligned}$$

Bei dieser Berechnung wurde ausgenutzt, dass sich die komplexe Exponentialfunktion umschreiben lässt. Das Argument der Exponentialfunktion wird zum Argument der Sinus- beziehungsweise Kosinusfunktion. Für $k = 1$ wurde im Grunde untersucht, wie viel der Grundschwingung in jedem Samplewert steckt. Das Ergebnis dieser Analyse liefert die komplexe Zahl $-173.34 - 187.74i$ aus der dann die Amplitude berechnet wird:

$$|z| = |-173.34 - 187.74i| = \sqrt{(-173.34)^2 + (-187.74)^2} \approx 255.52$$

Dies ist die Gewichtung der Grundschwingung.

Im Hinblick auf die beschriebene Fouriertransformation, ist der Ablauf folgender:

1. p_n steht für die einzelnen Pixelwerte.

2. Über die Fourierreihe wird berechnet, wie viel von der Grundschwingung oder den Oberschwingungen in jedem Pixelwert p_n steckt. $k = 1$ bedeutet daher, dass die p_n für die Grundschwingung untersucht werden. $k = 2$ bedeutet analog, dass die p_n für die erste Harmonische untersucht werden.
3. Der komplexe Zahlenwert, der dann erhalten wird, beinhaltet die Amplitude der untersuchten Oberschwingung.
4. Die Amplitude berechnet sich aus: $|z| = |a + bi| = \sqrt{a^2 + b^2}$
5. Diese Analyse wird für jedes Pixel durchgeführt.

Wie angemerkt, ist F_0 ein Sonderfall:

$$\begin{aligned}
 F_0 &= 12 \cdot e^{-\frac{i2\pi \cdot 0 \cdot 0}{8}} + 36 \cdot e^{-\frac{i2\pi \cdot 0 \cdot 1}{8}} + \dots + 17 \cdot e^{-\frac{i2\pi \cdot 0 \cdot 7}{8}} \\
 &= 12 \cdot [\cos(0) + i \sin(0)] + 36 \cdot [\cos(0) + i \sin(0)] + \dots + 17 \cdot [\cos(0) + i \sin(0)] \\
 &= 12 \cdot 1 + 36 \cdot 1 + \dots + 17 \cdot 1 \\
 &= 699
 \end{aligned}$$

Was hier berechnet wurde, ist der sogenannten Gleichanteil. In ihm steckt die Information der durchschnittlichen Helligkeit der Pixelwerte. Das kommt daher, dass nur die einzelnen Pixelwerte aufaddiert werden. Wird dieser Wert durch N geteilt ergibt sich 87,375, das arithmetische Mittel der Pixelwerte. Der Faktor $\frac{1}{N}$ der inversen Fouriertransformation bildet demnach das Mittel der Pixelwerte, was bei der Rücktransformation benötigt wird. Für die anderen Samples sind die Berechnungen analog. Das Amplitudenspektrum, also die konkrete Darstellung aller berechneten Amplitudenwerte der komplexen Zahlen F_k , sieht wie in Abbildung 3.12 aus:

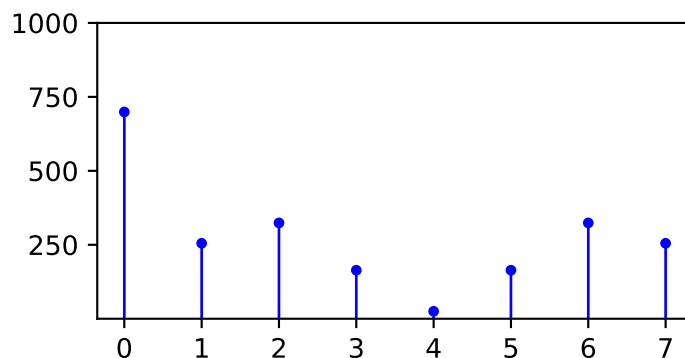


Abbildung 3.12: Amplitudenspektrum der Pixelwerte

Auf der x -Achse befinden sich die Einträge für alle F_k und auf der y -Achse ist die Amplitudenhöhe festgelegt. Für F_1 wurde der Wert $173.34 - 187.74i$ ermittelt, aus dem wiederum die Amplitude 255.52 ermittelt wurde. Um die inverse Fouriertransformation anwenden zu können, ist es also notwendig, aus der Amplitude die kartesische komplexe Zahl zu ermitteln. Dies ist jedoch nicht möglich. An dieser Stelle wurde die Tatsache unterschlagen, dass eine komplexe Zahl erst dann eindeutig bestimmt ist, wenn sowohl die Phase als auch die Amplitude bekannt sind. Die Phase für F_1 entspricht:

$$\alpha = \arctan(-187.74/173.34) \approx 312.7^\circ$$

Jede komplexe Zahl ist über ihre Polarkoordinaten darstellbar,

$$z = a \cdot \cos(\alpha) + a \cdot i \sin(\alpha)$$

wobei a der Amplitudenwert ist und α der Phase entspricht. Eingesetzt folgt hieraus:

$$z = 255.52 \cdot \cos(312.7^\circ) + 255.52 \cdot i \sin(312.7^\circ) = 173.34 - 187.74i$$

Dies entspricht exakt dem Wert von F_1 . Bei einer Senkung der Amplitude auf 200 würde F_1 einen anderen Wert annehmen.

$$z = 200 \cdot \cos(177.716^\circ) + 200 \cdot i \sin(312.773^\circ) \approx 135.82 - 146.81i$$

Dies ist faktisch, was bei der Fouriertransformation ausgenutzt wird. Diese Änderung kann für alle F_k durchgeführt werden. Dadurch ergeben sich die neuen beziehungsweise ursprünglichen komplexen Werte.

$F_0: 699 + 0i$	$F_4: 25 + 0i$
$F_1: 135.82 - 146.81i$	$F_5: 153.34 + 59.74i$
$F_2: 294 + 137i$	$F_6: 294 - 137i$
$F_3: 153.34 - 59.74i$	$F_7: 135.82 + 146.81i$

F_1 und F_7 sind gemäß des neuen Ergebnisses abgeändert. Es ist auffällig, dass sich die komplexen Werte ab einem gewissen Abstand wiederholen, wobei beim Imaginärteil ein Vorzeichenwechsel stattfindet. Das liegt an der Periodizität der Funktionen, die verwendet werden, und führt zu einer Symmetrie, die an anderer Stelle ausgenutzt werden wird. Da die Berechnung der inversen DFT sehr ähnlich zur normalen Fouriertransformation ist, wird hier nur der Ansatz für p_1 gegeben:

$$p_1 = 699 \cdot e^{\frac{i2\pi \cdot 1 \cdot 0}{8}} + (135.82 - 146.81) \cdot e^{\frac{i2\pi \cdot 1 \cdot 1}{8}} + \dots = 83.42$$

Im Folgenden werden die Werte für alle p_n aufgelistet:

p_0 : 236.29	p_4 : 91.70
p_1 : 83.42	p_5 : 16.58
p_2 : 38.76	p_6 : -4.76
p_3 : 158.11	p_7 : 78.88

Abbildung 3.13 zeigt die neuen eingetragenen Werte. Zur Erinnerung: p_n waren die ursprünglichen Pixelwerte der Folge 17, 132, 64, 22, 220, 196, 36 und 12 (blau). Obwohl nur F_1 in seiner Amplitude geändert wurde, haben sich alle anderen Pixelwerte (grün) geändert. Diese Änderungen sind daher groß, da F_1 die Information der Grundschwingung enthält, also jener Schwingung, welche die Pixelwerte am besten approximiert. Da diese Werte nicht alle im Intervall von $[0, 255]$ liegen, werden sie betragsmäßig angepasst:

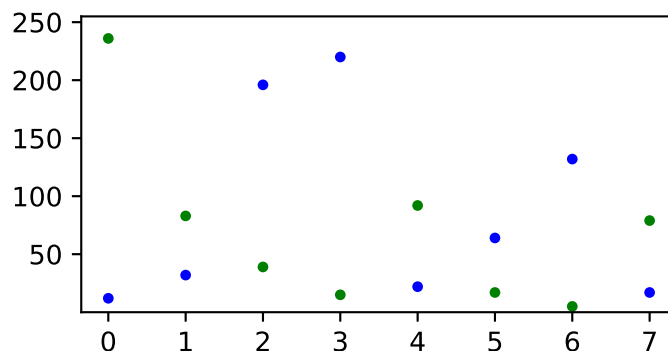


Abbildung 3.13: Vergleich der alten und neuen Pixelwerte

Es wurde ein Zeilenvektor der Vandermonde Matrix mit dem Bildspaltenvektor multipliziert. Als Resultat ergibt sich eine komplexe Zahl, die die Information dieser Multiplikation enthält. Diese Information kann als Ähnlichkeit zwischen dem Zeilen- und dem Spaltenvektor interpretiert werden. Die komplexe Zahl ist in ihrer kartesischen Form nicht nützlich. Daher wird aus ihr die Amplitude berechnet, die dann verändert werden kann. Die Änderung einer Amplitude beeinflusst alle Pixelwerte. Die Abbildung über alle Amplitudenwerte nennt man Amplitudenspektrum. Zur Rücktransformation werden beide Informationen, Phase und Amplitude, der komplexen Zahl benötigt. Abänderungen tiefer Frequenzen (zum Beispiel der Grundschwingung) haben große Auswirkungen auf die rücktransformierten Werte.

3.1.6 2D-DFT im Zusammenhang mit der Bildverarbeitung

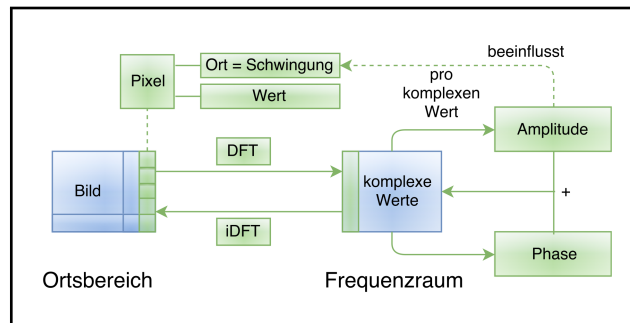


Abbildung 3.14: Übersicht

In diesem Kapitel wird das angeeignete Wissen auf Bilder übertragen. Da Bilder zweidimensional sind, müssen Überlegungen der erläuterten Sachverhalte angepasst werden. Im eindimensionalen Fall wird eine Funktion durch Schwingungen dargestellt und in diese zerlegt. Dies verhält sich im zweidimensionalen Fall nicht anders. Ein Bild wird als Funktion interpretiert, die in ihre Basisfunktionen zerlegt werden kann. Für ein 4×4 -Bild wären dies Kombinationen dieser Funktionen (siehe Abbildung 3.15):

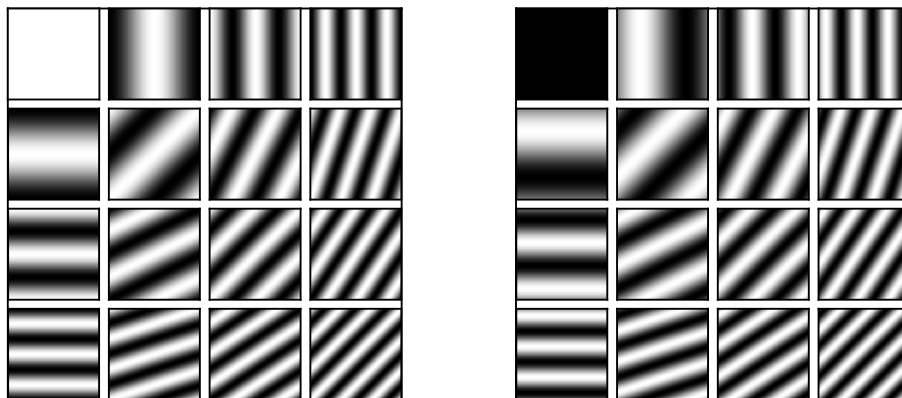


Abbildung 3.15: v.l.n.r. Kosinusbasisfunktionen; Sinusbasisfunktionen

Die Schwingungen werden aus der Vogelperspektive betrachtet. Schwarz bedeutet, dass sich die Schwingung am unteren Ende ihres Wertebereichs befindet, für weiß analog am oberen Ende. In der ersten Reihe und Spalte wird deutlich, was im Prinzip im Beispiel 3.1.5 berechnet wurde. Von oben nach unten beziehungsweise links nach rechts nimmt die Frequenz der Schwingungen zu. In der Diagonalen liegen die Schwingungen,

die zwischen den anderen korrelieren. Für den Pixel $p_{1,1}$ bedeutet das, dass die dort abgebildete Schwingung genau zwischen $p_{0,1}$ und $p_{1,0}$ liegt und daher um 45° rotiert ist. Die Amplitude des berechneten F_k gibt nun genau wieder, zu welchem Anteil die jeweiligen Schwingungen (in diesem Fall die einzelnen Bilder) vertreten sind. Der Pixelort gibt also nicht nur den Wert des Pixels an, sondern auch die Schwingung, zu der die Amplitude berechnet wird. Würde ein größeres Bild untersucht werden, so würde, wie hier bereits zu sehen, für jede weitere Basisfunktion die Frequenz um eins zunehmen.

Als nächstes wurde überlegt, wie einzelne Vektoren transformiert werden können. Auf ein Bild angewandt, ändert sich dank Linearität bei der Berechnung nicht viel:

$$f(x) = \sum_{n=0}^{N-1} \left(\sum_{m=0}^{M-1} (e^{-\frac{ilm2\pi}{N}}) p_{n,m} \right) e^{-\frac{ikn2\pi}{N}}$$

Die Formel besagt nichts anderes als dass für jeden komplexen Wert die Fouriertransformation einmal auf jeden Spaltenvektor des Bildes und anschließend auf jeden Zeilenvektor anzuwenden ist. Durch weitere Überlegungen ist festzustellen, dass sich die 2D-DFT wie folgt berechnen lässt:

1. Multipliziere das Bild von links mit der DFT-Matrix.
2. Invertiere die DFT-Matrix.
3. Multipliziere die invertierte DFT-Matrix von rechts an das Bild.

Das führt dazu, dass jede Basisfunktion auf ihren komplexen Wert $F_{k,l}$ abgebildet wird. Für die erste Oberschwingung in horizontaler Richtung kann dies wie folgt veranschaulicht werden 3.16:

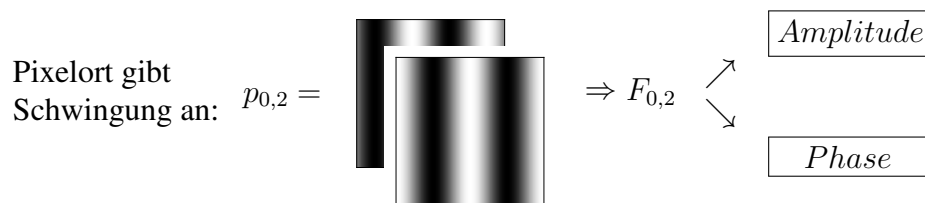


Abbildung 3.16: Pixelort zeigt welche Frequenz untersucht wird.

In Abbildung 3.16 wird noch einmal der Zusammenhang zwischen der Aufspaltung einer reellen Zahl in einen Real- und Imaginärteil im Eindimensionalen und der Aufspaltung eines Bildes in zwei Bilder im Zweidimensionalen deutlich. Das Bild, das die Information der Phase beinhaltet, wird Phasenbild genannt. Es drückt aus „wo“ sich etwas im Ursprungsbild befunden hat. Das Amplitudenbild drückt aus, zu welchem Anteil

die Frequenzen dort vorkommen. Beide Bilder werden benötigt, um das ursprüngliche Bild wiederherzustellen. Dies verläuft analog wie zum Eindimensionalen. Das Phasenbild wird normalerweise nicht bearbeitet.

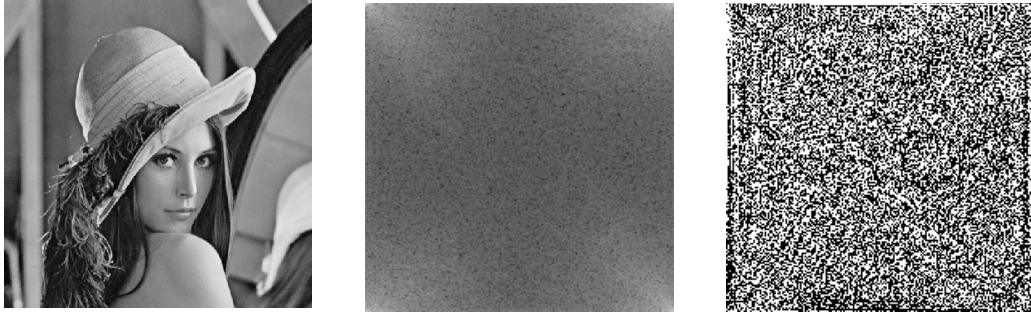


Abbildung 3.17: v.l.n.r. Lenna; Amplitudenspektrum; Phasenbild

Auf Grund von vertikal und horizontal auftretenden Symmetrien, wie sie im Berechnungsbeispiel (Kapitel 3.1.5) erläutert wurden, ist das Amplitudenspektrum in vier Quadranten aufgeteilt, von denen jeder die gleichen Informationen enthält, wie in Abbildung 3.17 zu sehen ist. Das Amplitudenspektrum wird ebenfalls aus der Vogelperspektive betrachtet. Jedes Pixel enthält demnach die Amplitude einer bestimmten Schwingung. Wie aus Abbildung 3.15 zu erkennen ist, befinden sich die niedrigsten Frequenzen in den Ecken dieser Quadranten. Konventionsgemäß werden diese in die Mitte des Bildes gedreht, sodass eine einfachere Bearbeitung möglich ist (siehe Abbildung 3.18):

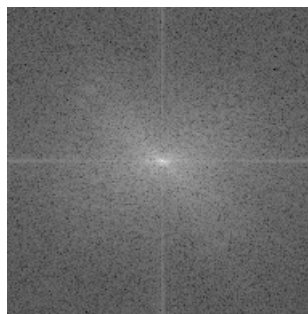


Abbildung 3.18: Amplitudenspektrum (tiefe Frequenzen in die Mitte gedreht)

Werden jeweils nur das Amplitudenspektrum beziehungsweise das Phasenspektrum zurücktransformiert, ergeben sich die zwei Bilder der Abbildung 3.19:

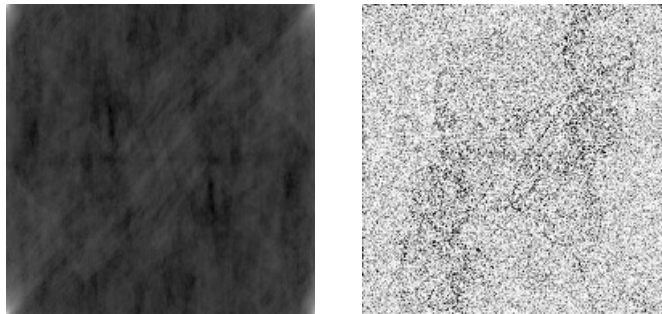


Abbildung 3.19: v.l.n.r. zurücktrans. Amplitudenspektrum; zurücktrans. Phasenbild

Beide Bilder sind alleine nicht nützlich. Im Phasenbild lassen sich zumindest die Umrisse von Lenna erahnen. Dies verdeutlicht, dass zur Rücktransformation sowohl Phase als auch Amplitude benötigt werden.

Es wurde festgestellt, dass sich die DFT auf Bilder anwenden lässt. Dazu wird jedes Bild in zweidimensionale Basisfunktionen zerlegt, die aus Sinus- und Kosinuswellen zusammengesetzt sind. Ein Bild wird durch die Transformation zweikanalig, spaltet sich also in sein Amplitudenspektrum und Phasenspektrum auf, wovon meistens das Amplitudenspektrum bearbeitet wird. Für die Rücktransformation sind beide Bilder erforderlich.

3.1.7 Anwendungen der DFT auf Bilder

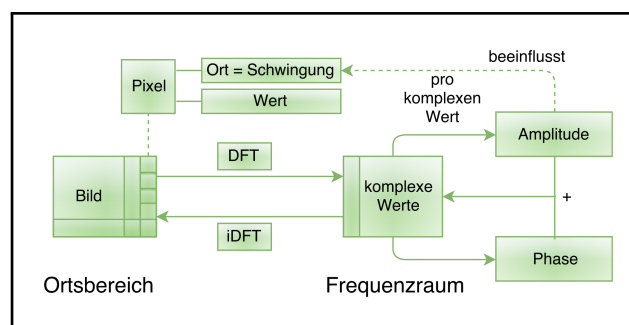


Abbildung 3.20: Übersicht

Als Motivation der Fouriertransformation wurde gezeigt, dass die DFT auf Bilder angewandt zu einer Glättung beziehungsweise Kantendetektion führen kann (siehe Kapitel 3.1). Eine Glättung kann im Ortsbereich durch die Anwendung eines Gaußfilters herbeigeführt werden. Die Glättung auf ein transformiertes Bild entspricht der Anwendung

des Ergebnisses des transformierten Gaußfilters auf das Amplitudenbild, wie den Abbildungen 3.21 und 3.22 zu entnehmen ist:

$$\begin{pmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{pmatrix} \xrightarrow{\text{DFT}} \text{Bild}$$


Abbildung 3.21: Transformierter Gaußfilter

Das Ergebnis der Abbildung 3.21 kann als Maske interpretiert werden, die, über das Amplitudenspektrum gelegt, die Amplituden tiefer Frequenzen mit dem Faktor 1 gewichtet und die Amplituden hoher Frequenzen mit dem Faktor 0. Dies wird ebenfalls als Tiefpassfilter bezeichnet.

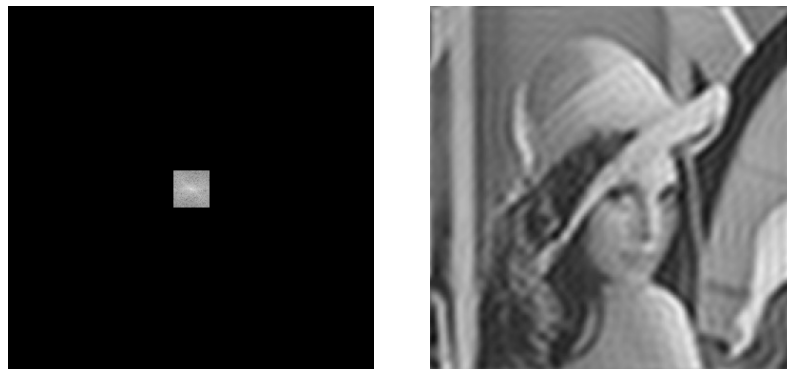
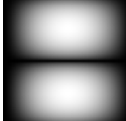


Abbildung 3.22: v.l.n.r. verändertes Amplitudenspektrum; Rücktransformation (Glättung)

Wie im Berechnungsbeispiel 3.1.5 gesehen, führt eine Veränderung der Amplitude zu einer Änderung der anderen Pixelwerte. Dieser Zusammenhang gilt ebenfalls umgekehrt. Amplituden, die außerhalb des Zentrums des Amplitudenspektrums liegen, führen zu kleinen Änderungen im Bild. Dies bedeutet, dass die tiefen Frequenzen den größten Informationsgehalt des Bildes mit sich tragen. Das kommt der Annäherung der Rechteckfunktion aus Abbildung 3.3 gleich. Ohne die erste Oberschwingung kann der Großteil der Rechteckschwingung nicht rekonstruiert werden. Analog dazu werden feinere Änderungen des Signals mit kleinen Frequenzen angeglichen. Auf das Bild 3.22

bezogen bedeutet das, dass die Details beim Löschen der Amplituden der hohen Frequenzen aus dem Bild entfernt werden. Da die Amplitude mit 0 multipliziert wird, ist die dazugehörige Schwingung nicht mehr vorhanden. Analog ergibt sich für die beiden Sobelmasken Abbildung 3.23:

$$\begin{pmatrix} 1 & 0 & -1 \\ 2 & 0 & -2 \\ 1 & 0 & -1 \end{pmatrix} \xrightarrow{\text{DFT}} \text{Bild}$$


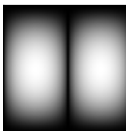
$$\begin{pmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{pmatrix} \xrightarrow{\text{DFT}} \text{Bild}$$


Abbildung 3.23: Transformierte Sobelfilter

Beide Filter haben in ihrer transformierten Variante gemeinsam, dass die Amplituden des Zentrums nicht berücksichtigt werden. Daraus ergibt sich Abbildung 3.24:

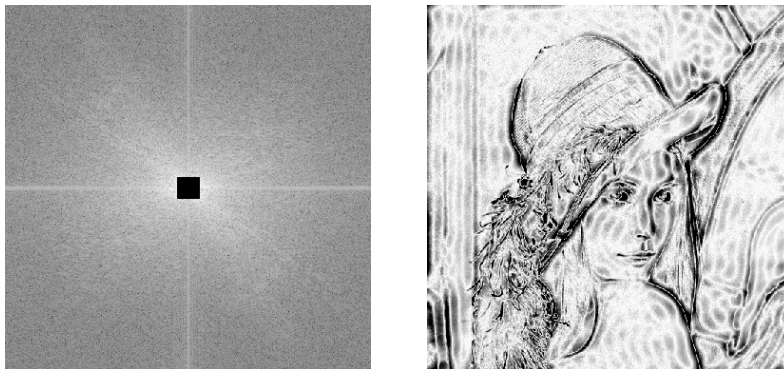


Abbildung 3.24: v.l.n.r. verändertes Amplitudenspektrum; Kantendetektion (invertiert)

Die Kantendetektion entspricht dem Löschen der Amplituden der tiefen Frequenzen, was ebenfalls als Hochpassfilter bezeichnet wird. Der Großteil der Informationen des Bildes geht hierbei verloren, sodass nur noch die Überlagerungen der hohen Frequenzen bleiben. Diese Überlagerungen sind die Kanten. Wie erwähnt kann das Löschen der Amplituden als Multiplikation mit 0 und 1 interpretiert werden. Daher entspricht

die Faltung im Ortsraum einer Multiplikation im Frequenzraum und dies entspricht der Aussage des Faltungstheorems [Fol92], welches in Abbildung 3.25 ([10] Quelle: Eigene Darstellung) dargestellt ist:

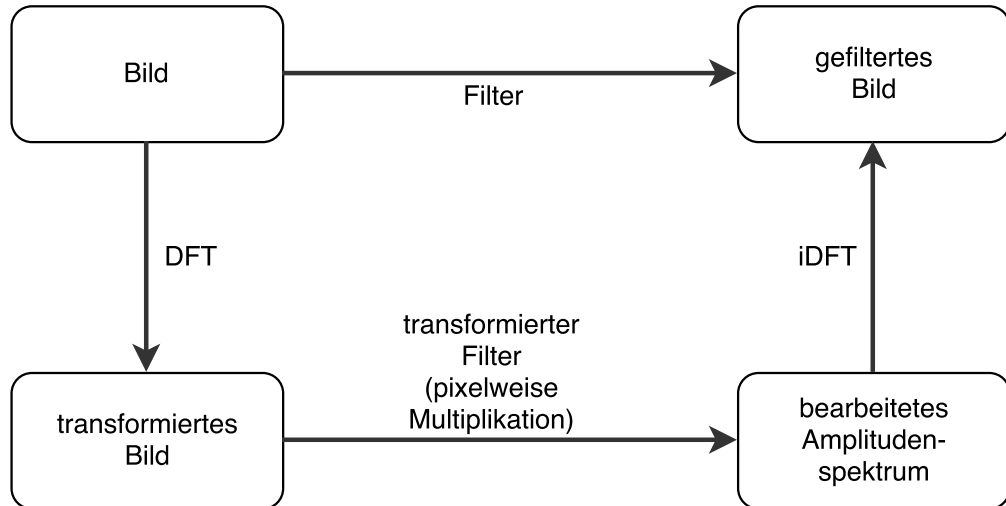


Abbildung 3.25: Veranschaulichung des Faltungstheorems

Wie bereits erläutert, wird beim Löschen der tiefen Frequenzen im Frequenzraum eine Kantendetektion im Ortsbereich erzeugt. In Abbildung 3.26 ist die Holzverkleidung einer Wand zu erkennen:



Abbildung 3.26: Holzverkleidung

Die Grundaussage ist, dass das Phasenbild die Ortsinformationen des transformierten Bildes enthält. Es soll versucht werden, die schwarzen Streifen des Bildes nur mit Hilfe des Amplitudenspektrums verschwinden zu lassen. Angenommen das Bild werde horizontal abgelaufen und die Punkte zu einem gegebenen Zeitpunkt in ein Koordinatensystem eingetragen, so wäre zu erwarten, dass für jede Bildzeile weitestgehend eine konstante Funktion entsteht. Wird das Bild vertikal abgelaufen, folgen auf helle Flächen

dunkle Flächen. Vereinfacht könnte angenommen werden, dass von weiß zu schwarz gesprungen wird und umgekehrt. Diese Sprünge wurden im eindimensionalen Fall als Rechteckfunktionen vorgestellt, die über Sinusfunktionen unterschiedlicher Amplitude angenähert wurden, wie Abbildung 3.27 zeigt:

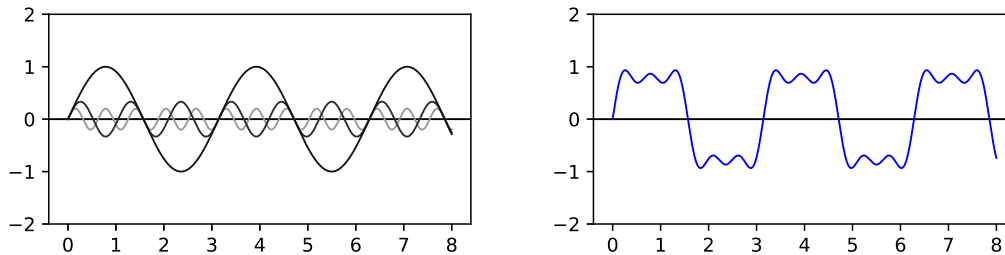


Abbildung 3.27: v.l.n.r. Schwingungen; addierte Schwingungen

Die Approximation des Rechtecksignals ist nur gegeben, da die Schwingung mit der höchsten Amplitude den Großteil des Signals approximiert. Wenn diese Schwingung nicht existiert, kann das Signal schlechter angenähert werden, wie Abbildung 3.28 verdeutlicht:

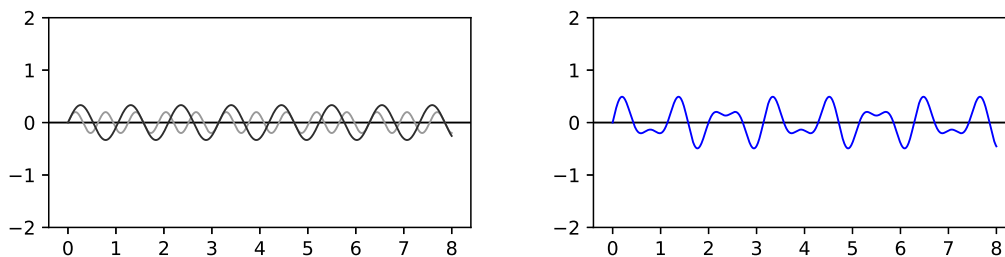


Abbildung 3.28: v.l.n.r. Schwingungen; addierte Schwingungen

Um im Bild die schwarzen Streifen zu entfernen, müssen demnach die Basisfunktionen gefunden und gelöscht werden, die das dem Bild unterliegende zweidimensionale Rechtecksignal am besten annähern. Da diese Rechtecksignale, wie im Bild auch, im Amplitudenspektrum horizontal verlaufen müssen, werden diese entfernt (siehe Abbildung 3.29).

Das Ergebnis bestätigt die Annahme nicht eindeutig. Die invertierten schwarzen Streifen sind noch deutlich zu erkennen. Werden die Frequenzen großzügiger gelöscht, ergibt sich Abbildung 3.30.

Hier scheinen die ausschlaggebenden Frequenzen des Amplitudenspektrums besser entfernt worden zu sein. Die Methodik, die angewandt wurde, sorgt nicht dafür, dass

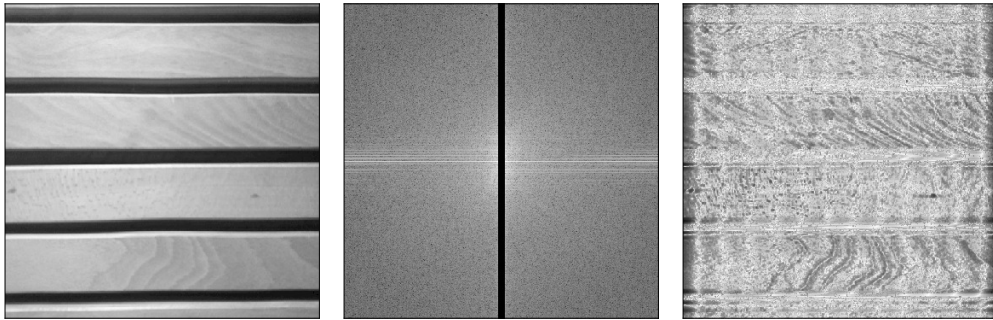


Abbildung 3.29: v.l.n.r. Holzverkleidung; bearbeitetes Amplitudenspektrum; inv. Ergebnis

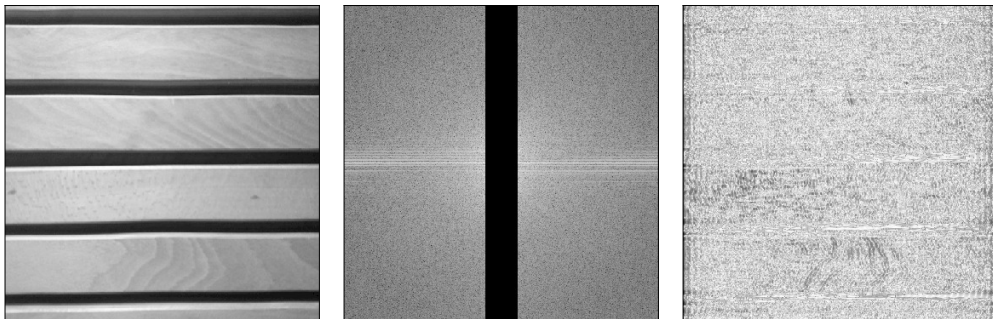


Abbildung 3.30: v.l.n.r. Holzverkleidung; bearbeitetes Amplitudenspektrum; inv. Ergebnis

nur die schwarzen Streifen verschwinden, sondern viel mehr dafür, dass die Übergänge von schwarz nach weiß und umgekehrt nicht mehr so hart ausfallen. Das gesamte Bild erscheint homogener. Bei einem geeigneten Bild scheint es möglich, Streifen im Amplitudenspektrum zu finden und diese zu entfernen. Durch diese beiden bearbeiteten Amplitudenspektren fällt auf, dass die Ergebnisbilder nicht gleich hell geblieben sind. Der Grund dafür ist, dass Bearbeitungen des Amplitudenspektrums das ursprüngliche Signal, oder in diesem Fall Bild, abschwächen oder verstärken können. Werden wie im Beispiel Frequenzen entfernt, so wird das Signal des Ergebnisses abgeschwächt, was in der Bildverarbeitung zur Folge hat, dass das Ergebnisbild dunkler wird. Je mehr Frequenzen entfernt werden, desto dunkler wird das Ergebnisbild. Diese Auffälligkeit ist die Interpretation des parsevalschen Theorems [14]:

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} |fou(x)|^2 dx = \sum_{n=-\infty}^{\infty} |p_n|^2$$

Die Funktion $f_{ou}(x)$ bezeichnet hierbei die komplexe Fourierreihe. An einer Bildspalte interpretiert, lässt sich die rechte Seite der Gleichung als Summe des Ortsbereichs über alle quadrierten Pixelwerte verstehen, wobei der Laufindex der Summe n auf Grund der formalen periodischen Fortsetzung der Bildspalte von $-\infty$ bis ∞ läuft. Auf der linken Seite kann das Integralsymbol als die Summe kontinuierlicher Werte verstanden werden. Konkret ausgedrückt: Es werden alle Amplituden der Schwingungen, die in der zu untersuchenden Funktion vorkommen, berechnet, diese jeweils quadriert und deren Summe durch die Integration berechnet. Bekannt ist, dass der gemittelte Wert von F_0 (siehe Kapitel 3.1.5) der Wert der durchschnittlichen Helligkeit des Bildes ist. Die rechte Seite lässt sich ebenfalls auf diese Art interpretieren. Wird ein Pixelwert gelöscht, sinkt demnach die durchschnittliche Helligkeit des Bildes, was sich wiederum auf den Frequenzbereich auswirkt. Ist das Pixel nicht mehr existent, verändert sich ebenfalls die Summe aller Amplitudenwerte, was zu einem Helligkeitsverlust führt. Dies ist auch umgekehrt der Fall.

Die Bildvergrößerung stellt ein weiteres praktisches Anwendungsgebiet der Fouriertransformation dar. Um ein Bild zu vergrößern ist es eine Möglichkeit bilinear zu interpolieren, wodurch dem Bild Informationen hinzugefügt werden, die ursprünglich nicht vorhanden waren. Eine andere Lösung bietet die DFT:



Abbildung 3.31: v.l.n.r. Lenna; Amplitudenspektrum

In Abbildung 3.31 ist eine verkleinerte Abbildung des Beispielbilds und dessen Amplitudenspektrums zu sehen. Die Idee der Bildvergrößerung mittels DFT ist, die Matrix des Amplitudenspektrums mit Nullen aufzufüllen, bis die gewünschte Größe erreicht ist, wie Abbildung 3.32 zeigt:

Das Amplitudenspektrum geht vom Ursprung aus, damit keine neuen Frequenzen in das Bild aufgenommen werden. Da das Amplitudenspektrum so groß wie das Bild selbst ist, resultiert aus der Rücktransformation ein vergrößertes Bild, wie Abbildung 3.33 zu entnehmen ist:

Das Amplitudenspektrum dieses Bildes ist allerdings nicht, wie zu erwarten, gestreckt oder im Bildursprung geblieben (siehe Abbildung 3.34):

Offensichtlich wurde bei der Rücktransformation das Amplitudenspektrum mit dem Informationsgehalt automatisch in die Mitte verschoben. Die Informationen wurden also nicht vermehrt, wie es bei der bilinearen Interpolation der Fall wäre.

Es lässt sich feststellen, dass die fouriertransformierten Sobel- und Gaußfilter Aussagen darüber treffen, wie das Amplitudenspektrum zu verändern ist, um den gleichen Ef-

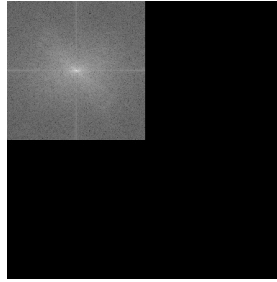


Abbildung 3.32: aufgefülltes Amplitudenspektrum



Abbildung 3.33: Ergebnis nach Rücktransformation

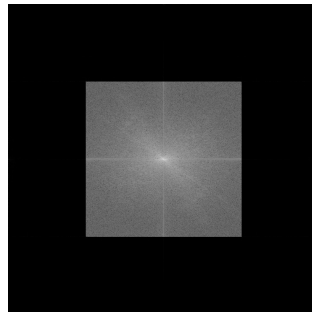


Abbildung 3.34: Ergebnis nach erneuter Transformation

fekt bei der Rücktransformation zu erhalten. Die Erkenntnis führt zum Faltungstheorem, welches besagt, dass Faltungen im Ortsbereich durch Multiplikationen im Frequenzbereich bewirkt werden können. Weiterhin zeigt sich, dass sich harte Kanten eines Bildes in bestimmten Fällen mit Hilfe der Veränderung des Amplitudenspektrums beseitigen lassen. Bei der Rücktransformation dieser Spektren wird deutlich, dass die Bilder einem Helligkeitsverlust unterliegen, was auf das parsevalsche Theorem zurückzuführen ist. Es besagt, dass Auswirkungen im Frequenzbereich, Auswirkungen im Ortsbereich

nach sich ziehen und umgekehrt. Basierend auf der angewandten Methodik kann eine Bildvergrößerung mittels der DFT herbeigeführt werden, die zum Vorteil hat, dass dem Bild keine weiteren Informationen hinzugefügt werden.

3.1.8 Die Fast Fourier Transform (FFT)

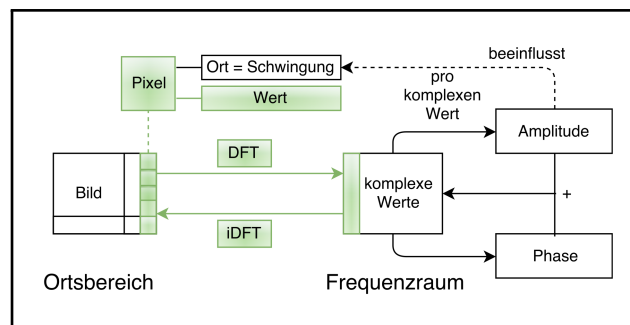


Abbildung 3.35: Übersicht

Die schnelle Fouriertransformation ist ein Algorithmus zur schnellen Berechnung der DFT [Pri15]. Wie bei der Anwendung implizit angedeutet, ist die Berechnung der DFT aufwendig. Für zwei Werte p_1 und p_2 berechnen sich die komplexen Werte aus (wobei exp die Exponentialfunktion an der jeweiligen Stelle beschreibt):

$$F_1 = p_1 \cdot exp_0 + p_2 \cdot exp_1$$

$$F_2 = p_1 \cdot exp_0 + p_2 \cdot exp_1$$

Für einen Vektor mit drei Samplewerten wären dies bereits drei Multiplikationen pro Wert:

$$F_1 = p_1 \cdot exp_0 + p_2 \cdot exp_1 + p_3 \cdot exp_2$$

$$F_2 = p_1 \cdot exp_0 + p_2 \cdot exp_1 + p_3 \cdot exp_2$$

$$F_3 = p_1 \cdot exp_0 + p_2 \cdot exp_1 + p_3 \cdot exp_2$$

Die Anzahl der benötigten Operationen steigt quadratisch, was bei großen Bildern zu langen Laufzeiten führt. Die Idee der FFT ist es, Symmetrien der Transformationsmatrix auszunutzen. Die 4-DFT Matrix ist definiert als:

$$\mathbf{T}_4 = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -i & -1 & i \\ 1 & -1 & 1 & -1 \\ 1 & i & -1 & -i \end{pmatrix}$$

Werden die Werte der geraden und ungeraden Indizes der zweiten und vierten Zeile verglichen, so ist zu erkennen, dass die Summe jeweils zu dem gleichen Ergebnis führt:

$$R_{1_{\text{even}}} = 1 - 1 = -1 + 1 = R_{3_{\text{even}}}$$

$$R_{1_{\text{odd}}} = i - i = -i + i = R_{3_{\text{odd}}}$$

Diese Symmetrie ist durch die Periodizität der Basisfunktionen bedingt, was dazu führt, dass eine Zeile mit jeweils einer anderen ein Pärchen bildet, sodass die Ergebnisse der gebildeten Summen ausgenutzt werden können [Pri15].

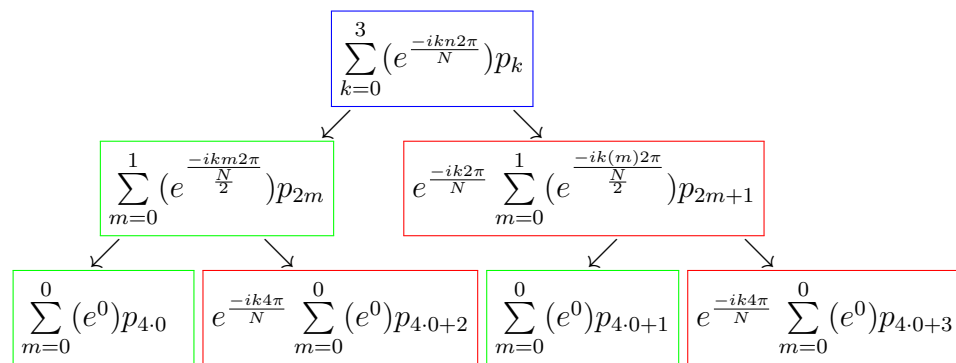
Dies kann ausgedrückt werden indem die Summe der Fourierreihe gemäß gerader und ungerader Indizes aufteilt wird:

$$\sum_{k=0}^{N-1} (e^{\frac{-ikn2\pi}{N}}) p_k = \sum_{m=0}^{\frac{N}{2}-1} (e^{\frac{-ikm2\pi}{N}}) p_{2m} + \sum_{m=0}^{\frac{N}{2}-1} (e^{\frac{-ik(2m+1)2\pi}{N}}) p_{2m+1}$$

Bei der Summe der ungeraden Indizes ist es möglich, einen Teil der e -Funktion auszuklammern und, da dieser Teil in keiner Abhängigkeit zur Summe steht, vor die Summe zu ziehen:

$$\sum_{m=0}^{\frac{N}{2}-1} (e^{\frac{-ik(2m+1)2\pi}{N}}) p_{2m+1} = e^{-\frac{-ik2\pi}{N}} \sum_{m=0}^{\frac{N}{2}-1} (e^{\frac{-ik(m)2\pi}{N}}) p_{2m+1}$$

Diese Umformung ist bei weiteren rekursiven Aufspaltungen der Summen in gerade und ungerade Indizes möglich. Für die Matrix vierer Samplewerte ergibt sich:



Mit rot sind die Summen mit ungeraden Indizes gekennzeichnet und mit grün die mit geraden Indizes. Im ersten Schritt wird für F_0 die Summe der Blätter der Baumstruktur gebildet:

$$F_0 = 4 + 2 + 3 + 1 = 10$$

Aus aufgezeigten Gründen werden die Summen der geraden und ungeraden Indizes explizit gebildet und sich gemerkt:

$$F_{0_{\text{odd}}} = 4 + 2 = 6$$

$$F_{0_{\text{even}}} = 3 + 1 = 4$$

F_1 bildet sich aus der Summe der darüberliegenden Ebene:

$$F_1 = p_0 + e^{-\frac{ik4\pi}{4}}p_2 + e^{-\frac{ik2\pi}{4}}(p_1 + p_3e^{-\frac{ik4\pi}{4}}) = 4 + (-1) \cdot 2 + (-i)(3 + (-1) \cdot 1) = 2 - 2i$$

Erneut werden sich die Werte der geraden und ungeraden Summen gemerkt:

$$F_{1_{\text{odd}}} = 2$$

$$F_{1_{\text{even}}} = 2$$

Aus Symmetriegründen gilt:

$$F_{0_{\text{even}}} = 6 = F_{2_{\text{even}}}$$

$$F_{0_{\text{odd}}} = 4 = F_{2_{\text{odd}}}$$

$$F_{1_{\text{even}}} = 2 = F_{3_{\text{even}}}$$

$$F_{1_{\text{odd}}} = 2 = F_{3_{\text{odd}}}$$

Daraus ergibt sich:

$$F_2 = F_{0_{\text{even}}} + e^{\frac{-ik2\pi}{N}} F_{0_{\text{odd}}} = 6 + (-1) \cdot 4 = 2$$
$$F_3 = F_{1_{\text{even}}} + e^{\frac{-ik2\pi}{N}} F_{1_{\text{odd}}} = 2 + 2i$$

Insgesamt ergeben sich $F_0 = 10$, $F_1 = 2 - 2i$, $F_2 = 2$ und $F_3 = 2 + 2i$.

Es wurde eine Methode vorgestellt, mit der die DFT schnell berechnet werden kann. Die Grundidee dabei ist, die Symmetrien der jeweiligen DFT-Matrix auszunutzen, indem die Summen zur Berechnung der einzelnen Frequenzen rekursiv entsprechend gerader und ungerader Indizes aufteilt werden. Dies hat den Vorteil, dass zuvor berechnete Ergebnisse ausgenutzt werden können und sich die Laufzeit für die Berechnung verkürzt.

3.2 Diskrete Kosinustransformation

In diesem Kapitel wird die diskrete Kosinustransformation (DCT) am Beispiel der Glättung und Kantendetektion als Alternative zur DFT vorgestellt.

3.2.1 Ähnlichkeiten zwischen DCT und DFT

Bei der DFT wurde festgestellt, dass es sich bei den Basisfunktionen der Transformationsmatrix um Exponentialfunktionen handelt, die jeweils als eine Verknüpfung von Sinus- und Kosinusschwingungen ausgedrückt werden können. Wie gezeigt wurde, bilden diese Funktionen, in ihrer diskretisierten Form, die Zeilen dieser Matrix, mit der ein Vektor in den Frequenzraum abgebildet wird. Bei der diskreten Kosinustransformation finden nur noch Kosinusschwingungen Anwendung. Da die Imaginärteile, die Sinusfunktionen, nicht existieren, wird die Darstellung in der Exponentialform redundant.

Analog zu Abbildung 3.6 kann diese Transformation ebenfalls anschaulich dargestellt werden:

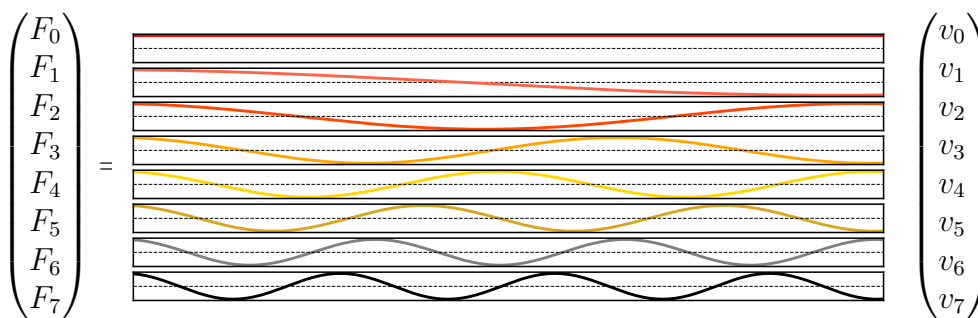


Abbildung 3.36: Transformationsmatrix mit kontinuierlichen Basisfunktionen (DCT)

Wie bekannt, ist in Abbildung 3.36 für jede Zeile der 8×8 -Matrix die Kosinusfunktion dargestellt, die bei der Multiplikation mit dem Vektor v in äquidistanten Schritten diskretisiert wird. Die DCT ist im eindimensionalen Fall gegeben als:

$$F_k = \sum_{n=0}^{N-1} (\cos(\frac{\pi}{N}(nk + \frac{k}{2})))p_n$$

Es wird die durch p_n (Pixelwerte) gewichtete Summe der Kosinusschwingungen gebildet. Diese Linearkombination kann anwendungsspezifisch eine Zerlegung oder eine Approximation einer gegebenen Funktion sein. k gibt an, welche Frequenz untersucht

wird. Innerhalb des Arguments der Kosinusfunktion taucht der Faktor $\frac{1}{2}$ auf, der im Kapitel 3.2.2 näher untersucht wird. Aus dieser Formel, kann analog zur DFT, eine Transformationsmatrix hergeleitet werden.

Die DCT kann als eine DFT, die mit reellen Zahlen und geraden Symmetrien arbeitet, interpretiert werden. Dies erklärt die Ähnlichkeiten dieser beiden Transformationen.

3.2.2 Wichtige Unterschiede zur DFT

Im Kapitel 3.1.6 wurde erklärt, dass das Phasenbild, die visualisierten Phasenverschiebungen der Schwingungen im Bild, die Ortsinformationen des ursprünglichen Bildes beinhaltet. Durch die additive Kombination von Sinus- und Kosinusschwingungen können alle Phasen dargestellt werden, wie Abbildung 3.37 verdeutlicht:

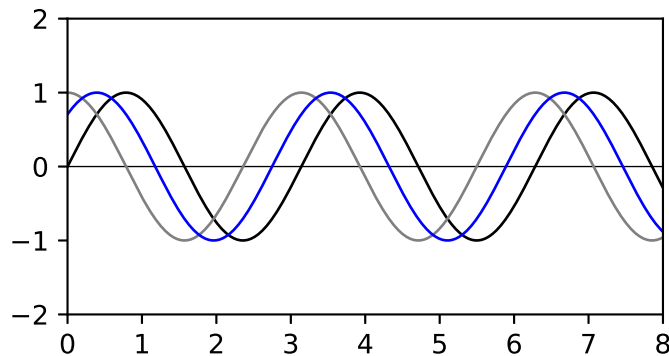


Abbildung 3.37: grau: Kosinus; schwarz: Sinus; blau: normierte Addition

Die normierte blaue Funktion ergibt sich aus der Addition der Sinus- und Kosinusschwingung und liegt, um $\frac{\pi}{4}$ verschoben, zwischen den beiden ursprünglichen Funktionen. Eine Phasenverschiebung durch die additive Verknüpfung der durch die DCT gegebenen Kosinusfunktionen ist nicht möglich. Es wird sich damit geholfen, die Schrittweite um den Faktor $\frac{1}{2}$ innerhalb des Kosinusarguments zu erhöhen, was zur Folge hat, dass nicht wie bei der DFT ganzschrittige, sondern halbschrittige Frequenzen untersucht werden. Aus diesem Grund ist es nicht möglich alle verschiedenen Phasenverschiebungen abzubilden. Durch die Erhöhung halber Frequenzschritte wird dies allerdings ausgeglichen. Die 2D-DCT ist gegeben als:

$$F_{k,l} = \sum_{n=0}^{N-1} \sum_{m=0}^{M-1} \left(\cos\left(\frac{\pi}{N}\left(mk + \frac{k}{2}\right)\right) \right) p_{m,n} \left(\cos\left(\frac{\pi}{M}\left(nl + \frac{l}{2}\right)\right) \right)$$

Aus ihr ergeben sich die Basisfunktionen der 2D-DCT für ein 4×4 -Bild (siehe Abbildung 3.38):

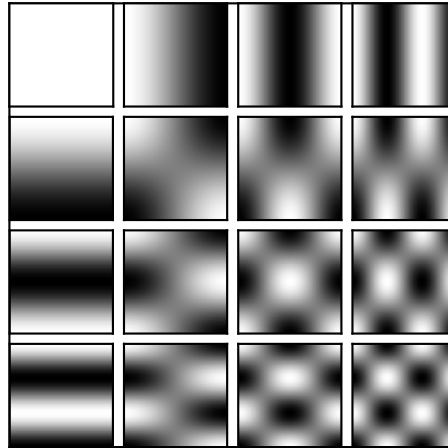


Abbildung 3.38: 2D-DCT Basisfunktionen

Die tiefen Frequenzen gehen vom Ursprung des transformierten Bildes aus und nehmen gemäß des Faktors $\frac{1}{2}$ halbschrittig zu. Aus der Linearkombination dieser Funktionen besteht das zweidimensionale Bild. Ein durch die DFT zerlegtes Bild zerfällt in sein Amplituden- und Phasenspektrum (wie Abbildung 3.39 zeigt),

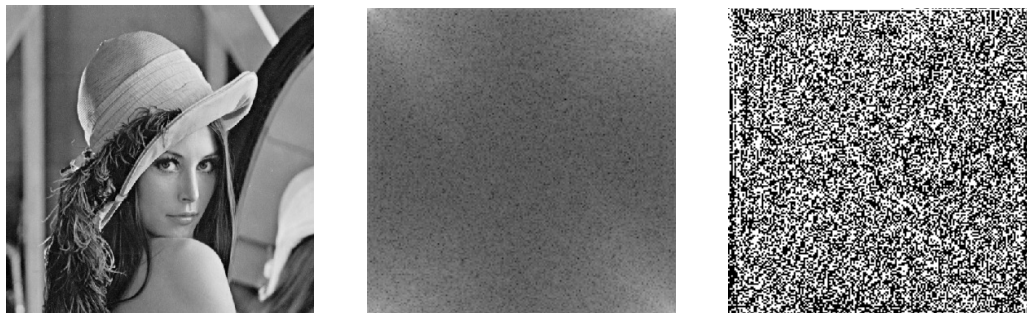


Abbildung 3.39: v.l.n.r. Lenna; Amplitudenspektrum; Phasenbild

wobei die Eingangswerte reelle Pixelwerte sind, die auf komplexe Zahlen abgebildet werden, aus denen das Amplituden- beziehungsweise Phasenspektrum berechnet wird.

Da es bei der DCT keinen Imaginärteil gibt, werden die reellen Zahlen des zu transformierenden Bildes ebenfalls auf reelle Zahlen abgebildet. Daher ist das Ergebnis ein Bild, in dem Phase und Amplitude nicht separiert sind. Für das Beispielbild ergibt sich unter Verwendung der DCT die Abbildung 3.40:



Abbildung 3.40: v.l.n.r. Lenna; transformiertes Bild (DCT)

Da die durch die Exponentialfunktion erzeugte Periodizität entfällt, enthält das Ergebnisbild jede Information einmal, wodurch sich beispielsweise das Vorkommen der tiefen Frequenzen auf den Ursprung links oben im Bild beschränkt. Werden die tiefen Frequenzen des transformierten Bildes gelöscht (Hochpassfilter) ergibt sich Abbildung 3.41:

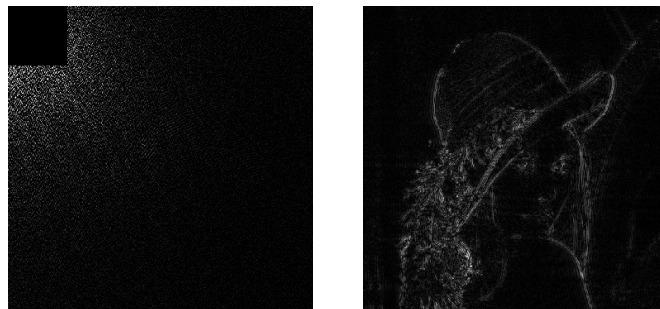


Abbildung 3.41: Hochpassfilter: 100×100

Analog zur DFT wird beim Löschen der Amplitude die Phasenverschiebung gelöscht, da die zugehörige Schwingung nicht mehr existiert. Für einen Tiefpassfilter liefert die DCT Abbildung 3.42:



Abbildung 3.42: Tiefpassfilter: 100×100

Das Ergebnis ist eine Glättung. Bei Beibehaltung der tiefen Frequenzen und Löschen der hohen Frequenzen, wird ein Großteil des transformierten Bildes gelöscht, dem zurücktransformiertem Bild fehlen jedoch nur Details. Für weitere Tiefpassfilter mit unterschiedlichen Fenstern ergeben sich die Bilder der Abbildung 3.43:



Abbildung 3.43: Tiefpassfilter: 25×25 , 50×50 , 75×75 und 100×100

Es wird deutlich, dass erst bei einem hohen Verlust der tiefen Frequenzen für das menschliche Auge erkennbar wird, dass Informationen des Bildes fehlen. Dies wird bei der JPEG Bildkompression ausgenutzt. Die gegebenen Qualitätsstufen geben an, wie die unterschiedlichen Frequenzen des Frequenzspektrums des DCT-Bildes gewichtet werden. Frequenzen werden hierbei reduziert und bei schlechten Qualitätsstufen vermehrt entfernt als bei guten Qualitätsstufen, was im Bild zu einem größeren Detailver-

lust führt. Dieser Informationsverlust kann, je nach Anwendungsfall, gewünscht sein, da auf diese Weise zum Beispiel Speicherplatz gespart werden kann [2].

Im Kapitel 3.1.1 wurde versucht die Rechteckfunktionen, die bei der Wandverkleidung auftauchen, mit Hilfe des Amplitudenspektrums herauszurechnen. Abbildung 3.44 zeigt den Versuch, dies mit Hilfe der DCT zu bewerkstelligen:

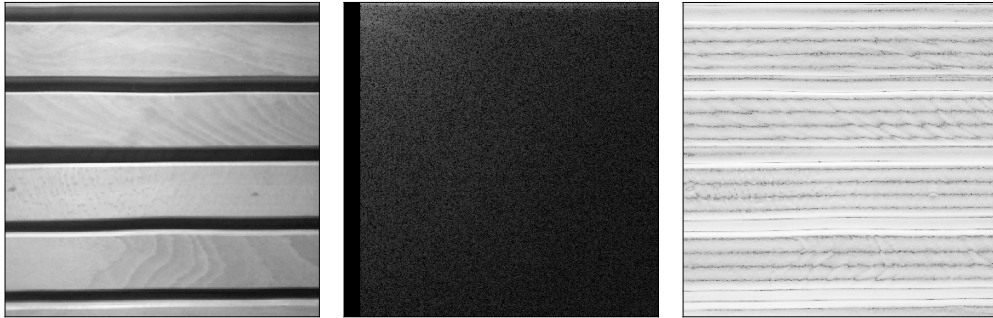


Abbildung 3.44: v.l.n.r. Holzverkleidung; bearbeitetes Amplitudenspektrum; inv. Ergebnis

Im Ergebnis sind die schwarzen Streifen noch sichtbar. Der Grund hierfür könnte sein, dass sich die DCT-Basisfunktionen schlecht für die Löschung von Rechteckfunktionen eignen, wie auch eine Überlagerung von Kosinusschwingungen deutlich macht:

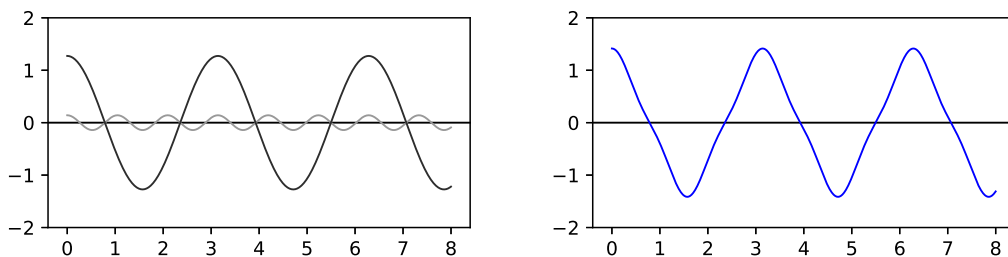


Abbildung 3.45: v.l.n.r. Kosinusschwingungen; addierte Kosinusschwingungen

Aus Abbildung 3.45 lässt sich ableiten, dass Kosinusschwingungen bei Überlagerung eher zackige Funktionen ergeben, was erklären könnte, dass die schwarzen Streifen nicht verschwinden. Allerdings, ist wie bei der DFT auch, das Bild homogener geworden.

Zusammenfassend ist zu sagen, dass die DCT ausschließlich auf reellen Zahlen beruht und, durch die Kosinusfunktionen bedingt, auf geraden Symmetrien basiert. Durch die Nutzung der gegebenen Kosinusschwingungen können nicht alle Phasenverschiebungen dargestellt werden, was dadurch ausgeglichen wird, dass halbschrittige Frequenzen berechnet werden. Obwohl es nicht möglich ist Amplituden- und Phasenspektrum getrennt voneinander zu bearbeiten, können durch Hoch- beziehungsweise Tiefpassfilter Kantendetektionen und Glättungen erzielt werden.

3.3 Hadamard-Walsh-Transformation

In diesem Kapitel wird die diskrete Hadamard-Walsh-Transformation (DHTW) als eine lineare Transformation vorgestellt, die nicht mit der DFT beziehungsweise DCT verwandt ist.

3.3.1 Aufbau der Basisfunktionen

Die Basisfunktionen der Hadamard-Walsh-Matrix bilden Rechteckfunktionen und unterscheiden sich daher von den zuvor betrachteten sinusförmigen Schwingungen [Fol92]. Anschaulich ergeben sich für die 8×8 -Transformationsmatrix folgende Funktionen 3.46:

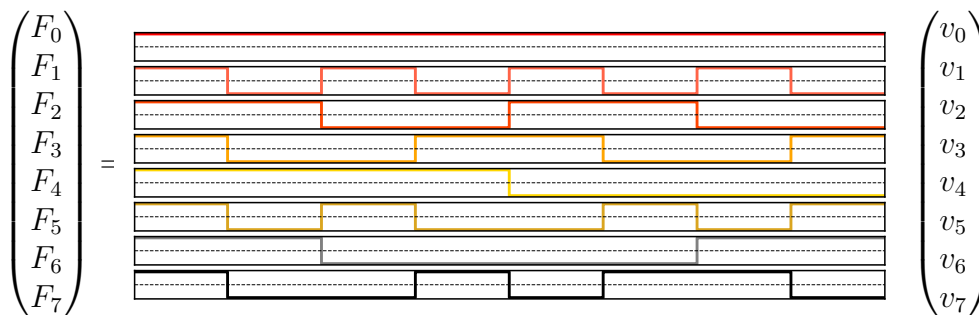


Abbildung 3.46: Transformationsmatrix (Hadamard-Walsh-Basisfunktionen)

Diese Matrix ist das Ergebnis einer Rekursion [TK08]. Sei \mathbf{H}_2 definiert als

$$\mathbf{H}_2 = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$

dann berechnet sich \mathbf{H}_n durch:

$$\mathbf{H}_n = \begin{pmatrix} \mathbf{H}_{n-1} & \mathbf{H}_{n-1} \\ \mathbf{H}_{n-1} & -\mathbf{H}_{n-1} \end{pmatrix}$$

Den in Abbildung 3.46 gezeigten Funktionen liegt demnach diese Matrix zu Grunde:

$$\mathbf{H}_8 = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 & 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 \\ 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 \\ 1 & -1 & 1 & -1 & -1 & 1 & -1 & 1 \\ 1 & 1 & -1 & -1 & -1 & -1 & 1 & 1 \\ 1 & -1 & -1 & 1 & -1 & 1 & 1 & -1 \end{pmatrix}$$

Alternativ kann diese Art der Transformationsmatrix über das Kroeneker-Produkt berechnet werden:

$$\mathbf{H}_4 = \mathbf{H}_2 \otimes \mathbf{H}_2 = \begin{pmatrix} 1 \cdot \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} & 1 \cdot \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \\ 1 \cdot \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} & (-1) \cdot \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \end{pmatrix}$$

Es wird jede Komponente der linken Matrix (blau) jeweils mit der kompletten rechten Matrix verknüpft und die Ergebnisse in eine neue Matrix eingetragen [TK08]. Einige Eigenschaften dieser Transformation werden unmittelbar deutlich:

- 1.) Die Transformationsmatrix besteht aus den reellen Werten 1 und -1 .
- 2.) Es wird auf reelle Werte abgebildet.
- 3.) Durch den rekursiven Aufbau ergeben sich Matrizen der Größe 2^n (wobei $n \in \mathbb{N}$).
- 4.) Die Berechnung der Transformationsmatrix ist einfach.

Die Hadamard-Walsh-Transformation wurde als eine weitere Möglichkeit kennengelernt Vektoren über Recheckfunktionen in einen Frequenzraum abzubilden.

3.3.2 Anwendung in der Bildverarbeitung

Im weiteren Verlauf wird behandelt inwieweit die Hadamard-Walsh-Transformation für die Bildverarbeitung geeignet ist. Die Anwendung auf ein Bild erfolgt analog zur DCT und DFT:

- 1.) Hintransformation: Jeder Zeilen- und Spaltenvektor des Bildes wird mit der Matrix multipliziert.
- 2.) Rücktransformation: Jeder transformierte Zeilen- und Spaltenvektor wird mit der inversen Matrix multipliziert.
- 3.) Anschließend werden die Werte gemittelt.

In Abbildung 3.47 wurde aus dem Beispielbild das transformierte Bild berechnet:



Abbildung 3.47: v.l.n.r. Lenna; transformiertes Bild (ungeordnete DHTW)

Ein angewandter „Hochpassfilter“ liefert folgendes Ergebnis (siehe Abbildung 3.48):

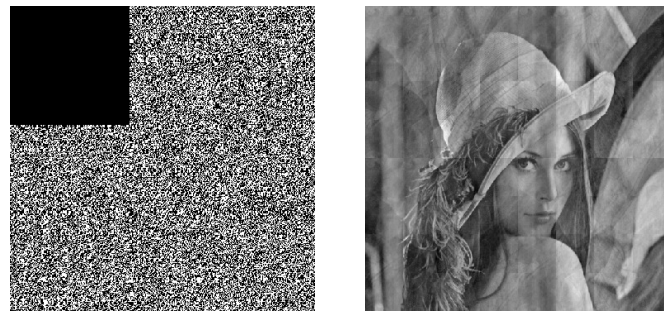


Abbildung 3.48: v.l.n.r. „Hochpassfilter“; Ergebnis

Ausgehend von der DCT und DFT wäre eine Kantendetektion zu erwarten. Problematisch ist, dass die Frequenzen nicht wie bei den zuvor behandelten Transformationen geordnet sind:

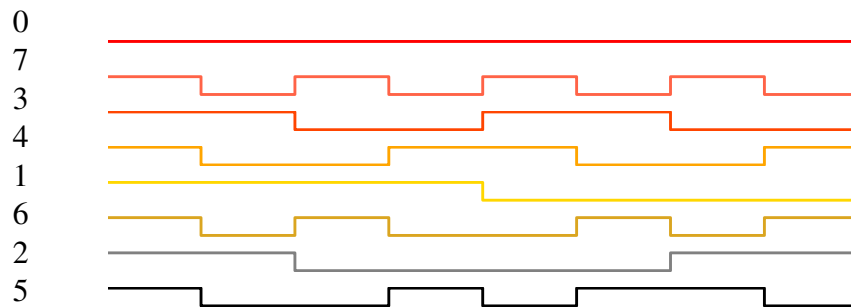


Abbildung 3.49: ungeordnete Hadamard-Walsh-Funktionen

In Abbildung 3.49 ist beispielsweise zu erkennen, dass die letzte Frequenz nicht gleichmäßig ist, da sie mal länger und mal kürzer den Wert 1 beziehungsweise -1 beibehält. Das schlechte Ergebnis der Hochpassfilterung beruht darauf, dass nicht zuverlässig tiefe Frequenzen entfernt werden, sondern ebenfalls hohe, die beibehalten werden sollen. Für jede Zeile der Matrix sind die Vorzeichenwechsel gezählt worden. Die Idee ist, die vorhandenen Frequenzen der Transformationsmatrix anhand dieser Vorzeichenwechsel zu sortieren, wie Abbildung 3.50 zeigt:

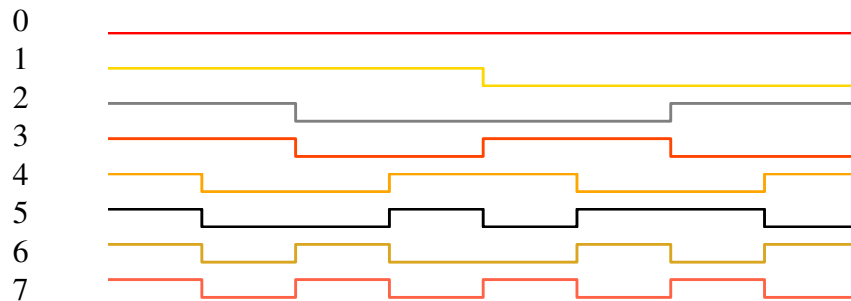


Abbildung 3.50: geordnete Hadamard-Walsh-Funktionen (nach Vorzeichenwechsel)

Gleichmäßige Schwingungen der anderen behandelten Transformationen können durch die Ordnung nicht ausgeglichen werden, jedoch ergibt sich zumindest eine Reihenfolge, die der frequentiellen Ordnung der anderen Transformationen ähnelt.

Abbildung 3.51 zeigt die Basisfunktionen für ein 4×4 -Bild:

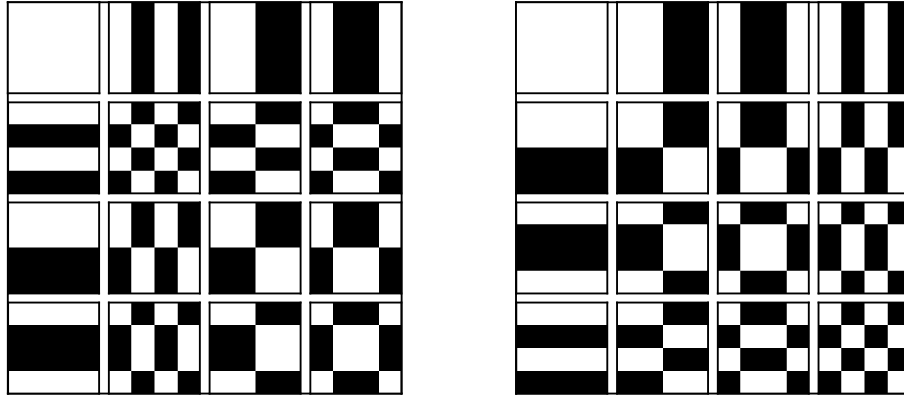


Abbildung 3.51: l.: ungeordnete Basisfunktionen; r.: geordnete Basisfunktionen

Wird auf die geordneten Frequenzen des Spektrums eine Hochpassfilterung angewandt, ergeben sich die Bilder 3.52:

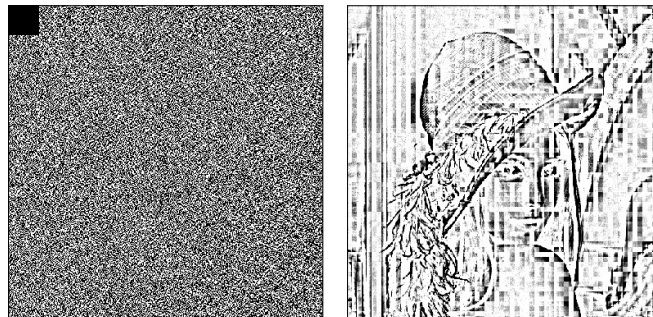


Abbildung 3.52: Hochpassfilter: 50×50 ; Ergebnis (invertiert)

Analog ergeben sich die Ergebnisse mehrerer Tiefpassfilter (siehe Abbildung 3.53).

Es ist deutlich zu erkennen, dass sich die Rechteckmuster der Basisfunktionen in den gefilterten Bildern widerspiegeln. So fallen bei kleinen Hoch- und Tiefpassfiltern die Kanten ins Gewicht und liefern daher Ergebnisse, die sich von der DFT und DCT unterscheiden.



Abbildung 3.53: Hochpassfilter: 25×25 , 50×50 , 75×75 und 100×100

Zuvor wurde für das Bild der Wandverkleidung untersucht, inwieweit die auftauchenden Rechteckfunktionen über das Amplitudenspektrum herauszurechnen sind. Für die Hadamard-Walsh-Transformation ergibt sich durch die Bearbeitung des Frequenzspektrums folgendes Ergebnis 3.54:

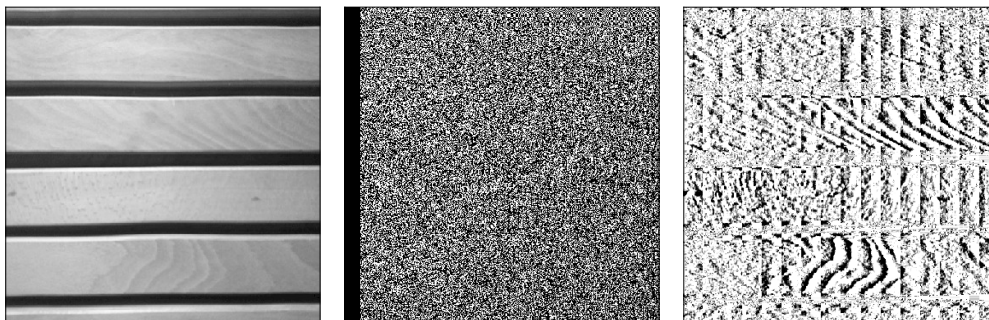


Abbildung 3.54: v.l.n.r. Holzverkleidung; bearbeitetes Amplitudenspektrum; inv. Ergebnis

In dem Ergebnisbild sind nach wie vor die Streifen zu erkennen. Obwohl zu erwarten wäre, dass die Bearbeitung durch die Hadamard-Walsh-Transformation auf Grund der rechteckigen Basisfunktionen gut klappen könnte, ist das Ergebnis ernüchternd.

Die Hadamard-Walsh-Transformation hat durch die rechteckigen Basisfunktionen also eine Sonderstellung innerhalb der linearen Transformationen. Dennoch können der

DCT und DFT ähnliche Ergebnisse erzielt werden. In diesen Ergebnissen bilden sich für kleinere Filter deutlich die rechteckigen Basisfunktionen ab, die bei zunehmender Filtergröße immer geringer ausfallen.

Kapitel 4

Vergleich der Transformationen

In diesem Kapitel werden die Transformationen im Hinblick auf ihre vorgestellten Merkmale verglichen. Zuerst werden einige Eigenschaften der Transformationen wiederholt. Anschließend werden Gemeinsamkeiten hervorgehoben, Unterschiede untersucht und abschließend in der Bildverarbeitung erzielte Ergebnisse miteinander verglichen.

	Eingangswerte	Ausgangswerte	Basisfunktionen	Anz. Spektren
DFT	reelle Zahlen	komplexe Zahlen	Exponentialfunktionen	2
DCT	reelle Zahlen	reelle Zahlen	Kosinusfunktionen	1
DHTW	reelle Zahlen	reelle Zahlen	Rechteckfunktionen	1

Tabelle 4.1: Vergleich einiger Eigenschaften der behandelten Transformationen

In der Tabelle 4.1 sind verschiedene Eigenschaften der Transformationen aufgeführt. Jede Transformation erhält reelle Zahlen als Eingangswerte. Der Wertebereich ist für die DCT und die Hadamard-Walsh-Transformation ebenfalls reell. Die DFT hingegen bildet auf komplexe Werte ab, was in der Folge dazu führt, dass ein Phasenspektrum und ein Amplitudenspektrum entstehen. Die Basisfunktionen jeder behandelten Transformation sind verschiedene, diskretisierte, schwingungsförmige Funktionen, wobei es sich bei der DFT um komplexe Exponentialfunktionen, bei der DCT um reelle Kosinusfunktionen und bei der Hadamard-Walsh-Transformation um reelle Rechteckfunktionen handelt. Da alle drei Transformationen linear sind, existiert jeweils eine Transformationsmatrix, bestehend aus den jeweiligen Basisfunktionen, durch die in einen Frequenzraum abgebildet werden kann. Dieser Vorgang ist umkehrbar. Die Matrix-Vektor-Multiplikation kann als Ähnlichkeit zwischen den Zeilenvektoren und dem zu transformierenden Spaltenvektor interpretiert werden.

In Abbildung 4.1 sind die Ergebnisse eines kleinen Tiefpassfilters mit Hilfe der verschiedenen Transformationen nebeneinander aufgezeigt:



Abbildung 4.1: v.l.n.r. 15×15 Tiefpassfilter: DHTW (geordnet); DCT; DFT

Bei diesen Glättungen fällt auf, dass die Fouriertransformation (rechts) das deutlich beste Ergebnis liefert. Während sich bei Anwendung der DCT der Bildinhalt noch erahnen lässt, können für die geordnete Hadamard-Walsh-Transformation (links) nur noch Pixel wahrgenommen werden. Diese Unterschiede liegen darin begründet, dass die DFT auf Exponentialfunktionen aufbaut, wodurch jede Phasenverschiebung dargestellt werden kann. Bei der DCT hingegen wird dies durch halbschrittige Frequenzschritte ausgeglichen, was zu einem Ergebnis führt, welches sich qualitativ mittig einordnet. Innerhalb des Vergleichs dieser drei Transformationen schneidet die Hadamard-Walsh-Transformation offenbar bei der Anwendung kleiner Tiefpassfilter am schlechtesten ab. Dies ist der Tatsache geschuldet, dass keine Frequenzen im eigentlichen Sinne existieren, sondern innerhalb einer Zeile der Transformationsmatrix eher Takte angegeben werden. Den Vergleich kleiner Hochpassfilter liefert Abbildung 4.2:

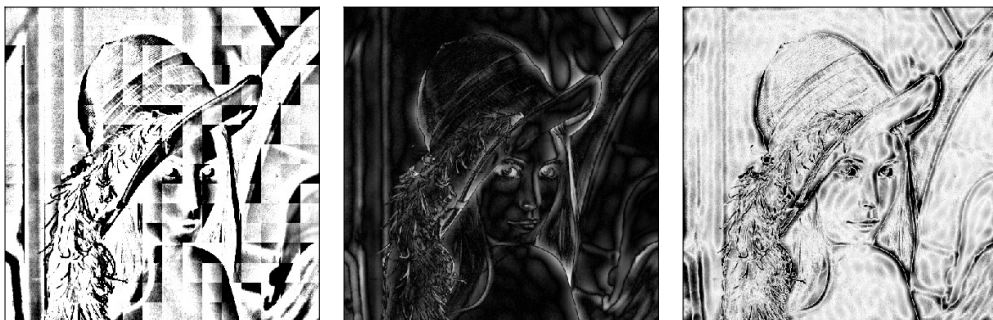


Abbildung 4.2: v.l.n.r. 15×15 Hochpassfilter: DHTW (geordnet, inv.); DCT; DFT (inv.)

Die zuvor dargelegte Argumentation scheint ebenfalls für Hochpassfilter zu gelten. Für ein gutes Resultat müssen viele im Bild vorhandene Kanten detektiert werden, wobei im Ergebnisbild die Kantenverläufe nicht zu großzügig ausfallen sollten. Auch hier fällt das schlechtere Ergebnis der geordneten Hadamard-Walsh-Transformation, das

durch harte Kanten der Rechteckfunktionen gezeichnet ist, auf. Diese Transformation findet zwar Kanten, aber ihr fehlt die Feinheit diese korrekt darzustellen. Beim Resultat der DCT werden Kanten besser detektiert, können jedoch nur ungenau dargestellt werden. Die DFT besticht durch die Feinheit, mit der Kanten aufgezeigt werden können und liefert dadurch das beste Ergebnis. Für einen größeren Tiefpassfilter ergeben sich die Effekte der Abbildung 4.3:



Abbildung 4.3: v.l.n.r. 100×100 Tiefpassfilter: DHTW (geordnet) ; DCT; DFT

Das Bild der geordneten Hadamard-Walsh-Transformation entspricht eher dem, was von einer Glättung erwartet wird. Die rechteckigen Basisfunktionen fallen nicht mehr so sehr ins Gewicht, wie bei der Anwendung des kleinen Tiefpassfilters. Interessanterweise ist das Resultat der DCT für diese Größe des Filters dem der DFT nahezu identisch. Erst bei genauer Betrachtung können feine Unterschiede wahrgenommen werden. Dies führt zu der Annahme, dass die geordnete Hadamard-Walsh-Transformation und die DCT besser mit größeren Tiefpassfiltern arbeiten, als mit kleinen. Diese These soll nun unter Anwendung eines größeren Hochpassfilters überprüft werden. Abbildung 4.4 veranschaulicht den äußerst verschiedenen Effekt der Transformationen bei Nutzung eines größeren Hochpassfilters.

Auch hier fällt auf, dass, wenn die Details nicht mehr so ins Gewicht fallen, alle Transformationen ein brauchbares Ergebnis liefern. Die geordnete Hadamard-Walsh-Transformation kann auf Grund des größeren Filters die Kanten des Bildes genauer beschreiben, als bei einem Kleineren, wodurch Kanten gefunden werden, wie sie auch im Bild vorhanden sind. Die DFT und DCT liefern wiederum fast identische Resultate.

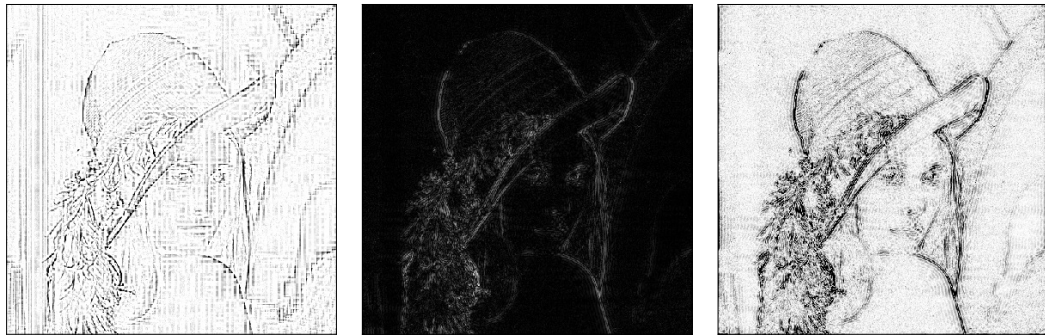


Abbildung 4.4: v.l.n.r. 100×100 Hochpassfilter: DHTW (geordnet, inv.); DCT; DFT inv.

Schließlich werden die Bilder der Wandverkleidung miteinander verglichen (siehe Abbildung 4.5):

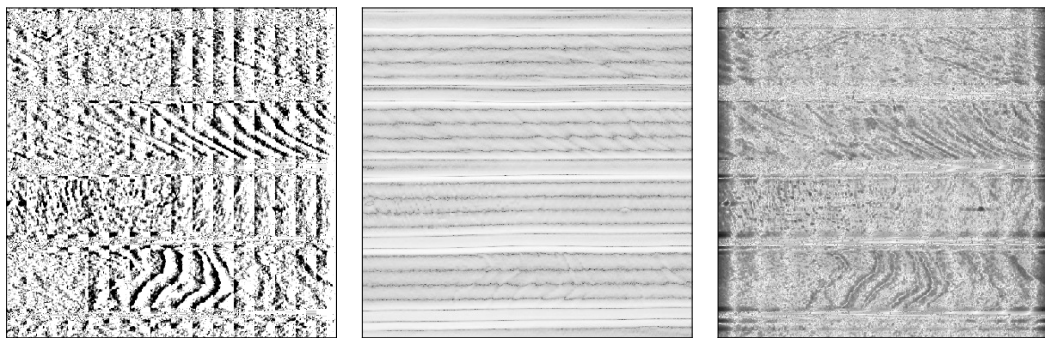


Abbildung 4.5: v.l.n.r. DHTW (geordnet, inv.); DCT (inv.); DFT (invertiert)

Untersucht werden sollte, in welchem Ausmaße sich schwarze Streifen des Bildes mit Hilfe der Frequenzspektren herausrechnen lassen. Die Idee dahinter war, dass der Übergang von hell nach dunkel beziehungsweise dunkel nach hell in vertikaler Richtung des Bildes als Rechteckfunktion interpretiert werden kann, die sich ebenfalls im Frequenzspektrum vertikal erstreckt. Die Ergebnisse spiegeln das Löschen dieser Frequenzen wieder, wobei für die DFT nur das Amplitudenspektrum bearbeitet wurde. In allen drei Fällen sind die invertierten schwarzen Streifen deutlich zu erkennen, wobei diese aus dem DCT Ergebnisbild am besten hervorgehen. In allen drei Fällen konnte und sollte erreicht werden, dass die Ergebnisbilder in ihren Graustufen einheitlicher verlaufen, was durch das Löschen dieser Frequenzen bewirkt wurde.

Abschließend ist festzuhalten, dass alle drei behandelten linearen Transformationen sich in ihren Basisfunktionen unterscheiden und somit unterschiedlich gute beziehungsweise schlechte Ergebnisse mittels Tief- oder Hochpassfiltern liefern. Allgemein

scheint zu gelten, dass die DFT am besten, sowohl bei kleinen als auch bei großen Filtern, abschneidet. Auf Grund der Basisfunktionen, welche die DCT und die geordnete Hadamard-Walsh-Transformation bieten, können diese bei kleinen Filtern nicht konkurrieren, da Details nicht hinreichend beschrieben werden können. Bei größeren Filtern liefert die geordnete Hadamard-Walsh-Transformation im Bezug auf die Einfachheit ihrer Basisfunktionen bei Hoch- als auch bei Tiefpassfiltern, gute Ergebnisse. Die Resultate der DCT und DFT sind für das menschliche Auge für genügend große Filter kaum noch zu unterscheiden. Zusätzlich wurde festgestellt, dass sich die angewandte Methodik auf die Frequenzspektren am Beispielbild der Wandverkleidung offenbar nicht gut dazu eignet die Struktur des Bildinhalts zu bearbeiten. Da beim Löschen der Amplituden immer die zugehörigen Schwingungen gelöscht werden, geht dies zu lasten der Genauigkeit des Bildinhalts, wodurch sich kein gutes Ergebnis erzielen lässt. Zu einer Zeit rechenschwacher Computer war die Existenz der Hadamard-Walsh-Transformation berechtigt, da sich über einen Algorithmus, ähnlich der FFT, die Berechnungen reduzieren lassen. Die DCT hat, auf Grund ihrer Fähigkeit Daten ohne deutliche Qualitätsverluste einsparen zu können, ihren Platz in der Bildkompression gefunden. Da leistungsstarke Computer heute gegeben sind und die DFT über den FFT Algorithmus ebenfalls effizient berechnet werden kann, wird die Transformation den anderen in vielen Bereichen vorgezogen.

Tabellenverzeichnis

4.1	Vergleich einiger Eigenschaften der behandelten Transformationen . . .	71
-----	--	----

Abbildungsverzeichnis

2.1	Konvergenzverhalten einer Folge a_n	14
2.2	Komplexe Zahlenebene	16
2.3	Komplexer Einheitskreis	18
2.4	Graphische Darstellung der Zusammensetzung eines Vektorraums	19
2.5	Beispiel einer Linearkombination	21
2.6	Transformationszyklus von zwei Basen	22
3.1	v.l.n.r. Lenna; Kantendetektion (invertiert); Glättung	27
3.2	Grundschiwingung und Oberschwingungen einer Sinusfunktion	28
3.3	Rechtecksignal	29
3.4	Approximationen eines Rechtecksignals	29
3.5	Übersicht	30
3.6	Transformationsmatrix	31
3.7	Übersicht	32
3.8	Übersicht	36
3.9	Eingetragene Pixelwerte	36
3.10	Korrelation zwischen Basisfunktion und Pixelwerten	37
3.11	Übersicht	38
3.12	Amplitudenspektrum der Pixelwerte	39
3.13	Vergleich der alten und neuen Pixelwerte	41
3.14	Übersicht	42
3.15	v.l.n.r. Kosinusbasisfunktionen; Sinusbasisfunktionen	42
3.16	Pixelort zeigt welche Frequenz untersucht wird.	43
3.17	v.l.n.r. Lenna; Amplitudenspektrum; Phasenbild	44
3.18	Amplitudenspektrum (tiefe Frequenzen in die Mitte gedreht)	44
3.19	v.l.n.r. zurücktrans. Amplitudenspektrum; zurücktrans. Phasenbild	45
3.20	Übersicht	45
3.21	Transformierter Gaußfilter	46
3.22	v.l.n.r. verändertes Amplitudenspektrum; Rücktransformation (Glättung)	46
3.23	Transformierte Sobelfilter	47
3.24	v.l.n.r. verändertes Amplitudenspektrum; Kantendetektion (invertiert)	47

3.25	Veranschaulichung des Faltungstheorems	48
3.26	Holzverkleidung	48
3.27	v.l.n.r. Schwingungen; addierte Schwingungen	49
3.28	v.l.n.r. Schwingungen; addierte Schwingungen	49
3.29	v.l.n.r. Holzverkleidung; bearbeitetes Amplitudenspektrum; inv. Ergebnis	50
3.30	v.l.n.r. Holzverkleidung; bearbeitetes Amplitudenspektrum; inv. Ergebnis	50
3.31	v.l.n.r. Lenna; Amplitudenspektrum	51
3.32	aufgefülltes Amplitudenspektrum	52
3.33	Ergebnis nach Rücktransformation	52
3.34	Ergebnis nach erneuter Transformation	52
3.35	Übersicht	53
3.36	Transformationsmatrix mit kontinuierlichen Basisfunktionen (DCT) . .	57
3.37	grau: Kosinus; schwarz: Sinus; blau: normierte Addition	58
3.38	2D-DCT Basisfunktionen	59
3.39	v.l.n.r. Lenna; Amplitudenspektrum; Phasenbild	59
3.40	v.l.n.r. Lenna; transformiertes Bild (DCT)	60
3.41	Hochpassfilter: 100×100	60
3.42	Tiefpassfilter: 100×100	61
3.43	Tiefpassfilter: 25×25 , 50×50 , 75×75 und 100×100	61
3.44	v.l.n.r. Holzverkleidung; bearbeitetes Amplitudenspektrum; inv. Ergebnis	62
3.45	v.l.n.r. Kosinusschwingungen; addierte Kosinusschwingungen	62
3.46	Transformationsmatrix (Hadamard-Walsh-Basisfunktionen)	64
3.47	v.l.n.r. Lenna; transformiertes Bild (ungeordnete DHTW)	66
3.48	v.l.n.r. „Hochpassfilter“; Ergebnis	66
3.49	ungeordnete Hadamard-Walsh-Funktionen	67
3.50	geordnete Hadamard-Walsh-Funktionen (nach Vorzeichenwechsel) . . .	67
3.51	l.: ungeordnete Basisfunktionen; r.: geordnete Basisfunktionen	68
3.52	Hochpassfilter: 50×50 ; Ergebnis (invertiert)	68
3.53	Hochpassfilter: 25×25 , 50×50 , 75×75 und 100×100	69
3.54	v.l.n.r. Holzverkleidung; bearbeitetes Amplitudenspektrum; inv. Ergebnis	69
4.1	v.l.n.r. 15×15 Tiefpassfilter: DHTW (geordnet); DCT; DFT	72
4.2	v.l.n.r. 15×15 Hochpassfilter: DHTW (geordnet, inv.); DCT; DFT (inv.)	72
4.3	v.l.n.r. 100×100 Tiefpassfilter: DHTW (geordnet) ; DCT; DFT	73
4.4	v.l.n.r. 100×100 Hochpassfilter: DHTW (geordnet, inv.); DCT; DFT inv.	74
4.5	v.l.n.r. DHTW (geordnet, inv.); DCT (inv.); DFT (invertiert)	74

Notationstabelle

Symbol	Erklärung
n	natürliche Zahl
m	natürliche Zahl
a_n	Folge
g	Grenzwert
x	Funktionsparameter
$f(x)$	Funktion
e	eulersche Zahl
\in	enthalten in
\mathbb{N}	Menge der natürlichen Zahlen
\mathbb{R}	Menge der reellen Zahlen
\mathbb{C}	Menge der komplexen Zahlen
z	komplexe Zahl
a	Realteil
b	Imaginärteil
i	imaginäre Einheit
α	Winkel
π	Kreiszahl
P	Punkt in der komplexen Zahlenebene
t	Zeit
w	Vektor
v	Vektor
u	Vektor
s	Skalar
B	Basis
T	Transformationsmatrix
F	Korrelation des Skalarprodukts von Zeilen- und Spaltenvektor
N	abgetastete Werte (Samples)
M	abgetastete Werte (Samples)
p	Pixelwerte
l	zu untersuchende Frequenz
k	zu untersuchende Frequenz
ω	Substituierte e-Funktion
fou	komplexe Fourierreihe
exp	Platzhalter einer Exponentialfunktion
H	Hadamard-Walsh-Transformationsmatrix

Literatur

- [AHK⁺09] ARENS, T. ; HETTLICH, F. ; KARPFIGER, C. ; KOCKELKORN, U. ; LICHTENEGGER, K. ; STACHEL, H.: *Mathematik*. Auflage 2. Springer Berlin Heidelberg, 2009. – 114; 574–576; 1044–1060 S. – ISBN 9783642449192
- [Beu09] BEUTELSPACHER, A.: *Lineare Algebra*:. Auflage 7. Vieweg+Teubner Verlag, 2009 (Mathematik für Studienanfänger). – 48–49 S. – ISBN 9783528665081
- [Bov09] BOVIK, A.C.: *The Essential Guide to Video Processing*. Elsevier Science, 2009. – 101 S. – ISBN 9780080922508
- [Fol92] FOLLAND, G.B.: *Fourier Analysis and Its Applications*. Wadsworth & Brooks/Cole Advanced Books & Software, 1992 (Advanced Mathematics Series). – 201; 251 S. – ISBN 9780534170943
- [MW06] MERZIGER, Gerhard. ; WIRTH, Thomas.: *Repetitorium der höheren Mathematik*. Springer : Binomi, 2006. – ISBN 9783923923335
- [Pap09] PAPULA, L.: *Mathematik für Ingenieure und Naturwissenschaftler Band 1: Ein Lehr- und Arbeitsbuch für das Grundstudium*. 12.Auflage. Vieweg+Teubner Verlag, 2009 (Viewegs Fachbücher der Technik). – 176–181 S. – ISBN 9783834817495
- [Pri15] PRIESE, L.: *Computer Vision: Einführung in die Verarbeitung und Analyse digitaler Bilder*. Springer Berlin Heidelberg, 2015 (eXamen.press). – 185 S. – ISBN 9783662451281
- [Sch10] SCHULZ, R.H.: *Repetitorium Bachelor Mathematik: Zur Vorbereitung auf Modulprüfungen in der mathematischen Grundausbildung*. Auflage 7. Vieweg+Teubner Verlag, 2010 (Vieweg Studium). – 6 S. – ISBN 9783834896322
- [TK08] THEODORIDIS, S. ; KOUTROUMBAS, K.: *Pattern Recognition*. Auflage 4. Elsevier Science, 2008. – 368–369 S. – ISBN 9780080513621

Internet-Quellen

- [1] Scott E. Brodie. <https://www.cut-the-knot.org/arithmetic/algebra/Scott.shtml>, letztmals besucht am 27.2.2018.
- [2] A.M. Raid; W.M. Khed; M.A. El dosuky; Whesam Ahmed. <http://airccse.org/journal/ijcses/papers/5214ijcses04.pdf>, letztmals besucht am 27.2.2018.
- [3] Jean-Baptiste Joseph Fourier. https://www.irphe.fr/~clanet/otherpaperfile/articles/Fourier/N0029061_PDF_1_676.pdf, letztmals besucht am 27.2.2018.
- [4] Professor Colin H. Hansen. http://www.who.int/occupational_health/publications/noise1.pdf, letztmals besucht am 27.2.2018.
- [5] Elisha Rhodes Huggins. http://physics2000.com/PDF/Text/Ch_16_FOURIER_ANALYSIS,_NORMAL_MODES_AND_SOUND.pdf, letztmals besucht am 27.2.2018.
- [6] Itseez. <https://opencv-python-tutroals.readthedocs.io/en/latest/>, letztmals besucht am 1.3.2018.
- [7] Dr. Jürgen Pöschel. <http://www.mathematik.uni-stuttgart.de/studium/infomat/HM-Poeschel-WS1415/skript/5-c.pdf>, letztmals besucht am 27.2.2018.
- [8] Wolfgang Stöbel. <https://rd.springer.com/content/pdf/bfm%3A978-3-662-01618-3%2F1.pdf>, letztmals besucht am 27.2.2018.
- [9] Unbekannt. <https://www.geogebra.org/?lang=de>, letztmals besucht am 1.3.2018.
- [10] Unbekannt. <https://www.draw.io/>, letztmals besucht am 1.3.2018.
- [11] Unbekannt. <http://links.uwaterloo.ca/Repository.html>, letztmals besucht am 27.2.2018.

- [12] Unbekannt. <http://mathworld.wolfram.com/VandermondeMatrix.html>, **letztmals besucht am 1.3.2018.**
- [13] Unbekannt. <http://mathworld.wolfram.com/NyquistFrequency.html>, **letztmals besucht am 1.3.2018.**
- [14] Unbekannt. <http://mathworld.wolfram.com/ParsevalsTheorem.html>, **letztmals besucht am 27.2.2018.**
- [15] Ron Yorita. <http://digitalcommons.calpoly.edu/cgi/viewcontent.cgi?article=2422&context=theses>, **letztmals besucht am 27.2.2018.**