Institut für Neuroinformatik

Ruhr-Universität Bochum

Internal Report 2008–01

Untersuchung der Datenredundanz von Bündelgraphen zur Gesichtserkennung

by

Darius Malysiak

Ruhr-Universität Bochum
Institut für Neuroinformatik
44780 Bochum

IR-INI 2008–01
Juli 2008
ISSN 0943-2752

© 2008 Institut für Neuroinformatik, Ruhr-Universität Bochum, FRG

Untersuchung der Datenredundanz von Bündelgraphen zur Gesichtserkennung

Darius Malysiak
Institut für Neuroinformatik, Ruhr-Universität Bochum, FRG

Abstract

Die vorliegende Arbeit stellt die Untersuchung so genannter Bündelgraphen dar, welche im Bereich der Gesichtserkennung angewandt werden. Der hierbei verwendete EBGM-Algorithmus nutzt Graphen dieser Art zur Bestimmung markanter Punkte innerhalb eines gegebenen Bildes, dieser Vorgang stellt einen wichtigen Schritt vor der anschliessenden Gesichtserkennung dar. Zur Lokalisierung der erwähnten Punkte, werden die an jedem Knoten des Bündelgraphen vorhandenen Informationen genutzt. In dieser Arbeit wird nun der Versuch unternommen die Datenredundanz dieser Informationen zu bestimmen. Der hierzu gewählte Ansatz ist die Untersuchung der Zusammenhänge zwischen Genauigkeit der Lokalisierung und dazugehöriger Erkennungsrate, diese werden bestimmt während ausgewählte Information der Knoten entfernt werden.

Inhaltsverzeichnis

1 Einleitung **4**
1.1 Motivation . . . 5

2 Das menschliche Sehen **6**

3 Fourier Analysis **10**
3.1 Fourier-Reihen . . . 10
3.1.1 Orthogonalität / Orthonormalität . . . 10
3.1.2 Vektorräume . . . 10
3.1.3 Fourier-Reihen . . . 12
3.2 Fourier-Transformation . . . 16
3.3 Abtast-Theorem und Aliasing . . . 18
3.4 Diskrete Fourier Transformation . . . 20

4 Heisenbergsche Unschärferelation **22**
4.1 Lokalisierung im Frequenzraum . . . 22
4.2 Mathematischer Hintergrund . . . 23

5 Wavelet-Transformation **28**
5.1 Gefensterte Fourier Transformation . . . 28
5.2 Wavelet-Transformation . . . 31

6 Gabor Wavelets **36**
6.1 Gabor-Funktionen . . . 36
6.2 Gabor Funktionen als Wavelet-Transformation . . . 36

7 Datenredundanz und Bilddatenkompression **42**

8 Elastic Bunch Graph Matching **46**
8.1 Einleitung . . . 46
8.2 Merkmalsrepräsentation . . . 46
8.2.1 FeaSt / Jet . . . 46
8.2.2 Modellgraphen / Bündelgraphen . . . 47
8.3 Elastic Bunch Graph Matching . . . 48
8.3.1 Elastic Graph Matching . . . 48
8.3.2 Moves . . . 49

9 Redundanz - Analyse **52**
9.1 Transformation eines Bündelgraphen in Matrixform . . . 52
9.2 Analyse-Algorithmen . . . 53
9.2.1 Redundanz-Analyse . . . 53
9.2.2 Ausdünnen / Komprimieren eines Bündelgraphen . . . 55
9.3 Beurteilung der Qualität eines Bündelgraphen . . . 57
9.3.1 Lokalisierungsfehler . . . 57
9.3.2 Kreuzvergleich . . . 59
9.4 Verwendete Bilddatenbanken und Untersuchungsaufbau . . . 60
9.4.1 FaceGen Test-Sets . . . 61
9.4.2 CAS-PEAL Test-Sets . . . 61
9.4.3 Zusammengestellte Test-Sets . . . 61
9.5 Parametrisierung der Moves und topologische Kosten . . . 62
9.6 Untersuchungsablauf . . . 63

9.6.1 Untersuchung der Erkennungsrate 63
9.6.2 Untersuchung des Lokalisierungsfehlers 63
9.7 Zusammenfassung 64

10 Resultate **65**
10.1 Einleitung 65
10.2 FaceGen Test-Sets 65
10.2.1 Signifikanz des Komprimierungsverfahrens 66
10.2.2 Zusammenhang von Lokalisierungsfehler und Erkennungsrate 66
10.2.3 Vollständige Ausdünnung 68
10.3 CAS-PEAL Test-Sets 68
10.4 Visualisierte Ergebnisse 70

11 Diskussion **73**

A Test-Sets **75**

1. Einleitung

Die alltägliche Aufgabe Objekte im persönlichen Umfeld zu identifizieren, wird meist unterbewusst durchgeführt. Obwohl es einem Menschen sehr leicht fallen kann, bekannte Gesichter, Gebäude, Umgebungen etc zu identifizieren, so wird bei genauerer Analyse dieser Aufgabe deutlich, dass es sich dabei um kein triviales Problem handelt. Wir erkennen ein und dasselbe Objekt unter verschiedenen Konditionen, hierbei sind u.a. unterschiedliche Beleuchtungsverhältnisse, Rotation, Translation und Skalierung des Objektes gemeint. Die Leichtigkeit, mit welcher wir diese Aufgabe meistern, soll nicht implizieren, dass diese sehr einfach ist. Dieses komplexe Problem wird vom Gehirn mit Bravour gemeistert, doch wie geht es dabei vor bzw. welche Probleme sind zu bewältigen?
An dieser Stelle sollen die Probleme hinsichtlich technischer Adaption der zuvor erwähnten Objekterkennung kurz angesprochen werden. Gehen wir von 2 Grauwertbildern aus welche das zu erkennende Objekt beinhalten. Betrachten wir den Ansatz des reinen Pixelwertvergleichs, [12] führt hierzu das sogenannte Korrespondenzproblem auf

> **Korrespondenzproblem:** Sind 2 Bilder des selben Objektes gegeben, so gilt es zu entscheiden welche Paare von Punkten auf ein und demselben physikalischen Objekt liegen. Die Möglichkeit, dass Punkte ohne korrespondierenden Partner existieren muss ebenfalls berücksichtigt werden.

Betrachten wir weiterhin den Begriff der Textur, es existiert keine einheitliche Definition dieses Wortes, [4] führt beispielsweise folgende Beschreibung auf:

> ”Texturen sind 2-dimensionale Felder mit variierender Intensität, wobei die Regeln zur Variation der Anordnung beliebig sind, solange eine charakteristische Wiederholung der Grundelemente vorhanden ist”

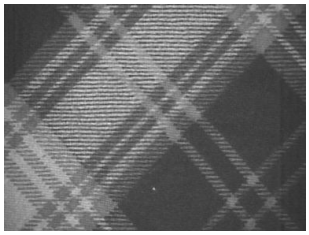


Abb. 1: Textil

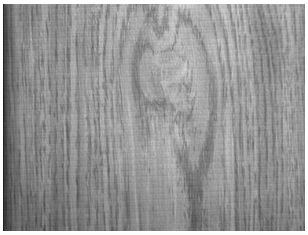


Abb. 2: Holz

Abb. 3: menschliches Gesicht

In Abb.(1) bis (3) sind verschiedene Beispiele für Texturen dargestellt, um Fehlinterpretationen zu vermeiden, soll hier nicht weiter versucht werden, den Begriff der Textur zu definieren. Im weiteren Verlauf dieser Arbeit sei der Begriff ”Textur” als ein Synonym für Muster aufzufassen.
Betrachten wir den zuvor erwähnten Fall zweier Grauwertbilder, nennen diese jeweils Bild A bzw. Bild B und nehmen weiterhin an, dass das zu identifizierende bzw. zu vergleichende Objekt ein menschliches Gesicht ist. In diesem Fall kann ein markanter Punkt z.B. im Bereich des linken Auges einer Person im Bild A liegen. Ein markanter Punkt bzw. eine Landmarke stellt hierbei die Position, eines für die Objektklasse wichtigen Merkmals im Bild dar. Das in dieser Arbeit verwendete Vergleichsverfahren basiert auf einem Texturvergleich. D.h. die Textur in einem begrenzten Bereich, um den zuvor genannten Punkt, wird mit der Textur eines korrespondierenden Punktes im anderen Bild verglichen. Jedoch tritt hierbei das zuvor erwähnte Korrespondenzproblem auf, schließlich wird zu jedem markanten Punkt im Bild A, der Zugehörige im Bild B benötigt. Mit anderen Worten ausgedrückt muss für den Punkt im Bild A ein weiterer im anderen Bild gefunden werden, dieser muss natürlich das selbe Merkmal kennzeichnen (in unserem Beispiel das linke Auge).
Zum lösen dieses Problems existiert das so genannte Graph Matching. Dieses Verfahren lokalisiert Landmarken mit Hilfe von Graphen, deren Knoten jeweils einen markanten Punkt symbolisieren (zum besseren Verständniss sei auf die Bilder 4,5 verwiesen). Solch ein Graph wird als Modellgraph bezeichnet.

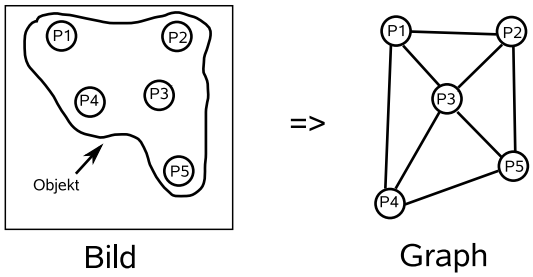


Abb. 4: Landmarken allgemein

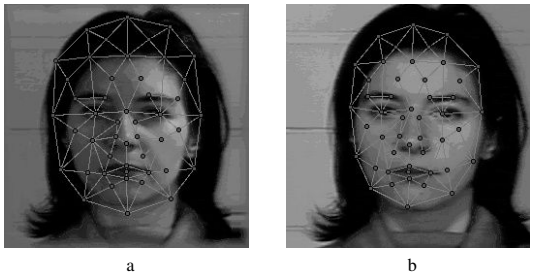


Abb. 5: Landmarken eines Gesichts

Im Bild 4 wird dargestellt wie ein Modellgraph aussehen kann, das Bild 5 zeigt die Anwendung des Verfahrens im Bereich der Gesichtserkennung. D.h. ein solcher Graph stellt eine neue Form der Gesichtsrepräsentation dar. Jedoch werden für das Lokalisieren der markanten Punkte innerhalb eines gegebenen Bildes, mehrere solcher Graphen benötigt. Sobald genug davon zur Verfügung stehen, wird aus diesen ein so genannter Bündelgraph gebildet. Dies ist ein neuer Graph welcher eine ganze Gesichtsklasse repräsentiert. Zum Finden der markanten Punkte wird der Bündelgraph über das gegebene Bild bewegt bis die entsprechenden Landmarken gefunden wurden, eine nähere Erläuterung des Verfahrens wird in Kapitel 8 gegeben.
Zum Abschluss dieses Abschnittes, soll ein kurzer Überblick über den Aufbau dieser Arbeit gegeben werden. In Kapitel 2 wird der neurobiologische Hintergrund des in dieser Arbeit verwendeten Verfahrens erklärt, und eine Verbindung zwischen in späteren Abschnitten erläuterten mathematischen Operationen hergestellt. Kapitel 3 beschäftigt sich mit benötigten Grundlagen aus der Signalverarbeitung, diese werden weitestgehend mathematisch für den Fall 1-dimensionaler Signale vertieft. Ein wichtiger Aspekt dieses Kapitels ist die Fourier-Transformation, welche die Grundlagen der in späteren Kapiteln verwendeten Textur-Analyse veranschaulichen soll. Aufbauend auf den angesprochenen Grundlagen, wird in Kapitel 4 die Heisenbergsche Unschärferelation hergeleitet und veranschaulicht. Mit diesen Kenntnissen werden anschließend im 5. Kapitel die Grundlagen der Wavelet-Analyse anhand der Haar-Transformation erklärt. Im nachfolgenden Kapitel wird schließlich der Übergang zur 2-dimensionalen Gabor-Wavelet-Transformation gemacht, welche für das im Kapitel 8 erläuterte EBGM-Verfahren sehr wichtig ist. Dieses Verfahren dient der Lokalisierung von markanten Punkten in unbekannten Bildern. Alle bisher erwähnten Kapitel stellen bereits vorhandene Verfahren dar bzw. beschreiben diese, die nachfolgenden Kapitel 9-11 beschäftigen sich mit der im Rahmen dieser Arbeit durchgeführten Untersuchung (unter Nutzung der zuvor erwähnten Verfahren). Diese 3 Kapitel repräsentieren die vom Autor dieser Arbeit durchgeführte Untersuchung. In diesen werden die verwendeten Untersuchungsmethoden, der Untersuchungsablauf, die Resultate sowie die Feststellungen beschrieben.

1.1. Motivation

Ein signifikanter Schritt des Objektvergleichs bzw. der Objekterkennung ist das bereits erwähnte Finden bestimmter Punkte / Landmarken. Die Lokalisation solcher Punkte in einem gegebenen Bild erfolgt mit Hilfe des Bündelgraphen. Jedoch existiert für jede spezifische Personengruppe (z.B. Personen gleichen Geschlechts) ein darauf abgestimmter Bündelgraph (welcher die Eigenschaft hat, die gesuchten Landmarken möglichst genau zu bestimmen). Ein solcher Bündelgraph verfügt an seinen Knoten über so genannte Texturdeskriptoren, welche benutzt werden um ein Gesicht innerhalb eines unbekannten Bildes lokalisieren zu können. In dieser Arbeit wird die Redundanz dieser Deskriptoren untersucht, bzw. es wird der Frage nachgegangen, wieviele und welche dieser Deskriptoren nötig sind, um eine gute Lokalisierung der Landmarken zu ermöglichen.

2. Das menschliche Sehen

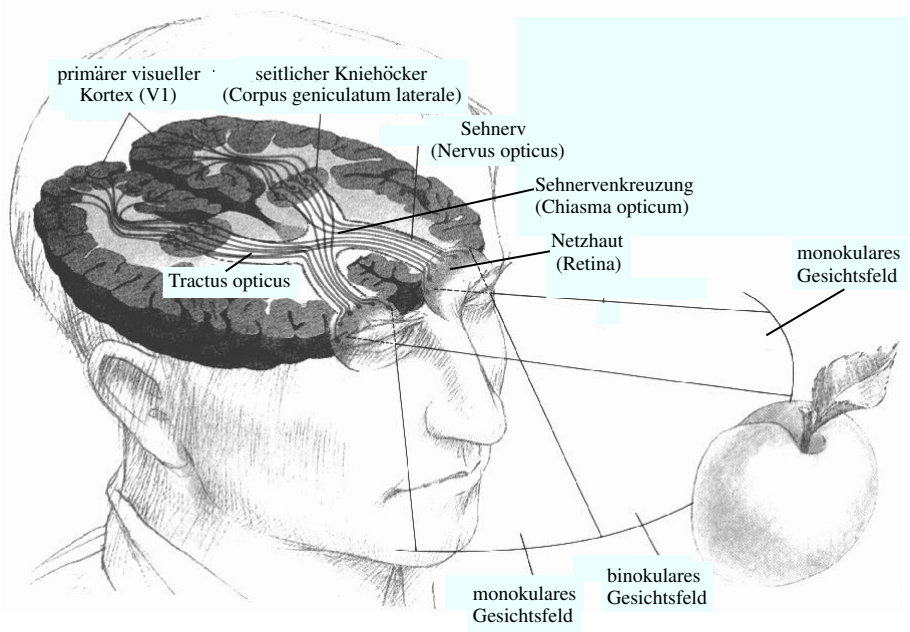


Abb. 6: Prinzipieller Aufbau des menschlichen Sehapparates

In diesem Abschnitt soll der Aufbau des menschlichen Sehapparates und seine Funktion näher erläutert werden. Bei Betrachtung der Abbildung 6 wird der physische Aufbau deutlich. Die Retina leitet über den Sehnerv ihre Signale weiter an die seitlichen Kniehöcker ”corpus geniculatem laterale ” (im weiteren mit CGL bezeichnet). Bevor die Informationen der Retina jedoch dort ankommen, passieren sie die Sehnervenkreuzung ”chiasma optikum”. Diese Stelle erfüllt die Funktion der Signalaufteilung zu jedem Kniehöcker hin, d.h. jeder Sehnerv wird mit jedem Kniehöcker verbunden (wie eine Weiche die Signale einer Quelle an 2 Senken weiterleitet). Die Funktionen des CGL sind nach [5],[12] u.a. das Zusammenführen der Informationen beider Augen, Aktivitätsregelung der visuellen Verbindung zum visuellen Cortex während Schlaf- und Wachphase sowie Kontrolle der kortikalen Verbindung. Der CGL kann weiterhin als eine Weiterleitungsstation angesehen werden welche die Informationen der Augen an den primären visuellen Cortex ”V1” weiterleitet, dort findet die eigentliche Verarbeitung der Informationen statt. Da das stereoskopische Sehen für die eigentliche Objekterkennung keine signifikante Rolle spielt, wird nicht näher darauf eingegangen bzw. es soll nur 1 Auge und dessen Informationsfluss über den CGL betrachtet werden.
Beginnen wir mit der Retina. Zum besseren Verständnis ihrer Funktion soll zuerst der Aufbau einer Nervenzelle bzw. eines Neurons näher erläutert werden. In Abbildung 7 sind 2 miteinander verbundene Nevenzellen abgebildet. Jede Nervenzelle besteht hauptsächlich aus einem Zellkern, Dendriten (Ausläufer von Nervenzellen) und einem Axon. Die Dendriten dienen dem Signalempfang bzw. bilden den synaptischen Eingang, der Zellkern entscheidet wann eine Antwort erfolgen soll. Mit anderen Worten gesagt, es existiert ein ”Schwellwert” in jeder Zelle, bei dessen Überschreitung eine Antwort über das Axon erfolgt. Es wird auch von einem Netzwerk gesprochen, da ein Axon mit den Dendriten weiterer Zellen verbunden sein kann, diese Verbindungsstellen werden als Synapsen bezeichnet (nach dem griechischen Wort für ”Vereinigung”).
Mit der Erläuterung der wichtigsten Begriffe soll nun die Retina hinsichtlich ihrer Funktion genauer erklärt werden. An dieser Stelle sei bereits erwähnt, dass der Aufbau der Retina nachfolgend stark idealisiert betrachtet wird (für genauere Erläuterungen siehe [14]). Die Netzhaut bildet eine Fläche von Stäbchen und Zapfen (welche als Rezeptoren bezeichnet werden), diese sind wiederum mit so genannten Ganglienzellen verbunden. Stäbchen dienen der Vermittlung von Helligkeitswerten und Zapfen dem Farbsehen. Ein Ganglion ist eine

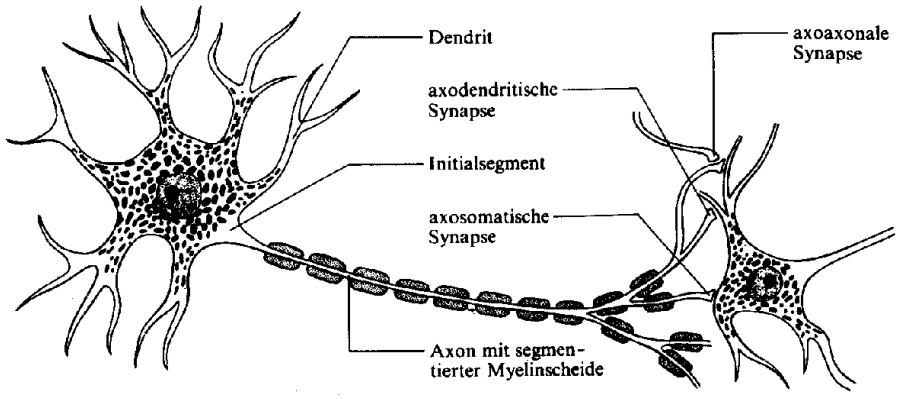


Abb. 7: 2 verbundene Nervenzellen

Ansammlung von Nervenzellen welche als Ganglienzellen bezeichnet werden. Jedoch besteht keine 1:1 Verbindung zwischen jeweils einem Rezeptor und einer Ganglienzelle. Mit anderen Worten ausgedrückt bedeutet dies, dass Ganglienzellen der Retina durchaus von mehreren Rezeptoren beeinflusst werden. Der Bereich von zusammenhängenden Rezeptoren durch welchen eine Zelle stimuliert werden kann, wird als rezeptives Feld der Zelle bezeichnet. Rezeptive Felder nehmen hinsichtlich ihrer Größe vom Zentrum der Retina nach Außen hin zu, somit existieren im Zentrum Felder mit der kleinsten Auflösung.
Unter dem Begriff der Stimulation soll nicht verstanden werden, dass eine Stimulation nur antwort-fördernd wirken kann. Es hängt von der stimulierten Zelle ab, wie diese auf eine Erregung an ihren Dendriten reagiert. Im Bereich der Neurobiologie existieren die Worte "Inhibition" und "Exhibition", der Begriff der Inhibition bezeichnet eine hemmende Stimulation der Nervenzelle wobei Exhibition eine anregende Stimulation kennzeichnet.
Rezeptive Felder retinaler Ganglienzellen stellen keine disjunkten Mengen von Zellen dar, d.h. rezeptive Felder können sich durchaus überlappen, weiterhin sind diese Felder in ihrer Form kreissymmetrisch mit einem inhibitorischen sowie exhibitorischen Bereich. An dieser Stelle soll der Begriff des rezeptiven Profils([5]) erläutert werden, darunter wird eine mathematische Funktion verstanden welche, das rezeptive Feld einer Zelle nachbildet. Anders ausgedrückt kann mit Hilfe des Profils die Stimulation einer Zelle, bei Erregung beliebiger Stellen ihres rezeptiven Feldes, untersucht werden. Das rezeptive Feld einer retinalen Ganglienzelle kann durch

$$\gamma(\vec{x}) = \sigma_- \exp(-\frac{\vec{x}^T\vec{x}}{2\sigma_+}) - \sigma_+ \exp(-\frac{\vec{x}^T\vec{x}}{2\sigma_-}) \tag{1}$$

beschrieben werden ($\vec{x} = (x, y)^T$ stellt den Positionsvektor dar). Zur Erklärung bzw. zum besseren Verständnis von σ_+ und σ_-, müssen die Begriffe der rezeptiven "on-center"- bzw "off-center"-Felder betrachtet werden. Unter einem on-center-Feld wird ein rezeptives Feld verstanden, welches in seinem Zentrum einen exhibitorischen Bereich aufweist und auf der verbleibenden Fläche einen inhibitorischen, off-center-Felder verhalten sich genau umgekehrt. Die Parameter σ_+ und σ_- legen fest ob es sich um ein off-center- bzw. on-center-Feld handelt, falls $\sigma_+ < \sigma_-$ so handelt es sich um ein on-center-Feld und für $\sigma_+ > \sigma_-$ um ein off-center-Feld (siehe Abbildungen 8 und 9).
Bei einer Intensitätsverteilung (im Sinne eines Bildes) I kann die Stimulation S der Zelle im Punkt $\vec{x}$ ausgedrückt werden als eine Faltung der Form:

$$S(\vec{x}) = (\gamma * I)(\vec{x}) = \int \gamma(\vec{\tilde{x}}) I(\vec{\tilde{x}} - \vec{x}) d\vec{\tilde{x}} \,. \tag{2}$$

Jeder Funktionswert $\gamma(\vec{x})$ kann als ein Faktor für den im selben Punkt liegenden Intensitätswert betrachtet werden. Erst die Summe aller gewichteten Intensitäten innerhalb des rezeptiven Feldes entscheidet, ob eine Zelle antwortet bzw. ob der Schwellwert für eine Antwort überschritten wird. Die Antworten der Ganglienzellen

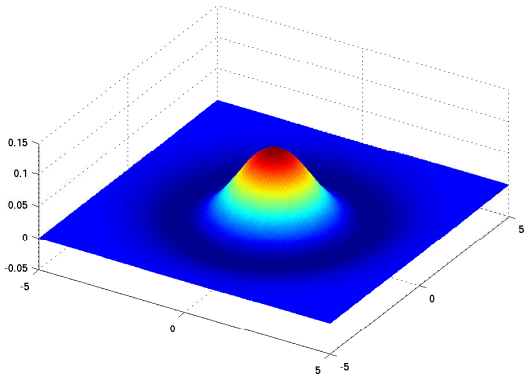


Abb. 8: on-center-Feld $\sigma_+ = 2.0, \sigma_- = 2.1$

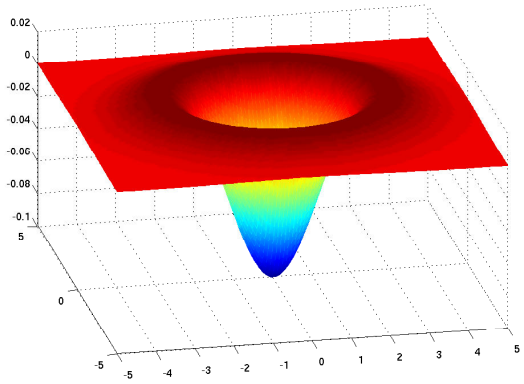


Abb. 9: off-center-Feld $\sigma_+ = 2.1, \sigma_- = 2.0$

werden an den CGL weitergeleitet, hier sei noch erwähnt, dass die Verbindung zwischen Retina und CGL bzw. CGL und V1 retinotop ist. Mit retinotop wird die Eigenschaft einer neuronalen Verbindung bezeichnet, die Aktivitätsmuster Topologie erhaltend abzubilden (benachbarte Erregungsmuster in der Retina werden auf benachbarte Neuronen im CGL abgebildet). Da die weiteren Funktionen des CGL hier nicht näher betrachtet werden, soll diesen betreffend davon ausgegangen werden, dass er die retinale Aktivität einfach an den primären visuellen Cortex weiterleitet.
Der Cortex besitzt besitzt in seiner Eingangsschicht Zellen die den retinalen Ganglienzellen sehr ähnlich sind. Es existieren natürlich verschiedene Zelltypen in dieser Schicht, da jedoch nur die so genannten "einfachen Zellen" für die in dieser Arbeit beschriebenen Verfahren wichtig sind, soll auch nur auf diese eingegangen werden. Eine einfache Zelle besitzt nach [12] folgende Eigenschaften:

1. Linearität)
Wird die Zelle als System betrachtet, so ist dieses in guter Näherung linear.

2. Zulässigkeit)
Es erfolgt keine Antwort auf räumlich konstante Stimulation.

3. Lokalisierbarkeit im Ortsraum)
Außerhalb des rezeptiven Feldes erfolgt keine Stimulation.

4. Orientierungsabhängigkeit)
Eine Antwort erfolgt selektiv auf die Orientierung einer Kante bzw. eines Balkens im rezeptiven Feld. Jede Zelle besitzt eine Vorzugsrichtung, falls Kante und Vorzugsrichtung orthogonal zueinander stehen, erfolgt keinerlei Stimulation.

5. Lokalisierbarkeit im Frequenzraum)
Wird die Zelle mit einer sinusförmigen Intensitätsverteilung stimuliert, so zeigt sich, dass sie eine bevorzugte Ortsfrequenz besitzt.

Das Profil einer einfachen Zelle wird nach [12] mit einem Gabor-Wavelet (siehe Kapitel 6) der Form

$$\psi(\vec{x}) = \frac{\vec{k}^T\vec{k}}{\sigma^2}\exp(-\frac{\vec{k}^T\vec{k}\vec{x}^T\vec{x}}{2\sigma^2})\left[\exp(i\vec{k}^T\vec{x}) - \exp(-\frac{\sigma^2}{2})\right] \tag{3}$$

beschrieben. Auf Gabor Wavelets sowie die oben genannten Forderungen wird im erwähnten Abschnitt näher eingegangen. Ein weiterer Unterschied des V1 zum CGL besteht darin, dass die Verbindung zwischen beiden zwar retinotop ist, jedoch kann ein retinales Aktivitätsmuster im Cortex mehrmals abgebildet werden.

In [12] und [5] wird der Begriff der Hypersäule erwähnt, darunter wird eine Menge von Zellantworten bestimmter Zellen verstanden. Diese besitzen ein gleich großes rezeptives Feld und sind auf dem selben Bereich der Retina zentriert, jedoch können sie sich hinsichtlich der Orientierung unterscheiden. Die Antworten solcher Zellmengen werden als Merkmalsvektoren bezeichnet.

3. Fourier Analysis

3.1. Fourier-Reihen

Zum besseren Verständnis der Fourier-Transformation, beginnt dieses Kapitel zunächst mit dem Thema der Fourier-Reihen. Aufbauend auf diesen, sollen die nötigen Grundlagen zum Verständnis der Fourier-Transformation (sowie darauf basierender Verfahren) geschaffen werden. Auf tiefgehende Konvergenzbeweise wird soweit wie möglich verzichtet und stattdessen auf entsprechende Literatur verwiesen. Die in diesem Kapitel vermittelten Grundlagen werden immer für den 1-dimensionalen Fall erklärt, sind jedoch analog auf höhere Dimensionen übertragbar. Des Weiteren wird bei Integralen sofern keine Integrationsgrenzen angegeben werden, das Integral über ganz $\mathbb{R}$ gemeint, d.h. $\int f(t)\,dt := \int_{-\infty}^{\infty} f(t)\,dt$

3.1.1. Orthogonalität / Orthonormalität

Für das elementare Verständniss der nachfolgenden Transformationen sind die Begriffe der Orthogonalität bzw. Orthonormalität sehr wichtig. Seien $u, v \in \mathbb{R}^n$, $n \in \mathbb{N}$, so werden diese als orthogonal zueinander bezeichnet falls $< u, v >= u^T v = 0$. Der Ausdruck $< u, v >$ stellt das innere bzw. skalare Produkt der Vektoren dar. Eine weitere Berechnungsvorschrift für das Skalarprodukt lautet

$$< u, v >= ||u||_2 \cdot ||v||_2 \cos(\alpha)\ . \tag{4}$$

Mit anderen Worten bedeutet dies, dass falls $||u||_2 \neq 0$ und $||v||_2 \neq 0$, die Gleichung $< u, v >= 0$ nur erfüllt sein kann falls u und v senkrecht aufeinander stehen (d.h. $\cos(\alpha) = 0$). In Abbildung 11 wird diese Berechnung veranschaulicht, v' stellt die vektorielle Komponente von v in Richtung von u dar. Zur Bestimmung des Skalarproduktes werden nun die Längen der Vektoren u, v' miteinander multipliziert. Betrachten wir den konkreten Fall mit $u = (1, 0)^T$, $v = (0, 1)^T$, das Skalarprodukt von u und v ist $< u, v >= (1, 0)(0, 1)^T = 1 \cdot 0 + 0 \cdot 1 = 0$ bzw. $< u, v >= 1 \cdot 1 \cdot \cos(\frac{\pi}{2})$ (Abb.(10)). Somit sind beide Vektoren orthogonal zueinander, dies ist analog auf höhere Dimensionen übertragbar.

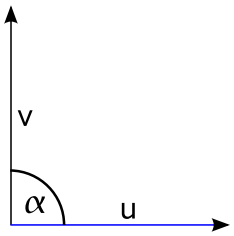


Abb. 10: Orthogonale Vektoren

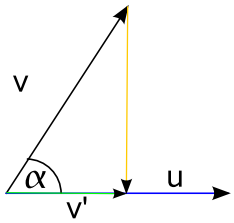


Abb. 11: Skalarprodukt

Falls zusätzlich $||u||_2 = 1$, $||v||_2 = 1$ gilt, so werden diese Vektoren als orthonormal zueinander bezeichnet. Die im obigen Beispiel erwähnten Vektoren sind ebenfalls orthonormal. Der Begriff der Orthonormalität ist ein fundamentales Element der Fourier Analysis.

3.1.2. Vektorräume

Sei V eine nicht leere Menge, so wird eine additive abelsche Gruppe $G = (V, +, 0)$ über einem Körper K als Vektorraum bezeichnet sofern auf ihr eine Skalarmultiplikation $\circ : K \times V \rightarrow V$ definiert ist. Vereinfacht gesagt bedeutet dies, dass V eine Menge darstellt, deren Elemente addiert und mit Elementen von K multipliziert werden können. Eine Menge U für welche gilt $U \subseteq G$ wird als Vektorteilraum bzw. Vektorunterraum

bezeichnet, sofern folgende Aussagen zutreffen:

$$u, v \in U \rightarrow (u + v) \in U \ \ , \forall u, v \in U \tag{5}$$

$$\lambda \in K, u \in U \rightarrow \lambda \cdot u \in U \ \ , \forall \lambda \in K, \ \forall u \in U \ . \tag{6}$$

Hieraus folgt, dass U hinsichtlich der Multiplikation und Addition abgeschlossen ist. Diese Definition eines Vektorraumes soll hier nicht weiter vertieft werden.
Ein weiterer Begriff der linearen Algebra ist die Vektorraumbasis. Eine Menge $V' \subseteq G$ stellt eine Vektorraumbasis dar sofern

$$\sum_{i=0}^{n} \alpha_i \cdot v_i' = 0 \Rightarrow \forall \alpha_i : \alpha_i = 0 \ \ , \forall v_i' \in V' \ , \ \ \forall \alpha_i \in K \ , \ \ i = 1, ..., n \tag{7}$$

gilt. Eine Interpretation obiger Aussage ist, dass sich kein Element aus V' als eine Linearkombination der übrigen Elemente darstellen lässt. Alle Elemente der Vektorraumbasis werden als Basisvektoren bezeichnet, durch Bildung einer Linearkombination dieser Vektoren lassen sich alle übrigen Elemente des Vektorraumes darstellen.
Zur Veranschaulichung eines Vektorunterraums betrachten wir folgende Mengen $V = \{(x_1, x_2)^T | \ x_1, x_2 \in \mathbb{R}\}$ und $V' = \{(x_1, x_2)^T | \ x_1, x_2 \in \mathbb{R}, \ x_1 = x_2\}$. Es ist zu erkennen, dass V' eine Teilmenge des Vektorraumes V ist. Versuchen wir nun eine Vektorraumbasis für V zu finden, d.h. eine Menge welche ausschließlich voneinander linear unabhängige Vektoren beinhaltet. Um diese Frage zu klären, muss V' mit Hilfe von 7 geprüft werden. Betrachten wir hierfür 3 Elemente von V, $(0, 1), (1, 0), (-2, -2)$. Es stellt sich heraus, dass

$$(-2, -2)^T = -2 \cdot (1, 0)^T + -2 \cdot (0, 1)^T \tag{8}$$

gilt. Mit anderen Worten, der Vektor $(2, 2)^T$ kann als eine Linearkombination von $(1, 0)^T$ und $(0, 1)^T$ dargestellt werden (siehe Abb. 12).

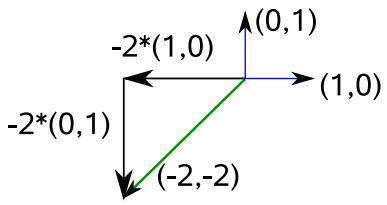


Abb. 12: Linearkombination von 2 Vektoren

Für die Vektoren$(1, 0)^T$ und $(0, 1)^T$ gibt es jedoch keine solche Möglichkeit, somit kann für V die Vektorraumbasis $V_B = \{(0, 1)^T, (1, 0)^T\}$ definiert werden.
Dies ist ein wichtiger Schritt bei der Analyse eines Vektorraumes, jedoch stellt sich als nächstes die Frage nach den Koeffizienten innerhalb der Linearkombination 7.
Bezeichnen wir nun die Basisvektoren mit $v_{B_1} = (1, 0)^T$, $v_{B_2} = (0, 1)^T$ und versuchen die Koeffizienten in der Gleichung $\alpha_1 \cdot v_{B_1} + \alpha_2 \cdot v_{B2} = v$ zu bestimmen. Hierbei hilft die Eigenschaft der normierten (orthonormalen) Vektorraumbasis V_B. Es gilt

$$v = \sum_{i=1}^{N} \alpha_i \cdot v_{B_i} \tag{9}$$

mit

$$< v_{B_i}, v_{B_j} >= \begin{cases} 0 & i \neq j \\ 1 & i = j \end{cases} \ . \tag{10}$$

Unter Ausnutzung von 10 ergibt sich der Zusammenhang

$$< v, v_{B_j} > = \sum_{i=1}^{N} \alpha_i \cdot < v_{B_i}, v_{B_j} > \tag{11}$$

$$= \alpha_j \ . \tag{12}$$

Auf diese Art können sämtliche Koeffizienten der Linearkombination bestimmt werden. Zur Verdeutlichung sei das vorherige Beispiel betrachtet, d.h. mit $v = (-2, -2)^T$ ergeben sich die Koeffizienten

$$\alpha_1 = < (-2, -2)^T, (1, 0)^T > \tag{13}$$

$$= -2 \tag{14}$$

$$\alpha_2 = < (-2, -2)^T, (0, 1)^T > \tag{15}$$

$$= -2 \ . \tag{16}$$

Diese Berechnungsvorschrift gilt analog für höhere Dimensionen.
Abschliessend zu den Grundlagen noch eine kurze Zusammenfassung. Ein Vektorraum stellt eine Menge dar, deren Elemente bestimmte Axiome erfüllen. Elemente des Vektorraumes werden als orthogonal zueinander bezeichnet, sofern sie senkrecht aufeinander stehen, diese können in einer als Vektorraumbasis bezeichneten Teilmenge des Vektorraumes zusammengefasst werden. Jedes Element des Vektorraumes exklusiv der Vektorraumbasis, kann durch eine Linearkombination der Basisvektoren "konstruiert" werden. Falls alle Basisvektoren orthonormiert sind, so wird von einer orthonomalen Vektorraumbasis gesprochen. Zur Bestimmung der Koeffizienten in der Linearkombination ist eine orthonormale Basis sehr vorteilhaft (siehe Gleichung (12)).

3.1.3. Fourier-Reihen

Wie zuvor gezeigt kann ein Vektorraum eine Menge von Vektoren $\in \mathbb{R}^n$ beinhalten, d.h. nicht, dass ein Vektorraum ausschliesslich Elemente dieser Form beinhalten muss.
Sei M eine nicht leere Menge, $\mathbb{K}$ ein Körper und $Abb(M, \mathbb{K})$ die Menge aller Funktionen von M nach $\mathbb{K}$, dann wird die Abbildungsmenge $Abb(M, \mathbb{K})$ zu einem Vektorraum über $\mathbb{K}$, falls für die Funktionen $f, g \in Abb(M, \mathbb{K})$ und einen Skalar $\lambda \in \mathbb{K},\ t \in M$

$$(f + g) : M \to \mathbb{K}, t \mapsto (f(t) + g(t)) \tag{17}$$

$$\lambda f : M \to \mathbb{K}, t \mapsto \lambda f(t) \tag{18}$$

gilt. Solche Vektorräume werden auch als Funktionenräume bezeichnet. Der für diese Arbeit wichtige Funktionenraum ist der $L^2(\mathbb{R})$ (d.h. ein Raum über dem Körper der reellen Zahlen). Der L^2 stellt einen so genannten Hilbertraum dar. Vereinfacht gesagt ist ein Hilbertraum ein linearer Vektorraum über dem Körper der komplexen bzw. reellen Zahlen, in dem ein Skalarprodukt definiert ist und der in Bezug auf die durch dieses Skalarprodukt erzeugte Norm vollständig ist (für tiefer gehende Informationen hierzu sei auf [10] verwiesen). Die Elemente des L^2 sind Funktionen für welche das Integral

$$\int |f(t)|^2 dt =: ||f||^2 \tag{19}$$

endlich ist.
Weiterhin ist das Skalarprodukt zweier Funktionen $f, g \in L^2$ folgenderweise definiert

$$< f, g >= \int_{-\infty}^{\infty} f(t) \cdot \overline{g(t)} dt \ . \tag{20}$$

Die Norm für Funktionen $f \in L^2$ mit D als Definitionsbereich lautet

$$||f||_{L^2(D)} = \left(\int_D |f(t)|^2 dt \right)^{\frac{1}{2}} \ , \tag{21}$$

somit ist L^2 ein Hilbertraum.
Fourier-Reihen betreffen 2π periodische Funktionen der Form

$$f : \mathbb{R} \to \mathbb{C}, \quad f(t + 2\pi) \equiv f(t) \,. \tag{22}$$

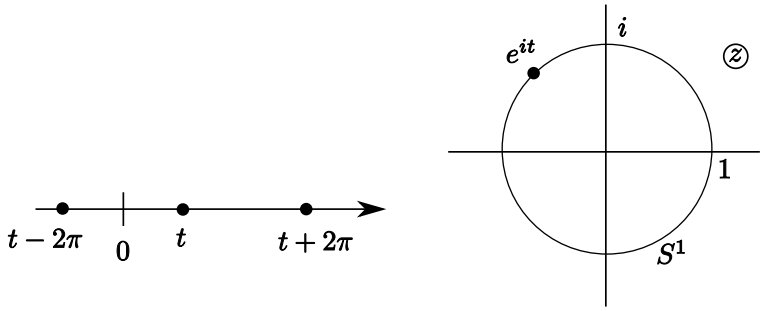


Abb. 13: Funktionen auf dem Einheitskreis

Eine weitere Schreibweise für solche Funktionen ist nach [2] $f : \mathbb{R}/2\pi \to \mathbb{C}$. Der Einheitskreis S^1 in der komplexen z-Ebene kann als Definitionsbereich für diese angesehen werden (siehe Abb.(13)). Da die Funktionen 2π-periodisch sind, besitzen sie unendlich viele modulo 2π äquivalente Punkte. D.h. es existieren unendlich viele Punkte $t + 2k\pi,\ k \in \mathbb{Z}$ für welche gilt: $f(t + 2k\pi) = f(t)$. Auf S^1 erscheinen diese Punkte jedoch als einziger Punkt e^{it}. Betrachten wir hierzu die Potenzfunktionen

$$\chi_k : S^1 \to S^1, \quad z \mapsto z^k \,. \tag{23}$$

Werden diese durch die Variable t ausgedrückt, so resultieren daraus die periodischen Grundfunktionen bzw. die reinen Schwingungen

$$e_k : \mathbb{R} \to \mathbb{C}, \quad t \mapsto e^{ikt} \quad (k \in \mathbb{Z}) \,. \tag{24}$$

Somit sind die Funktionen $e_k \in L^2(\mathbb{R})$, weiterhin stellt für diese

$$< f, g >= \frac{1}{2\pi} \int f(t) g(t)\, dt \tag{25}$$

ein natürliches Skalarprodukt dar.
Analog zum vorherigen Beispiel betrachten wir nun einen konkreten Fall, jedoch definieren wir eine neue Vektorraumbasis $V_B := \{\mathbf{e}_k\}$ mit $\mathbf{e}_k : \mathbb{R} \to \mathbb{C}, t \mapsto e^{ikt}$ und $k \in \mathbb{Z}$. Die Orthonormalität dieser Basis ist gewährleistet da

$$\begin{aligned}
< \mathbf{e}_m, \mathbf{e}_n > &= \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{imt} \cdot \overline{e^{int}} dt && (26)\\
&= \frac{1}{2\pi} \int_{0}^{2\pi} e^{imt} \cdot e^{-int} dt && (27)\\
&= \frac{1}{2\pi} \int_{0}^{2\pi} e^{imt - int} dt && (28)\\
&= \frac{1}{2\pi} \int_{0}^{2\pi} e^{it(m-n)} dt && (29)\\
&= \begin{cases} 1 & m = n \\ \frac{1}{2\pi(m-n)} e^{it(m-n)} \Big|_0^{2\pi} = 0 & m \neq n \end{cases} && (30)
\end{aligned}$$

erfüllt ist. Somit bildet V_B eine aus Funktionen bestehende Basis, analog zum Beispiel mit reellen Zahlen. Ebenfalls analog zum vorhergehenden Beispiel lässt sich eine Funktion durch eine Linearkombination der jeweiligen Basis-Elemente darstellen. In mathematischer Schreibweise bedeutet dies

$$f \rightsquigarrow \sum_{k=-n}^{n} c_k \mathbf{e}_k \ . \tag{31}$$

Der zuvor erwähnte Begriff der Darstellung sei hierbei, wie sich anhand der Gleichung erkennen lässt, nicht als Identität der Funktion zu verstehen. Der Ausdruck (31) stellt eine Approximation der Funktion f dar, welche im Grenzwert $n \to \infty$ gegen f geht. Betrachten wir die Partialsumme

$$s_n = \sum_{k=-n}^{n} c_k \mathbf{e}_k \qquad \text{bzw.} \qquad s_n(t) = \sum_{k=-n}^{n} c_k e^{ikt} \tag{32}$$

und erinnern uns an das Beispiel mit reellen Zahlen. Wie zuvor (12) lassen sich die Koeffizienten in der Summe unter Ausnutzung der Orthogonalität von V_B leicht berechnen.

$$c_k := < f, \mathbf{e}_k > = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) e^{ikt} dt \ . \tag{33}$$

Diese Koeffizienten werden als Fourier-Koeffizienten bezeichnet, wobei jeweils der k-te Koeffizient die k-te Koordinate bezüglich der Basis e_k darstellt. Der Vorgang der Koeffizientenbestimmung wird auch als Analyse einer Funktion bezeichnet. Je mehr Basis-Funktionen hinzukommen, desto mehrdimensionaler wird der aufgespannte Unterraum

$$U_n = span(e_{-n}, ..., 1, ..., e_n) \ . \tag{34}$$

Somit stellt (32) eine Orthogonalprojektion von f auf U_n dar. Geometrisch betrachtet ist s_n der f am nächsten gelegene Punkt von U_n. Für die Bestimmung des Abstandes im $L^2(\mathbb{R})$ wird die Rechenvorschrift

$$d(f, g) := ||f - g|| = \left(\int |f(t) - g(t)|^2 dt \right)^{\frac{1}{2}} \tag{35}$$

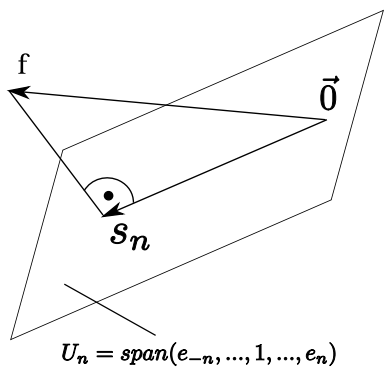


Abb. 14: Projektion einer Funktion f auf den aufgespannten Unterraum U_n

verwendet. Für beliebige Funktionen $f, g \in L^2$ gilt die Parsevalsche Formel

$$\sum_{k=-\infty}^{\infty} c_{k,f} c_{k,g} = < f, g > \tag{36}$$

(wobei es sich bei $c_{k,f}$ und $c_{k,g}$ um die Fourier-Koeffizienten der Funktionen f bzw. g handelt), insb. gilt weiterhin $\sum_{k=-\infty}^{\infty} |c_k|^2 = ||f||^2$. Die Frage nach der Genauigkeit der Darstellung einer Funktion f durch ihre entsprechende Fourier-Reihe soll nun geklärt werden. Für die Partialsumme (32) folgt nach dem Satz von Pythagoras:

$$||f - s_n||^2 = ||f||^2 - ||s_n||^2 = ||f||^2 - \sum_{k=-n}^{n} |c_k|^2 \,. \tag{37}$$

Wie in [2] nachzulesen ist, gilt $\lim_{n\to\infty} ||f - s_n||^2 = 0$. Somit konvergiert die Fourier-Reihe einer Funktion $f \in L^2(\mathbb{R})$ im Sinne der L^2 Metrik gegen f. Für tiefer gehende Informationen diesbezüglich sei auf [2] verwiesen.
Nicht 2π-periodische Funktionen sind ebenfalls durch eine Fourier-Reihe darstellbar. Sei hierzu $f : \mathbb{R} \to \mathbb{C}$ eine T-periodische Funktion mit $T > 0$ und $\int_0^T |f(t)|^2 dt < \infty$. So gilt

$$f(t) \rightsquigarrow \sum_{k=-\infty}^{\infty} c_k e^{2k\pi it/T} \tag{38}$$

und für die Koeffizienten

$$c_k = \frac{1}{T} \int_0^T f(t) e^{-2k\pi it/T} \, dt \,. \tag{39}$$

Oft wird der Faktor $\frac{2\pi}{T}$ im Exponenten der komplexen Exponentialfunktion durch den Bezeichner $\omega := \frac{2\pi}{T}$ ausgedrückt. An dieser Stelle betrachten wir ein konkretes Beispiel, es soll eine Funktion

$$f(t) := \begin{cases} t & t \in [0,1] \\ 0 & t \in (1,2] \end{cases} \tag{40}$$

mit Hilfe ihrer Fourier-Reihe darstellt werden. Hierzu werden zunächst die Fourier-Koeffizienten c_k nach oben angegebener Formel (33) bestimmt und anschließend eine Synthese der Funktion durchgeführt.

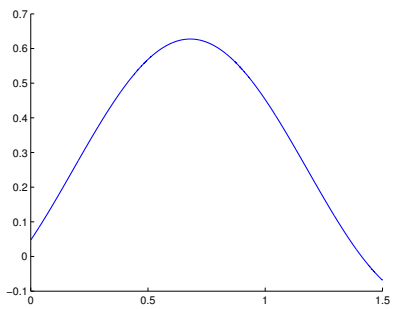


Abb. 15: Synthese einer Funktionmit 1 Summand

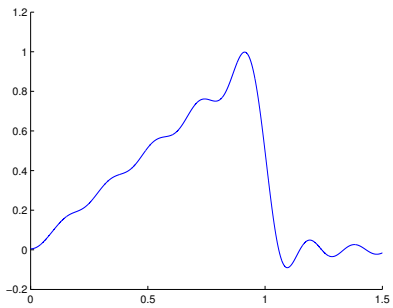


Abb. 16: Synthese einer Funktion mit 10 Summanden

Anhand der Abbildungen (15) bis (19) ist die "Qualität" der Rekonstruktion (Synthese) einer Funktion deutlich zu erkennen. Bei genauerer Beobachtung der Abbildungen wird eine gewisse Ungenauigkeit an scharfen Übergängen deutlich, es scheint dort zu einer Häufung von "Überschwingern" zu kommen. Dies ist das sogenannte Gibbsche Phänomen, welches an unstetigen Stellen der zu rekonstruierenden Funktion auftritt. Der Funktionswert an diesen Stellen besitzt eine Abweichung von ca. 18%.

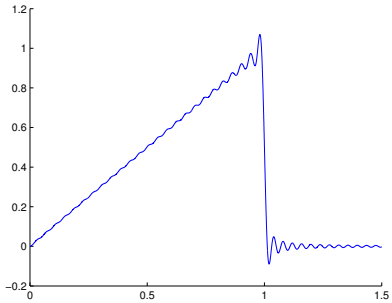


Abb. 17: Synthese einer Funktion mit 50 Summanden

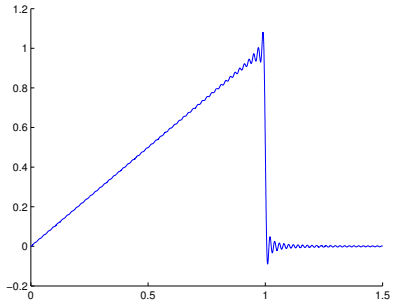


Abb. 18: Synthese einer Funktion mit 100 Summanden

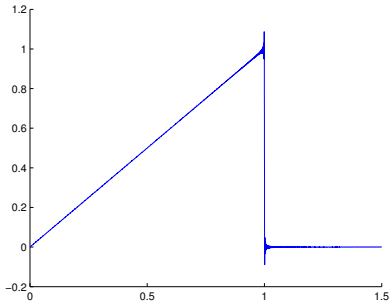


Abb. 19: Synthese einer Funktion mit 1000 Summanden

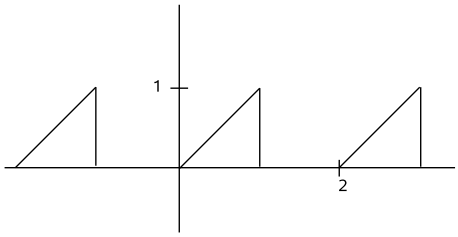


Abb. 20: Original Funktion

3.2. Fourier-Transformation

Bisher wurden Funktionen aus dem L^2 betrachtet. Die ist jedoch nicht der einzige Funktionenraum welcher für die Bestimmung von Fourier-Reihen wichtig ist. Der Vollständigkeit halber und da für den Nachweis der Heisenbergschen Unschärferelation benötigt, sollen an dieser Stelle 3 Funktionenräume kurz erläutert werden. In Abb.(21) werden die Funktionenräume L^1,L^2 und S dargestellt, hierbei beinhalten diese Räume Funktionen der Form $f : \mathbb{R} \to \mathbb{C}$. Diese Funktionen werden von nun an als Funktionen im Zeitbereich bzw. als Zeitsignale bezeichnet. Weiterhin gibt die Abbildung Aufschluss über die Inklusion der Räume. Der Raum L^1 beinhaltet alle solche Funktionen für die das Integral

$$\int |f(t)|dt =: ||f||_1 \tag{41}$$

endlich ist. Analog hierzu besteht der Raum L^2 aus Funktionen für welche das Integral

$$\int |f(t)|^2 dt =: ||f||^2 \tag{42}$$

endlich ist. Mit S wird der sogenannte Schwartzsche Raum bezeichnet, er enthält beliebig oft differenzierbare Funktionen, deren Ableitungen mit $|t| \to \infty$ schneller als jede Potenz $\frac{1}{|t|^n}$ gegen 0 gehen, tiefer gehende Informationen hierzu sind in [2] zu finden.
Bisher wurden T-periodische Funktionen mit Hilfe der Fourier-Reihen dargestellt, wie verhält sich jedoch die Fourier-Reihe einer Funktion, für deren Periodendauer $T \to \infty$ gilt. Betrachten wir Gleichungen (38) und (39),

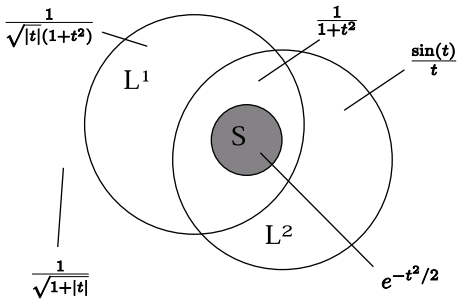


Abb. 21: Funktionenräume

und führen folgende Bezeichner ein

$$\xi_k = \frac{k}{T}, \text{ sowie } \quad \Delta\xi_k = \xi_{k+1} - \xi_k = \frac{k+1}{T} - \frac{k}{T} \,. \tag{43}$$

Die erwähnten Gleichungen lassen sich somit ein wenig umgeformt schreiben als

$$f(t) = \sum_{k=-\infty}^{\infty} c_k e^{2k\pi it/T} T \frac{1}{T} = \sum_{k=-\infty}^{\infty} (Tc_k) e^{2\pi it\xi_k} \frac{1}{T} \tag{44}$$

und

$$Tc_k = \int_{-T/2}^{T/2} f(t) e^{-2\pi it\xi_k} \, dt \,. \tag{45}$$

Da das Produkt Tc_k von T, f und ξ_k abhängig ist, fassen wir dieses als

$$Tc_k =: \widehat{f_T}(\xi_k) \tag{46}$$

zusammen und es ergibt sich

$$f(t) = \sum_{k=-\infty}^{\infty} \widehat{f_T}(\xi_k) e^{2\pi it\xi_k} \Delta\xi_k \,. \tag{47}$$

Nun folgt die Betrachtung für $T \to \infty$, hierbei geht $\widehat{f_T}(\xi_k)$ über in das Integral

$$\lim_{T\to\infty} \widehat{f_T}(\xi_k) = \int_{-\infty}^{\infty} f(t) e^{-2\pi it\xi} dt \tag{48}$$

und für die Summe analog

$$\lim_{T\to\infty} \sum_{k=-\infty}^{\infty} \widehat{f_T}(\xi_k) e^{2\pi it\xi_k} \Delta\xi_k = \int_{-\infty}^{\infty} \widehat{f}(\xi) e^{2\pi it\xi} d\xi \,. \tag{49}$$

An dieser Stelle sei nun $\widehat{f}$ eingeführt, diese Funktion wird als Fourier-Transformierte der Funktion f bezeichnet und ist nach [3] definiert durch

$$\widehat{f}(\omega) = \int_{-\infty}^{\infty} f(t) e^{-i\omega t} dt \,, \tag{50}$$

für die inverse Fourier-Transformation gilt entsprechend

$$f(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \widehat{f}(\omega) e^{it\omega} d\omega \,, \tag{51}$$

hierbei ist $\omega = 2\pi\xi$. In der Literatur ist die Definition für die Fourier-Transformation nicht eindeutig, je nach gewählter Normierung wird die Transformationsvorschrift mit einem entsprechenden Faktor versehen. Analog zu den Fourier-Koeffizienten bei der Fourier-Reihe von f, stellt $\widehat{f}(\xi)$ die komplexe Amplitude der Frequenz ξ dar.
Analog zu den Fourier-Reihen gilt auch hier für beliebige $f, g \in L^2$ die Formel von Parseval-Plancherel

$$< \widehat{f}, \widehat{g} >=< f, g > \text{ bzw. } \int \widehat{f}\ \overline{\widehat{g}} = \int f\ \overline{g}\,. \tag{52}$$

Insbesonders gilt dadurch

$$||\widehat{f}||^2 = ||f||^2 \text{ bzw. } \int |\widehat{f}(\xi)|^2 d\xi = \int |f(t)|^2 dt\,. \tag{53}$$

An dieser Stelle seien 2 Rechenregeln hinsichtlich der Fourier-transformation aufgeführt, welche für diese Arbeit von Bedeutung sind. Erstens sei die Linearität der Transformation erwähnt, seien $f_1, f_2 \in L^2$ und $k_1, k_2 \in \mathbb{R}$ so gilt

$$(k_1 f_1 + k_2 f_2)\widehat{} = k_1 \widehat{f_1} + k_2 \widehat{f_2} \tag{54}$$

und zweitens der Faltungssatz,

$$(f_1 * f_2)(t) = \mathcal{F}^{-1}(\widehat{f_1} \cdot \widehat{f_2})(t)\,. \tag{55}$$

In Worten bedeutet der Faltungssatz, dass die Faltung zweier Funktionen im Zeitbereich, einer Multiplikation der entsprechenden Fourier-Transformierten im Frequenzbereich entspricht.

3.3. Abtast-Theorem und Aliasing

Da letztenendes die entsprechenden Berechnungen zur Analyse und Synthese von Funktionen auf einem Rechner (mit endlicher Genauigkeit und in endlicher Zeit) ausgeführt werden sollen, drängt sich die Frage nach der Diskretisierung der beschriebenen Verfahren auf. Die zentrale Frage welcher in diesem Abschnitt nachgegangen wird, lautet "An wievielen Stellen muss eine Funktion ausgewertet werden, um eine Rekonstruierbarkeit dieser zu garantieren". Diese Frage wird durch Abb.(22) verdeutlicht. In dieser Abbildung sind 2 Sinusfunktionen mit unterschiedlicher Frequenzskalierung dargestellt, benennen wir diese Funktionen als $f(t) = \sin(\frac{2\pi}{5}t)$ (niedrige Frequenz) sowie $g(t) = \sin(2\pi t)$ (hohe Frequenz). Nehmen wir weiterhin an, dass wir den Funktionsverlauf nicht kennen und nur über die Abtastwerte verfügen, welche an den entsprechend gekennzeichneten Stellen aufgezeichnet wurden. Weiterhin sei g die gewünschte (zu untersuchende), unbekannte Funktion. Der Versuch, den Graphen der Funktion aus den Abtastwerten zu rekonstruieren, wird bei entsprechender Interpolation, anstelle von g den Verlauf von f liefern. Dieses Phänomen wird als Aliasing bezeichnet. Für weitere Interpretationen und Beispiele sei an dieser Stelle auf [8] verwiesen. Zur Vermeidung des Aliasing kann durch das Shannon'sche Abtast-Theorem eine Abtastrate bzw. Abtastfrequenz bestimmt werden welche diese Problematik umgeht.
Sei $f \in L^2(\mathbb{R})$, so ist das Resultat der Abtastung eine komplexe bzw. reelle Zahlenfolge $\mathbf{u} = (u_n)_{n\in\mathbb{Z}}$. Diese Folge besitzt die Eigenschaft $\sum_{n\in\mathbb{Z}} |u_n|^2 < \infty$, Folgen welche diese Eigenschaft erfüllen werden als quadratsummierbar bezeichnet. In [3] wird die Menge solcher Folgen durch $l^2(\mathbb{Z})$ ausgedrückt. Analog zu Funktionen aus Hilberträumen, gilt hier für das Skalarprodukt zweier Folgen $\mathbf{u}, \mathbf{v} \in l^2(\mathbb{Z})$

$$< \mathbf{u}, \mathbf{v} >= \sum_{n\in\mathbb{Z}} \overline{u_n} v_n\,. \tag{56}$$

Betrachten wir nun $\mathbf{u} = (f(n\Delta t))_{n\in\mathbb{Z}}$, so liegt die Vermutung nahe, dass die Funktion f aus ihren Abtastwerten rekonstruiert werden kann, sofern Δt nur hinreichend klein gewählt wird. D.h. falls die Abtastrequenz hoch genug ist um auch schnelle Änderungen im Funktionsverlauf erfassen zu können. Vor genauerer Betrachtung des Abtast-Theorems sei zunächst der Begriff einer bandbegrenzten Funktion erläutert.

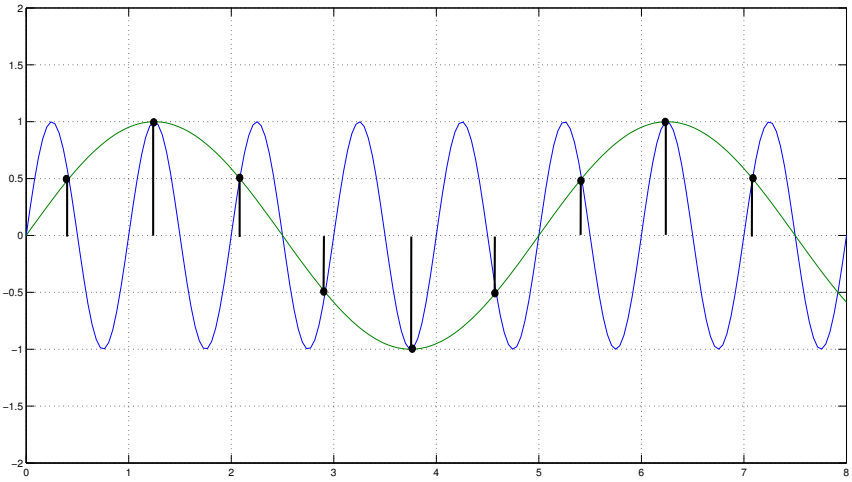


Abb. 22: Aliasing

Eine Funktion f heisst bandbegrenzt, falls es eine Grenzfrequenz ξ_g gibt mit $\widehat{f}(\xi) = 0$ für $|\xi| > \xi_g$, ξ_g wird auch als Nyquist-Frequenz bezeichnet. Somit lässt sich das Theorem folgenderweise formulieren:
Ist $f \in L^2(\mathbb{R})$ eine durch die Frequenz ξ_g bandbegrenzte Funktion, und gilt $\Delta t \leq \frac{1}{2\xi_g}$, so ist f aus den Abtastwerten $u_n = f(n\Delta t) \quad , n \in \mathbb{Z}$ exakt rekonstruierbar:

$$f(t) = \sum_{k\in\mathbb{Z}} f(k\Delta t)\frac{\sin(2\pi\xi_g(t-k\Delta t))}{2\pi\xi_g(t-k\Delta t)} \ . \tag{57}$$

In Worten bedeutet dies, dass die Abtast-Intervalle kleiner gleich dem Kehrwert der doppelten Grenzfrequenz sein müssen um Aliasing zu vermeiden. Aufgrund der Signifikanz des Abtast-Theorems soll an dieser Stelle der Beweis nach [3] geführt werden.
Nehmen wir $\Delta t = \frac{1}{2\xi_g}$ an, so lässt sich der entsprechende Nachweis durch Bestimmung der Fourier-Reihe zur Fouriertransformierten $\widehat{f}(\xi)$ für $-\xi_g \leq \xi \leq \xi_g$ führen

$$\widehat{f}(\xi) = \sum_{k\in\mathbb{Z}} c_k e^{ik\omega\xi} = \sum_{k\in\mathbb{Z}} c_k e^{ik\frac{2\pi}{2\xi_g}\xi} = \sum_{k\in\mathbb{Z}} c_k e^{i2\pi k\Delta t\xi} \ , \tag{58}$$

hierbei wurde ausgenutzt, dass die Periodendauer $2\xi_g$ beträgt (siehe Periodisierung von nicht 2π periodischen Funktion in (38)). Es folgt die Bestimmung der komplexen Koeffizienten c_k

$$c_k = \frac{1}{2\xi_g}\int_{-\xi_g}^{\xi_g} \widehat{f}(\xi)e^{ik\frac{\pi}{\xi_g}\xi}d\xi = \frac{1}{2\xi_g}\int_{-\infty}^{\infty} \widehat{f}(\xi)e^{ik\frac{\pi}{\xi_g}\xi}d\xi = \frac{1}{2\xi_g}f(-k\frac{1}{2\xi_g}) = \frac{1}{2\xi_g}f(-k\Delta t) \ . \tag{59}$$

Die Umformungen in diesem Schritt beruhten auf Gleichung (33) und (50) Mit Hilfe dieser Gleichung transformieren wir nun $\widehat{f}(\xi)$ zurück in den Zeitbereich

$$f(t) = \int_{-\infty}^{\infty} \widehat{f}(\xi)e^{i2\pi\xi t}d\xi = \int_{-\xi_g}^{\xi_g} \widehat{f}(\xi)e^{i2\pi\xi t}d\xi = \frac{1}{2\xi_g}\sum_{k\in\mathbb{Z}} f(-k\Delta t)\int_{-\xi_g}^{\xi_g} e^{i2\pi\xi t}e^{i2\pi k\Delta t\xi}d\xi \tag{60}$$

und bestimmen das Integral

$$\int_{-\xi_g}^{\xi_g} e^{i2\pi\xi t}e^{i2\pi k\Delta t\xi}d\xi = \int_{-\xi_g}^{\xi_g} e^{i2\pi\xi(t+k\Delta t)}d\xi = \frac{\sin(2\pi\xi_g(t+k\Delta t))}{\pi(t+k\Delta t)} \ . \tag{61}$$

Mit Anschliessender Substitution von $k = -k$ ergibt sich nach einsetzen (57).

3.4. Diskrete Fourier Transformation

Nachdem im vorherigen Abschnitt erörtert wurde, in welchen Abständen ein Signal abgetastet werden muss, um eine exakte Rekonstruierbarkeit zu gewährleisten, soll nun untersucht werden, wie weit die Fouriertransformation auf abgetastete Signale anwendbar ist. [19] führt hierbei 2 wichtige Punkte diesbezüglich auf

- Gleichung (50) gilt nur für kontinuierliche Signale
- Gleichung (50) setzt ein zeitlich unendlich ausgedehntes Signal voraus

Betrachten wir bezüglich des ersten Punktes die Folge von äquidistanten Abtastwerten des Signals f. D.h.

$$f_A(t) = f(t) \sum_{n \in \mathbb{Z}} \delta(t - n\Delta t) \,, \tag{62}$$

so ergibt sich für dessen Frequenzraumdarstellung nach Gleichung (50)

$$\widehat{f_A}(\omega) = \int f(t) \sum_{n \in \mathbb{Z}} \delta(t - n\Delta t) e^{-i\omega t} dt = \sum_{n \in \mathbb{Z}} f_A(n\Delta t) e^{-i\omega n \Delta t} \,. \tag{63}$$

Mit Hilfe dieser Transformation ergibt sich für ein unendlich ausgedehntes diskretes Zeitsignal, eine kontinuieriche Frequenzraumdarstellung mit periodischer Struktur. D.h.

$$\widehat{f_A}(\omega + \frac{2\pi}{\Delta t} k) = \widehat{f_A}(\omega) \ \ , \ k \in \mathbb{Z} \,. \tag{64}$$

Nach dem Abtast-Theorem muss $\widehat{f}$ auf einem endlichen Träger definiert sein, d.h. es muss eine Grenzfrequenz existieren. Die bisherigen Feststellungen lauten also:

- f ist im Ortsraum unendlich ausgedehnt und diskret
- $\widehat{f}$ ist im Frequenzraum lokalisiert jedoch kontinuierlich

Diese Aussage gilt analog falls Orts- und Frequenzraum vertauscht werden. Dieser Umstand lässt sich zur Formulierung der vollständigen Diskretisierung nutzen. Falls nun $\widehat{f}$ im Frequenzraum im Intervall $\Delta\omega$ abgetastet wird, so muss f im Ortsraum einen endlichen Träger besitzen.

$$\widehat{f_A}(\rho\Delta\omega) = \sum_{n \in \mathbb{Z}} f_A(n\Delta t) e^{-i\rho\Delta\omega n \Delta t} \,. \tag{65}$$

Da die Fourier-Transformation nach der Diskretisierung immer noch eine bijektive Abbildung darstellen soll, muss die Anzahl N der Abtastwerte im Orts- und Frequenzraum übereinstimmen. Somit folgt die Bedingung $\frac{2\pi}{\Delta x} = N\Delta\omega$.
Betrachten wir die Gleichung (39) für die komplexen Fourier-Koeffizienten leicht modifiziert, so lässt sich das Integral durch folgende Rechtecksumme approximieren

$$c_\rho \approx \gamma_\rho := \frac{1}{T} \sum_{n=0}^{N-1} f(n\Delta t) e^{-i\rho\omega n \Delta t} \Delta t \,. \tag{66}$$

Werden nun die folgenden Zusammenhänge verwendet, $T = N\Delta t,\ \omega T = 2\pi$, so ergibt sich daraus

$$\gamma_\rho = \frac{1}{N} \sum_{n=0}^{N-1} f(n\Delta t) e^{-i\rho 2\pi n \Delta t} = \frac{1}{N} \sum_{n=0}^{N-1} u_n e^{-i\rho \frac{2\pi}{N} n} = \frac{1}{N} \sum_{n=0}^{N-1} u_n w_N^{\rho n} \,, \tag{67}$$

mit $w_N := e^{-i\frac{2\pi}{N}}$ Daraus folgen die Gleichungen für die diskrete Fourier-Transformation

$$(DFT\ f)(\rho) = \breve{f}(\rho) = \frac{1}{N}\sum_{n=0}^{N-1} f(n) w_N^{\rho n} \tag{68}$$

mit $f_A(n) := f(t)\delta(t - n\Delta t)$ und die entsprechende Rücktransformation (inverse Diskrete Fourier-Transformation)

$$(IDFT\ \breve{f})(n) = f(n) = \frac{1}{N}\sum_{\rho=0}^{N-1} \breve{f}(n) w_N^{-\rho n}\ . \tag{69}$$

In der Praxis werden diese Gleichungen jedoch nicht in dieser Form verwendet, es existiert ein schneller Algorithmus zur Berechnung der DFT, die so genannte FFT (Fast Fourier Transformation). Der Algorithmus nutzt Symmetrien der Potenzen von w_N aus und ist besonders schnell, wenn N eine Potenz von 2 darstellt. Auf die Herleitung sowie den mathematischen Nachweis der Effektivität der FFT wird hier verzichtet und auf [19],[3],[17] und [8] verwiesen.

4. Heisenbergsche Unschärferelation

4.1. Lokalisierung im Frequenzraum

Einen für diese Arbeit sehr wichtigen Aspekt stellt die sogenannte Heisenbergsche Unschärferelation dar. Hierbei handelt es sich um eine quantitative Formulierung der Antwort auf die Frage "Wie gut ist die Fourier-Transformierte einer Funktion im Frequenzraum lokalisiert?". Der Begriff der Lokalisierung kann als "Ausbreitung" einer Funktion angesehen werden, d.h. bei einem breiten Träger einer Funktion wird von einer schlechten Lokalisierung gesprochen. Analog hierzu bedeutet ein schmaler Träger eine gute bzw. scharfe Lokalisierung. An dieser Stelle sei der Begriff der Trägerbreite genauer betrachtet, dieser kann als Länge des Intervalls $[\min(supp^*(f)), \max(supp^*(f))]$ aufgefasst werden. Mit $supp^*(f)$ wird die Menge $\{x| \, |f(x)| >> 0\}$ bezeichnet, andernfalls hätte beispielsweise die Gauss-Funktion eine Intervalllänge von ∞. Betrachten wir hierzu Abb.(23) und (24).

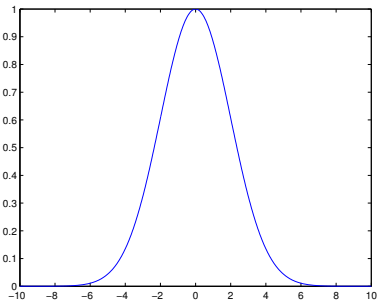


Abb. 23: Schlechte zeitliche Lokalisierung

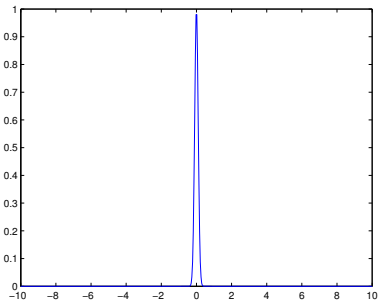


Abb. 24: Gute zeitliche Lokalisierung

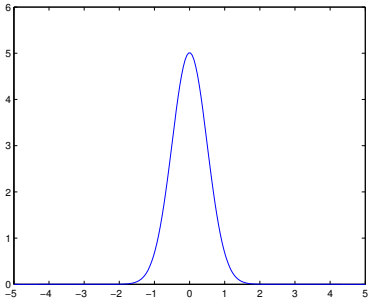


Abb. 25: Gute Frequenz Lokalisierung

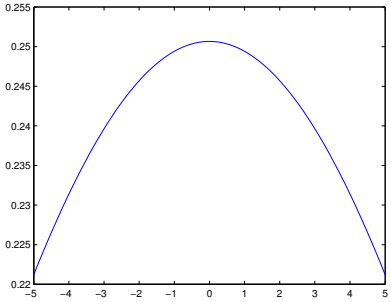


Abb. 26: Schlechte Frequenz Lokalisierung

In beiden Abbildungen ist jeweils dieselbe Funktion $f(t) := e^{-a\frac{t^2}{2}}$ dargestellt, jedoch unterscheiden sich die Graphen hinsichtlich des Parameters a. In Abb. (23) beträgt $a = 0.25$ wohingegen in Abb.(24) $a = 100$ gewählt wurde. Analog hierzu wird ebenfalls im Frequenzraum von einer entsprechenden Lokalisierung gesprochen, Abb(25) und Abb(26) stellen die entsprechenden Fourier-Transformierten der Funktionen dar, schon hier wird der Zusammenhang deutlich, je besser die zeitliche Lokalisierung ist desto schlechter die diese im Frequenzbereich.
Der Hintergrund der Frage nach der "Ausbreitung" einer Funktion, ist der Wunsch feststellen zu können, wann welche Frequenzen auftreten. Betrachten wir dazu das vorherige Beispiel der Analyse und Synthese der Funktion (40). Die Koeffizienten c_k stellen die komplexe Amplitude dar, mit welcher die Frequenz k in f vertreten

ist, jedoch kann keine Aussage über den Zeitpunkt getroffen werden, wann diese Frequenz auftritt. Einzig die Aussage, dass diese Frequenz überhaupt auftritt, kann getroffen werden.
Ein naiver Ansatz wäre der Versuch, die zu analysierende Funktion mit einem verschobenen δ-Impuls an einer Stelle zu multiplizieren (d.h. die Abtasteigenschaft der δ-Distribution ausnutzen). Somit wäre die Funktion zeitlich perfekt lokalisiert. Jedoch gilt $\widehat{\delta}(\xi) = 1$, somit ist die entsprechende Fourier-Transformierte im Frequenzraum nicht lokalisiert.
Die Unschärferelation trifft über diesen Sachverhalt eine quantitative Aussage und liefert darüber hinaus eine Menge von Funktion, welche eine optimale Lokalisierbarkeit im Zeit- und Frequenzbereich bieten.

4.2. Mathematischer Hintergrund

An dieser Stelle folgt die Herleitung der Heisenbergschen Unschärferelation. Zum besseren Verständnis des Nachweises werden zunächst einige wichtige Rechenregeln hergeleitet.

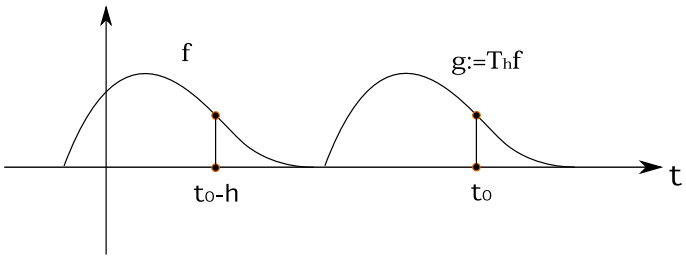


Abb. 27: Verschiebung einer Funktion f

Zunächst definieren wir den Translations-Operator T_h, der auf eine beliebige Funktion $f : \mathbb{R} \to \mathbb{C}$ angewandt, diese um den Wert $h \in \mathbb{R}$ entlang des Definitionsbereichs von f verschiebt (siehe Abb(27)).

$$T_h f(t) := f(t-h) \tag{70}$$

Anschaulich bedeutet dies, dass für positive h, der Graph von f um h nach rechts verschoben wird. Für negative Werte entsprechend nach links. Betrachten wir nun den Zusammenhang zwischen einer Verschiebung im Zeitbereich und dem Resultat im Frequenzbereich. Sei $f \in L^2$ und $g(t) := T_h f(t)$, so lautet die Fourier-Transformierte von g

$$\widehat{g}(\xi) = \int f(t-h) e^{i\xi t}\, dt\,. \tag{71}$$

Mit Hilfe der Substitution von $t - h = t'$ folgt daraus

$$\widehat{g}(\xi) = \int f(t') e^{i\xi(t'+h)}\, dt' = e^{i\xi h} \widehat{f}(\xi)\,. \tag{72}$$

Dies führt zur ersten wichtigen Formel (R1) für Herleitung der Unschärferelation.

$$\mathcal{F}(T_h f)(\xi) = e^{i\xi h} \widehat{f}(\xi) \tag{R1}$$

Kurz zusammengefasst bedeutet dies, dass aus einer Verschiebung der Funktion im Zeitbereich um den Wert h, eine um den Faktor $e^{i\xi h}$ skalierte Fouriertransformierte dieser Funktion resultiert.
Selbstverständlich kann auch die Fourier-Transformierte einer Funktion im Frequenzbereich verschoben werden. Dazu betrachten wir die mit $e^{i\omega t}$ modulierte Funktion f und definieren $g(t) := e^{i\omega t} f(t)$. Aus der Fourier-Transformierten

$$\widehat{g}(\xi) = \int e^{i\omega t} f(t) e^{-i\xi t}\, dt = \int f(t) e^{-i(\xi-\omega)t}\, dt = \widehat{f}(\xi - \omega) \tag{73}$$

folgt, dass eine Modulation der Zeitfunktion mit $e^{i\omega t}$ bzw. e_ω in einer Verschiebung im Frequenzbereich resultiert. Hieraus folgt die Formel (R2).

$$\mathcal{F}(\mathbf{e}_\omega f)(\xi) = \widehat{f}(\xi - \omega) \tag{R2}$$

In Worten ausgedrückt, bedeutet dies, dass durch die Modulation einer Funktion f mit einer weiteren Funktion e_ω, die Fouriertransformierte $\widehat{f}$ entlang der ξ-Achse verschoben wird.
Für die nächste Rechenregel wird ein weiterer Operator D_a benötigt, welcher wie folgt definiert ist,

$$D_a f(t) = f\left(\frac{t}{a}\right), a \in \mathbb{R}\backslash\{0\} \,. \tag{74}$$

Dieser Operator soll nun benutzt werden, um den Zusammenhang zwischen einer dilatierten Funktion f und ihrer Fourier-Transformierten zu untersuchen. Zuvor jedoch einige Worte hinsichtlich der Interpretation von D_a. Der Operator dient der "'Stauchung"' bzw. "'Streckung"' einer Funktion, in Abb.(28) wird der Fall einer Streckung verdeutlicht. Für $|a| > 1$ wird der Graph der Funktion gestreckt und für $|a| < 1$ wird dieser gestaucht. Weitere triviale Implikationen sind, dass zum einen für $a < 0$ der Graph der Funktion an der vertikalen Achse gespiegelt wird und zum anderen für $a = 1$ unverändert bleibt. Wie in den vorherigen Herleitungen definieren wir eine Funktion $g(t) := D_a f(t)$ und bestimmen deren Fourier-Transformierte nach bekannter Regel

$$\widehat{g}(\xi) = \int f\left(\frac{t}{a}\right) e^{-i\xi t} dt \,. \tag{75}$$

Nach der Substitution $t' = \frac{t}{a}$ ergeben sich

$$t = t'a, \;\; t' \in \mathbb{R}, \qquad dt = a\, dt' \tag{76}$$

und damit folgt für $\widehat{g}$

$$\widehat{g}(\xi) = \frac{|a|}{\sqrt{2\pi}} \int f(t') e^{-i\xi a t'} dt = |a| \widehat{f}(a\xi) \,. \tag{77}$$

Bei Betrachtung der Skalierung der Parameter t und ξ fällt direkt auf, dass diese immer umgekehrt proportional skaliert werden. In Worten ausgedrückt heißt dies, dass falls der Graph von f gestreckt wird, so wird entsprechend der Graph der Fourier-Transformierten gestaucht. Es ist ebenfalls zu erkennen, dass die Fourier-Transformierte in diesem Fall zusätzlich vertikal gestreckt wird. Diese Eigenschaft wird durch (R3) ausgedrückt.

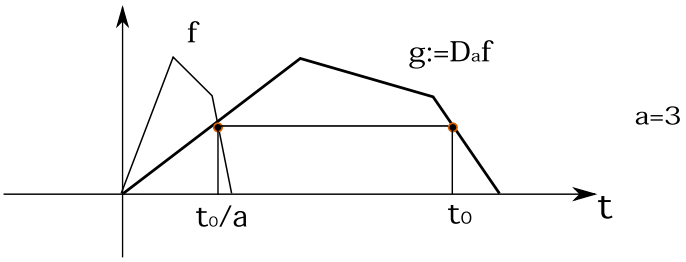


Abb. 28: Dilatation einer Funktion f

Schon anhand dieser Regel lässt sich erkennen, dass eine Funktion f und ihre Fourier-Transformierte $\widehat{f}$ nicht gleichzeitig auf einem beliebig kleinen Bereich der $t-$ bzw. ξ-Achse lokalisiert sein können.

$$\mathcal{F}(D_a f)(\xi) = |a| D_{\frac{1}{a}} \widehat{f}(\xi) \tag{R3}$$

Zum Abschluss soll der Zusammenhang zwischen der Ableitung einer Funktion und deren resultierender Fourier-Transformierten betrachtet werden. Sei $f \in C^1$ und f sowie f' integrabel, dann gilt $\lim\limits_{t \to \pm\infty} f(t) = 0$.

Wir bestimmen die Fourier-Transformierte von f' mit Hilfe der partiellen Integration

$$\int f'(t)e^{-i\xi t}dt = f(t)e^{-i\xi t}\Big|_{-\infty}^{\infty} + i\xi \int f(t)e^{-i\xi t}dt\,. \tag{78}$$

Hieraus folgt die Regel (R4).

$$\widehat{f'}(\xi) = i\xi\widehat{f} \tag{R4}$$

Weiter fortfahrend ergibt sich für beliebig hohe Ableitungen ($r \geq 0$) die Formel

$$\widehat{f^{(r)}}(\xi) = (i\xi)^r\widehat{f}\,. \tag{79}$$

Wie bereits anhand von (R3) zu erkennen war, kann eine Funktion nicht gleichzeitig beliebig genau im Frequenz- und Zeitbereich lokalisiert sein. Die Heisenbergsche Unschärferelation beschreibt diesen Sachverhalt quantitativ. In der Quantenmechanik wird die Bewegung eines Teilchens mit Hilfe einer Funktion ψ beschrieben, wobei $f_X(x) := |\psi(x)|^2$ die Wahrscheinlichkeitsdichte für den Ort und $f_P(\xi) = |\widehat{\psi}(\xi)|^2$ die Wahrscheinlichkeitsdichte für den Impuls des Teilchens darstellt. Der Erwartungswert der Zufallsvariablen X^2 wird bestimmt durch

$$\int x^2 f_X(x)dx = \int x^2|\psi(x)|^2dx = ||x\psi||^2\,, \tag{80}$$

analog hierzu der Erwartungswert der Zufallsvariable P^2

$$\int \xi^2 f_P(\xi)d\xi = \int \xi^2|\widehat{\psi}(\xi)|^2d\xi = ||\xi\widehat{\psi}||^2\,. \tag{81}$$

Diese beiden Formeln können nach [2] als Maß, für die (horizontale und bei 0 beginnende) Ausbreitung der Funktionen über die x- bzw. ξ-Achse betrachtet werden. Die Unschärferelation trifft nun für beliebige Funktionen $\psi \in L^2$ folgende Aussage

$$||x\psi|| \cdot ||\xi\widehat{\psi}|| \geq \frac{1}{2}||\psi||^2\,. \tag{82}$$

In Worten bedeutet Gleichung (82), dass die "Ausbreitungen" beider Funktionen miteinander korrelieren. D.h. falls eine dieser Funktionen beliebig genau lokalisiert wird, so "kompensiert" die andere diese Genauigkeit. Eine weitere Folgerung aus der Unschärferelation ist, dass eine der beiden Zahlen $||x\psi(x)||$ und $||\xi\widehat{\psi}(\xi)||$ immer $\geq ||\psi(x)||^2/2$ ist. Das Gleichheitszeichen gilt für konstante Vielfache der Funktion $x \mapsto e^{-cx^2}$ mit $c \geq 0$. Es folgt nun der mathematische Beweis dieser Ungleichung, für zu umfangreiche Teile wird auf die entsprechende Literatur verwiesen.
Um den Nachweis hinsichtlich Konvergenzfragen zu vereinfachen, sei die Annahme getroffen, dass $\psi \in S$ ist, somit gilt insb. $\lim\limits_{t\to\pm\infty} x|\psi(x)|^2 = 0$. Mit Hilfe von (R4) und der Parsevalschen Formel (36) lässt sich die Fourier-Transformierte aus (82) entfernen.

$$||\xi\widehat{\psi}|| = ||\widehat{\psi'}|| = ||\psi'|| \tag{83}$$

Die Unschärferelation lässt sich somit zu

$$||x\psi|| \cdot ||\psi'|| \geq \frac{1}{2}||\psi||^2 \tag{84}$$

umschreiben. Weiterhin gilt die Cauchy-Schwarzsche Ungleichung

$$||x\psi|| \cdot ||\psi'|| \geq |<x\psi, \psi'>| \geq |\Re(<x\psi, \psi'>)|\,. \tag{85}$$

Die Rechte Seite der Gleichung lässt sich nun auf folgende Weise berechnen

$$2\Re(<x\psi, \psi'>) =<x\psi, \psi'> + <\psi', x\psi>= \int x(\psi(x)\overline{\psi'(x)} + \psi'(x)\overline{\psi(x)})dx$$

$$= x|\psi(x)|^2\big|_{-\infty}^{\infty} - \int |\psi(x)|^2 dx = -||\psi|| \,. \tag{86}$$

Durch einsetzen dieses Ergebnisses in (85) ergibt sich (84). Wie in [2] und [6] nachzulesen ist, lässt sich aufgrund der Tatsache, dass S in L^2 dicht liegt, die Gleichung (84) ohne Änderung auf Funktionen $\psi \in L^2$ anwenden. Damit in (84) an beiden Stellen das Gleichheitszeichen gilt, müssen die Vektoren $x\psi$ und ψ' linear abhängig sein. Mathematisch ausgedrückt führt dies zur folgenden Differentialgleichung

$$\psi'(x) \equiv (\mu + i\nu)x\psi(x) \ \ , \ x \in \mathbb{R} \ \ , \ \mu + i\nu \in \mathbb{C} \,. \tag{87}$$

Als Lösungen dieser Gleichung ergeben sich Funktionen

$$\psi(x) = Ce^{(\mu+i\nu)x^2/2} \ \ , \ C \in \mathbb{C} \,, \tag{88}$$

diese Funktionen sind genau dann $\in L^2$ falls μ negativ ist. Weiterhin muss natürlich $< x\psi, \psi' >$ reell sein, d.h. mit (87) ergibt sich

$$< x\psi, \psi' > = < x\psi, (\mu + i\nu)x\psi > = (\mu + i\nu)||x\psi||^2 \in \mathbb{R} \,, \tag{89}$$

woraus $\nu = 0$ folgt.
Die Unschärferelation besagt also, dass die Funktionen ψ und $\widehat{\psi}$ an den Stellen $x = 0$ bzw. $\xi = 0$ nicht gleichzeitig scharf (d.h. gering ausgebreitet) lokalisiert sein können. Dies gilt auch für ein beliebiges Zahlenpaar (x_0, ξ_0). D.h. für Funktionen $\psi \in L^2$ gilt

$$||(x - x_0)\psi|| \cdot ||(\xi - \xi_0)\widehat{\psi}|| \geq \frac{1}{2}||\psi||^2 \,. \tag{90}$$

Für den Nachweis dieser Behauptung, betrachten wir die Funktion

$$g(t) := e^{-i\xi_0 t}\psi(t + x_0) \tag{91}$$

und bestimmen hierzu

$$||g||^2 = \int |\psi(t + x_0)|^2 dt = ||\psi||^2 \tag{92}$$

$$||tg||^2 = \int t^2|\psi(t + x_0)|^2 dt = \int (x - x_0)^2|\psi(x)|^2 dx \,, \tag{93}$$

wobei bei der letzten Gleichung die Substitution $x = t + x_0$ benutzt wurde. Nun lässt sich g umschreiben

$$g(t) = e^{-i\xi_0 t}h(t) \ \ , \ h(t) := f(t + x_0) \tag{94}$$

und mit Hilfe der Regeln (R1) und (R2) folgt

$$\widehat{g}(\tau) = \widehat{h}(\tau + \xi_0) = e^{ix_0(\tau+\xi_0)}\widehat{f}(\tau + \xi_0) \,. \tag{95}$$

Daraus folgt analog

$$||\tau g||^2 = \int \tau^2|\widehat{f}(\tau + \xi_0)|^2 d\tau = \int (\xi - \xi_0)|\widehat{f}(\xi)|^2 d\xi \,. \tag{96}$$

Wird nun Gleichung (82) auf die Funktion g angewandt, so ergibt sich nach einsetzen von $||g||$,$||tg||$ und $||\tau\widehat{g}||$ die Gleichung (90).
Die Kernaussage der Unschärferelation kann mit Hilfe einer so genannten Heisenberg-Box (Abb.(29)) dargestellt werden. Die genannte Grafik veranschaulicht die Zeit- und Frequenzlokalisierung der Funktion $f := e^{-x^2/2}$. Der Mittelpunkt bzw. die Position des Rechtecks wird durch die Funktion f und deren Fourier-Transformierten $\widehat{f}$ bestimmt, Höhe und Breite des Rechtecks stellen die "Ausbreitung" beider Funktionen über der entsprechenden Achse dar. Dieses Rechteck wird als Heisenberg-Box bezeichnet und besitzt eine konstante Fläche. Dies bedeutet für eine Streckung der Funktion f um den Faktor a (d.h $D_a f$), dass die Breite der Box

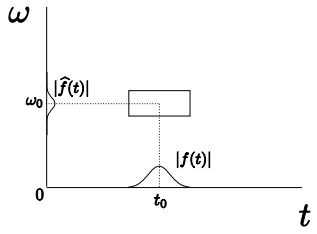


Abb. 29: Heisenberg Box

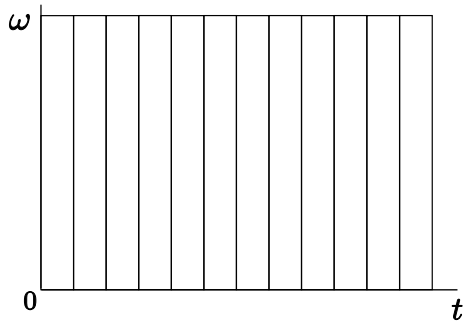


Abb. 30: Unschärfe der Sample Werte

Abb. 31: Unschärfe der Basisfunktionen e_k

zunimmt und die Höhe entsprechend abnimmt. Die Darstellung der Unschärfe einer Funktion mit Hilfe der Heisenberg-Box wird in Kapitel 5 eine wichtige Rolle spielen.
Betrachten wir an dieser Stelle die Basisfunktionen e_k der Fourier-Reihe, welche in diesem Zusammenhang auch als Zeit-Frequenz-Atome bezeichnet werden. Diese sind natürlich im Frequenzraum absolut präzise lokalisiert, dementsprechend im Zeitbereich überhaupt nicht (Abb.(31)). Die mit Hilfe der Fourier-Reihe darzustellende Funktion ist hinsichtlich ihrer Funktionswerte im Zeitbereich perfekt lokalisiert, besitzt jedoch im Frequenzbereich eine maximale Unschärfe (Abb.(30)).

5. Wavelet-Transformation

5.1. Gefensterte Fourier Transformation

Im Bereich der Bildverarbeitung wird die 2-dimensionale Fourier-Transformation zur Texturanalyse benutzt. Ein Bild kann als 2-dimensionales Signal angesehen werden, somit ist es möglich, für dieses auch eine Fourier-Transformierte zu bestimmen (ausgehend von Funktionen $L^2(\mathbb{R}^2)$). Texturen innerhalb eines Bildes erzeugen charakteristische Muster in ihrer Fourier-Transformierten. Wie in [4] erwähnt wird, kann eine Textur oft durch ihre "Granularität" beschrieben werden, betrachten wir hierzu 2 Bilder von jeweils einer Sand- und Kiesoberfläche. Natürlich weist die Sandoberfläche eine feinere Struktur auf als die Kiesoberfläche, die Fourier-Transformierte des Bildes der Sandoberfläche wird im Vergleich zur Fourier-Transformierten des anderen Bildes wesentlich stärker im höheren Frequenzbereich vertreten sein. Darunter sei zu verstehen, dass der Graph des Betragsspektrums nur im höheren Frequenzbereich signifikante Werte (d.h. Werte $>> 0$) darstellt. Leider kann ohne weiteres keine Aussage über die für ein solches Muster zugehörige Position im Bild getroffen werden. Es kann höchstens eine Aussage darüber getroffen werden, dass bestimmte Texturen überhaupt in einem Bild vorkommen. In diesem Kapitel wird die Idee der Zeit-Frequenz-Analyse vorgestellt, welche einen Ansatz zur Lösung des Zeit-Frequenz- bzw. Ort-Frequenz-Zuordnungsproblems darstellt. Zur Vereinfachung der Erläuterung wird auch hier nur auf den 1-dimensionalen Fall eingegangen.
Wie im Kapitel über die Heisenbergsche Unschärferelation beschrieben wurde, kann eine Funktion nicht gleichzeitig beliebig genau im Zeit- und Frequenzbereich lokalisiert sein. Es sei an die zuvor erwähnte Idee der Funktionsabtastung mit Hilfe von δ-Impulsen erinnert. Leider wäre in einem solchen Fall überhaupt keine Lokalisierung im Frequenzbereich möglich (ganz im Gegensatz zum Zeitbereich). Wird jedoch anstelle des δ-Impulses eine andere "Abtast"-Funktion gewählt, so kann ein Kompromiss in dieser Problematik getroffen werden. Betrachten wir hierzu die Gauss-Funktion

$$w(t,\sigma) := e^{-\frac{1}{2}\frac{t^2}{\sigma^2}} , \tag{97}$$

wobei der Parameter σ die Breite des Funktionsgraphen bestimmt. Weiterhin sei $g(t) := chirp_{t_0,f_0}(t)$ ein

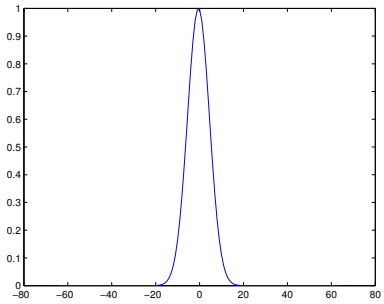


Abb. 32: Gauss Funktion mit $\sigma = 5$

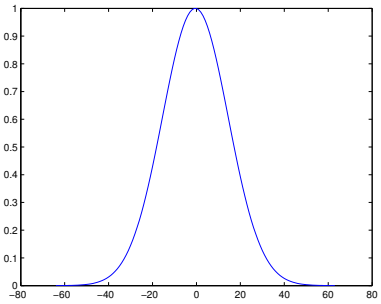


Abb. 33: Gauss Funktion mit $\sigma = 15$

Chirp-Signal[1], dessen Frequenz linear mit der Zeit wächst und bei t_0 (in Sekunden) die Frequenz f_0 (in Hz) erreicht. In Abb.34 wird ein Beispiel eines solchen Chirp-Signals dargestellt, hierbei erreicht die Frequenz des Signals zum Zeitpunkt $t = 1s$ die Frequenz $f = 5$Hz. Wird nun die Funktion w mit diesem Signal multipliziert, so entsteht ein neues Signal $h_{\vec{k}}$, welches zeitlich besser lokalisiert ist. In mathematischer Schreibweise ausgedrückt bedeutet dies, $h_{\vec{k}}(t) := g(t)w(t-1, 0.15)$ (siehe Abb.35) wobei $\vec{k} = (1, 0.15)$ den Parameter-Vektor darstellt. Die Funktion w wird als Fensterfunktion bezeichnet, da diese bei Multiplikation mit einer weiteren Funktion g immer einen Teilausschnitt dieser Funktion liefert, die Bildung eines solchen Ausschnitts wird auch

[1] ein Chirp-Signal stellt ein Signal dar dessen Frequenz sich zeitlich ändert

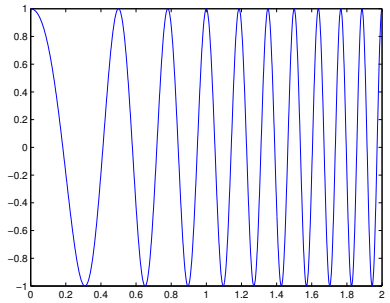


Abb. 34: Chirp Signal $chirp_{1,5}(t)$

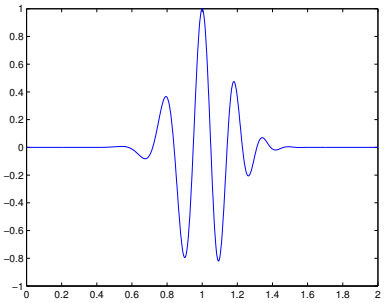


Abb. 35: Gefenstertes Chirp Signal

als "einblenden" bezeichnet. Mit Hilfe der Fensterfunktion läßt sich somit ein bestimmter Bereich einer Funktion einblenden, hierzu muss die Fensterfunktion natürlich entsprechend verschoben und dilatiert werden.
Da w nach der Unschärferelation im Zeit- und Frequenzbereich optimal lokalisiert ist (siehe Kapitel 4), soll an dieser Stelle die Fourier-Transformierte, eines mit dieser Funktion gefensterten Signals betrachtet werden. Es sei $h_{\vec{k}}(t) := g(t)w(t-k_1,k_2)$, so lautet die Transformation von $h_{\vec{k}}$ in den Frequenzbereich

$$\widehat{h_{\vec{k}}}(\omega) = |k_2|e^{i\omega k_1} \cdot (\widehat{g} * D_{k_2}\widehat{w})(\omega) \tag{98}$$

mit $\widehat{w}(\omega, \frac{1}{k_2}) = \frac{1}{\sqrt{2\pi}}e^{-\frac{\omega^2 k_2^2}{2}}$, auch an dieser Gleichung lassen sich die Rechenregeln (R2) und (R3) erkennen. Eine solche Transformation wird als gefensterte Fourier-Transformation bzw. Short-Time-Fourier-Transformation (STFT) bezeichnet. Wird die STFT für alle Translationen von w mit fester Fensterbreite σ bestimmt, so lässt sich daraus ein Spektrogramm erstellen. Die Erstellung eines Spektrogramms stellt die Zeit-Frequenz-Analyse eines Signals dar. Bei dieser Untersuchung wird versucht festzustellen, zu welchen Zeiten welche Frequenzen im Signal vorkommen.
Als Beispiel als auch zur Veranschaulichung eines Spektrogramms seien die Abbildungen (36) und (37) betrachtet, hierbei wurde ein Signal $g(t) := chirp_{1,200}(t)$ mit jeweils einer Variante der Gauss-Fensterfunktion w multipliziert und die STFT für alle Translationen von w berechnet. Abb. (36) stellt die Zeit-Frequenz-Analyse für $\sigma = 0.125$ und Abb. (37) jene für $\sigma = 0.25$ dar (hierbei wurde ein logarithmischer Maßstab in der Einheit "dB" verwendet). In beiden Grafiken werden die zu den entsprechenden Zeiten auftretenden Amplituden

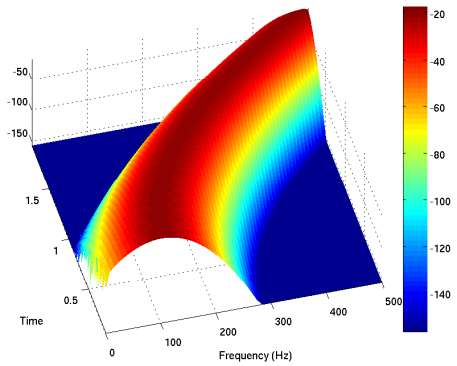


Abb. 36: STFT($chirp_{1,200}$), $\sigma = 0.125$

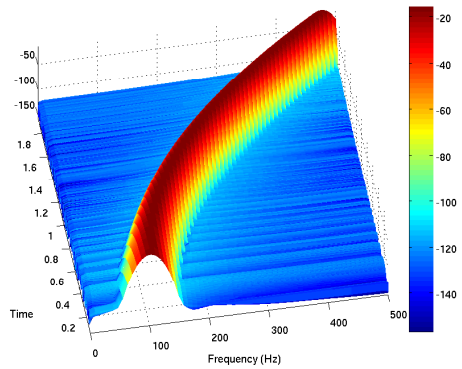


Abb. 37: STFT($chirp_{1,200}$), $\sigma = 0.25$

der Frequenzen innerhalb des Signals g dargestellt. Der Wert $\sigma = 0.125$ bedeutet eine gute zeitliche Lokalisierung für die Funktion $h(t) := g(t)w(t, 0.125)$, d.h. die Fourier-Transformierte $\widehat{h}(\omega)$ ist dem entsprechend

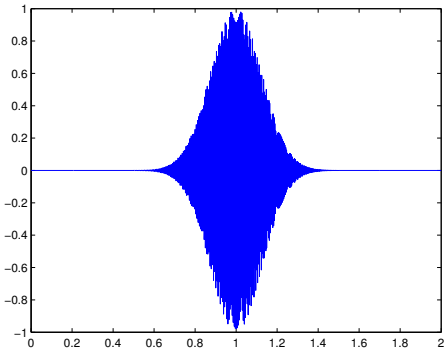


Abb. 38: $chirp_{1,200}$, $\sigma = 0.125$

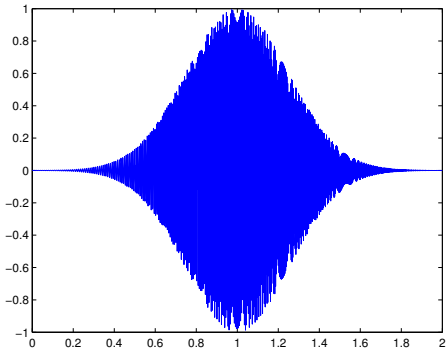


Abb. 39: $chirp_{1,200}$, $\sigma = 0.25$

schlecht im Frequenzbereich lokalisiert. Hingegen ist bei $\sigma = 0.25$ die Aussage umzukehren, die Fourier-Transformierte ist im direkten Vergleich zum vorherigen Fall wesentlich besser lokalisiert, leider lässt sich aufgrund der schlechten zeitlichen Lokalisierung (siehe hierzu die Ausbreitung der Funktionsgraphen in Abb.(38) und (39)) nicht genau erkennen wann diese Frequenz erscheint .
Mit Hilfe des Spektrogramms lässt sich somit abschätzen wann welche Frequenzen auftreten. Bei Betrachtung des Spektrogramms (36) ist erkennbar, dass der Graph entlang der Frequenz-Achse sehr ausgebreitet ist im Vergleich zum Spektrogramm (37). Verantwortlich hierfür ist die Fensterfunktion w, da diese die Ausbreitung einer Funktion entlang ihres Trägers im Zeitbereich bestimmt. Das Signal g besitzt zum Zeitpunkt $t = 1$ die Frequenz $f = 200$Hz, wird nun versucht diesen Wert mit Hilfe beider Spektrogramme zu ermitteln, so ergeben sich folgende Probleme:

- Für $\sigma = 0.125$ ist h zeitlich recht gut lokalisiert, jedoch besitzt $\widehat{h}$ zu jedem Zeitpunkt eine große Ausbreitung über die Frequenz-Achse. D.h. es kommen viele Frequenzen auf diesem Zeitintervall vor

- Für $\sigma = 0.25$ ist h zeitlich wesentlich schlechter lokalisiert als zuvor, dafür ist die Fourier-Transformierte $\widehat{h}$ deutlich besser im Frequenzbereich lokalisiert

Anhand der Unschärferelation lässt sich die Entwicklung des Dilemmas für eine entsprechende Änderung von σ vorhersagen. Würden wir für ein unendlich ausgebreitetes Signal ein Spektrogramm erstellen, bei welchem zusätzlich eine Fensterfunktion mit unendlicher Breite verwendet wird, so wäre anhand des Spektrogramms absolut deutlich erkennbar, dass zum Zeitpunkt t_0 nur eine Frequenz f_0 erscheint. Jedoch wäre diese Erkenntnis nichts weiter als die Tatsache, dass diese Frequenz überhaupt im Signal auftaucht. Mit anderen Worten gesagt ist die Intepretation eines Spektrogramms abhängig von der gewählten Fensterfunktion (bzw. der Breite des eingeblendeten Bereichs).
Zur genaueren Abschätzung des Ortes weiterer Frequenzen müssen mehrere gefensterte Fourier-Transformationen durchgeführt werden, jeweils mit verschiedenen Werten für σ. Somit ist die STFT aufgrund der festen Fenster größe für eine Zeit-Frequenz-Analyse nur bedingt geeignet. Jedoch besitzt die STFT einen großen Vorteil gegenüber der Fourier-Transformation, mit ihrer Hilfe ist es möglich Aussagen über lokale Signaleigenschaften zu machen. Dies ist z.B. bei einer Texturanalyse im Bereich der Bildverarbeitung von Vorteil, da mit Hilfe der STFT Bildsegmente hinsichtlich der dort auftretenden Textur untersucht werden können.
An dieser Stelle sei die Auflösung der STFT mit Hilfe der Heisenberg-Box betrachtet. Für eine konstante Fensterbreite σ stellen die Abb.(40) sowie (41) die Zeit-Frequenz-Unschärfe der jeweiligen gefensterten Fourier-Transformation dar.

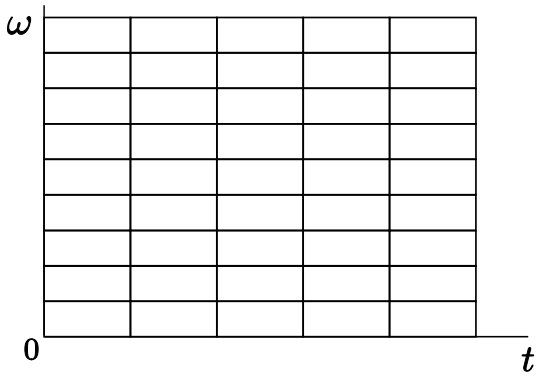


Abb. 40: Unschärfe für $\sigma > 1$

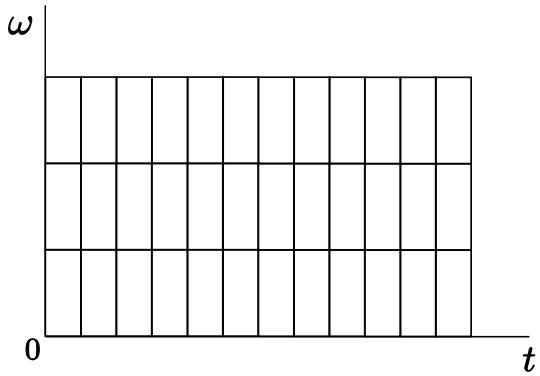


Abb. 41: Unschärfe für $0 < \sigma < 1$

5.2. Wavelet-Transformation

In diesem Abschnitt wird die Wavelet-Transformation als Alternative zur STFT vorgestellt und ihre Vor- und Nachteile behandelt. Das EBGM-Verfahren, welches im Kapitel 8 vorgestellt wird, nutzt die so genannte Gabor-Wavelet-Transformation zur abstrakten Beschreibung bestimmter Bildpunkte. Wie in den vorherigen Kapiteln wird auch hier nur der 1-dimensionale Fall behandelt, da gerade bei der Wavelet-Transformation eine Darstellung von 2-dimensionalen Signalen aufgrund der resultierenden Parametermenge recht heikel ist. Eine Funktion $\psi : \mathbb{R} \to \mathbb{C}$ welche folgende Bedingungen erfüllt

$$\psi \in L^2, \quad ||\psi|| = 1 \tag{1}$$

$$2\pi \int_{\mathbb{R}^*} \frac{|\widehat{\psi}(\xi)|^2}{|\xi|} d\xi := C_\psi < \infty \,, \;\; \mathbb{R}^* = \mathbb{R}\backslash\{0\} \,, \tag{2}$$

wird als Wavelet bezeichnet. Bedingung (2) ist äquivalent mit

$$\int \psi(t)dt = 0 \quad \text{bzw.} \quad \widehat{\psi}(0) = 0 \,, \tag{3}$$

was für Funktionen $\psi \in L^2$ und $t\psi \in L^1$ (d.h. $\int |t||\psi(t)|dt < \infty$), äquivalent mit (2) ist. Der erste Teil von (3) ist gleichbedeutend damit, dass die Funktion ψ den Mittelwert 0 hat, das Kürzel "bzw." in (3) bezieht sich auf die Betrachtung des Zeit- oder Frequenzbereichs. Für die Darstellung der Funktion im Frequenzbereich muss somit $\widehat{\psi}(0) = 0$ gelten, schliesslich stellt $\widehat{\psi}(0)$ (konstanter Wert / Schwingung mit Frequenz 0) den Gleichanteil der Funktion dar. Somit verläuft der Graph der Funktion im Zeitbereich, teilweise ober- und unterhalb der t-Achse. Die erwähnten Bedingungen (1) bis (3) können in [1] sowie [2] detailiert nachgelesen werden.
Für ein gewähltes Wavelet ψ und eine Funktion $f \in L^2$ wird allgemein mit

$$(W_\psi f)(a,b) := \frac{1}{\sqrt{|a|}} \int f(t)\overline{\psi(\frac{t-b}{a})}dt \,, \quad a \neq 0 \tag{99}$$

die Wavelet-Transformation definiert, $W_\psi f$ ist die so genannte Wavelet-Transformierte der Funktion f. Hierbei handelt es sich um die kontinuierliche Wavelet-Transformation (CWT), der Definitionsbereich von $W_\psi f$ lautet somit $\mathbb{R}^2_- := \{(a,b)|a \in \mathbb{R}^*, b \in \mathbb{R}\}$. Bei näherer Betrachtung von (99) lässt sich feststellen, dass es sich bei $\psi(\frac{t-b}{a}) = D_a(T_b f)(t)$ um eine dilatierte und verschobene Version von ψ handelt. Der Parameter a wird als Skalen- oder Dilatationsparameter und b als Translations-Parameter bezeichnet. Die Funktion ψ stellt somit ein Ursprungs-Wavelet dar, aus welchem sich neue Wavelets "generieren" lassen. In der Literatur wird ein solches Wavelet auch als Mother-Wavelet bezeichnet. An dieser Stelle sei noch die Bezeichnung $\psi_{a,b} := D_a(T_b f)$ für Wavelets eingeführt, weiterhin stellt $\Psi := \{\psi_{a,b}|(a,b) \in \mathbb{R}^2_-\}$ eine Wavelet-Basis dar, welche die Funktionsmenge aller dilatierten und verschobenen Versionen von ψ beinhaltet. Der Begriff der Basis sei hierbei nicht

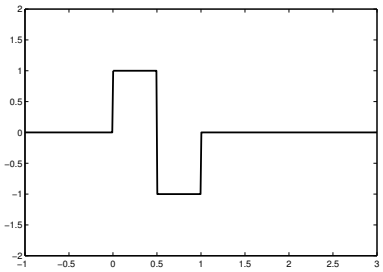
als Vektorraumbasis zu verstehen, es ist nur die Menge aller aus dem Mother-Wavelet konstruierbarer Funktionen gemeint. Wie am Anfang des Abschnitts erwähnt, ist die Darstellung der Wavelet-Transformierten recht schwierig, da diese bereits bei 1-dimensionalen Signalen 2 Parameter aufweist. Im weiteren Verlauf der Arbeit wird gezeigt, dass 2-dimensionale Signale eine von 3 Parametern abhängige Wavelet-Basis besitzen können.
Die Grundidee der Wavelet-Transformation besteht darin, eine Funktion $f \in L^2$ als Linearkombination bezüglich einer Orthonormalen Basis des L^2 darzustellen. Jedoch werden hierbei keine 2π-periodischen komplexen Exponential-Funktionen benutzt, sondern dilatierte und verschobene Versionen einer Ursprungs-Funktion (Mother-Wavelet). Intuitiv gedeutet stellen die Funktionswerte $(W_\psi f)(a,b)$, analog zur Fourier-Transformation, die Koeffizienten einer Linearkombination zur Synthese von f dar. Bevor wir ein konkretes Beispiel für ein Wavelet und die dazugehörige Wavelet-Transformierte eines Signals betrachten, soll ein sehr wichtiger Punkt hinsichtlich der eigentlichen Wavelet-Transformation erwähnt werden.
Die Transformations-Vorschrift zur Bestimmung der Wavelet-Koeffizienten einer Funktion lässt sich allgemein bezüglich der zeitlichen Translation auch als Faltung interpretieren. D.h.

$$(W_\psi f)(a,b) = \frac{1}{\sqrt{|a|}}(f * D_{-a}\overline{\psi})(b)\ . \tag{100}$$

Hierzu sei die Herleitung nach [1] und [11] gegeben

$$\begin{aligned}
(W_\psi f)(a,b) &= \frac{1}{\sqrt{|a|}}\int f(t)\overline{\psi(\frac{t-b}{a})}dt && (101)\\
&= \frac{1}{\sqrt{|a|}}\int f(t)(D_{-a}\overline{\psi})(b-t)dt && (102)\\
&= \frac{1}{\sqrt{|a|}}(f * D_{-a}\overline{\psi})(b)\ . && (103)
\end{aligned}$$

Im Kapitel 6 werden wir diese Bedingung für eine bestimmte Gruppe von Wavelets umformen, jedoch ändert dies nichts an der Tatsache, dass die Wavelet-Koeffizienten mit Hilfe der Faltung bestimmt werden können. Der Vorteil dieser Darstellung ist bei Betrachtung der Gleichung (55) erkennbar, die Bestimmung der Wavelet-Koeffizienten im Zeitbereich entspricht somit einer Multiplikation der Fourier-Transformierten von f und $\overline{\psi}$ im Frequenzbereich. Weiterhin wird durch die Darstellung als Faltung eine zusätzliche Eigenschaft von Wavelets deutlich. Da die Fourier-Transformierte eines Wavelets nach Gleichung (3) bei $\omega_0 = 0$ den Wert 0 besitzt, stellt ein Wavelet nach (100) ein Bandpass-Filter dar. Diese Eigenschaft wird in Kapitel 6 näher erläutert.
Betrachten wir nun das so genannte Haar-Mother-Wavelet, welches definiert ist durch (104) . Das Haarsche



$$\psi(t) := \begin{cases} 1 & 0 \le t < \frac{1}{2} \\ -1 & \frac{1}{2} \le t < 1 \\ 0 & \text{sonst} \end{cases} \tag{104}$$

Abb. 42: Das Haarsche Mother-Wavelet

Wavelet wird nun definiert als

$$\psi_{m,n}(t) := 2^{-\frac{m}{2}}\psi(2^{-m}t - n),\ \ m,n \in \mathbb{Z}\ , \tag{105}$$

wobei natürlich $\psi_{0,0} = \psi$. Zur Bedeutung des Skalen-Parameters sei gesagt, dass je kleiner (insb. für negative Werte) dieser gewählt wird, umso stärker wird der Graph der Funktion gestaucht. Prüfen wir nun die

Orthonormalität der Haarschen Basis $\psi_{a,b}$, für $n \neq n'$ und $m = m'$ gilt

$$< \psi_{m,n}, \psi_{m',n'} >=< \psi_{m,n}, \psi_{m,n'} >= \int \psi_{m,n}(t)\psi_{m,n'}(t)dt = 0 \,, \tag{106}$$

da $\psi_{m,n}$ und $\psi_{m,n'}$ in diesem Fall disjunkte (sich nicht überlappende) Träger besitzen. Für unterschiedliche Skalen-Parameter, d.h. $m \neq m'$ überlappen sich die Träger beider Funktionen. Dies ist jedoch kein Problem, da eine der Funktionen auf dem Träger der anderen einen konstanten Wert besitzt (Abb.(43) veranschaulicht dies). Somit folgt für $m \neq m'$ oder $n \neq n'$

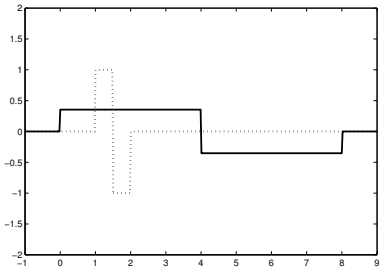


Abb. 43: $< \psi_{m,n}, \psi_{m',n'} >$ für $m \neq m'$ und $n \neq n'$

$$< \psi_{m',n}, \psi_{m',n'} >= \int \psi_{m,n}(t)\psi_{m',n'}(t)dt = 0 \,, \tag{107}$$

weiterhin gilt offensichtlich

$$||\psi_{m,n}||^2 = \int |\psi_{m,n}|^2 dt = 1 \,, \tag{108}$$

somit ist die Funktionsmenge Ψ eine orthonormale Basis des $L^2(\mathbb{R})$. Nun gilt es nur noch zu prüfen, ob $\widehat{\psi}(0) = 0$ ist, d.h. wir bestimmen die Fourier-Transformierte von ψ an der Stelle $\omega = 0$

$$\widehat{\psi}(0) = \int_0^{1/2} e^{-i\cdot 0\cdot t}dt - \int_{1/2}^{1} e^{-i\cdot 0\cdot t}dt = 0 \tag{109}$$

somit erfüllt die Haarsche Basis alle Bedingungen für Wavelets (ausführliche Nachweise hierzu sind in [2] zu finden).
Betrachten wir an dieser Stelle ein Beispiel nach [3], in welchem die Haar-Approximation der Funktion $f(t) := e^{-0.3t^2}(4\sin(2t) + 2\cos(3t))$ bestimmt wird. Es gilt für die Haar-Approximation einer Funktion

$$f \approx \sum_{m=m_0+1}^{m_1} \sum_{n\in\mathbb{Z}} \nu_{m,n}\psi_{m,n} \tag{110}$$

mit m_0 als Anfangs-Skalenwert und m_1 End-Skalenwert. Die Koeffizienten $v_{m,n}$ lassen sich durch

$$\nu_{m,n} =< f, \psi_{m,n} >= \int_{n2^m}^{(n+1)2^m} f(t)\psi_{m,n}(t)dt = 2^{-\frac{m}{2}} \int_{n2^m}^{(n+0.5)2^m} f(t)dt - \int_{(n+0.5)2^m}^{(n+1)2^m} f(t)dt \tag{111}$$

bestimmen. In Abb.(44) ist das Resultat der Rekonstruktion von f zu sehen, hierbei wurden die Werte $m0 = -2$ bis $m_1 = 6$ verwendet.
Zum Abschluss dieses Kapitels sollen die Vorteile und Unterschiede der Wavelet-Transformation zur STFT kurz erörtert werden. Der wesentliche Unterschied zur STFT ist die Tatsache, dass die Fenstergröße bei der

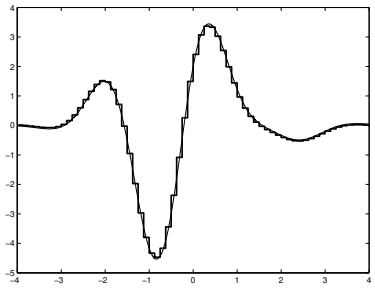


Abb. 44: Haar-Approximation für Skalenwerte -2 bis 6

Wavelet-Transformation variabel ist. Betrachten wir zum besseren Verständnis die Bilder (45) bis (47), diese stellen die Graphen verschiedener Funktionen aus der Haarschen Basis dar. Wir wissen, dass der Graph eines (Mother-)Wavelets immer eine Form von Schwingung darstellt, wie verhält sich jedoch die Frequenz der Schwingung, bei einer Dilatation dieser Funktion. In den erwähnten Abbildungen wird dieser Sachverhalt verdeutlicht, je stärker ein Wavelet gestaucht wird, umso größer wird dessen Frequenz auf seinem Träger. Für die Dilatation eines Wavelets gilt analog, dass sich dessen Frequenz verringert. Weiterhin stellt die Wavelet-

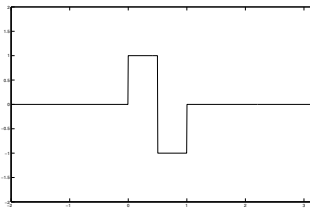


Abb. 45: $\psi_{0,0}$

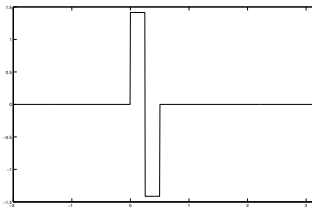


Abb. 46: $\psi_{-1,0}$

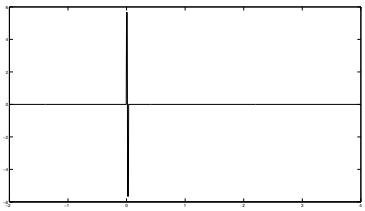


Abb. 47: $\psi_{-5,0}$

Transformation (99) eine Korrelation des Signals mit einem Wavelet an einer bestimmten Stelle dar, d.h. je besser ein Signal mit einem Wavelet an entsprechender Stelle b korreliert (hinsichtlich der von a abhängigen Frequenz), umso größer wird der Wert $(W_\psi f)(a, b)$. Natürlich ist ersichtlich, dass die Wavelet-Transformation keine direkte Zeit-Frequenz-Analyse ermöglicht, schließlich ist $W_\psi f$ nur von a und b abhängig, wobei a keine Frequenz darstellt. Es ist jedoch möglich, einen Bezug zwischen Skalenwert und Frequenz herzustellen, jedoch wird auf eine Ausführung dieses Zusammenhangs an dieser Stelle verzichtet und stattdessen auf [1] verwiesen. In dieser Hinsicht folgt zusammengefasst für die Wavelet-Transformation, dass für große Werte von a nur niedrige Frequenzen und für kleine Werte nur hohe Frequenzen erfasst werden. Als Beispiel hierzu sei die Wavelet-Transformierte des Signals $g(t) := chirp_{2,6}(t)$ aufgeführt. In Abbildung (48) wird das Signal und seine Wavelet-Transformierte dargestellt. Der oberen Teil der Grafik stellt das eigentliche Signal g dar, im unteren Teil wird die dazugehörige 2-dimensionale Wavelet-Transformierte dargestellt. Entlang der x-Achse beider Graphen wird die Zeit aufgetragen, in diesem Fall von 0sec bis 4sec. Die y-Achse des Graphen der Wavelet-Transformierten stellt die jeweilige Skala dar. Um das ablesen des Graphen zu vereinfachen, wird dieser 2-dimensional dargestellt, d.h. die Funktionswerte $(W_\psi f)(a, b)$ werden an den entsprechenden Stellen durch einen Farbwert dargestellt, je dunkler dieser ist umso kleiner ist der korrespondierende Funktionswert.
Wie leicht zu erkennen ist, besitzt $(W_\psi f)(a, b)$ bei großen Skalenwerten nur zur Anfangszeit des Signals signifikante Werte (d.h. Werte $>> 0$). Dies ist entspricht der Erwartung, da die Wavelet-Transformierte für große Skalen-Werte, eine Korrelation des Signals mit einem "langgezogenen" und somit niederfrequenten Wavelet darstellt. Da das Signal eine mit der Zeit linear steigende Frequenz aufweist, erreicht die Korrelation mit einem niederfrequenten Wavelet nur zu Beginn des Signals signifikante Werte. Wie oben bereits festgestellt wurde,

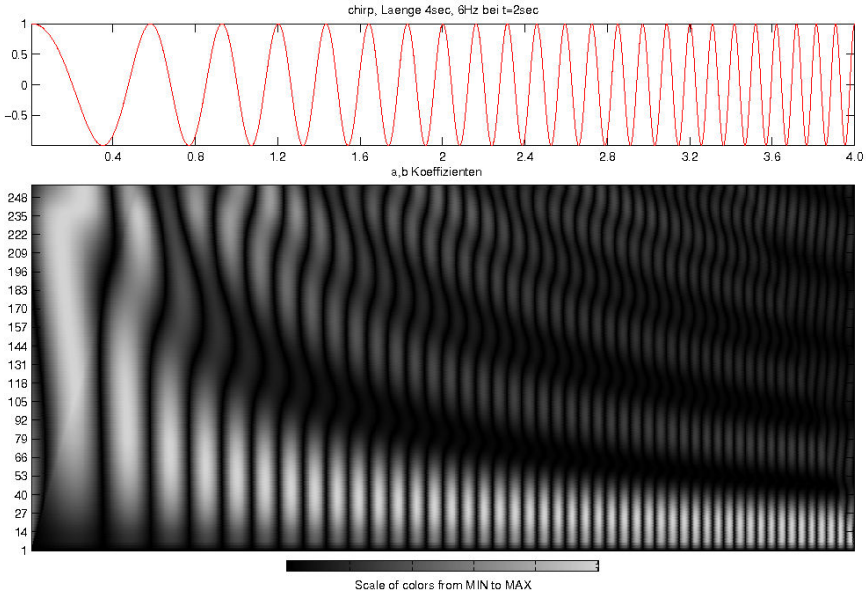


Abb. 48: CWT($chirp_{2,6}(t)$)

kann die Wavelet-Transformation als Faltung aufgefasst werden. Somit stellt $(W_\psi f)(a_0, b)$, für einen festen Skalen-Parameter a_0, die Filterantworten des Wavelets $D_{a_0}\psi$ an den Stellen b dar.

Dies mag auf den ersten Blick ein großer Vorteil gegenüber der STFT sein, jedoch entsteht bei der CWT eine große Redundanz hinsichtlich der Wavelet-Koeffzienten, d.h. es werden nicht alle Koeffizienten der CWT benötigt, um das analysierte Signal rekonstruieren zu können. Eine Lösung dieses Problems ist die Diskretisierung hinsichtlich der Parameter a und b, im Kapitel 6 wird diese für 2-dimensionale Signale und eine spezielle Wavelet-Basis vorgestellt.

6. Gabor Wavelets

6.1. Gabor-Funktionen

Mit Beginn dieses Abschnittes wenden wir uns mit der Wavelet- und Fourier-Transformation vollständig dem für diese Arbeit wichtigen Anwendungsgebiet zu, nämlich der digitalen Bildverarbeitung. Von nun an werden neu eingeführte Funktionen und Transformationen direkt auf den 2-dimensionalen Fall bezogen. Bevor der Begriff einer Gabor-Funktion erläutert wird, sollen die bisherigen und für die kommenden Abschnitte wichtigen Transformations-Vorschriften für den Fall 2-dimensionaler Signale angegeben werden. Die bekannte Fourier-Transformation für ein Signal $I \in L^2(\mathbb{R}^2)$ lautet somit

$$(\mathcal{F}I)(\vec{\omega}) = \int_{\mathbb{R}^2} I(\vec{x}) e^{-i\vec{\omega}^T \vec{x}} d\vec{x} \tag{112}$$

und für deren Umkehrung ebenfalls analog zum 1-dimensionalen Fall

$$(\mathcal{F}^{-1}\widehat{I})(\vec{x}) = \frac{1}{2\pi} \int_{\mathbb{R}^2} \widehat{I}(\vec{\omega}) e^{i\vec{\omega}^T \vec{x}} d\vec{\omega} \; . \tag{113}$$

Für die letztendliche Implementierung der Fourier-Transformation wird, wie bereits erwähnt, eine Diskretisierung dieser benötigt. Auch im Fall der DFT ist diese analog auf den höher dimensionalen Bereich anwendbar, für $\tau = \{0, ..., N-1\}^2$ und $W_N = e^{-i\frac{2\pi}{N}}$ gilt somit

$$(DFT\, I)(\vec{\rho}) = \frac{1}{N} \sum_{\vec{n} \in \tau} I(\vec{n}) W_N^{\vec{\rho}^T \vec{n}} \; , \tag{114}$$

für die inverse Transformation entsprechend

$$(IDFT\, \breve{I})(\vec{n}) = \frac{1}{N} \sum_{\vec{\rho} \in \tau} \breve{I}(\vec{\rho}) W_N^{-\vec{\rho}^T \vec{n}} \; . \tag{115}$$

Da nun ausschliesslich 2-dimensionale reelle Signale in Form von Bildern betrachtet werden, wird nicht mehr vom Zeitbereich sondern vom Ortsraum gesprochen (mit $\vec{x}$ als Positionsangabe).
Unter einer Gabor-Funktion wird eine Funktion der Form

$$\gamma(\vec{x}) = C e^{-\frac{1}{2}(\vec{x}-\vec{x_0})^T D (\vec{x}-\vec{x_0})} \cdot e^{-i\omega_0^T \vec{x}} \tag{116}$$

verstanden, D stellt eine Diagonalmatrix mit den Zahlen σ_1^{-2} und σ_2^{-2} dar, weiterhin gilt $C \in \mathbb{C}$. Der Vektor ω_0 legt die Frequenz der komplexen 2-dimensionalen Exponentialfunktion fest. Eine Gabor-Funktion ist also nichts anderes als eine gefensterte komplexe Exponential-Funktion. Als Fenster-Funktion wird hier wieder die Gauss-Funktion aus Kapitel 4 verwendet, wie nun bekannt ist, garantiert diese eine (im Sinne der Unschärferelation) optimale Ort-Frequenz-Lokalisierung. Jedoch sind dies nicht die einzigen Vorteile dieser Funktionen, im Hinblick auf die Nachbildung des Antwort-Verhaltens einfacher Zellen im menschlichen Sehapparat haben sich, wie in [12] erläutert wird, Gabor-Funktionen sehr bewährt. Der Begriff der Nachbildung sei hier so zu verstehen, dass Gabor-Funktionen als Profil einfacher Zellen angesehen werden können. Für diese Gruppe von Zellen wurden in Kapitel 2 insgesamt 5 Punkte aufgeführt, ohne Modifikation erfüllen Gabor-Funktionen 4 dieser Punkte. Im Sinne der Nachbildung des Verhaltens einfacher Zellen wird die Bedingung der Zulässigkeit leider nicht erfüllt, da der reelle Teil einer Gabor-Funktion durchaus auf konstante Illumination anspricht.

6.2. Gabor Funktionen als Wavelet-Transformation

Für die Zulässigkeit einer Gabor-Funktion ψ muss also gelten $\widehat{\psi}(0) = 0$. Eine Lösung dieses Problems ist nach [12], die Subtraktion einer bei $\omega_0 = 0$ zentrierten und entsprechend skalierten Gabor-Funktion derselben Breite

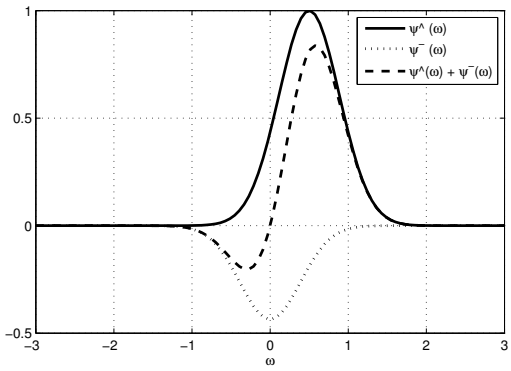


Abb. 49: Fourier-Transformierte einer Gauss-Funktion ψ mit $\widehat{\psi}(0) = 0$

wie $\widehat{\psi}$.
In Abb(49) wird ein Schnitt durch den Graph einer fouriertransformierten 2-dimensionalen Gabor-Funktion $\widehat{\psi}$ dargestellt. Es ist erkennbar, dass diese Funktion bei $\omega = 0$ ungleich 0 ist. Somit wird eine weitere Gabor-Funktion von gleicher Breite wie $\widehat{\psi}$, so skaliert, dass sie einen Maximalwert von $\widehat{\psi}(0)$ besitzt. Anschließend wird diese Funktion an die entsprechende Stelle verschoben und von der Funktion $\widehat{\psi}$ subtrahiert. Nach [12] ergibt sich somit für die Fourier-Transformierte

$$\widehat{\psi}(\vec{\omega}) = \frac{C2\pi}{\sqrt{\det D}}(e^{-\frac{1}{2}||\vec{\omega}-\vec{\omega_0}||_{D^{-1}}} - e^{-\frac{1}{2}(||\vec{\omega_0}||_{D^{-1}}+||\vec{\omega}||_{D^{-1}})}) , \tag{117}$$

hierbei wird $||\vec{\omega}||_D := \vec{\omega}^T D\vec{\omega}$ als abkürzende Schreibweise verwendet. Daraus folgt für die Funktion ψ im Ortsraum

$$\psi(\vec{x}) = Ce^{-\frac{1}{2}||\vec{x}||_D}(e^{-i\vec{\omega}_0\vec{x}} - e^{-\frac{1}{2}||\vec{\omega}_0||_{D^{-1}}}) . \tag{118}$$

Mit Hilfe auf dieser Weise modifizierter Gabor-Funktionen ist es nun möglich, das Antwort-Verhalten einfacher Zellen nachzubilden. Dem nicht genug, wie in [19] und [12] nachzulesen ist, erfüllen diese Gabor-Funktionen zusätzlich die Anforderungen an Wavelets. Dies impliziert natürlich, dass diese nun als Filter angesehen werden dürfen. Im weiteren Verlauf wird daher die Bezeichnung Gabor-Filter synonym mit Gabor-Wavelet verwendet. Die obige Erklärung stellt ein allgemeines Schema dar, um Gabor-Funktionen hinsichtlich der Anforderungen zulässig zu machen. Jedoch wird für die in dieser Arbeit angewandten Verfahren, folgendes Gabor-(Mother)Wavelet verwendet

$$\psi(\vec{x}) = \frac{1}{\sigma^2}e^{-\frac{1}{2\sigma^2}||\vec{x}||^2}(e^{i\vec{x}^T e_1} - e^{-\frac{\sigma^2}{2}}) , \tag{119}$$

mit $e_1 = (1, 0)^T$. Für die hier getroffene Wahl der Normalisierungsfaktoren wird auf [12] verwiesen.

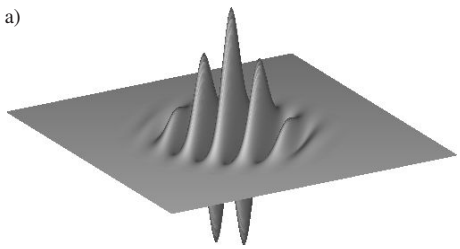


Abb. 50: Gabor-Filter Realteil

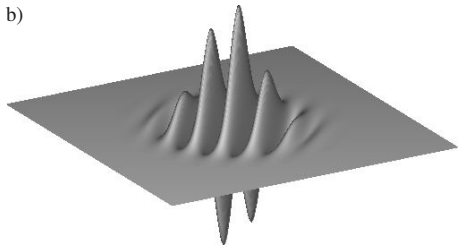


Abb. 51: Gabor-Filter Imaginärteil

Aus vorherigen Abschnitten ist bekannt, dass die Wavelet-Transformation mit Hilfe einer Faltung beschrieben

werden kann. Jedoch wurde zuvor ebenfalls darauf hingewiesen, dass es hinsichtlich des in diesem Kapitel benutzten Mother-Wavelets eine Besonderheit gibt. Für einen Gabor-Filter ψ gilt

$$\psi(-\vec{x}) = \overline{\psi(\vec{x})}\,, \tag{120}$$

in Worten bedeutet dies, dass $\Re\psi$ eine gerade und $\Im\psi$ eine ungerade Funktion darstellt. Eine Folge davon ist, dass ψ eine reelle Fourier-Transformierte besitzt. Das Hauptaugenmerk sei jedoch auf die Darstellung der Wavelet-Transformation als Faltung gerichtet, es gilt unter Nutzung der oben genannten Eigenschaft und $\mathcal{Q}(\vartheta) = \begin{pmatrix} \cos(\vartheta) & \sin(\vartheta) \\ -\sin(\vartheta) & \cos(\vartheta) \end{pmatrix}$

$$\begin{aligned}
(W_\psi I)(a, \vec{x}_0, \vartheta) &= \frac{1}{a^2}\int_{\mathbb{R}^2} I(\vec{x})\overline{\psi(\frac{1}{a}\mathcal{Q}(\vartheta)(\vec{x}-\vec{x_0}))}d\vec{x} && (121)\\
&= \frac{1}{a^2}\int_{\mathbb{R}^2} I(\vec{x})\psi(\frac{1}{a}\mathcal{Q}(\vartheta)(\vec{x_0}-\vec{x}))d\vec{x} && (122)\\
&= (\psi_{a,\vec{0},\vartheta} * I)(\vec{x}_0)\,, && (123)
\end{aligned}$$

somit ändert sich die Faltungsvorschrift geringfügig bezüglich der allgemeinen Definition. Hierbei wurde hinsichtlich der Wavelet-Transformation ein 3. Parameter ϑ eingeführt. Dieser wird als Rotations-Parameter bezeichnet und dient der Rotation des Gabor-Filters. Dies ist nötig da, einfache Zellen im menschlichen Sehapparat, wie bereits erwähnt, eine Orientierungsselektivität aufweisen.
Wie schon im vorherigen Beispiel der CWT, ist auch in diesem Fall die Redundanz hinsichtlich der Filterantworten sehr groß. Es gibt jedoch einen Weg, diese Redundanz zu vermeiden ohne nennenswerten Informationsverlust hinnehmen zu müssen. Hierzu sei daran erinnert, dass ein Wavelet einen Bandpass-Filter darstellt, d.h. jede Funktion aus der Wavelet-Basis lässt nur ein spezifisches Frequenzband hindurch. Da prinzipiell bei der CWT unendlich viele Funktionen benutzt werden, liegt die Vermutung nah, dass sich deren Frequenz-Passbänder überlappen. Weiterhin besitzt das zu transformierende Signal natürlich eine Grenzfrequenz, die Grundidee besteht nun darin, mit möglichst wenig Basisfunktionen den Frequenzbereich bis zur Grenzfrequenz abzudecken. Zur besseren Veranschaulichung sei hier Abb.(52) erwähnt, welche mehrere Gabor-Filter im Frequenzraum darstellt. Wie zu erkennen ist, kann mit entsprechender Anordnung der Kerne im Frequenzraum, ein ausreichend großer Frequenzbereich abgedeckt werden. Um dies zu erreichen wird die Wavelet-Transformation

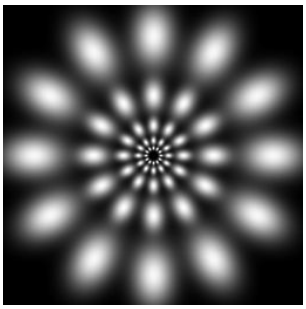


Abb. 52: Gabor-Kerne im Frequenzraum

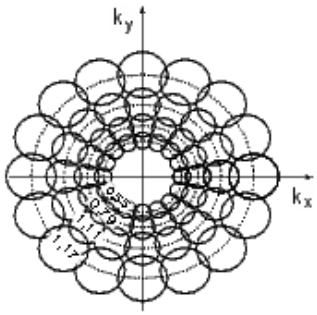


Abb. 53: Diskretisierung der Parameter ϑ und a

bezüglich der Parameter ϑ und a diskretisiert. Bevor nun das hierzu gehörende Diskretisierungs-Schema angegeben wird, sollen die Hintergründe für dessen Wahl betrachtet werden. Schreiben wir hierzu zunächst den Ausdruck $\psi(\frac{1}{a}\mathcal{Q}(\vartheta)\vec{x})$ äquivalent um in

$$\Psi_{\vec{k}}(\vec{x}) = \frac{||\vec{k}||^2}{\sigma^2} e^{\frac{||\vec{k}||^2||\vec{x}||^2}{2\sigma^2}} \left[e^{i\vec{k}^T\vec{x}} - e^{-\frac{\sigma^2}{2}} \right]\,, \tag{124}$$

hierbei wird $\vec{k}$ als Wellenvektor bezeichnet. Jedoch ist für die Wahl der Diskretisierung, die Betrachtung des Frequenzraums nötig, somit wird die Fourier-Transformierte zu $\Psi_{\vec{k}}$ bestimmt

$$\widehat{\Psi}_{\vec{k}}(\vec{\omega}) = e^{-\frac{\sigma^2}{2||\vec{k}||^2}(\vec{k}-\vec{\omega})^2} - e^{-\frac{\sigma^2}{2||\vec{k}||^2}(||\vec{k}||^2-||\vec{\omega}||^2)} \ . \tag{125}$$

Es ist bereits ersichtlich, dass $\vec{k}$ eine Verschiebung im Frequenzraum bewirkt. Hierbei gilt für den Wellenvektor

$$\vec{k} = k_\nu \begin{pmatrix} \cos(\phi_\mu) \\ \sin(\phi_\mu) \end{pmatrix} \ . \tag{126}$$

Die Zahlen k_ν sind definiert als

$$k_\nu = a_0^\nu a_{\text{min}} \ , \quad \nu = 0, ..., M-1 \ , \tag{127}$$

wobei a_{min} den Startwert des Dilatations-Parameters a darstellt und a_0 die Schrittweite (d.h. die Erhöhung des Parameters) hinsichtlich jeder Inkrementierung von ν bestimmt. Weiter gilt für ϕ_μ

$$\phi_\mu = \mu \frac{\pi}{L} \ , \quad \mu = 0, ..., L-1 \ , \tag{128}$$

somit resultieren aus einer Gabor-Transformation insgesamt $M \cdot L$ 2-dimensionale Filterantworten. An dieser Stelle seien nochmals die Gleichungen (124) und (125) betrachtet. Wie aus der Regel (R3) in Kapitel 3 hervorgeht, resultiert aus einer Dilatation im Ortsraum mit Faktor $\frac{1}{a}$ eine reziproke Wirkung im Frequenzraum. Weiterhin stellt $\vec{k}$ Punkte entlang eines Kreises mit Radius k_ν dar, in Abb.(53) wären dies jene Zentren der dargestellten Kreise. Da die Funktionen $\Psi_{\vec{k}}$ lokalisiert und um $\vec{k}$ verschoben sind, kann jeder der dargestellten Kreise als Passband einer solchen Funktion angesehen werden. Unter Berücksichtigung obiger Regel ist somit ersichtlich, dass nach dem Diskretisierungs-Schema alle Kreise für einen festen Winkel dicht bzw. überlappend aneinander liegen können. In erwähnter Grafik bezeichnet k_x die x-Komponente und k_y die y-Komponente des Vektors $\vec{k}$. Wird diese Transformation zusätzlich hinsichtlich der Signalwerte diskretisiert, so wird von einer diskreten Wavelet-Transformation (DWT) gesprochen. Eine sehr gute Darstellung des dazugehörigen Samplings wird in [12] gegeben.
Weitergehende Informationen im Bezug auf die Gabor-Transformation sind in [19],[13] und [12] zu finden.
Wie die STFT, kann auch die Gabor-Transformation zur Texturanalyse eingesetzt werden, d.h. durch $(W_\psi I)(a, \vec{x}, \vartheta)$ werden lokale Texturmerkmale repräsentiert. Jedoch fällt bei der Gabor-Transformation (bzw. bei Wavelet-Transformationen allgemein) die einschränkende feste Fenstergröße der STFT weg. War die Darstellung der CWT in Kapitel 5 schon schwierig, so ist sie in diesem Fall nicht mehr anschaulich in einem Graph realisierbar, schließlich wird jeder Funktionswert von 3 Parametern bestimmt (abgesehen davon, dass ein Bild ein 2-dimensionales Signal darstellt). Es ist natürlich möglich $W_\psi I$ in mehreren Graphen darzustellen, betrachten wir hierzu die Gabor-Filterantworten in Abbildung (55). In dieser Grafik wird die Gabor-Transformierte bezüglich 4 Skalen und 4 Richtungen zum Bild (54) dargestellt. Es sei noch erwähnt, dass in dieser Abbildung jeweils der Betrag einer Filterantwort gezeigt wird, da ein Gabor-Wavelet im Gegensatz zu einem Haar-Wavelet, eine komplexwertige Funktion bildet. In der praktischen Anwendung werden natürlich deutlich mehr Skalen und Richtungen berücksichtigt, die dargestellten Filterantworten sollen nur der Veranschaulichung dienen.

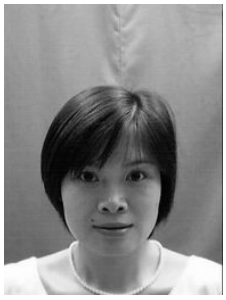


Abb. 54: Ausgangsbild für Gabor-Transformation

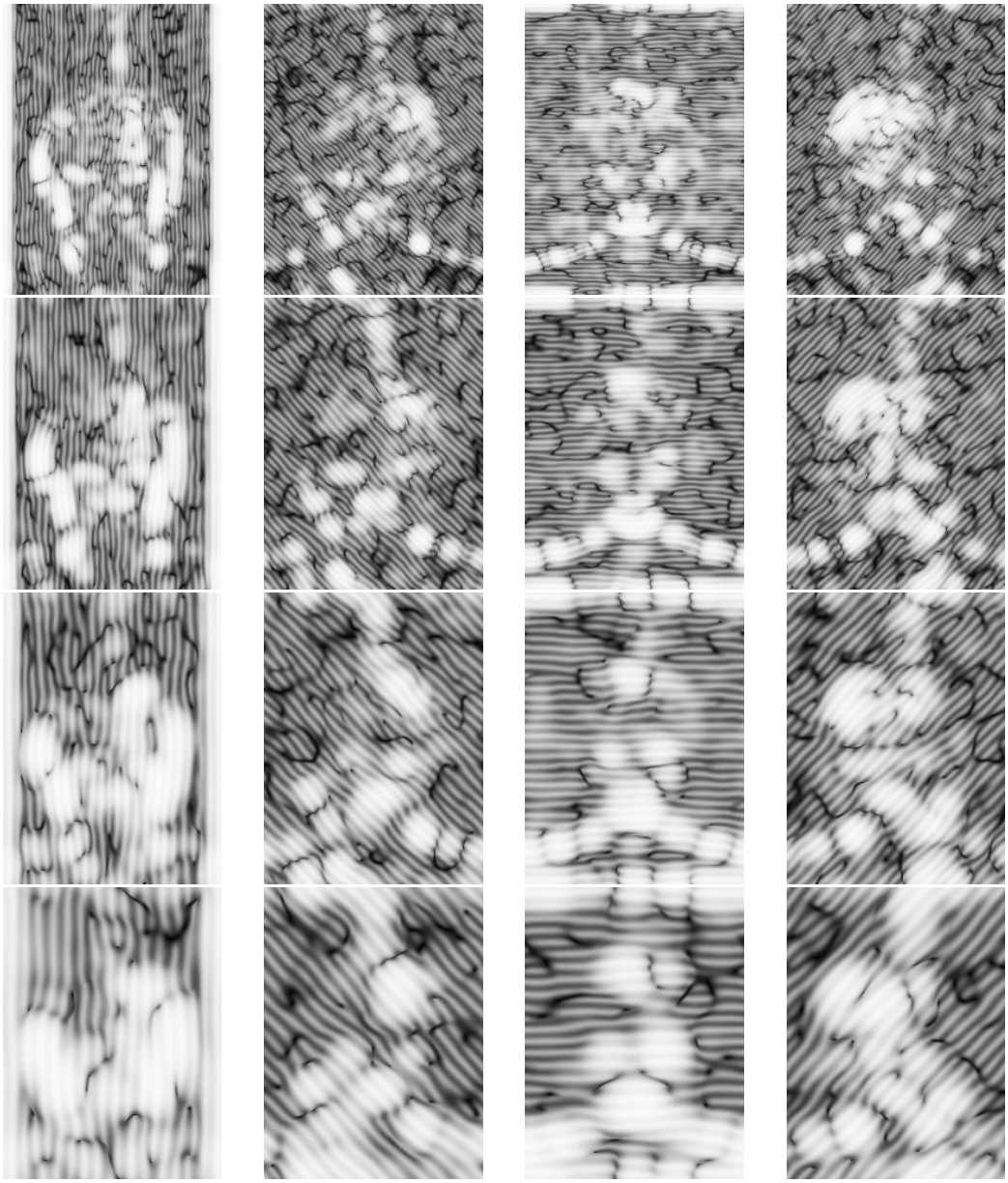


Abb. 55: Diskrete Gabor-Transformation bezüglich 4 Skalen und 4 Orientierungen (Winkel), in jedem Bild ist das Betragsspekrum dargestellt. Von links nach rechts betrachtet nimmt die Rotation des Gabor-Wavelets beginnend bei 0 jeweils um $\frac{\pi}{4}$ zu. Die Skala nimmt von oben nach unten um $2(\sqrt{2})^i$, $i = 0, ..., 3$ zu.

7. Datenredundanz und Bilddatenkompression

Die STFT bzw. DWT ist nicht nur zur Textur-Analyse geeignet, ein weiteres sehr verbreitetes Einsatzgebiet im Bereich der Bildverarbeitung ist die Bilddatenkompression. In diesem Kapitel wird nochmals der Analyse und Synthese von Signalen nachgegangen, jedoch mit dem Unterschied, dass nicht alle bei der Analyse bestimmten Koeffizienten zur Funktions-Synthese herangezogen werden. Um diesen Grundgedanken zu veranschaulichen, betrachten wir Abb.(7). Hier sind insgesamt 4 Bilder einer Person dargestellt, das erste von links stellt das

Abb. 56: Kompression mit DWT und Haar-Wavelet, Schwellwert-Quantisierung

Ausgangsfoto dar, die weiteren wurden mit Hilfe der DWT unter Benutzung des Haar-Wavelets verlustbehaftet komprimiert. Hierfür wurde die so genannte Multi-Skalen-Analyse (MSA) eingesetzt, dies sei nur der Vollständigkeit halber erwähnt und für Details auf [1],[3],[7] und [2] verwiesen. Es ist zu erkennen, dass die Qualität nach rechts hin abnimmt, jedoch sinkt ebenfalls das Datenvolumen des Bildes. Die Einsatzgebiete für Bilddatenkompression sind hauptsächlich

- Digitale Bildübertragung: Ohne Reduzierung des Datenvolumens hochauflösender Bilder, kann die Kapazität des Übertragungsmediums überschritten werden.
- Archivierung digitaler Bilder: Eine Bessere Ausnutzung des vorhandenen Speichers ist möglich

Der eigentliche Vorgang zur verlustbehafteten Bilddatenkompression, so wie sie im Allgemeinen verwendet wird, lässt sich durch Abb.(57) beschreiben. Nachfolgend wird kurz auf die dort dargestellten Komponenten eingegangen .

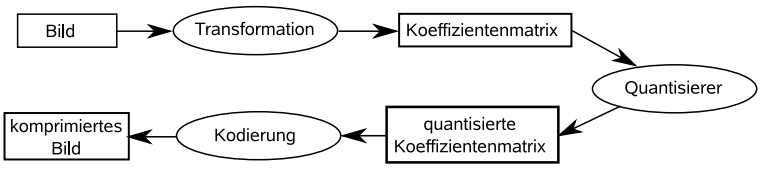


Abb. 57: Komponenten eines Kompressionsalgorithmus

Zunächst sei erwähnt, dass es sich bei Bildern um Signale aus $\mathbb{R}^{m\times n}$, $m, n \in \mathbb{N}$ handelt, welche somit durch Matrizen dargestellt werden können. Im weiteren Verlauf dieses Kapitels wird ein Bild bzw. das Bildsignal mit I bezeichnet. Der erste Schritt im angegebenen Schema ist die Transformation des Bildes. Hierfür wird eine lineare und bijektive Transformation der Form

$$T : \mathbb{R}^{m\times n} \to \mathbb{R}^{m\times n} \text{ bzw. } T : \mathbb{R}^{m\times n} \to \mathbb{C}^{m\times n} \tag{129}$$

benötigt. Etablierte Kompressionsverfahren nutzen hierfür u.A. diskretisierte Formen der Fourier-, Kosinus- und Wavelet-Transformation.

Die nach erfolgreicher Transformation resultierenden Koeffizienten $I_T = TI$ werden anschließend quantisiert. Unter dem Begriff der Quantisierung wird das Abbilden der Koeffizienten I_T auf ganzzahlige Werte I_Q verstanden. Sehr häufig werden in dieser Hinsicht so genannte Look-Up-Tabellen verwendet. Diese dienen dazu, die Koeffizienten auf bestimmte Intervalle abzubilden. Um ein Beispiel zu nennen, kann ein Quantisierungsalgorithmus alle Elemente von I_T, welche einen festgelegten Schwellwert unterschreiten, auf 0 setzen und den restlichen eine ganze Zahl zuordnen. Dies wurde im oben angegebenen Beispiel der Kompression angewandt und stellt einen irreversiblen Prozess dar. Um die Irreversibilität der Quantisierung zu veranschaulichen, betrachten wir diesbezüglich die Abb.(59), welche eine Folge von diskreten Werten dargestellt. Werden diese unter Benutzung der in Abb.(58) angegebenen Look-Up-Tabelle quantisiert, so ergeben sich die in Abb.(60) gezeigten Werte. Der Quantisierungsalgorithmus bestimmt somit die Qualität des komprimierten Bildes, schließlich zieht

Alter Wert	Neuer Wert
1	1
2	1
3	2
4	2
5	3
6	3
7	4
8	4

Abb. 58: Look-Up-Tabelle

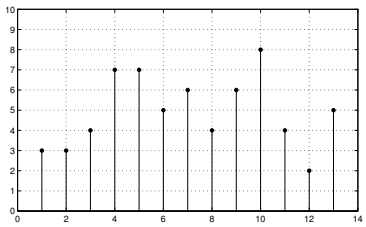


Abb. 59: Ausgangswerte

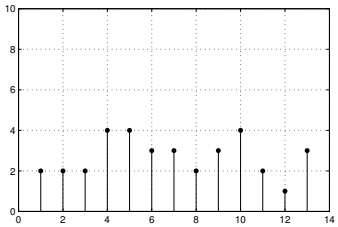


Abb. 60: Quantisierte Werte

der Einsatz einer Look-Up-Tabelle einen Informationsverlust nach sich.
Natürlich wird das Datenvolumen durch einfaches Quantisieren nicht reduziert, d.h. die Anzahl der Koeffizienten bleibt gleich. Aus dem Prozess des Quantisierens resultiert eine verkleinerte Entropie H der Datenmenge, welche nach folgender Formel bestimmt wird

$$H = -\sum_{i \in I_T} p(i) \log_2(i) \; . \tag{130}$$

Der Ausdruck $p(.)$ stellt die Wahrscheinlichkeit mit welcher der Signalwert auftreten wird dar. Für den Fall des oben angegebenen Beispiels beträgt die Entropie der Originalwerte $H_{orig} = 2.71$ und die der neuen Werte $H_Q = 1.82$. Nach [15] und [16] kann der Begriff der Redundanz als Maß für die Menge zusätzlicher Daten angesehen werden, welche zur Repräsentation einer Information nicht benötigt werden. Ohne an dieser Stelle tief auf die Grundlagen der Kodierungstheorie einzugehen, sei nur erwähnt, dass eine Reduktion der Entropie eine Erhöhung der Datenredundanz bewirkt. Dieser Umstand wird im anschließenden und letzten Schritt, der Kodierung, zur Reduktion der eigentlichen Datenmenge ausgenutzt. Hierbei übernehmen meist Entropie- bzw. Lauflängenkodierer die eigentliche Reduktion des Datenvolumens. Auch in diesem Schritt soll ein kleines Beispiel das Verständnis vereinfachen. Wir betrachten die Huffman-Kodierung einer Zahlenfolge

$$9 \; , \; 123 \; , \; 17 \; , \; 63 \; , \; 129 \; , \; 17 \; , \; 123 \; , \; 123 \; , \; 52 \; .$$

Wird für jede Zahl ein Byte verwendet, so werden hierfür 72-Bit benötigt. Mit Hilfe der Huffman-Kodierung werden nun Code-Wörter für jedes Element dieser Folge bestimmt. Es ergeben sich

$$\begin{aligned} 123 &\mapsto 0 \\ 17 &\mapsto 10 \\ 63 &\mapsto 1100 \\ 9 &\mapsto 1101 \\ 129 &\mapsto 1110 \\ 52 &\mapsto 1111\,. \end{aligned}$$

Werden nun die Elemente der Folge durch die entsprechenden Code-Wörter dargestellt, so resultiert daraus folgende Bitfolge

$$110101011001110100011 11\,.$$

Der benötigte Speicherplatz wurde somit auf 23 Bits reduziert, nach [1] wird die Kompressionsrate k definiert als

$$k = \frac{S_{\text{orig}}}{S_{\text{komp}}}\,, \tag{131}$$

mit S_{orig} als Datenvolumen des Originalbildes und S_{komp} der des Komprimierten. Im dargestellten Beispiel, ergibt sich eine Kompressionsrate von 3.13 und es werden im Schnitt nur $2.56\frac{\text{Bits}}{\text{Zahl}}$ benötigt.
Bisher wurde noch keine allgemeine Rekonstruktionsformel für die CWT bzw. DWT angegeben, dies liegt darin begründet, dass die in dieser Arbeit benutzten Verfahren das Ursprungsbild nicht mehr rekonstruieren. Vielmehr werden die Wavelet-Koeffizienten bzw. Gabor-Filterantworten als neue Bildrepräsentation verwendet. Nun liegt die Frage nahe, warum überhaupt das Thema der Bilddatenkompression angesprochen wird. Ohne zu sehr auf nachfolgende Kapitel vorzugreifen, sei an dieser Stelle gesagt, dass in dieser Arbeit ebenfalls Wavelet-Koeffizienten auf Redundanz untersucht werden (siehe Kapitel 9). Jedoch liegen diese Koeffizienten in einer anderen Darstellung vor bzw. der Kontext, d.h. die Motivation, ist unterschiedlich. Wurden bisher Kompressionsverfahren hinsichtlich der Kompressionsrate bzw. des Fehlers zum Originalbild beurteilt, so werden für diese Arbeit andere Kriterien zur Beurteilung der Kompression benutzt.
Diese kurze Einführung in die allgemeine Bilddatenkompression sollte den Begriff der Redundanz verdeutlichen, da im bereits angesprochenen Kapitel 9 ein Verfahren vorgestellt wird, welches eine Menge von Gabor-Filterantworten auf Redundanz prüft. Diese Wavelet-Koeffizienten werden im EBGM-Algorithmus zur Lokalisierung markanter Bildpunkte verwendet, welche anschließend für die Objekterkennung eine wichtige Rolle spielen.

8. Elastic Bunch Graph Matching

8.1. Einleitung

Wie zu Beginn bereits erwähnt wurde, stellt das Korrespondenzproblem ein signifikantes Problem im Bereich der Gesichtserkennung dar. Das Elastic-Bunch-Graph-Matching (EBGM) bietet eine Lösung hierzu, es nutzt dabei die zuvor erläuterten Vorteile der Gabor-Transformation zur Texturanalyse aus. Jede Gabor-Filterantwort stellt einen Texturdeskriptor dar, welcher die Textur an der entsprechenden Stelle eines Bildes repräsentiert. Natürlich spielen hier auch die neurobiologischen Aspekte eine wichtige Rolle, diesbezüglich wird jedoch auf [12],[18] und [14] verwiesen.
Zum Objektvergleich müssen zunächst die entsprechenden Objekte innerhalb der vorgegebenen Bilder lokalisiert werden. Andernfalls wäre es nicht möglich, einen Objektvergleich durchzuführen, da z.B. ein Gesicht im Bild A mit einem Bereich des Hintergrundes von Bild B verglichen würde. Um dieses Problem zu lösen werden so genannte Landmarken gesetzt, eine Landmarke repräsentiert die Position eines wichtigen Merkmals des gesuchten Objektes. So werden im Fall eines Gesichtes Landmarken an Stellen wie beispielsweise Augen,Kinn,Mund,... gesetzt, Landmarken werden ebenfalls als markante Punkte bezeichnet. Beim EBGM-Verfahren werden diese Landmarken durch die Knoten eines Graphen repräsentiert. Für den anschließenden Objektvergleich werden diese Graphen herangezogen. Jedoch setzt dieser Ansatz voraus, dass bereits entsprechende Graphen für jedes Bild eines Sets existieren, hierbei sei der Begriff des Sets als Bildmenge zu verstehen. Somit müssen für jedes unbekannte Bild zunächst die entsprechenden Landmarken gesetzt werden. Die eigentliche Aufgabe des EBGM-Verfahrens besteht darin, für ein Set bei welchem noch keine Landmarken gesetzt wurden, diese zu bestimmen. D.h. mit Hilfe dieses Verfahrens ist es nicht nötig, für ein unbekanntes Bild die Landmarken manuell zu setzen. An dieser Stelle sei erwähnt, dass der EBGM-Algorithmus somit kein Verfahren zur Objekterkennung bzw. zum Objektvergleich darstellt, er dient ausschließlich der Lokalisierung markanter Punkte.
Bevor der eigentliche Algorithmus erklärt wird, sollen die für diesen wichtige Begriffe kurz erläutert werden.

8.2. Merkmalsrepräsentation

8.2.1. FeaSt / Jet

Die Bezeichnung FeaSt ist eine Abkürzung für ”Feature Set”. Ein FeaSt kann unterschiedliche Informationen beinhalten, wie z.B. die RGB-Werte eines Farbbildes. Im Bezug auf das EBGM-Verfahren, beinhaltet ein FeaSt die Gabor-Filterantwort zu einer bestimmten Bildposition. Hat das FeaSt eine vektorielle Form, so wird es als Jet bezeichnet. Betrachten wir zum besseren Verständnis Abb.(61) , hier wird der Zusammenhang zwischen

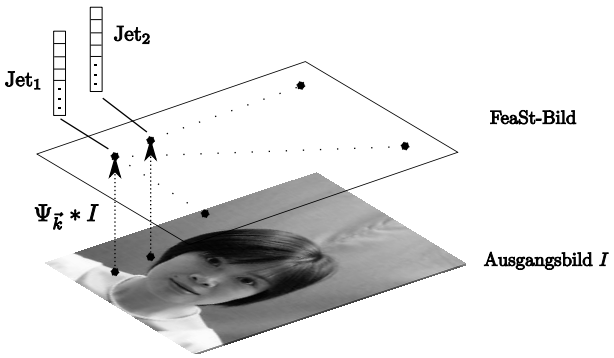


Abb. 61: Feast-Bild

Ausgangsbild und dem so genannten FeaSt-Bild $\mathcal{I}$ dargestellt. Ein FeaSt-Bild hat die gleichen Dimensionen wie das Ausgangsbild, jedoch beinhaltet es keine Intensitäts-/Farbwerte an den entsprechenden Stellen. Stattdessen

werden zu jeder Bildposition die entsprechenden Gabor-Filterantworten berechnet und in Form eines Jets $\mathcal{J}$ gespeichert. Mathematisch betrachtet, stellt das FeaSt-Bild eine $X \times Y \times M \cdot L$ Matrix dar, wobei X und Y die Dimensionen des Ausgangsbildes sind und $M \cdot L$ für die Anzahl der verschiedenen Filterparameter steht (M stellt die Skalenanzahl und L die Menge der Ausrichtungen dar). Ein FeaSt-Bild stellt somit eine abstrahierte Darstellung des eigentlichen Fotos dar, Intensitätswerte wurden durch Texturdeskriptoren ersetzt. Der EBGM-Algorithmus verwendet ausschließlich diese Bildrepräsentation zur Lokalisierung markanter Punkte.

8.2.2. Modellgraphen / Bündelgraphen

Zur Lokalisierung markanter Punkte wird ein so genannter Bündelgraph (englisch: Bunchgraph) verwendet, wobei dieser aus einzelnen Modellgraphen gebildet wird. Ein Modellgraph stellt einen beschrifteten Graphen dar und repräsentiert jeweils ein Gesicht. Ausgehend von Landmarken $\vec{x}_n,\ n = 1, ..., N$, wird jeder Knoten eines solchen Graphen mit einem Jet $\mathcal{J}_n$, bezogen auf die jeweilige Landmarke, beschriftet. Die Kanten eines Modellgraphen werden mit 2-dimensionalen Abstandsvektoren $\Delta\vec{x}_e = \vec{x}_n - \vec{x}_{n'},\ e = 1, ..., E$, versehen, wobei e die Beschriftung der Verbindungskante zwischen Knoten n und n' darstellt. Jeder Modellgraph besitzt somit N Knoten und E Kanten (siehe Abb.(62)). Natürlich kann nun argumentiert werden, dass ein solcher Graph kein Gesicht repräsentiert, sondern vielmehr das entsprechende Bild des Gesichtes, schließlich hängen an jedem Knoten nur auf das Bild bezogene Texturdeskriptoren. Jedoch kann ein Gesicht innerhalb eines Fotos ebenfalls unterschiedlich repräsentiert werden, beispielsweise durch unterschiedliche Expressionen (offene/geschlossene Augen, lachen, ...). Der im weiteren Verlauf dieser Arbeit benutzte Begriff der Gesichtsdarstellung mittels Graphen, sei auf diese Argumentation bezogen. Der Bündelgraph wird nun aus P voneinander verschiedenen Mo-

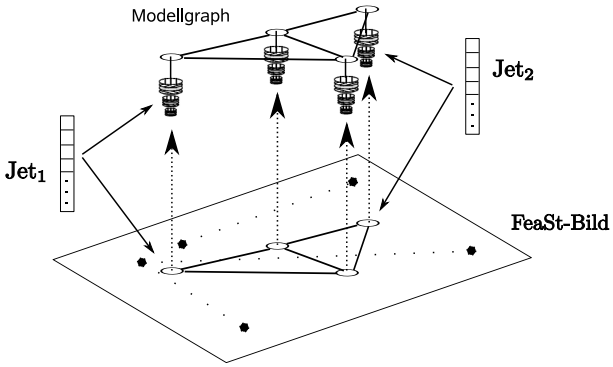


Abb. 62: Modellgraph

dellgraphen $\mathcal{G}^{\mathcal{B}p},\ p = 1, ..., P$ gebildet, natürlich sind die benutzten Modellgraphen hinsichtlich ihrer Struktur identisch, nur die Jets der Knoten unterscheiden sich so, dass keine identischen Graphen verwendet werden. Weiterhin sind diese Modellgraphen immer noch auf die entsprechende Gesichtsklasse bezogen. D.h. es werden Modellgraphen verwendet, welche z.B. nur männliche/weibliche Gesichter mit geschlossenen Augen repräsentieren, sich jedoch in anderen Arten der Expression durchaus unterscheiden können. In Abb.(63) wird das Prinzip der Erstellung eines Bündelgraphen verdeutlicht, der entstehende Graph besitzt natürlich die selbe Anzahl von Knoten und Kanten wie die verwendeten Modellgraphen. Eine Menge von Jets, welche sich auf dieselbe, d.h. in ihrer Bedeutung gleiche Landmarke beziehen, werden als ein Jet-Bündel $\mathcal{J}_n^{\mathcal{B}p}$ bezeichnet. Jeder Knoten n des Bündelgraphen wird mit dem entsprechenden Jet-Bündel beschriftet, die Kanten des Bündelgraphen tragen die gemittelte Distanz $\Delta\vec{x}_e^{\mathcal{B}} = \sum_p \Delta\vec{x}_e^{\mathcal{B}p}/P$ der Kanten jedes Modellgraphen. Ein Bündelgraph deckt somit eine große Variation von Gesichtern einer bestimmten Gesichtsklasse ab. Der EGBM-Algorithmus nutzt dies zur Lokalisierung von Landmarken aus, indem er den Punkt eines neuen FeaSt-Bildes bestimmt, welcher am besten mit einem bestimmten Jet-Bündel korrespondiert. Auf diese Weise wird ein Gesicht innerhalb eines unbekannten Bildes, mit Hilfe der Kombination von Elementen bekannter Gesichter lokalisiert. Diese Prozedur wird im nachfolgenden Abschnitt des Elastic Graph Matchings erläutert.

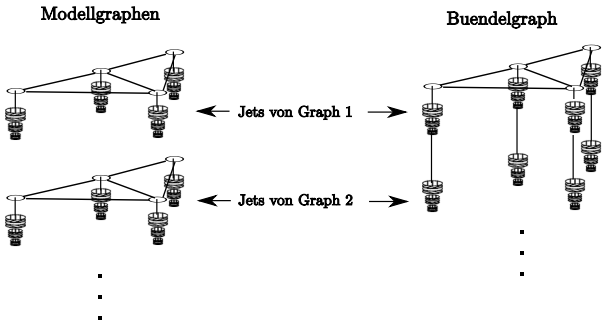


Abb. 63: Bündelgraph

8.3. Elastic Bunch Graph Matching

8.3.1. Elastic Graph Matching

Ausgehend von einem Bündelgraphen, welcher genügend Informationen bezüglich der benutzten Gesichtsklasse enthält, können mit Hilfe des Elastic Graph Matchings Landmarken in unbekannten Bildern dieser Klasse lokalisiert werden. Hierzu wird für das zu untersuchende Bild, zunächst ein entsprechendes FeaSt-Bild generiert. Im Anschluss wird der Bündelgraph in so genannten Moves über dieses FeaSt-Bild bewegt, dabei wird nach jeder erfolgten Bewegung die Ähnlichkeit jedes Jet-Bündels mit dem "darunter liegenden" Jet bestimmt. Der Begriff der Moves wird später erläutert und sei an dieser Stelle als Bezeichnung für eine bestimmte Art der Bewegung eines Graphen zu verstehen. Da es sich bei den Elementen eines Jets im Allgemeinen um komplexe Werte handelt, lässt sich ein solches Jet-Element ebenfalls in der Form $\mathcal{J}_j = a_j e^{i\phi_j},\ j = 1, ..., M \cdot L$ schreiben. Wobei hier die Phase $\phi_j(\vec{x})$ und der zugehörige Betrag $a_j(\vec{x})$ natürlich von der Position im Bild abhängig sind. Nach [9] werden folgende Ähnlichkeitsfunktionen zum Vergleich zweier Jets verwendet,

$$\mathcal{S}_a(\mathcal{J}, \mathcal{J}') := \frac{\sum_j a_j a'_j}{\sqrt{\sum_j a_j^2 \sum_j a'^2_j}} \tag{132}$$

und

$$\mathcal{S}_\phi(\mathcal{J}, \mathcal{J}') := \frac{\sum_j a_j a'_j \cos(\phi_j - \phi'_j - \vec{d}^T\ \vec{k}_j)}{\sqrt{\sum_j a_j^2 \sum_j a'^2_j}}\ . \tag{133}$$

(132) stellt hierbei ein Maß für die Ähnlichkeit der Beträge eines Jet-Paars dar, wobei analog dazu wird mit (133) ein Maß für die Ähnlichkeit hinsichtlich der Phasenlage definiert. In letzterer Gleichung wird der Term $\vec{d}^T\ \vec{k}_j$ verwendet, dieser dient der Kompensation der schnell variierenden Phasenlage, für Details hinsichtlich der Bestimmung dieses Terms sei auf [9] verwiesen. In Abb.(64) wird der Vorgang des so genannten Graph Matchings illustriert. Unter dem Begriff des Graph Matchings wird allgemein das Maximieren der Ähnlichkeitsfunktion verstanden, dies geschieht wie bereits erwähnt durch das Bewegen des Bündelgraphen über das jeweilige FeaSt-Bild. Jedoch werden hierbei pro Knoten nicht einzelne Jets miteinander verglichen, sondern jeweils ein Jet-Bündel mit einem Jet. Die hierfür eingesetzten Gleichungen lauten

$$\mathcal{S}_\mathcal{B}(\mathcal{G}^\mathcal{I}, \mathcal{B}) = \frac{1}{N} \sum_n max_p(\mathcal{S}_\phi(\mathcal{J}_n^\mathcal{I}, \mathcal{J}_n^{\mathcal{B}p})) - \frac{\lambda}{E} \sum_e \frac{(\Delta\vec{x}_e^\mathcal{I} - \Delta\vec{x}_e^\mathcal{B})^2}{(\Delta\vec{x}_e^\mathcal{B})^2} \tag{134}$$

$$\mathcal{S}_\mathcal{B}(\mathcal{G}^\mathcal{I}, \mathcal{B}) = \frac{1}{N} \sum_n max_p(\mathcal{S}_a(\mathcal{J}_n^\mathcal{I}, \mathcal{J}_n^{\mathcal{B}p})) - \frac{\lambda}{E} \sum_e \frac{(\Delta\vec{x}_e^\mathcal{I} - \Delta\vec{x}_e^\mathcal{B})^2}{(\Delta\vec{x}_e^\mathcal{B})^2}\ , \tag{135}$$

je nachdem ob die Ähnlichkeit mit Hilfe der Phase oder des Betrags bestimmt werden sollen. Die letzte Summe innerhalb beider Gleichungen wird als topologischer Term bzw. als topologische Kosten bezeichnet, wobei der

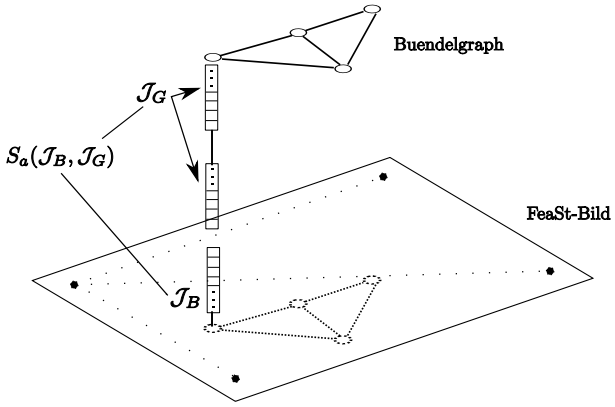


Abb. 64: Graph-Matching

Faktor λ die Signifikanz dieses Terms bestimmt, dies wird in der Erläuterung der Moves verdeutlicht. Weiterhin sei vermerkt, dass Gleichung (134) abhängig ist vom benutzten Bündelgraph $\mathcal{B}$ und einem Graph $\mathcal{G}^{\mathcal{I}}$, letzterer stellt die entsprechenden Punkte des FeaSt-Bildes $\mathcal{I}$ "unterhalb" des bewegten Graphen $\mathcal{B}$ dar. Die Maximum-Funktion innerhalb der ersten Summe dient der Auswahl des sogenannten "local expert" für den jeweiligen Knoten n. Der "local expert" ist jener Jet des Bündelgraphen, welcher dem entsprechenden Jet $\mathcal{J}_n^{\mathcal{I}}$ des Graphen $\mathcal{G}^{\mathcal{I}}$ am ähnlichsten ist. Der Begriff des Elastic Bunch Graph Matchings liegt in den zuvor erwähnten Moves begründet, in welchen der Bündelgraph nicht nur über das Bild geschoben, sondern zusätzlich skaliert und lokal verzerrt wird.

8.3.2. Moves

Moves bezeichnen die Art der Bewegung des Bündelgraphen über das FeaSt-Bild, innerhalb des EBGM-Verfahrens ist der so genannte Scan-Global-Move meist die erste Bewegungsform des Graphen. Dieser Move dient zur groben Lokalisierung des gesuchten Gesichtes. Hierbei wird der Bündelgraph in diskreten Schritten über das FeaSt-Bild geschoben (Abb.65), wobei für jede Verschiebung die Ähnlichkeit des Bündelgraphen mit den entsprechenden Jets im FeaSt-Bild bestimmt wird. Hierfür kann jeweils eine der oben erwähnten Ähnlichkeitsfunktionen verwendet werden, mit anderen Worten gesagt, kann ein und der selbe Move zuerst mit (134) und im Anschluss dazu nochmals, jedoch mit (135) durchgeführt werden.

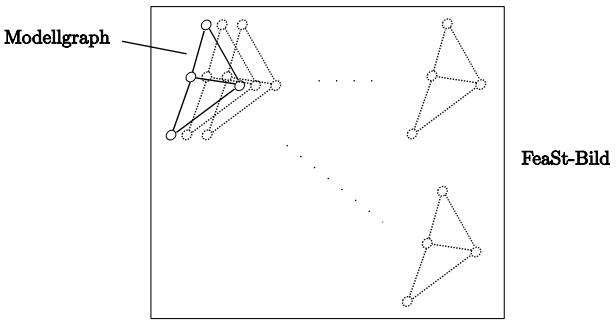


Abb. 65: Scan-Global-Move

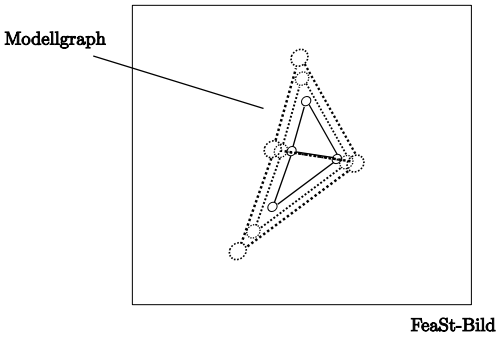


Abb. 66: Scan-Scale-Move

Beim Scan-Global-Move werden die topologischen Kosten bei einer Bewegung nicht beeinflusst, da die an den Kanten hängenden Distanzvektoren unverändert bleiben. Selbiges gilt jedoch nicht für den Scan-Scale-Move, bei welchem der Bündelgraph an der zuvor bestimmten Stelle skaliert wird (Abb. 66). Die Skalierung kann für

x- und y-Richtung unabhängig geschehen. Anschließend besteht die Möglichkeit, mit Hilfe des Scan-Local-Moves eine "Feinjustierung" vorzunehmen. Die Knoten des Bündelgraphen werden hierbei nacheinander entlang eines Rasters um ihre jeweiligen Ursprungspunkte bewegt (Abb.67). Diese Bewegungsart hat ebenfalls Auswirkungen auf die topologischen Kosten in (134), je nach Variation der relativen Distanz eines Knotens zu seinen Nachbarn ändern sich die topologischen Kosten bzw. der topologische Term.
Dieser und alle weiteren Moves können hinsichtlich ihrer Bewegung parametrisiert werden, beispielsweise kann beim Scan-Global-Move das Raster festgelegt werden, entlang welchem der Graph verschoben wird. Nach Anwendung einer Folge dieser Moves, ist das Gesicht im Bild mit Hilfe der Landmarken lokalisiert, nicht mit derselben Präzision wie beim manuellen setzen der Landmarken, jedoch in guter Näherung dazu. Mit Hilfe der bestimmten markanten Punkte wird nun ein Modellgraph des Gesichtes erstellt, hierzu werden nach zuvor beschriebenen Prinzip die Jets an den Landmarken in einem neuen Graphen gespeichert. Das Erstellen dieses Modellgraphen für jeweils ein Gesicht wird auch als Extraktion eines Modellgraphen bezeichnet. Im anschliessenden Gesichtsvergleich werden (132) und (133) verwendet, um den neuen Modellgraphen mit einem anderen zu vergleichen. Auf diesen so genannten Kreuzvergleich zwischen Modellgraphen wird in Kapitel 9 eingegangen.

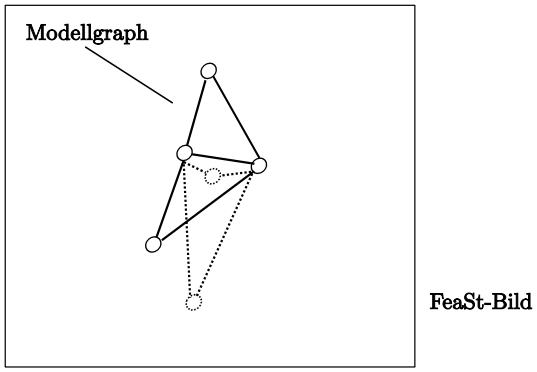


Abb. 67: Scan-Local-Move

9. Redundanz - Analyse

9.1. Transformation eines Bündelgraphen in Matrixform

In diesem Kapitel wird das Konzept vorgestellt, welches verwendet wurde, um Bündelgraphen auf Datenredundanz hinsichtlich der darin enthaltenen Texturdeskriptoren zu untersuchen. Die nachfolgend genauer erläuterte Transformation wurde im Rahmen dieser Arbeit entwickelt und implementiert.
Zunächst wird der zu untersuchende Bündelgraph in eine andere Darstellungsart überführt. Die an den Knoten gespeicherten Daten werden dazu in einer 2 dimensionalen Matrix abgelegt, dieser Vorgang wird im weiteren Verlauf des Dokumentes auch als Transformation und die dazugehörige Matrix als Bündelgraph-Matrix bezeichnet. Die folgenden Bezeichnungen sind aus Kapitel 6 und 8 übernommen, n ist die Knotenanzahl der Modellgraphen und M, L die Anzahl der Skalierungen bzw. Orientierungen der Gabor-Transformation dar. Abb.(68) veranschaulicht das beschriebene Konzept, jede Spalte der Matrix stellt einen der Modellgraphen dar, welcher zur Bildung des Bündelgraphen verwendet wurde, in Abb.(72) wird eine solche Matrix dargestellt. In jeder Spalte werden die Jets der einzelnen Knoten des Graphen abgelegt, somit besitzt jede Spalte $M \cdot L \cdot n$ reelle Einträge.

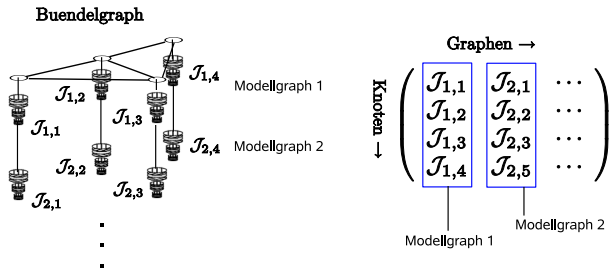


Abb. 68: Repräsentation eines Bündelgraphen durch eine Matrix

Um die Bedeutung der Größen M, L, n sowie den Zusammenhang zwischen Bündelgraph und Bündelgraph-Matrix zu verdeutlichen sei folgendes Beispiel betrachtet. Gegeben ist ein aus $P = 2$ Modellgraphen gebildeter Bündelgraph, dieser hat $n = 4$ Knoten an welchen jeweils Jet-Bündel hängen. Ein Jet-Bündel besteht wie bereits erwähnt aus konkatenierten Jets einzelner Modellgraphen, nehmen wir weiterhin an, dass jeder Jet die Gabor-Filterantworten zu $M = 5$ Skalen und $L = 8$ Orientierungen enthält (d.h. insgesamt 40 komplexe Werte). In Abb.(69) wird dieser Bündelgraph dargestellt, die Jets sind für eine bessere Darstellung entsprechend gekürzt worden, weiterhin wird neben dem Bündelgraph die dazuhörige Bündelgraph-Matrix abgebildet. Im Bündelgraph sind 2 Modellgraphen vorhanden, d.h. die Matrix besitzt 2 Spalten oder mit anderen Worten gesagt repräsentiert eine Spalte der Matrix die Knoteninformationen eines Modellgraphen. Betrachten wir nun die Spalten genauer, jeder Modellgraph besitzt 4 Knoten mit jeweils 40 Gabor-Filterantworten, somit enthält jede Spalte $40 \cdot 4$ Elemente oder anders ausgedrückt besitzt die Matrix 160 Zeilen. Der Zusammenhang zwischen einer Matrix-Zeile und Bündelgraph-Position wird durch die mit einem Kreis umrandeten Elemente verdeutlicht.
Zu Beginn dieser Arbeit wurden erste Untersuchungen auf Bilddatenbanken ausgeführt, für welche keine Phaseninformationen hinsichtlich der Gabor-Filterantworten bereitstanden d.h. es existierten nur reelle Werte innerhalb der Jets. Für die Endresultate hingegen, wurden Datenbanken verwendet in welchen Phaseninformationen enthalten waren. Das zuvor erwähnte Schema der Bündelgraph-Transformation konnte in dieser Form nicht mehr angewandt werden, da es keine komplexen Werte berücksichtigte. Um die Implementierung des Algorithmus nicht verändern zu müssen, wurde die Transformationsregel geringfügig verändert.
Es wird nun unterschieden zwischen Bündelgraphen mit komplexen Werten an ihren Knoten und solchen mit reellen Werten. Für den letzten Fall wird das oben angegebene Schema unverändert weiter verwendet. Besitzt der Bündelgraph jedoch komplexe Werte, so werden die Realteile und Imaginärteile der Werte getrennt in einer Bündelgraph-Matrix gespeichert.
Hierzu sei Abb.(70) betrachtet, die Spaltenzahl der Matrix wurde verdoppelt, wird diese nun vertikal halbiert

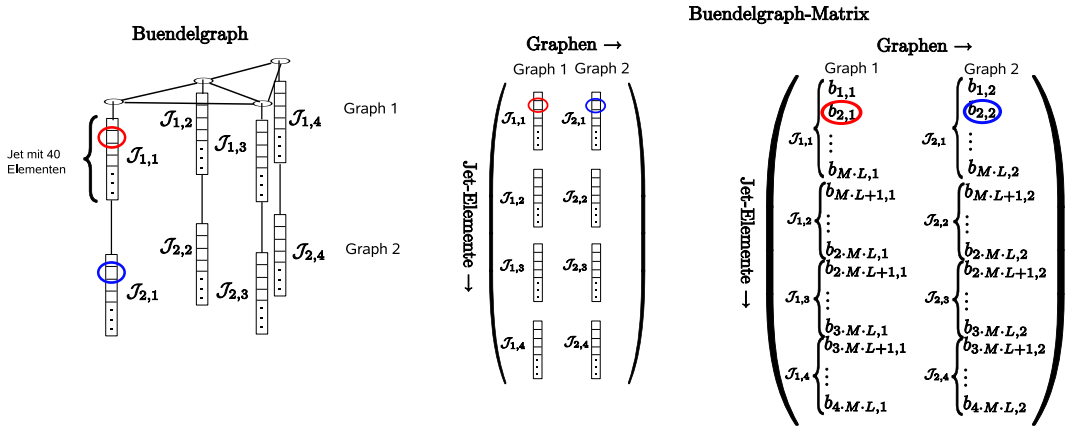


Abb. 69: Zusammenhang zwischen Bündelgraph und dazugehöriger Bündelgraph-Matrix

$$\begin{array}{cc} & \text{Graphen} \rightarrow \\ \text{Knoten} \rightarrow & \begin{pmatrix} \Re\mathcal{J}_{1,1} & \Re\mathcal{J}_{2,1} & \cdots & \Re\mathcal{J}_{P,1} & \Im\mathcal{J}_{1,1} & \Im\mathcal{J}_{2,1} & \cdots & \Im\mathcal{J}_{P,1} \\ \Re\mathcal{J}_{1,2} & \Re\mathcal{J}_{2,2} & \cdots & \Re\mathcal{J}_{P,2} & \Im\mathcal{J}_{1,2} & \Im\mathcal{J}_{2,2} & \cdots & \Im\mathcal{J}_{P,2} \\ \Re\mathcal{J}_{1,3} & \Re\mathcal{J}_{2,3} & \cdots & \Re\mathcal{J}_{P,3} & \Im\mathcal{J}_{1,3} & \Im\mathcal{J}_{2,3} & \cdots & \Im\mathcal{J}_{P,3} \\ \Re\mathcal{J}_{1,4} & \Re\mathcal{J}_{2,4} & \cdots & \Re\mathcal{J}_{P,4} & \Im\mathcal{J}_{1,4} & \Im\mathcal{J}_{2,4} & \cdots & \Im\mathcal{J}_{P,4} \end{pmatrix} \end{array}$$

Abb. 70: Bündelgraph-Matrix für komplexwertige Jets eines Bündelgraphen

betrachtet, so enthält die linke Hälfte die Realteile und die rechte Hälfte die dazugehörigen Imaginärteile. Im weiteren Verlauf dieses Kapitels wird deutlich werden, dass diese Darstellungen des Bündelgraphen für den eigentlichen Kompressionsalgorithmus vollkommen transparent sind. Damit ist gemeint, dass der Algorithmus nicht verändert werden muss um beide Darstellungen verarbeiten zu können.
Diese Darstellung lässt sich problemlos in die ursprüngliche (Bündelgraph-)Form zurück transformieren, dieser Vorgang wird auch als Rekonstruktion eines Bündelgraphen bezeichnet.

9.2. Analyse-Algorithmen

9.2.1. Redundanz-Analyse

Betrachten wir zunächst Abb.(69), es ist zu erkennen, dass die Elemente jeder Matrix-Zeile jeweils einen Wert verschiedener Jets am selben Knoten darstellen. Es wird auch deutlich, dass diese Werte jeweils die selbe Position innerhalb des jeweiligen Jets aufweisen. Zur Veranschaulichung sei ein Beispiel betrachtet; wird jedes Element der 1. Zeile innerhalb der Bündelgraph-Matrix auf 0 gesetzt, so entspricht dies dem 0-setzen des ersten Elementes, des ersten Knotens aller Modellgraphen (siehe Abb.(69)) im Bündelgraph, eine auf diese Art behandelte Matrix ist in Abb.(73) zu sehen. Bei Betrachtung von Abb.(71) wird die Darstellung in Abb.(72) und (73) verdeutlicht.
In den folgenden Abschnitten wird das 0-setzen der Elemente einer Matrix-Zeile auch als Löschen bzw. Entfernen einer Zeile bezeichnet.
Hier wird auch der Vorteil der Bündelgraph-Transformation bei komplexen Werten deutlich, durch die Anordnung der Realteile und Imaginärteile innerhalb einer Zeile, bleibt beim Entfernen einer solchen, die Zuordnung von Real- und Imaginärteil der jeweiligen Gabor-Filterantworten erhalten. Mit anderen Worten gesagt bedeutet dies, dass beim Entfernen eines komplexen Jet-Elements stets der Real- und Imaginärteil entfernt wird. Der Algorithmus musste somit nicht erweitert werden um komplexe Werte zu berücksichtigen, es wird wie zuvor eine Matrix mit reellen Elementen verwendet.

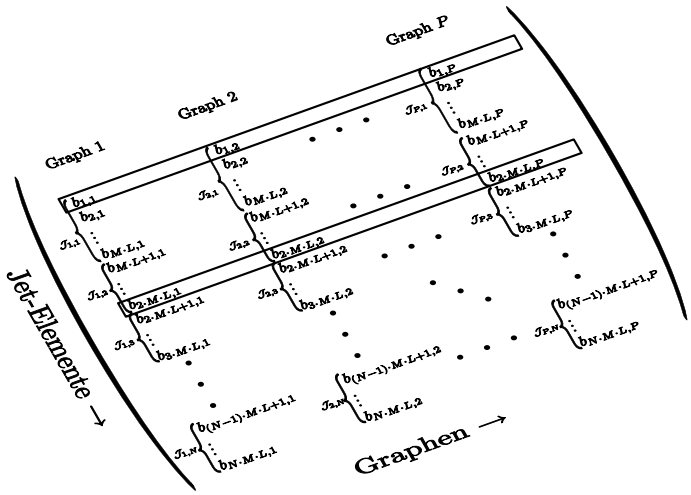


Abb. 71: Bündelgraph-Matrix mit markierten Zeilen

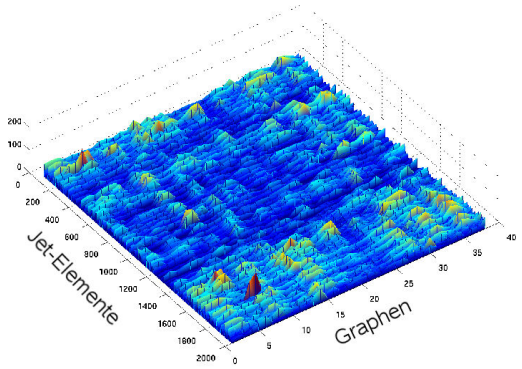


Abb. 72: Bündelgraph-Matrix der Form $1920x38$ ($M = 5, L = 8, n = 48$, 38 Modellgraphen)

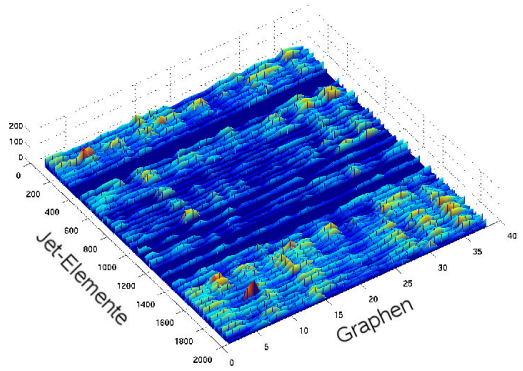


Abb. 73: Ausgedünnte Bündelgraph-Matrix

Das Löschen einer Zeile wird nun zur Redundanz-Analyse genutzt, im weiteren Verlauf wird dieser Vorgang auch als Ausdünnung bzw. verlustbehaftete Komprimierung der Matrix / des Bündelgraphen bezeichnet (hierbei beziehen sich die Begriffe natürlich nur auf die Knoteninformationen des Graphen). Ein vollständig komprimierter Bündelgraph besitzt Jet-Bündel welche ausschließlich den Wert 0 für jedes Jet-Element beinhalten.
Zur Erklärung der nachfolgenden Punkte ist es nötig, den Begriff eines Test-Sets einzuführen. In dieser Arbeit kann ein Test-Set als eine Menge von bekannten bzw. unbekannten Bildern verstanden werden, welche jeweils ein menschliches Gesicht beinhalten. Zu jedem Test-Set können Modellgraphen vorliegen, d.h. für jedes Bild des Sets können bereits (manuell gesetzte) Landmarken existieren.
An dieser Stelle sei der verwendete Analyse-Algorithmus grob betrachtet, auf die einzelnen Punkte wird anschließend näher eingegangen,

1. Transformieren des Bündelgraphen in Matrixform
2. Entfernen einer oder mehrerer Zeilen der Matrix
3. Rekonstruieren des Bündelgraphen

4. Nutzung des Bündelgraphen zum Matching auf einem Test-Set
5. Beurteilung der Qualität des Bündelgraphen
6. solange nicht alle Zeilen gelöscht wurden, weiter mit dem 1.Schritt.

Im 1. Schritt wird die zum Bündelgraph gehörende Bündelgraph-Matrix gebildet. Diese Darstellung wird nun ausgenutzt um Informationen aus dem Bündelgraph zu entfernen. Dazu werden im 2. Schritt einzelne Zeilen der Matrix gelöscht. Die zu löschenden Zeilen werden hierbei nach jeweils einem der folgenden 2 Schemata ausgewählt

- beginnend bei der Zeile welche die kleinste euklidische Norm besitzt, anschließend die Zeile mit der nächstgrößeren Norm
- beginnend bei der Zeile welche die größte euklidische Norm besitzt, anschließend die Zeile mit der nächstkleineren Norm.

Dieser Vorgang stellt die eigentliche verlustbehaftete Komprimierung des Bündelgraphen dar und wird im nächsten Abschnitt genauer erläutert. Nach erfolgreicher Entfernung der Zeilen werden die Informationen der Bündelgraph-Matrix zurück in den Bündelgraph geschrieben. Dieser Bündelgraph wird anschliessend wie in Kapitel 8 beschrieben wurde zur Lokalisierung markanter Punkte verwendet. Im nächsten Schritt wird der Begriff der Qualität eines Bündelgraphen erwähnt, die genauen Kriterien zur Qualitätsbeurteilung eines Bündelgraphen werden später erläutert, an dieser Stelle sei gesagt, dass die Position der lokalisierten Landmarken eine wichtige Rolle hinsichtlich dieser Beurteilung spielt. Sofern noch Informationen im Bündelgraph enthalten sind, d.h. falls noch nicht alle Jet-Sets den Wert 0 innerhalb ihrer Elemente enthalten so wird wieder beim 1. Schritt begonnen. Dieser Algorithmus stellt das zentrale Element der in dieser Arbeit durchgeführten Untersuchung dar. Der Algorithmus sowie das Konzept zur Entfernung der Zeilen, wurden für diese Arbeit entwickelt und im Rahmen dieser implementiert.

9.2.2. Ausdünnen / Komprimieren eines Bündelgraphen

Während der zuvor beschriebenen Redundanz-Analyse wird der Bündelgraph in den Schritten 1-3 komprimiert, in diesem Abschnitt wird der zugehörige Komprimierungs-Algorithmus genauer betrachtet. Wie zuvor beschrieben wurde wird ein Bündelgraph $\mathcal{B}$ komprimiert indem Knoteninformationen entfernt werden.
Die Knoteninformationen werden zunächst in einer Bündelgraph-Matrix M angeordnet, dies wird in Abb.(74) dargestellt. Der 1. Schritt zur Komprimierung ist die Berechnung der euklidischen Normen aller Zeilen m_r , $r = 1, ..., R$ von M, wobei $R = M \cdot L \cdot n$. Nun wird eine so genannte Look-Up Tabelle (LUT) erzeugt, diese Tabelle besitzt 2 Spalten, in der linken wird die Zeilennummer r eingetragen und in der rechten Spalte die dazugehörige Norm $||m_r||$ (siehe Abb.(75)).
Die Einträge (Zeilen) der Tabelle werden nun bezüglich der rechten Spalte, d.h. der Zeilennorm $||m_r||$ sortiert. Diese Sortierung kann aufsteigend oder absteigend erfolgen. Nach erfolgter Sortierung wird die LUT in K gleich große Segmente unterteilt (siehe Abb.(76)), der Wert K stellt die Anzahl der Komprimierungsstufen dar und wird vor Beginn der Redundanz-Analyse festgelegt. Die Look-Up Tabelle wird nur einmal zu Beginn des Analyse-Algorithmus berechnet sowie sortiert und für alle wiederholten Komprimierungen innerhalb des Analyse-Algorithmus weiter verwendet.
Für den anschließenden Schritt betrachten wir das 1. Segment der LUT, es werden nun jene Zeilen m_r aus M gelöscht deren Index r in der linken Spalte des 1. Segments liegt. Die Matrix-Elemente werden anschließend zurück in $\mathcal{B}$ geschrieben, dieser Vorgang stellt die Rekonstruktion des Bündelgraphen dar.
$\mathcal{B}$ ist nun um $\frac{1}{K} \cdot 100\%$ komprimiert worden, diese prozentuale Angabe gibt den Anteil der entfernten Zeilen der Bündelgraph-Matrix wieder. An dieser Stelle soll der Begriff der Kompressionsstufen verdeutlicht werden. Bei einer erneuten Durchführung des beschriebenen Kompressions-Algorithmus wird das nächste Segment der LUT verwendet. Dadurch werden weitere Zeilen aus M gelöscht, die Komprimierung beträgt dann $2 \cdot \frac{1}{K} \cdot 100\%$. Die prozentuale Kompression $\mathcal{K}$ eines Bündelgraphen in Abhängigkeit des Komprimierungsschrittes kann verallgemeinert werden als

$$\mathcal{K}_K(k) := k \cdot \frac{1}{K} \cdot 100\% \quad , \quad k = 0, ..., K \tag{136}$$

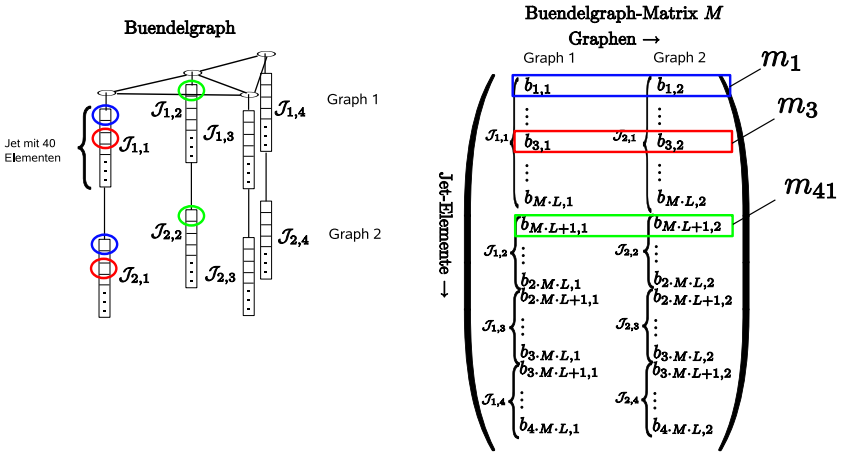


Abb. 74: Bündelgraph und Zeilen der zugehörigen Bündelgraph-Matrix M

r	$\|\|m_r\|\|$
1	$\|\|m_1\|\|$
2	$\|\|m_2\|\|$
⋮	⋮
R	$\|\|m_R\|\|$

Abb. 75: Look-Up Tabelle für Zeilen-Index und Zeilennorm

Hierbei bedeutet $\mathcal{K}_K(0)$, dass keine Komprimierung stattgefunden hat. Im weiteren Verlauf dieser Arbeit wird von der Komprimierung nach "größter Norm" und "kleinster Norm" gesprochen, damit ist die Wahl der Sortierung während der Redundanz-Analyse gemeint. Eine Komprimierung nach größter Norm entspricht einer absteigenden Sortierung der LUT, im angegeben Beispiel bedeutet dies für die Zeilennorm

$$||m_3|| \geq ||m_5|| \geq ... \tag{137}$$

Analog gilt bei einer Komprimierung nach kleinster Norm

$$||m_3|| \leq ||m_5|| \leq ... \tag{138}$$

d.h. es werden zunächst alle Matrix-Zeilen mit kleinster Norm entfernt. Die Art der Sortierung bzw. der Kompression wird vor Beginn der Redundanz-Analyse festgelegt, d.h. es wird während einer Untersuchung nicht zwischen der Komprimierung nach größter und kleinster Norm gewechselt. Die Komprimierung des Bündel-

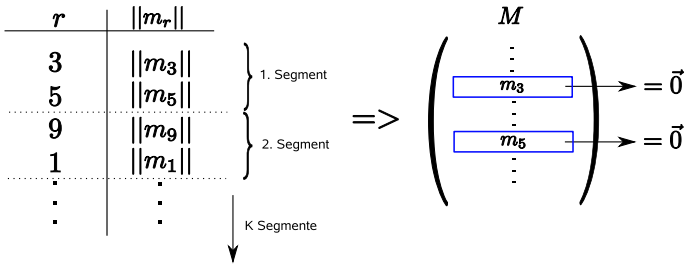


Abb. 76: sortierte Look-Up Tabelle und Zusammenhang zwischen Segment und Matrix-Zeile

graphen hat Einfluß auf die Lokalisierung der Landmarken während des EBGM-Verfahrens. Je weniger Daten

ein Bündelgraph enthält umso geringer wird seine Fähigkeit Landmarken in unbekannten Bildern zu finden. Nicht korrekt bestimmte Landmarken können ein starkes Absinken der Erkennungsrate bewirken, schließlich spielt in solch einem Fall das Korrespondenzproblem eine wichtige Rolle. Es würden beispielsweise keine Gesichtsmerkmale mehr miteinander verglichen sondern u.U. Teile des Bildhintergrundes.

9.3. Beurteilung der Qualität eines Bündelgraphen

In diesem Abschnitt wird die zuvor erwähnte Qualitätsbewertung eines Bündelgraphen genauer beschrieben. Der 5. Schritt des Analyse-Algorithmus nutzt den komprimierten Bündelgraph zum Matchen auf einem bestimmten Test-Set. Die daraus resultierenden Modellgraphen werden nun hinsichtlich zweier Eigenschaften bewertet

- Fehler hinsichtlich der Position der erkannten Landmarken
- Erkennungsrate bei Verwendung der extrahierten Modellgraphen zum Kreuzvergleich,

diese 2 Punkte dienen der Qualitätsbewertung des Bündelgraphen.

9.3.1. Lokalisierungsfehler

Es sei daran erinnert, dass das EBGM-Verfahren der Lokalisierung markanter Punkte dient, mit Hilfe des Bündelgraphen werden hierbei die Positionen dieser im Bild bestimmt und daraus ein neuer Modellgraph gebildet. Sofern für ein Test-Set bereits "optimale" Modellgraphen existieren, d.h. im Sinne von manuell gesetzten markanten Punkten, so kann die Abweichung zwischen optimalen und extrahierten Modellgraphen bestimmt werden.
Das erste Bewertungskriterium ist der so genannte Lokalisierungsfehler, hierbei wird die zuvor erwähnte Abweichung bezüglich der Position berechnet. Zur Bestimmung dieses Fehlers wird der vektorielle Abstand zwischen korrespondierenden Knoten bestimmt, dies wird in Abb.(77) veranschaulicht. Natürlich müssen hierbei die Knoten nicht wie im Bild dargestellt den gleichen Abstand zueinander haben, die Abstände zwischen den jeweiligen Knoten können durchaus verschieden sein.

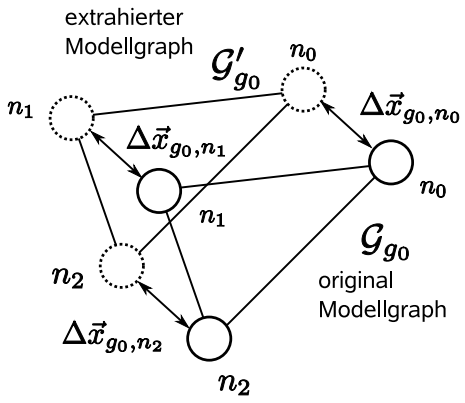


Abb. 77: Lokalisierungsfehler

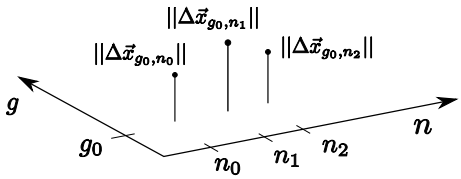


Abb. 78: Visualisierung des Lokalisierungsfehlers

An dieser Stelle soll die Auswertung des Lokalisierungsfehlers beschrieben werden, wie bereits erwähnt, wird dieser für alle Bilder des Test-Sets berechnet. Gehen wir weiterhin davon aus, dass G Elemente im Test-Set enthalten sind, so ergeben sich hinsichtlich der Bestimmung des Lokalisierungsfehlers insgesamt $G \cdot N$ reelle Distanzvektoren

$$\Delta\vec{x}_{g,n} \in \mathbb{R}^2 \ , \ \ g = 1, ..., G \tag{139}$$

schließlich besitzt jeder Modellgraph N Knoten. Da für die quantitative Auswertung nur der Abstand zwischen den Knoten wichtig ist, wird hierzu der Betrag $||\Delta\vec{x}_{g,n}||$ jedes Differenzvektors bestimmt. Für die Anordnung der Werte bietet sich natürlich eine Matrix $\in \mathbb{R}^{G\times N}$ an, in welcher der Lokalisierungsfehler für einen Graph als Spaltenvektor dargestellt wird, Abb.(78) zeigt die prinzipielle Anordnung und Visualisierung dieser Matrixwerte.
Jedoch ist es schwer, den Fehler subjektiv zu beurteilen, schließlich ist es nicht leicht innerhalb der verstreuten Werte die lokale (d.h. auf Knoten bzw. Graphen bezogene) Entwicklung der Positionsabweichung zu erkennen (siehe Abb.(81)). Zur visuellen Aufbereitung in dieser Hinsicht, werden die Werte innerhalb der Matrix sortiert. Hierzu bieten sich wiederum 2 Methoden an, der erste und einfachste Ansatz ist die elementweise Sortierung, hierbei handelt es sich um eine einfache 2-dimensionale Sortierung, diese wird in Abb.(79) veranschaulicht. In Abb.(82) wird eine nach diesem Verfahren sortierte Matrix dargestellt, der Verlauf der Fläche ist zwar glatter und erlaubt einen guten Vergleich der Fehlerentwickung, jedoch ist die Zuordnung zwischen Graphen und dazugehörigen Knoten vollständig verloren gegangen.
Eine bessere Alternative hierzu ist die 2-dimensionale Sortierung bezüglich der Norm, d.h. im Gegensatz zur vorherigen Sortierung wird nun nicht mehr elementweise sondern hinsichtlich der Spalten- und Zeilen-Norm sortiert, siehe Abb.(80). Der Algorithmus für diese Sortierung kann in 4 Schritte unterteilt werden. Im ersten Schritt wird zunächst die euklidische Norm der Spalten berechnet, im anschliessenden 2. Schritt werden die Spalten anhand ihrer Norm aufsteigend sortiert. Der 3. Schritt stellt die Berechnung der Norm jeder Zeile der sortierten Matrix dar. Im letzten Schritt wird die Matrix bezüglich der Zeilennorm absteigend sortiert.

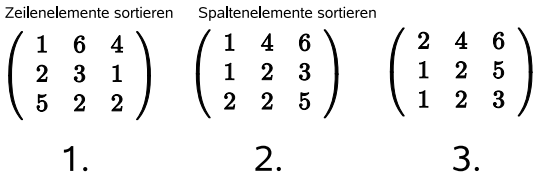


Abb. 79: Matrix elementweise sortiert

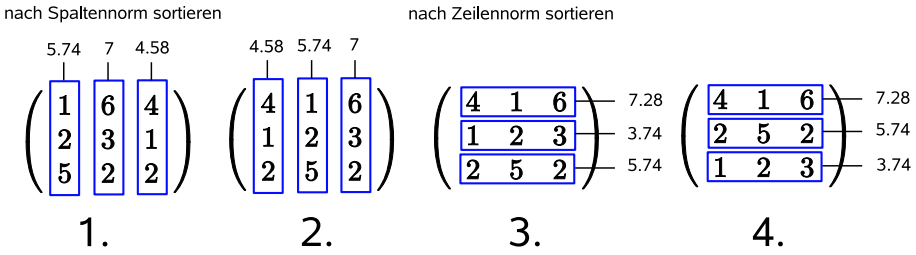


Abb. 80: Matrix sortiert nach Spalten- und Zeilennorm

Hierbei bleibt die Zuordnung von Knoten zu ihren entsprechenden Graphen erhalten, jedoch geht die Reihenfolge der Knoten und Graphen verloren. In Abb.(83) wird eine nach diesem Verfahren sortierte Matrix dargestellt. Es sei zusätzlich angemerkt, dass anhand der Diagramme nur die Beträge der vektoriellen Größen erkennbar sind, d.h. es lässt sich nicht erkennen in welcher Richtung ein Knoten von der idealen Position abweicht.
Für die Visualisierung des Lokalisierungsfehlers werden in Kapitel 10 beide Darstellungen eingesetzt. Das Verfahren zur Untersuchung des Lokalisierungsfehlers sowie dessen Darstellung wurden ebenfalls im Rahmen dieser Arbeit entwickelt und implementiert.

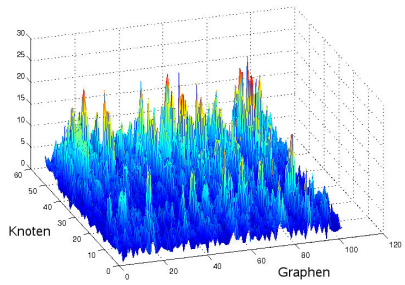


Abb. 81: Lokalisierungsfehler unsortiert

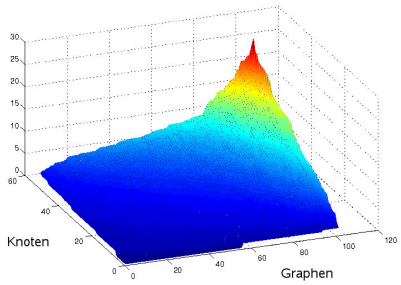


Abb. 82: Lokalisierungsfehler elementweise sortiert

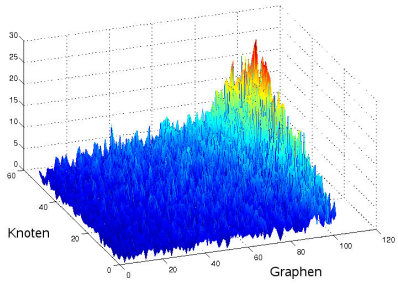


Abb. 83: Lokalisierungsfehler nach Norm sortiert

9.3.2. Kreuzvergleich

Das 2. Kriterium zur Beurteilung der Qualität des Bündelgraphen ist die bereits erwähnte Erkennungsrate, zur Bestimmung dieser wird ein so genannter Kreuzvergleich eingesetzt. Unter diesem Begriff wird die eigentliche Gesichtserkennung verstanden oder anders ausgedrückt, der Versuch ein Gesicht in einer Menge unbekannter Bilder zu finden. Bei einem Kreuzvergleich werden alle Modellgraphen des Test-Sets unter Ausnutzung der Ähnlichkeitsfunktion (134) bzw. (135) mit dem Graphen des gesuchten Gesichtes verglichen (Abb.(84)). Das Bild dessen zugehöriger Graph die genannte Funktion maximiert, ist der beste Kandidat welcher das gesuchte Gesicht enthalten kann. An dieser Stelle soll kurz daran erinnert werden, dass ein extrahierter Modellgraph nicht unbedingt ein Gesicht repräsentieren muss, schließlich können beim EBGM-Verfahren durchaus Positionierungsfehler hinsichtlich der Landmarken auftreten. Die Knoten eines extrahierten Modellgraphen würden in diesem Fall Jets von unbedeutenden Bildpositionen enthalten.
Mit dem Begriff der Erkennungsrate R ist der Quotient

$$R = \frac{\text{Anzahl korrekter Erkennungen}}{\text{Anzahl der möglichen korrekten Erkennungen}} \tag{140}$$

gemeint. In dieser Arbeit wird nun untersucht, wie stark die Erkennungsrate variiert, sofern der eingesetzte Bündelgraph komprimiert wird. In Kapitel 10 wird deshalb die Erkennungsrate in Abhängigkeit von der Ausdünnung des Bündelgraphen dargestellt.

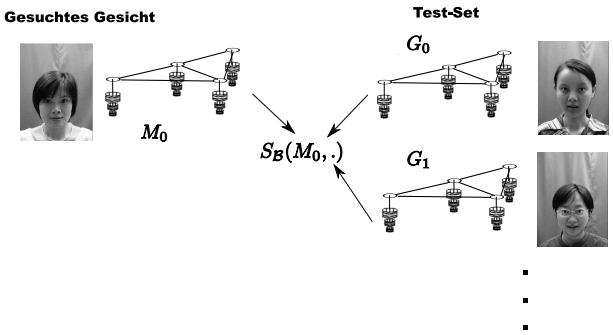


Abb. 84: Kreuzvergleich

9.4. Verwendete Bilddatenbanken und Untersuchungsaufbau

Nachdem die Untersuchungsmethoden illustriert wurden soll nun der Versuchsaufbau betrachtet werden. Kommen wir zunächst auf die eingesetzten Bild-Datenbanken zu sprechen. Es wurden 2 solcher Datenbanken zur Analyse verwendet, hierbei handelt es sich erstens um die CAS-PEAL Datenbank und zweitens um die FaceGen Datenbank.
In der CAS-PEAL Datenbank sind Bilder von Männern und Frauen chinesischer Abstammung gespeichert. Für die Untersuchungen wurden bezüglich jeder verwendeten Person 3 unterschiedliche Bilder benutzt, jeweils eines für geschlossene Augen, lächelnden Gesichtsausdruck und ein Foto auf welchem die Person mit einem neutralem Gesichtsausdruck zu sehen ist.
Die FaceGen Datenbank beinhaltet nur Bilder von künstlich generierten Gesichtern, hier wird zwischen der Position des Gesichtes (im Sinne einer Rotation) und dem Gesichtsausdruck (offene/geschlossene Augen, lachen usw.) unterschieden (siehe Abb.(85)).
Für jedes Bild dieser Datenbanken sind bereits extrahierte bzw. manuell bestimmte Modellgraphen vorhanden. Die Untersuchung hinsichtlich des Positionierungsfehlers wurde jedoch nur mit den FaceGen Bildern durchgeführt, dies liegt darin begründet, dass für jedes künstlich erzeugte Bild die Landmarken natürlich optimal bei der Generierung gesetzt werden können. Solche optimal bestimmten Graphen werden im weiteren Verlauf dieser Arbeit auch als originale Modellgraphen bezeichnet. Die vorhandenen Graphen innerhalb der CAS-PEAL Datenbank wurden mit Hilfe des EBGM-Verfahrens bestimmt und sind somit für die Analyse nur bedingt geeignet. Weiterhin sei erwähnt, dass alle Modellgraphen komplexe Gabor-Filterantworten enthalten, somit kann beim EBGM-Verfahren zusätzlich die Phase mitberücksichtigt werden.

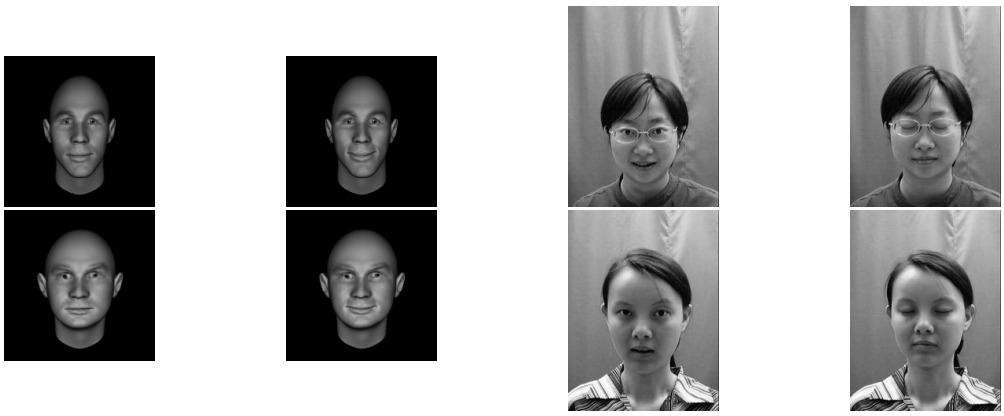


Abb. 85: Beispielbilder aus der FaceGen und CAS-PEAL Datenbank

Kommen wir nun zur Aufstellung der Test-Sets beider Datenbanken, d.h. wir betrachten die zur Untersuchung ausgewählten Bilder und ihre entsprechende Gegenüberstellung.

9.4.1. FaceGen Test-Sets

Zunächst sei die FaceGen Datenbank betrachtet, hierbei wurden Bilder aus den Gruppen "E0P0" und "E1P1" ausgewählt. Die erste Gruppe enthält frontale Bilder von Gesichtern mit neutralem Ausdruck, wohingegen die andere Gruppe leicht rotierte Bilder derselben Gesichter beinhaltet, welche zusätzlich einen lächelnden Gesichtsausdruck besitzen. Für den weiteren Verlauf werden wir die Test-Sets A und B benutzen, wobei Test-Set A 100 Bilder der Gruppe "E0P0" und Test-Set B 100 Bilder der Gruppe "E1P1" umfasst. Der für die Modellgraph-Extraktion eingesetzte Bündelgraph enthält insgesamt 70 Modellgraphen, 35 aus "E0P0" und 35 aus "E1P1", natürlich sind diese Graphen nicht in den Test-Sets A bzw. B enthalten.

9.4.2. CAS-PEAL Test-Sets

Bei der CAS-PEAL Datenbank wurde analog zur FaceGen Datenbank vorgegangen, hierbei wurden jedoch 3 Test-Sets und 2 Bündelgraphen konstruiert. Die CAS-PEAL Datenbank wird in 3 Gruppen unterteilt:

- Gruppe "EN" beinhaltet nur Bilder von Gesichtern mit neutralem Ausdruck
- Gruppe "EC" umfasst Bilder derselben Personen mit geschlossenen Augen
- Gruppe "ES" enthält Bilder der zuvor genannten Personen mit dem Unterschied, dass jede dieser einen lächelnden Gesichtsausdruck besitzt

Auf Basis jeder dieser Gruppen wurde ein Test-Set gebildet, um eine Verwechslung mit den zuvor definierten Test-Sets zu vermeiden, bezeichnen wir hierbei die erste Bildmenge mit Test-Set C, diese enthält 100 Elemente aus der Gruppe "EN". Weiterhin wurden die Test-Sets D und E verwendet, welche ebenfalls 100 Bilder der Gruppen "EC" bzw. "ES" enthalten. Wie zuvor erwähnt wurden 2 unterschiedliche Bündelgraphen benutzt, dies liegt darin begründet, dass 2 unterschiedliche Anordnungen der Test-Sets verwendet wurden. Unter Anordnung sei zu verstehen, dass zum einen, Test-Set C mit D und zum anderen, Test-Set C mit E zur Untersuchung herangezogen wurde. Jeder dieser Bündelgraphen enthält 35 Elemente aus der Gruppe "EN", schließlich ist das Test-Set in beiden zuvor erwähnten Anordnungen enthalten. Neben diesen Modellgraphen sind noch 35 weitere im Bündelgraph enthalten, je nach Anordnung stammen diese aus den Gruppen "EC" bzw. "ES". Die Bündelgraphen welche für beide Datenbanken eingesetzt wurden, besitzen 52 Knoten mit jeweils 40 komplexen Gabor-Filterantworten. Hierbei resultiert die Anzahl der Filterantworten aus einer Gabor-Transformation mit 8 Orientierungen und 5 Skalierungen.

9.4.3. Zusammengestellte Test-Sets

An dieser Stelle sei ein zusammenfassender Überblick über die gewählten Test-Sets und verwendete Bündelgraphen gegeben, eine Identität stellt hierbei eine Person dar.

- **Test-Set A**: 100 Bilder der FaceGen Datenbank welche jeweils ein künstlich generiertes Gesicht mit neutralem Ausdruck darstellen.
  Identitäten: random200_e0_c0_p0_i0_s1.00 bis "random299_e0_c0_p0_i0_s1.00"
- **Test-Set B**: 100 Bilder der FaceGen Datenbank welche die selben Gesichter wie in Test-Set A darstellen, jedoch mit lächelnden Gesichtsausdruck und leichter horizontaler Rotation.
  Identitäten: "random200_e1_c0_p1_i0_s1.00" bis "random299_e1_c0_p1_i0_s1.00"

- **Bündelgraph für Test-Set A<->B**: Identitäten: ”random000_e0_c0_p0_i0_s1.00” bis ”random034_e0_c0_p0_i0_s1.00” und ”random000_e1_c0_p1_i0_s1.00” bis ”random034_e1_c0_p1_i0_s1.00”

- **Test-Set C**: 100 Bilder der CAS-PEAL Datenbank welche Personen chinesischer Abstammung mit neutralem Gesichtsausdruck zeigen
  Identitäten: siehe Anhang

- **Test-Set D**: 100 Bilder der selben Personen wie in Test-Set C, jedoch mit geschlossenen Augen
  Identitäten: siehe Anhang

- **Test-Set E**: 100 Bilder der selben Personen wie in Test-Set C, jedoch mit einem lächelnden Gesichtsausdruck
  Identitäten: siehe Anhang

- **Bündelgraph für Test-Set C<->D**: jeweils 35 Modellgraphen aus der EN- und EC-Gruppe der CAS-PEAL Datenbank.
  Identitäten: siehe Anhang

- **Bündelgraph für Test-Set C<->E**: jeweils 35 Modellgraphen aus der EN- und ES-Gruppe der CAS-PEAL Datenbank.
  Identitäten: siehe Anhang

9.5. Parametrisierung der Moves und topologische Kosten

Die Parametrisierung der Moves welche während des EBGM-Verfahrens eingesetzt wurden ist für beide Bewertungskriterien und Datenbanken identisch, hierbei wurde folgende Aufstellung verwendet:

1. Scan-Global-Move über das gesamte Bild in Intervallen von 4 Pixeln in X- und Y-Richtung (Ähnlichkeitsfunktion (135))

2. Scan-Global-Move im 4×4-Umfeld der zuletzt bestimmten Position in Intervallen von 1 Pixeln in X- und Y-Richtung (Ähnlichkeitsfunktion (135))

3. Scan-Scale-Move mit einer Skalierung von 0.8 bis 1.2 in Intervallen der Breite 0.05 (Ähnlichkeitsfunktion (135))

4. Scan-Global-Move im 4×4-Umfeld der zuletzt bestimmten Position in Intervallen von 1 Pixeln in X- und Y-Richtung (Ähnlichkeitsfunktion (135))

5. Scan-Scale-Move mit einer Skalierung von 0.9 bis 1.1 in Intervallen der Breite 0.05, jedoch nur in X-Richtung (Ähnlichkeitsfunktion (134))

6. Scan-Scale-Move mit einer Skalierung von 0.9 bis 1.1 in Intervallen der Breite 0.05, jedoch nur in Y-Richtung (Ähnlichkeitsfunktion (134))

7. Scan-Local-Move im 2×2-Umfeld der zuletzt bestimmten Position in Intervallen von 1 Pixeln in X- und Y-Richtung (Ähnlichkeitsfunktion (134))

, weiterhin wurde die Signifikanz der topologischen Kosten mit $\lambda = 0.04$ festgelegt.

	Untersuchung der Erkennungsrate	**Untersuchung des Lokalisierungsfehlers**	**Anzahl der Kompressions-schritte**
Komprimierung nach kleinster Norm	A ↔ B	A ↔ B	12
	C ↔ D		12
	C ↔ E		12
Komprimierung nach größter Norm	A ↔ B	A ↔ B	12
	C ↔ D		12
	C ↔ E		12

Tabelle 1: Gegenüberstellung der Test-Sets und dafür durchführte Untersuchungen

9.6. Untersuchungsablauf

Zum Abschluss dieses Kapitels wird der Ablauf der durchgeführten Untersuchung erläutert, die entsprechenden Resultate sind in Kapitel 10 zu finden. Zum besseren Verständnis sei zunächst Tabelle (1) betrachtet, hier sind die entsprechenden Gegenüberstellungen der Test-Sets für die jeweiligen Bewertungskriterien darstellt.
Zur Bewertung der Erkennungsrate werden die Test-Sets A,B,C,D und E herangezogen, hierbei werden jeweils Paare dieser Bildmengen gebildet. Um die Übersicht zu vereinfachen, wird das in der Anordnung links stehende Set als Galerie-Set und das rechte als Probe-Set bezeichnet. In der letzten Spalte der Tabelle (1) ist die jeweilige Anzahl von Kompressionsschritten verzeichnet, in diesem Fall wird der Bündelgraph 12 mal zu jeweils $\frac{100}{12}\%$ ausgedünnt.

9.6.1. Untersuchung der Erkennungsrate

Die Untersuchung der Erkennungsrate wird auf folgende Art durchgeführt:

1. Komprimiere den Bündelgraph (nach der Methode der kleinsten oder größten Norm)
2. Benutze diesen Bündelgraph und extrahiere für jedes Bild aus beiden Sets einen Modellgraphen
3. Führe einen Kreuzvergleich durch, d.h. versuche, für jedes Bild aus dem Probe-Set das entsprechende aus dem Galerie-Set zu finden
4. Bestimme die Erkennungsrate R für diese Kompressionsstufe
5. solange der Bündelgraph nicht vollständig komprimiert ist, gehe zu Schritt 1.)

9.6.2. Untersuchung des Lokalisierungsfehlers

Die Untersuchung hinsichtlich des Lokalisierungsfehlers wird folgenderweise durchgeführt:

1. Komprimiere den Bündelgraph (nach der Methode der kleinsten oder größten Norm)
2. Benutze diesen Bündelgraph und extrahiere für jedes Bild aus beiden Sets einen Modellgraphen
3. Bestimme für jeden dieser Modellgraphen den Lokalisierungsfehler im Bezug auf die jeweiligen originalen Modellgraphen (mit originalen Modellgraphen werden die optimal bestimmten Graphen gemeint)

4. solange der Bündelgraph nicht vollständig komprimiert ist, gehe zu Schritt 1.)

Anhand der Aufstellung in der zuvor genannten Tabelle ist ebenfalls erkennbar, dass für diese Untersuchung nur Sets aus der FaceGen Datenbank verwendet wurden (aus zuvor genannten Gründen).

9.7. Zusammenfassung

Zur Redundanzuntersuchung der Daten innerhalb eines Bündelgraphen werden die Informationen aller Knoten dieses Bündelgraphen in einer Matrix angeordnet. Anschliessend werden schrittweise Zeilen aus dieser Matrix gelöscht, d.h. alle Elemente dieser Zeile werden auf 0 gesetzt. Dies geschieht solange bis keine Zeilen mehr "übrig" sind. Bei jedem der zuvor erwähnten Schritte können durchaus mehrere Zeilen gelöscht werden, die Wahl der zu löschenden Zeilen erfolgt nach der Größe ihrer euklidischen Norm. Weiterhin werden nach jedem Schritt die Informationen aus der Matrix zurück in den Bündelgraph geschrieben. Der Bündelgraph hat nun Elemente an seinen Knoten "verloren", dies wird auch als Ausdünnung bzw. Komprimierung des Graphen bezeichnet. Innerhalb des EBGM-Verfahrens wird der Bündelgraph zum Auffinden markanter Punkte innerhalb eines Bildes verwendet, in dieser Arbeit wird untersucht wie gut ein komprimierter Bündelgraph zur Lokalisierung der genannten Punkte geeignet ist. Die Beurteilung in dieser Hinsicht erfolgt erstens durch Bestimmung der Positionsabweichung der ermittelten Landmarken, und zweitens durch Beurteilung der Erkennungsrate unter Verwendung dieser Landmarken. Das für diese Untersuchung eingesetzte Bildmaterial stammt aus 2 verschiedenen Datenbanken, eine davon beinhaltet Bilder künstlicher Gesichter, wohingegen die andere nur Bilder von Personen chinesischer Abstammung enthält. Aus diesen Datenbanken werden so genannte Test-Sets gebildet, wobei ein Test-Set eine Bildmenge darstellt, in welcher sich nach bestimmten Kriterien ausgewählte Bilder befinden. Für die zuvor beschriebene Untersuchung wurden insgesamt 5 solcher Test-Sets (A,B,C,D,E) gebildet.

10. Resultate

10.1. Einleitung

Im Rahmen dieser Arbeit wurden die in Kapitel 9 beschriebenen Untersuchungen durchgeführt. In diesem Kapitel betrachten wir die Resultate der zuvor beschriebenen Untersuchungsmethoden, bevor wir im nächsten Abschnitt die Ergebnisse bezüglich der FaceGen-Sets (Test-Set A und B) betrachten, soll kurz die Motivation der Untersuchung hervorgehoben und ein Überblick über die kommenden Abschnitte gegeben werden.
Mit Hilfe der erwähnten Untersuchungsverfahren soll geklärt werden ob eine allgemeine Aussage über die Datenredundanz der Knoteninformationen eines Bündelgraphen möglich ist, d.h. ob alle Bündelgraphen eine ähnliche Redundanz aufweisen und ob diese unabhängig von den verwendeten Test-Sets ist. Weiterhin wird untersucht in wie weit die Lokalisierung der Landmarken mit der anschliessenden Objekterkennung / Erkennungsrate zusammenhängt. Eine hohe Datenredundanz könnte ebenfalls dazu genutzt werden das EBGM-Verfahren zu beschleunigen. Die redundanten Daten könnten hierzu aus dem Bündelgraph entfernt werden, aufgrund der geringeren Datenmenge würde der Vergleich einzelner Jets mit Jet-Bündeln beschleunigt werden. Der Ausgangspunkt der Untersuchung ist ein Bündelgraph $\mathcal{B}$ welcher aus P Modellgraphen gebildet wird. Wie zuvor erläutert wurde besitzt ein Bündelgraph insgesamt $P \cdot N$ Jets

$$\mathcal{J}_n^{\mathcal{B}p} \in \mathbb{C}^{M \cdot L} \;\; n = 1, ..., N \;\; , \;\; p = 1, ..., P \tag{141}$$

N stellt hierbei die Knotenanzahl eines der verwendeten Modellgraphen dar und P die Anzahl der verwendeten Modellgraphen. Jeder Jet ist ein komplexwertiger Vektor mit $M \cdot L$ Elementen (Gabor-Filterantworten). Nach dem in Kapitel 9 beschriebenen Verfahren der Ausdünnung wurden nun sukzessiv Filterantworten aus den Jets entfernt, das Entfernen sei als 0-setzen zu verstehen, d.h. die jeweiligen komplexen Zahlen werden durch den Wert 0 ersetzt. Für die Untersuchung wurde eine Anzahl von Komprimierungsschritten (Ausdünnungsstufen) K festgelegt, der Bündelgraph wurde somit K-mal ausgedünnt und zwar jeweils um $\frac{100}{K}\%$. Bei vollständiger Ausdünnung d.h. nach K Komprimierungsschritten ergibt sich

$$\forall \mathcal{J}_n^{\mathcal{B}p} : \mathcal{J}_n^{\mathcal{B}p} = \vec{0} \;\; , \;\; n = 1, ..., N \;\; , \;\; p = 1, ..., P \tag{142}$$

mit anderen Worten bedeutet dies, dass alle an den Knoten hängenden Jets nur 0-Werte beinhalten. Nach jeder durchgeführten Komprimierung wird die Qualität des Bündelgraphen geprüft (siehe Kapitel 9). Sofern der Bündelgraph eine Datenredundanz aufweist, sollte das Entfernen von Knoteninformationen keinen deutlichen Einfluß auf die Qualität des Graphen haben.
An dieser Stelle seien die zuvor erwähnten Größen für die durchgeführte Untersuchung angegeben

$$P = 70, \; N = 52, \; M = 5, \; L = 8, \; K = 12 \tag{143}$$

Diese Werte sind für alle verwendeten Bündelgraphen gleich, d.h. unabhängig vom jeweiligen Test-Set.
In den folgenden Abschnitten werden die Ergebnisse bezogen auf die FaceGen und CAS-PEAL Test-Sets aufgeführt und analysiert, als "Anhang" zu diesem Kapitel sind im letzten Abschnitt die Visualisierungen der Ergebnisse zu sehen.

10.2. FaceGen Test-Sets

In den Abbildungen (89) und (90) werden die Lokalisierungsfehler für jeweils 3 Ausdünnungsstufen dargestellt. Die erste Grafik (89) zeigt die Positionsabweichung der extrahierten Graphen für die Ausdünnung nach der "kleinsten Norm", wohingegen die zweite Grafik (90) diese Abweichung für die Methode der "größten Norm" visualisiert. Hinsichtlich dieser Abbildungen sei erwähnt, dass dort jeweils die Positionierungsfehler für 200 Graphen dargestellt werden. Diese Anzahl von Graphen liegt in der Darstellung der unsortierten Fehlerwerte begründet, es wurde hierbei der Lokalisierungsfehler für die Test-Sets A und B in einem Diagramm dargestellt. Die Graphen wurden hierbei folgenderweise aufgeteilt

- Graphen 0-99 gehören zu Test-Set A

- Graphen 100-199 gehören zu Test-Set B

Bei den sortierten Darstellungen wie sie in Abb.(89) und (90) vorkommen ist diese Reihenfolge jedoch nicht mehr gegeben (aufgrund der Eigenschaften der Sortierverfahren).
Beim Vergleich der einzelnen Diagramme ist zu erkennen, daß die jeweilige Maximalabweichung, d.h. der maximale Wert in dieser Grafik, bei der zuletzt genannten Komprimierungsmethode größer ist. Um eine Aussage über die Fehlerentwicklung treffen zu können, müssen die linken Diagramme in den Abbildungen (91) und (92) betrachtet werden. In diesen wird die durchschnittliche Abweichung eines Modellgraph-Knotens in Pixeln betrachtet, dies in Abhängigkeit von der jeweiligen Ausdünnung. Es ist zu erkennen, dass sich der durchschnittliche Fehler bei der Komprimierung nach kleinster Norm, relativ konstant auf einem geringen Wert hält. Im Vergleich dazu, wächst der Fehler bei einer Ausdünnung nach größter Norm mit steigender Kompressionsstufe. Weiterhin wird in beiden Abbildungen ebenfalls der Lokalisierungsfehler bei vollständiger Ausdünnung dargestellt.

10.2.1. Signifikanz des Komprimierungsverfahrens

Zur weiteren Analyse des Bündelgraphen sollen nun die rechtsseitigen Diagramme in den Grafiken (91) und (92) untersucht werden. Hier wird die zur jeweiligen Ausdünnungsmethode gehörende Erkennungsrate dargestellt. Es ist zu erkennen, dass für eine Ausdünnung nach kleinster Norm, die Erkennungsrate durchgehend bei 100% liegt, während diese bei der Methode nach größter Norm im mittleren Bereich des Graphen deutlich schwankt. Eine mögliche Erklärung hierzu kann durch Betrachtung des durchschnittlichen Lokalisierungsfehlers gegeben werden. Anhand der entsprechenden Graphen ist zu erkennen, dass bei der Kompression nach größter Norm, dieser Fehler rasch ansteigt. Dieser Beobachtung nach spielen Zeilen der Bündelgraph-Matrix mit größter Norm, eine wichtige Rolle bei der korrekten Lokalisierung der markanten Punkte. Diese Vermutung wird auch durch Betrachtung des Fehlers hinsichtlich der kleinsten Norm unterstützt, dieser hält sich hierbei auch bei höherer Kompression auf einem niedrigen Niveau. Dies stimmt mit der These überein, dass die Zeilen der Bündelgraph-Matrix mit größter Norm einen wichtigen Stellungswert bei der Lokalisierung von Landmarken einnehmen. Schließlich werden bei der Ausdünnung nach kleinster Norm, die für dieses Test-Set signifikanten Zeilen (jene mit größter Norm) zunächst erhalten bleiben.

10.2.2. Zusammenhang von Lokalisierungsfehler und Erkennungsrate

Die Festellungen hinsichtlich der Wichtigkeit des Komprimierungsverfahrens sind jedoch keine Erklärung für die angesprochenen Erkennungsraten, schließlich treten innerhalb beider Kompressionsmethoden stellenweise dieselben Fehler auf, wobei sich die zugehörigen Erkennungsraten unterscheiden. Es enstand die Vermutung, dass der Lokalisierungsfehler bei einer Ausdünnung nach größter Norm stellenweise sehr hoch ist, d.h. es wurde vermutet, dass der Lokalisierungsfehler nur für wenige extrahierte Modellgraphen sehr hoch ist. Dieser Ansatz würde den ähnlichen Durchschnittsfehler bei unterschiedlichen Erkennungsraten erklären, schliesslich könnte für 2 vollkommen verschiedene Fehlerverteilungen ein und derselbe Durchschnittsfehler existieren.
Um dies zu untersuchen, wurden die Lokalisierungsfehler für eine Kompression von $\mathcal{K}_{12}(3)$ betrachtet. Der durchschnittliche Fehler ist bei beiden Ausdünnungsmethoden sehr ähnlich, für die Methode der kleinsten Ausdünnung beträgt dieser $\approx 2.3\frac{\text{pixel}}{\text{knoten}}$ und für die Methode der größten Ausdünnung $\approx 2.8\frac{\text{pixel}}{\text{knoten}}$. Die Erkennungsrate beträgt bei der zuerst genannten Methode 100%, wohingegen diese bei der Methode nach größter Norm auf 98% absinkt. Hier sei nochmals darauf hingewiesen, dass der geringfügig höhere Durchschnittsfehler keine Erklärung hierfür ist, schliesslich ergibt sich bei der Ausdünnung nach kleinster Norm und einem durchschnittlichen Lokalisierungsfehler von $\approx 8.2\frac{\text{pixel}}{\text{knoten}}$ immer noch eine 100%ige Erkennungsrate.
Da jedes Test-Set 100 Bilder enthält, folgt aus einer Erkennungsrate von 98%, dass 2 Bilder des Sets nicht korrekt erkannt wurden. Diese 2 Bilder werden in Abb.(86) sowie Abb.(87) darstellt, ebenfalls ist dort der extrahierte Modellgraph zu erkennen. Bei genauem Hinsehen können die Positionierungsfehler des Graphen erkannt werden, jedoch kann bei subjektiver Beurteilung kein deutlicher Unterschied zwischen der Ausdünnung nach kleinster bzw. größter Norm gemacht werden.
Zur weiteren Analyse des Sachverhaltes wurden die unsortierten Darstellungen des Lokalisierungsfehlers für

eine Kompression von $\mathcal{K}_{12}(3)$ herangezogen (siehe Abb.(88)). Unter Nutzung der nicht sortierten Darstellung ist erkennbar welche Graphen u.U. große Lokalisierungsfehler aufweisen, da deren Reihenfolge innerhalb einer unsortierten Darstellung erhalten bleibt. Es ist zu erkennen, dass bei einer Ausdünnung nach kleinster Norm, kein Modellgraph einen im Bezug zu den anderen Graphen, signifikanten Fehler aufweist. Im Gegensatz dazu, ist bei der Methode der größten Norm deutlich zu erkennen, dass es in dieser Hinsicht mehrere sich deutlich abhebender Modellgraphen gibt. Dies bestätigt die zu Beginn erwähnte Vermutung, obwohl die durchschnitt-

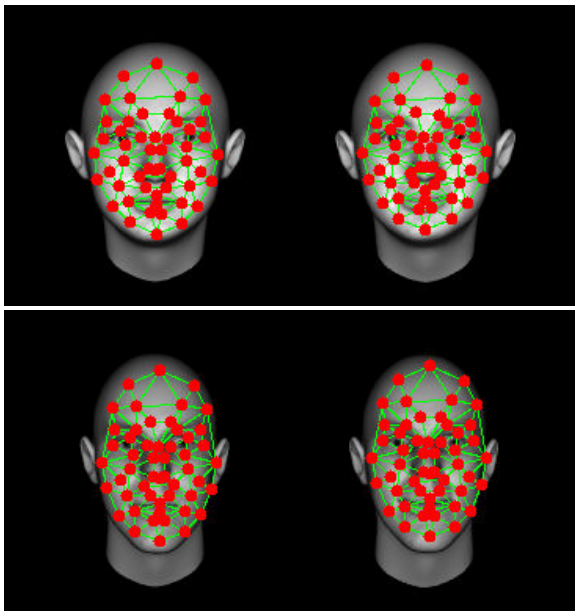


Abb. 86: Matching für Test-Sets A und B bei einer Kompression von $\mathcal{K}_{12}(3)$ nach kleinster Norm. Die linke Hälfte beinhaltet Bilder aus Test-Set A, die rechte Hälfte hingegen stellt die korrespondierenden Bilder aus Test-Set B dar. Die über den Gesichtern positionierten Graphen stellen die extrahierten Modellgraphen für das jeweilige Bild dar.

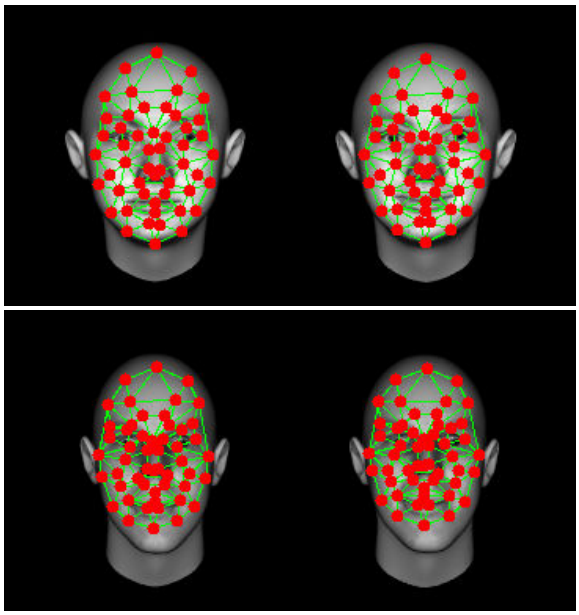


Abb. 87: Matching für Test-Sets A und B bei einer Kompression von $\mathcal{K}_{12}(3)$ nach größter Norm. Die linke Hälfte beinhaltet Bilder aus Test-Set A, die rechte Hälfte hingegen stellt die korrespondierenden Bilder aus Test-Set B dar. Die über den Gesichtern positionierten Graphen stellen die extrahierten Modellgraphen für das jeweilige Bild dar.

lichen Fehler beider Ausdünnungsvefahren für eine Kompression von $\mathcal{K}_{12}(3)$ dicht zusammen liegen, können durchaus "extreme" Fehlpositionierungen einzelner Graphen auftreten.
Es wurde weiterhin festgestellt, dass die Graphen mit signifikantem Lokalisierungsfehler zu den falsch erkannten Bildern gehören. Die Entwicklung des Positionierungsfehlers ist auch in den linken Diagrammen der Abb.(89) und (90) zu erkennen. In zuerst genannter Abbildung hebt sich kein Modellgraph besonders hervor, d.h. es kann gesagt werden, dass alle Graphen das Gesicht im Bild "gleich gut" lokalisieren (damit ist die Position der gefundenen Landmarken gemeint). Werden diese extrahierten Modellgraphen zum Kreuzvergleich eingesetzt, so kann durchaus eine gute Erkennungsrate resultieren, schliesslich werden weiterhin Graphen miteinander verglichen, welche dieselben Merkmale eines Gesichts repräsentieren.
Dies ist bei der Kompression nach größter Norm nicht unbedingt der Fall, da ein Modellgraph mit signifikantem Lokalisierungsfehler, nicht zwangsläufig dieselben Bereiche eines Gesichts repräsentieren muss. In Abb.(90) ist erkennbar, dass mindestens 1 Modellgraph einen im Vergleich mit den restlichen Modellgraphen signifikanten Lokalisierungsfehler aufweist. Mit anderen Worten ausgedrückt bedeutet dies, dass Modellgraphen mit deutlicher Abweichung von durchschnittlichen Lokalisierungsfehler, nicht dieselben Merkmale eines Gesichts

repräsentieren müssen wie die übrigen Graphen.

10.2.3. Vollständige Ausdünnung

Bei einer 100%igen Komprimierung des Bündelgraphen beträgt die Erkennungsrate hinsichtlich beider Ausdünnungsmethoden 100%. Dies liegt zum einen an der Eigenschaft der verwendeten Bilder, denn innerhalb dieser liegen die Gesichter zentriert, zum anderen wurde bei den Untersuchungen festgestellt, dass bei vollständiger Ausdünnung keine individuelle Verformung des Bündelgraphen mehr stattfindet. D.h. die bestimmten Landmarken sind bei allen Bildern der Test-Sets identisch, natürlich bedeutet dies nicht, dass daraus eine Erkennungsrate von 100% folgt, jedoch ist diese für beide Verfahren unter 100%iger Ausdünnung und den verwendeten Test-Sets (A und B) stets identisch. Die Positionierung des Bündelgraphen, ist in solch einem Fall nur von seinen Initialwerten abhängig, d.h. die Ausrichtung auf dem Bild erfolgt anhand der ursprünglichen Knotenpositionen. Ein weiterer wichtiger Punkt ist der Hintergrund jedes FaceGen-Bildes, dieser ist vollständig schwarz, somit ist es wesentlich einfacher, auch mit wenigen Informationen im Bündelgraph, das Gesicht mit Hilfe eines Scan-Global-Moves zu finden.

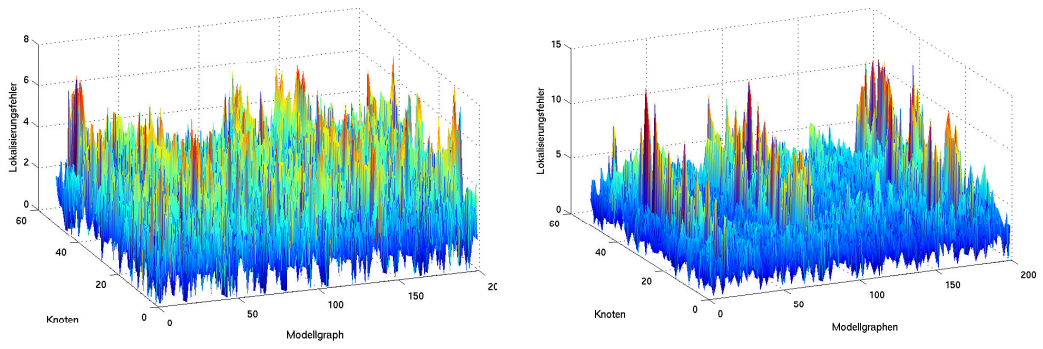


Abb. 88: unsortierte Darstellung des Lokalisierungsfehlers für die Test-Sets A und B bei einer Kompression von $\mathcal{K}_{12}(3)$, links für die Methode der kleinsten Norm und rechts für die der größten Norm

Bevor die Ergebnisse bezüglich der CAS-PEAL Sets betrachtet werden, soll noch eine kurze Zusammenfassung der bisherigen Feststellungen gegeben werden

- hinsichtlich beider Ausdünnungsmethoden wird eine gute Erkennungsrate erreicht, jedoch liegt diese in der Art der verwendeten Bilder begründet (schwarzer Hintergrund und zentrierte Gesichter)
- Zeilen mit größter Norm innerhalb der Bündelgraph-Matrix nehmen einen wichtigen Stellenwert bei der Bestimmung der Landmarken ein
- die "Verformung" des Bündelgraphen ist bei vollständiger Komprimierung nicht mehr vom zugrunde liegenden Bild abhängig

10.3. CAS-PEAL Test-Sets

Betrachten wir nun die Untersuchungsergebnisse unter Verwendung der CAS-PEAL Test-Sets C,D und E. In Abb.(93) wird die Erkennungsrate bei Benutzung der Sets C und D dargestellt, wobei der linke Graph diese für die Methode der kleinsten Norm darstellt, analog dazu bezieht sich der rechte Graph auf die Methode der größten Norm.
Im Vergleich mit den Ergebnissen und Feststellungen hinsichtlich der FaceGen Test-Sets, wird hier deutlich, dass die Erkennungsrate bei steigender Ausdünnung rapide absinkt. Ein weiterer wesentlicher Unterschied ist

die Tatsache, dass hierbei unter Verwendung der Methode der größten Norm eine bis zum 8. Ausdünnungsschritt bessere Erkennungsrate resultiert.
Ähnliches gilt für die Test-Sets C und E, auch hierbei zeichnet sich, bei Betrachtung von Abb.(94), dieselbe Entwicklung hinsichtlich der fallenden Erkennungsrate ab. Ein nennnenswerter Unterschied zum Vergleich der Sets C und D ist die Tatsache, dass die Ausdünnung nach größter Norm eine schlechtere Erkennungsrate bewirkt als die Ausdünnung nach kleinster Norm.
Analog zur Analyse der FaceGen Sets, ist die Erkennungsrate bei 100%iger Ausdünnung in beiden Fällen der Kompression identisch, jedoch beträgt diese nicht mehr 100%. Dies liegt u.A. darin begründet, dass die Gesichter in den verwendeten Bildern nicht mehr zentriert sind und der Bündelgraph wie bereits erwähnt, keine auf die Bilder bezogene individuelle Verformung durchführt. Ein weiterer Grund in diesem Zusammenhang ist der Umstand, dass die Bilder der CAS-PEAL Datenbank teilweise verschiedene Hintergründe aufweisen.

10.4. Visualisierte Ergebnisse

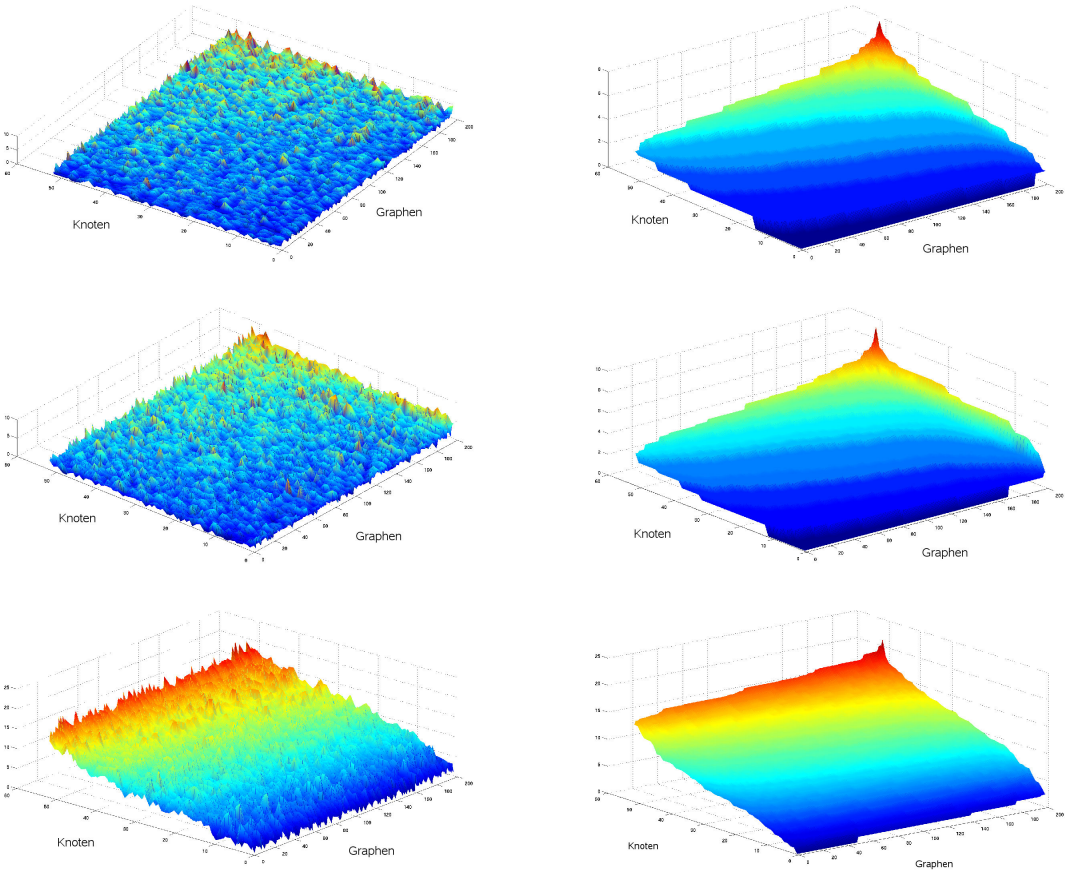


Abb. 89: Lokalisierungsfehler unter Nutzung der FaceGen Datenbank, die dargestellten Graphen gelten für die Kompression von $\mathcal{K}_{12}(2)$, $\mathcal{K}_{12}(7)$ und $\mathcal{K}_{12}(12)$ (kleinste Norm)

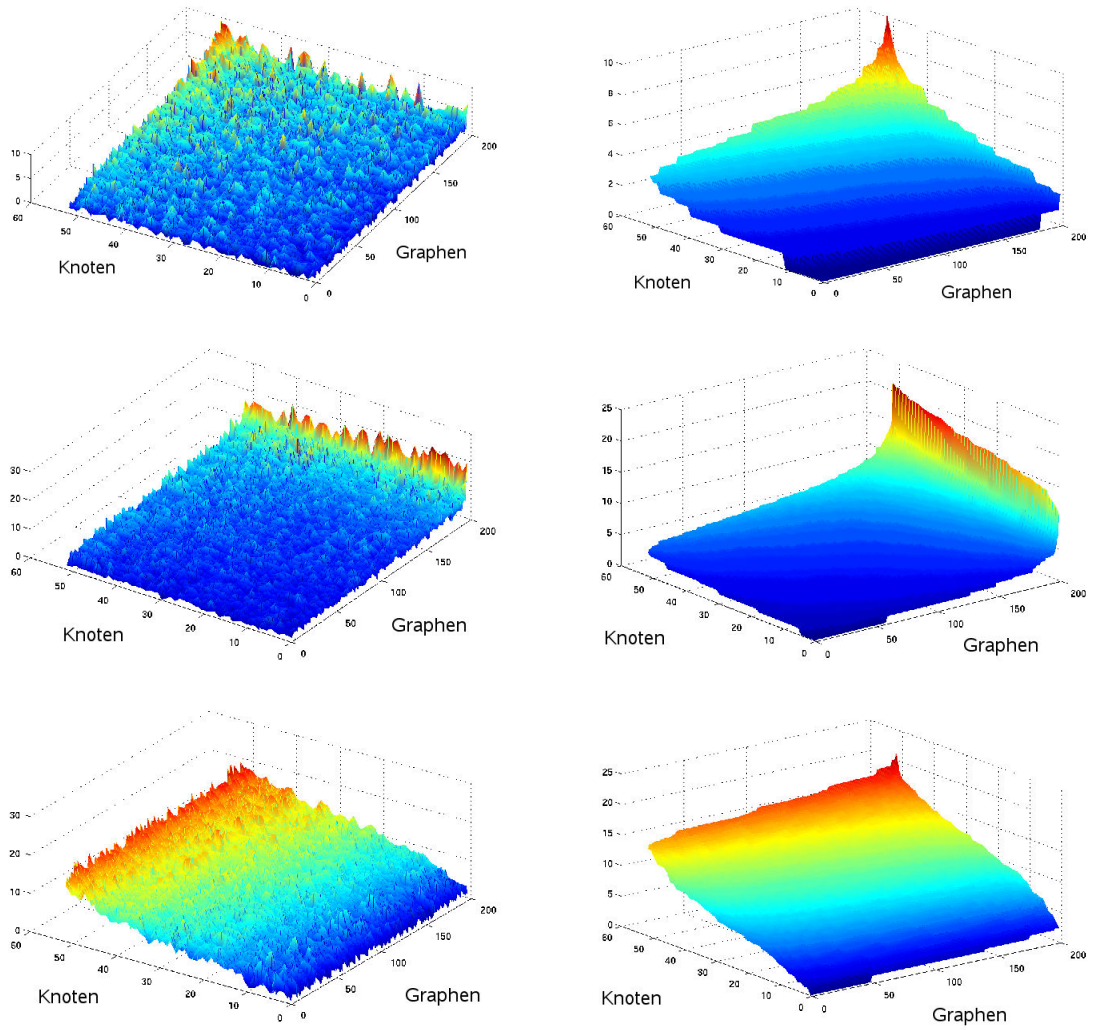


Abb. 90: Lokalisierungsfehler unter Nutzung der FaceGen Datenbank, die dargestellten Graphen gelten für die Ausdünnung von $\mathcal{K}_{12}(2)$, $\mathcal{K}_{12}(7)$ und $\mathcal{K}_{12}(12)$ (größte Norm)

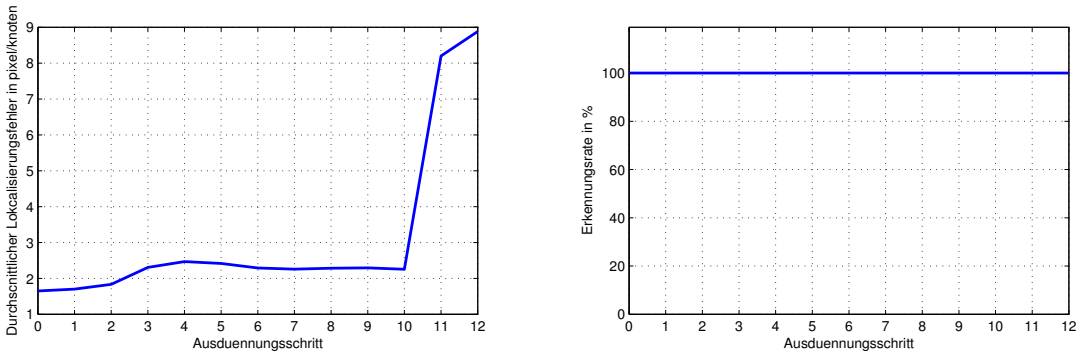


Abb. 91: Gegenüberstellung von durchschnittlichen Lokalisierungsfehler und Erkennungsrate unter Nutzung der FaceGen Datenbank (kleinste Norm)

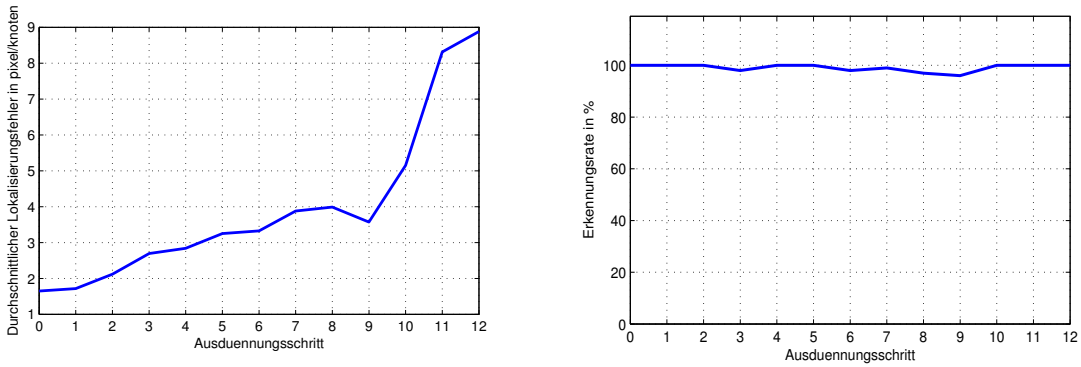


Abb. 92: Gegenüberstellung von durchschnittlichen Lokalisierungsfehler und Erkennungsrate unter Nutzung der FaceGen Datenbank (größte Norm)

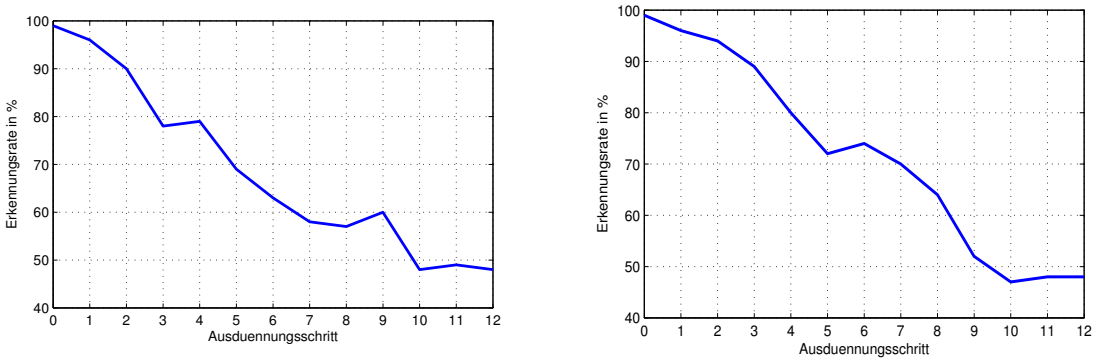


Abb. 93: Erkennungsrate für Anordnung $C \leftrightarrow D$, links nach kleinster Norm ausgedünnt, rechts nach größter Norm

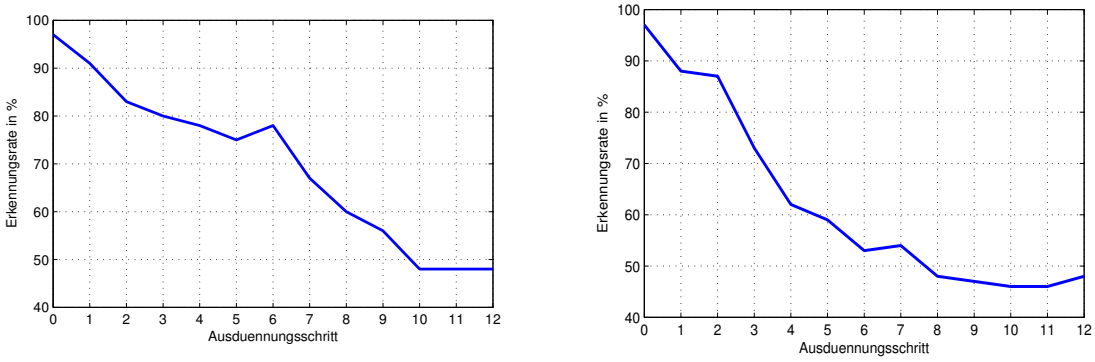


Abb. 94: Erkennungsrate für Anordnung $C \leftrightarrow E$, links nach kleinster Norm ausgedünnt, rechts nach größter Norm

11. Diskussion

Das Ziel dieser Arbeit bestand darin, Bündelgraphen hinsichtlich der Datenredundanz ihrer Knoteninformationen zu untersuchen, d.h. es wurde der Frage nachgegangen, wieviele Texturinformationen nötig sind, um weiterhin eine gute Lokalisierung der Landmarken in einem Bild zu ermöglichen. Anhand der Untersuchungsresultate ist jedoch erkennbar, dass keine allgemeine Aussage über die Datenredundanz innerhalb eines Bündelgraphen getroffen werden kann.
Bei der Untersuchung wurde ein Algorithmus zur verlustbehafteten Kompression der Knoteninformationen eines Bündelgraphen verwendet. Wie bei allen verlustbehafteten Kompressionsverfahren stellt sich auch hier die Frage, wann der Datenverlust noch akzeptabel ist. In dieser Arbeit wurden 2 verschiedene Ansätze zur Kompression verwendet, zunächst wurden die Knoteninformationen des Bündelgraphen in einer Matrix dargestellt, anschliessend wurden aus dieser Werte "gelöscht" (hierbei sei unter dem Begriff des löschens, das "0-setzen" des entsprechenden Elementes verstanden). Die zuvor erwähnten Ansätze beziehen sich auf die Wahl der zu löschenden Matrix-Elemente, zum einen wurden nur Zeilen der Matrix gelöscht, deren Zeilenvektor die geringste Norm aufwies, zum anderen wurden nur Zeilen gelöscht, deren Zeilenvektor die größte Norm aufwies. Nach Durchführung dieser Prozedur wurden die Werte der Matrix wieder in den Bündelgraphen geschrieben, welcher daraufhin zum Matching eingesetzt wurde. Mit Hilfe dieses Vorgehens wurde versucht, nicht signifikante bzw. redundante Daten aus der Informationsmenge zu entfernen. Bei Verwendung der künstlich generierten Bilder aus der FaceGen Datenbank hatte sich gezeigt, dass bei einer Komprimierung nach der Methode der kleinsten als auch der größten Norm, durchaus noch akzeptable Ergebnisse zu erwarten sind. Jedoch ergaben sich bei Verwendung natürlicher Bilder rapide fallende Erkennungsraten. Dies liegt darin begründet, dass die idealisierten Bedingungen wie z.b. ein schwarzer Hintergrund bei natürlichen Bildern nicht unbedingt gegeben sind. Im Bezug auf die FaceGen Datenbank kann von einer hohen Datenredundanz gesprochen werden, da selbst bei starker Kompression des Bündelgraphen, immer noch akzeptable Ergebnisse erzielt werden. Mit anderen Worten gesagt bedeutet dies, dass viele Knoteninformationen nicht "benötigt" werden um letztendlich eine gute Erkennungsrate zu ermöglichen.
Ebenfalls kann aus den Ergebnissen abgelesen werden, dass Zeilen der Bündelgraphmatrix mit größter Norm, eine wichtige Rolle bei der Lokalisierung von Landmarken spielen. Der durchschnittliche Lokalisierungsfehler wuchs beim Entfernen der Zeilen mit größter Norm wesentlich schneller als beim löschen jener mit kleinster Norm. Während der Untersuchung des Lokalisierungsfehlers wurde auch dessen Beziehung zur Erkennungsrate deutlich. Die Ausprägung des durchschnittlichen Lokalisierungsfehlers impliziert keine bestimmte Erkennungsrate, es ist durchaus möglich, dass beide Kompressionsmethoden für einen Ausdünnungsschritt denselben durchschnittlichen Lokalisierungsfehler, jedoch unterschiedliche Erkennungsraten aufweisen.
Die CAS-PEAL Datenbank betreffend, können die zuvor erwähnten Feststellungen nicht erneut getroffen werden. Im Vergleich zu den FaceGen Untersuchungsergebnissen, nimmt die Erkennungsrate bei beiden Kompressionsmethoden schon bei geringer Ausdünnung stark ab. Im Gegensatz zur vorherigen Untersuchung, wurden im Bezug auf die CAS-PEAL Datenbank insgesamt 3 Test-Sets verwendet. Bei der Gegenüberstellung der Sets C und E ist die zuvor festgestellte Signifikanz der Zeilen mit größter Norm, weiterhin erkennbar, jedoch zeichnet sich bei Verwendung der Sets C und D das genaue Gegenteil ab, in diesem Fall tragen die Zeilen mit kleinster Norm mehr Relevanz für eine "gute" Erkennungsrate. Hieraus wird deutlich, dass auch die Wahl der Test-Sets innerhalb einer Datenbank eine wichtige Rolle zur Untersuchung der Redundanz spielt.
Weiterhin stellt die Wahl, Reihenfolge und Parametrisierung der Moves einen wichtigen Punkt dar, beispielsweise könnten mit einem sehr detailierten Scan-Global-Move, d.h. mit einer geringen Schrittweite, die Landmarken u.U. besser lokalisiert werden. Analog dazu nimmt die Wahl der Ausdünnungsmethode einen anderen wichtigen Stellenwert ein, schliesslich ist anhand der Ergebnisse der FaceGen Sets zu erkennen, dass die Ausdünnung nach kleinster Norm bessere Ergebnisse produziert. Möglicherweise existieren bessere Konzepte zum Entfernen der Knoteninformationen eines Bündelgraphen, d.h. Verfahren welche auf die jeweiligen Test-Sets angepasst sind.
In dieser Hinsicht sei zusammenfassend gesagt, dass die Analyse der Datenredundanz innerhalb eines Bündelgraphen, immer von den ausgewählten Test-Sets, Matching-Parametern und der Komprimierungsmethode abhängig ist. Die Daten des Graphen können durchaus eine Redundanz aufweisen, jedoch ist eine allgemeine Aussage in dieser Hinsicht nicht möglich.

A. Test-Sets

Test-Set C
MY_000102_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R0_S0
MY_000103_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R1_S0
MY_000104_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R0_S0
MY_000105_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R1_S0
MY_000114_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R0_S0
MY_000123_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R1_S0
MY_000124_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R1_S0
MY_000125_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R1_S0
MY_000126_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R1_S0
MY_000133_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R1_S0
MY_000137_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R1_S0
MY_000138_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R0_S0
MY_000140_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R0_S0
MY_000149_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R1_S0
MY_000153_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R0_S0
MY_000154_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R0_S0
MY_000155_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R0_S0
MY_000159_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R0_S0
MY_000160_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R0_S0
MY_000166_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R0_S0
MY_000168_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R1_S0
MY_000169_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R1_S0
MY_000170_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R1_S0
MY_000171_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R1_S0
MY_000177_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R0_S0
MY_000178_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R1_S0
MY_000180_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R0_S0
MY_000183_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R0_S0
MY_000185_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R0_S0
MY_000190_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R0_S0
MY_000191_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R0_S0
MY_000195_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R0_S0
MY_000196_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R0_S0
MY_000197_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R0_S0
MY_000198_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R0_S0
MY_000203_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R0_S0
MY_000204_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R0_S0
MY_000207_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R0_S0
MY_000227_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R1_S0
MY_000558_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R0_S0
MY_000616_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R1_S0
MY_000617_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R1_S0
MY_000623_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R0_S0
MY_000627_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R1_S0
MY_000629_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R1_S0
MY_000630_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R0_S0

MY_000633_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R1_S0
MY_000634_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R0_S0
MY_000635_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R0_S0
MY_000636_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R0_S0
MY_000641_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R0_S0
MY_000651_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R1_S0
MY_000652_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R0_S0
MY_000653_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R0_S0
MY_000654_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R0_S0
MY_000655_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R0_S0
MY_000657_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R1_S0
MY_000658_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R1_S0
MY_000660_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R1_S0
MY_000661_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R1_S0
MY_000663_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R0_S0
MY_000667_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R1_S0
MY_000668_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R1_S0
MY_000669_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R1_S0
MY_000670_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R1_S0
MY_000671_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R0_S0
MY_000673_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R0_S0
MY_000679_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R1_S0
MY_000680_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R1_S0
MY_000681_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R0_S0
MY_000682_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R1_S0
MY_000683_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R1_S0
MY_000684_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R1_S0
MY_000685_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R1_S0
MY_000686_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R1_S0
MY_000689_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R0_S0
MY_000690_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R0_S0
MY_000691_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R0_S0
MY_000693_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R0_S0
MY_000694_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R1_S0
MY_000695_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R1_S0
MY_000696_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R1_S0
MY_000697_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R0_S0
MY_000698_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R1_S0
MY_000699_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R1_S0
MY_000700_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R1_S0
MY_000701_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R1_S0
MY_000702_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R1_S0
MY_000703_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R1_S0
MY_000705_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R0_S0
MY_000710_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R0_S0
MY_000715_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R1_S0
MY_000716_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R1_S0
MY_000717_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R1_S0
MY_000718_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R1_S0

MY_000719_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R0_S0
MY_000720_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R0_S0
MY_000721_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R0_S0
MY_000722_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R0_S0
MY_000723_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R1_S0

Tabelle 2: Test-Sets C

Test-Set D
MY_000102_IEU+00_PM+00_EC_A0_D0_T0_BB_M0_R0_S0
MY_000103_IEU+00_PM+00_EC_A0_D0_T0_BB_M0_R1_S0
MY_000104_IEU+00_PM+00_EC_A0_D0_T0_BB_M0_R0_S0
MY_000105_IEU+00_PM+00_EC_A0_D0_T0_BB_M0_R1_S0
MY_000114_IEU+00_PM+00_EC_A0_D0_T0_BB_M0_R0_S0
MY_000123_IEU+00_PM+00_EC_A0_D0_T0_BB_M0_R1_S0
MY_000124_IEU+00_PM+00_EC_A0_D0_T0_BB_M0_R1_S0
MY_000125_IEU+00_PM+00_EC_A0_D0_T0_BB_M0_R1_S0
MY_000126_IEU+00_PM+00_EC_A0_D0_T0_BB_M0_R1_S0
MY_000133_IEU+00_PM+00_EC_A0_D0_T0_BB_M0_R1_S0
MY_000137_IEU+00_PM+00_EC_A0_D0_T0_BB_M0_R1_S0
MY_000138_IEU+00_PM+00_EC_A0_D0_T0_BB_M0_R0_S0
MY_000140_IEU+00_PM+00_EC_A0_D0_T0_BB_M0_R0_S0
MY_000149_IEU+00_PM+00_EC_A0_D0_T0_BB_M0_R1_S0
MY_000153_IEU+00_PM+00_EC_A0_D0_T0_BB_M0_R0_S0
MY_000154_IEU+00_PM+00_EC_A0_D0_T0_BB_M0_R0_S0
MY_000155_IEU+00_PM+00_EC_A0_D0_T0_BB_M0_R0_S0
MY_000159_IEU+00_PM+00_EC_A0_D0_T0_BB_M0_R0_S0
MY_000160_IEU+00_PM+00_EC_A0_D0_T0_BB_M0_R0_S0
MY_000166_IEU+00_PM+00_EC_A0_D0_T0_BB_M0_R0_S0
MY_000168_IEU+00_PM+00_EC_A0_D0_T0_BB_M0_R1_S0
MY_000169_IEU+00_PM+00_EC_A0_D0_T0_BB_M0_R1_S0
MY_000170_IEU+00_PM+00_EC_A0_D0_T0_BB_M0_R1_S0
MY_000171_IEU+00_PM+00_EC_A0_D0_T0_BB_M0_R1_S0
MY_000177_IEU+00_PM+00_EC_A0_D0_T0_BB_M0_R0_S0
MY_000178_IEU+00_PM+00_EC_A0_D0_T0_BB_M0_R1_S0
MY_000180_IEU+00_PM+00_EC_A0_D0_T0_BB_M0_R0_S0
MY_000183_IEU+00_PM+00_EC_A0_D0_T0_BB_M0_R0_S0
MY_000185_IEU+00_PM+00_EC_A0_D0_T0_BB_M0_R0_S0
MY_000190_IEU+00_PM+00_EC_A0_D0_T0_BB_M0_R0_S0
MY_000191_IEU+00_PM+00_EC_A0_D0_T0_BB_M0_R0_S0
MY_000195_IEU+00_PM+00_EC_A0_D0_T0_BB_M0_R0_S0
MY_000196_IEU+00_PM+00_EC_A0_D0_T0_BB_M0_R0_S0
MY_000197_IEU+00_PM+00_EC_A0_D0_T0_BB_M0_R0_S0
MY_000198_IEU+00_PM+00_EC_A0_D0_T0_BB_M0_R0_S0
MY_000203_IEU+00_PM+00_EC_A0_D0_T0_BB_M0_R0_S0
MY_000204_IEU+00_PM+00_EC_A0_D0_T0_BB_M0_R0_S0
MY_000207_IEU+00_PM+00_EC_A0_D0_T0_BB_M0_R0_S0
MY_000227_IEU+00_PM+00_EC_A0_D0_T0_BB_M0_R1_S0

MY_000558_IEU+00_PM+00_EC_A0_D0_T0_BB_M0_R0_S0
MY_000616_IEU+00_PM+00_EC_A0_D0_T0_BB_M0_R1_S0
MY_000617_IEU+00_PM+00_EC_A0_D0_T0_BB_M0_R1_S0
MY_000623_IEU+00_PM+00_EC_A0_D0_T0_BB_M0_R0_S0
MY_000627_IEU+00_PM+00_EC_A0_D0_T0_BB_M0_R1_S0
MY_000629_IEU+00_PM+00_EC_A0_D0_T0_BB_M0_R1_S0
MY_000630_IEU+00_PM+00_EC_A0_D0_T0_BB_M0_R0_S0
MY_000633_IEU+00_PM+00_EC_A0_D0_T0_BB_M0_R1_S0
MY_000634_IEU+00_PM+00_EC_A0_D0_T0_BB_M0_R0_S0
MY_000635_IEU+00_PM+00_EC_A0_D0_T0_BB_M0_R0_S0
MY_000636_IEU+00_PM+00_EC_A0_D0_T0_BB_M0_R0_S0
MY_000641_IEU+00_PM+00_EC_A0_D0_T0_BB_M0_R0_S0
MY_000651_IEU+00_PM+00_EC_A0_D0_T0_BB_M0_R1_S0
MY_000652_IEU+00_PM+00_EC_A0_D0_T0_BB_M0_R0_S0
MY_000653_IEU+00_PM+00_EC_A0_D0_T0_BB_M0_R0_S0
MY_000654_IEU+00_PM+00_EC_A0_D0_T0_BB_M0_R0_S0
MY_000655_IEU+00_PM+00_EC_A0_D0_T0_BB_M0_R0_S0
MY_000657_IEU+00_PM+00_EC_A0_D0_T0_BB_M0_R1_S0
MY_000658_IEU+00_PM+00_EC_A0_D0_T0_BB_M0_R1_S0
MY_000660_IEU+00_PM+00_EC_A0_D0_T0_BB_M0_R1_S0
MY_000661_IEU+00_PM+00_EC_A0_D0_T0_BB_M0_R1_S0
MY_000663_IEU+00_PM+00_EC_A0_D0_T0_BB_M0_R0_S0
MY_000667_IEU+00_PM+00_EC_A0_D0_T0_BB_M0_R1_S0
MY_000668_IEU+00_PM+00_EC_A0_D0_T0_BB_M0_R1_S0
MY_000669_IEU+00_PM+00_EC_A0_D0_T0_BB_M0_R1_S0
MY_000670_IEU+00_PM+00_EC_A0_D0_T0_BB_M0_R1_S0
MY_000671_IEU+00_PM+00_EC_A0_D0_T0_BB_M0_R0_S0
MY_000673_IEU+00_PM+00_EC_A0_D0_T0_BB_M0_R0_S0
MY_000679_IEU+00_PM+00_EC_A0_D0_T0_BB_M0_R1_S0
MY_000680_IEU+00_PM+00_EC_A0_D0_T0_BB_M0_R1_S0
MY_000681_IEU+00_PM+00_EC_A0_D0_T0_BB_M0_R0_S0
MY_000682_IEU+00_PM+00_EC_A0_D0_T0_BB_M0_R1_S0
MY_000683_IEU+00_PM+00_EC_A0_D0_T0_BB_M0_R1_S0
MY_000684_IEU+00_PM+00_EC_A0_D0_T0_BB_M0_R1_S0
MY_000685_IEU+00_PM+00_EC_A0_D0_T0_BB_M0_R1_S0
MY_000686_IEU+00_PM+00_EC_A0_D0_T0_BB_M0_R1_S0
MY_000689_IEU+00_PM+00_EC_A0_D0_T0_BB_M0_R0_S0
MY_000690_IEU+00_PM+00_EC_A0_D0_T0_BB_M0_R0_S0
MY_000691_IEU+00_PM+00_EC_A0_D0_T0_BB_M0_R0_S0
MY_000693_IEU+00_PM+00_EC_A0_D0_T0_BB_M0_R0_S0
MY_000694_IEU+00_PM+00_EC_A0_D0_T0_BB_M0_R1_S0
MY_000695_IEU+00_PM+00_EC_A0_D0_T0_BB_M0_R1_S0
MY_000696_IEU+00_PM+00_EC_A0_D0_T0_BB_M0_R1_S0
MY_000697_IEU+00_PM+00_EC_A0_D0_T0_BB_M0_R0_S0
MY_000698_IEU+00_PM+00_EC_A0_D0_T0_BB_M0_R1_S0
MY_000699_IEU+00_PM+00_EC_A0_D0_T0_BB_M0_R1_S0
MY_000700_IEU+00_PM+00_EC_A0_D0_T0_BB_M0_R1_S0
MY_000701_IEU+00_PM+00_EC_A0_D0_T0_BB_M0_R1_S0
MY_000702_IEU+00_PM+00_EC_A0_D0_T0_BB_M0_R1_S0

MY_000703_IEU+00_PM+00_EC_A0_D0_T0_BB_M0_R1_S0
MY_000705_IEU+00_PM+00_EC_A0_D0_T0_BB_M0_R0_S0
MY_000710_IEU+00_PM+00_EC_A0_D0_T0_BB_M0_R0_S0
MY_000715_IEU+00_PM+00_EC_A0_D0_T0_BB_M0_R1_S0
MY_000716_IEU+00_PM+00_EC_A0_D0_T0_BB_M0_R1_S0
MY_000717_IEU+00_PM+00_EC_A0_D0_T0_BB_M0_R1_S0
MY_000718_IEU+00_PM+00_EC_A0_D0_T0_BB_M0_R1_S0
MY_000719_IEU+00_PM+00_EC_A0_D0_T0_BB_M0_R0_S0
MY_000720_IEU+00_PM+00_EC_A0_D0_T0_BB_M0_R0_S0
MY_000721_IEU+00_PM+00_EC_A0_D0_T0_BB_M0_R0_S0
MY_000722_IEU+00_PM+00_EC_A0_D0_T0_BB_M0_R0_S0
MY_000723_IEU+00_PM+00_EC_A0_D0_T0_BB_M0_R1_S0

Tabelle 3: Test-Sets D

Test-Set E
MY_000102_IEU+00_PM+00_ES_A0_D0_T0_BB_M0_R0_S0
MY_000103_IEU+00_PM+00_ES_A0_D0_T0_BB_M0_R1_S0
MY_000104_IEU+00_PM+00_ES_A0_D0_T0_BB_M0_R0_S0
MY_000105_IEU+00_PM+00_ES_A0_D0_T0_BB_M0_R1_S0
MY_000114_IEU+00_PM+00_ES_A0_D0_T0_BB_M0_R0_S0
MY_000123_IEU+00_PM+00_ES_A0_D0_T0_BB_M0_R1_S0
MY_000124_IEU+00_PM+00_ES_A0_D0_T0_BB_M0_R1_S0
MY_000125_IEU+00_PM+00_ES_A0_D0_T0_BB_M0_R1_S0
MY_000126_IEU+00_PM+00_ES_A0_D0_T0_BB_M0_R1_S0
MY_000133_IEU+00_PM+00_ES_A0_D0_T0_BB_M0_R1_S0
MY_000137_IEU+00_PM+00_ES_A0_D0_T0_BB_M0_R1_S0
MY_000138_IEU+00_PM+00_ES_A0_D0_T0_BB_M0_R0_S0
MY_000140_IEU+00_PM+00_ES_A0_D0_T0_BB_M0_R0_S0
MY_000149_IEU+00_PM+00_ES_A0_D0_T0_BB_M0_R1_S0
MY_000153_IEU+00_PM+00_ES_A0_D0_T0_BB_M0_R0_S0
MY_000154_IEU+00_PM+00_ES_A0_D0_T0_BB_M0_R0_S0
MY_000155_IEU+00_PM+00_ES_A0_D0_T0_BB_M0_R0_S0
MY_000159_IEU+00_PM+00_ES_A0_D0_T0_BB_M0_R0_S0
MY_000160_IEU+00_PM+00_ES_A0_D0_T0_BB_M0_R0_S0
MY_000166_IEU+00_PM+00_ES_A0_D0_T0_BB_M0_R0_S0
MY_000168_IEU+00_PM+00_ES_A0_D0_T0_BB_M0_R1_S0
MY_000169_IEU+00_PM+00_ES_A0_D0_T0_BB_M0_R1_S0
MY_000170_IEU+00_PM+00_ES_A0_D0_T0_BB_M0_R1_S0
MY_000171_IEU+00_PM+00_ES_A0_D0_T0_BB_M0_R1_S0
MY_000177_IEU+00_PM+00_ES_A0_D0_T0_BB_M0_R0_S0
MY_000178_IEU+00_PM+00_ES_A0_D0_T0_BB_M0_R1_S0
MY_000180_IEU+00_PM+00_ES_A0_D0_T0_BB_M0_R0_S0
MY_000183_IEU+00_PM+00_ES_A0_D0_T0_BB_M0_R0_S0
MY_000185_IEU+00_PM+00_ES_A0_D0_T0_BB_M0_R0_S0
MY_000190_IEU+00_PM+00_ES_A0_D0_T0_BB_M0_R0_S0
MY_000191_IEU+00_PM+00_ES_A0_D0_T0_BB_M0_R0_S0
MY_000195_IEU+00_PM+00_ES_A0_D0_T0_BB_M0_R0_S0

MY_000196_IEU+00_PM+00_ES_A0_D0_T0_BB_M0_R0_S0
MY_000197_IEU+00_PM+00_ES_A0_D0_T0_BB_M0_R0_S0
MY_000198_IEU+00_PM+00_ES_A0_D0_T0_BB_M0_R0_S0
MY_000203_IEU+00_PM+00_ES_A0_D0_T0_BB_M0_R0_S0
MY_000204_IEU+00_PM+00_ES_A0_D0_T0_BB_M0_R0_S0
MY_000207_IEU+00_PM+00_ES_A0_D0_T0_BB_M0_R0_S0
MY_000227_IEU+00_PM+00_ES_A0_D0_T0_BB_M0_R1_S0
MY_000558_IEU+00_PM+00_ES_A0_D0_T0_BB_M0_R0_S0
MY_000616_IEU+00_PM+00_ES_A0_D0_T0_BB_M0_R1_S0
MY_000617_IEU+00_PM+00_ES_A0_D0_T0_BB_M0_R1_S0
MY_000623_IEU+00_PM+00_ES_A0_D0_T0_BB_M0_R0_S0
MY_000627_IEU+00_PM+00_ES_A0_D0_T0_BB_M0_R1_S0
MY_000629_IEU+00_PM+00_ES_A0_D0_T0_BB_M0_R1_S0
MY_000630_IEU+00_PM+00_ES_A0_D0_T0_BB_M0_R0_S0
MY_000633_IEU+00_PM+00_ES_A0_D0_T0_BB_M0_R1_S0
MY_000634_IEU+00_PM+00_ES_A0_D0_T0_BB_M0_R0_S0
MY_000635_IEU+00_PM+00_ES_A0_D0_T0_BB_M0_R0_S0
MY_000636_IEU+00_PM+00_ES_A0_D0_T0_BB_M0_R0_S0
MY_000641_IEU+00_PM+00_ES_A0_D0_T0_BB_M0_R0_S0
MY_000651_IEU+00_PM+00_ES_A0_D0_T0_BB_M0_R1_S0
MY_000652_IEU+00_PM+00_ES_A0_D0_T0_BB_M0_R0_S0
MY_000653_IEU+00_PM+00_ES_A0_D0_T0_BB_M0_R0_S0
MY_000654_IEU+00_PM+00_ES_A0_D0_T0_BB_M0_R0_S0
MY_000655_IEU+00_PM+00_ES_A0_D0_T0_BB_M0_R0_S0
MY_000657_IEU+00_PM+00_ES_A0_D0_T0_BB_M0_R1_S0
MY_000658_IEU+00_PM+00_ES_A0_D0_T0_BB_M0_R1_S0
MY_000660_IEU+00_PM+00_ES_A0_D0_T0_BB_M0_R1_S0
MY_000661_IEU+00_PM+00_ES_A0_D0_T0_BB_M0_R1_S0
MY_000663_IEU+00_PM+00_ES_A0_D0_T0_BB_M0_R0_S0
MY_000667_IEU+00_PM+00_ES_A0_D0_T0_BB_M0_R1_S0
MY_000668_IEU+00_PM+00_ES_A0_D0_T0_BB_M0_R1_S0
MY_000669_IEU+00_PM+00_ES_A0_D0_T0_BB_M0_R1_S0
MY_000670_IEU+00_PM+00_ES_A0_D0_T0_BB_M0_R1_S0
MY_000671_IEU+00_PM+00_ES_A0_D0_T0_BB_M0_R0_S0
MY_000673_IEU+00_PM+00_ES_A0_D0_T0_BB_M0_R0_S0
MY_000679_IEU+00_PM+00_ES_A0_D0_T0_BB_M0_R1_S0
MY_000680_IEU+00_PM+00_ES_A0_D0_T0_BB_M0_R1_S0
MY_000681_IEU+00_PM+00_ES_A0_D0_T0_BB_M0_R0_S0
MY_000682_IEU+00_PM+00_ES_A0_D0_T0_BB_M0_R1_S0
MY_000683_IEU+00_PM+00_ES_A0_D0_T0_BB_M0_R1_S0
MY_000684_IEU+00_PM+00_ES_A0_D0_T0_BB_M0_R1_S0
MY_000685_IEU+00_PM+00_ES_A0_D0_T0_BB_M0_R1_S0
MY_000686_IEU+00_PM+00_ES_A0_D0_T0_BB_M0_R1_S0
MY_000689_IEU+00_PM+00_ES_A0_D0_T0_BB_M0_R0_S0
MY_000690_IEU+00_PM+00_ES_A0_D0_T0_BB_M0_R0_S0
MY_000691_IEU+00_PM+00_ES_A0_D0_T0_BB_M0_R0_S0
MY_000693_IEU+00_PM+00_ES_A0_D0_T0_BB_M0_R0_S0
MY_000694_IEU+00_PM+00_ES_A0_D0_T0_BB_M0_R1_S0
MY_000695_IEU+00_PM+00_ES_A0_D0_T0_BB_M0_R1_S0

MY_000696_IEU+00_PM+00_ES_A0_D0_T0_BB_M0_R1_S0
MY_000697_IEU+00_PM+00_ES_A0_D0_T0_BB_M0_R0_S0
MY_000698_IEU+00_PM+00_ES_A0_D0_T0_BB_M0_R1_S0
MY_000699_IEU+00_PM+00_ES_A0_D0_T0_BB_M0_R1_S0
MY_000700_IEU+00_PM+00_ES_A0_D0_T0_BB_M0_R1_S0
MY_000701_IEU+00_PM+00_ES_A0_D0_T0_BB_M0_R1_S0
MY_000702_IEU+00_PM+00_ES_A0_D0_T0_BB_M0_R1_S0
MY_000703_IEU+00_PM+00_ES_A0_D0_T0_BB_M0_R1_S0
MY_000705_IEU+00_PM+00_ES_A0_D0_T0_BB_M0_R0_S0
MY_000710_IEU+00_PM+00_ES_A0_D0_T0_BB_M0_R0_S0
MY_000715_IEU+00_PM+00_ES_A0_D0_T0_BB_M0_R1_S0
MY_000716_IEU+00_PM+00_ES_A0_D0_T0_BB_M0_R1_S0
MY_000717_IEU+00_PM+00_ES_A0_D0_T0_BB_M0_R1_S0
MY_000718_IEU+00_PM+00_ES_A0_D0_T0_BB_M0_R1_S0
MY_000719_IEU+00_PM+00_ES_A0_D0_T0_BB_M0_R0_S0
MY_000720_IEU+00_PM+00_ES_A0_D0_T0_BB_M0_R0_S0
MY_000721_IEU+00_PM+00_ES_A0_D0_T0_BB_M0_R0_S0
MY_000722_IEU+00_PM+00_ES_A0_D0_T0_BB_M0_R0_S0
MY_000723_IEU+00_PM+00_ES_A0_D0_T0_BB_M0_R1_S0

Tabelle 4: Test-Sets E

Bündelgraph für Test-Set C<->D
MY_000001_IEU+00_PM+00_EC_A0_D0_T0_BB_M0_R1_S0
MY_000002_IEU+00_PM+00_EC_A0_D0_T0_BB_M0_R1_S0
MY_000003_IEU+00_PM+00_EC_A0_D0_T0_BB_M0_R1_S0
MY_000005_IEU+00_PM+00_EC_A0_D0_T0_BB_M0_R1_S0
MY_000006_IEU+00_PM+00_EC_A0_D0_T0_BB_M0_R1_S0
MY_000008_IEU+00_PM+00_EC_A0_D0_T0_BB_M0_R1_S0
MY_000009_IEU+00_PM+00_EC_A0_D0_T0_BB_M0_R1_S0
MY_000011_IEU+00_PM+00_EC_A0_D0_T0_BB_M0_R1_S0
MY_000012_IEU+00_PM+00_EC_A0_D0_T0_BB_M0_R1_S0
MY_000014_IEU+00_PM+00_EC_A0_D0_T0_BB_M0_R1_S0
MY_000016_IEU+00_PM+00_EC_A0_D0_T0_BB_M0_R1_S0
MY_000017_IEU+00_PM+00_EC_A0_D0_T0_BB_M0_R1_S0
MY_000018_IEU+00_PM+00_EC_A0_D0_T0_BB_M0_R1_S0
MY_000019_IEU+00_PM+00_EC_A0_D0_T0_BB_M0_R1_S0
MY_000020_IEU+00_PM+00_EC_A0_D0_T0_BB_M0_R1_S0
MY_000021_IEU+00_PM+00_EC_A0_D0_T0_BB_M0_R1_S0
MY_000022_IEU+00_PM+00_EC_A0_D0_T0_BB_M0_R1_S0
MY_000024_IEU+00_PM+00_EC_A0_D0_T0_BB_M0_R1_S0
MY_000026_IEU+00_PM+00_EC_A0_D0_T0_BB_M0_R0_S0
MY_000027_IEU+00_PM+00_EC_A0_D0_T0_BB_M0_R1_S0
MY_000028_IEU+00_PM+00_EC_A0_D0_T0_BB_M0_R1_S0
MY_000029_IEU+00_PM+00_EC_A0_D0_T0_BB_M0_R1_S0
MY_000030_IEU+00_PM+00_EC_A0_D0_T0_BB_M0_R1_S0
MY_000031_IEU+00_PM+00_EC_A0_D0_T0_BB_M0_R1_S0
MY_000033_IEU+00_PM+00_EC_A0_D0_T0_BB_M0_R1_S0

MY_000034_IEU+00_PM+00_EC_A0_D0_T0_BB_M0_R1_S0
MY_000036_IEU+00_PM+00_EC_A0_D0_T0_BB_M0_R0_S0
MY_000037_IEU+00_PM+00_EC_A0_D0_T0_BB_M0_R1_S0
MY_000038_IEU+00_PM+00_EC_A0_D0_T0_BB_M0_R1_S0
MY_000039_IEU+00_PM+00_EC_A0_D0_T0_BB_M0_R1_S0
MY_000040_IEU+00_PM+00_EC_A0_D0_T0_BB_M0_R1_S0
MY_000041_IEU+00_PM+00_EC_A0_D0_T0_BB_M0_R0_S0
MY_000042_IEU+00_PM+00_EC_A0_D0_T0_BB_M0_R1_S0
MY_000043_IEU+00_PM+00_EC_A0_D0_T0_BB_M0_R1_S0
MY_000044_IEU+00_PM+00_EC_A0_D0_T0_BB_M0_R1_S0
MY_000001_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R1_S0
MY_000002_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R1_S0
MY_000003_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R1_S0
MY_000004_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R1_S0
MY_000005_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R1_S0
MY_000006_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R1_S0
MY_000008_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R1_S0
MY_000009_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R1_S0
MY_000011_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R1_S0
MY_000012_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R1_S0
MY_000014_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R1_S0
MY_000016_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R1_S0
MY_000017_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R1_S0
MY_000018_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R1_S0
MY_000019_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R1_S0
MY_000020_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R1_S0
MY_000021_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R1_S0
MY_000022_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R1_S0
MY_000024_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R1_S0
MY_000026_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R0_S0
MY_000027_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R1_S0
MY_000028_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R1_S0
MY_000029_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R1_S0
MY_000030_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R1_S0
MY_000031_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R1_S0
MY_000033_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R1_S0
MY_000034_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R1_S0
MY_000036_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R0_S0
MY_000037_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R1_S0
MY_000038_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R1_S0
MY_000039_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R1_S0
MY_000040_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R1_S0
MY_000041_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R0_S0
MY_000042_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R1_S0
MY_000043_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R1_S0

Tabelle 5: Bündelgraph für Test-Set C<->D

Bündelgraph für Test-Set C<-->E
MY_000001_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R1_S0
MY_000002_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R1_S0
MY_000003_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R1_S0
MY_000004_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R1_S0
MY_000005_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R1_S0
MY_000006_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R1_S0
MY_000008_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R1_S0
MY_000009_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R1_S0
MY_000011_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R1_S0
MY_000012_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R1_S0
MY_000014_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R1_S0
MY_000016_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R1_S0
MY_000017_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R1_S0
MY_000018_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R1_S0
MY_000019_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R1_S0
MY_000020_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R1_S0
MY_000021_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R1_S0
MY_000022_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R1_S0
MY_000024_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R1_S0
MY_000026_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R0_S0
MY_000027_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R1_S0
MY_000028_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R1_S0
MY_000029_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R1_S0
MY_000030_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R1_S0
MY_000031_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R1_S0
MY_000033_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R1_S0
MY_000034_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R1_S0
MY_000036_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R0_S0
MY_000037_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R1_S0
MY_000038_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R1_S0
MY_000039_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R1_S0
MY_000040_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R1_S0
MY_000041_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R0_S0
MY_000042_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R1_S0
MY_000043_IEU+00_PM+00_EN_A0_D0_T0_BB_M0_R1_S0
MY_000001_IEU+00_PM+00_ES_A0_D0_T0_BB_M0_R1_S0
MY_000002_IEU+00_PM+00_ES_A0_D0_T0_BB_M0_R1_S0
MY_000003_IEU+00_PM+00_ES_A0_D0_T0_BB_M0_R1_S0
MY_000005_IEU+00_PM+00_ES_A0_D0_T0_BB_M0_R1_S0
MY_000006_IEU+00_PM+00_ES_A0_D0_T0_BB_M0_R1_S0
MY_000008_IEU+00_PM+00_ES_A0_D0_T0_BB_M0_R1_S0
MY_000009_IEU+00_PM+00_ES_A0_D0_T0_BB_M0_R1_S0
MY_000011_IEU+00_PM+00_ES_A0_D0_T0_BB_M0_R1_S0
MY_000012_IEU+00_PM+00_ES_A0_D0_T0_BB_M0_R1_S0
MY_000014_IEU+00_PM+00_ES_A0_D0_T0_BB_M0_R1_S0
MY_000016_IEU+00_PM+00_ES_A0_D0_T0_BB_M0_R1_S0
MY_000017_IEU+00_PM+00_ES_A0_D0_T0_BB_M0_R1_S0
MY_000018_IEU+00_PM+00_ES_A0_D0_T0_BB_M0_R1_S0

MY_000019_IEU+00_PM+00_ES_A0_D0_T0_BB_M0_R1_S0
MY_000020_IEU+00_PM+00_ES_A0_D0_T0_BB_M0_R1_S0
MY_000021_IEU+00_PM+00_ES_A0_D0_T0_BB_M0_R1_S0
MY_000022_IEU+00_PM+00_ES_A0_D0_T0_BB_M0_R1_S0
MY_000024_IEU+00_PM+00_ES_A0_D0_T0_BB_M0_R1_S0
MY_000026_IEU+00_PM+00_ES_A0_D0_T0_BB_M0_R0_S0
MY_000027_IEU+00_PM+00_ES_A0_D0_T0_BB_M0_R1_S0
MY_000028_IEU+00_PM+00_ES_A0_D0_T0_BB_M0_R1_S0
MY_000029_IEU+00_PM+00_ES_A0_D0_T0_BB_M0_R1_S0
MY_000030_IEU+00_PM+00_ES_A0_D0_T0_BB_M0_R1_S0
MY_000031_IEU+00_PM+00_ES_A0_D0_T0_BB_M0_R1_S0
MY_000033_IEU+00_PM+00_ES_A0_D0_T0_BB_M0_R1_S0
MY_000034_IEU+00_PM+00_ES_A0_D0_T0_BB_M0_R1_S0
MY_000036_IEU+00_PM+00_ES_A0_D0_T0_BB_M0_R0_S0
MY_000037_IEU+00_PM+00_ES_A0_D0_T0_BB_M0_R1_S0
MY_000038_IEU+00_PM+00_ES_A0_D0_T0_BB_M0_R1_S0
MY_000039_IEU+00_PM+00_ES_A0_D0_T0_BB_M0_R1_S0
MY_000040_IEU+00_PM+00_ES_A0_D0_T0_BB_M0_R1_S0
MY_000041_IEU+00_PM+00_ES_A0_D0_T0_BB_M0_R0_S0
MY_000042_IEU+00_PM+00_ES_A0_D0_T0_BB_M0_R1_S0
MY_000043_IEU+00_PM+00_ES_A0_D0_T0_BB_M0_R1_S0
MY_000044_IEU+00_PM+00_ES_A0_D0_T0_BB_M0_R1_S0

Tabelle 6: Bündelgraph für Test-Set C<->E

Literatur

[1] Andreas Rieder Alfred Karl Louis, Peter Maaß. *Wavelets - Theorie und Anwendungen*. Teubner, 1998.

[2] Christian Blatter. *Wavelets - Eine Einführung*. Vieweg, 2003.

[3] W. Bäni. *Wavelets - Eine Einführung für Ingenieure*. Oldenbourg Verlag, 2005.

[4] Martin Goik. *Modellbasierte Objekterkennung mittels Neuronaler Netze unter Verwendung textureller Merkmale*. PhD thesis, 2007.

[5] Florian Hardt. *Robuste Bilderkennung mit lokalen linearen Abbildungen und elastischer Graphenanpassung*. PhD thesis, Universität Stuttgart, 2006.

[6] Herbert Amann, Joachim Escher. *Analysis 3*. Birkhäuser, 2001.

[7] Martin Lindberg Jöran Bergh, Fredrik Ekstedt. *Wavelets mit Anwendungen in Signal- und Bildverarbeitung*. Springer, 2007.

[8] Khalid Sayood. *An introduction to data compression*. Morgan Kaufmann, 2006.

[9] Laurenz Wiskott and Jean-Marc Fellous and Norbert Krüger and Christoph von der Malsburg. Face recognition by elastic bunchgraph matching. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1997.

[10] C.C. Noack. Hilbertraum-kompendium. *Universität Bremen*, 2003.

[11] Joachim Wehler Otto Forster. Fourier-transformation und wavelets. *LMU München*, 2000.

[12] Rolf P. Würtz. *Multilayer dynamic link networks for establishing image point correspondences and visual object recognition*. Verlag Harri Deutsch, 1995.

[13] Andreas H.J. Tewes. *A flexible Object Model for Encoding and Matching Human Faces*. PhD thesis, 2005.

[14] Richard F. Thompson. *Das Gehirn - Von der Nervenzelle zur Verhaltenssteuerung*. Spektrum Akademischer Verlag, 2001.

[15] Tilo Strutz. *Bilddatenkompression*. Vieweg, 2005.

[16] Martin Werner. *Information und Codierung*. Vieweg, 2002.

[17] William H. Press, Saul A. Teukolsky, William T. Vetterling. *Numerical recipes*. Cambridge University Press, 2007.

[18] Laurenz Wiskott, Jean-Marc Fellous, Norbert Krüger, and Christoph von der Malsburg. Distortion invariant object recognition in the dynamic link architecture. *IEEE Transactions on Computers*, 1993.

[19] Rolf P. Würtz. Praktikum Objekterkennung. *Institut für Neuroinformatik - Ruhr-Universität Bochum*, 2007.

Index

Ähnlichkeitsfunktionen, 48
2D - Fourier-Transformation, 36

Abelsche Gruppe, 10
Ableitung im Zeitbereich, 24
Abtast-Intervall, 19
Abtast-Theorem, 18
Aliasing, 18
Analyse, 14
Ausbreitung einer Funktion, 25
Axon, 6

Bündelgraph, 47
Bandbegrenzung, 18
Bijektivität der DFT, 20
Bilddatenkompression, 42

Cauchy-Schwarzsche-Ungleichung, 25
Chirp-Signal, 28
corpus geniculatem laterale, 6

Datenredundanz, 43
Dendriten, 6
Dilatation einer Funktion, 24
Diskrete Fourier Transformation (DFT), 20, 21
diskrete Wavelet-Transformation (DWT), 39
Diskretisierung, 18

einfache Zelle, 8
Einheitskreis, Funktionen auf, 13
Elastic Graph Matching, 48
Elastic-Bunch-Graph-Matching, 46
Entropie, 43
Erwartungswert, 25
Exhibition, 7

Faltungssatz, 18
Fast Fourier Transformation (FFT), 21
FeaSt, 46
FeaSt-Bild, 46
Fensterfunktion, 28
Filterantwort eines Wavelets, 35
Fourier-Koeffizienten, 14
Fourier-Reihen, 10, 12
Fourier-Transformation, 16
Fourier-Transformierte, 17
Funktionenräume, 12

Gabor-Funktion, 36
Gabor-Wavelet, 37
Ganglienzellen, 7
Ganglion, 6
Gesichtsklasse, 47

Haar-Mother-Wavelet, 32
Heisenberg-Box, 26
Heisenbergsche Unschärferelation, 22
Hilbertraum, 12
Huffman-Kodierung, 43
Hypersäule, 9

Inhibition, 7
Intensitätsverteilung, 7
Inverse Diskrete Fourier Transformation (IDFT), 21

Jet, 46
Jet-Bündel, 47

Kompressionsalgorithmus, Komponenten eines, 42
Kompressionsrate, 44
kontinuierliche Wavelet-Transformation (CWT), 31
Korrelation, 34

Landmarken, 46
lineare Unabhängigkeit, 11
Linearität Fourier-Transformation, 18
Linearkombination, 11
local expert, 49
Lokalisierung, 22
Look-Up-Tabelle, 43

Modellgraph, 47
Mother-Wavelet, 31
Moves, 48
Multi-Skalen-Analyse (MSA), 42

Nervenzelle, 6
Neuron, 6

off-center-Feld, 7
on-center-Feld, 7
Orthogonalität, 10
Orthogonalprojektion, 14
Orthonormalität, 10
Ortsraum, 36

Parsevalsche Formel, 14, 18
Profil, 7

Quantisierung, 43

Redundanz der CWT, 35
Retina, 6
retinotop, 8
rezeptives Feld, 7

Scan-Global-Move, 49
Scan-Local-Move, 50
Scan-Scale-Move, 49
Schwartzsche Raum, 16
Short-Time-Fourier-Transformation (STFT), 29
Skalarprodukt, 10
Spektrogramm, 29
Synapse, 6
Synthese, 15

Texturanalyse, 28
topologischer Term / topologische Kosten, 48
Translation einer Funktion, 23

Vektorraum, 10
Vektorraumbasis, 11
Vektorunterraum, 10

Wahrscheinlichkeitsdichte, 25
Wavelet, 31
Wavelet als Bandpass-Filter, 32
Wavelet-Transformation, 31
Wavelet-Transformation als Faltung, 32
Wellenvektor, 39

Zeit-Frequenz-Analyse, 28
Zeitsignale, 16

Abbildungsverzeichnis

1 Textil . . . 4
2 Holz . . . 4
3 menschliches Gesicht . . . 4
4 Landmarken allgemein . . . 5
5 Landmarken eines Gesichts . . . 5
6 Prinzipieller Aufbau des menschlichen Sehapparates . . . 6
7 2 verbundene Nervenzellen . . . 7
8 on-center-Feld $\sigma_+ = 2.0, \sigma_- = 2.1$ . . . 8
9 off-center-Feld $\sigma_+ = 2.1, \sigma_- = 2.0$ . . . 8
10 Orthogonale Vektoren . . . 10
11 Skalarprodukt . . . 10
12 Linearkombination von 2 Vektoren . . . 11
13 Funktionen auf dem Einheitskreis . . . 13
14 Projektion einer Funktion f auf den aufgespannten Unterraum U_n . . . 14
15 Synthese einer Funktion mit 1 Summand . . . 15
16 Synthese einer Funktion mit 10 Summanden . . . 15
17 Synthese einer Funktion mit 50 Summanden . . . 16
18 Synthese einer Funktion mit 100 Summanden . . . 16
19 Synthese einer Funktion mit 1000 Summanden . . . 16
20 Original Funktion . . . 16
21 Funktionenräume . . . 17
22 Aliasing . . . 19
23 Schlechte zeitliche Lokalisierung . . . 22
24 Gute zeitliche Lokalisierung . . . 22
25 Gute Frequenz Lokalisierung . . . 22
26 Schlechte Frequenz Lokalisierung . . . 22
27 Verschiebung einer Funktion f . . . 23
28 Dilatation einer Funktion f . . . 24
29 Heisenberg Box . . . 27
30 Unschärfe der Sample Werte . . . 27
31 Unschärfe der Basisfunktionen e_k . . . 27
32 Gauss Funktion mit $\sigma = 5$ . . . 28
33 Gauss Funktion mit $\sigma = 15$ . . . 28
34 Chirp Signal $chirp_{1,5}(t)$ . . . 29
35 Gefenstertes Chirp Signal . . . 29
36 STFT($chirp_{1,200}$), $\sigma = 0.125$ . . . 29
37 STFT($chirp_{1,200}$), $\sigma = 0.25$ . . . 29
38 $chirp_{1,200}$, $\sigma = 0.125$ . . . 30
39 $chirp_{1,200}$, $\sigma = 0.25$ . . . 30
40 Unschärfe für $\sigma > 1$ . . . 31
41 Unschärfe für $0 < \sigma < 1$ . . . 31
42 Das Haarsche Mother-Wavelet . . . 32
43 $< \psi_{m,n}, \psi_{m',n'} >$ für $m \neq m'$ und $n \neq n'$ . . . 33
44 Haar-Approximation für Skalenwerte -2 bis 6 . . . 34
45 $\psi_{0,0}$ . . . 34
46 $\psi_{-1,0}$ . . . 34
47 $\psi_{-5,0}$ . . . 34

48 CWT($chirp_{2,6}(t)$) . . . 35
49 Fourier-Transformierte einer Gauss-Funktion ψ mit $\widehat{\psi}(0) = 0$ . . . 37
50 Gabor-Filter Realteil . . . 37
51 Gabor-Filter Imaginärteil . . . 37
52 Gabor-Kerne im Frequenzraum . . . 38
53 Diskretisierung der Parameter ϑ und a . . . 38
54 Ausgangsbild für Gabor-Transformation . . . 40
55 Diskrete Gabor-Transformation . . . 41
56 Kompression mit DWT und Haar-Wavelet, Schwellwert-Quantisierung . . . 42
57 Komponenten eines Kompressionsalgorithmus . . . 42
58 Look-Up-Tabelle . . . 43
59 Ausgangswerte . . . 43
60 Quantisierte Werte . . . 43
61 Feast-Bild . . . 46
62 Modellgraph . . . 47
63 Bündelgraph . . . 48
64 Graph-Matching . . . 49
65 Scan-Global-Move . . . 49
66 Scan-Scale-Move . . . 49
67 Scan-Local-Move . . . 50
68 Repräsentation eines Bündelgraphen durch eine Matrix . . . 52
69 Zusammenhang zwischen Bündelgraph und dazugehöriger Bündelgraph-Matrix . . . 53
70 Bündelgraph-Matrix für komplexwertige Jets eines Bündelgraphen . . . 53
71 Bündelgraph-Matrix mit markierten Zeilen . . . 54
72 Bündelgraph-Matrix der Form $1920x38$ ($M = 5, L = 8, n = 48$, 38 Modellgraphen) . . . 54
73 Ausgedünnte Bündelgraph-Matrix . . . 54
74 Bündelgraph und Zeilen der zugehörigen Bündelgraph-Matrix M . . . 56
75 Look-Up Tabelle für Zeilen-Index und Zeilennorm . . . 56
76 sortierte Look-Up Tabelle und Zusammenhang zwischen Segment und Matrix-Zeile . . . 56
77 Lokalisierungsfehler . . . 57
78 Visualisierung des Lokalisierungsfehlers . . . 57
79 Matrix elementweise sortiert . . . 58
80 Matrix sortiert nach Spalten- und Zeilennorm . . . 58
81 Lokalisierungsfehler unsortiert . . . 59
82 Lokalisierungsfehler elementweise sortiert . . . 59
83 Lokalisierungsfehler nach Norm sortiert . . . 59
84 Kreuzvergleich . . . 60
85 Beispielbilder aus der FaceGen und CAS-PEAL Datenbank . . . 60
86 Matching für Test-Sets A und B (komprimiert nach größter Norm) . . . 67
87 Matching für Test-Sets A und B (komprimiert nach größter Norm) . . . 67
88 Lokalisierungsfehler für die Test-Sets A und B . . . 68
89 Lokalisierungsfehler (kleinste Norm) unter Nutzung der FaceGen Datenbank . . . 70
90 Lokalisierungsfehler (größte Norm) unter Nutzung der FaceGen Datenbank . . . 71
91 Gegenüberstellung von durchschnittlichen Lokalisierungsfehler und Erkennungsrate unter Nutzung der FaceGen Datenbank (kleinste Norm) . . . 71
92 Gegenüberstellung von durchschnittlichen Lokalisierungsfehler und Erkennungsrate unter Nutzung der FaceGen Datenbank (größte Norm) . . . 72
93 Erkennungsrate für Anordnung $C \leftrightarrow D$, links nach kleinster Norm ausgedünnt, rechts nach größter Norm . . . 72
94 Erkennungsrate für Anordnung $C \leftrightarrow E$ . . . 72
...