

EBERHARD KARLS UNIVERSITÄT TÜBINGEN
LEHRSTUHL PROFESSOR DR. MALLOT

Ein Detektor für dynamische Verdeckung und Aufdeckung

BACHELORARBEIT IN KOGNITIONSWISSENSCHAFTEN

Autor:

LÉON, DANKERT

4151528

leon.dankert@student.uni-tuebingen.de

Betreuer:

PROF. DR. H. A. MALLOT

3. Februar 2021

Abstract

Der Effekt der erhöhten Tiefenwahrnehmung durch dynamische Verdeckung oder Aufdeckung ist ein messbares psycho-physiologisches Phänomen. Beschreibt man eine Verdeckungskante im $x-t$ -Raum mithilfe der Heaviside-Funktion, erzeugt dies im Fourierraum einen Phasensprung entlang des Frequenzbandes der zugrundeliegenden Bewegung. Mithilfe zweier Gabor-Filtern, die leicht neben diesem Frequenzband liegen, und an diesem gespiegelt werden, kann dieser Phasensprung detektiert werden. Durch die Gleichförmigkeit einer Bewegung, die für beide Filter genau dieselbe Energie produziert, negieren sich diese durchs subtrahieren. Eine Ver-/Aufdeckung, die einen Phasensprung und Vorzeichenwechsel auslöst, verstärkt sich hingegen bei dieser Subtraktion. Dadurch lässt sich ein Filterpaar erstellen, das nur bei dynamischer Ver-/Aufdeckung reagiert. Mehrere dieser spezifischen Filterpaare können zu einer Filterbank kombiniert werden, die einen großen Bereich des visuellen Spektrums abdecken. Diese Filterbank ist in der Lage zu bestimmen, wie stark eine Ver- bzw. Aufdeckung innerhalb eines jeglichen $x-t$ -Stimuli vorhanden ist. Dies lässt die Annahme folgen, dass der Effekt der dynamischen Ver-/Aufdeckung durch rezeptive Felder detektierbar ist, und somit ein Teil der frühen visuellen Wahrnehmung sein könnte.

Inhaltsverzeichnis

1	Einleitung	2
1.1	Zugrundeliegender psycho-physiologischer Effekt	2
1.2	Weitere Phänomene der frühen visuellen Wahrnehmung	3
1.3	Vorgehen	4
2	Theoretische Grundlagen	5
2.1	Darstellungen und Repräsentationen in $x-y-t$	5
2.1.1	Darstellung von Bewegung	5
2.1.2	Darstellung von Kanten	6
2.1.3	Darstellung von Verdeckung und Aufdeckung	6
2.2	Vereinfachende Annahmen	8
3	Ideen für einen Detektor	11
3.1	Fouriertransformierung des Effekts	11
3.2	Der eigentliche Detektor	12
4	Untersuchungen in MATLAB	14
4.1	Fouriertransformieren in MATLAB	14
4.2	Auswertungen in MATLAB	14
4.3	Bewegungsdetektion in MATLAB	16
4.4	Verdeckungsdetektion in MATLAB	17
4.5	Überprüfung der Ergebnisse	19
5	Erstellung einer Filterbank	20
5.1	Energieberechnung	21
5.2	Weitere Stimuli	23
6	Zusammenfassung und Ausblick	24
	Literatur	26
	Anhang	28

1 Einleitung

Tagtäglich nehmen wir unzählige Objekte in unterschiedlichen Anordnungen, Größen und Formen wahr und es fällt uns nahezu spielerisch einfach zu unterscheiden, welches dieser Objekte sich im Vordergrund und welches sich im Hintergrund befindet. Es wird nicht ein Gedanke daran verschwendet, ob das Medium, auf dem diese Arbeit zu lesen ist, nun im Vordergrund ist oder eine ebene Fläche mit dem Hintergrund bildet. Ganz offensichtlich nehmen wir den Bildschirm oder das Blatt Papier als näher an unserem visuellen System liegend wahr, als die dahinter bestehende Oberfläche. Das mag zwar ganz logisch klingen, dennoch verbirgt sich hier, wie immer wenn es um visuelle Wahrnehmung und etwas vermeintlich Intuitives geht, ein ziemlich komplexes System. Dieser Effekt der Objektsegmentierung und Tiefenwahrnehmung hat viele unterschiedliche Ursachen. Dies wird noch zusätzlich verstärkt, wenn Bewegung im Spiel ist; das heißt, dass eine texturierte Oberfläche eine andere ver- oder aufdeckt. Demonstrierend vorstellen kann man sich dies, wie in Abbildung 1 dargestellt, wenn ein Auto in der Ferne hinter eine Mauer fährt und scheinbar dahinter verschwindet. Intuitiv nehmen wir wahr, dass sich das Auto hinter der Wand befinden muss, obwohl das Bild, das auf unsere Retina trifft, durch seine Zweidimensionalität keinerlei Tiefeninformationen besitzt.

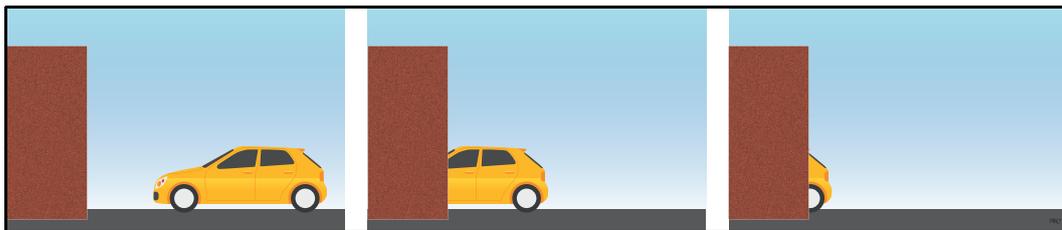


Abbildung 1: Versinnbildlichung des Verdeckungseffektes. Es sind drei aufeinanderfolgende Zeitpunkte dargestellt, wenn aus Sicht des Betrachters ein Auto hinter einer Mauer verschwindet. Die Mauer ist dabei die verdeckende Textur und das Auto der Hintergrund. (Eigene Darstellung)

1.1 Zugrundeliegender psycho-physiologischer Effekt

Das dies ein spezieller Fall der Tiefenwahrnehmung ist, der durch dynamische Ver-/Aufdeckung in einem zweidimensionalen Bild entsteht, haben erstmals Gibson, Kaplan, Reynolds, & Wheeler, 1969 in Erwägung gezogen. Experimentell nachgewiesen wurde er wenige Zeit später von Kaplan (1969). Kaplan nutzte einen Bildschirm, auf dem ein zufällig erstelltes schwarzweißes Rauschen dargestellt wurde. Dieses Rauschen wurde mithilfe einer nicht zu sehenden, rechtwinklig verlaufenden Kante in zwei Regionen separiert, welche bei keiner Bewegung wie eine kontinuierliche Fläche aussahen. Sobald sich aber mindestens eine der Seiten bewegte, wurde die abgrenzende Kante sichtbar und eine dynamische Ver- oder Aufdeckung war zu erkennen. Das Experimentaldesign sah nun vor, dass sich sowohl die beiden Regionen horizontal nach rechts oder links und ebenso die Abgrenzungskante mit kontinuierlicher Geschwindigkeit bewegen konnten. Den Studienteilnehmern wurden nun Videosequenzen präsentiert, in denen die verschiedenen Bedingungen mit unterschiedlichen Geschwindigkeiten kombiniert waren. Im Anschluss an jede Sequenz mussten die Teilnehmer entscheiden, welche der Regionen sie dem Vordergrund und welche dem Hintergrund zuordnen würden. Das Ergebnis zeigt, dass die Region, die in Besitz der auf- oder verdeckenden Kante war,

signifikant oft als im Vordergrund liegend wahrgenommen wurde. Der angrenzenden Region wurde dementsprechend zugeordnet, sich im Hintergrund zu befinden. Damit bestätigt sich die von Kaplan angenommene Hypothese.

Als weitere Befunde für diesen Effekt der Verdeckung/Aufdeckung sind unter anderem Studien von Yonas, Craton, & Thompson, 1987, Hildreth & Royden, 2011 und Tanrikulu, Froyen, Feldman, & Singh, 2018 aufzuführen. Yonas et al. (1987) haben gezeigt, dass die Bewegungsrichtung und -geschwindigkeit einer nicht statischen Regionsgrenze einen höheren Einfluss auf die Tiefenreihenfolge der Regionen, unabhängig von der Ver- bzw. Aufdeckungsrichtung der Texturen, hat. Hildreth & Royden (2011) hingegen konnten nachweisen, dass der Effekt durch dynamische Verdeckung sogar die Informationen vom binokularem Sehen überschreiben kann. Froyen, Feldman, & Singh stellten 2013 fest, dass der Tiefeneffekt bei Rotation von dreidimensionalen Objekten auf Selbstauf- bzw. verdeckung zurückzuführen ist. Auf dieser Grundlage konnten Tanrikulu et al. (2018) zeigen, dass die Konvexität der Aufdeckungs- oder Verdeckungskante ein wichtiger Faktor bei der Interpretation von Aufdeckung/Verdeckung ist.

1.2 Weitere Phänomene der frühen visuellen Wahrnehmung

Der Segmentierung von Regionen sind aber auch noch einige andere Faktoren zuzuschreiben. Angefangen mit der Gestaltpsychologie (Koffka, 1935/2013) und dem damit einhergehenden Anfang der Studien zu früher visueller Wahrnehmung, lassen sich einige Faktoren finden, die scheinbar einen Einfluss auf die Objektsegmentierung haben. So können beispielsweise umrahmende Kanten (Rubin, 1915/1958), die Symmetrie und Konvexität der Kantenanordnung (Kanizsa & Gerbino, 1976) oder das teilweise Hervorheben von Regionen (Froyen, Feldman, & Singh, 2010) die Unterscheidung in verschiedene Bereiche unterstützen. Eine detailliertere Zusammenfassung hierüber lässt sich bei Wagemans et al. (2012) finden.

Der frühen visuellen Wahrnehmung sind neben dem Segmentieren in verschiedene Regionen und damit in Objekte durch Ver- und Aufdeckung, auch noch andere Effekte zuzuschreiben. Besonders relevant für die Objekterkennung sind unter anderem die Kanten- und Bewegungsdetektion. Diese lassen sich durch rezeptive Felder beschreiben. Im Auge konvergieren hier mehrere Fotorezeptoren auf einzelnen oder wenigen Ganglienzellen. Dabei sind die Fotorezeptoren und die nachfolgende Ganglienzelle so miteinander verschaltet, dass diese nur ein Signal weiterleitet, wenn ein ganz spezieller, auf das rezeptive Feld passender, visueller Reiz auf die Retina trifft. Dies kann unter anderem eine bestimmte Geschwindigkeit aus einer bestimmten Richtung, oder eine Kante mit bestimmter Orientierung sein. Diese Art der Zellenverschaltung kann aber auch in anderen Bereichen der Sinnesrezeptoren auftauchen. (vgl. Mallot, 2013 und Wandell & Thomas, 1997)

Theoretisch lassen sich rezeptive Felder mit Gabor-Filtern approximieren (Cope, Blakeslee, & McCourt, 2009). Für die Erstellung solcher Gabor-Filter wird eine sinusförmige Ebene mit einer 'Gaußschen Glockenkurve' multipliziert, sodass der Filter auf ein bestimmtes Signal konvergiert. Adelson & Bergen erstellten so 1985 eine Gabor-Filterbank, mit der sich zuverlässig Bewegung detektieren ließ. Rechtsgerichtete Bewegung erzeugte ein positives Signal, linksläufige ein äquivalent negatives. Außerdem war die Art des Stimulus nicht von Relevanz, das heißt, dass Signal war für konträre Stimuli, schwarzer Balken vor weißem Grund und weißer Balken auf schwarzem Grund, gleich. Um diese Filterbank zu erzeugen, kombinierten Adelson & Bergen zwei räumlich orientierte und

zwei zeitlich abhängige Impulsantworten zu zwei kosinus- (geraden) und zwei sinusförmigen (ungeraden) Gabor-Filtern. Durch Addition bzw. Subtraktion eines geraden mit einem ungeraden Filter entstanden dann jeweils zwei links- und rechtsläufig selektive Filter. Diese wurden dann wiederum
75 kombiniert, indem die Linksgerichteten von den Rechtsgerichteten subtrahiert wurden und erzeugten so das oben beschriebene Verhalten - denn wenn der rechtsläufige Filter feuerte, war der linksläufige Filter gleich null und somit die Ausgabe positiv. Umgekehrt ergab sich eine negative Ausgabe bei Reaktion des linksläufigen Filters. Es ließ sich also ein einfacher Bewegungsdetektor für die frühe visuelle Wahrnehmung approximieren. Diesem Ansatz folgend, wird im Laufe dieser Arbeit eine
80 ähnliche Filterbank erstellt, die sich im Grunde nur durch einige Kleinigkeiten unterscheidet. Diese Filterbank soll dann nicht Bewegungsdetektion realisieren, sondern die Detektion von dynamischer Verdeckung oder Aufdeckung ausführen.

1.3 Vorgehen

Sollte der anfänglich beschriebene, kognitiv psycho-physiologische Effekt der Aufdeckung/Verdeckung wie zu erwarten ebenfalls in der frühen visuellen Wahrnehmung stattfinden, müsste sich ein
85 effektspezifisches, rezeptives Feld finden lassen. Im Laufe dieser Bachelorarbeit wird versucht genau dieses rezeptive Feld, durch Kombinationen aus Gabor-Filtern zu einer Filterbank, zu simulieren. Hierzu wird es zu Beginn eine kurze Einführung in die theoretischen Hintergründe geben. Dazu gehören die formelle Beschreibung von Bewegung allgemein, von Kanten und schließlich die Kom-
90 bination der beiden zum Effekt der Verdeckung/Aufdeckung. Nach der Definition des Effekts in Raum und Zeit werden einige vereinfachende Annahmen getroffen, die den Umgang und die Untersuchung einfacher und übersichtlicher gestalten sollen. Darauf folgt die Herleitung und Vorstellung der Idee, auf der der Detektor aufbaut. Diese erfolgt mithilfe der fouriertransformierten des vorher vereinfachten Stimulus. Nachdem diese rein theoretischen Überlegungen getroffen worden sind, wird
95 diese im Anschluss in MATLAB (2020) implementiert und untersucht. Mit den Resultaten wird dann eine Gabor-Filterbank implementiert, die der von Adelson & Bergen (1985) ähnelt. Zuerst wird ein einzelnes, spezifisches Filterpaar gebildet. Mit mehreren dieser Filterpaare wird dann die Filterbank für verschiedene Stimuli mit verschiedenen Geschwindigkeiten erzeugt. Danach werden die Ergebnisse kritisch hinterfragt und diskutiert und zum Abschluss wird es einige Ausblicke für
100 weitere Arbeiten geben.

2 Theoretische Grundlagen

2.1 Darstellungen und Repräsentationen in x-y-t

Für jede Kamera und ebenso unser Auge kann man Bewegung als dreidimensional in Raum und Zeit interpretieren. Dabei können die x - und y -Achsen als räumliche, und somit als Bild beschreibende, die z -Achse als zeitliche Dimension angenommen werden. Im weiteren Verlauf wird daher nur noch von x , y und t gesprochen. Um die Verdeckung und Aufdeckung eines zweidimensionalen Bildes mathematisch zu beschreiben, benötigt man aufgrund der zeitlichen Komponente eine dreidimensionale Darstellung. Daher werden Bewegung und Kanten so beschrieben und kombiniert, dass sie eine x - y - t -Darstellung der Verdeckung/Aufdeckung ergeben.

2.1.1 Darstellung von Bewegung

Nimmt man einen Bildausschnitt mit der bildgebenden Funktion $I(x, y)$ an, der sich mit der Geschwindigkeit $v \in \mathbb{R}^2$ bewegt, ergibt sich

$$I(x, y, t) = I(x + v_x \cdot t, y + v_y \cdot t), \quad (1)$$

wobei v_x die Bewegungsgeschwindigkeit in x -Richtung und v_y in y -Richtung beschreiben. Diese Bewegung ist in Abbildung 2 dargestellt. Hier bewegt sich ein Bildausschnitt mit schwarz-weißem Zufallsmuster mit kontinuierlicher Geschwindigkeit von links nach rechts. Das Zufallsmuster ist dabei immer dasselbe, nur die Position des Balkens innerhalb des gesamten Bildes verändert sich über die Zeit. Darstellung 2a zeigt die zweidimensionale Bildfunktion $I(x, y)$ und 2b die dreidimensionale Bewegung $I(x, y, t)$. Da es im gezeigten Beispiel nur eine Bewegung in x -Richtung, nicht jedoch in y gibt, beträgt $v_y = 0$. Für dieses Beispiel gilt, dass ein positives v_x eine rechtsgerichtete Bewegung; ein negatives v_x eine linksgerichtete Bewegung zur Folge hat.

Unser Auge nimmt jedoch nicht nur kontinuierliche Bewegung als solche wahr. Wenn die zeitliche Frequenz und räumliche Anordnung stimmen, wird auch Bewegung in diskreten Zeitschritten als kontinuierlich wahrgenommen. Jeder Film und jeder Bildschirm macht von diesem Phänomen Gebrauch. Wann es zu einer Wahrnehmung von Bewegung kommt, unter welchen Umständen sie als

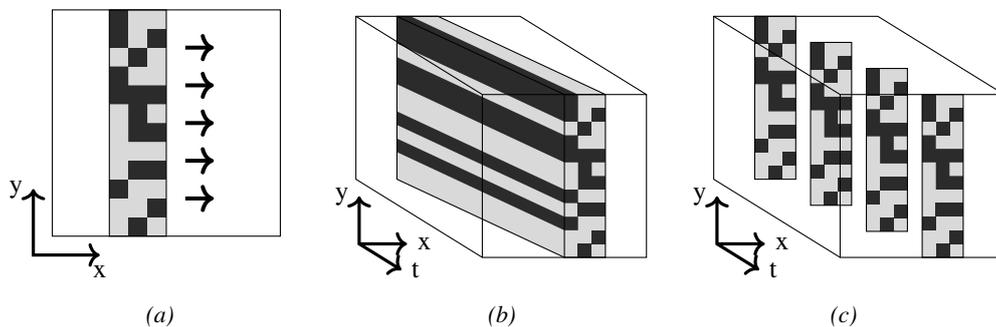


Abbildung 2: Darstellung von Bewegung. **2a** visualisiert die Wahrnehmung einer bewegten Textur. Die Pfeile beschreiben die Bewegungsrichtung. Der Balken bewegt sich nur in x -Richtung vor einem weißen Hintergrund. **2b** stellt dieselbe kontinuierliche Bewegung als x - y - t -Plot dar. **2c** zeigt die Bewegung in diskreten Zeitschritten. (Eigene Darstellung in Anlehnung an Adelson & Bergen, 1985)

125 kontinuierlich erscheint und wie Frequenz, Anordnung und Interpretation genauer zusammenhän-
 gen werden an dieser Stelle jedoch nicht weiter ausgeführt. Eine gute Zusammenfassung über diese
 Phänomene und Zusammenhänge bietet Sekuler, Watamaniuk, & Blake (2002). Für diese Thesis ist
 es ausreichend anzunehmen, dass auch Bewegung durch diskrete Schritte als kontinuierlich wahrgে-
 130 nommen werden kann (siehe Abbildung 2c), sodass wir diese folglich mit diskreten Zeitintervallen
 behandeln können.

2.1.2 Darstellung von Kanten

Eine Kante kann durch verschiedene Faktoren gesehen werden. Der triviale Fall ist der einer Kante in
 einem Bild. Hier verändert sich eine sonst stetige Bildfunktion schlagartig, also eine Textur wechselt
 zu einer anderen. In Figur 2 wäre das der Sprung des Rauschens zum weißen Hintergrund. Dieser
 135 Sprung kann oft auch als Treppenfunktion beschrieben werden.

Etwas schwieriger ist hingegen der Fall, wenn die Kante erst durch Bewegung wahrgenommen
 wird. Hier befindet sich die Kanteninformation nicht in der bildgebenden Funktion selbst, sondern
 nur in Abhängigkeit von t und somit beim Differenzieren zu vorherigen Bildern. Da im weiteren
 Verlauf jedoch nur mit Bildern gearbeitet wird, die starke Unterschiede in der Textur aufweisen
 140 und dementsprechend die Positionsinformation der Kanten unabhängig von t ist, wird dieser Fall
 ebenfalls nicht genauer erläutert. Bei Wunsch nach genaueren Informationen kann hier Ziou and
 Tabbone (1998) empfohlen werden.

Für den ersten beschriebenen Fall kann man folglich eine Kante gut mit einer Treppenfunktion
 approximieren. Für eine zur x -Achse orthogonalen Kante an der Stelle $x = 0$ bietet sich die Heaviside-
 145 Funktion \mathcal{H} mit

$$\mathcal{H}(x, y) = \begin{cases} 0 & : x < 0 \\ 1 & : x \geq 0 \end{cases} \quad (2)$$

an. Hat man nun ein Bild, welches aus zwei unterschiedlichen Bildfunktionen $I_1(x, y)$ und $I_2(x, y)$
 besteht, die an der Stelle $x = 0$ zusammengesetzt sind, wird die gesamte Bildfunktion $\mathcal{I}(x, y)$ durch

$$\mathcal{I}(x, y) = \begin{cases} I_1(x, y) & : x < 0 \\ I_2(x, y) & : x \geq 0 \end{cases} \quad (3)$$

dargestellt. Ausgedrückt mithilfe der Heaviside-Funktion aus 2 ergibt sich

$$\mathcal{I}(x, y) = I_1 \times (1 - \mathcal{H}(x, y)) + I_2 \times \mathcal{H}(x, y) \quad (4)$$

wobei \times punktweise Multiplikation symbolisiert. Bei dieser Formulierung wird der Fall angenom-
 150 men, dass es in y -Richtung zu keiner Kante kommt. Figur 3 soll die Entstehung der endgültigen
 Bildfunktion visualisieren. Die Formel 2 der Heaviside-Funktion ist in Figur 3a und zwei beispiel-
 hafte Bildfunktionen $I_1(x, y)$ und $I_2(x, y)$ sind in 3b und 3c dargestellt. Die kombinierte Bildfunktion
 $\mathcal{I}(x, y)$, die mit den Formeln 3 und 4 beschrieben wird, ist außerdem in Darstellung 3d zu sehen.

2.1.3 Darstellung von Verdeckung und Aufdeckung

155 Der Effekt der Verdeckung oder Aufdeckung entsteht nun aus einer Kombination einer Kante und
 Bewegung. Man hat die in x und y herrschenden Kanten sowie die in t stattfindende Bewegung.

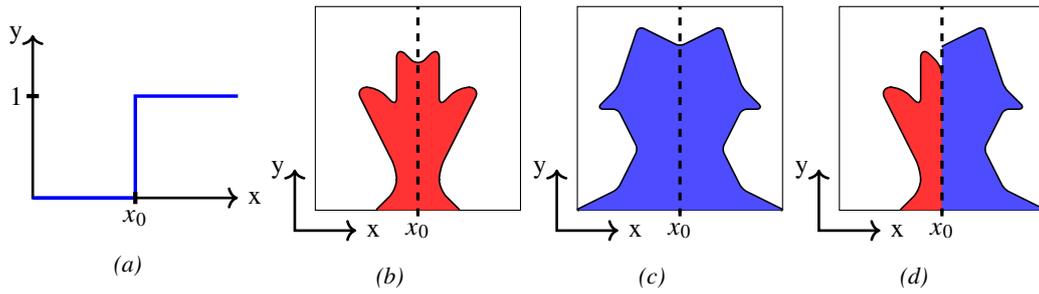


Abbildung 3: Entstehung einer kombinierten Bildfunktion nach Gleichung 4. **3a** ist die Heaviside-Funktion \mathcal{H} . **3b** und **3c** zeigen zwei unabhängige Bildfunktionen. **3d** visualisiert die Kombination der beiden Bildfunktionen. Dabei ist **3b** mit $1 - \mathcal{H}$ und **3c** mit \mathcal{H} multipliziert worden und anschließend zusammen addiert. (Eigene Darstellung)

Die endgültige Darstellung ist also wieder eine in x - y - t , aber mit zwei Bildfunktionen $I_1(x, y)$ und $I_2(x, y)$, die sich mit den Geschwindigkeiten $v_1, v_2 \in \mathbb{R}^2$ bewegen:

$$I_1(x, y, t) = I_1(x + v_1(x) \cdot t, y + v_1(y) \cdot t) \quad (5)$$

$$I_2(x, y, t) = I_2(x + v_2(x) \cdot t, y + v_2(y) \cdot t) \quad (6)$$

Das Vorzeichen der Bewegung determiniert die Richtung der Bewegung. Die Erzeugung einer links- bzw. rechtsgerichteten Bewegung bedarf ein $v(x) < 0$ bzw. $v(x) > 0$. Für eine Aufwärtsbewegung benötigt man dementsprechend ein $v(y) > 0$ und für eine nach unten verlaufende Bewegung ein $v(y) < 0$.

Um nun eine Verdeckung oder Aufdeckung zu erzeugen, müssen die Bewegungen orthogonal zur Kante, also entlang der x -Achse derselben Richtung, dasselbe Vorzeichen besitzen. Ansonsten kann eine Kombination aus Verdeckung und Aufdeckung entstehen, die zwar auch eine Tiefenwahrnehmung erzeugen kann, oft aber nur zu einem Flickern führt und an dieser Stelle nicht untersucht werden soll. Daher wird festgelegt, dass

$$v_1(x) \cdot v_2(x) \geq 0 \quad (7)$$

gelten muss. Die Richtung der Bewegung legt außerdem fest, welche der Texturen als Ver- bzw. Aufdecker und welche als Hintergrund wahrgenommen werden. Bei Bewegung nach links ist immer die linke Oberfläche im Hintergrund und wird aufgedeckt oder verdeckt, bei Bewegung nach rechts genau umgekehrt. Ob es sich um eine Verdeckung oder Aufdeckung handelt, bestimmt das Verhältnis der Geschwindigkeiten. Ist $v_1(x) < v_2(x)$ findet eine Aufdeckung, bei $v_1(x) > v_2(x)$ immer eine Verdeckung statt, unabhängig von der allgemeinen Bewegungsrichtung.

Da sich die Kante verschiebt, wenn beide Geschwindigkeiten $v_1(x), v_2(x) \neq 0$ sind, verschiebt sich die Funktion, mit der die Kante beschrieben wird; in diesem Fall die Heaviside-Funktion. Die Kante bewegt sich nur in x -Richtung mit

$$s = \begin{cases} v_1(x) & : v_1(x) > 0 \\ v_2(x) & : v_2(x) < 0 \\ 0 & : v_1(x) = 0, v_2(x) = 0. \end{cases} \quad (8)$$

Es soll keine Kantenbewegung geben, wenn sich eine der Texturen nicht bewegt, daher ist in diesen Fällen $s = 0$. Die neue Heaviside-Funktion ergibt sich aus Gleichung 2 kombiniert mit s zu

$$\mathcal{H}(x, y, t) = \mathcal{H}(x + s \cdot t). \quad (9)$$

Das endgültige, dreidimensionale Bild $\mathcal{I}(x, y, t)$ setzt sich aus den beiden bewegten Bildern $I_1(x, y, t)$ und $I_2(x, y, t)$, die durch die Heaviside-Funktion getrennt werden, zusammen. Die Bildfunktion der rechten Seite, also I_2 wird mit der normalen Heaviside-Funktion, die linke I_1 mit 1-Heaviside-Funktion multipliziert.

$$\mathcal{I}(x, y, t) = I_1(x, y, t) \times (1 - \mathcal{H}(x, y, t)) + I_2(x, y, t) \times \mathcal{H}(x, y, t) \quad (10)$$

Die Gleichung 10 beschreibt den zu untersuchenden Effekt der Aufdeckung und Verdeckung für zwei unterschiedliche, zweidimensionale Texturen, die sich gegenseitig verdecken oder aufdecken in Abhängigkeit von t .

2.2 Vereinfachende Annahmen

Um den Effekt einfacher und umfangreicher untersuchen zu können, wurden im Folgenden einige vereinfachende Annahmen getroffen. Die verwendete Heaviside-Funktion ist bereits eine dieser Vereinfachungen, da sie nur eine ganz spezifische Kantenart zulässt. Konvexität innerhalb der Kante wird in dieser Annahme nicht mit berücksichtigt. Weitergehend werden einige weitere Annahmen gemacht, um den Effekt für eine erste Untersuchung handlicher zu gestalten. Dabei wurde darauf geachtet, dass möglichst die Teile dem Vereinfachen zum Opfer fallen, die wenig oder gar nichts mit dem eigentlichen Effekt zu tun haben. Dazu gehören unter anderem die Bewegung der Texturen in y -Richtung und die gleichzeitige Bewegung beider Texturen generell. Damit ist gemeint, dass sich mit lediglich einer bewegten Textur beide Effekte, also Aufdeckung und Verdeckung, darstellen lassen. Wie in Abschnitt 2.1.3 erklärt, kommt es auf das Verhältnis der Bewegungen an. Somit lässt sich für $v_1(x) = 0$ mit $v_2(x) > 0$ eine Aufdeckung und mit $v_2(x) < 0$ eine Verdeckung erzeugen. Nimmt man also für $v(y) = 0$ an, kann man anstelle von $v_1(x)$ und $v_2(x)$ nur noch von v_1 und v_2 sprechen. Außerdem ergibt sich für die Bildfunktionen

$$I(x, y, t) = I(x + vt). \quad (11)$$

Dies für I_1 und I_2 zusammen mit dem Ergebnis von Gleichung 9 in 10 eingesetzt führt zu

$$\mathcal{I}(x, y, t) = I_1(x + v_1t) \times (1 - \mathcal{H}(x + st)) + I_2(x + v_2t) \times \mathcal{H}(x + st). \quad (12)$$

Damit eine Textur, in diesem Fall fällt die Wahl auf die Linke I_1 , statisch wird, muss lediglich $v_1 = 0$ gesetzt werden. Somit muss auch nicht mehr in v_1 und v_2 unterschieden werden, und es wird im Folgenden nur v verwendet. Hierdurch ergibt sich aber auch nach Gleichung 8, dass immer $s = 0$ gilt. Bildlich gesprochen heißt dies, dass sich die Ver-/Aufdeckungskante nicht mehr verschiebt. Diese beiden Änderungen zur 10. Gleichung hinzugefügt ergeben

$$\mathcal{I}(x, y, t) = I_1(x) \times (1 - \mathcal{H}(x)) + I_2(x + vt) \times \mathcal{H}(x). \quad (13)$$

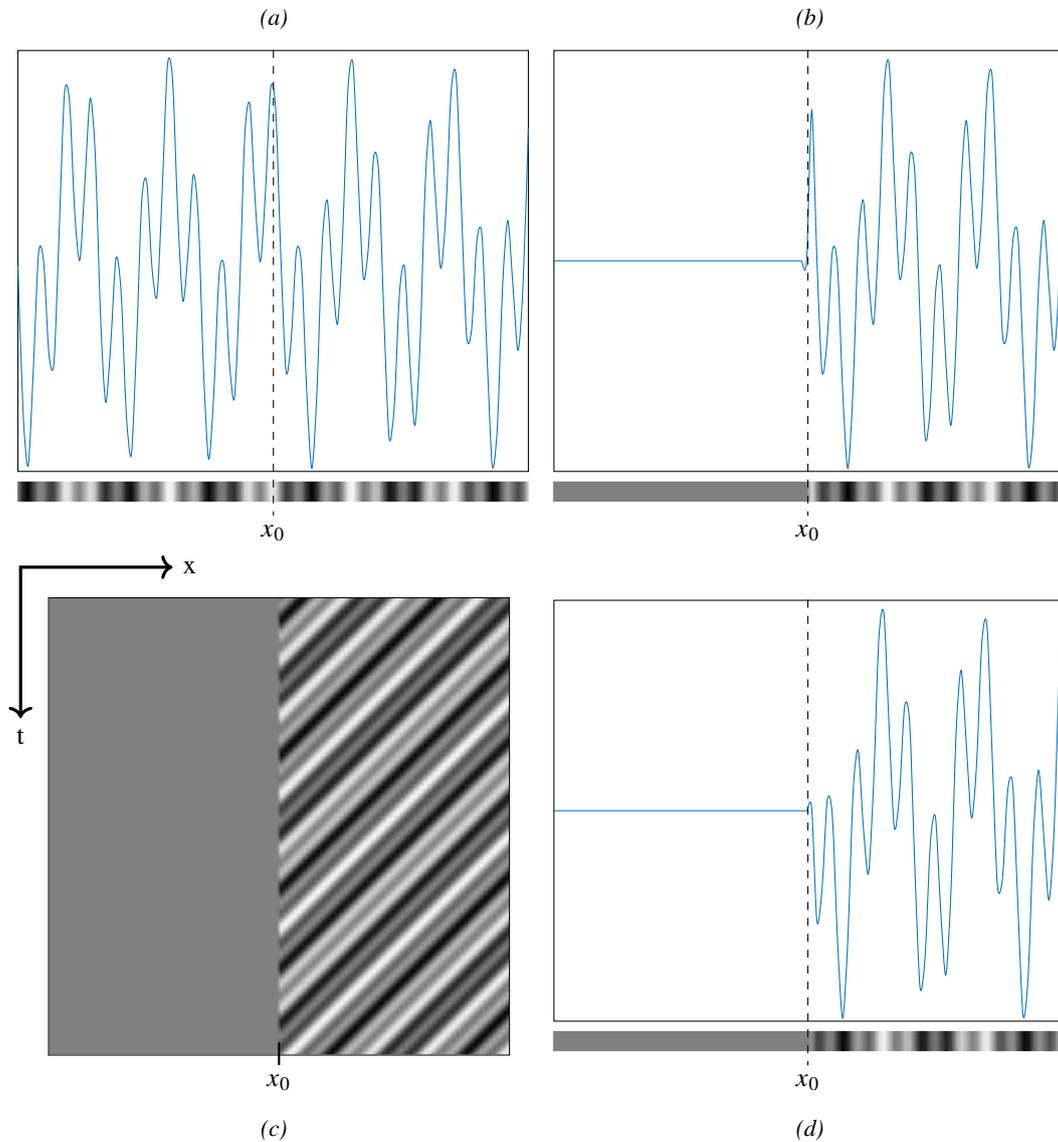


Abbildung 4: Erstellung des verwendeten $x-t$ -Stimulus. **4a** ist das verwendete Sinus-Kosinus Rauschen. Einmal als klassischer zweidimensionaler Graph und direkt unterhalb als Interpretation eines Grauwert-Streifens. **4b** zeigt dasselbe Rauschen mit einem Verdeckter auf der linken Hälfte, ebenfalls klassisch und als Grauwert-Streifen. **4d** zeigt den darauf folgenden Schritt $t + 1$, wenn man das Rauschen nach Links hinter den statischen Verdeckter rückt. **4c** zeigt den endgültigen Stimulus, der aus mehreren dieser von t abhängigen Grauwert-Streifen besteht, mit konstanter Geschwindigkeit v der Bewegung und konstanter Zeit. Das Spektrum bei allen reicht von -1 (schwarz) bis 1 (weiß). Dabei hat der Verdeckter immer den Wert 0. (Eigene Darstellung)

Als weiteren Schritt kann man die Textur der unbewegten Fläche entfernen, also $I_1 = 0$, dies entspricht einem schlichten, einfarbigen Aufdecker/Verdecker. Somit kann I_2 durch I substituiert werden.

$$\mathcal{I}(x, y, t) = 0 \times (1 - \mathcal{H}(x)) + I_2(x + vt) \times \mathcal{H}(x) \quad (14)$$

$$= I(x + vt) \times \mathcal{H}(x) \quad (15)$$

Nun ist nichts mehr von y abhängig und dementsprechend kann der gesamte Effekt auf einen zweidimensionalen x - t -Zusammenhang vereinfacht werden.

$$\mathcal{I}(x, t) = I(x + vt) \times \mathcal{H}(x) \quad (16)$$

210 In diesem Fall kann man sich $I(x, t)$ als einen einfachen, texturierten Streifen vorstellen (Figur 4a), der in diskreten Zeitschritten nach links oder rechts geschoben wird. In Figur 4b und 4d sind zwei aufeinanderfolgende Schritte $\mathcal{I}(x, t)$ und $\mathcal{I}(x, t + 1)$ dargestellt, die bei einer Verschiebung um genau eine x -Koordinate nach links entsteht. Die Geschwindigkeit ist also $v = \frac{\Delta x}{t} = -1$. Die Funktionswerte können, wie unter den Diagrammen zu sehen, auch als skaliertes Grauwert dargestellt werden, sodass
 215 man einen eindimensionalen Streifen pro Schritt erhält. Reiht man diese für folgende t 's gedanklich untereinander, erhält man ein raumzeitliches x - t -Grauwertbild, dass in Figur 4c visualisiert ist. Die gesamte Zeit t über ist die Geschwindigkeit konstant. Als bildgebende Funktion für x wurde ein einfaches Sinus-Kosinus Rauschen verwendet.

3 Ideen für einen Detektor

220 Die Zielsetzung dieser Arbeit verlangt nun eine Fouriertransformation der Gleichung 16. Die Fouriertransformation ist eine Transformation, mit der jedes periodische Signal in ihre einzelnen Frequenzen zerlegt werden kann. Demnach besteht jedes periodische Signal aus einer Kombination verschiedener Sinus- und Kosinusfunktionen. Die Fourierreihe \mathcal{F} kann jedes periodische Signal $f(t)$ durch

$$\mathcal{F} f(t) = \frac{a_0}{2} + \sum_{k=1}^{\infty} [a_k \cdot \cos(k \cdot 2\pi \cdot t) + b_k \cdot \sin(k \cdot 2\pi \cdot t)] \quad (17)$$

mit den Vorfaktoren a_k und b_k approximieren. So kann beispielsweise ein Akkord in seine einzelnen
225 Töne aufgegliedert werden. Die Fouriertransformation wird auch als Integralfunktion mit

$$\mathcal{F} f(t) = \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} f(x) e^{-it \cdot x} dx \quad (18)$$

beschrieben, wobei i die imaginäre Einheit mit $i^2 = -1$ und $t \cdot x$ das Standardskalarprodukt der Vektoren x und t ist. Für die Bildanalyse kann man einen Bildausschnitt als eine Periode festlegen. Dadurch lässt sich auch ein Bild fouriertransformieren und somit das Frequenzspektrum untersuchen. Übersichtliche Erklärungen, Beschreibungen und Beweise lassen sich bei Mallot (2013) und James
230 (2011) finden, weshalb eine genauere Erläuterung dieses komplizierten mathematischen Prozesses an dieser Stelle ausgespart wird. Wichtig ist nur, dass man durch die Fouriertransformation ein Signal in seine Frequenzkomponenten zerlegen kann. Im Folgenden wird dies ein raumzeitliches Signal sein, dass in seine räumlichen und seine zeitlichen Frequenzen isoliert wird.

3.1 Fouriertransformierung des Effekts

235 Da die Fouriertransformation \mathcal{F} eine lineare Operation ist, kann man die Transformation der Gesamtbildfunktion \mathcal{I} nach Gleichung 16 folgendermaßen definieren:

$$\mathcal{F}\mathcal{I}(x, t) = \mathcal{F}(I(x + vt) \times \mathcal{H}(x)) \quad (19)$$

$$= \mathcal{F}I(x + vt) * \mathcal{F}\mathcal{H}(x) \quad (20)$$

Durch die Transformation in den Fourierraum wird die punktweise Multiplikation (\times) durch eine Faltung ($*$) ersetzt. Nun kann die Bildfunktion und die Heaviside-Funktion unabhängig voneinander transformiert und danach durch eine Faltung wieder zusammengeführt werden.

240 Um die Funktion der Bewegung $I(x + vt)$ geeignet transformieren zu können, wird diese durch eine andere Schreibweise definiert,

$$I(x + vt) = \int I(x') \delta(x + vt + x') dx' \quad (21)$$

$$= I(x) * \delta(x + vt) \quad (22)$$

wobei δ die Delta-Distribution oder auch Impulsfunktion darstellt. Diese liefert

$$\forall x \neq 0, \delta(x) = 0. \quad (23)$$

Mit Hinblick auf die Linearität und das eben beschriebene Faltungs-Theorem kann man die Bestandteile der Gleichung 21 wieder unabhängig voneinander transformieren.

$$\mathcal{F}I(x + vt) = \mathcal{F}(I(x) * \delta(x + vt)) \quad (24)$$

$$= \mathcal{F}I(x) \times \mathcal{F}\delta(x + vt) \quad (25)$$

245 Die Fouriertransformation von $\delta(x + vt)$ ist $\delta(\omega + v/\nu)$ mit ω als räumliche Frequenz des Bildes $I(x)$ und ν als zeitliche Frequenz der Bewegung in t (vgl. Korn & von Seelen, 1972). Dementsprechend gilt:

$$\mathcal{F}I(x + vt) = \mathcal{F}I(x) \times \delta(\omega + \nu/\nu). \quad (26)$$

Dies ist somit die Fouriertransformation einer kontinuierlichen Bewegung ohne verdeckende Kante. Eine Geschwindigkeit von $\nu = 0$ ist logischerweise nicht zulässig, dementsprechend beschreibt diese Gleichung ausnahmslos ein bewegtes Bild. Für die eindimensionale Heaviside-Funktion gilt die Fouriertransformation

$$\mathcal{F}\mathcal{H}(x) = \frac{1}{2} \left(\delta(\omega) - \frac{i}{\pi\omega} \right), \quad (27)$$

wobei i die imaginäre Einheit mit der Eigenschaft $i^2 = -1$ repräsentiert. Somit ergeben die Gleichungen 26 und 27, in 19 eingesetzt, die Fouriertransformation von $\mathcal{F}I(x, t) = \mathcal{A}(\omega, \nu)$, in Abhängigkeit der Ortsfrequenz ω und der Zeitfrequenz ν ,

$$\mathcal{A}(\omega, \nu) = \frac{1}{2} \mathcal{F}I(\omega) \times \delta\left(\omega + \frac{\nu}{\nu}\right) * \left(\delta(\omega) - \frac{i}{\pi\omega}\right). \quad (28)$$

255 Zur Verdeutlichung: dies ist die Fouriertransformation für einen uniformen, statischen Verdeckter oder Aufdecker und eine kontinuierlich bewegte Textur mit \times als punktweise Multiplikationssymbolisierung und $*$ als Faltung.

3.2 Der eigentliche Detektor

Nach Korn & von Seelen (1972) verläuft das Frequenzband einer Bewegung innerhalb eines ω - ν -Systems entlang einer Geraden, die von ν abhängig ist. Dies beschreibt den in 26 eingeführten Term $\delta(\omega + \nu/\nu)$. An diesem Band kann man die einzelnen Orts- und Zeitfrequenzen der Bewegung ablesen. Das Verhältnis von ω und ν ergibt dabei Rückschluss auf die Geschwindigkeit mit $\nu = -\omega/\nu$. Für eine positive Geschwindigkeit, also in diesem Fall rechtsläufige Bewegung und somit Aufdeckung, wird eine negative Orts- oder Zeitfrequenz benötigt. Diese ört- und zeitliche Abhängigkeit und damit die Geschwindigkeit, konnten Adelson & Bergen (1985) durch Abtasten der Orts- und Zeitfrequenz mithilfe von Gabor-Filtern abschätzen.

Betrachtet man den letzten Teil von Gleichung 28, also die in 27 eingeführte fouriertransformierte Heaviside-Funktion, fällt ein wichtiger Zusammenhang auf. Wenn $\omega \rightarrow 0$ strebt, entsteht ein schlagartiger Vorzeichenwechsel, da dieser Teil für $\omega = 0$ nicht definiert ist. Es folgen also

$$\lim_{\omega \rightarrow 0^-} \left(\delta(\omega) - \frac{i}{\pi\omega} \right) = \infty + i\infty \quad (29)$$

270 und

$$\lim_{\omega \rightarrow 0^+} \left(\delta(\omega) - \frac{i}{\pi\omega} \right) = \infty - i\infty. \quad (30)$$

Durch die Multiplikation der Bewegung mit der Heaviside-Funktion im bildgebenden Raum ergibt sich eine Faltung im Frequenzraum. Durch diese Faltung findet dieser Sprung ins konjugiert Komplexe entlang des Frequenzbandes der Bewegung statt. In Figur 5 ist der beschriebene Zusammenhang von

$$\delta(\omega + \nu/\nu) * \mathcal{FH} \quad (31)$$

dargestellt ohne Berücksichtigung der Bildfunktion I . Es ist die Formel 31, also der Funktionswert der δ -Verteilung nach Faltung mit der Heaviside-Funktion in Abhängigkeit von ω und ν aufgetragen. Dargestellt ist außerdem als Linie die Achse des Frequenzbandes $\omega = \nu/-\nu$, mit der Geschwindigkeit $\nu = -1$. Zu sehen ist der beschriebene Vorzeichensprung im Frequenzspektrum entlang des Frequenzbandes. Immer wenn $\delta(\omega + \nu/\nu)$ eine Impulsantwort gibt, erzeugt die Heaviside-Funktion einen Sprung. Dies passiert genau dann, wenn $\omega + \nu/\nu = 0$ ist. Da es sich bei der Delta-Distribution um eine Verteilung handelt, reicht es für eine Impulsantwort der Null ganz nah zu kommen. Dies ist wichtig, da $\mathcal{FH}(0)$ nicht definiert und somit nicht stetig ist.

Da bei einer einfachen Bewegung ohne Verdeckung/Aufdeckung der letzte Teil, die Heaviside-Funktion, nicht vorhanden ist, sollte dieser Sprung ebenfalls nicht vorhanden sein. Um eine Verdeckung/Aufdeckung zuverlässig detektieren zu können, muss man folglich die An- bzw. Abwesenheit dieses Sprunges untersuchen. Ist der Sprung vorhanden, ist auch eine Verdeckung oder Aufdeckung vorhanden. Sollte dies nicht der Fall sein, handelt es sich um eine einfache Bewegung.

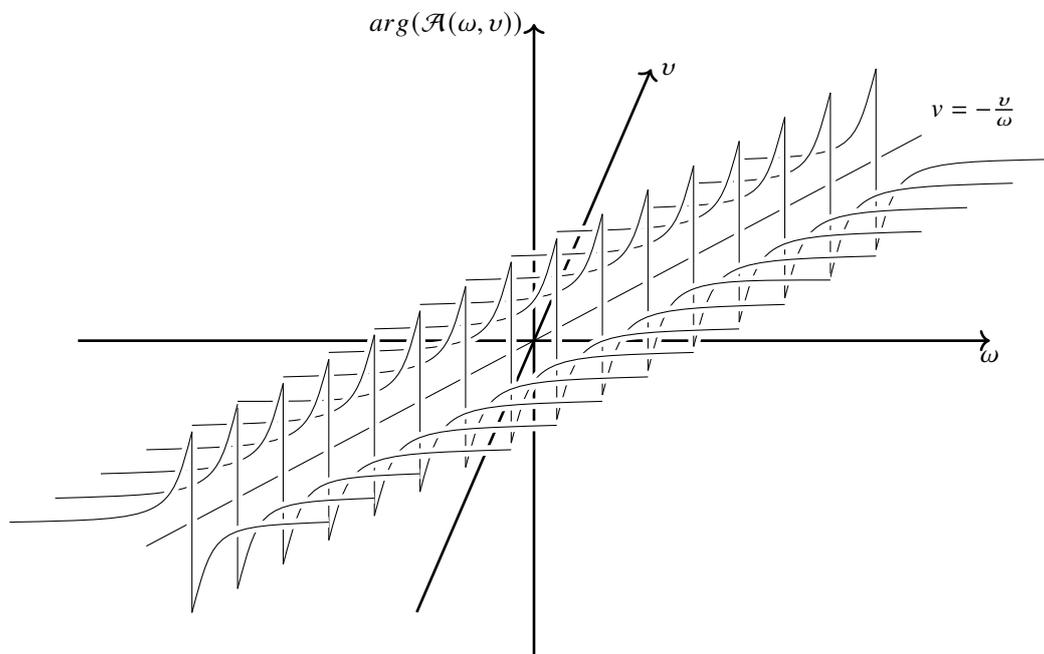


Abbildung 5: Schematische Darstellung des Sprunges ins konjugiert Komplexe. Dargestellt ist der Funktionswert der δ -Funktion nach Faltung mit der Heaviside-Funktion entlang des Frequenzbandes einer Bewegung mit $\nu = -1$ ohne Berücksichtigung einer Bildfunktion. (Eigene Darstellung)

4 Untersuchungen in MATLAB

Alle Berechnungen und Untersuchungen, die im Folgenden unternommen werden, wurden mit und in MATLAB (2020) durchgeführt. MATLAB ist eine Rechensoftware, mit der sich mathematische Probleme berechnen und grafisch darstellen lassen. Es ist hauptsächlich auf das numerische Berechnen von Matrizen ausgelegt. Da Bilder im Allgemeinen als große Matrizen dargestellt werden können, bei denen jeder Wert einen Farbwert widerspiegelt, eignet es sich hervorragend zur Untersuchung dieser. Der verwendete Code mit allen Berechnungen findet sich im Anhang und als ausführbares Skript auf einer beigefügten CD.

4.1 Fouriertransformieren in MATLAB

Um Evidenzen für die Theorie, die in Abschnitt 3 hergeleitet wurde zu finden, wurde der Stimulus aus Abbildung 4c in MATLAB fouriertransformiert und untersucht. Wird in MATLAB ein Bild - in diesem Fall ein x - t -Bild - mit der Fouriertransformation transformiert, entsteht ein Array derselben Größe. Im Falle eines zweidimensionalen Bildes können die Achsen des Arrays als Frequenzen in x - und y -Richtung interpretiert werden, also zwei räumliche Frequenzen, zum Beispiel bei einer zweidimensionalen Sinuskurve die Schwingung entlang der jeweiligen Achsen. Jedoch ist im vorliegenden Fall eine dieser Frequenzen als zeitlich zu interpretieren (mit ν symbolisiert), die andere Frequenz als räumlich (mit ω).

Um etwaige Sprünge entlang der umgrenzenden Kanten dieses Bildes zu vermeiden, muss das Bild mit einer 'Gaußschen-Glockenebene' multipliziert werden, sodass die Sprünge an den Rändern nicht miteinbezogen werden. Durch die Gaußglocke werden Werte, je weiter sie vom Zentrum entfernt sind, immer kleiner und somit entsteht ein Fokus auf das Zentrum des Bildes. Dies ist nötig, da die Fouriertransformation in MATLAB für Punkte außerhalb des Bildausschnitts den Wert null annimmt. Bei dieser Transformierung in MATLAB wird die Null-Frequenz-Komponente auch an den Anfang des Arrays gelegt. Aufgrund der Symmetrie der Fouriertransformation ins konjugiert Komplexe, sollte diese Komponente jedoch in der Mitte liegen. Daher wurde die Transformation umarrangiert, indem der erste mit dem dritten und der zweite mit dem vierten Quadranten vertauscht wurden. Das Ergebnis ist eine Matrix komplexer Zahlen, mit den Null-Frequenz-Komponenten genau in der Mitte. Der Imaginär- bzw. Sinusanteil dieser komplexen Zahlen entspricht der Phase der Frequenz an der jeweiligen Stelle. Der Real- bzw. Kosinusanteil entspricht hingegen der Amplitude der Frequenz an der jeweiligen Stelle.

4.2 Auswertungen in MATLAB

Um die Frequenzen und deren Vorzeichen zu visualisieren, können diese beiden Informationen über Phasen und Amplituden miteinander multipliziert werden. Dadurch geht zwar die Information der genauen Phase verloren, für einen ersten Eindruck ist es jedoch nur wichtig, ob sich das Vorzeichen bei vorhandenen Frequenzen ändert. In Abbildung 6 ist die beschriebene Visualisierung für Bewegung mit (Abb. 6a) und ohne Verdeckung (Abb. 6b) in Abhängigkeit von ω und ν dargestellt. An roten Stellen ist das Argument sehr groß (positiv) und an blauen sehr klein (negativ). In der großen Abbildung 6a ist eine dynamische Verdeckung, in der kleineren 6b links unten eine einfache Bewegung dargestellt. Es wurden nur Werte dargestellt, die größer als 10 % des größten Wertes waren,

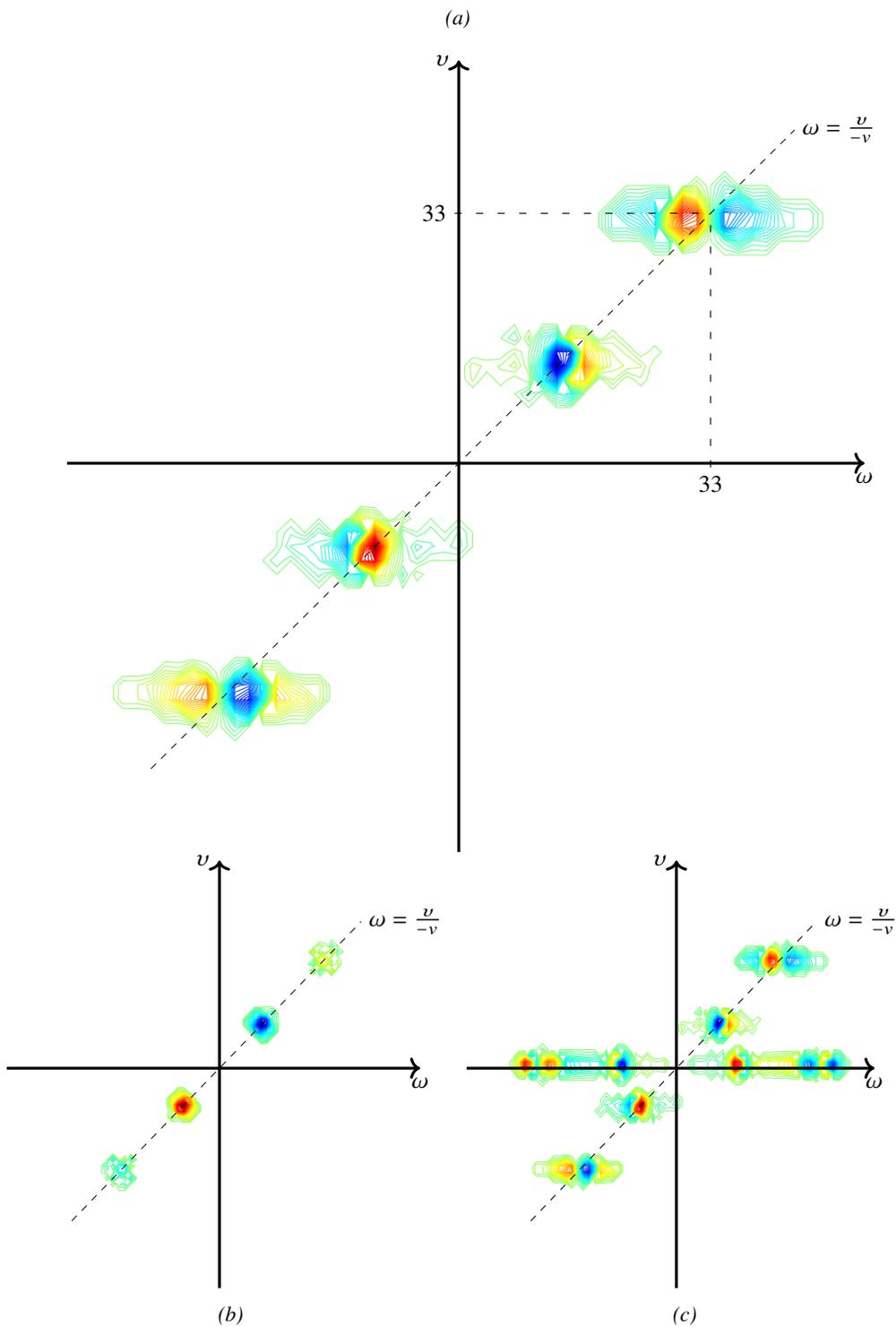


Abbildung 6: Darstellung der Fouriertransformation. Zu sehen ist die Multiplikation von Phase und Amplitude dreier fouriertransformierten Stimuli in Abhängigkeit zu räumlicher (ω) und zeitlicher (v) Frequenz. Werte $< 10\%$ der maximalen Power sind entfernt. Rot-Gelb stellt positive Werte dar, Blau-Grün negative. Die gestrichelte Linie stellt die Frequenzachse für die Bewegung dar. **6a** zeigt den beschriebenen Zusammenhang für dynamische Verdeckung. Außerdem sind passende Orts- und Zeitfrequenz markiert, bei denen ein Gabor-Filter anschlagen würde. **6b** zeigt selbigen Zusammenhang für eine einfache Bewegung. **6c** zusätzlich mit Textur auf dem Verdeckter. (Eigene Darstellung)

damit unnötiges Rauschen nicht zu sehen ist. Da das zugrundeliegende Sinus-Kosinus Rauschen aus mehreren Sinus und Kosinus zusammengesetzt ist, kann man auch mehrere starke Ausschläge erkennen, einen für jede Frequenz des Rauschens. Für ein natürliches Bild oder ein Zufallsmuster wären überall entlang des eingezeichneten Frequenzbandes Ausschläge zu registrieren. Wie erhofft kann bei der Verdeckung ein Vorzeichensprung entlang der Frequenzachse festgestellt werden, welcher bei der reinen Bewegung nicht vorhanden ist. Dies bestätigt die Idee, dass sich Verdeckung - und somit auch Aufdeckung - durch Analyse der Fouriertransformation feststellen lässt. Um nun einen Detektor zu erstellen, muss genau dieser Phasensprung detektiert werden. Dies muss jedoch mithilfe einer für Neuronen möglichen Struktur geschehen. Also suchen wir nach einem rezeptiven Feld, das diesen Phasensprung detektieren kann. Wie man von der Fouriertransformation zu einem passenden rezeptiven Feld in Form eines Gabor-Filters gelangt, wird vorerst am Beispiel der reinen Bewegungsdetektion gezeigt.

4.3 Bewegungsdetektion in MATLAB

Durch Analyse der Fouriertransformation kann genau festgestellt werden, welche Frequenzen vertreten sind. Diese kann dann genutzt werden, um selektive Gabor-Filter zu erstellen, die nur bei der passenden Frequenz ‚anspringen‘. Im Beispiel der reinen Bewegungsdetektion kann aus 6b die genaue Orts- und Zeitfrequenz der Bewegung abgelesen werden. Für den passenden Gabor-Filter werden nun zwei Wellenebenen, eine ungerade (Sinus) und gerade (Kosinus) erstellt,

$$\sin(\omega \cdot x + \nu \cdot t) \tag{32}$$

$$\cos(\omega \cdot x + \nu \cdot t), \tag{33}$$

mit Ortsfrequenz ω und Zeitfrequenz ν , welche dann anschließend mit einer Gaußglocke multipliziert wird. Das Resultat sind zwei räumlich-zeitlich selektive Gabor-Funktionen. Faltet man nun diese Gabor-Funktionen mit dem ursprünglichen Stimulus, wird die passende Bewegung herausgefiltert. Ein Problem entsteht, wenn man eine Energie dieser Bewegung bestimmen möchte, da der Filter während einer Bewegung immer wieder zwischen positiven und negativen Ausgaben oszilliert. Um nun eine Energie bestimmen zu können, kann man die beiden Ausgaben des geraden und ungeraden Filters quadrieren und anschließend aufsummieren (Adelson & Bergen, 1985). Dies macht sich die Eigenschaft

$$\sin(x)^2 + \cos(x)^2 = 1 \tag{34}$$

zu Nutze. Die Energie der Bewegung kann man auch im Frequenzbereich, also in Abhängigkeit von ω und ν , berechnen, da die Energie eines Signals gleich seiner Energie im Frequenzbereich ist. Beschrieben wird dies durch den Satz von Parseval, der sich durch die Formel

$$\int_{-\infty}^{\infty} |I(x, t)|^2 dx dt = \int_{-\infty}^{\infty} |FI(\omega, \nu)|^2 d\omega d\nu \tag{35}$$

ausdrückt. Deshalb kann die Energie über die quadrierten Ausgaben der beiden Filter bestimmt werden (James, 2011).

Bei umgekehrter Bewegung oder einer anderen Geschwindigkeit würde der Filter kaum oder gar nicht reagieren. Um diese Bewegung ebenfalls detektieren zu können, müssten weitere Gabor-Filter

erstellt werden, die für andere Frequenzen selektiv wären. Dadurch kann man eine Gabor-Filterbank
 360 erstellen, die den gesamten möglichen Frequenzbereich abtastet (Adelson & Bergen, 1985).

4.4 Verdeckungsdetektion in MATLAB

Bei genauer Untersuchung des Frequenzbereichs wird sichtbar, dass auch sehr ähnliche Frequenzen
 immer noch eine große Energie bei einem selektiven Gabor-Filter haben. Ein Filter, der also nahe der
 eigentlichen Frequenz sitzt, erzeugt dementsprechend ebenfalls eine starke Ausgabe. Im vorliegenden
 365 Fall entspricht dies einem Filter, der sich nur mit sehr geringem $\Delta\omega$ und $\Delta\nu$ vom eigentlichen Filter
 unterscheidet, und ebenfalls eine starke Reaktion auf den Stimulus erzeugt. Spiegelt man diesen Filter
 entlang der Frequenzachse mit $\omega + \nu/\nu$, hat dieser bei einer gleichförmigen Bewegung wie in 6b zu
 sehen, die gleiche Energie. Ein Filterpaar ist aus jeweils einem ungeraden und einem geraden Filter

$$\sin_{+/-} = \sin((\omega \pm \Delta\omega \cdot -\nu) \cdot x + (\nu \mp \Delta\nu) \cdot t) \quad (36)$$

$$\cos_{+/-} = \cos((\omega \pm \Delta\omega \cdot -\nu) \cdot x + (\nu \mp \Delta\nu) \cdot t) \quad (37)$$

zusammengesetzt. Damit die Differenz der Filter auch für andere Geschwindigkeit weiterhin orthogo-
 370 nal zum Frequenzbande liegt, wird $\Delta\omega \cdot -\nu$ hinzugefügt. Wird dieses Filterpaar dann mit dem Stimulus
 S_B einer gleichförmigen Bewegung ohne Verdeckung gefaltet (*), ergibt sich also dieselbe Energie
 für beide Filter des Filterpaares:

$$(S_B * \sin_+)^2 + (S_B * \cos_+)^2 = (S_B * \sin_-)^2 + (S_B * \cos_-)^2. \quad (38)$$

Für eine gleichförmige Bewegung sollte also gelten:

$$(S_B * \sin_+)^2 + (S_B * \cos_+)^2 - (S_B * \sin_-)^2 + (S_B * \cos_-)^2 = 0. \quad (39)$$

Bei einer Bewegung mit Verdeckung S_V gilt dieser Zusammenhang, aufgrund des Vorzeichenwech-
 375 sels der Phase genau an der Frequenzachse, nicht mehr. Wie in 6a zu sehen ist, sollte die Annahme

$$(S_V * \sin_+)^2 + (S_V * \cos_+)^2 = ((S_V * \sin_-)^2 + (S_V * \cos_-)^2) \cdot -1 \quad (40)$$

und somit

$$(S_V * \sin_+)^2 + (S_V * \cos_+)^2 - (S_V * \sin_-)^2 + (S_V * \cos_-)^2 \neq 0 \quad (41)$$

also halten. Um nun die Anwesenheit einer Verdeckung oder Aufdeckung für eine bestimmte Orts-
 Zeitfrequenz zu untersuchen, genügt es zwei leicht abweichende symmetrische Filter, einen positiven
 und einen negativen, zu erstellen und nach der Faltung mit dem Stimulus deren Energien zu sub-
 380 trahieren. Bei einem Ergebnis nahe Null ist keine Verdeckung/Aufdeckung vorhanden, bei einem
 anderen Ergebnis schon. Dies liegt daran, dass sich bei reiner Bewegung die Energien der beiden
 Filter durch die Subtraktion gegenseitig aufheben; bei Bewegung mit Verdeckung oder Aufdeckung
 hingegen aufaddieren.

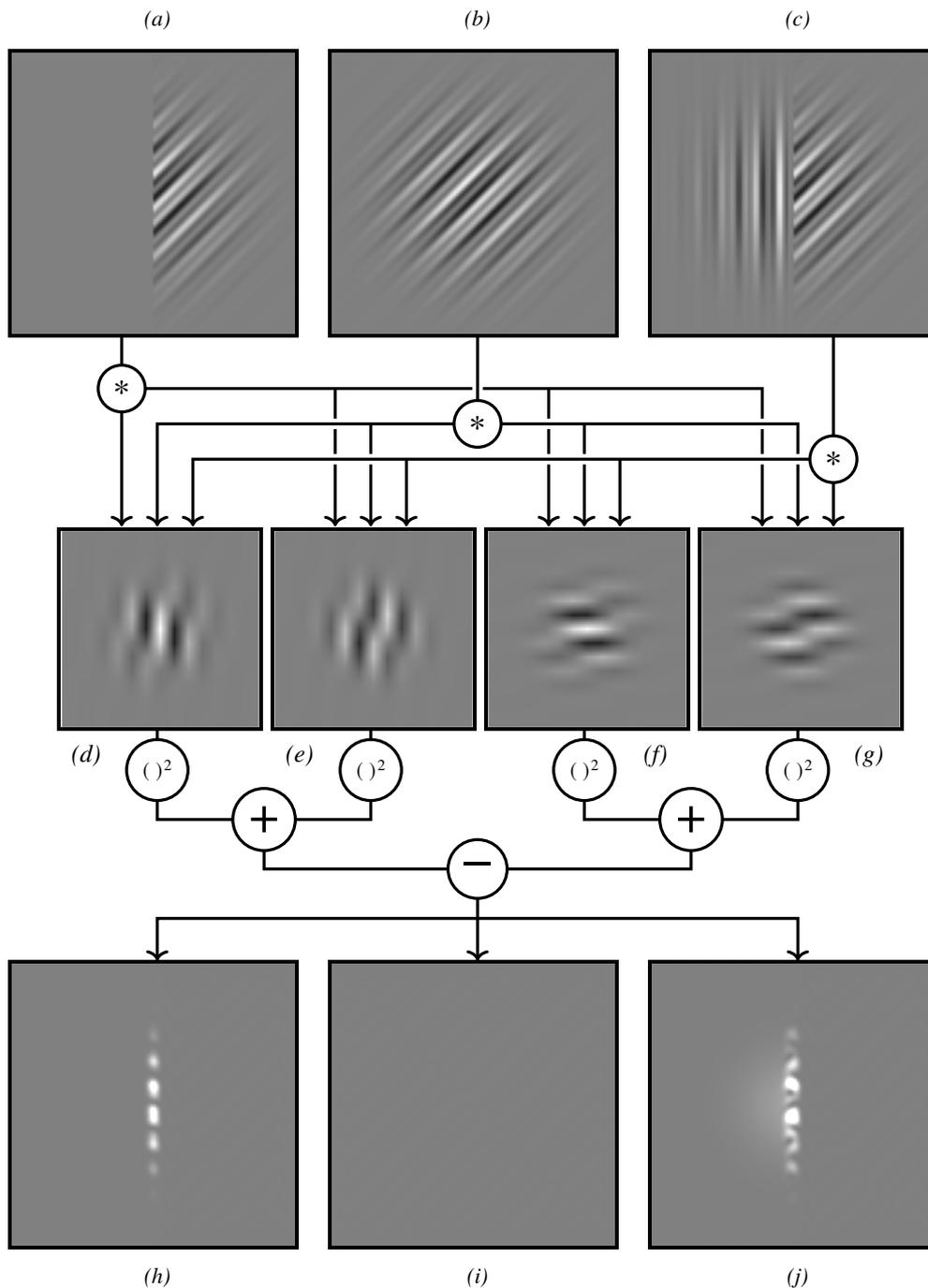


Abbildung 7: Durchlaufen eines Filterpaares. Die drei präsentierten Stimuli 7a-7c werden jeweils mit den Sinus- und Kosinusfiltern des Filterpaares 7d-7g gefaltet. Anschließend quadriert, aufsummiert und die Paare subtrahiert (dargestellt in 7h-7j). 7a zeigt den Verdeckter-Stimulus aus 4c mit einer Gaußglocke multipliziert. 7b zeigt denselben Stimulus ohne Verdeckung. 7c zeigt ebenfalls denselben Stimulus, diesmal mit einem texturierten Verdeckter. 7d zeigt den ersten Kosinus basierten Gabor-Filter cos_+ und 7e den zugehörigen Sinus basierten sin_+ . 7f und 7g zeigen analog die Gabor-Filter cos_- und sin_- . 7h zeigt das Ergebnis nach Subtraktion der beiden Filter des Filterpaares, die wiederum durch Quadrierung und Addition der mit dem Stimulus aus 7a gefalteten Gabor-Filter entstanden sind. 7i zeigt denselben Zusammenhang auf die reine Bewegung ohne Verdeckung bezogen. 7j zeigt ebenfalls denselben Zusammenhang, diesmal mit Bezug auf die Verdeckung durch eine texturierte Oberfläche. (Eigene Darstellung)

4.5 Überprüfung der Ergebnisse

385 Zum Überprüfen dieses Zusammenhangs auf Tauglichkeit wurden zwei Stimuli erstellt. Beide mit derselben Geschwindigkeit $v = -1$ und demselben zugrunde liegenden Sinusrauschen. Ein Stimulus wurde außerdem zusätzlich mit einem Verdeckter versehen. Somit hat man zwei gleiche Stimuli, die sich lediglich durch An- bzw. Abwesenheit eines Verdeckers unterscheiden. Dargestellt sind diese in Abbildung 7a und 7b. Genau unterhalb sind die vier Gabor-Filter, die für das Filterpaar benötigt
390 werden, dargestellt. Jeweils ein Kosinus (Fig. 7d und 7f) und ein Sinus (Fig. 7e und 7g) für jeden Filter. Die beiden Stimuli wurden daraufhin jeweils mit allen Filtern gefaltet und die jeweils zusammengehörigen Paare von einem ungeraden und einem geraden Filter quadriert und aufsummiert. Die daraus entstandenen Energien wurden dann durch Subtraktion zu einer Gesamtenergie der Verdeckung zusammengefügt. Diese Gesamtenergie ist für den verdeckten Stimulus in Abbildung 7h
395 und ohne Verdeckung in 7i visualisiert. Man sieht, dass es bei dem Stimulus ohne Verdeckter keine Reaktion gibt, wohingegen mit Verdeckung eine Reaktion entlang der Verdeckungskante zu erkennen ist. Dies legt nahe, dass es möglich ist, über die Differenz zweier nahe angrenzender Gabor-Filter eine Verdeckung und somit auch Aufdeckung, zu detektieren.

Dieselbe Überprüfung wurde auch mit einem texturierten Verdeckter (Abbildung 7c) getätigt. Eine
400 statische Textur auf dem Verdeckter verändert sich nicht über die Zeit hinweg. Es sollte also auch keine Veränderung entlang des Frequenzbandes der Bewegung geben, sondern lediglich eine zusätzliche Komponente entlang der Ortsfrequenzachse, ohne Abhängigkeit zur Zeit. Betrachtet man wieder die Multiplikation von Phase und Amplitude, sieht man diese Annahme bestätigt (vgl. 6c). Da es keine Änderung entlang des Frequenzbandes der Bewegung gibt, sollte das Hinzufügen der Textur auch
405 keinen Einfluss auf den Detektor haben. Das Durchlaufen der Filter mit diesem Stimulus ist wie bei den vorherigen beiden Stimuli ebenfalls in der letzten Zeile der Abbildung 7, in 7j, zu sehen. Man sieht, dass es nicht zu merklichen Abweichungen im Vergleich zur einfachen Verdeckung kommt. Dementsprechend stabil ist der Detektor auch für einen texturierten Verdeckter.

5 Erstellung einer Filterbank

410 Im vorherigen Abschnitt wurde ein Filterpaar für eine spezifische Orts- und Zeitfrequenz erstellt. Dies entspricht jedoch nicht wirklich einer natürlichen Herangehensweise, da es in der Natur unendlich viele Kombinationen von Frequenzen gibt und es somit unmöglich ist alle abzudecken. Die Frequenz-spezifischen Filter haben jedoch eine gewisse Umgebung, bei der sie noch ausreichend stark reagieren. Man kann also Filtercluster erstellen, die einen Bereich relativ flächendeckend abdecken.

415 Daher sollte eine Umgebung geschaffen werden, die kontinuierlich über alle Kombinationen von Frequenzen und somit Geschwindigkeiten und Richtungen der Verdeckung/Aufdeckung, zuverlässig reagiert. In Figur 8 ist eine mögliche Anordnung für 10 Filterpaare dargestellt. Entlang der x -Achse ist wieder der räumliche, entlang der y -Achse der zeitliche Frequenzbereich dargestellt. Jeweils zwei Filter gehören zu einem Paar und sind um dieselben $\Delta\omega$ und $\Delta\nu$ in positive und negative Richtung verschoben. Je weiter die Paare vom Null-Frequenzbereich entfernt sind, desto größer wurden die Δ 's gewählt. Die gestrichelte Linie stellt annäherungsweise das sehbbare Spektrum dar (vgl. Adelson & Bergen, 1985 und Watson, Ahumada, & Farrell, 1986). Die Rautenförmigkeit reflektiert die arithmetische Wechselbeziehung zwischen räumlichen und zeitlichen Frequenzen. Wird die zeitliche Frequenz erhöht, muss die räumliche Frequenz zwangsläufig vergleichbar reduziert werden, damit der Stimulus sichtbar bleibt (Robson, 1966 und Watson et al., 1986). Die Filter im oberen Quadranten benötigen eine negative Geschwindigkeit, also linksläufige Bewegung und somit Verdeckung. Im unteren Quadranten ist eine positive Geschwindigkeit und somit rechtsläufige Bewegung notwendig, die eine Aufdeckung zur Folge hat. Dies gilt aber natürlich nur in dem hier verwendeten Aufbau.

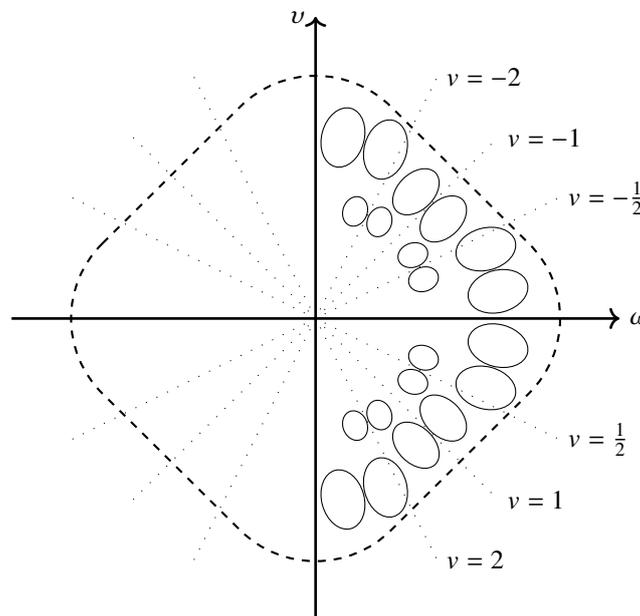


Abbildung 8: Zu sehen ist die schematische Anordnung der Filterpaare zu einer Filterbank in Abhängigkeit von ω und ν . Der erste Quadrant ist für linksgerichtete Bewegung, der vierte Quadrant für rechtsgerichtete Bewegung. Außerdem ist die Grenze des sichtbaren Spektrums eingezeichnet und schematisch die Frequenzachsen verschiedener Bewegungen. (Eigene Darstellung in Anlehnung an Adelson & Bergen, 1985)

430 Die Wahl der Positionen, die Größe der Differenzen und die Größe der Filter selbst wurden so gewählt, dass möglichst eine flächendeckende Abtastung erfolgt. Jede Geschwindigkeit soll möglichst in einem Filterpaar detektiert werden, ohne dass unendlich viele dieser Paare implementiert werden müssen. Gleichzeitig muss aber ebenfalls gewährleistet sein, dass es nicht zu Fehlreaktionen kommt, beispielsweise bei reiner Bewegung ohne Verdeckung. Dies könnte zum Beispiel dann passieren, 435 wenn der Stimulus einer reinen Bewegung einen Filter nur ganz am Rande streift und durch die Differenz mit dem zugehörigen Filterpartner eine Verdeckung mimt. Dies sollte durch geschickte Wahl der Parameter vermieden werden, indem darauf geachtet wird, dass die Filterumgebung bei der die Filter noch ausreichend stark reagieren, nicht zu groß ist. Außerdem ist eine Änderung bei der Berechnung der Gesamtenergie nötig, die aber im weiteren Verlauf näher erklärt wird. Bei den vorlie- 440 genden Annahmen bezüglich der Form und Position der Filterpaare, wurden jedoch keine genaueren Berechnungen angestellt, sondern es wurden lediglich die Häufigkeiten und Beschaffenheiten von in der Natur vorkommenden Reizen angeschaut, um eine vorerst geeignete Lösung zu finden.

5.1 Energieberechnung

Für die Berechnung eines falsifizierbaren Ergebnisses wurden, wie vorher in Abschnitt 4.4 gezeigt, 445 die Energien der einzelnen Filter der Filterpaare berechnet. Durch die möglichen Fehlreaktionen, die entstehen können, wenn einer der Filter zufällig genau auf der Frequenzachse sitzt, kann es sein, dass es bei einem Filterpaar dazu kommt, dass einer dieser Filter eine starke Energie hat, der andere aber keine oder eine sehr geringe. Dadurch kann es passieren, dass eine einfache Bewegung als Verdeckung wahrgenommen wird. Um dies zu vermeiden werden die Energien der zusammengehörigen Paare 450 in analytischen Zusammenhang gebracht. Wenn E_+ die Energie des einen Filters und E_- die des anderen ist, wird die Gesamtstärke der Energie \mathcal{E}_f des Filters berechnet mit

$$\mathcal{E}_f = (E_+ \times E_-^2) - (E_- \times E_+^2). \quad (42)$$

Dadurch werden weiterhin die gewünschten Verhältnisse beibehalten, um die Verdeckung oder Auf- 455 deckung zu detektieren, aber es werden ebenfalls die Fälle herausgenommen, bei denen einer der Filter eine sehr geringe Energie hat. Durch die Multiplikationen wird immer vorausgesetzt, dass beide Filter eine ausreichende Energie haben. Durch die Quadrierung verschwindet der Einfluss, den das Vorzeichen der Energie durch das Multiplizieren haben würde, sodass nichts an der An- bzw. Abwesenheit des Phasensprungs geändert wird. Für die endgültige Ausgabe der Filterbank \mathcal{E} werden nun einfach alle Energien der einzelnen Paare aufsummiert.

$$\mathcal{E} = \sum_i \mathcal{E}_{f_i} \quad (43)$$

wobei \mathcal{E}_{f_i} die Einzelenergien der zehn Filterpaare sind. Reine Bewegung erzeugt eine sehr geringe 460 Energie, da sich die Filter der einzelnen Paare gegenseitig auflösen, eine Verdeckung oder Aufdeckung hingegen eine wesentlich größere, da sich die Filter innerhalb der Paare aufsummieren.

Tauscht man die Seiten des Stimulus und des Verdeckers/Aufdeckers, sollte die vorgelegte Filterbank trotzdem funktionieren. Es drehen sich die Verhältnisse von Verdeckung und Aufdeckung; durch die Symmetrie in der Fouriertransformation ins konjugiert Komplexe, schlagen die Filter aber an 465 denselben Stellen an. Der daraus entstehende Vorzeichenwechsel macht aber durch die Quadrierung

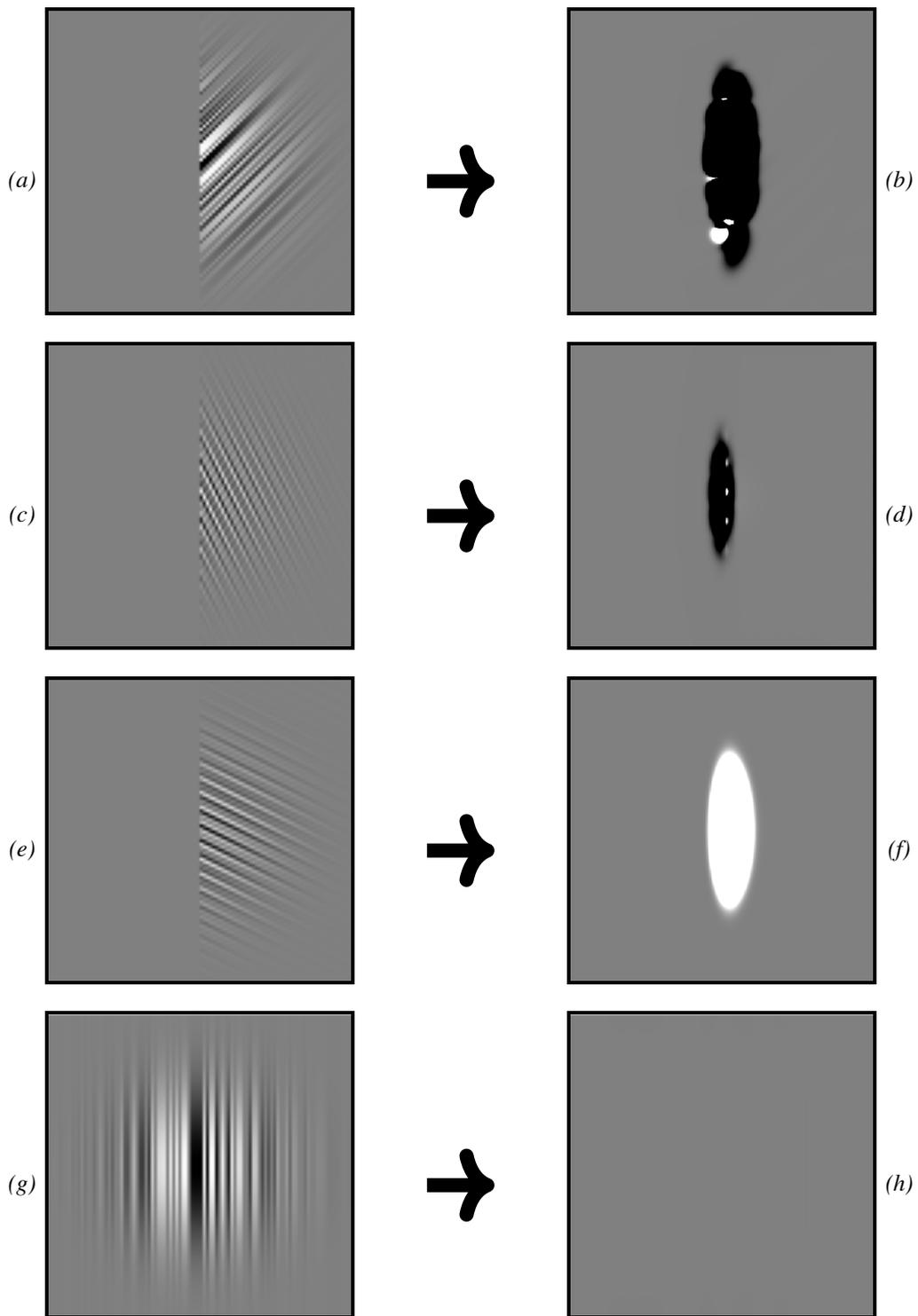


Abbildung 9: Darstellung eines Stimulus mit Ausgabe der Filterbank. **9a** zufälliges schwarzweißes Rauschen das verdeckt wird mit zugehöriger Ausgabe **9b** nach Durchlaufen der Filterbank. **9c** Sinus- Kosinus Rauschen mit Aufdeckung und erhöhter Geschwindigkeit, zugehörige Ausgabe in **9d**. **9e** ebenfalls Aufdeckung aber mit verminderter Geschwindigkeit und Ausgabe **9f**. **9g** stehendes schwarzweiß Rauschen ohne starke Ausgabe **9h**. Alle Bilder sind in MATLAB berechnet und nicht angepasst worden. (Eigene Darstellung)

beim Berechnen der Energie keinen Unterschied. Aufgrund dieser Symmetrie muss auch nicht das gesamte Fourierspektrum abgedeckt werden, sondern lediglich die Hälfte. Die hier erstellte Filterbank sollte also für die meisten Stimuli funktionieren, wenn gleich definitiv noch Verbesserungen und Abänderungen erfolgen müssen, um eine optimal arbeitende Filterbank zu erhalten.

470 **5.2 Weitere Stimuli**

In Figur 9 sind vier weitere Stimuli präsentiert, die durch die vorgestellte Filterbank gelaufen sind. Links in der Grafik sind immer die Stimuli zu sehen; auf derselben Höhe der rechten Seite die zugehörigen Ausgaben der Filterbank. Zuerst wird gezeigt, dass die Filterbank auch bei einfachem Rauschen funktioniert. Dann folgen zwei Beispiele mit anderen Geschwindigkeiten und Aufdeckung
475 anstelle von Verdeckung und schließlich der Sonderfall, dass sich gar nichts bewegt und es dadurch keine Verdeckung bzw. Aufdeckung gibt. Die schwarzen Energien sind negativ. Das heißt in diesem Fall, dass der Phasensprung in der Fouriertransformation umgekehrt ist. Zu interpretieren ist dies jedoch nicht anders als eine positive Energie. Auch hier sind alle Stimuli wieder als Matrix in
480 MATLAB implementiert worden und ohne jegliche Änderungen durch ein und dieselbe Filterbank gelaufen. Der verwendete MATLAB-Code für den einzelnen Detektor, als auch für die Filterbank, sind im Anhang zu finden und auch als ausführbares Skript auf einer beigefügten CD.

6 Zusammenfassung und Ausblick

Im Verlauf dieser Arbeit konnte gezeigt werden, dass es möglich ist, dynamische Verdeckung und Aufdeckung zu detektieren. Beim Vergleich von Bewegung ohne Ver-/Aufdeckung kann man einen Unterschied in der Fouriertransformation sehen. Dieser Phasensprung, der durch das Fouriertransformieren der Heaviside-Funktion entsteht, macht es möglich, diesen Unterschied festzustellen. Um ihn in der Natur feststellen zu können, muss ein Detektor diesen Phasensprung entlang der Frequenzachse der Bewegung erfassen und zuverlässig zwischen reiner Bewegung und Ver-/Aufdeckung unterscheiden. Hierfür wurden Gabor-Filterpaare erstellt, die eine hohe Energie bei der Reaktion auf Verdeckung/Aufdeckung aufweisen und eine niedrige bei reiner Bewegung. Dies wurde durch Differenzieren nahe benachbarter Frequenzen möglich, die bei Verdeckung oder Aufdeckung unterschiedliche Vorzeichen aufweisen und bei reiner Bewegung gleiche. Im Anschluss wurden mehrere dieser selektiven Filterpaare für verschiedene Frequenzen zu einer Filterbank kombiniert. Da sich dies nur mithilfe von Gabor-Filtern bewerkstelligen ließ, kann man annehmen, dass sich die Erkennung des Effekts der Verdeckung und Aufdeckung durch dynamische Bewegung der frühen visuellen Wahrnehmung zuordnen lässt. Die Filterbank kann zuverlässig für einen eindimensionalen, von der Zeit abhängigen Bildausschnitt feststellen, ob im präsentierten Stimulus eine Verdeckung oder Aufdeckung existiert.

Rückblickend könnte man einige Dinge kritisch hinterfragen. So ist zum Beispiel die Position der Filter innerhalb der Filterbank recht frei gewählt und man sollte definitiv überprüfen, ob und inwieweit diese Anordnung Sinn ergibt. So kam es beim Erstellen der Bank schon zu gewissen Ungereimtheiten bei einigen Kombinationen von Geschwindigkeiten und Sinus-Rauschen. Viele ließen sich zwar auf die Unnatürlichkeit der Kombinationen - zum Beispiel sehr hohe Geschwindigkeiten bei gleichzeitiger hoher Frequenz des Stimulus - zurückführen, sollten jedoch nicht ganz außer Acht gelassen werden. Zudem ergab sich ein Problem, wenn man den Verdecker nicht in die Mitte der Amplitude - Verdecker war 0 bei Rauschen das von -1 bis 1 reichte - setzte, sondern beispielsweise an einen Extrempunkt - Verdecker 0 bei Rauschen von 0 bis 1. Hierfür gab es extreme Werte in der Fouriertransformation an der Null-Frequenz-Komponente. Dies lässt sich vermutlich darauf zurückführen, dass fast jeder Wert im Rauschen somit extrem weit von diesem Extrempunkt entfernt ist. Deshalb ist an der Kante, und somit an der Orts-Zeitlichen-Null-Frequenz, ein extremer Sprung, der zu einer großen Power in der Fouriertransformation führt. In dieser Arbeit wurde daher der Verdecker in die Mitte der Rauschamplitude gelegt. Alternativ könnte man auch die Fouriertransformation ohne die beschriebenen Frequenzen betrachten. Aber dies ist nur eine Annahme und sollte ebenfalls noch genauer untersucht und besser hergeleitet werden. Daher sollte man die Filterbank noch einmal an natürlichen Stimuli testen und das künstliche Sinus-Rauschen durch natürliche Bildausschnitte ersetzen.

Außerdem sollte überprüft werden, ob die vorliegende Filterbank mit ihren zehn Filterpaaren ausreicht, um zuverlässig das gesamte visuelle Spektrum abzudecken. Auch die Wahl der Δ 's, also die Abdeckungsbreite der einzelnen Paare, bietet viel Spielraum für Verbesserungen und Erweiterungen. Die Ergebnisse sind außerdem nur eine erste Feststellung und es müssen weitere Untersuchungen getätigt werden. Der verwendete Stimulus ist sehr vereinfacht und sterilisiert worden. Diese Simplifizierungen sollten Stück für Stück rückgängig gemacht und überprüft werden. Eine dieser Vereinfachungen, der einfache schwarze Verdecker, wurde im Verlauf schon rückgängig gemacht und

durch einen texturierten Verdeckter ersetzt. Dies konnte nicht an der Annahme rütteln, und ließ das
525 Ergebnis gänzlich unbeeindruckt (vgl. 4.5). Es kann davon ausgegangen werden, dass dies auch für
andere komplexe Erweiterungen gilt. Dabei ist die wohl wichtigste, aber auch komplexeste Änderung,
die von einem eindimensionalen $x-t$ -Stimulus zu einem zweidimensionalen $x-y-t$ -Stimulus. Hierfür
müsste nämlich auch der Filterbank eine ganze Dimension hinzugefügt werden. Ist dies für einen
einfachen Verdeckter oder Aufdeckter gelungen, sollte man hier auch untersuchen, wie weitere Entde-
530 ckungen sich auf den Effekt und die Filterbank auswirken. So sollte man zum Beispiel prüfen, wie
die Ergebnisse von Tanrikulu et al. (2018) - also Verstärkung durch Konvexität der Verdeckungskante
- sich auf die Fouriertransformation auswirken. Ebenfalls sollte untersucht werden, wie die Fälle der
Eigenverdeckung/-aufdeckung durch Rotation sich auf die Reaktionen der Filter auswirken (Froyen
et al., 2013). Ein weiterer Fall der betrachtet werden müsste, ist die Bewegung beider Texturen. Dabei
535 kann sich nämlich die Verdeckungskante verschieben und somit die Detektion erschwert werden.
Am Ende sollte es möglich sein, einen Stimulus, zum Beispiel einen Film, durch den Filter laufen zu
lassen, sodass alle Kanten die eine Verdeckung oder Aufdeckung erzeugen hervorgehoben werden.
Abschließend kann man feststellen, dass es wahrscheinlich möglich sein sollte, dynamische Aufde-
ckung und Verdeckung als psycho-physiologischen Effekt der frühen visuellen Wahrnehmung durch
540 Gabor-Filter zu detektieren. Dennoch ist dies nur eine erste Untersuchung und es müssen noch einige
weitere unternommen werden, damit sich ein einheitlich klares Bild ergibt.

Literatur

- Adelson, E. H., & Bergen, J. R. (1985). Spatio-temporal energy models for the perception of motion. *Journal of the Optical Society of America*, 2(2), 284-299. doi: 10.1364/JOSAA.2.000284
- Cope, D., Blakeslee, B., & McCourt, M. E. (2009). Simple cell response properties imply receptive field structure: balanced gabor and/or bandlimited field functions. *JOSA A*, 26(9), 2067–2092. doi: 10.1364/JOSAA.26.002067
- Froyen, V., Feldman, J., & Singh, M. (2010). A bayesian framework for figure-ground interpretation. In J. Lafferty, C. Williams, J. Shawe-Taylor, R. Zemel, & A. Culotta (Eds.), *Advances in neural information processing systems* (Vol. 23, pp. 631–639). Curran Associates, Inc.
- Froyen, V., Feldman, J., & Singh, M. (2013, 08). Rotating columns: Relating structure-from-motion, accretion/deletion, and figure/ground. *Journal of Vision*, 13(10), 6-6. doi: 10.1167/13.10.6
- Gibson, J. J., Kaplan, G. A., Reynolds, H. N., & Wheeler, K. (1969). The change from visible to invisible. *Perception & Psychophysics*, 5, 113-116. doi: 10.3758/BF03210533
- Hildreth, E. C., & Royden, C. S. (2011). Integrating multiple cues to depth order at object boundaries. *Attention, Perception, & Psychophysics*, 73, 2218–2235. doi: 10.3758/s13414-011-0172-0
- James, J. F. (2011). *A student's guide to fourier transforms: With applications in physics and engineering* (3rd ed.). Cambridge University Press. doi: 10.1017/CBO9780511762307
- Kanizsa, G., & Gerbino, W. (1976). Convexity and symmetry in figure-ground organization. *Art and Artefacts*. Henle M.
- Kaplan, G. A. (1969). Kinetic disruption of optical texture: The perception of depth at an edge. *Perception & Psychophysics*, 6, 193-198. doi: 10.3758/BF03207015
- Koffka, K. (2013). *Principles of gestalt psychology* (Vol. 44). Routledge. (Original work published 1935)
- Korn, A., & von Seelen, W. (1972). Dynamische Eigenschaften von Nervennetzen im visuellen System. *Kybernetik*, 10. (PMID: 29804494) doi: 10.1007/BF00292232
- Mallot, H. A. (Ed.). (2013). *Computational neuroscience: a first course*. Springer.
- MATLAB. (2020). *version 9.8.0.1359463 (r2020a)*. Natick, Massachusetts: The MathWorks Inc.
- Robson, J. G. (1966, Aug). Spatial and temporal contrast-sensitivity functions of the visual system. *J. Opt. Soc. Am.*, 56(8), 1141–1142. doi: 10.1364/JOSA.56.001141
- Rubin, E. (1958). Figure and ground. *Readings in perception*, 194-203. (Original work published 1915)
- Sekuler, R., Watamaniuk, S. N., & Blake, R. (2002). Motion perception. *Stevens' Handbook of Experimental Psychology*. doi: 10.1002/0471214426.pas0104
- Tanrikulu, O. D., Froyen, V., Feldman, J., & Singh, M. (2018). When is accreting/deleting texture seen as in front? Interpretation of depth from texture motion. *Perception*, 47(7), 694-721. (PMID: 29804494) doi: 10.1177/0301006618776119
- Wagemans, J., Elder, J. H., Kubovy, M., Palmer, S. E., Peterson, M. A., Singh, M., & von der Heydt, R. (2012). A century of gestalt psychology in visual perception: I. perceptual grouping and figure-ground organization. *Psychological bulletin*, 138(6), 1172-1217. doi: 10.1037/a0029333
- Wandell, B., & Thomas, S. (1997). *Foundations of vision* (Vol. 42) (No. 7).
- Watson, A. B., Ahumada, A. J., & Farrell, J. E. (1986, Mar). Window of visibility: a psychophysical theory of fidelity in time-sampled visual motion displays. *J. Opt. Soc. Am. A*, 3(3), 300–307.

doi: 10.1364/JOSAA.3.000300

Yonas, A., Craton, L. G., & Thompson, W. B. (1987). Relative motion: Kinetic information for the order of depth at an edge. *Perception & Psychophysics*, *41*, 53–59. doi: 10.3758/BF03208213

Ziou, D., & Tabbone, S. (1998). Edge detection techniques-an overview. *Pattern Recognition and Image Analysis C/C of Raspoznavaniye Obrazov I Analiz Izobrazhenii*, *8*, 537–559.

Anhang

Anhang 1: SimpleFilter.m

```

1  %% MATLAB CODE ZUR BACHELORARBEIT VON LEON DANKERT
2  % SimpleFilter.m
3  clear
4  close all
5  %% Stimuli Erzeugung
6  v = -1;      % Geschwindigkeit: Minus = Links/Verdeckung, Plus =
   Rechts/Verdeckung
7  t = 100;    % Anzahl der Zeitschritte und Stimulibreite
8
9  A = (sin(-v*[1:t]*0.75 + transpose([1:t])*0.75) ... %Sinus
   Kosinus Rauschen
10     .* cos(-v*[1:t]*0.33 + transpose([1:t])*0.33));
11
12 % Erzeugung des texturlosen Verdeckers
13 % A(:, 1:round(t/2)) = 0 ;
14
15
16 % Erzeugung des Gaussfensters
17 Gausglocke = gausswin(t,3) * gausswin(t,3)';
18 A = A .* Gausglocke;    %Multiplikation mit der Gaussglocke
19
20 %% Fouriertransformation
21 ft = fftshift(fft2(A)); %Transformierung und Shiffting der
   Quadranten
22 ftang = angle(ft);      %Extraktion der Phase
23 ftabs = abs(ft);       %Extraktion der Amplitude
24
25 ftmax = max(max(ftabs));%Maximale Amplitude
26 a = ftabs > 0.1*ftmax; %Weniger als 10 Prozent von Max
27 ftabs10 = ftabs.*a;
28 weightedAng = ftabs10 .* ftang; %Phase x Amplitude
29 contour(weightedAng);  %Einfache Darstellung
30
31 %% Spezifisches Raum-Zeitlich-spezifisches Filterpaar
32 xG = [-7:7];           %Filtergroesse in x
33 tG = [-7:7]';         %Filtergroesse in t
34 omega0 = 33;          %Omega 0 des Filters
35 ny0 = 33;             %Ny 0 des Filter (\uspilon in Text)
36 delta = 1;           %Abstand der Filter vom Frequenzband

```

```
37 | deltav = delta*(omega0/ny0);%Delta * Geschwindigkeit, damit
    | weiterhin gespiegelt
38 | sigma = 1.8;          %SD der Gaussglocke
39 | gg = exp(-(xG.^2+tG.^2)/(2*sigma^2)); %Gaussglocke
40 | % Die Vier Gaborfilter eines Paares
41 | gabCP = gg .* (cos((omega0 + deltav)* xG +(ny0 - delta) * tG)); %
    | Cos_+
42 | gabCM = gg .* (cos((omega0 - deltav)* xG +(ny0 + delta) * tG)); %
    | Cos_-
43 | gabSP = gg .* (sin((omega0 + deltav)* xG +(ny0 - delta) * tG)); %
    | Sin_+
44 | gabSM = gg .* (sin((omega0 - deltav)* xG +(ny0 + delta) * tG)); %
    | Sin_-
45 | %Darstellungen der vier Filter
46 | nexttile; imshow(gabCP,[-1 1]);
47 | nexttile; imshow(gabCM,[-1 1]);
48 | nexttile; imshow(gabSP,[-1 1]);
49 | nexttile; imshow(gabSM,[-1 1]);
50 |
51 | %% Energie Berechnung
52 | EPlus = imfilter(A, gabCP).^2 + imfilter(A,gabSP).^2; %Erster
    | Filter
53 | EMinus = imfilter(A, gabCM).^2 + imfilter(A,gabSM).^2; %Zweiter
    | Filter
54 |
55 | EGes = EPlus - EMinus;
56 | %Darstellungen der drei Energien
57 | nexttile; imshow(EPlus,[-1 1]);
58 | nexttile; imshow(EMinus,[-1 1]);
59 | nexttile; imshow(EGes,[-1 1]);
```

Anhang 2: Filterbank.m

```

1 %Filterbank.m
2 % Benoetigt einen Stimulus A, vorgesehene Groesse: 100
3 %% Filter Definierungen für die Filterbank
4 % 10 Filter spatialFrequency(omega), temporalFrequency(upsilon),
   sigma, delta
5 % Erste Haelfte für linkslaufige
6 [f1cp,f1sp,f1cm,f1sm] = gabFilter(37.5,37.5, 3, 2);
7 [f2cp,f2sp,f2cm,f2sm] = gabFilter(17,34, 2, 1.5);
8 [f3cp,f3sp,f3cm,f3sm] = gabFilter(34,17, 2, 1.5);
9 [f4cp,f4sp,f4cm,f4sm] = gabFilter(58,16, 4, 2);
10 [f5cp,f5sp,f5cm,f5sm] = gabFilter(16,58, 4, 2);
11 % Zweite Haelfte für rechtslaufige
12 [f6cp,f6sp,f6cm,f6sm] = gabFilter(-37.5,37.5, 3, 2);
13 [f7cp,f7sp,f7cm,f7sm] = gabFilter(-17,34, 2, 1.5);
14 [f8cp,f8sp,f8cm,f8sm] = gabFilter(-34,17, 2, 1.5);
15 [f9cp,f9sp,f9cm,f9sm] = gabFilter(-58,16, 4, 2);
16 [f10cp,f10sp,f10cm,f10sm] = gabFilter(-16,58, 4, 2);
17 % Energienbestimmungen der einzelnen Filter
18 e1m = imfilter(A,f1cm).^2 + imfilter(A,f1sm).^2;
19 e1p = imfilter(A,f1cp).^2 + imfilter(A,f1sp).^2;
20 e2m = imfilter(A,f2cm).^2 + imfilter(A,f2sm).^2;
21 e2p = imfilter(A,f2cp).^2 + imfilter(A,f2sp).^2;
22 e3m = imfilter(A,f3cm).^2 + imfilter(A,f3sm).^2;
23 e3p = imfilter(A,f3cp).^2 + imfilter(A,f3sp).^2;
24 e4m = imfilter(A,f4cm).^2 + imfilter(A,f4sm).^2;
25 e4p = imfilter(A,f4cp).^2 + imfilter(A,f4sp).^2;
26 e5m = imfilter(A,f5cm).^2 + imfilter(A,f5sm).^2;
27 e5p = imfilter(A,f5cp).^2 + imfilter(A,f5sp).^2;
28 e6m = imfilter(A,f6cm).^2 + imfilter(A,f6sm).^2;
29 e6p = imfilter(A,f6cp).^2 + imfilter(A,f6sp).^2;
30 e7m = imfilter(A,f7cm).^2 + imfilter(A,f7sm).^2;
31 e7p = imfilter(A,f7cp).^2 + imfilter(A,f7sp).^2;
32 e8m = imfilter(A,f8cm).^2 + imfilter(A,f8sm).^2;
33 e8p = imfilter(A,f8cp).^2 + imfilter(A,f8sp).^2;
34 e9m = imfilter(A,f9cm).^2 + imfilter(A,f9sm).^2;
35 e9p = imfilter(A,f9cp).^2 + imfilter(A,f9sp).^2;
36 e10m = imfilter(A,f10cm).^2 + imfilter(A,f10sm).^2;
37 e10p = imfilter(A,f10cp).^2 + imfilter(A,f10sp).^2;
38 % Berechnung der Gesamtenergie der Filterpaare
39 f1 = e1m.*e1p.^2 - e1p.*e1m.^2;
40 f2 = e2m.*e2p.^2 - e2p.*e2m.^2;

```

```
41 f3 = e3m.*e3p.^2 - e3p.*e3m.^2;
42 f4 = e4m.*e4p.^2 - e4p.*e4m.^2;
43 f5 = e5m.*e5p.^2 - e5p.*e5m.^2;
44 f6 = e6m.*e6p.^2 - e6p.*e6m.^2;
45 f7 = e7m.*e7p.^2 - e7p.*e7m.^2;
46 f8 = e8m.*e8p.^2 - e8p.*e8m.^2;
47 f9 = e9m.*e9p.^2 - e9p.*e9m.^2;
48 f10 = e10m.*e10p.^2 - e10p.*e10m.^2;
49 f11 = e11m.*e11p.^2 - e11p.*e11m.^2;
50 f12 = e12m.*e12p.^2 - e12p.*e12m.^2;
51 % Entgeltige Gesamtenergie
52 final = (f1 + f2 + f3 + f4 + f5 + f6 + f8 + f9 + f10);
53 imshow(final,[-1 1]) % Darstellung
54
55
56 %% Funktionsbereich um jeweils einzelne Filterpaare zu erstellen
57 function [gabCP, gabSP, gabCM, gabSM] = gabFilter(spatialFrequency
    , temporalFrequency, sigma, delta)
58 xG = [-50:50];
59 tG = transpose([-50:50]);
60 omega0 = spatialFrequency;
61 ny0 = temporalFrequency;
62 delta = delta;
63 deltav = delta * -(temporalFrequency/spatialFrequency);
64 gg = exp(-(xG.^2+tG.^2)/(2*sigma^2));
65 gabCP = gg .* (cos((omega0 + deltav)* xG +(ny0 - delta) * tG));
66 gabCM = gg .* (cos((omega0 - deltav)* xG +(ny0 + delta) * tG));
67 gabSP = gg .* (sin((omega0 + deltav)* xG +(ny0 - delta) * tG));
68 gabSM = gg .* (sin((omega0 - deltav)* xG +(ny0 + delta) * tG));
69 end
```

Restlicher Matlab-Code auf beigefügtem Medium.

Erklärung

Hiermit erkläre ich, dass ich die Arbeit selbständig verfasst, keine anderen als die angegebenen Hilfsmittel und Quellen benutzt habe, alle wörtlich oder sinngemäß aus anderen Werken übernommenen Aussagen als solche gekennzeichnet habe und dass die Arbeit weder vollständig noch in wesentlichen Teilen Gegenstand eines anderen Prüfungsverfahrens gewesen ist und dass ich die Arbeit weder vollständig noch in wesentlichen Teilen bereits veröffentlicht habe sowie dass das in Dateiform eingereichte Exemplar mit eingereichten gebundenen Exemplaren übereinstimmt.

Tübingen, den 3. Februar 2021

LÉON DANKERT