

Intelligente Sehsysteme

7 Fourier-Transformation und Skalenräume

Orts- und Frequenzraum, Fourier-Transformation
Multiskalenstrategien und -räume

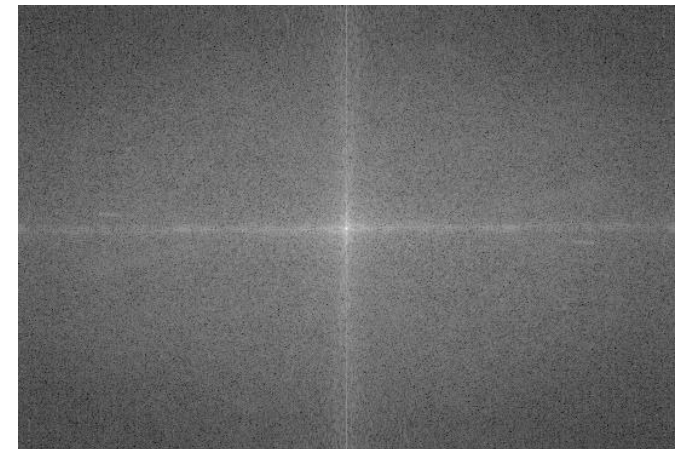
Volker Steinhage

Inhalt

- Basisfunktionen
- Orts- und Frequenzraum
- Fourier-Transformation
- Filter im Frequenzraum
- Multiskalen-Strategien
- Multiskalenraum
- Gauß-Pyramide
- Laplace-Pyramide

Fourier-Transformation

- **Fourier-Transformation (FT)**: Beschreibung einer beliebigen Funktion als Summe von gewichteten periodischen Funktionen (**Basisfunktionen**) mit unterschiedlichen Frequenzen (**Frequenzraumrepräsentation**).
- Anwendungen sind u.a.:
 - Beschreibung des Informationsverlusts bei Digitalisierung
 - Restauration von linearen Störungen
 - Schnelle Filterung
 - hier: **Motivation von Multiskalenansätzen**



Basisfunktionen

- Bilder wurden bereits als zweidimensionale Funktionen beschrieben.
- Allgemein:

Jede Funktion $f(n)$ kann als Summe von N gewichteten Basisfunktionen b_u aufgefasst werden:

$$f(n) = \sum_{u=0}^{N-1} w_u \cdot b_u(n)$$

Die Gewichte w_u bilden eine neue Funktion $w(u)$, die $f(n)$ zusammen mit den Basisfunktionen genau beschreibt.

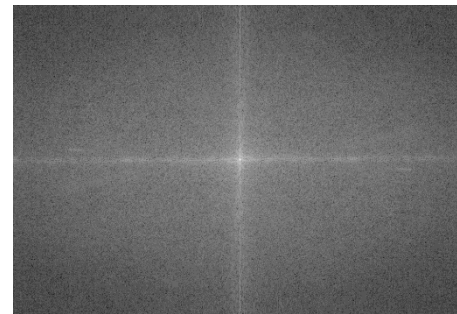
- Die Fourier-Transformation ist die Transformation von einer Ortsbasis in eine Frequenzbasis. Was heißt das?

Ortsraum und Frequenzraum

- Die Basisfunktionen der FT sind periodische Funktionen unterschiedlicher Frequenz und Amplitude. Durch diese wird die transformierte Bildfunktion repräsentiert. Man spricht von einer Repräsentation im **Frequenzraum**.
- Das ursprüngliche Bild wird Repräsentation im **Ortsraum** genannt.
- Funktionen im **Ortsraum** werden i.A. durch kleine Buchstaben (z.B. $f(i,j)$) und korrespondierende **Frequenzraumfunktionen** durch große Buchstaben (z.B. $F(u,v)$) bezeichnet.



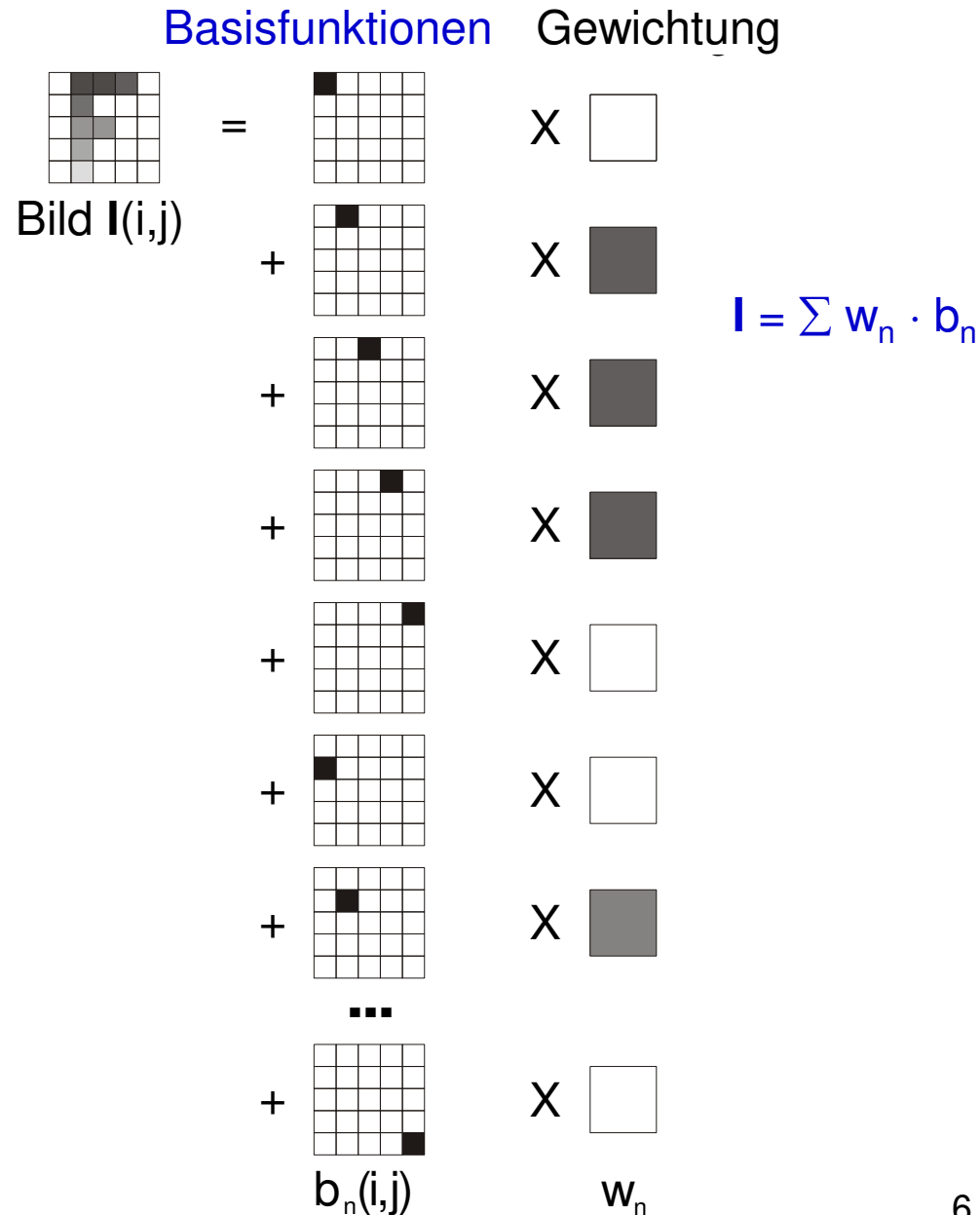
Repräsentation im Ortsraum



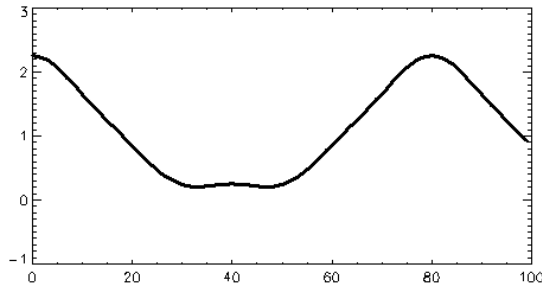
Repräsentation im Frequenzraum

Basisfunktionen im Ortsraum

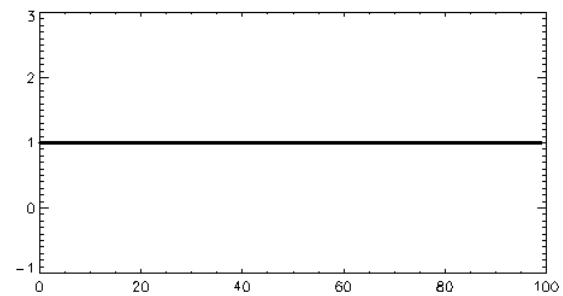
- Jede **Basisfunktion** ist ein Bild, das in genau einem Pixel den Wert 1 und in allen anderen Pixeln den Wert 0 hat.
- Zwei Basisfunktionen b_k und b_l unterscheiden sich dadurch, dass sie den „1-Pixel“ an unterschiedlichen Stellen haben.
- Die vorgeschlagene Basis ist vollständig und eindeutig.



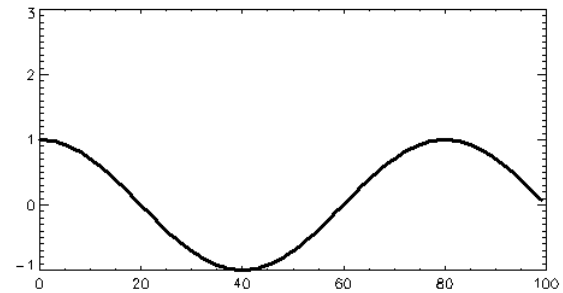
Basisfunktionen im Frequenzraum (1-D)



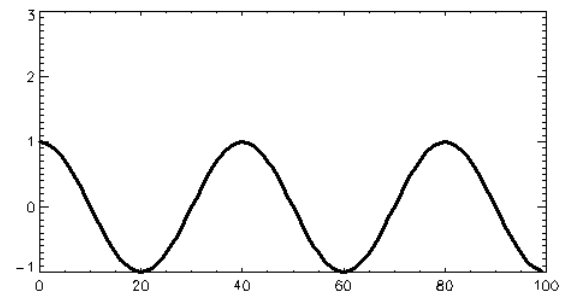
= $1.0 \times$



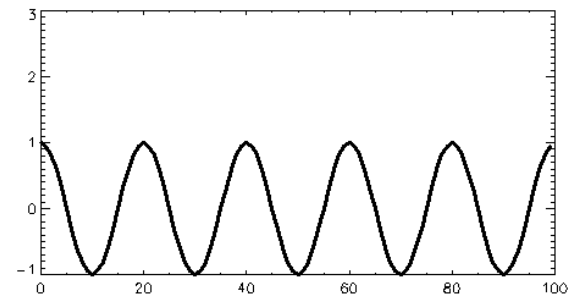
+ $1.0 \times$



+ $0.2 \times$



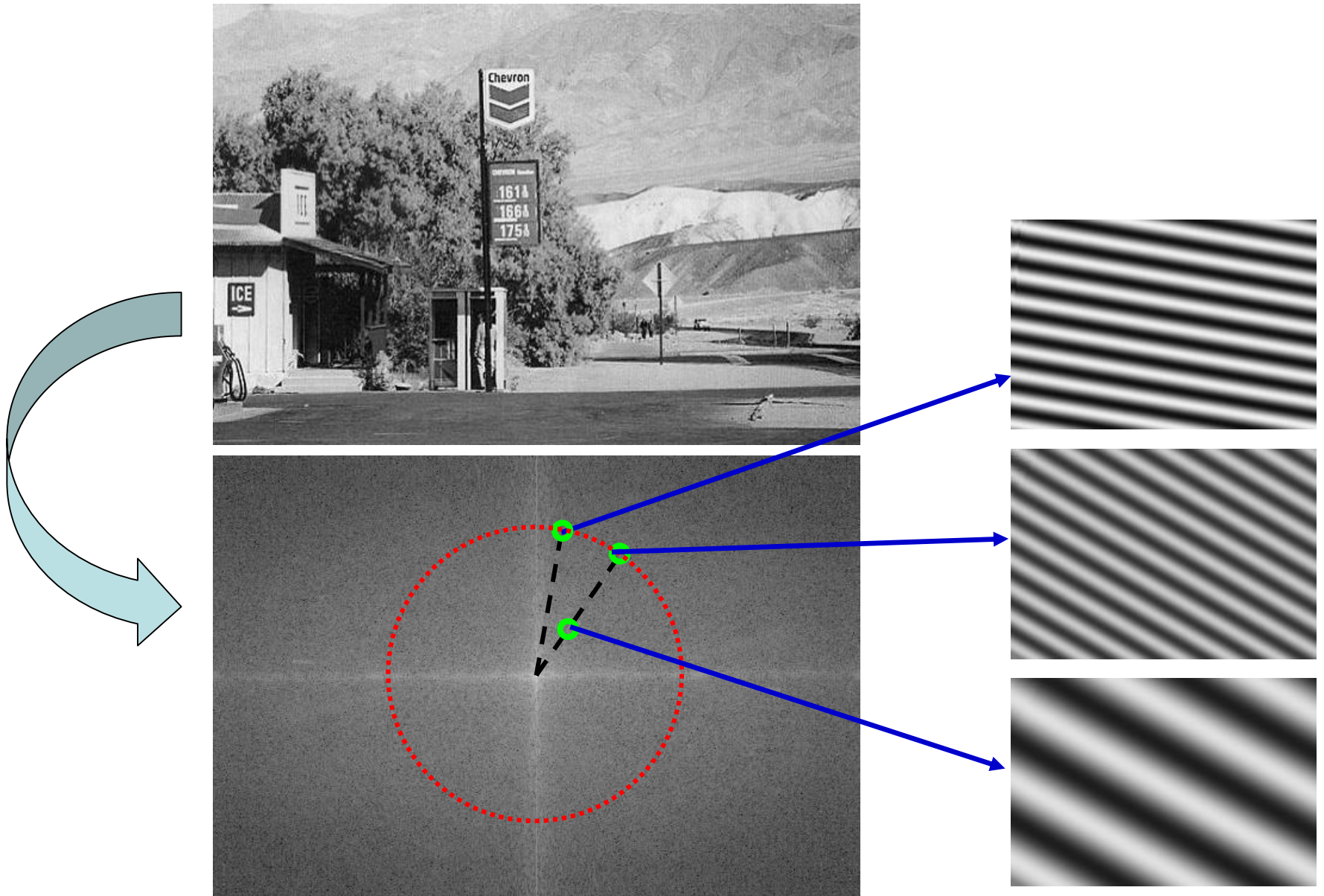
+ $0.05 \times$



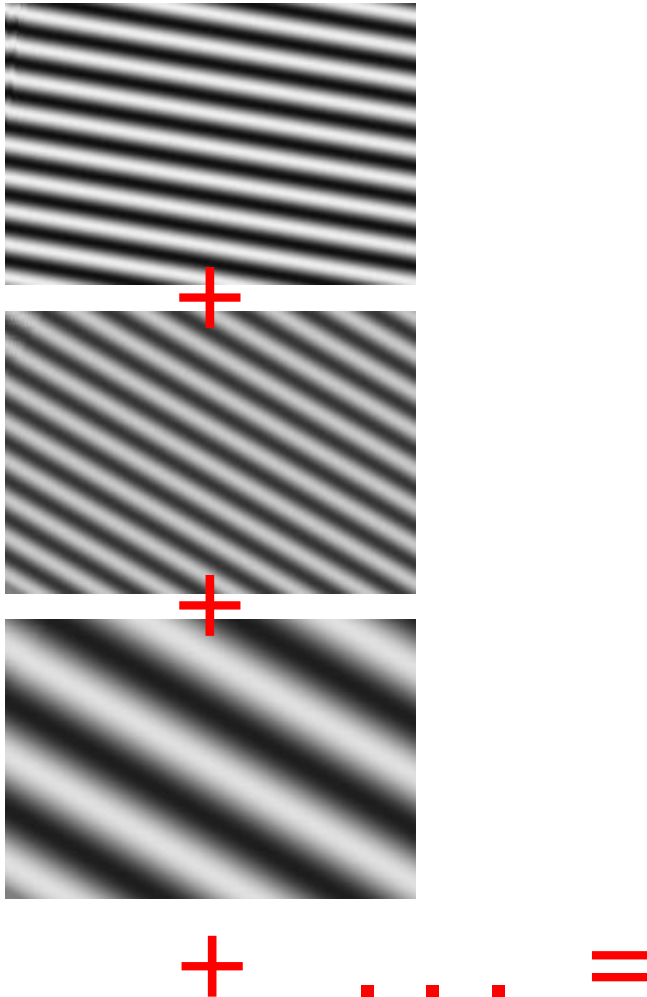
Bildquelle: Klaus Tönnies: Grundlagen der Bildverarbeitung, Pearson Studium, 2005.



Basisfunktionen im Frequenzraum (2-D)

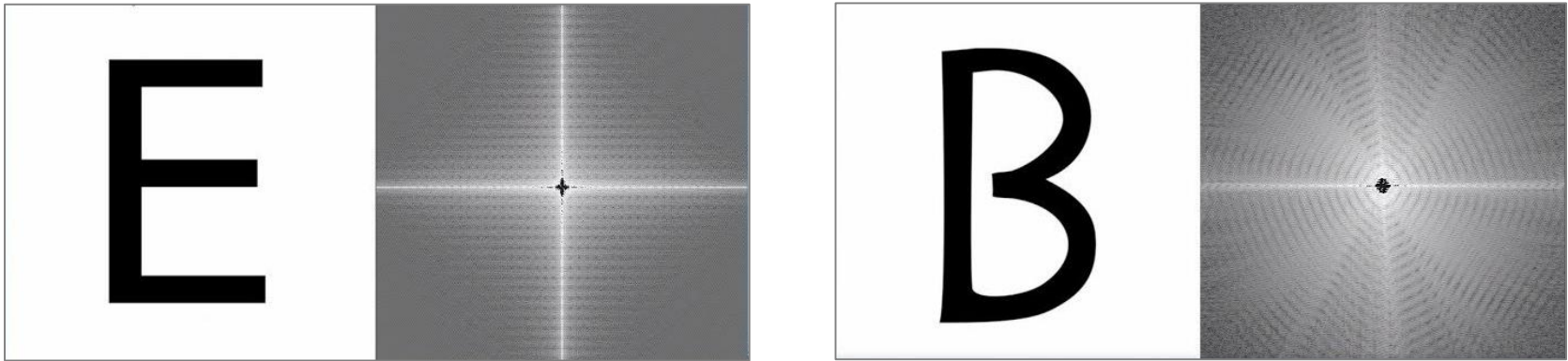


Frequenzbasis (2-D) – Forts.

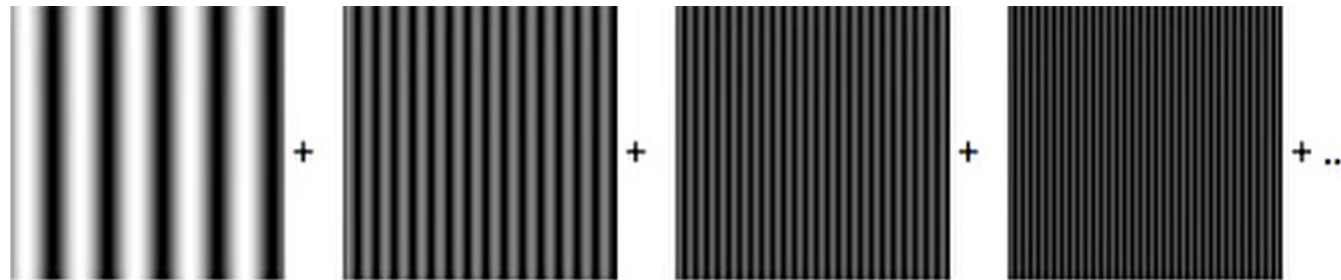


Bildquelle: Klaus Tönnies:
Grundlagen der Bildverarbeitung,
Pearson Studium, 2005.

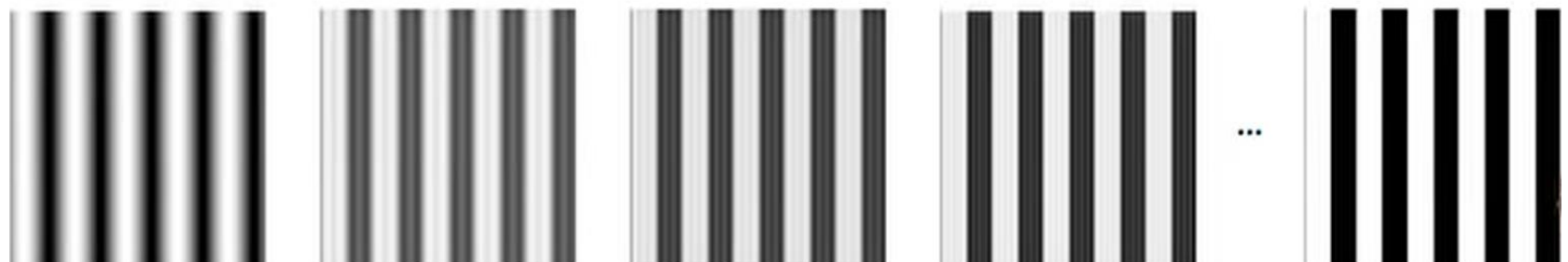
Frequenzbasis (2-D) – Forts.



Bildquelle: Sergio Canu, PySource: Fourier Transform – OpenCV 3.4 with python 3 Tutorial 35



Summed image



Bildquelle: Michelle Dunn, Swinburne Univ. of Technology, Fourier transforms in image processing

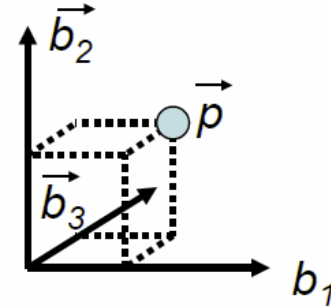
Invertierbarkeit

- Die Repräsentation und Verarbeitung von Bildern im Frequenzraum sollte auch im Ortsraum wiedergebar sein \leadsto Invertierbarkeit der Transformation
- Allg.:
 - **Transformation**: „Projektion“ der Bildfunktion auf die neue Basis
 - **Inverse Transformation**: „Rückprojektion“ auf die alte Basis
 - Eine Transformation ist invertierbar,
wenn die Basisfunktionen eine **orthogonale Basis** bilden
 - 1) Was ist Orthogonalität für Funktionen?
 - 2) Wann bilden Basisfunktionen eine orthogonale Basis?
- Orthogonalität, Projektion, Transformation etc. sind Begriffe aus der **Vektoralgebra** und können auch für Funktionen genutzt werden

Orthogonale Vektorbasis

- N Vektoren bilden eine Basis für einen N -dim. Raum, wenn sie alle **orthogonal** zueinander sind.
- Zwei Vektoren sind **orthogonal**, wenn ihr **Skalarprodukt** Null ist.
- Transformation: Projektion der Funktion auf die neue Basis.

Die Projektion eines Vektors \vec{p} auf einen Basisvektor \vec{b}_i ist durch das normierte **Skalarprodukt** zwischen ihnen gegeben.



$$\vec{b}_i \cdot \vec{b}_j = 0, \quad \text{für } i, j = 1, 2, 3 \wedge i \neq j$$

$$\vec{p}^B = \left(\frac{\vec{p} \cdot \vec{b}_1}{\|\vec{b}_1\|} \quad \frac{\vec{p} \cdot \vec{b}_2}{\|\vec{b}_2\|} \quad \frac{\vec{p} \cdot \vec{b}_3}{\|\vec{b}_3\|} \right) \quad \odot$$

Bildquelle: Klaus Tönnies:
Grundlagen der Bildverarbeitung,
Pearson Studium, 2005.

Funktionen statt Vektoren

- Eine Funktion $f(x)$ sei durch vollständige Aufzählung aller Werte eindeutig definiert: $f(x) = \{f(x_1), f(x_2), \dots\}$

- Skalarprodukt zwischen Funktionen f_1 und f_2 :

- Reeller, unbeschränkter Definitionsbereich: $\int_{-\infty}^{\infty} f_1(x)f_2(x)dx$

- Reeller, beschränkter Definitionsbereich: $\int_{x_{\min}}^{x_{\max}} f_1(x)f_2(x)dx$

- Ganzzahliger, unbeschränkter Definitionsbereich: $\sum_{n=-\infty}^{\infty} f_1(n)f_2(n)$

- Ganzzahliger, beschränkter Definitionsbereich: $\sum_{n=0}^{N-1} f_1(n)f_2(n)$

hier
relevant

- Orthogonale Basen können wie für Vektoren definiert und genutzt werden.

Für mehrdimensionale Funktionen

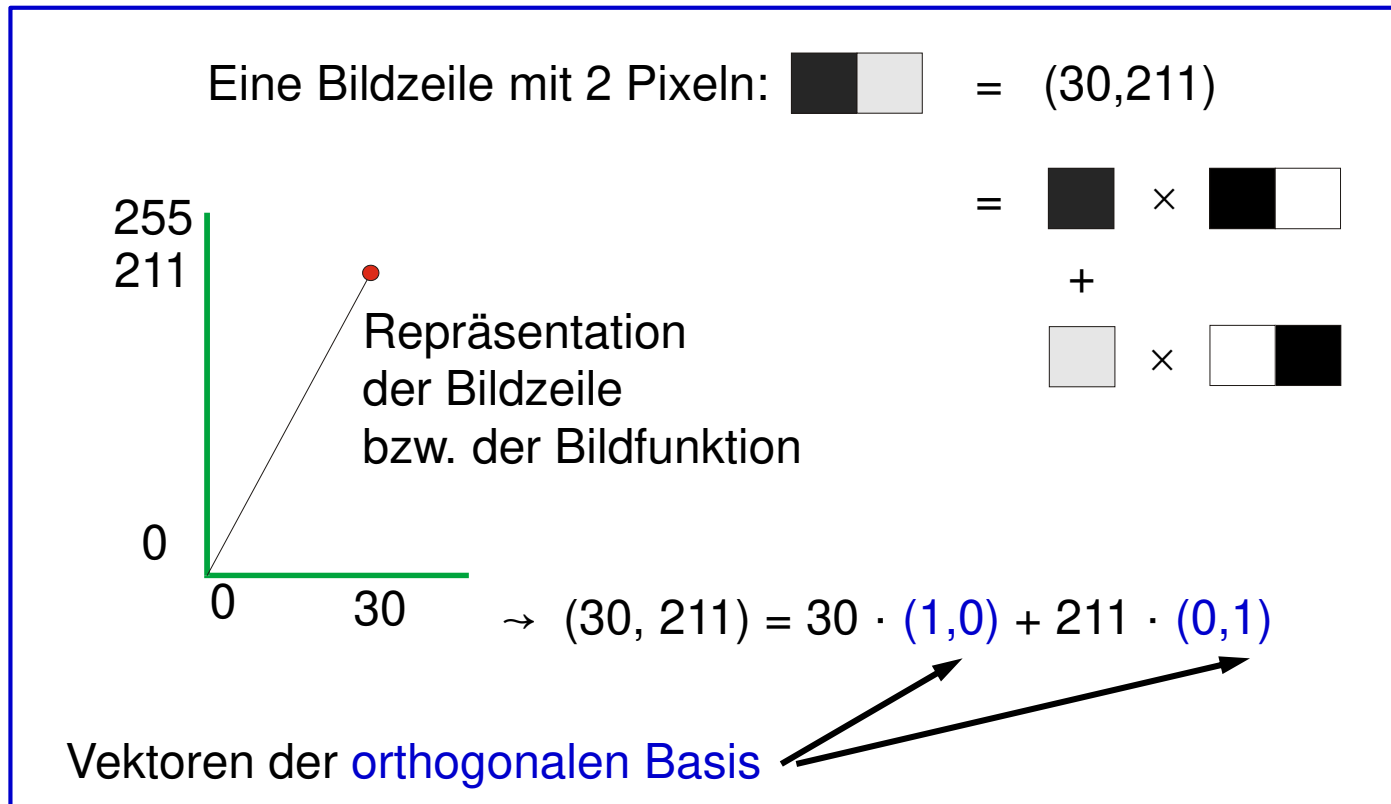
- Konzept der Basisfunktionen ist auf mehrdim. Funktionen erweiterbar. Nur das Skalarprodukt muss definiert werden.
- Für zwei Funktionen $f_1(m,n)$ und $f_2(m,n)$ mit ganzzahligem, beschränktem Definitionsbereich $(0,\dots,N-1)(1,\dots,M-1)$ ist das **Skalarprodukt** definiert als:

$$f_1 \bullet f_2 = \sum_{n=0}^{N-1} \sum_{m=0}^{M-1} f_1(m,n) \cdot f_2(m,n).$$

- Das Konzept lässt sich auf Funktionen mit reellwertigen und unbeschränkten Definitionsbereichen erweitern. Dies ist hier für digitale Bilder nicht relevant.

Funktion als Vektor (Beispiel)

- Ein Bild wird also als Funktion mit *endlichem Definitionsbereich* aufgefasst.
 - Anzahl der Pixel = Anzahl der Funktionswerte = Vektordimension
 - Intensitätsspektrum = Wertebereich für jedes Pixel.



Basis-Transformation und Invertierbarkeit

- Die Transformation T_b in eine andere Basis ist invertierbar,
 - wenn sie **orthogonal** ist
 - wenn die **Anzahl der Basisvektoren gleich** ist.
- Orthogonalität von zwei ganzzahligen Basisfunktionen mit beschränktem Definitionsbereich : $[b_1 \cdot b_2](n) = \sum_{n=0}^{N-1} b_1(n) \cdot b_2(n) = 0$
- Jede orthogonale Basis kann durch Rotation (und Translation) aus jeder anderen Basis für die gleiche Funktion erzeugt werden.

Transformation entspricht einer

Rotation (und ggf. Translation):

$$T_b[f](u) = \sum_{n=0}^1 f(n) \cdot b_u(n)$$

$$b_1(n) = (\cos \gamma \quad \sin \gamma)$$

$$b_2(n) = (-\sin \gamma \quad \cos \gamma)$$



Basis-Transformation: Vektor-Matrix-Darstellung

- Eine Funktion kann als Vektor \vec{f} ihrer Funktionswerte repräsentiert werden:

$$\vec{f} = (f(0) \ f(1) \ \dots \ f(N-1))$$

- Basisfunktionen b_u können in quadrat. Matrix **B** zusammengefasst werden:

$$\mathbf{B} = \begin{pmatrix} b_0(0) & b_1(0) & \dots & b_{N-1}(0) \\ b_0(1) & b_1(1) & & \dots \\ \dots & & & \\ b_0(N-1) & \dots & & b_{N-1}(N-1) \end{pmatrix}$$

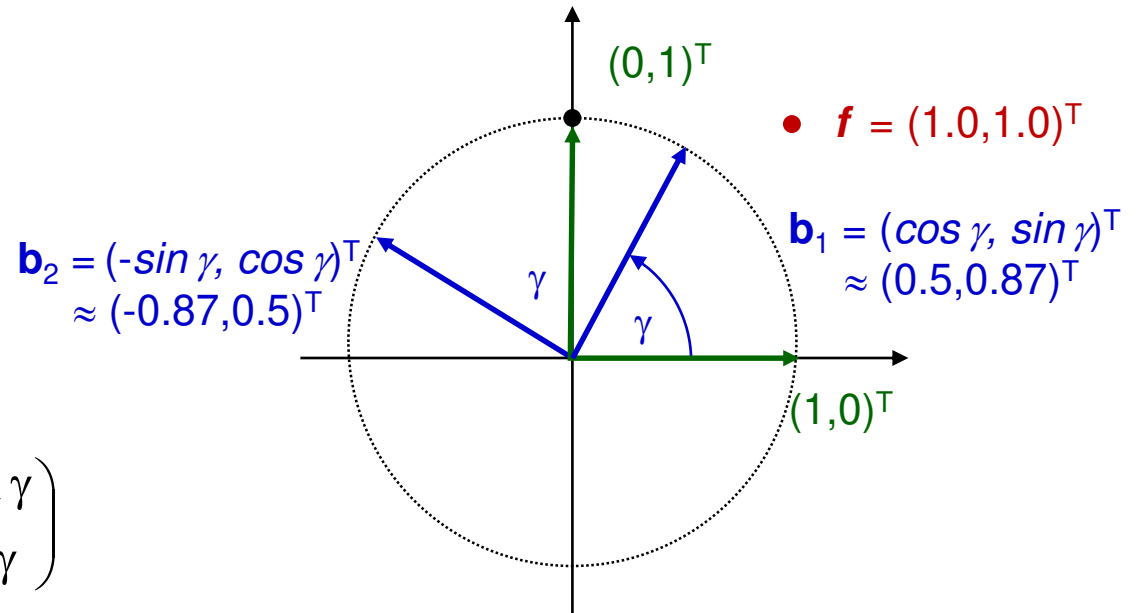
- Transformation auf die neue Basis ist die Multiplikation von \vec{f} mit **B**:

$$\vec{f'} = \vec{f} \times \mathbf{B}$$

Basis-Transformation: Beispiel

Neue Basisvektoren $\mathbf{b}_1, \mathbf{b}_2$ als um Winkel $\gamma = 60^\circ$ gegen den UZS rotierte Version der **alten Basis**:

$$\mathbf{b}_1 \cdot \mathbf{b}_2 = \begin{pmatrix} 0.5 \\ 0.87 \end{pmatrix} \cdot \begin{pmatrix} -0.87 \\ 0.5 \end{pmatrix} = 0$$

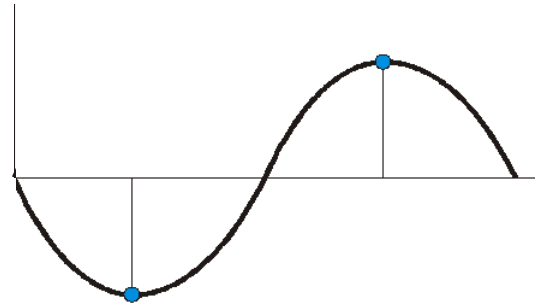
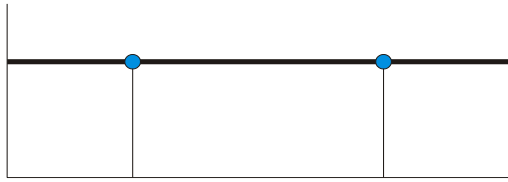


$$\mathbf{B} = (\mathbf{b}_1 \quad \mathbf{b}_2) = \begin{pmatrix} \cos \gamma & -\sin \gamma \\ \sin \gamma & \cos \gamma \end{pmatrix}$$

$$\mathbf{f}' = \mathbf{f} \cdot \mathbf{B} = \begin{pmatrix} 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} 0.5 & -0.87 \\ 0.87 & 0.5 \end{pmatrix} = \begin{pmatrix} 1.37 & -0.37 \end{pmatrix}$$

Orthogonale periodische Funktionen (1)

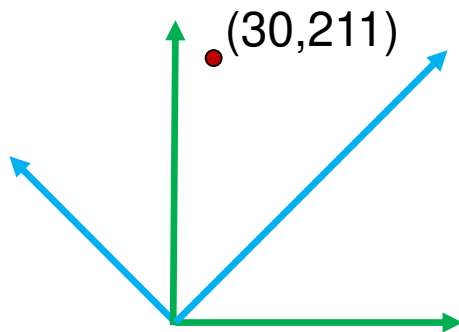
Zwei Basisfunktionen, die an zwei **diskreten Orten** definiert sind:



Basis:

$(1,1)$

$(-1,1)$



Transformation:

$$(30,211) \cdot \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} = (241,181)$$

Rücktransformation:

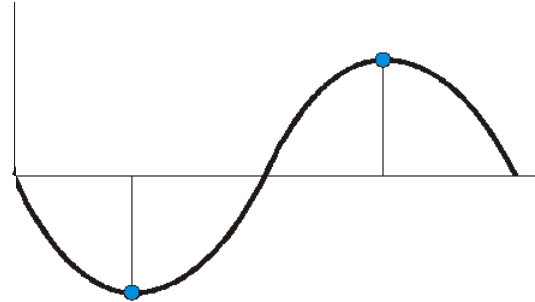
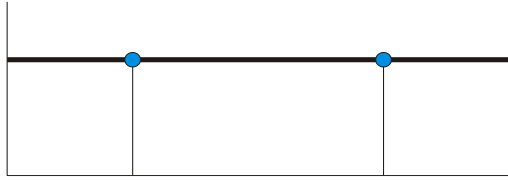
$$(241,181) \cdot \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} = (60,422)$$

Anmerkung:

Das Resultat der Rücktransformation muss skaliert werden, weil die Basis nicht *normiert* ist.

Orthogonale periodische Funktionen (2)

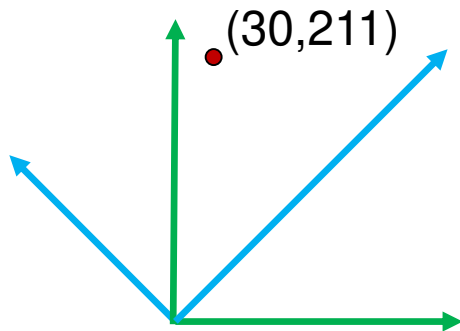
Dieselben zwei Basisfunktionen:



Orthonormalbasis:

$(0.7071, 0.7071)$

$(-0.7071, 0.7071)$



Transformation:

$$(30, 211) \cdot \begin{pmatrix} 0.707 & -0.707 \\ 0.707 & 0.707 \end{pmatrix} = (170.4, 128.0)$$

Rücktransformation:

$$(170.4, 128.0) \cdot \begin{pmatrix} 0.707 & 0.707 \\ -0.707 & 0.707 \end{pmatrix} = (30, 211)$$

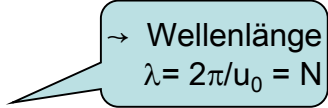
Vollständige Basis für den Frequenzraum (1)

Gegeben seien zwei ungleiche Kosinusfunktionen $\cos(u_1 \cdot n)$ und $\cos(u_2 \cdot n)$:

- diese sind an **diskreten Orten** $n = 0, \dots, N-1$ definiert (\rightarrow diskrete FT)
- die Frequenzen u_1 und u_2 sind **ganzzahl. Vielfache einer Basisfrequenz u_0**

Dann gilt:

Ist die Basisfrequenz u_0 gleich $2\pi/N$,



\rightarrow Wellenlänge
 $\lambda = 2\pi/u_0 = N$

so sind $\cos(u_1 \cdot n)$ und $\cos(u_2 \cdot n)$ orthogonal

Vollständige Basis für den Frequenzraum (2)

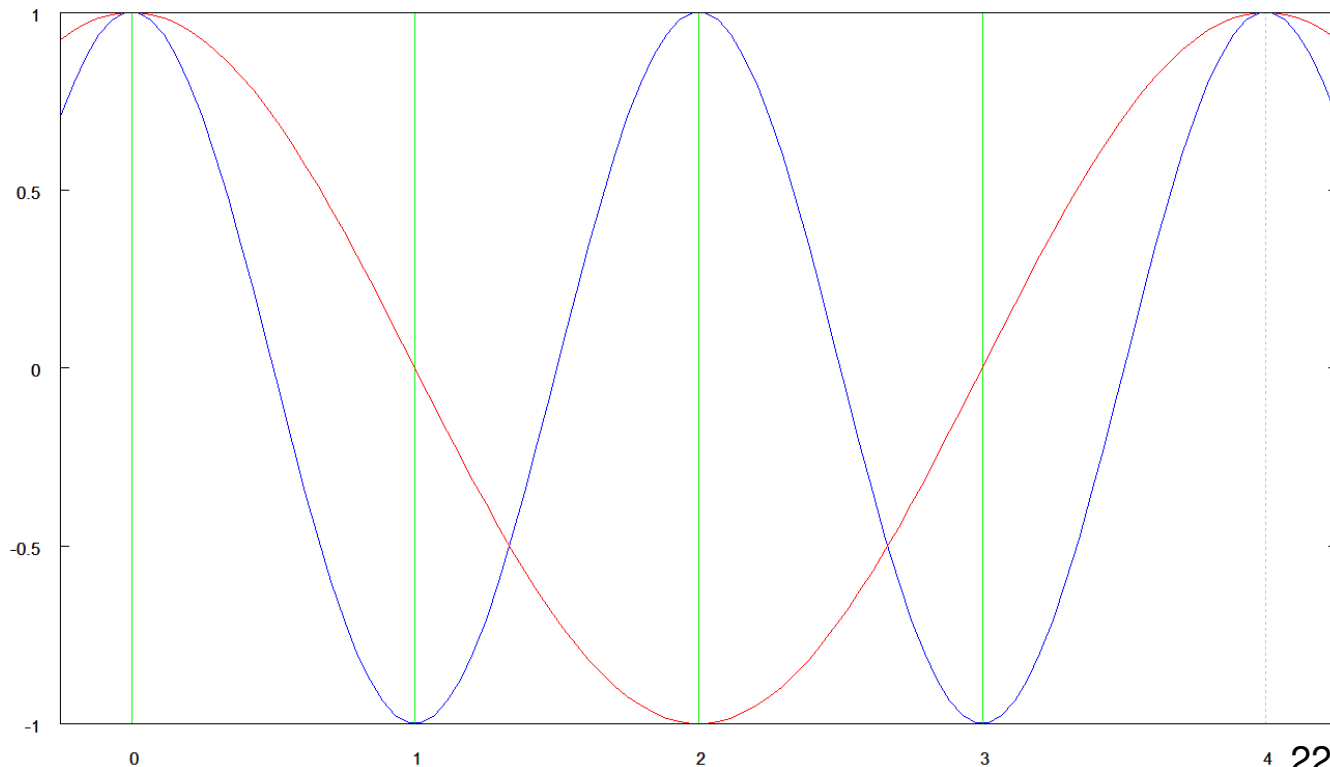
Bspl.: $N = 4$, $u_0 = 2\pi/N = \frac{1}{2}\pi$, $u_1 = 1 \cdot u_0$, $u_2 = 2 \cdot u_0$

$$f_1(n) = \cos(u_1 \cdot n) = \cos(2 \cdot \pi \cdot n / 4) = \cos(\frac{1}{2} \cdot \pi \cdot n) \text{ für } n = 0, \dots, 3$$

$$f_2(n) = \cos(u_2 \cdot n) = \cos(2 \cdot 2 \cdot \pi \cdot n / 4) = \cos(\pi \cdot n) \text{ für } n = 0, \dots, 3$$

$\rightarrow f_1(n)$ und $f_2(n)$ sind orthogonal: $[f_1 \cdot f_2](n) = \sum_{n=0}^{N-1} f_1(n) \cdot f_2(n) = 0$

n	$f_1(n)$	$f_2(n)$	$f_1(n) \cdot f_2(n)$
0	1	1	1
1	0	-1	0
2	-1	1	-1
3	0	-1	0
Σ			0



Abszisse : n
Ordinate: $f(n)$

Vollständige Basis für den Frequenzraum (3)

Aber für zwei an N diskreten Orten definierte Funktionen $\cos(u_1 \cdot n)$ und $\cos(u_2 \cdot n)$ mit $\frac{u_1}{u_0} = N - \frac{u_2}{u_0}$ gilt: $\cos(u_1 \cdot n) = \cos(u_2 \cdot n)$ für alle $n = 0, \dots, N-1$

Wellen mit Frequenzen $\leq N/2$ sind damit (an den definierten Stellen) nicht unterscheidbar von derart korrespondierenden Wellen mit Frequenzen $> N/2$ für alle $n = 0, \dots, N-1$

Damit stehen *nicht* hinreichend viele Kosinus-Funktionen
für eine Basis zur Verfügung

Vollständige Basis für den Frequenzraum (4)

Bspl.: $N = 4$, $u_0 = 2\pi/N = 1/2 \cdot \pi$, $u_1 = 3 \cdot u_0$, $u_2 = 1 \cdot u_0$

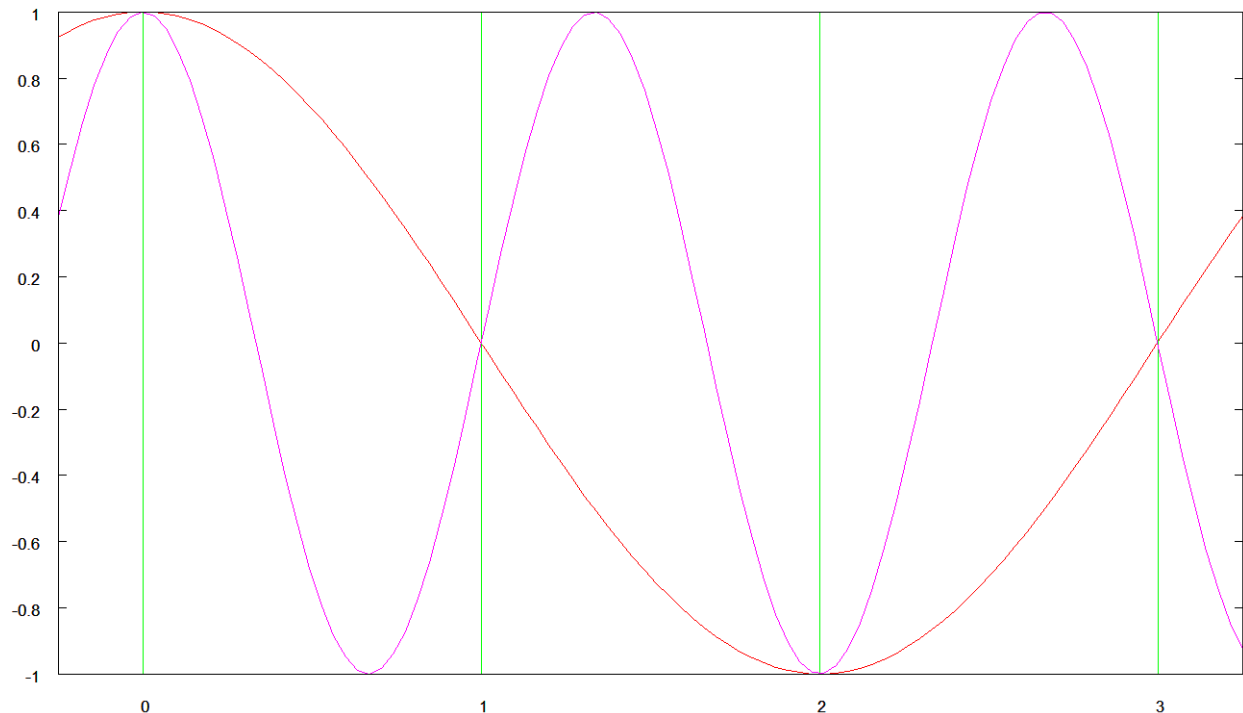
$$u_1/u_0 = 3 = N - u_2/u_0 = 4 - 1$$

$$f_1(n) = \cos(u_1 \cdot n) = \cos(3 \cdot 2 \cdot \pi \cdot n / 4) = \cos(1\frac{1}{2} \cdot \pi \cdot n) \text{ für } n = 0, \dots, 3$$

$$f_2(n) = \cos(u_2 \cdot n) = \cos(1 \cdot 2 \cdot \pi \cdot n / 4) = \cos(1/2 \cdot \pi \cdot n) \text{ für } n = 0, \dots, 3$$

$\rightarrow f_1(n)$ und $f_2(n)$ sind nicht unterscheidbar für $n = 0, \dots, 3$

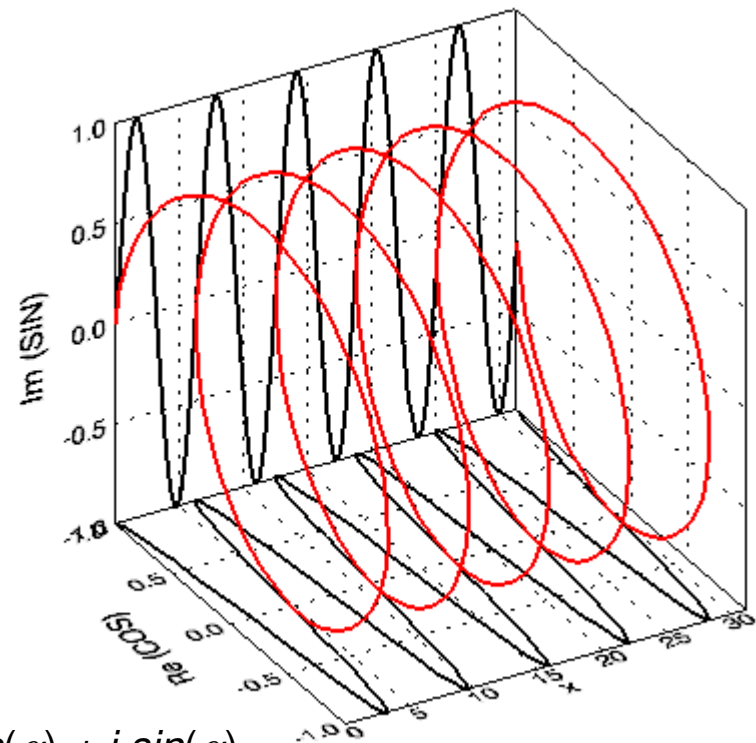
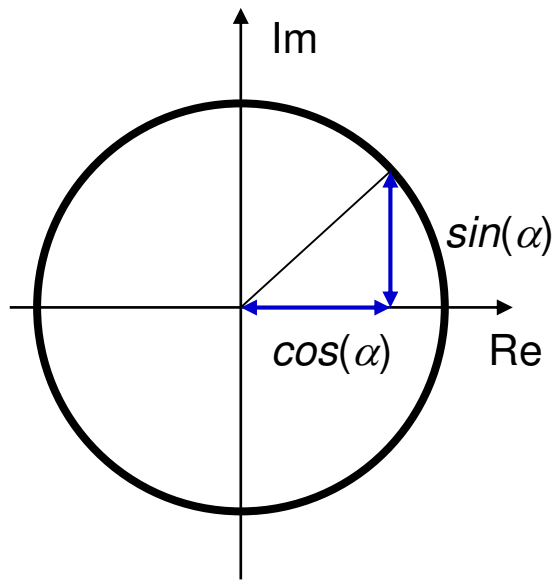
n	$f_1(n)$	$f_2(n)$
0	1	1
1	0	0
2	-1	-1
3	0	0



$$f_1(n) = \cos(u_2 \cdot n) = \cos(3 \cdot 2 \cdot \pi \cdot n / 4) = \cos(1\frac{1}{2} \cdot \pi \cdot n) \text{ und } f_2(n) = \cos(u_1 \cdot n) = \cos(2 \cdot \pi \cdot n / 4) = \cos(1/2 \cdot \pi \cdot n)$$

Vollständige Basis für den Frequenzraum (5)

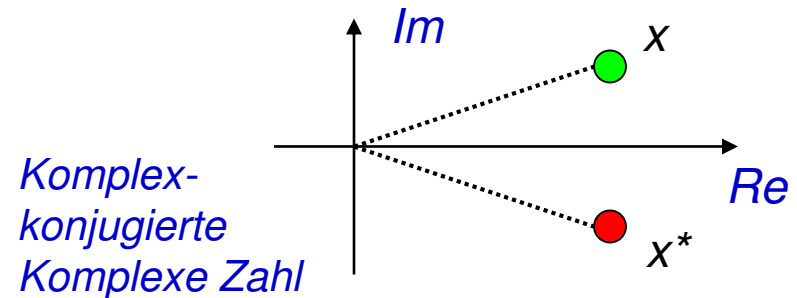
- Für eine vollständige orthogonale Basis sind also weitere Basisfunktionen zu finden, die möglichst der Semantik der Kosinuswellen folgen
- Anstelle reellwertiger Funktionen $\cos(u \cdot n)$ werden komplexe Funktionen $\cos(u \cdot n) + i \cdot \sin(u \cdot n)$ gewählt



Alle Werte für komplexe Zahlen der Form $\cos(\alpha) + i \sin(\alpha)$ liegen auf einem Kreis mit Abstand 1 in der komplexen Ebene.

Vollständige Basis für den Frequenzraum (6)

- Basisfunktionen sind also $\vec{b}_u(n) = [b_{u,\cos} \ b_{u,\sin}] = [\cos(n \cdot u \cdot 2\pi/N) \ \sin(n \cdot u \cdot 2\pi/N)]$
- Das Skalarprodukt zwischen zwei Vektoren mit komplexen Elementen:
 - Summe über den Produkten der Komponenten des ersten Vektors mit den komplex-konjugierten Komponenten des zweiten Vektors.
 - Die komplex-konjugierte komplexe Zahl zu $x = a + i \cdot b$ ist $x^* = a - i \cdot b$.
 - Also:



$$\vec{x} \bullet \vec{y} = \sum_{i=0}^{N-1} x_i \cdot y_i^* = \sum_{i=0}^{N-1} (\text{Re}(x_i) + i \text{Im}(x_i))(\text{Re}(y_i) - i \text{Im}(y_i))$$

Vollständige Basis für den Frequenzraum (7)

- Analog zu den reellwertigen Kosinusfunktionen lässt sich zeigen:
zwei komplexe Wellen $\cos(u_1 \cdot n) + i \sin(u_1 \cdot n)$ und $\cos(u_2 \cdot n) + i \sin(u_2 \cdot n)$
sind für ganzzahlige Vielfache $u_1 \neq u_2$ einer Basisfrequenz $u_0 = 2\pi/N$ im
Definitionsbereich $n = 0, \dots, N-1$ zueinander orthogonal
- Zusätzlich sind jetzt durch die zusätzlichen Sinuswellen bis zu N unterschiedliche Wellenlängen bei N diskreten Definitionsstellen auf dem Intervall unterschiedlich repräsentierbar
- Damit spannen die N Funktionen den gewünschten N -dim. Raum auf

Eindimensionale Fourier-Transformation (1)

- Basisfunktionen sind also $\vec{b}_u(n) = [b_{u,\cos} \ b_{u,\sin}] = [\cos(n \cdot u \cdot 2\pi/N) \ \sin(n \cdot u \cdot 2\pi/N)]$
- Bei der Fourier-Transformation wird eine Funktion $f(n)$ durch Konvolution mit den Basisfunktionen auf den N-dim. Frequenzraum projiziert gemäß:

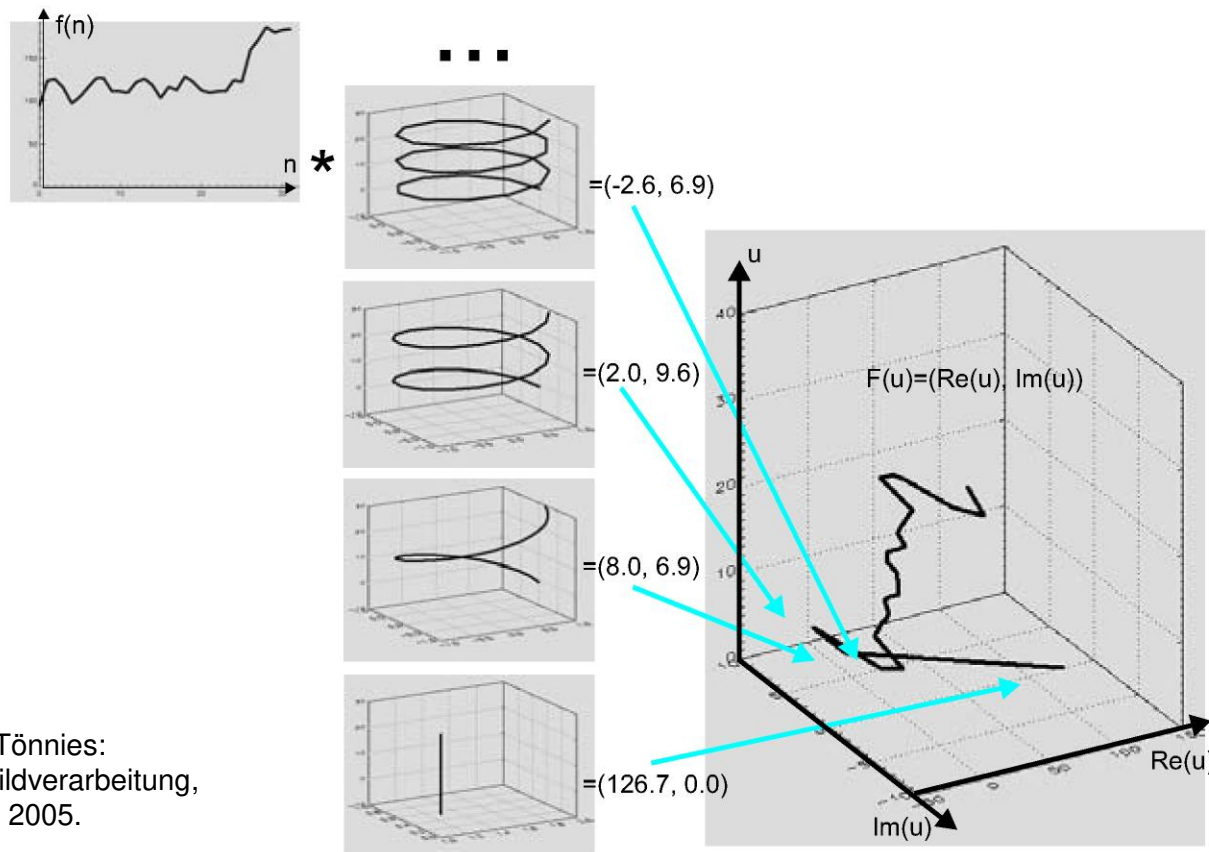
$$F(u) = \frac{1}{s} \sum_{n=0}^{N-1} f(n) \left[\cos\left(\frac{2\pi un}{N}\right) - i \cdot \sin\left(\frac{2\pi un}{N}\right) \right].$$

- Durch die Skalierung s werden die Basisfunktionen normiert. Der Betrag der Basisfunktionen ist dabei:

$$s = \|\vec{b}_u\| = \sqrt{\sum_{n=0}^{N-1} \cos^2\left(\frac{2\pi un}{N}\right) + \sin^2\left(\frac{2\pi un}{N}\right)} = \sqrt{N}.$$

Eindimensionale Fourier-Transformation (2)

Bspl.: Projektion einer 1-dim. Funktion auf eine Folge von orthogonalen, komplexen, periodischen Basisfunktionen. Das Ergebnis jeder einzelnen Projektion ist eine komplexe Zahl. Die Folge aller komplexen Zahlen ergibt die transformierte Funktion.



Bildquelle: Klaus Tönnies:
Grundlagen der Bildverarbeitung,
Pearson Studium, 2005.

Eindimensionale FT mit komplexen Basisfunktionen (1)

- Taylorreihenentwicklung für Kosinus und Sinus:

$$\cos(x) = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \dots \quad \sin(x) = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \dots$$

- Taylorreihenentwicklung für e^{ix} :

$$e^x = 1 + \frac{x}{1!} + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots \Rightarrow e^{ix} = 1 + \frac{ix}{1!} + \frac{(ix)^2}{2!} + \frac{(ix)^3}{3!} + \dots$$

- Wegen $i^2 = -1$ gilt daher die Eulersche Formel:

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

Eindimensionale FT mit komplexen Basisfunktionen (2)

Formuliert über Exponentialfunktion:

- Bildfunktion: $f(n)$, $n = 0, \dots, N-1$,

also: N Basisfunktionen

$$b_u(n) = e^{-i \cdot n \cdot u \cdot 2\pi/N} \text{ mit Frequenzfaktoren } u = 0, \dots, N-1,$$

$$\text{z.B. } b_0(n) = [(1,1), (1,1), \dots, (1,1)]$$

- Transformation FT: $\mathbf{FT}(\mathbf{f}) = \mathbf{F} = \mathbf{f} \cdot \mathbf{B}$ (Vektor-Matrix-Schreibweise)

$$F(u) = \frac{1}{\sqrt{N}} \cdot \sum_{n=0}^{N-1} f(n) \cdot \exp(-i \cdot n \cdot u \cdot 2\pi/N), u = 0, \dots, N-1$$

- Rücktransformation \mathbf{FT}^{-1} : $\mathbf{FT}(\mathbf{f}) = \mathbf{F} = \mathbf{f} \cdot \mathbf{B}$ (Vektor-Matrix-Schreibweise)

$$f(n) = \frac{1}{\sqrt{N}} \cdot \sum_{u=0}^{N-1} F(u) \cdot \exp(i \cdot n \cdot u \cdot 2\pi/N), n = 0, \dots, N-1$$



Skalierungsfaktoren, weil die Basisfunktionen nicht normiert sind

Zweidimensionale FT mit komplexen Basisfunktionen

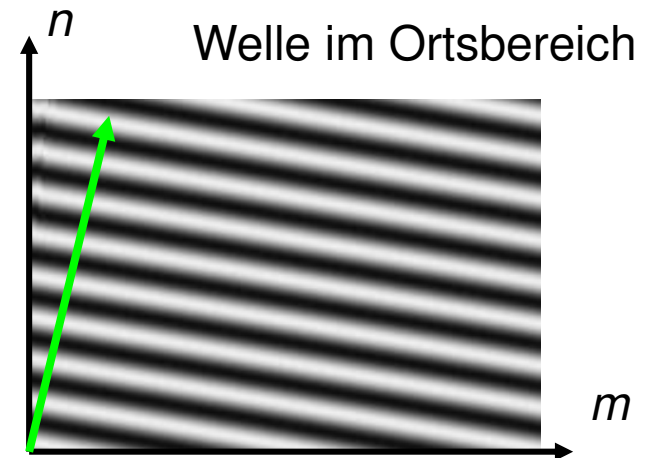
Basisfunktionen sind **Wellen**:

$$\exp((mu+nv) \cdot i \cdot 2\pi/N)$$

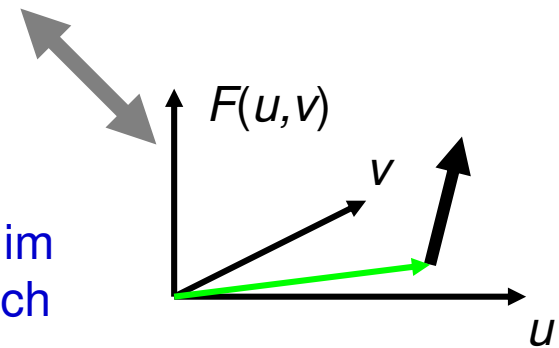
Richtung ist durch Vektor $(u \ v)$ gegeben.

Die Basisfunktionen der 2-D Fourier-Transformation sind **zerlegbar**:

$$\exp((m \cdot u + n \cdot v)) \cdot i \cdot 2\pi/N = \\ \exp(m \cdot u \cdot i \cdot 2\pi/N) \cdot \exp(n \cdot v \cdot i \cdot 2\pi/N)$$



(komplexer)
Funktionswert im
Frequenzbereich



2D Fourier-Transformationspaar

$$F(u, v) = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} f(m, n) \cdot e^{-i \cdot 2\pi \cdot \left(\frac{um}{M} + \frac{vn}{N} \right)}$$

$$f(m, n) = \frac{1}{MN} \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} F(u, v) \cdot e^{i \cdot 2\pi \cdot \left(\frac{um}{M} + \frac{vn}{N} \right)}$$

Transformationspaar für
Bilder der Größe $M \times N$

Transformationspaar für quadrat.

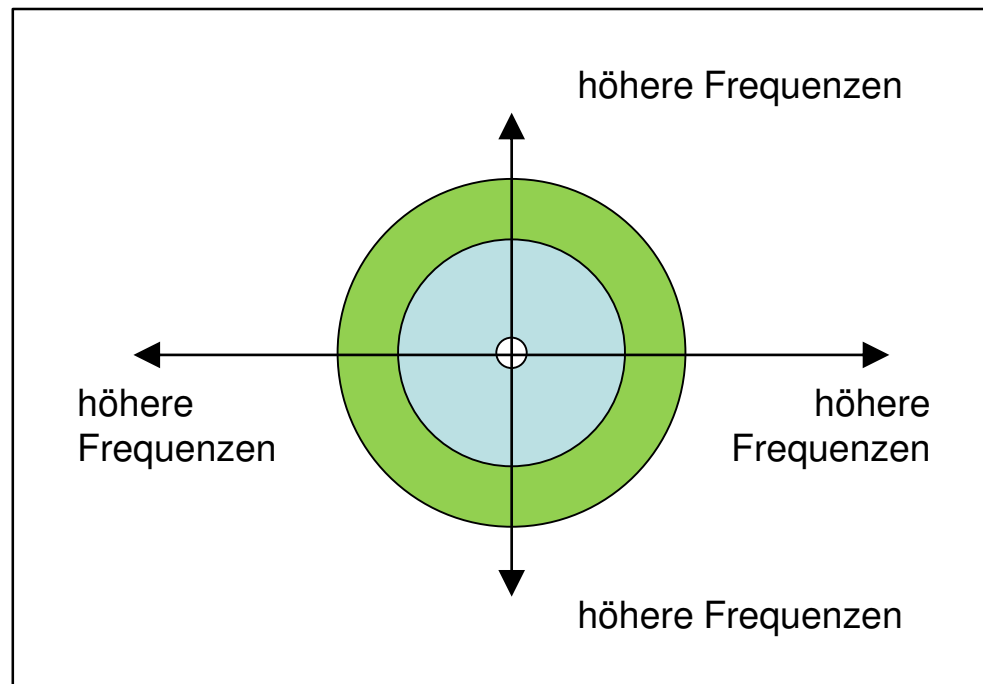
Bilder der Größe $N \times N$

$$F(u, v) = \frac{1}{N} \sum_{m=0}^{N-1} \sum_{n=0}^{N-1} f(m, n) \cdot e^{-i \cdot \frac{2\pi}{N} \cdot (um + vn)}$$

$$f(m, n) = \frac{1}{N} \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} F(u, v) \cdot e^{i \cdot \frac{2\pi}{N} \cdot (um + vn)}$$

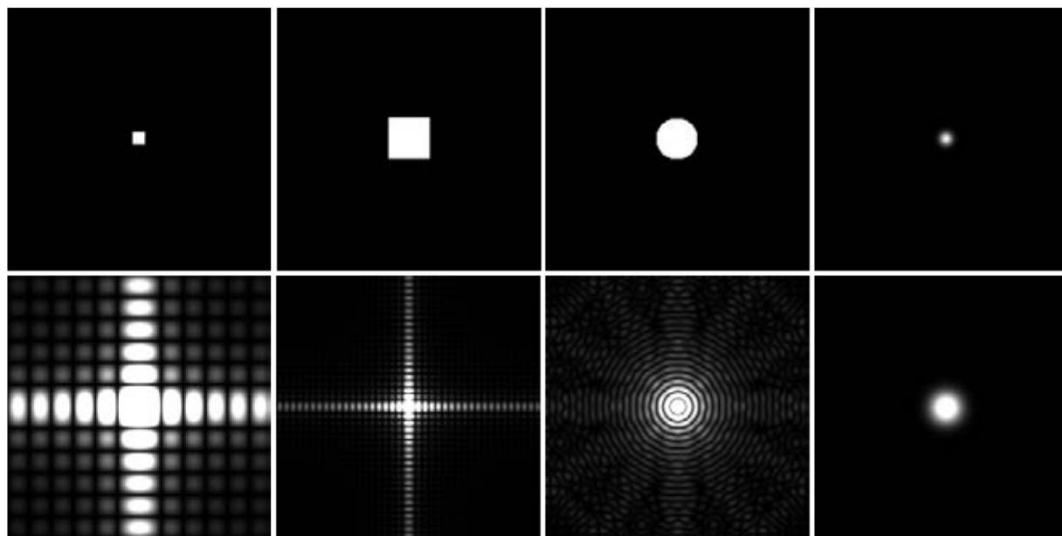
Bilder im Orts- und Frequenzraum (1)

- Das Ergebnis der FT ist komplex, daher wird zur Visualisierung des Amplitudenspektrums der Betrag genutzt: $|F(u, v)| = \sqrt{\left(\operatorname{Re}(F(u, v))\right)^2 + \left(\operatorname{Im}(F(u, v))\right)^2}$
- Bei der Darstellung des Amplitudenspektrums werden die Amplituden der tiefen Frequenzen in den Bildmittelpunkt verschoben
- Das Resultat ist eine im Bildmittelpunkt zentrierte Darstellung, bei dem der Ursprung in der Mitte und die hohen Frequenzen außen sind



Bilder im Orts- und Frequenzraum (2)

- Horizontale Strukturen im Originalbild resultieren in vertikalen Komponenten des Fourier-Spektrums. Vertikale Anordnungen im Originalbild resultieren in horizontalen Komponenten (z.B. Richtungspräferenz nach Ausführung der FT auf den beiden Bildern mit Quadraten).
- Die Repräsentation von scharfen Kanten erfordert sehr viele hohe Frequenzanteile. Nur im Bild rechts außen, das keine harte Kante zwischen Vordergrund und Hintergrund besitzt, sinken die Werte im Frequenzraum rasch mit steigender Frequenz.



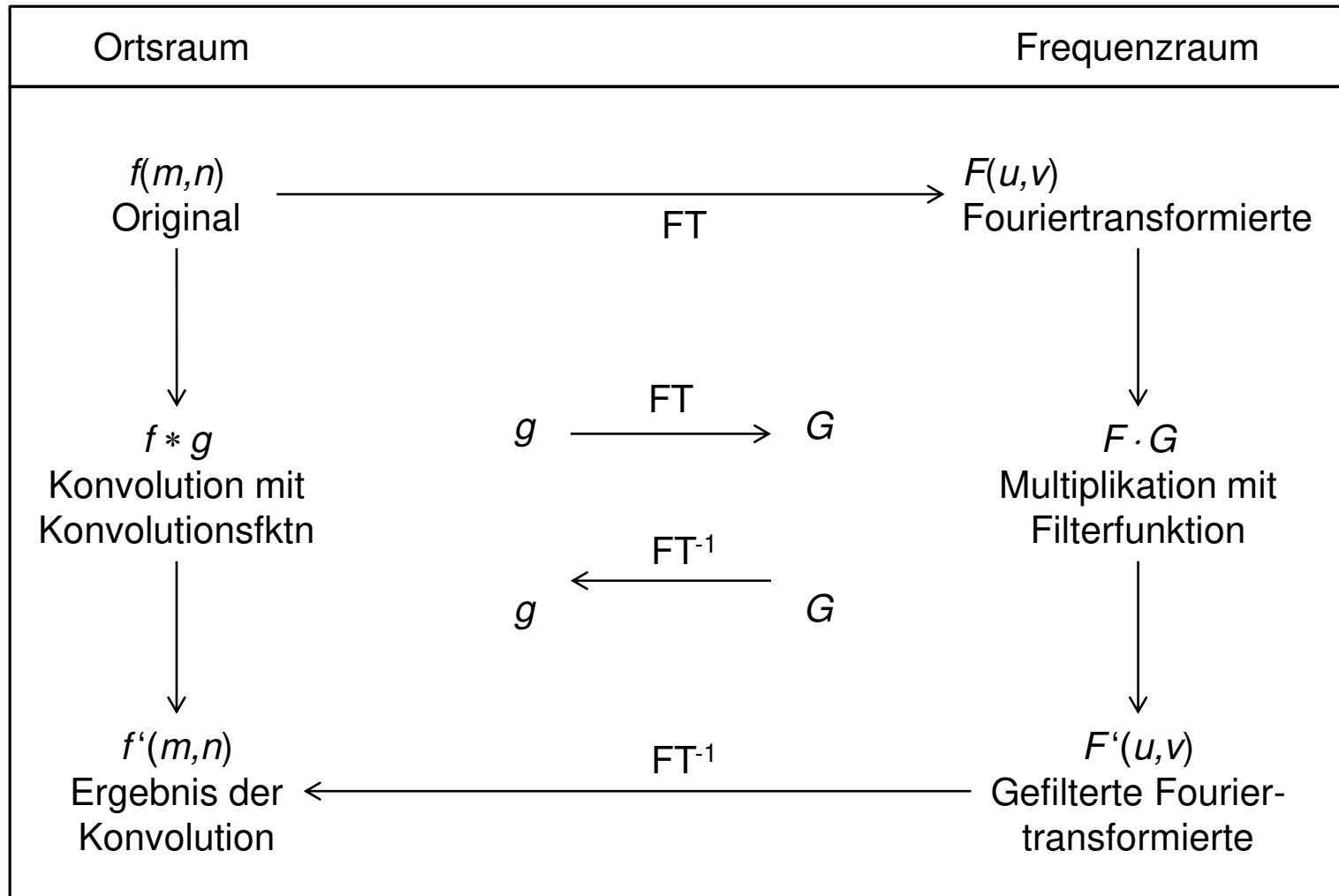
Bildquelle: Klaus Tönnies:
Grundlagen der Bildverarbeitung,
Pearson Studium,
2005.

Schnelle FT = Fast FT = FFT

- Geg.: Bild der Größe $N \times N$:
 - Berechnungsaufwand der FT für ein Pixel: $O(N^2)$
 - Berechnungsaufwand der FT für gesamtes Bild: $O(N^4)$
- Die FFT basiert auf einer Divide-and-Conquer-Strategie unter Nutzung der Periodizität ($F(u) = F(u + N)$, $f(n) = F(n + N)$), Symmetrie ($F(u) = {}^*F(-u)$, $f(u,v) = {}^*F(-u,-v)$) und Separabilität (2D-FT und $2D\text{-}FT^{-1}$ als Produkt von jeweils zwei 1D-FTs)
 - Berechnungsaufwand der FFT für ein Pixel: $O(\log N)$
 - Berechnungsaufwand der FFT für gesamtes Bild: $O(N^2 \log N)$

Konvolution und FT (1)

Die Fourier-Transformierte des Faltungsprodukts $f'(m,n) = f(m,n) * g(m,n)$ ist das Produkt der Fourier-Transformierten der einzelnen Funktionen (s. Anh.):



Konvolution und FT (2)

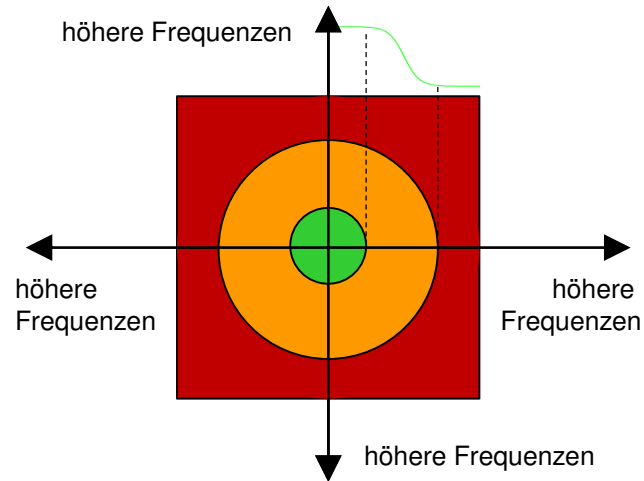
- Damit ist ein wichtiges Einsatzgebiet der FT angesprochen: Eine rechenintensive Faltung (Integration) im Ortsraum kann durch eine einfache Multiplikation im Frequenzraum durchgeführt werden
- Die Konvolution des Bildes der Größe $N \times N$ mit einem Faltungskern der Größe m (also $m \times m$ Gewichte) hat die Zeitkomplexität $O(m^2 \cdot N^2)$
- Über FFT ausgeführt braucht die Konvolution des Bildes $O(N^2 \cdot \log_2 N)$
- Für $\log_2 N < m^2$ ist die Konvolution also über FFT effizienter ausführbar
- Beispiele:
 - 1024×1024 -Bild: $\log_2 1024 = 10$, 512×512 -Bild: $\log_2 512 = 9$, 256×256 -Bild: $\log_2 256 = 8$
 - 3×3 -Filter: $3^2 = 9$, 5×5 -Filter: $5^2 = 25$, 7×7 -Filter: $7^2 = 49$, ...

Filterung im Frequenzbereich

- **Tiefpass:** lässt tiefe Frequenzen passieren, reduziert oder eliminiert hohe Frequenzen
- **Ideales Tiefpassfilter:** Alle Frequenzen kleiner/gleich dem *Cut Off* werden ohne Abschwächungen übernommen, alle anderen vollkommen unterdrückt
- **Hochpass:** lässt hohe Frequenzen passieren, reduziert oder eliminiert tiefe Frequenzen
- **Ideales Hochpassfilter:** Alle Frequenzen kleiner/gleich dem *Cut Off* werden ohne Abschwächungen übernommen, alle anderen vollkommen unterdrückt
- **Bandpass:** Kombination von Tiefpass und Hochpass

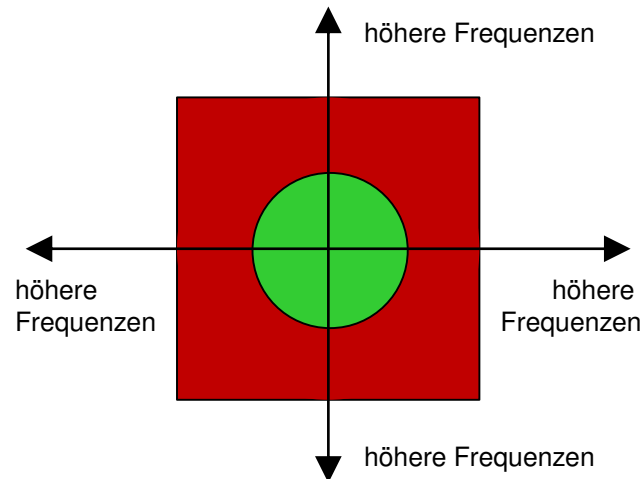
Tiefpass im Frequenzbereich

- Tiefpass:



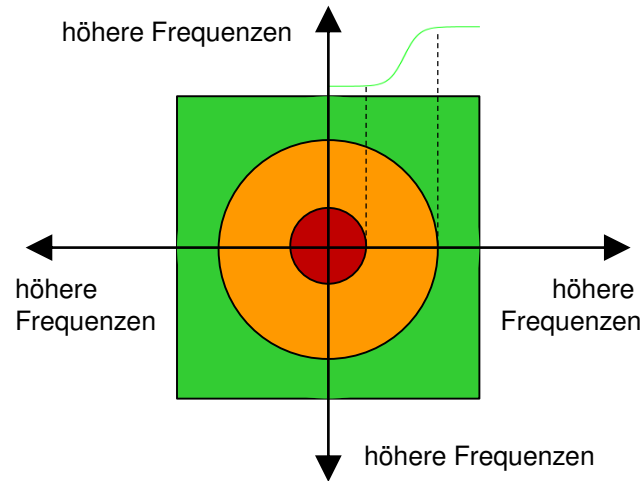
- Pass band
- Transition band
- Stop band

- Idealer Tiefpassfilter:



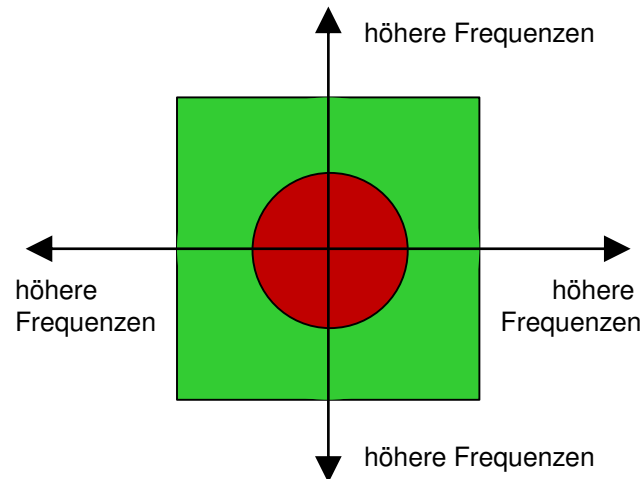
Hochpass im Frequenzbereich

- Hochpass:



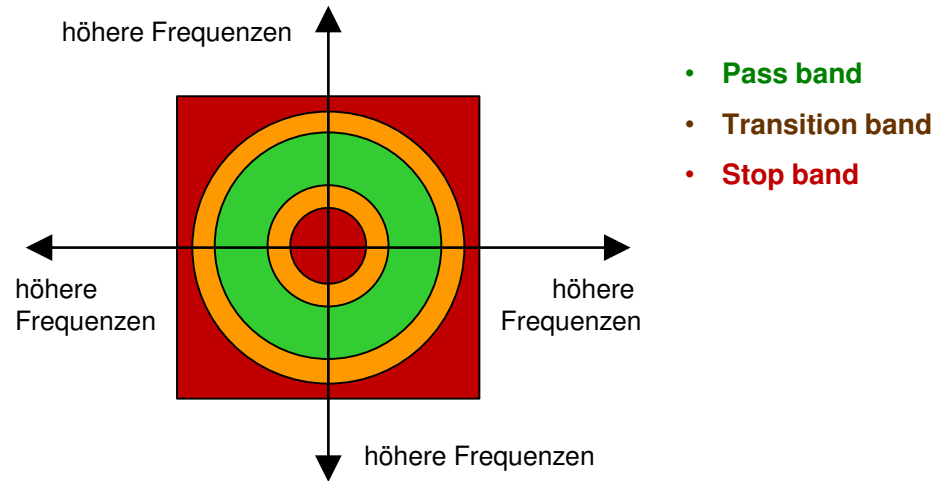
- **Pass band**
- **Transition band**
- **Stop band**

- Idealer Hochpassfilter:

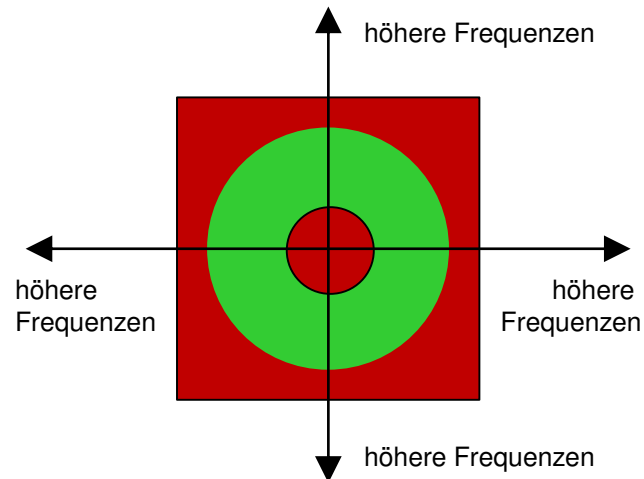


Bandpass im Frequenzbereich

- Bandpass:

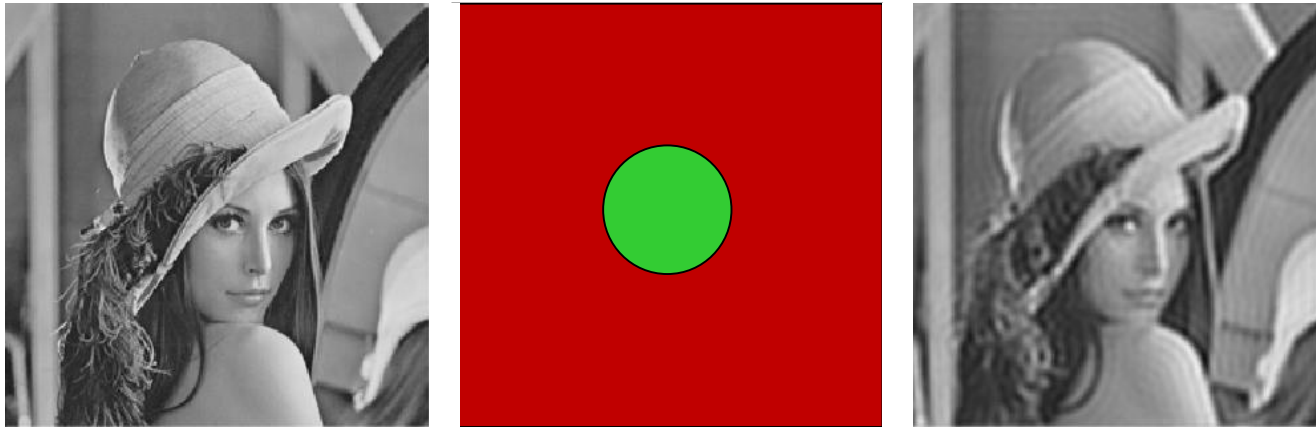


- Idealer Bandpassfilter:



Anwendung von Tiefpass

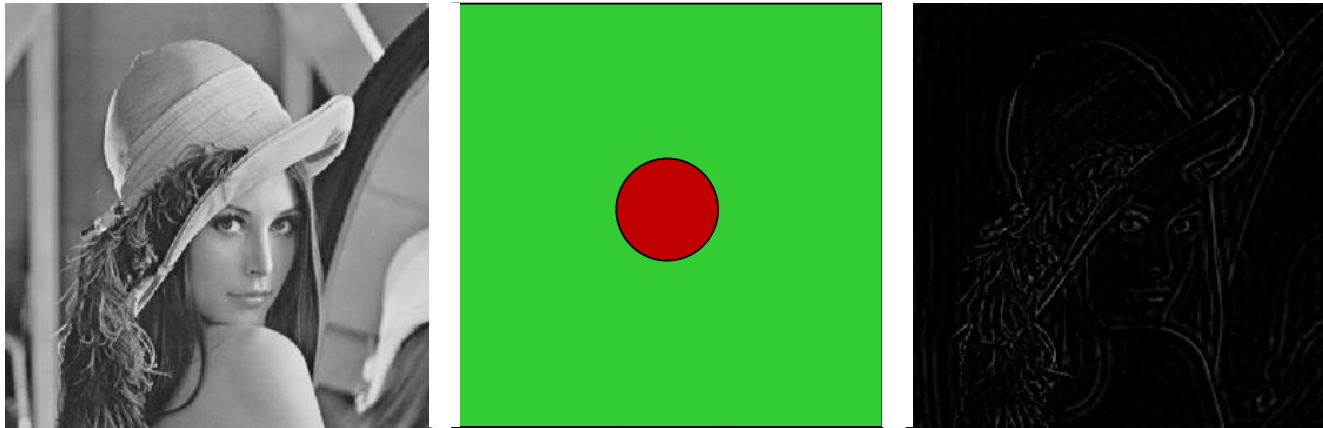
- Idealer Tiefpassfilter



- Tiefpassfilter sind glättend (entsprechen Glättungsfilttern im Ortsraum)
→ Verschmieren des Bildes und Abschwächen der Kanteninformation

Anwendung von Hochpass

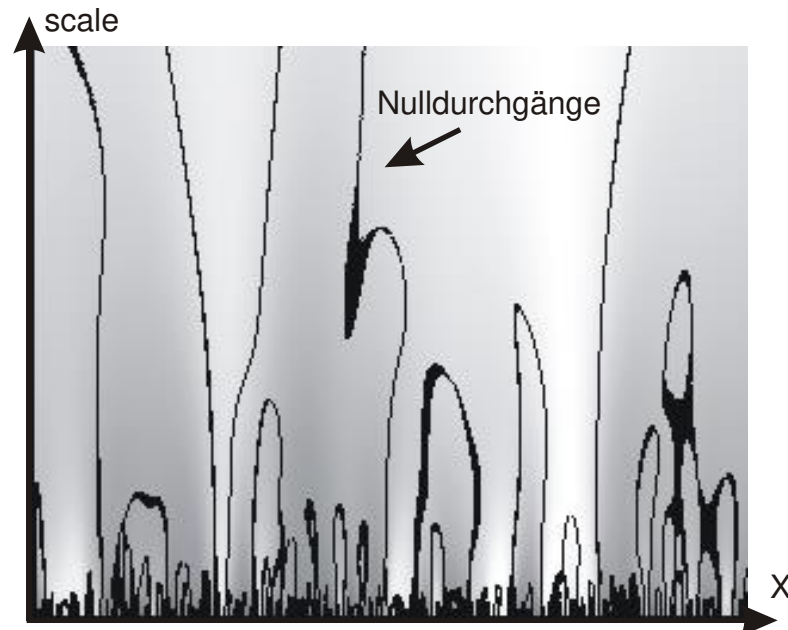
- Idealer Hochpassfilter



- Hochpassfilter heben Konturen hervor (entspr. Kantenfiltern im Ortsraum)
→ Hervorhebung der Kanteninformation

Multiskalenstrategien (1)

- Relative Kriterien für Homogenität (und damit auch Diskontinuitäten) können über unterschiedliche Entfernungen (Skalen) verschieden wirken.
- Segmentierung nach Multiskalenstrategie wertet Kriterien auf unterschiedlichen Skalierungen aus.



Multiskalenstrategien (2)

Multiskalenstrategie sind umsetzbar:

- auf einer expliziten Multiskalenrepräsentation

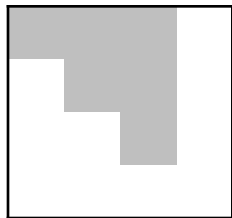
oder

- implizit durch Integration in den Segmentierungsalgorithmus

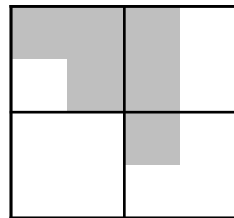
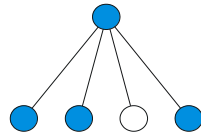
Implizit integrierte Multiskalenstrategien

Split-and-Merge (s. Vorl. 5) ist ein Beispiel für die **implizite Integration** einer Multiskalenstrategie in den Segmentierungsalgorithmus:

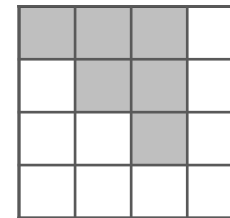
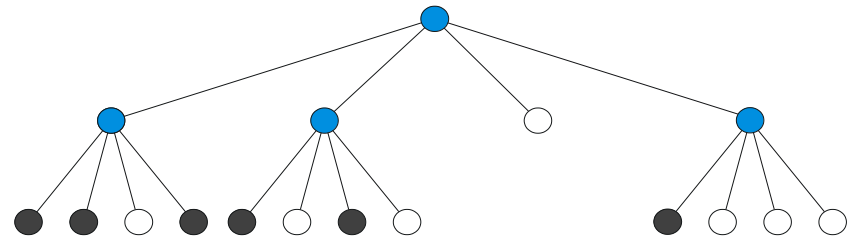
1. Skala:



2. Skala:



3. Skala:



Explizite Multiskalenrepräsentation (1)

- Expliziten Multiskalenrepräsentationen bilden das Originalbild in einen **Multiskalenraum** ab
- Die Idee des Multiskalenraums basiert auf der Erkenntnis, dass die menschliche Wahrnehmung Information in Detailebenen verwertet
- Anschaulich kann die Wahrnehmung von Detailebenen über unterschiedliche Entfernungen (Skalen) imaginiert werden

Die Seitenansicht desselben Autos wird bei unterschiedlichen Beobachtungsdistanzen unterschiedliche Detailebenen zeigen:



Seitenansicht eines PKWs in 307×108 bzw. 76×27 Pixeln (Bildquelle: Toyota Motor Corp.)

Explizite Multiskalenrepräsentation (2)

- Aber Ebenen unterschiedlicher Detailgenauigkeit sind oft über Modellwissen definiert. Die Wissen steht bei der Segmentierung i.A. nicht zur Verfügung.
- Ein Multiskalenraum ist ohne Modellwissen über spezifische Frequenzbänder bzw. spezifische Ebenen unterschiedlicher Auflösung definierbar.

Dazu können Glättungsfilter (im Ortsraum) bzw. Tiefpassfilter (im Frequenzraum) systematisch genutzt werden.

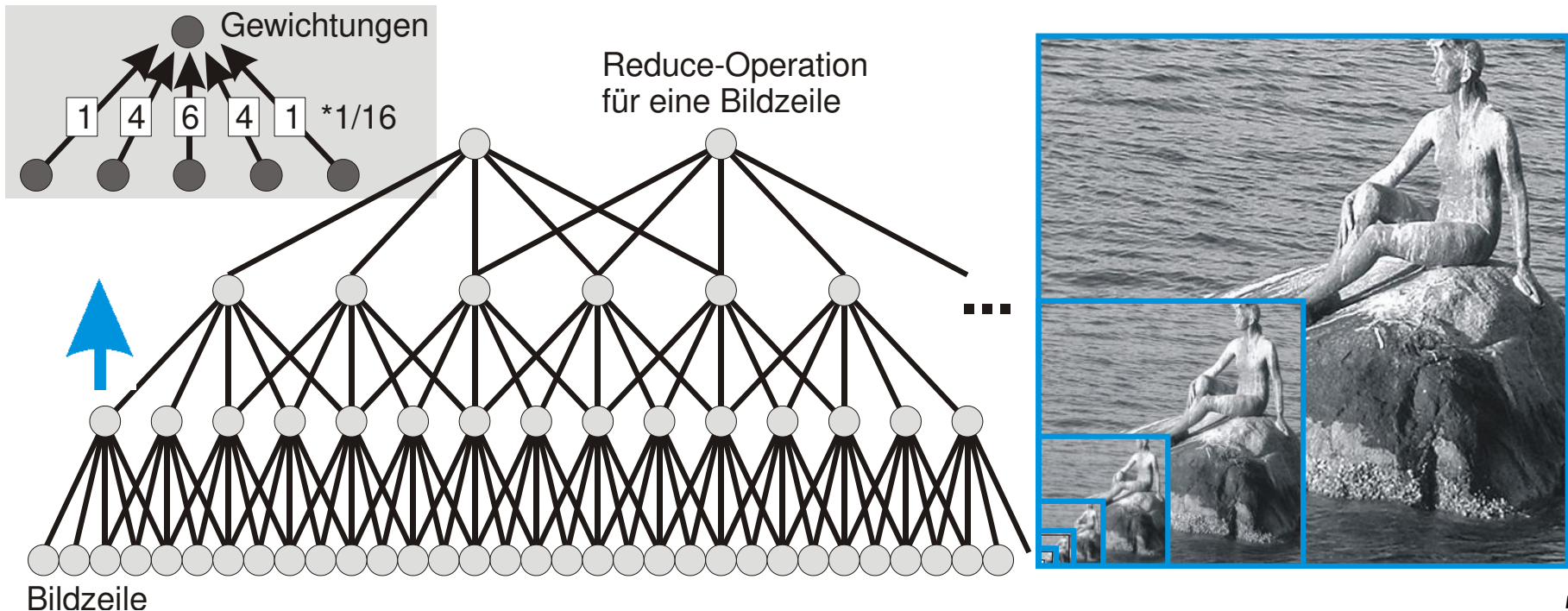


Seitenansicht eines PKWs in 307×108 bzw. 76×27 Pixeln sowie Ergebnis einer Tiefpassfilterung auf dem Originalbild (Bildquelle: Toyota Motor Corp.)

Gauß-Pyramide (1)

Eine Gauß-Pyramide ist eine Möglichkeit zur Abbildung eines Bildes in einen Multiskalenraum:

- Das Originalbild wird fortlaufend durch eine „**reduce**“-Operation in seiner Anzahl von Pixeln in jeder Spalte und jeder Zeile halbiert
- Jedes Pixel der nächsten Skalierungsstufe repräsentiert 4 Pixel der aktuellen Stufe

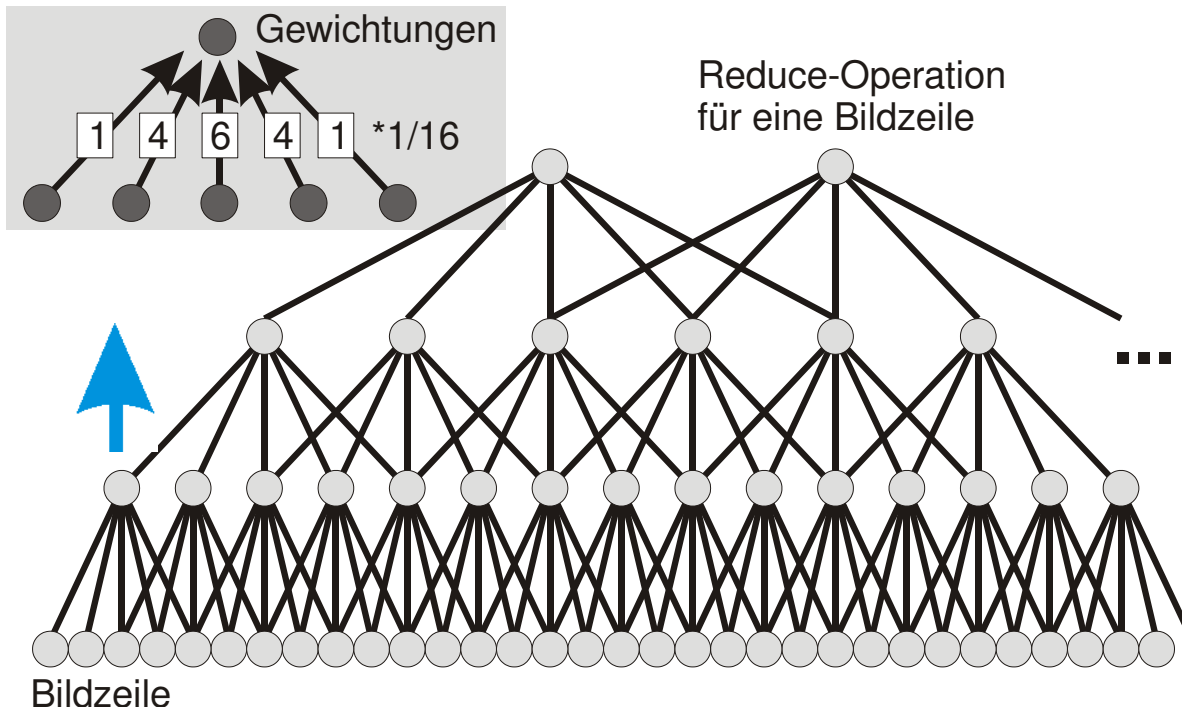


Gauß-Pyramide (2)

Vor der Reduktion wird der Frequenzumfang durch Filterung vermindert. Die Gauß-Filterung mit $\sigma = 1$ (bzw. deren Approx.) ist in der Praxis bewährt:

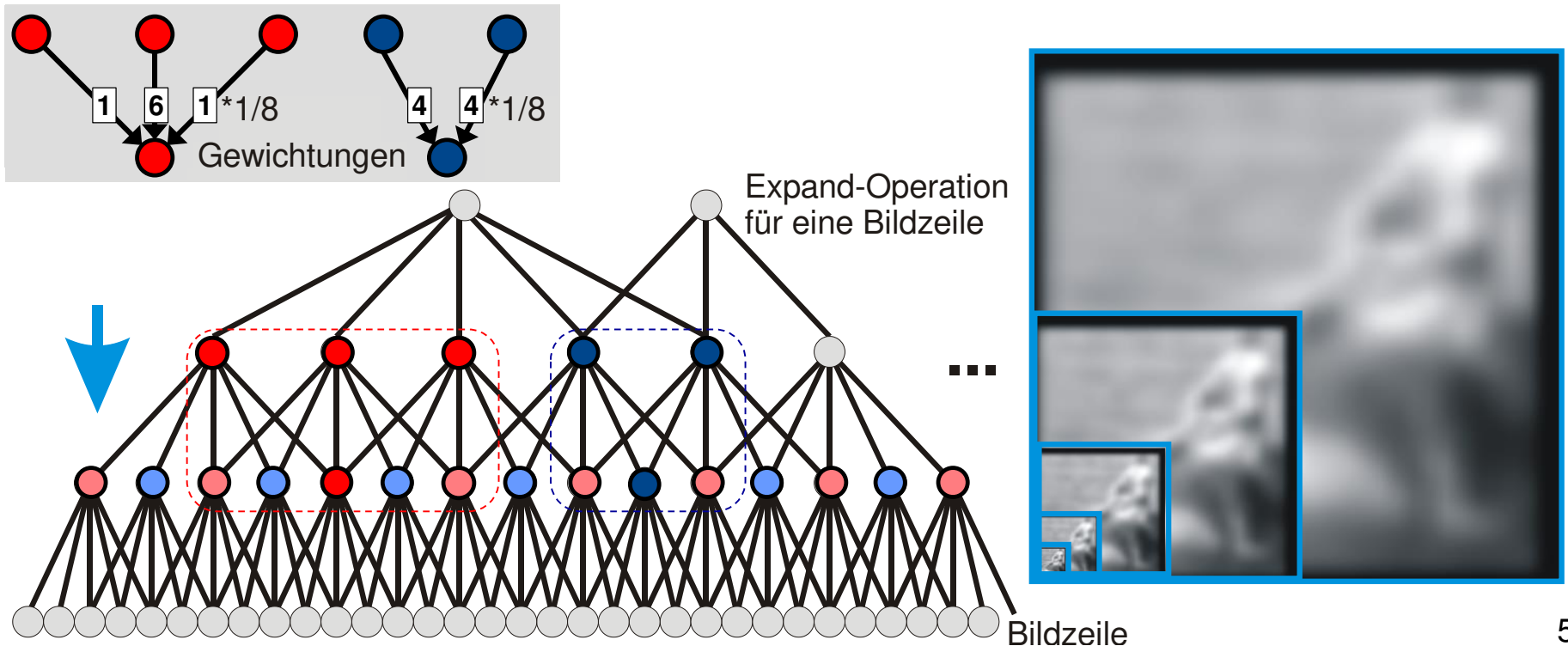
Gaußfilter: $\frac{1}{16} (0.87 \quad 3.91 \quad 6.44 \quad 3.91 \quad 0.87)$

Binomialfilter: $\frac{1}{16} (1 \quad 4 \quad 6 \quad 4 \quad 1)$



Gauß-Pyramide (3)

- Für Vergleiche verschiedener Skalierungsstufen müssen reduzierte Stufen wieder auf vorherige Stufen abbildbar sein
- Um die vorherige Skalierungsstufe aus der aktuellen Stufe zu erzeugen, wird daher eine „**expand**“-Operation definiert



Gauß-Pyramide (4)

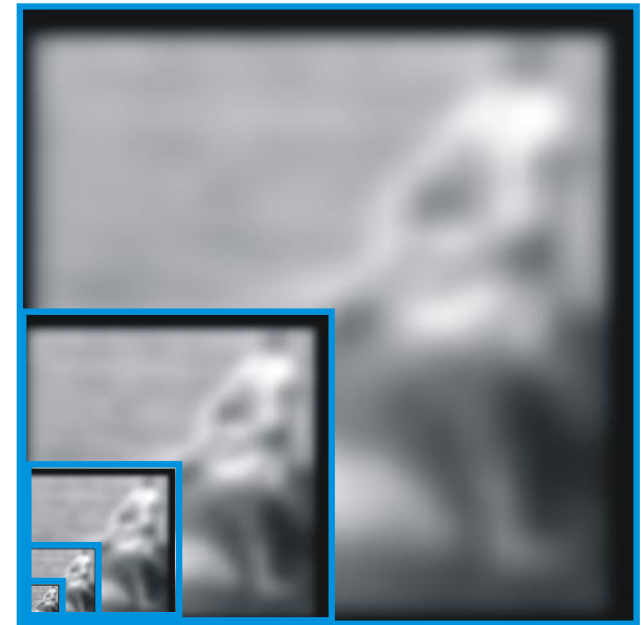
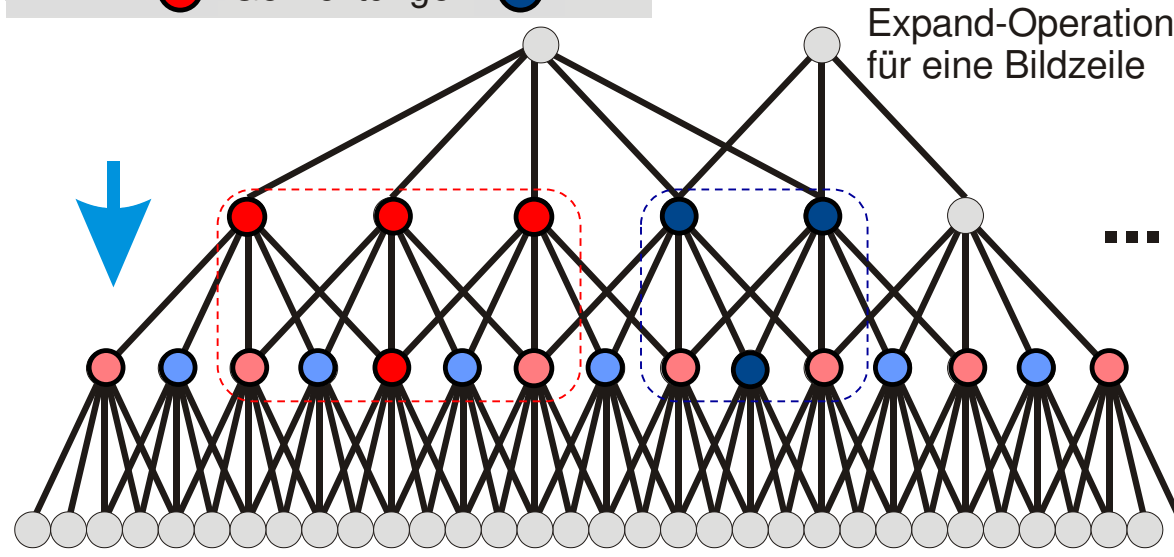
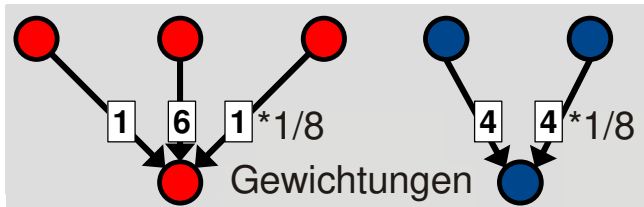
- Pixel der neuen Skalierungsstufe werden durch Interpolation erzeugt:

- Pixelorte, die auf beiden Skalierungsstufen existieren:

$$\frac{1}{8.18} \begin{pmatrix} 0.87 & 6.44 & 0.87 \end{pmatrix} \text{ bzw. } \frac{1}{8} \begin{pmatrix} 1 & 6 & 1 \end{pmatrix}$$

- Pixelorte, die nur auf der vorherigen Skalierungsstufe existieren:

$$\frac{1}{7.82} \begin{pmatrix} 3.91 & 3.91 \end{pmatrix} \text{ bzw. } \frac{1}{8} \begin{pmatrix} 4 & 4 \end{pmatrix}$$

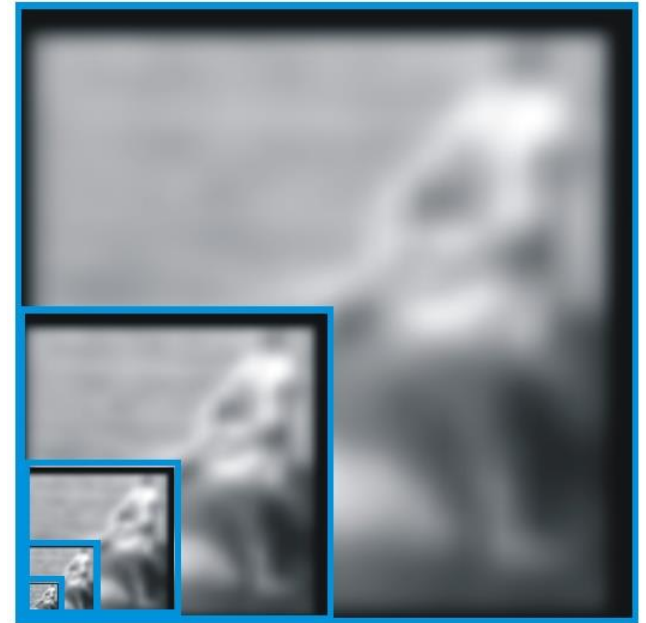


Gauß-Pyramide (5)

Der Aufbau der Gauß-Pyramide kann also wie folgt formal beschrieben werden:

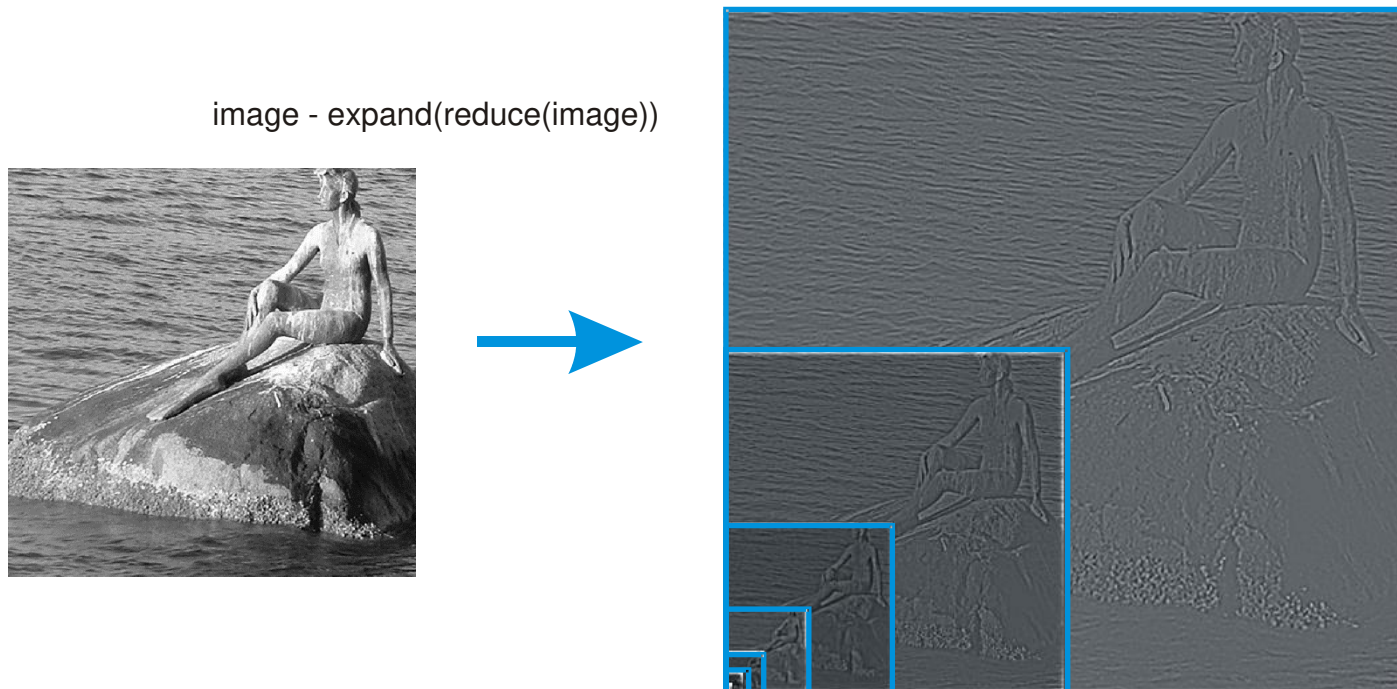
\mathbf{G}_0 = Originalbild $\mathbf{I}[x,y]$

\mathbf{G}_{k+1} = $Reduce(\mathbf{G}_k)$
für Stufen $k = 0, 1, \dots, r$



Laplace-Pyramide (1)

Eine nahezu redundanzfreie Repräsentation erhält man, wenn auf jeder Skalenstufe k nicht das reduzierte Bild \mathbf{G}_k , sondern das Differenzbild zwischen der Expansion des nächstreduzierten Bildes und G_k selbst gespeichert wird:



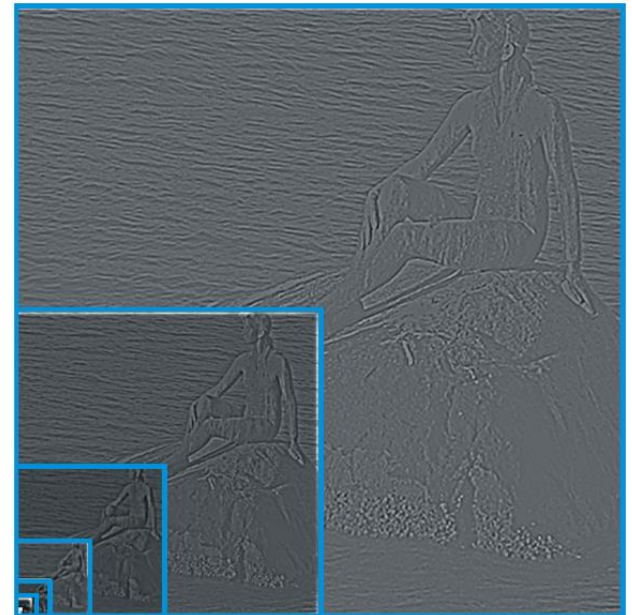
Laplace-Pyramide (2)

Der Aufbau der Laplace-Pyramide kann also wie folgt formal beschrieben werden:

$$\mathbf{L}_r = \mathbf{G}_r$$

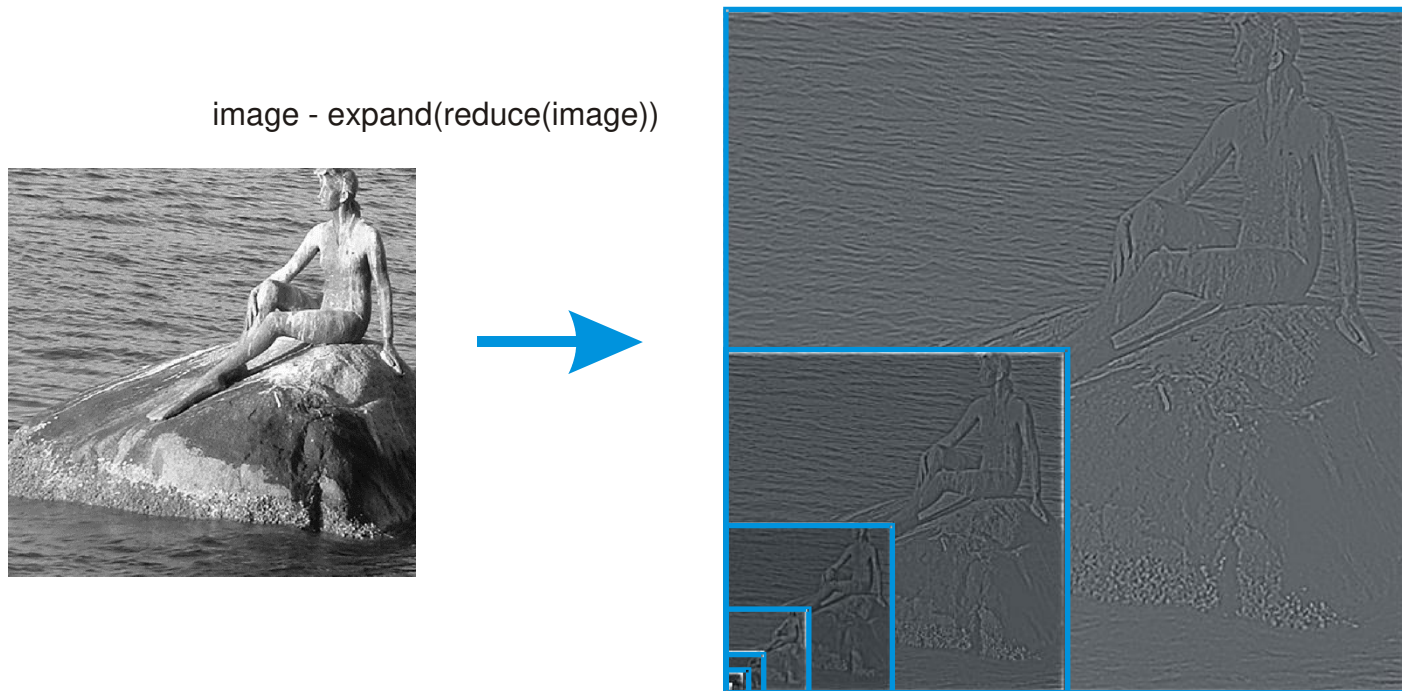
$$\mathbf{L}_k = \mathbf{G}_k - \text{Expand}(\mathbf{G}_{k+1})$$

für Stufen $k = 0, 1, \dots, r-1$



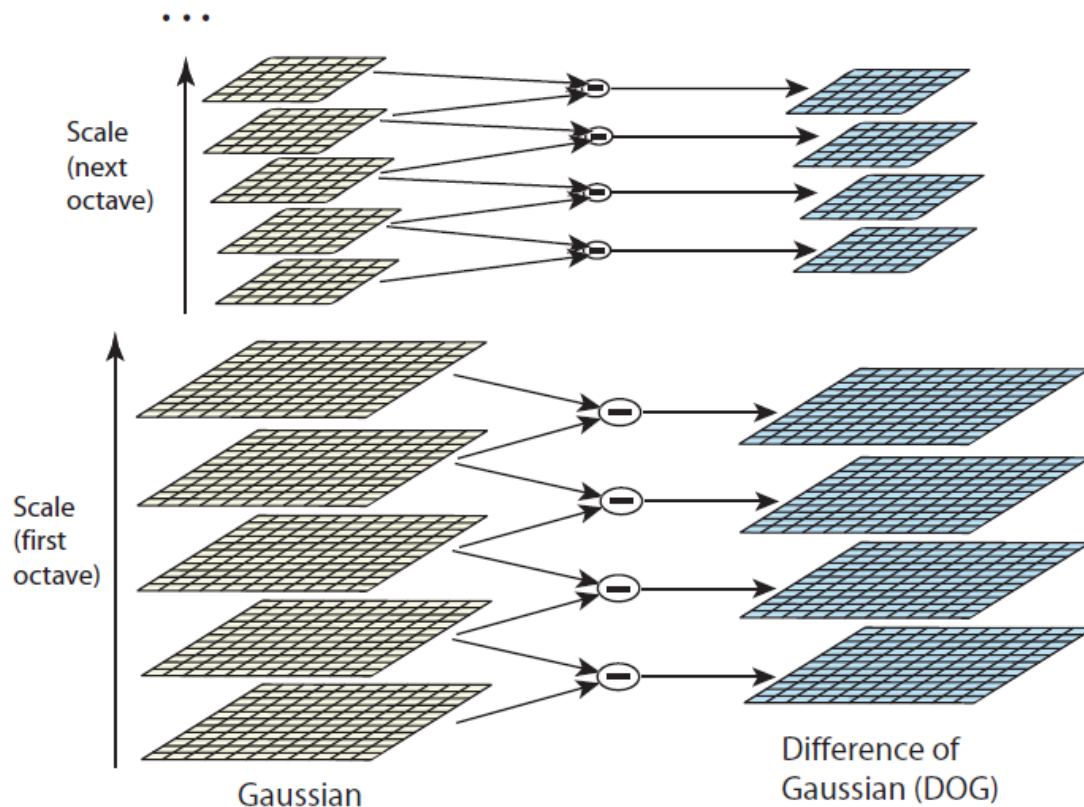
Laplace-Pyramide (3)

Faktisch wird hier der **Difference-of-Gaussian-Operator (DoG)** als Approximation des **Laplacian-of-Gaussian-Filter (LoG-Filter)** eingesetzt (s. Vorlesung 3).



Laplace-Pyramide (4)

- Gauß- und DoG-Hierarchien bilden wiederum die Basis von skaleninvarianten Ansätzen für das Finden und Beschreiben von sog. *Interest Points*. Einen grundlegenden Ansatz dazu stellt die Arbeit von D. Lowe dar.



Zusammenfassung (1)

- (Diskrete) Bildfunktionen (und damit Bilder) sind durch Basisfunktionen im Orts- und Frequenzraum darstellbar.
- Die diskrete Fourier-Transformation setzt periodische komplexe Basisfunktionen der Form $\cos(u \cdot n) + i \cdot \sin(u \cdot n)$ ein.
- Die Konvolution von Bildern mit Glättungs- und Kantenfilter im Ortsraum kann durch Multiplikation der entspr. Fourier-transformierten Bildfunktionen effizient umgesetzt werden.

Zusammenfassung (2)

- Multiskalen-Strategien versuchen, Merkmale oder Segmente in unterschiedlichen Frequenzbändern bzw. Auflösungsstufen zu ermitteln. Dies kann implizit erfolgen wie etwa in *Split-and-Merge* und der *hierarch. Wasserscheiden-transformation*.
- Andere Multiskalen-Strategien setzen explizite Repräsentation von Multiskalenräumen ein.
- Die bekanntesten Ansätze zur expliziten Repräsentation von Multiskalenräumen setzen Gauß-Pyramiden und Laplace-Pyramiden ein.

Anhang: Konvolution und FT

Multiplikation im Frequenzraum entspricht Konvolution im Ortsraum:

$$\begin{aligned} F(u) \cdot G(u) &= \sum_k f(k) \cdot \exp(-i2\pi uk/N) \cdot \sum_m g(m) \cdot \exp(-i2\pi um/N) \\ &= \sum_m \sum_k [f(k) \cdot \exp(-i2\pi uk/N) \cdot g(m) \cdot \exp(-i2\pi um/N)] \\ &= \sum_m \sum_k [f(k) \cdot g(m) \cdot \exp(-i2\pi uk/N) \cdot \exp(-i2\pi um/N)] \end{aligned}$$

(Verschiebeeigenschaft $g(m) \cdot \exp(-i2\pi uk/N) = g(m-k)$ in obigem Term)

$$\begin{aligned} &= \sum_m \sum_k f(k) \cdot g(m-k) \cdot \exp(-i2\pi um/N) \\ &= \sum_m \left[\sum_k f(k) \cdot g(m-k) \right] \cdot \exp(-i2\pi um/N) \\ &= \mathbf{FT} \left[\sum_k f(k) \cdot g(m-k) \right] = \mathbf{FT} [f * g](m) \end{aligned}$$