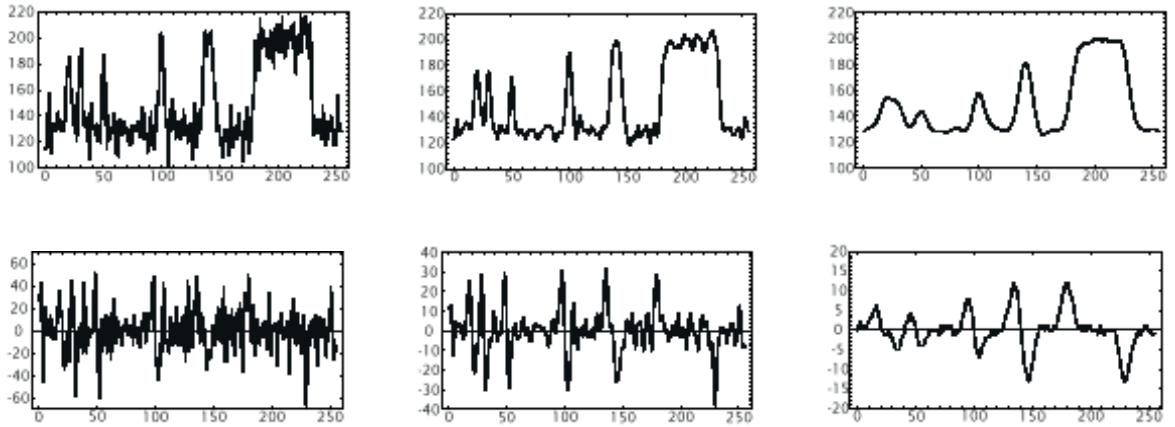
Multiskalenrepräsentation

Multiskalenrepräsentation

Einführung

- Warum Multiskalenrepräsentation?
  - Es gibt Merkmale, die sich nur aus großen Teilen eines Bildes extrahieren lassen, andere wirken sich nur in einer kleinen Umgebung aus.



Bilder aus Bernd Jähne, Digitale Bildverarbeitung, Springer 1997

Multiskalenrepräsentation

Prinzip der Multiskalenrepräsentation

- Bilderstapel mit verschieden gefilterten Versionen des Bildes (Parameter σ).
- Glättungsfilter sei Gaußfilter (Begründung später):

$$g(x, y, \sigma) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}}$$



Erzeugung von Skalenräumen

Erzeugung von Skalenräumen durch Diffusion

- Für die Flußdichte gilt:

$$\mathbf{j} = -D \operatorname{grad} f = -D\nabla f$$

- Die Kontinuitätsgleichung lautet:

$$\frac{\partial f}{\partial t} + \nabla \mathbf{j} = \frac{\partial f}{\partial t} + \operatorname{div} \mathbf{j} = 0$$

- Für homogene Diffusionsprozesse (D=const.) erhalten wir:

$$\frac{\partial f}{\partial t} = D \operatorname{div} \operatorname{grad} f = D\Delta f = D\left(\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2}\right)$$

Erzeugung von Skalenräumen

Fouriertransformation führt zu:

$$\frac{\partial F(\mathbf{u})}{\partial t} = -D\,|\mathbf{u}|^2\,F(\mathbf{u})$$

Die Lösung dieser Differentialgleichung ist:

$$F(\mathbf{u},t) = e^{-D|\mathbf{u}|^2 t} F(\mathbf{u},0)$$

bzw. mit $\xi = 2Dt$:

$$F(\mathbf{u},\xi) = e^{\frac{-|\mathbf{u}|^2 \xi}{2}} F(\mathbf{u},0)$$

Erzeugung von Skalenräumen

Dies entspricht einer Faltung des Originalbildes mit:

$$h(\mathbf{x},t) = \frac{1}{4\pi D t} e^{-\frac{|\mathbf{x}|^2}{4Dt}}$$

bzw. mit $\xi = 2Dt$:

$$h(\mathbf{x},\xi) = \frac{1}{2\pi\xi} e^{-\frac{|\mathbf{x}|^2}{2\xi}}$$

also

$$f(\mathbf{x},\xi) = h(\mathbf{x},\xi) \circ f(\mathbf{x},0)$$

Erzeugung von Skalenräumen

Vorteile dieser Methode:

- Keine neuen lokalen Minima/Maxima werden erzeugt
  - lokales Minimum wird nicht kleiner.
  - lokales Maximum wird nicht größer.
- Skaleninvarianz: Start von jedem beliebigen Startwert aus führt zu identischen Ergebnissen:

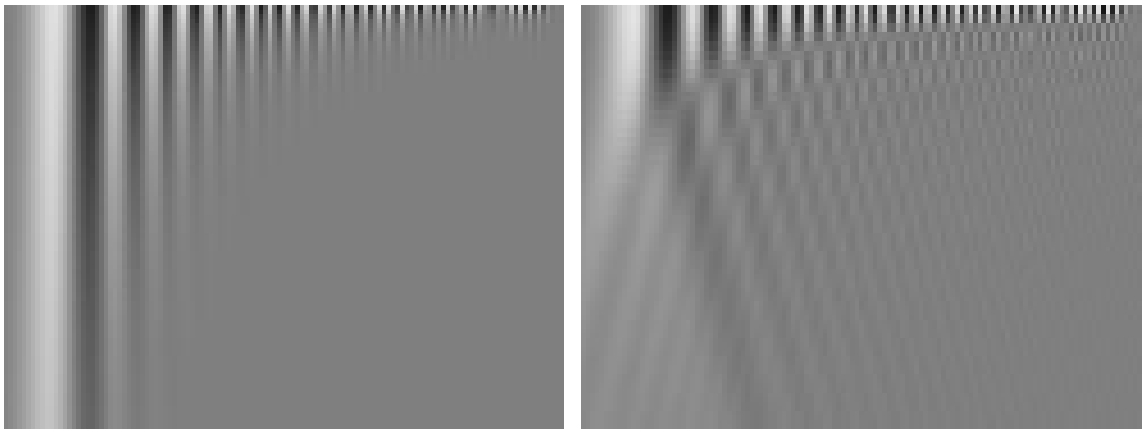
  wenn $\xi = \xi_1 + \xi_2$

  dann gilt: $f(\mathbf{x}, \xi) = h(\mathbf{x}, \xi_1) \circ h(\mathbf{x}, \xi_2) \circ f(\mathbf{x}, 0)$

Erzeugung von Skalenräumen

Vergleich mit (nicht so gut geeignetem) Boxfilter:

- Links: Glättung mit Gaußfilter
- Rechts: Glättung mit Rechteckfilter (Boxfilter):



Bilder aus Bernd Jähne, Digitale Bildverarbeitung, Springer 1997

Erzeugung von Skalenräumen

Nachteil: σ wächst nur mit der Wurzel der Zeit.

Lösung:

- Quadratischer Skalenraum:
  - Standardabweichung für die Glättung wächst linear mit der Zeit, der Skalenparameter ist proportional zum Quadrat der Zeit.

- Logarithmischer Skalenraum:
  - Skalenparameter ist proportional zu e^t

Differentielle Skalenräume

Wir betrachten die Bildänderungen, die sich als Folge der Änderung des Skalenparameters ergeben. Das Bild:

$$F(\mathbf{u},\xi) = e^{\frac{-|\mathbf{u}|^2 \xi}{2}} F(\mathbf{u},0)$$

wird nach ξ abgeleitet:

$$\frac{\partial F(\mathbf{u},\xi)}{\partial \xi} = \frac{-|\mathbf{u}|^2}{2} e^{\frac{-|\mathbf{u}|^2 \xi}{2}} F(\mathbf{u},0) = \frac{-|\mathbf{u}|^2}{2} F(\mathbf{u},\xi)$$

Differentielle Skalenräume

Im Ortsraum entspricht dies der Anwendung des Laplaceoperators:

$$\frac{\partial f(\mathbf{x},\xi)}{\partial \xi} = \frac{1}{2}\Delta f(\mathbf{x},\xi)$$

Zusammen mit der Tiefpaßfilterung entspricht das einer Bandpaßfilterung des Originalbildes mit der Transferfunktion:

$$H(\mathbf{u}) = \frac{-|\mathbf{u}|^2}{2} e^{\frac{-|\mathbf{u}|^2 \xi}{2}}$$

Diskrete Skalenräume

Werden die Filter diskret realisiert, erhält man unter bestimmten Bedingungen einen diskreten Skalenraum, der die o.g. Eigenschaften erfüllt. (Minima-Maxima erhaltend, Skaleninvarianz).
Diskrete differenzielle Skalenräume erhält man entsprechend durch Ersetzen der Ableitungen durch diskrete Differenzen.

Mehrgitterrepräsentation

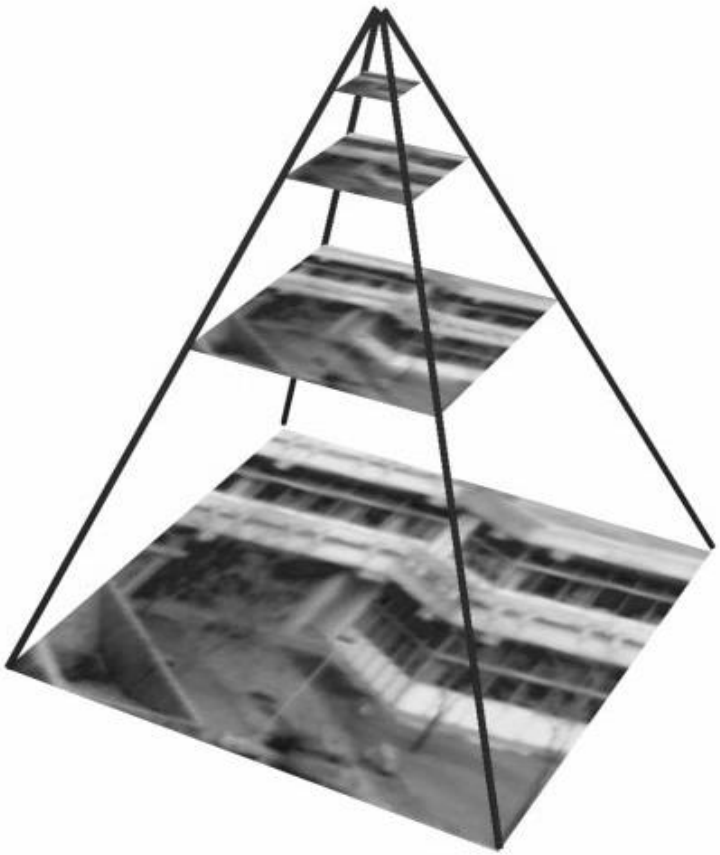
Prinzip: Grobe Abtastung bei großen Skalen.
Beispiele für Auflösungspyramiden

- Gaußpyramide
- Laplacepyramide

Gaußpyramide

Gaußpyramide:

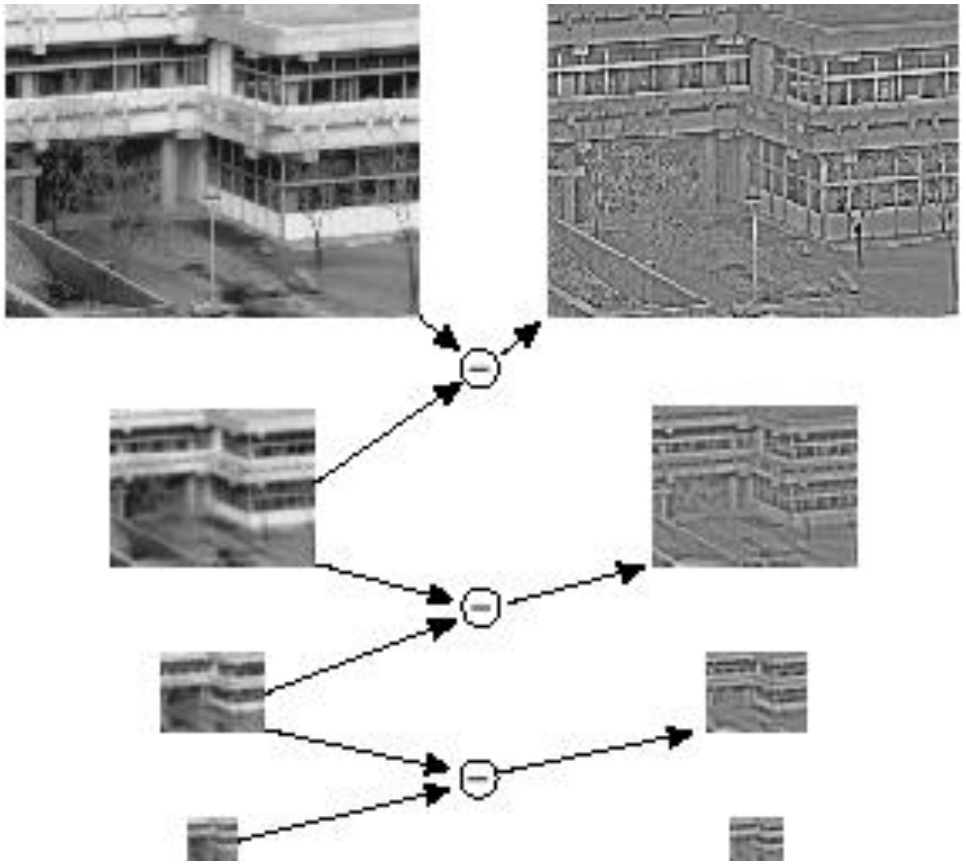
Filterung und Vergröberung der Abtastung als Rekursionsschritt



Bilder aus Bernd Jähne, Digitale Bildverarbeitung, Springer 1997

Laplacepyramide

Differentieller Skalenraum zu Gaußpyramide. Berechnung aus Gaußpyramide durch Expansion und Subtraktion.



Bilder aus Bernd Jähne, Digitale Bildverarbeitung, Springer 1997

Auflösungspyramiden

Eigenschaften:

- Speicherbedarf:

$$n \times m(1+\frac{1}{4}+\frac{1}{16}+\frac{1}{64}+...) < \frac{4}{3} n \times m$$

- Rechenbedarf (Anwendung der gleichen Operation auf immer kleinere Bilder): s.o.

- Rechenbedarf: auch für großskalige Operationen klein durch kleine Kerne.

Auflösungspyramiden

Anwendung: Feature-Detection
Verschiedene Features werden bei verschiedenen Auflösungen detektiert:

Auflösungspyramiden

Detektion von Kanten:

(u,v)-System:

u: Richtung senkrecht zum Gradienten

v: Richtung parallel zum Gradienten

Dann Bedingung f. Kante: lokales Maximum des Gradienten in v-Richtung:

$$L_{vv} = 0$$
$$L_{vvv} < 0$$

Auflösungspyramiden

Detektion von „Bergkämmen“ (ridges):

(p,q)-System:

p: erste Hauptkrümmungsrichtung

q: zweite Hauptkrümmungsrichtung

$$L_{pq} = 0 \; ; \; |L_{pp}| \geq |L_{qq}|$$

Dann Bedingung f. Kamm: lokales Maximum der Helligkeit in p-Richtung:

$$L_p = 0$$
$$L_{pp} < 0$$

Auflösungspyramiden

Verschiedene Features werden bei verschiedenen Auflösungen detektiert:

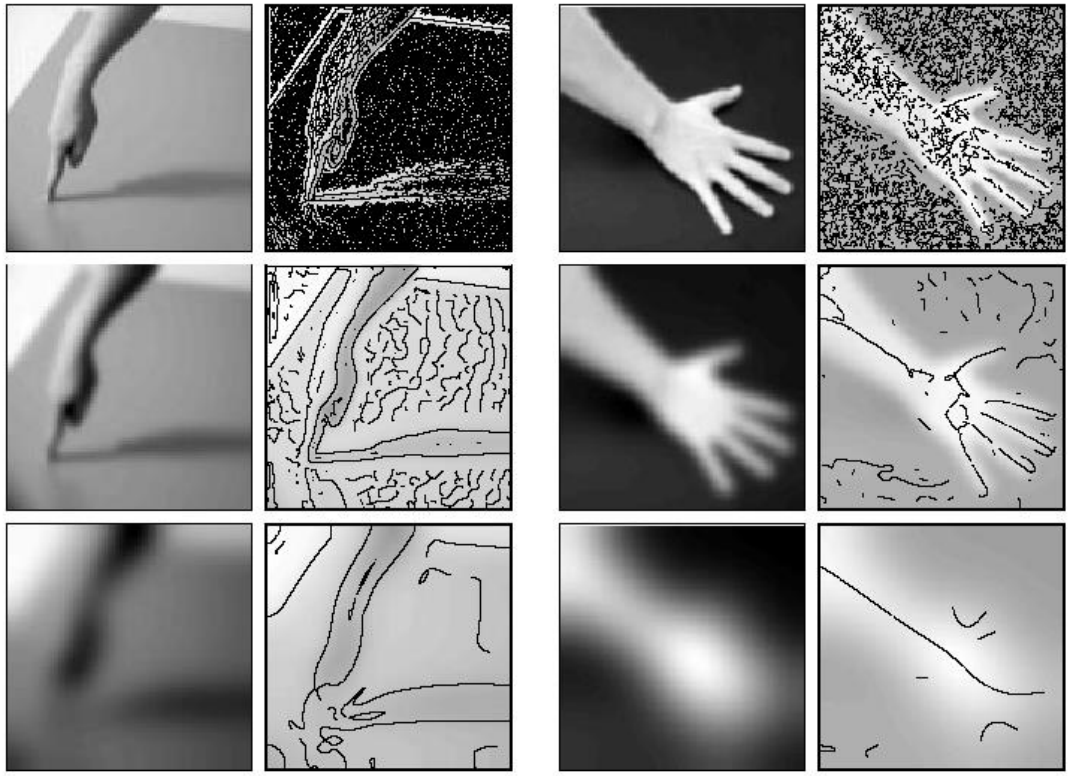


Figure 5: Edges and bright ridges detected at scale levels $t = 1.0$, 16.0 and 256.0, respectively.

Bilder aus T. Lindeberg, Scale-Space: A framework for handling image structures at multiple scales.

Auflösungspyramiden

Anwendung:

- Analyse von Bildern mit automatischer Wahl der Auflösung
  - Verwendung von normierten Ableitungen:

$$\partial_{x_i'} = \xi^{\frac{\gamma}{2}} \partial_{x_i}$$

  - γ wird dabei von Hand der Aufgabe angepaßt.
  - Detektor wird in allen Auflösungen berechnet.
  - Für jeden Punkt wird Auflösungsstufe mit maximaler Response gewählt.

Auflösungspyramiden

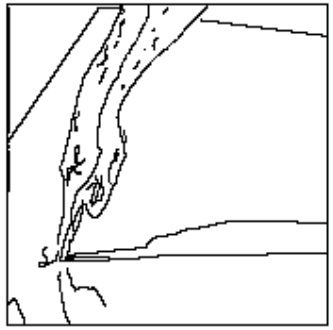
– Beispiele von Lindeberg:

Feature type	Normalized strength measure for scale selection	Value of γ
Edge	$t^{\gamma/2}\, L_v$	1/2
Ridge	$t^{2\gamma}(L_{pp} - L_{qq})^2$	3/4
Corner	$t^{2\gamma} L_v^2 L_{uu}$	1
Blob	$t^{\gamma} \nabla^2 L$	1





Figure 6: Results of edge detection with automatic scale selection on the arm image (all edges).



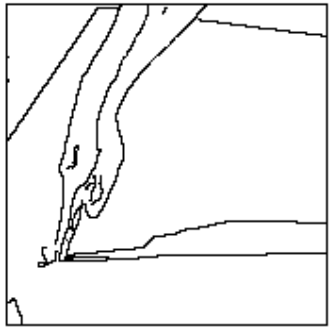


Figure 7: The 50 and 10 most significant edge curves as ranked on the integrated measure of edge strength along each connected edge curve.

Ableitungsoperatoren

Ableitungsoperatoren sind lineare Operatoren

- Beispiel die Ableitung $\frac{\partial}{\partial x}$

entspricht im Ortsfrequenzraum der Faltung mit

$$h(x) = \frac{1}{2\varepsilon} \lim_{\varepsilon \to 0} \delta(x - \varepsilon) - \delta(x + \varepsilon)$$

und im Ortsfrequenzraum der Multiplikation mit

$$H(u) = iu$$

Ableitungsoperatoren

- Die Ableitung zweiter Ordnung $\frac{\partial^2}{\partial x^2}$

entspricht im Ortsraum der Faltung mit

$$h(x) = \frac{1}{\varepsilon^2} \lim_{\varepsilon \to 0} \delta(x - \varepsilon) - 2\delta(x) + \delta(x + \varepsilon)$$

und im Ortsfrequenzraum der Multiplikation mit

$$H(u) = -u^2$$

Kantendetektion

- Ableitungen erster Ordnung sind richtungsabhängig - Betragsbildung nötig (nicht linear).
- Dagegen ist der Laplaceoperator richtungsunabhängig.
- Problem: Rauschen
  Abhilfe: Glätten
  Also: Glätten mit Gauss, dann ∇^2 (Laplace)

Laplacian of Gaussian (LoG)

Glätten mit Gauß:

OR: Faltung	$g(x) = \frac{1}{2\pi\sigma^2} e^{-\frac{\mathbf{x}^2}{2\sigma^2}}$
FR: Multiplikation	$G(\mathbf{u}) = e^{-\sigma^2\mathbf{u}^2}$

Laplaceoperator:

OR: Ableitung	$\Delta = \nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}$
FR: Multiplikation	$G(\mathbf{u}) = -4\pi^2\mathbf{u}^2$

Laplacian of Gaussian (LoG)

Zusammenfassung (Vertauschen bei Multiplikation möglich):

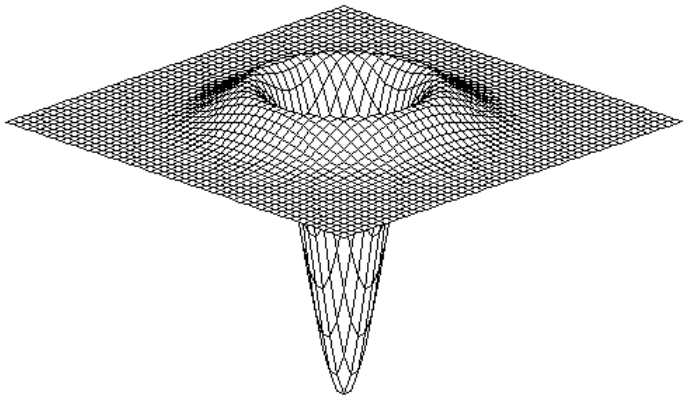
FR: Multiplikation $H(\mathbf{u}) = -4\pi^2\mathbf{u}^2 e^{-\sigma^2\mathbf{u}^2}$

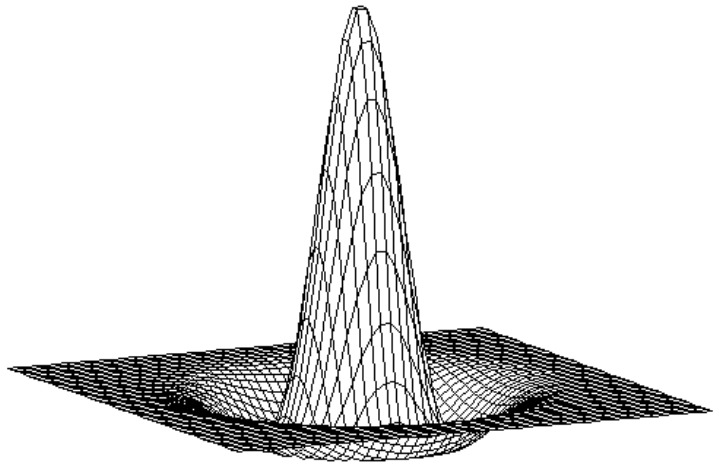
OR: Faltung $h(x) = \nabla^2 \frac{1}{2\pi\sigma^2} e^{-\frac{\mathbf{x}^2}{2\sigma^2}}$

$$= \frac{1}{2\pi\sigma^4}\left(\frac{\mathbf{x}^2}{\sigma^2} - 1\right) e^{-\frac{\mathbf{x}^2}{2\sigma^2}}$$

Laplacian of Gaussian (LoG)

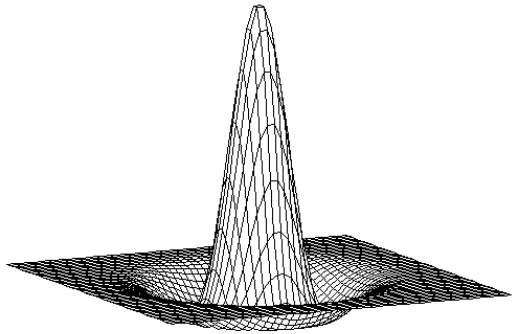
Mexican Hat:





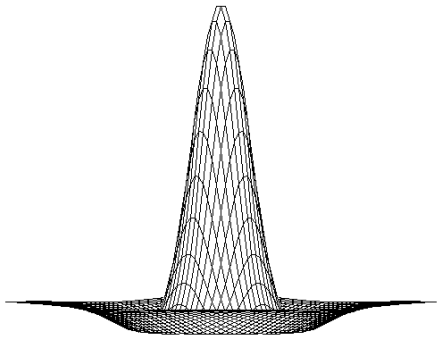
Laplacian of Gaussian (LoG)

Das ist ein
Mexican Hat:



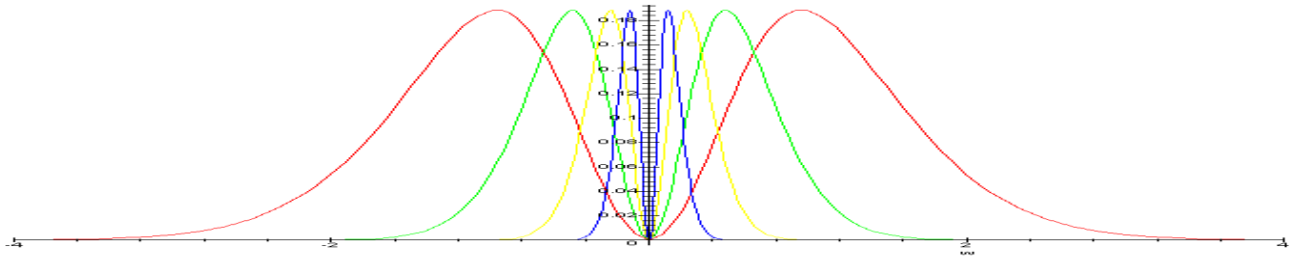
- Was ist das?

DoG!

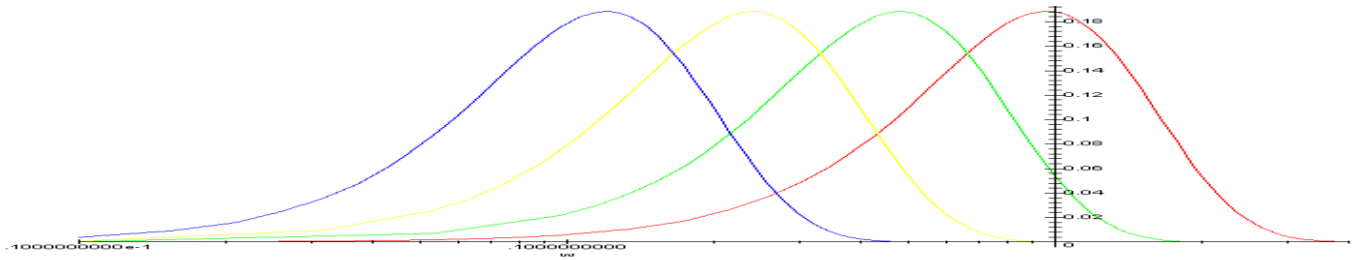


Difference of Gaussians (DoG)

DoG im Frequenzraum (Oktaven der Laplace-Pyramide, s.o., Betrag normiert):



in logarithmischer Darstellung:



Literatur

Bernd Jähne, Digitale Bildverarbeitung, Springer 1997

Tony Lindeberg: Scale-space theory: A framework for handling image structures at multiple scales
In: Proc. CERN School of Computing, Egmond aan Zee, The Netherlands, Sep. 8--21, 1996.
ftp://ftp.nada.kth.se/CVAP/reports/lin-cern-summ-school-96.pdf