

Discrete Scale-Space Formulation and Multiscale Edge Extraction Toward Higher Dimensions

Ji-Young Lim

Universität Hamburg
Fachbereich Informatik
Arbeitsbereich Kognitive Systeme

lim@informatik.uni-hamburg.de

Abstract: In den letzten Jahren hat sich die Skalenraum-Theorie in den verschiedensten Anwendungen bewährt. Die in der Literatur vorherrschenden Multiskalen-Verfahren zur Kantenextraktion basieren auf der linearen Skalenraum-Theorie, sind jedoch nur für kontinuierliche Signale gültig. Dieser Beitrag fasst die Arbeiten aus meiner Dissertation über eine Erweiterung zu einer mehrdimensionalen diskreten Skalenraum-Formulierung sowie die Ergebnisse theoretischer und experimenteller Untersuchungen von Multiskalen-Verfahren zur Kantenextraktion basierend auf mehrdimensionalen Kantenmodellen zusammen.

1 Einleitung

Die Extraktion von Kanten ist eines der zentralen Probleme im Forschungsgebiet Bildanalyse und Maschinensehen ([Jä97], [Ma76]). Wichtige Aspekte, die bei der Erforschung automatischer Verfahren zur Kantenextraktion zu berücksichtigen sind, sind die diskrete Bildsignalstruktur sowie die Tatsache, dass Kanten in Bildern i.a. komplexe mehrdimensionale Strukturen repräsentieren, die zudem auf unterschiedlichen Skalen erscheinen. Multiskalen-Verfahren zur Kantenextraktion werden zur vollständigen quantitativen Beschreibung der lokalen Eigenschaften von Kanten (Ort, Orientierung, Kontrast, Skala und Krümmung) eingesetzt. Demgemäß existiert eine sehr umfangreiche Literatur über dieses Gebiet (siehe z.B. [Be87], [Ca86], [Cl89], [El99], [Hi83], [Ko88], [Li98a], [MH80]).

Trotzdem gibt es aber gerade hier noch zahlreiche offene Fragen. Zunächst einmal befaßt sich der weitaus überwiegende Teil der auf diesem Gebiet aktiven Autoren mit Kanten in einem kontinuierlichen Kontext. Zudem wird das Problem in aller Regel auf eine im wesentlichen eindimensionale Situation reduziert. Nun treten aber die schwierigsten Fragen gerade im Mehrdimensionalen auf, eine brauchbare Definition einer Kante muss dann nämlich nicht nur der Steilheit des Grauwertüberganges Rechnung tragen, es muss auch noch die Form der Kante in Betracht gezogen werden. In der Bildverarbeitung ist ein kontinuierliches Modell nur von beschränktem Nutzen. Die in der Literatur vorherrschen-

den Multiskalen-Verfahren zur Kantenextraktion basieren auf der linearen Skalenraum-Theorie ([Li62], [Ko84], [SNFJ97], [Wi83]), die jedoch nur für kontinuierliche Signale gültig ist. Überdies basieren die meisten Verfahren auf eindimensionalen (1-D) Kantenmodellen, die wichtige Eigenschaften wie z.B. die Krümmung nicht explizit berücksichtigen. In meiner Dissertation ([Li03]) habe ich mich mit dem Problem einer Erweiterung auf eine mehrdimensionale diskrete Skalenraum-Formulierung sowie mit der Untersuchung von Multiskalen-Verfahren zur Kantenextraktion basierend auf mehrdimensionalen Kantenmodellen beschäftigt.

Der erste Teil der Dissertation konzentriert sich auf eine mehrdimensionale diskrete Skalenraum-Formulierung. Nach einer eingehenden Analyse über die überragende Rolle des Gaußkerns und der mit der Abtastung des Gaußkerns einhergehenden Probleme folgt eine Untersuchung des diskreten Skalenraum-Ansatzes von Lindeberg ([Li90], [Li94]). Zwar wurde in den Arbeiten von Lindeberg eine formal korrekte Diskretisierung der linearen Skalenraum-Theorie beschrieben, dies ist jedoch lediglich für 1-D Signale hinreichend detailliert ausgearbeitet. Als Ergebnis meiner Untersuchungen habe ich eine Verallgemeinerung auf 2-D und 3-D Signale vorgeschlagen. Eine nachfolgende theoretische Analyse der abgeleiteten Kerne sowie eine umfangreiche experimentelle Studie demonstriert anhand von synthetischen Bildern die Überlegenheit der von mir hergeleiteten Kerne gegenüber dem abgetasteten Gaußkern hinsichtlich der Glättungs- und Differentiationseigenschaften.

Der zweite Teil widmet sich Multiskalen-Verfahren zur Kantenextraktion basierend auf mehrdimensionalen Kantenmodellen. Ausgehend von einer Klassifikation mehrdimensionaler Kantenstrukturen bzgl. der lokalen Krümmungseigenschaften werden exemplarisch Modelle für geradlinige sowie für gekrümmte Kanten in 2-D Bildern vorgeschlagen. Die Modelle bilden die Grundlage für das von mir vorgeschlagene neuartige Verfahren zur optimalen Skalenbestimmung, wobei insbesondere der Einfluß der Kantenkrümmung im Skalenraum analysiert wird. Die Performanz des vorgeschlagenen Verfahrens zur optimalen Skalenbestimmung wird anhand von synthetischen Bildern untersucht. Die Studie zeigt, dass die experimentellen Ergebnisse im allgemeinen gut mit den theoretisch hergeleiteten Ergebnissen übereinstimmen.

2 Mehrdimensionale diskrete Skalenraum-Formulierung

Auf der Grundlage einer systemtheoretischen Analyse ([Br00], [F191]) habe ich erstmalig quantitativ gezeigt, dass der abgetastete Gaußkern für jede Dimension und insbesondere für kleine Skalen im Vergleich zum kontinuierlichen Gaußkern, der der eindeutige Kern für den linearen Skalenraum ist ([BWBD86], [WX90], [YP86]), mit einem unvermeidbaren Fehler behaftet ist. Basierend auf einer theoretischen Analyse der diskreten Skalenraum-Theorie nach Lindeberg für den 1-D Fall, die wiederum aus der Diskretisierung der linearen Diffusionsgleichung resultiert, ist erstmalig eine vollständige Herleitung für den 2-D und 3-D-Fall angegeben. Dabei bin ich auch ausführlich auf die Problematik der Zeitdiskretisierung der Diffusionsgleichung eingegangen. Eine besondere Rolle spielt hierbei die diskrete Umgebungsstruktur sowie die konkrete Diskretisierung des Laplace-Operators. Ein Kriterium ist dabei die Forderung nach Rotationssymmetrie, die auf ei-

nem diskreten Raum geeignet formuliert werden muss. Weiterhin soll die Halbgruppeneigenschaft erhalten bleiben. Dadurch konnte ich nachweisen, warum die von Lindeberg ([Li90], [Li94]) ohne weitere Begründung bzw. Herleitung angegebenen Parameter der Laplace-Operatoren (d.h. der linearen Operatoren in der zu diskretisierenden linearen Diffusionsgleichung) korrekt sind.

Diesen theoretischen Betrachtungen folgte die konsistente Herleitung der 2-D DSS-Kerne (engl. "discrete scale-space kernel"):

$$\begin{aligned} T(x, y; \Delta t) &= T(x; \Delta t) * T(y; \Delta t) \\ &= \begin{pmatrix} \frac{1}{6} & \frac{2}{3} & \frac{1}{6} \end{pmatrix} * \begin{pmatrix} \frac{1}{6} \\ \frac{2}{3} \\ \frac{1}{6} \end{pmatrix} = \begin{pmatrix} \frac{1}{36} & \frac{1}{9} & \frac{1}{36} \\ \frac{1}{9} & \frac{4}{9} & \frac{1}{9} \\ \frac{1}{36} & \frac{1}{9} & \frac{1}{36} \end{pmatrix}, \end{aligned}$$

wobei $\Delta t = \frac{1}{3}$. Der 3-D DSS-Kern sieht wie folgt aus:

$$\begin{aligned} T(x, y, z; \Delta t) &= T(x; \Delta t) * T(y; \Delta t) * T(z; \Delta t) \\ &= \begin{pmatrix} \frac{1}{6} & \frac{2}{3} & \frac{1}{6} \end{pmatrix}_x * \begin{pmatrix} \frac{1}{6} & \frac{2}{3} & \frac{1}{6} \end{pmatrix}_y * \begin{pmatrix} \frac{1}{6} & \frac{2}{3} & \frac{1}{6} \end{pmatrix}_z \\ &= \begin{pmatrix} \frac{1}{216} & \frac{1}{54} & \frac{1}{216} \\ \frac{1}{54} & \frac{2}{27} & \frac{1}{54} \\ \frac{1}{216} & \frac{1}{54} & \frac{1}{216} \end{pmatrix}_{z \pm 1}, \quad \begin{pmatrix} \frac{1}{54} & \frac{2}{27} & \frac{1}{54} \\ \frac{2}{27} & \frac{8}{27} & \frac{2}{27} \\ \frac{1}{54} & \frac{2}{27} & \frac{1}{54} \end{pmatrix}_z, \end{aligned}$$

wobei $\Delta t = \frac{1}{3}$.

Die Glättungs- und Differentiationseigenschaften der hergeleiteten DSS-Kerne wurden basierend auf der Z-Transformation untersucht. Diese theoretische Untersuchung zeigte, dass die Normalisierung der N-dimensionalen DSS-Kerne (gleiches gilt für die daraus hergeleiteten Differentiationskerne) für jede Skala immer garantiert ist und zudem die Konvolution mit Glättungs- und Differentiationskernen durch Ganzzahlarithmetik realisiert werden kann ([Li03], [LS03]).

Zum Beispiel sehen die geradzahigen 2-D Differentiationskerne erster Ordnung wie folgt

aus:

$$\begin{aligned}
T_{\Delta_{\text{even}, x}} \left(x, y; \frac{1}{3} \right) &= T_{\Delta_{\text{even}}} \left(x; \frac{1}{3} \right) * T \left(y; \frac{1}{3} \right) \\
&= \left(\frac{1}{6} \quad \frac{1}{2} \quad -\frac{1}{2} \quad -\frac{1}{6} \right)_x * \left(\frac{1}{6} \quad \frac{2}{3} \quad \frac{1}{6} \right)_y \\
&= \begin{pmatrix} \frac{1}{36} & \frac{1}{12} & -\frac{1}{12} & -\frac{1}{36} \\ \frac{1}{9} & \frac{1}{3} & -\frac{1}{3} & -\frac{1}{9} \\ \frac{1}{36} & \frac{1}{12} & -\frac{1}{12} & -\frac{1}{36} \end{pmatrix}.
\end{aligned}$$

Dagegen ergibt sich der ungeradzahlige Kern durch

$$\begin{aligned}
T_{\Delta_{\text{odd}, y}} \left(x, y; \frac{1}{3} \right) &= T \left(x; \frac{1}{3} \right) * T_{\Delta_{\text{odd}}} \left(y; \frac{1}{3} \right) \\
&= \left(\frac{1}{6} \quad \frac{2}{3} \quad \frac{1}{6} \right)_x * \left(\frac{1}{12} \quad \frac{1}{3} \quad 0 \quad -\frac{1}{3} \quad -\frac{1}{12} \right)_y \\
&= \begin{pmatrix} \frac{1}{72} & \frac{1}{18} & \frac{1}{72} \\ \frac{1}{18} & \frac{2}{9} & \frac{1}{18} \\ 0 & 0 & 0 \\ -\frac{1}{18} & -\frac{2}{9} & -\frac{1}{18} \\ -\frac{1}{72} & -\frac{1}{18} & -\frac{1}{72} \end{pmatrix}.
\end{aligned}$$

Der 1-D DSS-Kern und die entsprechenden Differentiationskerne erster Ordnung sind in Abbildung 1 dargestellt.

In meiner Arbeit habe ich gezeigt, dass die 2-D und 3-D DSS-Kerne

- die Halbgruppeneigenschaft erfüllen,
- die aus praktischer Sicht gewünschte Separabilität erfüllen,
- die Rotationsasymmetrie minimieren und
- durch iterative Konvolution einen diskreten linearen Skalenraum generieren, der analog aus der diskretisierten linearen Diffusionsgleichung resultiert.

Anhand von synthetischen Testbildern erfolgte eine Validierung der von mir hergeleiteten Kerne sowie ein Vergleich mit dem abgetasteten Gaußkern und seiner ersten Ableitung. Die Ergebnisse zeigten deutlich die Überlegenheit des DSS-Kerns und des daraus abgeleiteten ungeradzahligen Differentiationskerns.

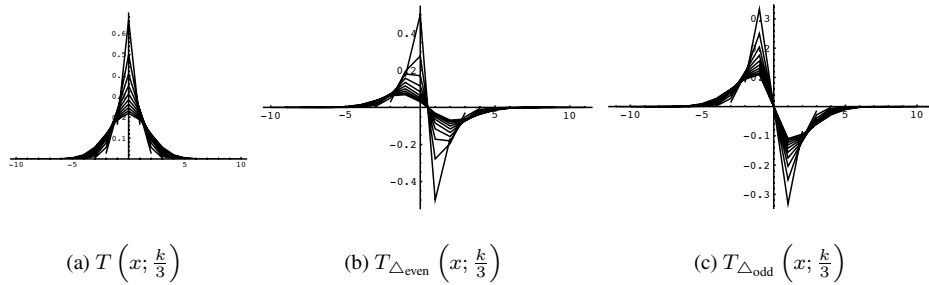


Abbildung 1: Der 1-D DSS-Kern und die entsprechenden Differentiationskerne erster Ordnung ($k = 1, 2, \dots, 10$)

3 Multiskalen-Verfahren zur Kantenextraktion

Fast allen Verfahren basierend auf der linearen Skalenraumtheorie ist gemein, dass ein spezielles Kantenmodell implizit angenommen bzw. sogar explizit berücksichtigt wird: die Heaviside-Funktion aus der Systemtheorie bzw. die daraus durch lineare Diffusion erzeugte sigmoide Funktion als glatte Funktion. Während die Behandlung des 1-D Falles unkritisch ist, hat sich die Behandlung des 2-D Falles als auch des mehrdimensionalen Falles u.a. wegen der notwendigerweise zu berücksichtigenden lokalen Krümmung als schwierig herausgestellt. Es sei denn, man trifft die Annahme einer linearen Replikation des 1-D Modells entlang der Tangentenrichtung auf dem kompakten Träger der Faltungskerne des Nabla-Operators. Da die ausschließlich geradlinige Kontur für reale digitale Bilder typischerweise nicht immer gilt, muss die Krümmung explizit in ein Multiskalen-Verfahren eingehen.

In meiner Arbeit wird bzgl. des sigmoiden Kantenmodells unter Bezugnahme auf Canny's "uncertainty principle" die Notwendigkeit der Betrachtung von multiplen Skalen alleine nur für den 1-D Fall gezeigt. Nach einer Erweiterung des 1-D Sigmoidenmodells auf den 2-D Fall habe ich sowohl für geradlinige als auch für konstant gekrümmte Konturen hergeleitet, wie das Problem der Skalenmessung bzw. -selektion gelöst werden kann (für geradlinige Konturen erfolgte dies unter Rückgriff auf Lindebergs Maximumsprinzip; siehe [Li98b]).

Die von mir hergeleitete Lösung für eine gekrümmte sigmoide Modellkontur basiert auf einer Betrachtung im Polarkoordinatensystem. Die konstant gekrümmte Kontur $L_{E_c}(x, y; t)$ in kartesischen Koordinaten kann zu $L_{E_c}(r; t)$ in Polarkoordinaten (mit $r^2 = x^2 + y^2$) transformiert werden:

$$L_{E_c}(r; t) = \mathcal{H}(R - r) * G(r; t_E + t),$$

wobei \mathcal{H} die Heaviside-Funktion bezeichnet, $G(\cdot)$ den Gaußkern, R den Krümmungsradius, t_E die lokale Breite der Intensitätsänderung der Kanten und t den Skalenparameter. Schließ-

lich ergibt sich für den Gradientenbetrag von $L_{E_c}(r; t)$

$$|\nabla L_{E_c}(r; t)| = \frac{R}{t_E + t} e^{-\frac{r^2 + R^2}{2(t_E + t)}} I_1\left(\frac{R \cdot r}{t_E + t}\right), \quad (1)$$

wobei $I_1(\cdot)$ die modifizierte Bessel-Funktion erster Ordnung ist ([AS72], [BS96]). In (1) wurde die lokale Krümmung als Kontureigenschaft explizit berücksichtigt, wobei angenommen wurde, dass der Krümmungsradius a priori bekannt ist. Dabei ist der sogenannte "curvature-scale constraint" zu beachten, welcher besagt, dass die lokale Krümmung der Kontur und die Skala der Kante (und somit die Operatorskala) in folgendem Konflikt zueinander stehen: Wenn der Krümmungsradius kleiner als die Kantenskala ist, ist ein Multiskalenverfahren prinzipiell nicht mehr anwendbar ([Li03]).

Eine experimentelle Validierung mit synthetischen 2-D Testdaten zeigte deutlich, dass die Theorie bzgl. geradliniger und konstant gekrümmter Konturen im Schnitt über alle Testbilder hinweg betrachtet sehr gut mit den experimentellen Ergebnissen übereinstimmt.

4 Zusammenfassung

Ausgehend von den Arbeiten von Lindeberg zu einer diskreten Skalenraum-Theorie, habe ich mich in meiner Dissertation mit dem Problem der Erweiterung auf eine mehrdimensionale diskrete Skalenraum-Formulierung sowie mit der Untersuchung von Multiskalen-Verfahren zur Kantenextraktion basierend auf mehrdimensionalen Kantenmodellen beschäftigt. Die theoretischen Untersuchungen sowie eine umfangreiche experimentelle Studie demonstrierten die Überlegenheit der von mir hergeleiteten DSS-Kerne gegenüber dem abgetasteten Gaußkern hinsichtlich der Glättungs- und Differentiationseigenschaften. Ich habe 2-D Kantenmodelle für ein neuartiges Verfahren zur optimalen Skalenbestimmung vorgeschlagen, wobei insbesondere der Einfluß der Kantenkrümmung im Skalenraum analysiert wurde. Die Performanz des vorgeschlagenen Verfahrens zur optimalen Skalenbestimmung wurde anhand von synthetischen Bildern untersucht. Die Studie zeigte, dass die experimentellen Ergebnisse im allgemeinen gut mit den theoretisch hergeleiteten Ergebnissen übereinstimmen.

Literatur

- [AS72] Abramowitz, M. und Stegun, I. A.: *Handbook of Mathematical Functions*. Dover Publisher, New York. 9th. 1972.
- [Be87] Bergholm, F.: Edge Focusing. *IEEE Trans. on Pattern Analysis and Machine Intelligence*. 9(6):726–741. 1987.
- [Br00] Bracewell, R. N.: *The Fourier Transform and Its Applications*. McGraw-Hill. 3rd. 2000.
- [BS96] Bronstein, I. N. und Semendjajew, K. A.: *Teubner-Taschenbuch der Mathematik*. B. G. Teubner, Stuttgart. 1996.

- [BWBD86] Babaud, J., Witkin, A. P., Baudin, M., und Duda, R. O.: Uniqueness of the Gaussian Kernel for Scale-Space Filtering. *IEEE Trans. on Pattern Analysis and Machine Intelligence*. 8(1):26–33. 1986.
- [Ca86] Canny, J. F.: A Computational Approach to Edge Detection. *IEEE Trans. on Pattern Analysis and Machine Intelligence*. 8(6):679–698. 1986.
- [Cl89] Clark, J. J.: Authenticating Edges Produced by Zero-Crossing Algorithms. *IEEE Trans. on Pattern Analysis and Machine Intelligence*. 11(1):43–57. 1989.
- [El99] Elder, J. H.: Are Edges Incomplete? *Internat. Journal of Computer Vision*. 34(2/3):97–122. 1999.
- [Fl91] Fliege, N.: *Systemtheorie*. Informationstechnik. Teubner, Stuttgart. 1991.
- [Hi83] Hildreth, E. C.: The Detection of Intensity Changes by Computer and Biological Vision Systems. *Computer Vision, Graphics, and Image Processing*. 22:1–27. 1983.
- [Ii62] Iijima, T.: Basic theory on normalization of pattern (in case of typical one-dimensional pattern). *Bulletin of the Electrotechnical Laboratory*. 26:368–388. 1962. in Japanese.
- [Jä97] Jähne, B.: *Digitale Bildverarbeitung*. Springer. 4th. 1997.
- [Ko84] Koenderink, J. J.: The Structure of Images. *Biological Cybernetics*. 50:363–370. 1984.
- [Ko88] Korn, A. F.: Toward a Symbolic Representation of Intensity Changes in Images. *IEEE Trans. on Pattern Analysis and Machine Intelligence*. 10(5):610–625. 1988.
- [Li90] Lindeberg, T.: Scale-Space for Discrete Signals. *IEEE Trans. on Pattern Analysis and Machine Intelligence*. 12(3):234–264. 1990.
- [Li94] Lindeberg, T.: *Scale-Space Theory in Computer Vision*. Kluwer Academic Publisher, Boston, MA. 1994.
- [Li98a] Lindeberg, T.: Edge Detection and Ridge Detection with Automatic Scale Selection. *Internat. Journal of Computer Vision*. 30(2):117–154. 1998.
- [Li98b] Lindeberg, T.: Feature Detection with Automatic Scale-Selection. *Internat. Journal of Computer Vision*. 30(2):77–116. 1998.
- [Li03] Lim, J. Y.: *Discrete Scale-Space Formulation and Multiscale Edge Extraction toward Higher Dimensions*. Dissertation. Fachbereich Informatik, Universität Hamburg, Germany. 2003. Published in Dissertationen zur Künstlichen Intelligenz, Band 272. Akademische Verlagsgesellschaft Aka GmbH, Berlin, 2003.
- [LS03] Lim, J. Y. und Stiehl, H. S.: A Generalized Discrete Scale-Space Formulation for 2-D and 3-D Signals. In: Griffin, L. D. und Lillholm, M. (Hrsg.), *Proc. 4th Int. Conf. Scale Space 2003, Isle of Skye, UK*. Lecture Notes in Computer Science 2695. S. 132–147. Springer-Verlag, Berlin. 2003.
- [Ma76] Marr, D.: Early Processing of Visual Information. *Phil. Royal Soc (B)*. 27:483–524. 1976.
- [MH80] Marr, D. und Hildreth, E.: Theory of Edge Detection. In: *Proc. Royal Society of London B 207*. S. 187–217. 1980.
- [SNFJ97] Sporring, J., Nielsen, M., Florack, L. M. J., und Johansen, P.: *Gaussian Scale-Space Theory*. Kluwer Academic Publishers. 1997.

- [Wi83] Witkin, A. P.: Scale-Space Filtering. In: *Proc. 8th Int. Joint Conf. Artificial Intelligence, Karlsruhe*. S. 1019–1021. 1983.
- [WX90] Wu, L. und Xie, Z.: Scaling Theorems for Zero-Crossings. *IEEE Trans. on Pattern Analysis and Machine Intelligence*. 12(1):46–54. 1990.
- [YP86] Yuille, A. L. und Poggio, T. A.: Scaling Theorems for Zero Crossings. *IEEE Trans. on Pattern Analysis and Machine Intelligence*. 8(1):15–25. 1986.

Ji-Young Lim studierte Computer Science zwischen 1991 und 1997 an der Ewha University in Seoul (Süd-Korea) und erlangte im Februar 1997 am Fachbereich Computer Science der Ewha University den Master-Abschluss. Von Oktober 1998 bis Mai 2003 gefördert mit einem Promotionsstipendium des DAAD (Deutscher Akademischer Austauschdienst), promovierte sie am Arbeitsbereich Kognitive Systeme des Fachbereichs Informatik an der Universität Hamburg. Im Anschluss daran war sie bis März 2004 wissenschaftliche Mitarbeiterin am Arbeitsbereich Kognitive Systeme des Fachbereichs Informatik an der Universität Hamburg. Seit April 2004 ist sie tätig als Software-Entwicklerin bei der Smartsoft GmbH in Hamburg.