A Discrete Scale Space Neighborhood for Robust Deep Structure Extraction

Martin Tschirsich[1] and Arjan Kuijper[1,2]

[1] Technische Universität Darmstadt, Germany
[2] Fraunhofer IGD, Darmstadt, Germany

Abstract. Linear or Gaussian scale space is a well known multi-scale representation for continuous signals. The exploration of its so-called deep structure by tracing critical points over scale has various theoretical applications and allows for the construction of a scale space hierarchy tree. However, implementational issues arise, caused by discretization and quantization errors. In order to develop more robust scale space based algorithms, the discrete nature of computer processed signals has to be taken into account. Aiming at a computationally practicable implementation of the discrete scale space framework, we investigated suitable neighborhoods, boundary conditions and sampling methods. We show that the resulting discrete scale space respects important topological invariants such as the Euler number, a key criterion for the successful implementation of algorithms operating on its deep structure. We discuss promising properties of topological graphs under the influence of smoothing, setting the stage for more robust deep structure extraction algorithms.

1 Introduction

In the field of computer vision, deriving information from observed images is a central problem. Various strategies have been invented to do so in a performant manner, usually by applying some kind of operator. Their performance depends on the *inner scale*, the sampling density or resolution of the image they operate on. To overcome this dependence between operator and inner scale, various strategies of multi-scale representations have been proposed. Almost all those strategies consist of transforming the given images into a scale independent representation first before applying an operator on this representation. A common requirement for such preliminary transformations is to mask as little information present in the original image as possible. The Gaussian scale space satisfies these requirements and can be thought of as the natural generalization of the lowpass pyramid. It is also popular for its theoretical foundation. A Gaussian scale space representation of a given signal is a family of derived signals, progressively smoothed with a Gaussian filter.

Its *deep structure* consists of critical points or zerocrossings traced over scale. An implementation of important scale space based algorithms can be found in the software tool ScaleSpaceViz [1]. ScaleSpaceViz has, according to the authors, "proven to be useful in exploring the deep structure of images and constructing applications involving scale space interest points, such as reconstruction and matching". Admittedly, this

holds true under certain conditions. Nevertheless, ScaleSpaceViz suffers from robustness problems. As it is the case with many such scale space applications, its implementation is based on a discretized continuous scale space.

The discrete scale space proposed by Lindeberg [2] takes the discrete nature of computer processed signals into account. It is based on equivalent assumptions and axioms that have been used to derive the continuous Gaussian scale space adapted to discrete signals. It is our belief that porting scale space algorithms from a discretized continuous to the discrete scale space will eventually lead to more accurate, robust and possibly faster implementations. The discrete scale space formalized by Lindeberg however *does not* respect important topological invariants such as the Euler number. Since most algorithms that operate on the deep structure of the Gaussian scale space require this topological invariant to hold, we present in this paper a modified definition of the discrete scale space respecting the Euler number. A subsequent investigation of various properties of this discrete scale space then results in a fast and robust sampling algorithm [3]. We propose the application of topological graphs together with adaptive sampling in order to reliably extract the deep structure of the discrete scale space.

2   Discrete Signals

Important discrete operators, including those needed to build a scale space over f, operate on functions of infinite domain $\mathbb{Z}^2$. Therefore, signals with bounded domain must be expanded to cover the whole range of $\mathbb{Z}^2$. Imposing a Dirichlet or zero-border condition implies assuming the signal to be of constant or zero-value outside the originally bounded domain. A zero-border condition may reduce the computational complexity of certain operations, namely discrete convolution with infinitely large kernels. However, a zero-border has drawbacks concerning the computation of a scale space over f. These can be avoided by assuming f to be periodic instead with period M for the first and N for the second dimension, thus imposing a periodic boundary condition.

Image structures or features are often defined in terms of limited subsets of image points, so called interest points. These include stationary points such as minima, maxima and saddles in the two-dimensional case. If the discrete signal f is interpreted as lattice height data, these critical points are then called peaks, pits and passes. They reveal important topological characteristics. Critical point detection has to be performed in order to find their positions.

Provided that f is a sampled continuous signal and the sampling operator is known, it is often possible to reconstruct and perform critical point detection on a functional representation of the original signal using the well known gradient or slope based definitions of critical points on continuous data [4]. In the general case however, this is not possible and we need a separate definition of critical points on discrete data.

2.1   Euler Number

A well known topological invariant stating a stable relation between the number of extrema and saddle points on closed continuous surfaces is the Euler number. The Euler number, also called Euler formula or mountaineer's equation [5] for a closed continuous
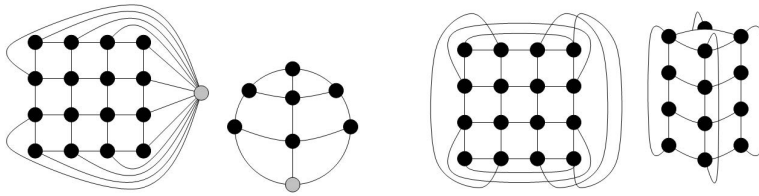
surface is given by $m^+ - s + m^-$, where m^+, s and m^- denote the number of local maxima, saddles and minima on that surface.

Various algorithms that exploit the deep structure of the continuous scale space rely on a stable relation between the number of these critical points. This invariant should therefore equally hold true for discrete critical points on discrete signals. Firstly, we need a definition of the Euler number for a discrete signal f and then a definition of discrete critical points that respects this invariant. An important aspect of this definition is the neighborhood of each point in $D(f)$. The neighborhood of a point consists of all those points it is adjacent to. This relation is usually expressed as a mapping from a point to the set of its neighbors $N : \mathbb{Z}^2 \to \mathcal{P}(\mathbb{Z}^2)$. In order to reason about critical points and the relations between the number of critical points in a given neighborhood, we resort to a graph representation of f.

The discrete signal f and a symmetric neighborhood N can be represented as an undirected graph $G_f = (V, E)$ with vertices $V = \{v_{x,y} : (x, y) \in D(f)\}$, edges $E = \{\{v_{x,y}, v_{x',y'}\} \in \mathcal{P}(V) : (x', y') \in N(x, y)\}$ and a function $f_G : V \to \mathbb{R}_+$ relating each vertex to a value $f_G(v_{x,y}) = f(x, y)$. A bijection between $D(f)$ and V relates each point $(x, y) \in D(f)$ to a vertex $v_{x,y} \in V$. For each pair of points (x, y), (x', y') that is adjacent given the neighborhood N, E contains an edge $\{v_{x,y}, v_{x',y'}\}$. Thus, the domain of f is represented via V, the neighborhood via E and the mapping f via the mapping f_G over the set of vertices.

If the graph G_f can be embedded into a closed surface such that its edges do not intersect, the graph genus g then denotes the minimal genus of the closed surface the graph can be embedded in [6]. For closed continuous surfaces, the Euler characteristic of that surface $\chi = 2 - 2g$ coincides with the Euler number [4]. The value of the Euler characteristic and thus the Euler number depends on the chosen neighborhood. A graph with no edges for example can always be embedded into a sphere, a closed surface with genus 0. By contrast, the complete graph K_5 on 5 vertices is not planar and can only be embedded into a torus, a closed surface with genus 1. Intuitively, the 4-neighborhood $N_4(x, y) = \{(x', y') \in \mathbb{Z}^2 : \|(x - x', y - y')\|_1 = 1\}$ seems to be a natural choice given the square lattice f is defined upon, but we will later see that a 6-neighborhood N_6 resulting from a Delaunay triangulation of the square lattice is a more practical choice. For bounded signals, G_f also has to reflect the chosen boundary condition. For positive signals, a zero-border condition can be modeled by introducing a virtual pit $v_o \in V$ with $f_G(v_0) = 0$ connected to all boundary points (Fig. 1a) [5]. In this case, the neighborhood of boundary points might deviate from N, but this does not matteras long as we are only interested in topological properties. Modeling a periodic boundary condition however requires a strict correspondence between the chosen neighborhood and the graph representation, modeled by duplication of $D(f)$ in all directions and connecting opposing boundary points through additional edges (Fig. 1b). A periodic replication of the square grid $D(f)$ does not affect the Euler number.

If we impose a zero-border and not a periodic boundary condition, the graph representation of a signal f with bounded domain $[1, 3] \times [1, 3] \subseteq D(f) \subset \mathbb{Z}^2$ and both 4-neighborhood N_4 or 6-neighborhood N_6 can be embedded into a sphere, i.e. G_f is planar (Fig. 1a). Intersection-free embedding into a sphere is not possible for periodic f.

(a) 4-neighborhood with virtual pit (gray), graph genus 0, embedded in a sphere (spherical polyhedron).

(b) periodic 4-neighborhood, graph genus 1, embedded in a torus (toroidal polyhedron).

Fig. 1. Embedding an undirected graph with 4-neighborhood N_4 into a closed surface

Proof. Using Kuratowski's theorem stating that a finite graph is not planar if it does contain a subgraph that is homeomorphic to K_5, it can be shown that the graph representation G_f of a periodic signal f with $D(f) = [1,3] \times [1,3] \subset \mathbb{Z}^2$ is not planar. Since the graph representation of a periodic signals f whose domain is a superset of $D(f)$ is a supergraph of G_f, it is homeomorphic to K_5 as well. The same holds for arbitrary neighborhoods N with $N_4(x,y) \subseteq N(x,y)$, since their graph representation is also a supergraph of G_f. □

The surface with the next higher genius is a torus. From Fig. 1b it becomes clear that it is possible to embed G_f into a torus for periodic f. For the Euler characteristic and therefore the Euler number then holds $\chi = 2 - 2g = 0 \Leftrightarrow m^+ + m^- = s$.

Now that we have transfered the invariant stated by the Euler number for closed continuous surfaces to the graph representation G_f and thus the periodic discrete signal f for the N_4 neighborhood and its triangulations N_6, we can reduce the number of reasonable discrete critical point definitions upon these neighborhoods to those that respect the Euler number. Eventually, such a critical point definition allows for easier adaption of algorithms that work well in the continuous case to discrete signals.

2.2 Detecting All Critical Points

There are principally two different ways to define critical points in the continuous case, gradient and local path based definitions [4]. This translates to the discrete case if proper definitions of gradient and local paths are given, both depending on the chosen neighborhood. According to Kovalevsky [7], a stationary (homogeneous) neighborhood suitable to formulate a definition of discrete critical points respecting the Euler number in the discrete n-dimensional space must contain $2(2^n - 1)$ points. Any stationary 6-neighborhood given by a triangulation of the square lattice meets this condition. Later, it was shown by Takahashi [5] that any arbitrary triangulation of the lattice is equally topologically consistent. However, a stationary neighborhood is inherently simpler and introduces no perturbations in the position of critical points, thus we will use a stationary hexagonal neighborhood $N_6(x,y) = \{(x \pm 1, y), (x, y \pm 1), (x \pm 1, y \mp 1)\}$ as proposed by Kuijper [8].

We can now give a definition of discrete critical points respecting the Euler number on the hexagonal neighborhood N_6: $(x_0, y_0) \in \mathbb{Z}^2$ is said to be a local *maximum (minimum)* for a signal $f : \mathbb{Z}^2 \to \mathbb{R}$ within neighborhood $N(x_0, y_0) \subseteq \mathbb{Z}^2$ iff

$\forall (x, y) \in N : (x_0, y_0) > (<) (x, y)$. This definition is only valid as long as there are no local extremal regions. In order to incorporate these plateaus, we introduce an arbitrary strict second ordering on points of equal value.

A common definition of saddle points is via sign-changes in the neighborhood, given a clockwise or counterclockwise order of the neighbors. Given a neighborhood N, the lexicographic order with minor adaptations to the boundary provides such an arrangement. We will use the definition given by Takahashi [5] who differs between regular and degenerated saddles. Four sign changes indicate a regular, six sign changes a degenerated saddle. At degenerated saddle points, more than two contour lines intersect. Degenerated saddles on N_6 can be split into two regular saddles.

Provided a neighborhood and a suitable critical point definition as outlined above, we can extract the surface network and other topological graphs such as the Reeb graph from f. This can be done using the algorithms outlined by Takahashi [5] and Scott [4] with minor adaptations to periodic signals whose representative graph is of genus 1.

3   Neighborhood and Discrete Scale Space

In this section, we will give a concise definition of the Gaussian scale space [9, 10] for two-dimensional discrete signals comprising a continuous scale parameter. The scale space representation of a discrete signal $f : \mathbb{Z}^2 \to \mathbb{R}$ is a one-parameter family of derived signals $L : \mathbb{Z}^2 \times \mathbb{R}^+$ defined by a set of discrete scale space axioms. It was first described by Lindeberg [2], who introduced a discrete version of the *non enhancement of local extrema* axiom based on *weak* local extrema within the 8-neighborhood (Fig. 2).

An equivalent definition via convolution with a scale space kernel yields $L (\cdot, \cdot; t) = k (\cdot, \cdot; t) * f (\cdot, \cdot)$ where $k : \mathbb{Z}^2 \times \mathbb{R}_+ \to \mathbb{R}$ is the discrete analogue of the continuous Gaussian scale space kernel.

$$L (x, y; t) = \sum_{m=-\infty}^{\infty} k (m; t) \sum_{n=-\infty}^{\infty} k(n; t) f(x - m, y - n)$$

The scale space representation L of a signal f is the solution of the diffusion equation. For the axioms chosen by Lindeberg, it takes the form

$$\partial_t L = \alpha \nabla_2^5 L + \beta \nabla_\times^2 L \tag{1}$$

with $L (\cdot, \cdot; 0) = f (\cdot, \cdot)$ as initial condition and some $\alpha, \beta \geq 0$. The five-point operator ∇_2^5 and the cross operator $\nabla_\times^2$ approximate the continuous Laplacian operator. They correspond to convolution of $L (\cdot, \cdot; t)$ with kernel $\begin{bmatrix} & 1 & \\ 1 & -4 & 1 \\ & 1 & \end{bmatrix}$ respectively $\begin{bmatrix} \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & -2 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} \end{bmatrix}$.

For $\alpha = 1$ and $\beta = 0$ Eq. (1) simplifies to $\partial_t L = \nabla_2^5 L$ and results in k being separable.

If one wishes to choose a strict local extrema definition and another neighborhood because of topological reasons, the axiomatic scale space definition may change. Only the strict local extrema definition within a 6-neighborhood has proven to respect the Euler number. However, not every spatial extremum in L within an arbitrary 6-neighborhood satisfies the aforementioned non-enhancement axiom based on weak local extrema
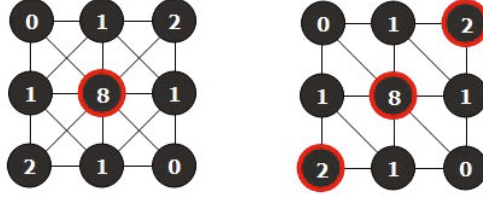
Fig. 2. Different neighborhoods (left: N_8, right: N_6) over the same signal. Only the central maximum satisfies the non-enhancement axiom $\partial_t L \leq 0$ for both neighborhoods and arbitrary $\alpha \geq 0$, $\beta > 0$.

within the 8-neighborhood chosen by Lindeberg. Thus, in order to incorporate the Euler number as topological invariant into the discrete scale space, either L and therefore the discretized diffusion equation has to be modified or the 6-neighborhood has to be chosen in a way that the extrema within such neighborhood represent a subset of the weak extrema within the 8-neighborhood.

3.1   Discretized Scale Space for N_6

We want to show that the discrete scale space over f with $N = N_6$ is defined by the solution of a discretized version of the differential equation as it is the case for $N = N_8$.

Necessity. Let $N = N_6$. Then, following the approach given by Lindeberg [2], it becomes evident that the scale space representation $L : \mathbb{Z}^2 \times \mathbb{R}_+ \to \mathbb{R}$ of signal $f : \mathbb{Z}^2 \to \mathbb{R}$ satisfies the differential equation $\partial_t L = \alpha \nabla_5^2 L$ for some $\alpha \geq 0$ with initial condition $L(\cdot, \cdot; 0) = f(\cdot, \cdot)$.

Proof. Strictly following the proof given by Lindeberg and resulting from his set of proposed discrete scale space axioms *without* the non-enhancement of local extrema axiom and thus not depending on the chosen neighborhood, $\partial_t L = AL$ holds for all $t \geq 0$ and a linear shift-invariant operator A. Because of the shift-invariance, this can be written as

$$(AL)(x, y; t) = \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} a_{m,n} L(x - m, y - n; t).$$

Let $N_6^+(x, y) = N_6(x, y) \cup \{(x, y)\}$. Now, from using the non-enhancement of local extrema axiom for extrema within N_6, it follows that $a_{m.n} = 0$ for $(m, n) \in N_6^+$. Otherwise, assuming that $a_{\tilde{m}.\tilde{n}} \neq 0$ for one $(\tilde{m}, \tilde{n}) \in N_6^+$, we could define a function $f_1 : \mathbb{Z}^2 \to \mathbb{R}$ with

$$f_1(x, y) = \begin{cases} \epsilon > 0 & \text{if } (x, y) = (0, 0) \\ 0 & \text{if } (x, y) \in N_6(0, 0) \\ 1 & \text{if } (x, y) = (\tilde{m}, \tilde{n}) \\ 0 & \text{otherwise} \end{cases}$$

and $\partial_t L (0,0;0) = \partial_t f (0,0) = \epsilon a_{o,o} + a_{\tilde{m},\tilde{n}} \geq 0$ for well-chosen ϵ, violating the non-enhancement axiom since $(0,0)$ is a local maximum of f within N_6. The same holds for local minima and $\epsilon < 0$. Thus, $\partial_t L$ can be written as

$$(AL)(x,y;t) = \sum_{m,n \in N_6^+} a_{m,n} L(x-m, y-n; t).$$

Imposing symmetry conditions in analogy to Lindeberg [2], AL equals a convolution of L with kernel $\begin{bmatrix} 0 & b & 0 \\ b & c & b \\ 0 & b & 0 \end{bmatrix}$ for some b and c. Considering another function $f_2(\cdot,\cdot) = 1$ with $f_2(0,0) = 1 + (-)\epsilon$ According to the strict definition of local extrema given earlier, the point $(0,0)$ is a local maximum (minimum), thus $\partial_t L (0,0;0) = \partial_t f_2(0,0) = 4b + c = 0$. Therefore, $\begin{bmatrix} 0 & b & 0 \\ b & c & b \\ 0 & b & 0 \end{bmatrix} = \alpha \begin{bmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{bmatrix}$ for some α. Finally, from $\partial_t L (0,0;0) = \partial_t f_1(0,0) = \alpha \epsilon c \leq 0$ it becomes evident that $\alpha \geq 0$. □

Following this proof, it becomes apparent that not only the hexagonal 6-neighborhood, but arbitrary 6-neighborhoods (arbitrary *triangulations*) as well as the 4-neighborhood N_4 result in the same differential equation $\partial_t L = \alpha \nabla_5^2 L$. For simplicity, the parameter α, which only affects the scaling of the scale parameter, is usually set to 1.

Sufficiency. The solution $L : \mathbb{Z}^2 \times \mathbb{R}_+ \to \mathbb{R}$ of the differential equation $\partial_t L = \alpha \nabla_5^2 L$ for some $\alpha \geq 0$ with initial condition $L(\cdot,\cdot;0) = f(\cdot,\cdot)$ for $f : \mathbb{Z}^2 \to \mathbb{R}$ is a scale space representation of f.

Proof. Let $\tilde{L}$ be a scale space representation of f. According to the previous definition, $\tilde{L}$ is then also the solution of $\partial_t L = \tilde{\alpha} \nabla_5^2 L$ for some $\tilde{\alpha} \geq 0$ with initial condition $L(\cdot,\cdot;0) = f(\cdot,\cdot)$. Since $\tilde{\alpha}$ is only a linear scaling parameter along the direction of t, $\tilde{\tilde{L}}(\cdot,\cdot;t) = \tilde{L}(\cdot,\cdot;\frac{\alpha}{\tilde{\alpha}}t)$ is the solution of $\partial_t L = \alpha \nabla_5^2 L$ and therefore $\tilde{\tilde{L}} = L$. It is obvious that $\tilde{\tilde{L}}$ still satisfies all the discrete scale scape axioms, thus L is a scale space representation of f. □

Linking. In order to extract critical curves, the discrete scale space has to be sampled along its continuous scale parameter, since we do not know how to compute the exact occurrences of the zerocrossings in the neighborhood as stated above. The known way to extract critical curves is to compute several scale images of L at preselected scales, then detecting spatially critical points on these scales and finally linking spatially close critical points on subsequent scales into critical curves. We need a criterion that tells us when further subsampling is required and when the sampling density is high enough to guarantee a correct result. Such a criterion is found in topological graphs or more precisely in the difference of topological graphs of subsequent scales. Tracking changes in the surface network over scale is a promising approach, since the set of possible changes between scale space events such as creations or annihilations of critical points is strictly limited.
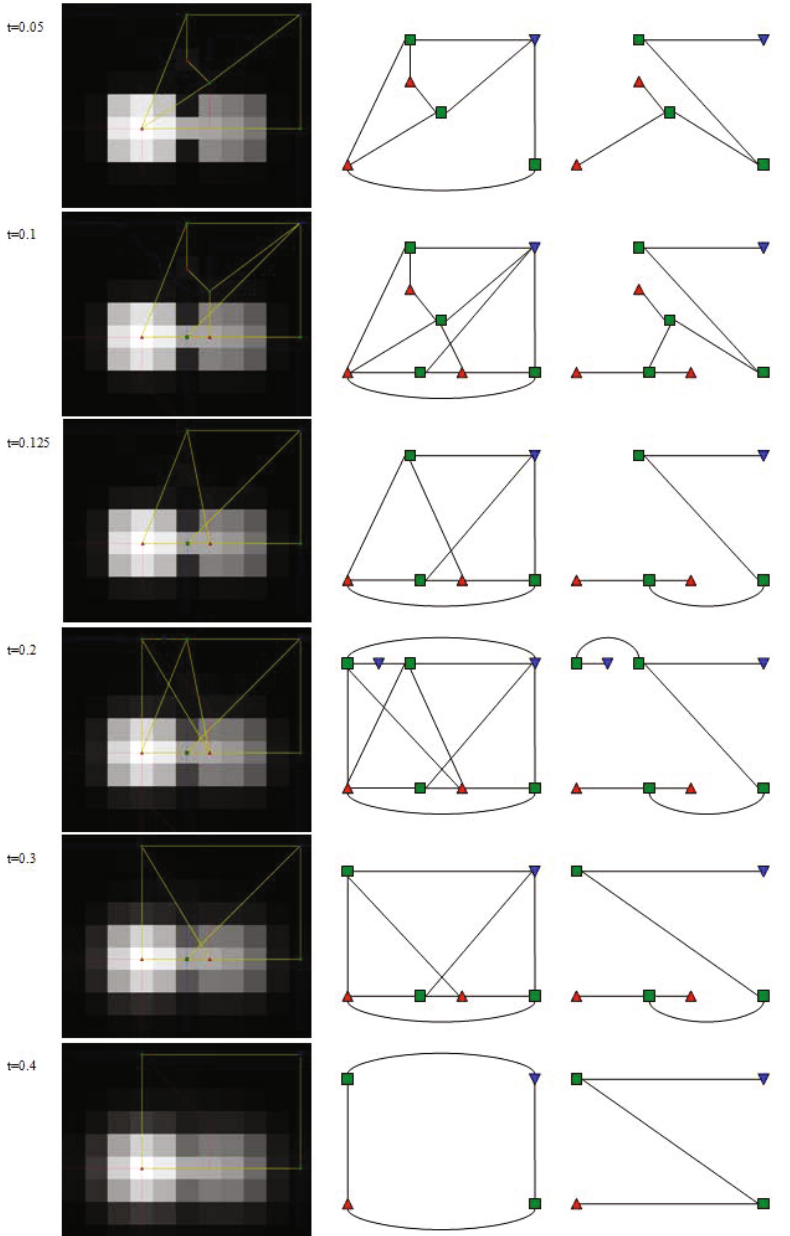
Fig. 3. A sampled discrete scale space. From the critical points alone (left), the subsequent annihilation and creation between scale $t_1 = 0.05$ and $t_2 = 0.125$ is not visible and might lead to incorrect linking. However, the surface networks (middle, incomplete) of these scales differ, thus providing a criterion whether further subsampling might be necessary. The Reeb graphs (right, incomplete) are identical, though.

4   Conclusion and Future Work

The discrete scale space as an equivalent to the two-dimensional Gaussian scale space has been discussed and some important properties have been derived. A computationally practicable implementation of the discrete scale space framework has been outlined. The regular 6-neighborhood, a periodic boundary condition and a suitable critical point definition respecting the Euler number have led to discrete scale space axioms that differ from those proposed by Lindeberg. It has been shown that the discretized diffusion equation inducing the discrete scale space derived from the modified axioms resembles the one found by Lindeberg but has one less degree of freedom. Using our computationally efficient sampling method, based on properties of the Laplacian kernel [3], we gave a first investigation of the deep structure of the discrete scale space illustrating the need for a more robust algorithm for critical curve extraction. Topological graphs have shown promising properties under the influence of changes in scale. However, further and more formal investigation of the deep structure of the discrete scale space is necessary.

References

1. Kanters, F., Florack, L., Duits, R., Platel, B., ter Haar Romeny, B.: Scalespaceviz: a-scale spaces in practice. Pattern Recognition and Image Analysis 17, 106–116 (2007)
2. Lindeberg, T.: Discrete Scale-Space Theory and the Scale-Space Primal Sketch. PhD thesis, Royal Institute of Technology (1991)
3. Tschirsich, M.: The discrete scale space as a base for robust scale space algorithms. Technical report, Department of Computer Science, Technical University of Darmstadt (June 2012)
4. Scott, P.J.: An algorithm to extract critical points from lattice height data. International Journal of Machine Tools and Manufacture 41(13-14), 1889–1897 (2001)
5. Takahashi, S., Ikeda, T., Shinagawa, Y., Kunii, T.L., Ueda, M.: Algorithms for extracting correct critical points and constructing topological graphs from discrete geographical elevation data. Computer Graphics Forum 14(3), 181–192 (1995)
6. Perez, A.: Determining the genus of a graph. HC Mathematics Review 1(2), 4–13 (2007)
7. Kovalevsky, V.A.: Discrete topology and contour definition. Pattern Recognition Letters 2(5), 281–288 (1984)
8. Kuijper, A.: On detecting all saddle points in 2d images. Pattern Recognition Letters 25(15), 1665–1672 (2004)
9. Koenderink, J.J.: The structure of images. Biological Cybernetics 50, 363–370 (1984)
10. Weickert, J., Ishikawa, S., Imiya, A.: Scale-space has been discovered in japan. Technical report, Department of Computer Science, University of Copenhagen (August 1997)