

The Morphological Structure of Images: The Differential Equations of Morphological Scale-Space

Rein van den Boomgaard and Arnold Smeulders

Abstract—In this paper we introduce a class of nonlinear differential equations that are solved using morphological operations. The erosion and dilation act as morphological propagators propagating the initial condition (original image in computer vision terminology) into the “scale-space,” much like the Gaussian convolution is the propagator for the linear diffusion equation.

Analysis starts in the set domain, resulting in the description of erosions and dilations in terms of contour propagation. We show that the structuring elements to be used must have the property that at each point of the contour there is a well-defined and unique normal vector. Then given the normal at a point of the dilated contour we can find the corresponding point (point-of-contact) on the original contour. In some situations we can even link the normal of the dilated contour with the normal in the point-of-contact of the original contour. The results of the set domain are then generalized to grey value images. The role of the normal is replaced with the function gradient. The same analysis also holds for the erosion.

Using a family of increasingly larger structuring functions we are then able to link infinitesimal changes in grey value (resulting from the use of an infinitesimally larger structuring function) with the gradient in the image. The obtained differential equations bear great resemblance to the nonlinear differential equation, Burgers’ equation, describing the propagation of a shock-wave.

In the discussion we indicate that the results of this paper provide the theoretical basis to analyze morphological scale-space in much greater depth.

I. INTRODUCTION

IN THIS PAPER, we investigate where points go when the object they are part of is repeatedly dilated with a small structuring element. The only relevant points to study are the points on the contour of course and we do so in Sections III-A and IV. For grey value morphology, where a 2-D function is represented as a 3-D set, the same analysis can be made (see Sections V and V-D). The goal of this paper is to derive the differential equations, for which morphological operations provide the solution (as discussed in Section VI). Such a differential equation is the mathematical morphological analogue to the heat diffusion equation that is the starting point in linear scale-space [1]. The use of morphological operations to solve partial differential equations is reported before (see [2], [3], [4]).

The use of morphological operators to solve differential equations is known from distance transform theory [5], [6].

Manuscript received March 25, 1993; revised October 14, 1994. Recommended for acceptance by Associate Editor D. Huttenlocher.

The authors are with the Faculty of Mathematics and Computer Science, University of Amsterdam, Kruislaan 403 1098 SJ, Amsterdam The Netherlands; e-mail: rein@fwi.uva.nl.

IEEE Log Number 9405784.

The common distance transform:

$$d_X(x) = \min_{y \in X^c} \|x - y\| \quad (1)$$

is a solution of the differential equation:

$$\|\nabla d_X\| = 1 \quad (2)$$

with initial condition $d_X(x) = 0$ for $x \in X^c$. The level sets of d_X can be found by eroding the original set X with a disk shaped structuring element. Note that the function d_X is not differentiable at all points $x \in X$. Matheron [5] showed that, because d_X is a Lipschitz function, d_X is *almost everywhere* differentiable. As a consequence the area (volume) of the set of all points where d_X is *not* differentiable is negligible. For those points where ∇d_X is defined, d_X indeed satisfies the differential equation. In mathematical terms we say that d_X is a *weak solution* (or a solution in the distribution sense) of the differential equation.

The wave propagation interpretation of the distance transform is introduced by Blum [7], and implemented by Verwer [6]. As a general method to solve differential equation(1), waves starting at the boundary of X , travel inwards with equal speed in all directions at all points of space. The arrival time of the wave at a point $x \in X$ is proportional to its distance from the boundary. The medium in which the waves propagate is thus isotropic and homogeneous.

Furthermore, Verwer considers the more complex, but closely related, differential equation:

$$\|\nabla f\|^2 = g^2$$

(known as the *Eikonal equation*). This differential equation is also solved with a distance transform, this time using an isotropic, *inhomogeneous* metric ($ds^2 = g(x_1, x_2)(dx_1^2 + dx_2^2)$). The Eikonal equation describes the propagation of waves in a medium where the propagation speed of the waves varies from position to position. In computer vision the propagation speed is proportional to the local grey value, hence the name “grey weighted distance transform.”

In this paper, we consider only the use of structuring function’s, which are equal for all positions in space. In mathematical morphology terms this means we use translational invariant transforms to solve the differential equation. The restriction to translation invariant transforms is the morphological interpretation of wave propagation in a homogeneous medium.

In Section VI we show that the erosion and dilation act as morphological propagators propagating the initial condition (original image in computer vision terminology) into scale-space, much as the Gaussian convolution is the propagator for the linear diffusion equation.

Some of the results derived in this paper bear resemblance to those from Gaussian scale-space theory, where Gauss convolutions are used to propagate the initial condition into the scale-space. For ease of reference we start in Section II with a short overview of the Gaussian scale-space theory (as far as it is relevant to this paper).

II. GAUSSIAN SCALE-SPACE, THE DIFFUSION EQUATION

An important property of the Gauss function is that it is Green's function (or propagator) of the *diffusion equation*. Let $x \in \mathbb{R}^2$ and $\rho \in \mathbb{R}^+$ and let $F(x, \rho)$ be a real valued function satisfying the diffusion equation:

$$F_\rho = \nabla^2 F$$

where $F_\rho = \frac{\partial F}{\partial \rho}$ and $\nabla^2 F$ is the Laplacian of F . With initial condition $F(x, 0) = f(x)$ (in computer vision f is the image to analyze) we can solve the diffusion equation using the Gauss function to propagate the initial condition into the "scale-space" $F(x, \rho)$:

$$F(x, \rho) = (f * g^\rho)(x)$$

where $*$ denotes the convolution and g^ρ is the two-dimensional Gauss function with scale parameter ρ :

$$g^\rho(x) = \frac{1}{4\pi\rho} e^{-\frac{\|x\|^2}{4\rho}}.$$

In computer vision the function $F(x, \rho)$ is interpreted as a family of images, where ρ indicates the level of resolution (or scale). The larger ρ , the more blurred the original image f is, finally showing only the larger structures in the image, until ultimately any image detail disappears.

The diffusion equation is apt to serve as a starting point to construct a scale-space (multiresolution representation) because it satisfies a *maximum principle* (see Hummel [8]). An immediate consequence of the maximum principle is that the scale-space "generated" by the diffusion equation *preserves causality in the resolution domain* [1], in the sense that moving towards larger scales no new details are introduced. It is also the maximum principle that explains the unique properties of the zero-crossings in the Laplacian of F .

Marr [9] proposed that the loci of all zero-crossings for all scales plus the gradient of F at those points provides a complete description of the original image f . This conjecture was later proven by Hummel [8]. This completeness theorem of the zero-crossings "scale-space signature" is important because it assures us that, without loss of information, we can use the signature to analyze the image instead of the original image itself.

Besides the fact that the Gauss function is the unique linear propagator that preserves causality in resolution domain (see

[1]), the Gauss function has several properties that are of practical importance. Firstly the Gauss convolution is the *unique* isotropic linear filter that can be dimensionally decomposed (separated in x - and y -direction). Secondly, the family of Gauss functions is closed with respect to convolution, i.e., the convolution of two Gauss functions is again a Gauss function: $g^\rho * g^\mu = g^{\rho+\mu}$.

In a previous paper [10] we have introduced the quadratic structuring functions (QSF) as the morphological equivalents of the Gauss functions when used as structuring functions in erosions or dilations. Two facts were shown in that paper. Firstly, the QSF's can be separated by dimension in a dilation along the x -axis followed by a dilation along the y -axis. Also the QSF is the *unique* isotropic structuring function that can be separated. Secondly, the class of quadratic structuring functions is closed with respect to dilation.

In an impressive paper, Koenderink [11] starts with the diffusion equation to derive receptive field families bearing great resemblance with the receptive fields observed in neuro physiological experiments. As a logical consequence of the linearity of the diffusion equation (i.e., linearity in the visual stimuli is assumed) the derived receptive fields form an orthogonal family of basis functions capable of describing the original image details through superposition of the receptive field responses.

All of this shows the relevance of Gaussian scale-space. However, as pointed out by Perona and Malik [12] Gaussian blurring "does not respect the natural boundaries of objects." Objects that are better left unmerged *are* merged. Furthermore edge junctions (corners) are destroyed. They introduced an inhomogeneous blurring scheme in which the amount of (infinitesimal) isotropic blurring needed to obtain $F(x, \rho + d\rho)$ from $F(x, \rho)$ is determined by the magnitude of the gradient. This leads to the following differential equation governing the scale-space: $F_\rho = \nabla \cdot (g(\|\nabla F\|) \nabla F)$ where $g(\cdot)$ is a decreasing, positive valued function. This means that blurring is small in the vicinity of edges, because there the gradient will be large. They also claim that the scale-space thus obtained satisfies a maximum principle and thus guarantees preservation of causality in the resolution domain.

Most images analyzed by either the human visual system or a computer vision system are projections of three dimensional reality on a two dimensional retina. Projective image formation makes linearity of visual stimuli a questionable assumption. The scheme of Perona & Malik may be advantageous over the original scheme, but still it does not tackle the questionable assumption of linearity in the visual stimuli.

In this paper we investigate the use of mathematical morphology to construct scale-spaces, without relying on the linearity assumption. The use of erosions and dilations to construct a scale-space is reported before [13]. These scale-spaces, however, are not—explicitly—based on a differential equation describing the evolution of images in scale-space. In this paper we will restrict ourselves to homogeneous—non-linear—differential equations, solved by morphological operators. Inhomogeneous "blurring" schemes, like the one introduced by Perona and Malik, are not considered here.

TABLE I
NOTATIONS AND DEFINITIONS OF THE
BASIC MORPHOLOGICAL SET OPERATIONS*

Name	Notation	Definition
Scaling	αX	$\alpha X = \{\alpha x \mid x \in X\}$
Translation	X_t	$X_t = \{x \mid x - t \in X\}$
Complement	X^c	$X^c = \{x \mid x \notin X\}$
Transpose	\tilde{X}	$\tilde{X} = \{x \mid -x \in X\}$
Union	$X \cup Y$	$X \cup Y = \{x \mid x \in X \text{ or } x \in Y\}$
Intersection	$X \cap Y$	$X \cap Y = \{x \mid x \in X \text{ and } x \in Y\}$
Difference	$X \setminus Y$	$X \setminus Y = X \cap Y^c$
Dilation	$X \oplus \tilde{S}$	$X \oplus \tilde{S} = \bigcup_{y \in \tilde{S}} X_y$
Erosion	$X \ominus \tilde{S}$	$X \ominus \tilde{S} = \bigcap_{y \in \tilde{S}} X_y$
Closing	$X \bullet S$	$X \bullet S = (X \oplus \tilde{S}) \ominus S$
Opening	$X \circ S$	$X \circ S = (X \ominus \tilde{S}) \oplus S$

*In this table and in the text we use X, Y and S to denote sets in n -dimensional space and we use x, y and t to denote a position vector in that space.

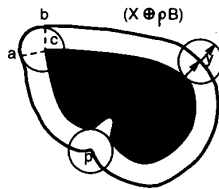


Fig. 1. Point of contact.

III. POINT OF CONTACT AND TRANSPORT IN SETS

In this section we specify the morphological notation used in this paper and discuss some properties needed. Table I summarizes the notations for the basic set operators (valid for sets in \mathbb{R}^n).

A. Point of Contact

In this section we focus attention on the calculation of the point of contact. The boundary of a set X is denoted as ∂X .

Proposition 1: Let X be a compact set and let S be a compact structuring element. Let $x \in \partial(X \oplus S)$ then \tilde{S}_x hits the boundary of X in at least one point-of-contact.

Proof: According to the definition of the dilation we have:

$$X \oplus S = \{x \mid X \cap \tilde{S}_x \neq \emptyset\}.$$

Thus for a point x on the boundary of $X \oplus S$ we have that $X \cap \tilde{S}_x \neq \emptyset$. It is evident that \tilde{S}_x must hit X in the boundary (i.e., $X \cap \tilde{S}_x \subset \partial X$). If this were not true then x could not be on the boundary of the dilated set. QED.

It is not necessarily true that \tilde{S}_x with $x \in X \oplus S$ hits ∂X in just one point. In fact it may hit ∂X in an infinite number of points or even along the entire contour of \tilde{S}_x .

As an example of the calculation of the point-of-contact, consider the dilation of a set $X \in \mathbb{R}^2$ with a disk of radius ρ (denoted as ρB), see Fig. 1.

The point $y \in \partial(X \oplus \rho B)$ is chosen such that $(\rho B)_y$ hits ∂X in one point, so that $X \oplus \rho B$ has a unique normal at y denoted as $N_{X \oplus \rho B}(y)$. It is evident that in case of a disk shaped structuring element, the point-of-contact is given by: $y - \rho N_{X \oplus \rho B}(y)$.

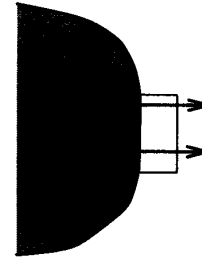


Fig. 2. Square-shaped structuring element.

It is important to note that the mapping of points on the contour of $X \oplus \rho B$ to their contact points on ∂X certainly is not a one-to-one mapping. For instance all points on $\partial(X \oplus \rho B)$ between a and b are mapped onto the same point $c \in \partial X$. Furthermore the set $(\rho B)_p$ hits ∂X in the points q and r . Note also that in p the contour $\partial(X \oplus \rho B)$ does not have a unique normal vector. Approaching p from the left, the point-of-contact approaches q while approaching p from the right gives r .

To be able to calculate the point-of-contact given the normal of the dilated set we have to limit the choice of the structuring element. As an example of a structuring element that does not allow us to calculate the point of contact, consider the square-shaped structuring element Q (see Fig. 2). Obviously the normals on one side of the square are all equal, so we cannot calculate the point-of-contact from the normal of the dilated set.

Definition 1—Strictly Convex: A closed set S is strictly convex if it is convex and if for all $x, y \in \partial S$, the line segment joining x and y only has its end points in ∂S .

Let B be the unit ball in \mathbb{R}^n : $B = \{x \mid \|x\| \leq 1\}$.

Definition 2—Regular: A compact set S is called regular if:

$$\exists \rho > 0 : S \circ \rho B = S = S \bullet \rho B$$

The existence of a $\rho > 0$ such that a set S is regular, sets a lower bound on the radius of curvature for convex ($S = S \circ \rho B$) and concave ($S = S \bullet \rho B$) portions of the boundary respectively.

Definition 3—Regular Strictly Convex: The class of all regular, strictly convex sets is denoted as S .

The notion of regular sets can be found in Serra [14]. There it is shown that all points on the boundary of a regular set have a well defined normal. In case the set is not only regular but also strictly convex we have the following more strong property.

Proposition 2—Unique Normal: Let $S \in S$ then:

$$\forall x, y \in \partial S, x \neq y : N_S(x) \neq N_S(y)$$

Proof: Because S is a convex set it can be written as the intersection of all half spaces that contain S . Furthermore because S is regular there is a well-defined normal at each point $x \in \partial S$. The normal evidently determines the tangential plane touching S in x . We define:

$$H(x) = \{z \mid N_S(x) \cdot (z - x) \geq 0\}$$

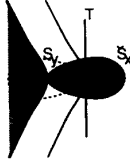


Fig. 3. Geometrical construction needed in the proof of Proposition 3.

such that:

$$S = \bigcap_{x \in \partial S} H(x).$$

Now consider two points $x, y \in \partial S$ such that $N_S(x) = N_S(y)$, then $S \subset H(x) = H(y)$. Because S is convex the line segment $L(x, y)$ connecting x with y must be in S . This line segment is entirely within $\partial H(x)$ and thus $L(x, y) \subset \partial S$. This contradicts the fact that S is strictly convex. QED.

The above proposition shows that if $S \in \mathcal{S}$ not only there is a well defined normal at each contour point, but also that the normal is unique. This means that we can define an "inverse normal" function.

Definition 4: The "inverse normal" function N_S^{-1} of a set $S \in \mathcal{S}$ maps a normal vector v ($\|v\| = 1$) onto a vector $x \in \partial S$ such that $v = N_S(x)$, i.e., $x = N_S^{-1}(v)$.

So we have a structuring element $S \in \mathcal{S}$ that has a unique normal everywhere. If that is the case, then a point on the contour of a set $X \oplus S$ can be uniquely traced back to a point on the contour of the original set X , as the following propositions say:

Proposition 3: Let $S \in \mathcal{S}$ and let X be a compact set. Consider a point $x \in \partial(X \oplus S)$ such that \check{S}_x hits ∂X in just one point y . Then:

$$N_{X \oplus S}(x) = N_S(x - y)q.$$

An equivalent property is also valid for the erosion. Consider a point $x \in \partial(X \ominus S)$ such that \check{S}_x hits ∂X in just one point y . Then:

$$N_{X \ominus S}(x) = -N_S(x - y)$$

Proof: According to the definition of the dilation we have $S_y \subset X \oplus S$. Note that $x \in \partial S_y$. Because $S \in \mathcal{S}$, S_y has a unique tangent—hyper—plane T in x (see Fig. 3). Because \check{S}_x hits ∂X in just one point y , $X \oplus S$ has a well-defined normal in x . The tangent plane T' of $X \oplus S$ in x is therefore also well-defined. It is easy to see that $T = T'$. If this were not the case then T' would intersect S_y , which would contradict the fact that $S_y \subset X \oplus S$. This means that the normal on $X \oplus S$ and S_y in point x are equal:

$$N_{X \oplus S}(x) = N_{S_y}(x)$$

Using the fact that $N_{S_y}(x) = N_S(x - y)$ this can be rewritten as:

$$N_{X \oplus S}(x) = N_S(x - y).$$

The proof for the erosion is a complete analogy of the above proof for the dilation. QED.

Now we are ready to compute the point where the structuring element makes contact.

Proposition 4—Point of Contact: Let X be a compact set and let $S \in \mathcal{S}$. If $x \in \partial(X \oplus S)$ then if the normal $N_{X \oplus S}(x)$ exists we can calculate the point of contact y as:

$$y = x - N_S^{-1}(N_{X \oplus S}(x)).$$

In case $x \in \partial(X \ominus S)$ then given the normal $N_{X \ominus S}(x)$ we can calculate the point of contact y as:

$$y = x - N_S^{-1}(-N_{X \ominus S}(x)).$$

Proof: This proposition is an immediate consequence of Proposition 3 together with the fact that $N_S(x)$ has an inverse in case $S \in \mathcal{S}$. QED.

Note very carefully that the above proposition does not give an answer for the case where $N_{X \oplus S}$ does not exist. In such a case we have a singularity in the contour.

B. Parallel Normal Transport

Thus far we have only demanded that the sets to be dilated or eroded are compact. In case we have the situation that X has a well-defined normal vector in the point-of-contact, we are able to link the normal on the dilated set with the normal vector on the original set.

This leads to the notion of parallel normal transport in the following proposition.

Proposition 5—Parallel Normal Transport: Let X be a compact set and let $S \in \mathcal{S}$. If $x \in \partial(X \oplus S)$ and the normal in x is well defined then if the original contour also has a well-defined normal in the point-of-contact $y = x - N_S^{-1}(N_{X \oplus S}(x))$ we have:

$$N_{X \oplus S}(x) = N_X(x - N_S^{-1}(N_{X \oplus S}(x))).$$

Equivalently for the erosion if the contour of X has a well defined normal in the point-of-contact $y = x - N_S^{-1}(-N_{X \ominus S}(x))$, we have:

$$N_{X \ominus S}(x) = N_X(x - N_S^{-1}(-N_{X \ominus S}(x))).$$

Proof: The proof of this proposition is almost identical to the proof of the point-of-contact propositions. Again we will only prove the proposition for the dilation. Referring to Fig. 3 we see that in this case the original contour has a well-defined normal in the point-of-contact y (and thus a tangent plane T''). Because \check{S}_x hits X in y the tangent plane of \check{S}_x in y must be equal to T'' . Thus we have:

$$N_X(y) = -N_{\check{S}_x}(y).$$

Note that the minus sign is necessary because the normals are pointing outward

$$\begin{aligned} N_X(y) &= -N_{\check{S}_x}(y - x) \\ &= N_S(x - y). \end{aligned}$$

The point-of-contact proposition states that $N_{X \oplus S}(x) = N_S(x - y)$. Using the above result we obtain:

$$N_{X \oplus S}(x) = N_X(y).$$

Substituting the expression for the point-of-contact given in Proposition 3 proves the proposition. QED.

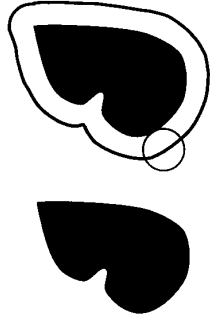


Fig. 4. Singularities created when dilating or eroding a smooth set X .

So from the proposition it appears that the normal is parallel transported to a normal on the original set: the one that is $N_S^{-1}(N_{X \oplus S}(x))$ more inwards.

Note that the normal transport is not necessarily in the direction of the normal. Only for spherical structuring elements this is true in general.

C. Characterization of Singularities

In general the dilation of a regular set X using a structuring element $S \in \mathcal{S}$ does not result in a set that is regular (see Fig. 4). From the figure it can be seen that the singularities in the dilated object are of a different character than the singularities formed in the eroded object. The dilation singularities are pointing inwards whereas the erosion singularities are pointing outwards.

To be able to treat the character of the singularities we need to define the notion of regular sets more carefully by introducing the notion of semiregular sets.

Definition 5—Semiregular Sets: A set S is called upper semi-regular (abbreviated as u.s.r.) if $\exists \rho > 0 : S \bullet \rho B = S$ and lower semi-regular (l.s.r.) if $\exists \rho > 0 : S \circ \rho B = S$. Note that S is regular if S is both upper and lower semiregular. Semiregularity is the property that is invariant under dilation as indicated in the following proposition.

Proposition 6: Let X be a compact set and let $S \in \mathcal{S}$ then $X \oplus S$ is lower semiregular and $X \ominus S$ is upper semiregular.

Proof: We only prove this for the dilation, the proof for the erosion follows from the duality. For $X \oplus S$ to be l.s.r. we must have that $\exists \rho : (X \oplus S) \circ \rho B = X \oplus S$:

$$\begin{aligned} (X \oplus S) \circ \rho B &= ((X \oplus S) \ominus \rho B) \oplus \rho B \\ &= ((S \oplus X) \ominus \rho B) \oplus \rho B \\ &\supset ((S \ominus \rho B) \oplus X) \oplus \rho B \\ &\supset ((S \ominus \rho B) \oplus \rho B) \oplus X \\ &\supset (S \circ \rho B) \oplus X \\ &\supset X \oplus (S \circ \rho B). \end{aligned}$$

Because S is regular: $\exists \rho : S \circ \rho B = S$ and thus

$$\exists \rho : (X \oplus S) \circ \rho B \supset X \oplus S.$$

On the other hand because the opening is antiextensive we know that:

$$(X \oplus S) \circ \rho B \subset X \oplus S.$$

Combining the above equations we finally obtain:

$$\exists \rho : (X \oplus S) \circ \rho B = X \oplus S$$

proving that $X \oplus S$ is l.s.r. QED.

This proposition indeed shows that if the dilation introduces singularities they must be pointing inwards. The erosion only introduces outwards pointing singularities (if any). The singularities are formed when the structuring element hits the original set in more than one point.

IV. WHERE DO POINTS GO?

In mathematical morphology any convex set is divisible with respect to the dilation (see Serra [14]), i.e.:

$$\alpha S \oplus \beta S = (\alpha + \beta)S. \quad (3)$$

It is this property of convex sets that allows us to determine where points on the contour of a set X go, when dilating X with increasingly larger structuring elements. First we trace a point on the contour of a dilated set back to a point on the original contour.

Proposition 7—Point Trace: Let X be a compact set and let $S \in \mathcal{S}$. Consider a point $x \in \partial(X \oplus S)$ such that \tilde{S}_x hits ∂X in just one point-of-contact y . Then for all $\rho \leq 1$ the set $(\rho \tilde{S})_x$ also hits $\partial(X \oplus (1 - \rho)S)$ in just one point-of-contact. Furthermore varying ρ from 0 to 1 the points-of-contact are on the straight line connecting x with y .

Proof: The point-of-contact y on ∂X is given by Proposition 3 as:

$$y = x - N_S^{-1}(N_{\partial(X \oplus S)}(x)).$$

Equivalently the point-of-contact z on $\partial(X \oplus (1 - \rho)S)$ is given by:

$$z = x - N_{\rho S}^{-1}(N_{\partial(X \oplus S)}(x)).$$

The inverse normal function $N_{\rho S}^{-1}$ equals ρN_S^{-1} as can be easily verified. This gives:

$$z = x - \rho N_S^{-1}(N_{\partial(X \oplus S)}(x)).$$

Obviously the points x , y and z are collinear. This is true for all values of $\rho \in [0, 1]$, which shows that the point y moves to x along a straight line. QED.

This proposition explains *where points come from*. In order to answer the question “where do points go?” we use the parallel normal transport proposition.

Proposition 8—Point Trajectory: Let X be a compact set and let $S \in \mathcal{S}$. Consider a point $x \in \partial X$ with a well-defined normal $N_X(x)$. Repeatedly dilating the set X with structuring element $d\rho S$ the point x “moves” in the direction $N_S^{-1}(N_X(x))$.

Proof: Let μS be a scaling of S such that there is a point y on $\partial(X \oplus \mu S)$ such that x is its unique point-of-contact. Then the point-of-contact proposition states that:

$$x = y - \mu N_S^{-1}(N_{X \oplus \mu S}(y)).$$

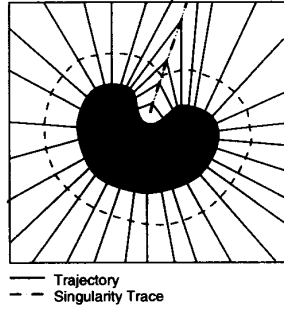


Fig. 5. Where do points go?

From the previous proposition we know that the point x moves towards y along a straight line, thus the vector from x pointing towards y is:

$$y - x = \mu N_S^{-1}(N_{X \oplus \mu S}(y)) \quad (4)$$

Because in x the normal on X is well-defined, we can use the parallel normal transport proposition:

$$N_X(x) = N_{X \oplus \mu S}(y).$$

Substituting this into (4) we obtain:

$$y - x = \mu N_S^{-1}(N_X(x)).$$

Note that the value of μ does not influence the *direction* in which the point x moves. Thus x moves in the direction: $N_S^{-1}(N_X(x))$ QED.

So each point travels with constant velocity along a straight line. Note that although the speed is constant along one trajectory, in general it will be different from trajectory to trajectory.

The above proposition shows that given the normal at a contour point x of X , the direction in which this point will move while dilating with ρS is given by the vector $N_S^{-1}(N_X(x))$. The direction of these contour point trajectories is thus completely determined by the normal on the original contour (which is a local property). However, how *far* a point will travel before it is "annihilated" by another contour point travelling in an intersecting direction, is a global property of the original set X .

The point where two (or more) contour point traces meet (for equal value of ρ) is a singularity point in the contour of $X \oplus \rho S$. In Fig. 5 the traces along which contour points move while dilating with a disk shaped structuring element are shown. In addition in the same figure we also indicate the traces of the singularities. Although we do not pursue the issue in this paper any further, we note that the figure indicates that singularity traces are very important in the description of the shape of the original set X . A singularity trace only reaches the contour of the original set in case the contour is not differentiable (no unique normal) and locally concave.

Once the trace of a point on the contour of the original set is annihilated by the trace of another contour point (say for dilation size ρ), both points no longer influence the dilation for larger sizes of the structuring element.

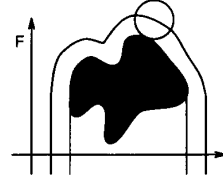


Fig. 6. Dilation of a 2-D set, using a parameterization of the contour.

Proposition 9: Let y be a singular point on the contour of $X \oplus \rho S$ then this point y never is the point-of-contact of a point x on the contour of $X \oplus \mu S$ with $\mu > \rho$.

Proof: Let $\eta = \mu - \rho$ then $X \oplus \mu S = (X \oplus \rho S) \oplus \eta S$. Using the property that $X \oplus S = (X \bullet S) \oplus S$ we have that:

$$X \oplus \mu S = ((X \oplus \rho S) \bullet \eta S) \oplus \eta S.$$

Thus for all points x on the contour of $X \oplus \mu S$ the corresponding point-of-contact y on the contour of $(X \oplus \rho S) \bullet \eta S$, must also be a point on the contour of $X \oplus \rho S$. Since $X \oplus \rho S$ is l.s.r. and $(X \oplus \rho S) \bullet \eta S$ is u.s.r. the point y has a well defined normal and therefore is not a singular point. QED.

Dilation of a set X with increasingly larger disks is also described by Kimia [15]. The contour of $X \oplus \rho S$ is described by a parameterization $C_{X,\rho}(s)$ where s is the path length. Kimia introduces the following differential equation, describing how the contour points are moved given an infinitesimal change in the "scale parameter" ρ :

$$\frac{\partial}{\partial \rho} C_{X,\rho} = N_{X,\rho}$$

where N is the outward pointing normal. This differential equation describes a dilation with a disk shaped structuring element. Locally we can always choose an extrinsic Euclidean frame to describe the contour as function F of x and ρ (see Fig. 6). Kimia shows that in this frame the above differential equation transforms to:

$$F_{x\rho} = \frac{F_x F_{xx}}{\sqrt{1 + F_x^2}}.$$

In Section VII we show that the above differential equation is a special case of the differential equations introduced in this paper (i.e., using a disk shaped structuring element).

V. POINT OF CONTACT AND TRANSPORT IN FUNCTIONS

A. Grey Value Morphology

The erosion of a grey-value image f using *structuring function* g , using \wedge as a shorthand for the infimum operator, is given by:

$$(f \ominus g)(x) = \bigwedge_{y \in \mathbb{R}^n} [f(x+y) - g(y)]. \quad (5)$$

The dilation $f \oplus g$ is defined as:

$$(f \oplus g)(x) = \bigvee_{y \in \mathbb{R}^n} [f(x+y) + g(y)], \quad (6)$$

TABLE II
NOTATION AND DEFINITIONS OF GREY-VALUE MORPHOLOGY*

Name	Notation	Definition
Translation	$f_{(x,t)}$	$f_{(x,t)}(x) = f(x-y) + t$
Complement	f^c	$f^c = -f$
Transpose	\bar{f}	$\bar{f}(x) = f(-x)$
Union	$f \vee g$	$(f \vee g)(x) = f(x) \vee g(x)$
Intersection	$f \wedge g$	$(f \wedge g)(x) = f(x) \wedge g(x)$
Dilation	$f \oplus g$	$(f \oplus g)(x) = \bigvee_{y \in \mathbf{R}^n} [f(x+y) + g(y)]$
Erosion	$f \ominus g$	$(f \ominus g)(x) = \bigwedge_{y \in \mathbf{R}^n} [f(x+y) - g(y)]$
Closing	$f \bullet g$	$(f \oplus g) \ominus g$
Opening	$f \circ g$	$(f \ominus g) \oplus g$

*In this table and in the text we use f, g and h to denote functions (e.g., $f: \mathbf{R}^n \mapsto \mathbf{R}$) and we use x and y to denote a position vector in \mathbf{R}^n , an element from \mathbf{R}^{n+1} will be denoted as $(x; t)$, where $x \in \mathbf{R}^n$ and $t \in \mathbf{R}$.

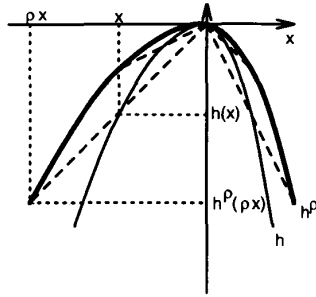


Fig. 7. Geometrical Construction needed in the proof of Proposition 10.

where \vee denotes the supremum operator. It should be noted that where in convolution kernels the pixels with value zero do not influence the convolution sum, in structuring function's these pixels have value $-\infty$. In Table II the definitions of grey-value morphology needed in this paper are summarized.

In Section IV the families of convex sets $\{\rho S\}_{\rho \in \mathbf{R}^+}$ that are closed under dilation were introduced. Evidently this property is also valid for structuring functions. The umbra $\mathbf{U}(f)$ of a function f is the $(n+1)$ -dimensional set of all points (x, t) under the graph of f : $\mathbf{U}(f) = \{(x, t) \mid f(x) \leq t\}$. We will call a structuring function h convex if its umbra $\mathbf{U}(h)$ is convex. Then we have that:

$$\alpha \mathbf{U}(h) \oplus \beta \mathbf{U}(h) = (\alpha + \beta) \mathbf{U}(h).$$

It is this property that allows us to define structuring function families that are closed under dilation.

Definition 6—Structuring Function Families: Let h be a convex structuring function, then we define the family of structuring function's $\{h^\rho\}_{\rho \in \mathbf{R}^+}$ with:

$$h^\rho(x) = \rho h\left(\frac{x}{\rho}\right).$$

Proposition 10—Closed under Dilation: A family of structuring function's $\{h^\rho\}_{\rho \in \mathbf{R}^+}$ "generated" by the convex function h is closed under dilation, i.e., $h^\rho \oplus h^\mu = h^{\rho+\mu}$.

Proof: It is easily verified that (see Fig. 7): $\mathbf{U}(h^\rho) = \rho \mathbf{U}(h)$, thus for the dilation $h^\rho \oplus h^\mu$ we have:

$$\mathbf{U}(h^\rho \oplus h^\mu) = \mathbf{U}(h^\rho) \oplus \mathbf{U}(h^\mu)$$

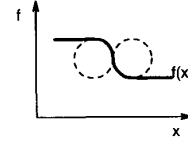


Fig. 8. A regular function need not be differentiable.

$$\begin{aligned} &= \rho \mathbf{U}(h) \oplus \mu \mathbf{U}(h^\mu) \\ &= (\rho + \mu) \mathbf{U}(h) \\ &= \mathbf{U}(h^{\rho+\mu}) \end{aligned}$$

QED.

B. Point of Contact

All propositions in Section III-A are based on the restriction to regular, strictly convex structuring elements. We start this section by generalizing these restrictions from sets to functions. Note that for functions we use B to denote the three-dimensional sphere of radius 1.

Definition 7—Regular Function: A function $f: \mathbf{R}^n \mapsto \mathbf{R}$ is called regular if $\exists \rho > 0: f = f \circ \rho B = f \bullet \rho B$. A function f for which $\exists \rho > 0: f = f \circ \rho B$ is called lower semiregular. If $\exists \rho > 0: f = f \bullet \rho B$ then f is called upper semiregular.

Note that a regular function need not be differentiable at all points. The fact that f is regular implies that the normal on $\mathbf{U}(f)$ is well defined everywhere, but the normal on the function surface can be horizontal. At those points the gradient is not defined (see Fig. 8). Also note when comparing this definition with the definition of regular sets, we have now dropped the compactness requirement.

Definition 8—Strictly Convex Structuring Functions: A structuring function f is strictly convex if its umbra $\mathbf{U}(f)$ is strictly convex.

Definition 9—Regular, Strictly Convex Structuring Functions: The class of all regular, strictly convex functions will be denoted as \mathcal{G} .

The following proposition links the normal on the set $\mathbf{U}(f)$ with the gradient of f according to standard differential geometry.

Proposition 11: Let f be continuous and differentiable at point x then the normal at point $(x; f(x)) \in \partial \mathbf{U}(f)$ is given by:

$$N_{\mathbf{U}(f)}(x; f(x)) = \frac{1}{\sqrt{1 + \|\nabla f\|^2}} (-\nabla f(x); 1)$$

Proof: The tangent plane at a point $(x; f(x))$ is described by the two direction vectors $(1, 0, f_x)$ and $(0, 1, f_y)$, with $f_x = \frac{\partial f}{\partial x}$ and $f_y = \frac{\partial f}{\partial y}$. The outward pointing normal of this plane is the vector $(-f_x, -f_y, 1)/\sqrt{1 + f_x^2 + f_y^2}$. QED.

Through this link we can generalize all propositions from set domain to the function domain. The proofs of the next three propositions are omitted as they are simple generalizations of the propositions in the set domain. The following proposition is a generalization of Proposition 3.

Proposition 12: Let $g \in \mathcal{G}$ and let $f \oplus g$ be differentiable at point x , then for the point of contact $(y; f(y))$ we have:

$$(\nabla(f \oplus g))(x) = (\nabla g)(x - y).$$

Equivalently for the erosion (demanding that $f \ominus g$ is differentiable in x) we have:

$$(\nabla(f \ominus g))(x) = -(\nabla g)(x - y).$$

The inverse normal function for strictly convex regular structuring element's is replaced with the inverse gradient function. If $g \in \mathcal{G}$ then $(\nabla g)^{-1}$ exists.

Proposition 13—Point of Contact: Let $g \in \mathcal{G}$ and let $f \oplus g$ be continuous and differentiable at point x , then the point-of-contact $(y; f(y))$ is given by:

$$\begin{aligned} y &= x - (\nabla g)^{-1}[\nabla(f \oplus g)(x)] \\ f(y) &= (f \oplus g)(x) - g(x - y). \end{aligned}$$

Equivalently for the erosion the point-of-contact is given by:

$$\begin{aligned} y &= x - (\nabla g)^{-1}[-\nabla(f \ominus g)(x)] \\ f(y) &= (f \ominus g)(x) + g(x - y). \end{aligned}$$

C. Parallel Normal Transport

Whereas the point-of-contact propositions only demand that $f \oplus g$ (respectively $f \ominus g$) are continuous and differentiable in x , the generalization of the parallel normal transport proposition requires in addition that the original function f is continuous and differentiable in the point-of-contact.

Proposition 14—Parallel Normal Transport: Let $g \in \mathcal{G}$ and let $f \oplus g$ be continuous and differentiable at point x , then if f is continuous and differentiable in the point-of-contact we have:

$$\nabla(f \oplus g)(x) = (\nabla f)[x - (\nabla g)^{-1}[\nabla(f \oplus g)(x)]].$$

Equivalently for the erosion we have:

$$\nabla(f \ominus g)(x) = (\nabla f)[x - (\nabla g)^{-1}[-\nabla(f \ominus g)(x)]].$$

D. Characterization of Singularities

In all propositions in the previous two sections we have demanded the eroded or dilated function to be differentiable at the point of interest. For the parallel normal transport we have also demanded differentiability at the point of contact of the original function. Note that we have not yet shown that these demands are met anywhere. Conditions allowing us to use these propositions at almost all points are given in this section.

We already indicated that regularity of a function does not imply that the function is differentiable (see Fig. 8). This means that singularities need not coincide with the nondifferentiable points of a function. In this paper we restrict ourselves to *Lipschitz* functions with the desirable property that they are almost everywhere differentiable (see [16]) and that the derivatives are bounded functions.

Definition 10—Lipschitz Functions: A function f is a member of the class of Lipschitz functions \mathcal{L} if $\forall x \neq y$:

$$\frac{|f(x) - f(y)|}{\|x - y\|} \leq k$$

where k is a finite real value.

The Lipschitz condition has a relationship with mathematical morphology through the cone-shaped structuring function.

Definition 11—Cone Structuring Function: The cone-shaped structuring function $\rho_c(x)$ with parameter of width ρ is defined as:

$$\rho_c(x) = -\rho\|x\|.$$

The cone function has a few straightforward properties we need in the following propositions:

Proposition 15—Cone Functions: The dilation of two cones is equal to the larger of the two: $\rho_c \oplus \mu_c = \rho \wedge \mu_c$. Closing an arbitrary function f with a cone is equal to the dilation of f with the cone: $f \bullet \rho_c = f \oplus \rho_c$. Opening an arbitrary function f with a cone is equal to the erosion of f with the cone: $f \circ \rho_c = f \ominus \rho_c$.

In this paper we need a more precise definition of the Lipschitz condition where we distinguish between upper and lower semi-Lipschitz, in the same way as done for regularity of sets and functions.

Definition 12—Semi-Lipschitz: A function f is called lower-semi-Lipschitz if $\exists \rho \in \mathbf{R}^+ : f = f \ominus \rho_c$. Equivalently, a function f is upper semi-Lipschitz if $\exists \rho \in \mathbf{R}^+ : f = f \oplus \rho_c$.

Proposition 16: A function f is Lipschitz iff it is both lower semi-Lipschitz as well as upper semi-Lipschitz.

Proof: If f is u.s.l. then:

$$\exists \rho : f \oplus \rho_c = f.$$

Let k^+ be that value of ρ , so we have:

$$\forall x : f(x) = \bigvee_y [f(y) - k^+\|x - y\|]$$

or equivalently:

$$\forall x, y : f(x) \geq f(y) - k^+\|x - y\|$$

or:

$$\forall x, y : \frac{f(x) - f(y)}{\|x - y\|} \geq -k^+. \quad (7)$$

Because f is also l.s.l. we also have: $\exists \rho : f \ominus \rho_c = f$. Let k^- be that value of ρ , then we have:

$$\forall x, y : f(x) \leq f(y) + k^-\|x - y\|$$

or:

$$\forall x, y : \frac{f(x) - f(y)}{\|x - y\|} \leq k^-. \quad (8)$$

Combining (7) and (8) and using $k = k^+ \vee k^-$ we obtain:

$$\forall x, y : \frac{f(x) - f(y)}{\|x - y\|} \leq k.$$

This is of course equivalent with the Lipschitz condition. QED.

Every Lipschitz function f is characterized with two scalar values $f_{\mathcal{L}}^- = \min\{\rho \mid f = f \ominus \rho_c\}$ and $f_{\mathcal{L}}^+ = \min\{\rho \mid f = f \oplus \rho_c\}$.

The importance of Lipschitz functions in mathematical morphology is that the class of Lipschitz functions is closed under erosions and dilations *irrespective of the structuring function that is used*.

Proposition 17: Let $f \in \mathcal{L}$ then for any structuring function g , $f \oplus g \in \mathcal{L}$ and $f \ominus g \in \mathcal{L}$.

Proof: We only give the proof for the dilation. Obviously if f is Lipschitz then also $-f$ is Lipschitz. Thus the duality can be simply used to show that $f \ominus g \in \mathcal{L}$ if we have shown that $f \oplus g \in \mathcal{L}$. For the dilation we have:

$$(f \oplus g) \oplus \rho c = (f \oplus \rho c) \oplus g.$$

Because $f \in \mathcal{L}$ we have:

$$\rho \geq f_{\mathcal{L}}^+ : f \oplus \rho c = f$$

and thus:

$$\rho \geq f_{\mathcal{L}}^+ : (f \oplus g) \oplus \rho c = (f \oplus \rho c) \oplus g = f \oplus g \quad (9)$$

showing that $f \oplus g$ is u.s.l.

For the erosion $(f \oplus g) \ominus \rho c$ we have:

$$(f \oplus g) \ominus \rho c = (f \oplus g) \circ \rho c \leq f \oplus g \quad (10)$$

where we use Proposition 15 and the antiextensivity of the opening. We also know that in general $(f \oplus g) \ominus h \geq (f \ominus h) \oplus g$. This means that:

$$(f \oplus g) \ominus \rho c \geq (f \ominus \rho c) \oplus g.$$

Because f is Lipschitz we have:

$$\rho \geq f_{\mathcal{L}}^- : f \ominus \rho c = f$$

and thus:

$$\rho \geq f_{\mathcal{L}}^- : (f \oplus g) \ominus \rho c \geq f \oplus g$$

combined with (10) this gives:

$$\rho \geq f_{\mathcal{L}}^- : (f \oplus g) \ominus \rho c = f \oplus g.$$

With (9) we have:

$$\rho \geq (f_{\mathcal{L}}^- \vee f_{\mathcal{L}}^+) : (f \oplus g) \oplus \rho c = (f \oplus g) \ominus \rho c = f \oplus g$$

showing that $f \oplus g$ is Lipschitz. QED.

The fact that the class of Lipschitz functions is invariant under morphological operations (erosions and dilations) and the fact that Lipschitz functions are differentiable almost everywhere, assures us that the point-of-contact propositions from the previous section can be used almost everywhere.

In Section III-C we have shown that when we dilate a set X with a structuring element $S \in \mathcal{S}$ the resulting set is lower semiregular, indicating that the singularities are pointing inwards. Equivalently we have shown that erosion of a set X results in an upper semiregular set, with the singularities pointing outwards. These results are easily generalized to the function domain:

Proposition 18: Let f be a Lipschitz function and let $g \in \mathcal{G}$ then $f \oplus g$ is lower semiregular, and $f \ominus g$ is upper semiregular.

Proof: The proof of this proposition is a complete analogy of the proof of Proposition 6. QED.

We see that dilation introduces only downwards pointing singularities and erosion introduces only upwards pointing singularities. The class of all lower semiregular Lipschitz functions is denoted as \mathcal{L}_l , equivalently \mathcal{L}_u denotes the class of all upper semiregular Lipschitz functions.

The Lipschitz condition assures us that the eroded and dilated functions are almost everywhere differentiable (allowing us to use the point-of-contact propositions), and the above proposition shows that all nondifferentiable points are upwards respectively downwards pointing singularities (if we started with a lower semiregular respectively upper semiregular function).

VI. THE EVOLUTION OF IMAGES: MORPHOLOGICAL PROPAGATORS

In this section we derive the differential equations, describing the changes in a function f when dilating or eroding f with a family of increasingly larger structuring function's $g^\rho \in \mathcal{G}$. We define the scale-functions F^\oplus and F^\ominus :

$$F^\oplus(x, \rho) = (f \oplus g^\rho)(x)$$

$$F^\ominus(x, \rho) = (f \ominus g^\rho)(x).$$

If F^\oplus and F^\ominus are the solutions of differential equations, then obviously they cannot be "classical solutions" because the erosion and dilation are not differentiable everywhere (the singular points being the exceptions). Such "almost-everywhere-solutions" are called solutions in the distribution sense (see Lax [17]) or *weak solutions* (see Smoller [18]).

In the previous paragraph we have introduced the collection \mathcal{G} of regular, strictly convex structuring functions. For a structuring function $g \in \mathcal{G}$ we know that $(\nabla g)^{-1}$ exists but furthermore:

Proposition 19: Let $g \in \mathcal{G}$ then:

$$(\nabla g^\rho)^{-1} = \rho(\nabla g)^{-1}.$$

Proof: Let $h(x) = g^\rho(x)$ i.e., $h(x) = \rho g(x/\rho)$. Thus $\nabla h(x) = (\nabla g^\rho)(x) = \nabla g(x/\rho)$. Let $y = \nabla g^\rho(x)$ then $x = (\nabla g^\rho)^{-1}(y)$, but also $y = \nabla g(x/\rho)$ thus $x/\rho = (\nabla g)^{-1}(y)$ or $x = \rho(\nabla g)^{-1}(y)$. Therefore: $(\nabla g^\rho)^{-1} = \rho(\nabla g)^{-1}$. QED.

Proposition 20: Let $f \in \mathcal{L}_l$ and $g \in \mathcal{G}$ then the function

$$F^\oplus(x, \rho) = (f \oplus g^\rho)(x)$$

is a weak solution of the differential equation:

$$F_\rho^\oplus = -(\nabla g)^{-1}(\nabla F^\oplus) \cdot \nabla F^\oplus + g((\nabla g)^{-1}(\nabla F^\oplus))$$

with initial data $F^\oplus(x, 0) = f(x)$.

Let $f \in \mathcal{L}_u$ and $g \in \mathcal{G}$ then the function

$$F^\ominus(x, \rho) = (f \ominus g^\rho)(x)$$

is a weak solution of the differential equation:

$$F_\rho^\ominus = -(\nabla g)^{-1}(-\nabla F^\ominus) \cdot \nabla F^\ominus - g((\nabla g)^{-1}(-\nabla F^\ominus))$$

with initial data $F^\ominus(x, 0) = f(x)$.

Proof: We start with the result from Proposition 13 using $h = f \oplus g^\rho$:

$$\begin{aligned} h(x - (\nabla g^{d\rho})^{-1}(\nabla(h \oplus g^{d\rho})(x))) \\ = (h \oplus g^{d\rho})(x) - g^{d\rho}((\nabla g^{d\rho})^{-1}(\nabla(h \oplus g^{d\rho})(x))). \end{aligned}$$

Note that this equation is valid almost everywhere because $f \in \mathcal{L}_l$ and $g \in \mathcal{G}$. Now let $F(x, \rho) = h(x) = (f \oplus g)(x)$ then $(h \oplus g^{d\rho})(x) = (f \oplus g^{\rho+d\rho})(x) = F^\oplus(x, \rho + d\rho)$, then we can rewrite the above expression as:

$$\begin{aligned} F^\oplus(x - (\nabla g^{d\rho})^{-1}(\nabla F^\oplus(x, \rho + d\rho)), \rho) \\ = F^\oplus(x, \rho + d\rho) - g^{d\rho}((\nabla g^{d\rho})^{-1}(\nabla F^\oplus(x, \rho + d\rho))) \quad (11) \end{aligned}$$

Definition 6 states that $g^{d\rho}(x) = d\rho g(x/d\rho)$. Proposition 19 states that $(\nabla g^{d\rho})^{-1} = d\rho(\nabla g)^{-1}$. Using the Taylor series of a function (and neglecting second and higher order terms):

$$F^\oplus(x + dx, \rho) = F^\oplus(x, \rho) + dx \cdot \nabla F^\oplus(x, \rho)$$

we can rewrite (11) (omitting all arguments (x, ρ) of the function F):

$$\begin{aligned} F^\oplus - d\rho(\nabla g)^{-1}(\nabla F^\oplus) \cdot \nabla F^\oplus \\ = F^\oplus + d\rho F_\rho^\oplus - d\rho g((\nabla g)^{-1}(\nabla F^\oplus)). \end{aligned}$$

Rewriting and division by $d\rho$ gives:

$$F_\rho^\oplus = -(\nabla g)^{-1}(\nabla F^\oplus) \cdot \nabla F^\oplus + g((\nabla g)^{-1}(\nabla F^\oplus)).$$

Using the duality between erosion and dilation (which amounts to the replacement of F^\oplus with $-F^\ominus$) proves the second part of the proposition. QED.

In a previous section we have shown that when dilating $f \in \mathcal{L}_l$, the structuring function g never hits f in a singularity point. This means that g hits f at a point where f is differentiable and thus we can use the parallel normal transport propositions. The matrix $(\nabla \nabla^T F)$ is known as the Hessian matrix:

$$(\nabla \nabla^T F) = \begin{bmatrix} F_{xx} & F_{xy} \\ F_{xy} & F_{yy} \end{bmatrix}.$$

This leads to the following differential equations:

Proposition 21: Let $f \in \mathcal{L}_l$ and $g \in \mathcal{G}$ then the function

$$F^\oplus(x, \rho) = (f \oplus g^\rho)(x)$$

is a weak solution of the differential equation:

$$(\nabla F^\oplus)_\rho = -(\nabla \nabla^T F^\oplus)(\nabla g)^{-1}(\nabla F^\oplus)$$

with initial data $F^\oplus(x, 0) = f(x)$.

Let $f \in \mathcal{L}_u$ and $g \in \mathcal{G}$ then the function

$$F^\ominus(x, \rho) = (f \ominus g^\rho)(x)$$

is a weak solution of the differential equation:

$$(\nabla F^\ominus)_\rho = -(\nabla \nabla^T F^\ominus)(\nabla g)^{-1}(-\nabla F^\ominus)$$

with initial data $F^\ominus(x, 0) = f(x)$.

Proof: The parallel normal transport proposition (Proposition 14) stated:

$$\nabla(h \oplus g^{d\rho})(x) = \nabla F^\oplus(x - (\nabla g^{d\rho})^{-1}(\nabla(f \oplus g)(x))).$$

Using $h(x) = F^\oplus(x, \rho) = (f \oplus g)(x)$ we can rewrite this as:

$$\nabla F^\oplus(x, \rho + d\rho) = \nabla F^\oplus(x - (\nabla g^{d\rho})^{-1}(\nabla F^\oplus(x, \rho + d\rho))).$$

Note that in this case ∇F^\oplus is a vector function and the Taylor expansion to be used is:

$$\nabla F^\oplus(x + dx, \rho) = \nabla F^\oplus(x, \rho) + (\nabla \nabla^T F^\oplus) dx.$$

Using this expansion we obtain (again omitting all arguments (x, ρ)):

$$\nabla F^\oplus + d\rho(\nabla F^\oplus)_\rho = \nabla F^\oplus - d\rho(\nabla \nabla^T F^\oplus)(\nabla g)^{-1}(\nabla F^\oplus).$$

Subtracting ∇F^\oplus from both sides and dividing by $d\rho$ we arrive at:

$$(\nabla F^\oplus)_\rho = -(\nabla \nabla^T F^\oplus)(\nabla g)^{-1}(\nabla F^\oplus).$$

Using the duality between the erosion and dilation the second part of this proposition easily follows from the above. QED.

VII. TWO EXAMPLES

This section uses two specific structuring functions to illustrate the differential equations derived in the previous section. First consider the spherically-shaped function $b(x)$:

$$b(x) = \begin{cases} \sqrt{1 - \|x\|^2} & : \|x\| \leq 1 \\ -\infty & : \|x\| > 1 \end{cases}.$$

Note that $b \in \mathcal{G}$ and that b has a compact support. The gradient in the interval $\|x\| < 1$ is given by:

$$\nabla b(x) = -\frac{x}{\sqrt{1 - \|x\|^2}}$$

with "inverse gradient":

$$(\nabla b)^{-1}(x) = -\frac{x}{\sqrt{1 + \|x\|^2}}.$$

Now define $F^\oplus = (f \oplus b^\rho)(x)$. Using the propositions from the previous section we can determine the differential equation's for which these functions are the weak solution. The dilation F^\oplus satisfies

$$F_\rho^\oplus = \sqrt{1 + \|\nabla F^\oplus\|^2}$$

as well as:

$$(\nabla F^\oplus)_\rho = \frac{(\nabla \nabla^T F^\oplus) \nabla F^\oplus}{\sqrt{1 + \|\nabla F^\oplus\|^2}}. \quad (12)$$

The spherical structuring element is used by Kimia [15] to dilate 2-D objects. Kimia derived the following differential equation (see Section IV):

$$F_{x\rho}^\oplus = \frac{F_x^\oplus F_{xx}^\oplus}{\sqrt{1 + (F_x^\oplus)^2}}.$$

Note that this differential equation is the simplification of (12) for a one-dimensional function.

In the second example we consider the important class of Quadratic Structuring Functions. The QSF's are important from a practical point of view because the parabolic QSF $q(x) = -\|x\|^2/4$ is the unique structuring function that can be separated by dimension. Defining $F^\oplus(x, \rho) = (f \oplus q^\rho)(x)$, this function satisfies:

$$F_\rho^\oplus = \|\nabla F^\oplus\|^2$$

and

$$(\nabla F^\oplus)_\rho = 2(\nabla \nabla^T F^\oplus) \nabla F^\oplus.$$

VIII. SUMMARY AND DISCUSSION

In this paper we have introduced the differential equations that are solved using morphological propagators. The class of nonlinear differential equations bears great resemblance with the Burgers equation governing the propagation of (shock) waves [19] (see [20]). Indeed the geometrical construction used by Burgers to solve "his" equation would nowadays be called a dilation¹.

Where Do Points Go? In Section IV we have shown that given the normal at a point on the contour of a set, we can determine the direction in which this point will move when dilating with increasingly larger structuring elements. To be able to do so, the structuring element must have a well-defined and unique normal at each point of its contour. Although we know the direction in which the point will move, we do not know *how far* it will move, before it is annihilated by another point moving in an intersecting direction.

The point-of-contact propositions from Section IV not only allow us to look forward, but given the normal at a contour point of the dilated set we can also calculate where that point came from (its point-of-contact on the original contour).

Algorithmic Consequences: The point-of-contact propositions give a theoretical foundation for the grass fire implementation of the repeated dilation, erosion [21] and skeleton [22]. These algorithms store the coordinates on the contour in a list. When dilating the object only the points on the list have to be processed. The point-of-contact propositions state that we only have to dilate a point in the direction in which that point will move. For discrete images we have to consider a cone of directions in which to dilate the point. This is necessary because the normal can not be accurately determined in a small neighborhood (say 3×3), secondly the contour itself is also discretized, and thirdly the set to be dilated may not have a unique normal. The worst case is a set containing only one point necessitating the dilation with the entire structuring element.

Construction of the Scale Function: The key notion in this paper is the scale function $F(x, \rho)$ and how it is derived from the image. Given the image $f(x)$, there is not just one unique $F(x, \rho)$, but its shape depends on the structuring elements with which the image is transformed. We demand that structuring

functions are members of a family generated by a convex function.

Such a family is closed under dilation. An important example is the family of the parabolic or quadratic structuring functions. The members of a structuring function family are characterized with one scale parameter ρ . For a parabolic function the value of ρ corresponds with the width of the parabola. For very narrow parabolas the scale parameter is close to 0. In fact, if the scale $\rho = 0$, the structuring function reduces to a delta needle and $F(x, 0) = f(x)$. For members with a wide range the scale parameter is high.

The erosion scale function $F^\ominus(x, \rho)$ is obtained by eroding the original image f with structuring function g^ρ : $F^\ominus(x, \rho) = (f \ominus g^\rho)(x)$. The dilation scale function $F^\oplus(x, \rho)$ is obtained by: $F^\oplus(x, \rho) = (f \oplus g^\rho)(x)$. Whereas the scale function obtained with Gaussian convolution treats object (light blobs) and background (dark blobs) alike, with the morphological scale functions we are able to make a distinction between the two.

For increasing values of ρ the scale in the resulting image, (either F^\oplus or F^\ominus) decreases as the original image f is processed with a structuring function with increasing size. So, F^\oplus and F^\ominus are two sequences of images each derived from the original image and ordered with respect to their internal scale.

Differentially Computable: The sequence F^\oplus is not just computable from the original image ($F^\oplus(x, \rho) = (f \oplus g^\rho)(x)$), but also from the previous entry in the scale sequence: $F^\oplus(x, \rho) = (f \oplus g^{\rho-d\rho}) \oplus g^{d\rho}$. This property is also valid for the scale sequence obtained by erosion. In conclusion, the scale sequences are differentially computable just as the Gaussian scale space is in the linear case.

The Way Function Values Evolve: Proposition 13 indicates how a specific configuration of $f(x)$ will develop under dilation with a structuring function g^ρ . Proposition 13 can be interpreted as follows: where a function value $f(x)$ at point x is touched by a structuring function g^ρ it causes another point y to assume a value $f(y)$. All values $f(y)$ together form $F(x, \rho)$.

Some points x in $f(x)$ have function values that cannot be reached by g^ρ . They do not lead to a new point in $F(x, \rho)$. In other words, x does not develop at scale ρ . Such a point forms a singularity of scale ρ . We show in this paper that such singularities always point downwards for dilation (upwards for an erosion) for any structuring function $g^\rho \in \mathcal{G}$.

The transport of a function value also holds when the initial image is not $f(x)$ but another entry in $F(x, \rho)$. From the differential computability forming $F(x, \rho)$ in small steps with structuring function $g^{d\rho}$, a trace can be formed starting at $f(x)$ and following the function values induced by $f(x)$ over ρ . The same can be done by following the location of the trace of singularities over scale, which appears to be an important notion in the characterization of $f(x)$.

In this regard the singularity trace takes the position of the traces of the minima and maxima in linear Gaussian scale space. The latter evolve over scale until a minimum merges into a maximum. In the dilation scale function, $F^\oplus(x, \rho)$, the local maxima do not move however. We conjecture that the

¹For this reason we would like to call the morphological propagators the Burgers propagators.

singularities always move to the local, nonsingular maxima to disappear there. The evolution of a regular point over scale is different from the linear Gaussian scale space. Here a point is generated by the value of one precisely known and unique other point. In Gaussian scale space the function value is the sum of all surrounding values weighted by the Gaussian blur.

Dilation and Erosion Scale-Space: Up to this point, $F(x, \rho)$ is introduced as a scale function. For $F(x, \rho)$ to be a scale space the evolution over the scale should be governed by a differential equation that in addition obeys the maximum principle. The differential equation depends on the shape of the structuring function g . The differential equation in its most general form is given in Propositions 20 and 21.

It is required that the inverse gradient function $(\nabla g)^{-1}$ exists. This requirement guarantees that differentiable points on the dilated surface can be uniquely traced back to the original contour. The requirement in this form excludes "flat" structuring functions. Because of the flat shape the normal is not unique, and therefore the dilation cannot be traced back.

However, the possibility to trace a point back over scale, is not necessary for a differential equation to exist. This is shown by Brocket and Maragos [4] who have derived the partial differential equations solved with morphological erosions and dilations with respect to specific structuring functions.

In a sequel paper [23] it will be shown that the right hand side in the differential equation:

$$F_\rho^\oplus = -(\nabla g)^{-1}(\nabla F^\oplus) \cdot \nabla F^\oplus + g((\nabla g)^{-1}(\nabla F^\oplus))$$

equals the Legendre transform of the strictly convex function g . For convex functions that are not strictly convex, the Legendre transform can be generalized in order to derive the differential equation's solved with erosions/dilations with respect to flat structuring functions. This result shows that the analysis in this paper can be easily generalized to flat structuring functions.

Maximum Principle: We conjecture that the differential equations governing the functions $F(x, \rho)$ satisfy a maximum principle. For the two examples given in Section VII this conjecture is easily verified because for the dilation with the parabolic structuring function we have calculated:

$$F_\rho^\oplus = \|\nabla F^\oplus\|^2$$

showing that F^\oplus only increases with increasing scale. The consequence of the maximum principle will be that the singularity traces with increasing scale will never split into two. In contrast with the Gaussian scale space, singularity traces may start at scale unequal zero. This is equivalent with the fact that skeleton branches do not need to start at the boundary of a set.

In Fig. 9 an example of a singularity trace through the dilation scale-space is shown. The singularities at each scale level are superimposed on the original image (the singularities are drawn 3 pixels thick so that they become clearly visible). Note that the thin leg of the chair (designed by G. T. Rietveld) only shows up at lower scales, it completely disappears at higher scales.

It is our belief that the results presented in this paper barely scratch the surface of the possible integration of differential

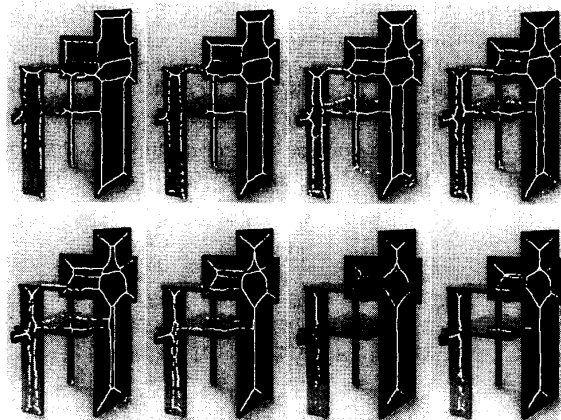


Fig. 9. Singularity trace through scale space. From left to right, top to bottom, the scales range from 1 to 8 (integer valued).

analysis (including differential geometry) with mathematical morphology. It is our hope that the equivalence between Gaussian scale-space theory and morphological scale-space can be extended even further.

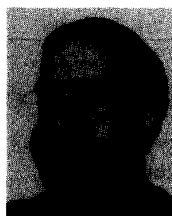
ACKNOWLEDGMENT

This paper could not have been written without the invaluable help of several people. First of all the authors would like to thank Henk Heijmans for being so kind as to share his mathematical expertise with us. The authors express their gratitude to Prof. E. M. de Jager for his help to make us understand the physics "behind" the Burgers equation. Also, the stimulating discussions with Marcel Worring about differential geometry are gratefully acknowledged.

REFERENCES

- [1] J. J. Koenderink, "The structure of images," *Biological Cybernetics*, vol. 50, no. 363-370, 1984.
- [2] R. van den Boomgaard and A. W. M. Smeulders, "The morphological structure of images," in *11th IAPR Int. Conf. Pattern Recognition*, The Hague, 1992, vol. III, pages 268-271.
- [3] R. van den Boomgaard and A. W. M. Smeulders, "Towards a morphological scale-space theory," *Proc. Workshop "Shape in Picture"*, Driebergen, The Netherlands, 7-11 September, 1992.
- [4] R. Brocket and P. Maragos, "Evolution equations for continuous-scale morphology," in *IEEE Int. Conf. Acoustics, Speech and Signal Processing*, San Francisco, March 1992.
- [5] G. Matheron, "Examples of topological properties of skeletons," in *Mathematical Morphology and Image Analysis*, vol. II, J. Serra, Ed. New York: Academic Press, 1988.
- [6] B. J. H. Verwer, "Distance transform: Metrics, algorithms and applications," Ph.D. thesis, Delft University of Technology, 1991.
- [7] H. Blum, "An associative machine for dealing with the visual field and some of its biological implications," in *Biological Prototypes and Synthetic Systems*, E. E. Bernard and M. R. Kare, Eds. New York: Plenum Press, 1961, pp. 244-260.
- [8] R. Hummel and R. Moniot, "Reconstructions from zero crossings in scale space," *IEEE ASSP mag.*, vol. 37, p. 2111, 1989.
- [9] D. Marr, *Vision*. San Francisco, CA: W. H. Freeman and Company, 1982.
- [10] R. van den Boomgaard, "Dimensional decomposition of structuring functions," submitted for publication.
- [11] J. J. Koenderink and A. J. van Doorn, "Receptive field families," *Biol. Cybern.*, 1990.

- [12] P. Perona and J. Malik, "Scale-space and edge detection using anisotropic diffusion," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 12, no. 7, p. 629, 1990.
- [13] M. H. Chen and P. F. Yan, "A multiscaling approach based on morphological filtering," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 11, no. 7, pp. 694-700, 1989.
- [14] J. Serra, *Image Analysis and Mathematical Morphology*. London: Academic Press, 1982.
- [15] B. B. Kimia, A. Tannenbaum, and S. W. Zucker, "Toward a computational theory of shape: An overview," in *Computer Vision—ECCV '90, First European Conference on Computer Vision*, Antibes, France, April 23-27, 1990, *Lecture Notes in Computer Science*, no. 427.
- [16] G. Matheron, *Random Sets and Integral Geometry*. New York: John Wiley and Sons, 1975.
- [17] P. D. Lax, "Hyperbolic Systems of Conservation Laws and the Mathematical Theory of Shock Waves," *SIAM*, 1973.
- [18] J. Smoller, *Shock-Waves and Reaction Diffusion Equations*. New York: Springer Verlag, 1983.
- [19] J. M. Burgers, *The Nonlinear Diffusion Equation: Asymptotic Solutions and Statistical Problems*. D. Reidel Publishers, 1974.
- [20] R. van den Boomgaard, "Mathematical morphology: Extensions towards computer vision," Ph.D. thesis, University of Amsterdam, March 1992.
- [21] B. J. H. Verwer and L. J. van Vliet, "A contour processing method for fast binary neighbourhood operations," *Pattern Recognition Letters*, 1988.
- [22] L. Vincent, "Efficient computation of various types of skeletons," in *SPIE Medical Imaging V*, San Jose, CA: SPIE, 1991.
- [23] L. Dorst and R. van den Boomgaard, "Morphological signal, processing and the slope transform," *Signal Processing*, vol. 38, pp 79-98, 1994.



eigenfunctions of mathematical morphology.

Rein van den Boomgaard received the M.Sc. degree from Delft University of Technology in 1988 and the Ph.D. degree from the university of Amsterdam in 1992. He is currently an associate professor of Computer Science at the University of Amsterdam. His main research interest is the theoretical foundation of mathematical morphology, notably the development of analytical techniques to bring it on a par with linear signal processing theory. He is currently working on morphological scale spaces and on the decomposition of shapes into the



aspects of image processing, as well as in the completion of the theory of mathematical morphology as a full theory of low-level vision. He has published a hundred papers since 1975.

Arnold W. M. Smeulders (S'80-M'82) received the M.Sc. degree in physics in 1975, and has been involved in image processing since that time. He is currently a professor of computer science, specializing in multimedia information systems, and a professor of biology, specializing in the processing of 3-D microscopical images. Initially his research interests were in accurate and precise measurements from digital images, but his current research interests are in the fundamentals of image databases, intelligent interaction systems, and system design