Theoretical Foundations for Spatially Discrete 1-D Shock Filtering

Martin Welk, Joachim Weickert, Irena Galić

Mathematical Image Analysis Group, Faculty of Mathematics and Computer Science, Bldg. E11, Saarland University, 66041 Saarbrücken, Germany
{welk, weickert, galic}@mia.uni-saarland.de

Abstract

While shock filters are popular morphological image enhancement methods, no well-posedness theory is available for their corresponding partial differential equations (PDEs). By analysing the dynamical system of ordinary differential equations that results from a space discretisation of a PDE for 1-D shock filtering, we derive an analytical solution and prove well-posedness. We show that the results carry over to the fully discrete case when an explicit time discretisation is applied. Finally we establish an equivalence result between discrete shock filtering and local mode filtering.

Keywords: Shock filters, analytical solution, well-posedness, dynamical systems, mode filters

1 Introduction

Shock filters are morphological image enhancement methods where dilation is performed around maxima and erosion around minima. Iterating this process leads to a segmentation with piecewise constant segments that are separated by discontinuities, so-called shocks. This makes shock filtering attractive for a number of applications where edge sharpening and a piecewise constant segmentation is desired.

In 1975 the first shock filters have been formulated by Kramer and Bruckner in a fully discrete manner [8], while first continuous formulations by means of partial differential equations (PDEs) have been developed in 1990 by Osher and Rudin [10]. The relation of these methods to the discrete Kramer–Bruckner filter became clear several years later [6, 14]. PDE-based shock filters have been investigated in a number of papers. Many of them proposed modifications with higher robustness under noise [1, 4, 7, 9, 14], but also coherence-enhancing shock filters [19] and numerical schemes have been studied [13].

Let us consider some continuous $d$-dimensional initial image $f : \mathbb{R}^d \to \mathbb{R}$. In the simplest case of a PDE-based shock filter [10], one obtains a filtered version $u(x,t)$ of $f(x)$ by solving the evolution equation

$$\partial_t u = - \text{sgn}(\Delta u) |\nabla u| \quad (t \geq 0)$$

with $f$ as initial condition, i.e. $u(x, 0) = f(x)$. Experimentally one observes that within finite “evolution time” $t$, a piecewise constant, segmentation-like result is obtained (see Fig. 1).
Specialising to the one-dimensional case, we obtain

\[
\partial_t u = -\text{sgn}(\partial_{xx} u) |\partial_x u| = \begin{cases} 
|\partial_x u|, & \partial_{xx} u < 0, \\
-|\partial_x u|, & \partial_{xx} u > 0, \\
0, & \partial_{xx} u = 0.
\end{cases}
\]

(1)

It is clearly visible that this filter performs dilation \(\partial_t u = |\partial_x u|\) in concave segments of \(u\), while in convex parts the erosion process \(\partial_t u = -|\partial_x u|\) takes place. The time \(t\) specifies the radius of the interval (a 1-D disk) \([-t, t]\) as structuring element. For a derivation of these PDE formulations for classical morphological operations, see e.g. \cite{2}.

While there is clear experimental evidence that shock filtering is a useful operation, no analytical solutions and well-posedness results are available for PDE-based shock filters. In general this problem is considered to be too difficult, since shock filters have some connections to classical ill-posed problems such as backward diffusion [10, 9].

The main goal of the present paper is to show that it is possible to establish analytical solutions and well-posedness as soon as we study the semidiscrete case with a spatial discretisation and a continuous time parameter \(t\). This case is of great practical relevance, since digital images already induce a natural space discretisation. For the sake of simplicity we restrict ourselves to the 1-D case. We also show that these results carry over to the fully discrete case with an explicit (Euler forward) time discretisation, and we establish an equivalence result between shock filtering and a specific image enhancement method called mode filtering.

Our paper is organised as follows: In Section 2 we present an analytical solution and a well-posedness proof for the semidiscrete case, whereas corresponding fully discrete results are given in Section 3. An equivalence result between shock and mode filters is presented in Section 4, and the paper is concluded with a summary in Section 5.

A shorter version of the current paper appeared at the 2005 International Symposium on Mathematical Morphology [20]. The present paper extends these results by investigations on signals with linear pixels and by establishing an equivalence between shock and mode filters.
2 The Semidiscrete Model

2.1 Problem Statement

In this section, we are concerned with a spatial discretisation of (1) which we will describe now. The time variable remains continuous here. Throughout the paper, the notion semidiscrete will refer to this setting.

**Problem.** Let \( \ldots, u_0(t), u_1(t), u_2(t), \ldots \) be a time-dependent bounded real-valued signal which evolves according to

\[
\dot{u}_i = \begin{cases} 
\max(u_{i+1} - u_i, u_{i-1} - u_i), & 2u_i > u_{i+1} + u_{i-1}, \\
\min(u_{i+1} - u_i, u_{i-1} - u_i), & 2u_i < u_{i+1} + u_{i-1}, \\
0, & 2u_i = u_{i+1} + u_{i-1}
\end{cases}
\]

with the initial conditions

\[
u_i(0) = f_i.
\]

Assume further that the signal is either of infinite length with compact support, or finite with reflecting boundary conditions.

Here, \( \dot{u}_i \) denotes the time derivative of \( u_i(t) \). Like (1), this filter switches between dilation and erosion depending on the local convexity or concavity of the signal. Dilation and erosion themselves are modeled by upwind-type discretisations [11], which are well established for dilation and erosion PDEs because they guarantee the detection of local extrema and stabilise the discretised process by adapting the discrete representation to the local directedness of the PDE evolution.

It should be noted that in case \( 2u_i > u_{i+1} + u_{i-1} \) the two neighbour differences \( u_{i+1} - u_i \) and \( u_{i-1} - u_i \) cannot be simultaneously positive; with the opposite inequality they cannot be simultaneously negative. In fact, always when the maximum or minimum in (2) does not select its third argument, zero, it returns the absolutely smaller of the neighbour differences.

No modification of (2) is needed for finite-length signals with reflecting boundary conditions. In this case, each boundary pixel has one vanishing neighbour difference.

In order to study the solution behaviour of this system, we have to specify the possible solutions, taking into account that the right-hand side of (2) may involve discontinuities. We call a time-dependent signal \( u(t) = (\ldots, u_1(t), u_2(t), u_3(t) \ldots) \) a solution of (2) if

1. each \( u_i \) is a piecewise differentiable function of \( t \),
2. each \( u_i \) satisfies (2) for all times \( t \) for which \( \dot{u}_i(t) \) exists,
3. for \( t = 0 \), the right-sided derivative \( \dot{u}_i^+(0) \) equals the right-hand side of (2) if \( 2u_i(0) \neq u_{i+1}(0) + u_{i-1}(0) \).

We also remark that throughout this paper, extremality of pixels is handled strictly local: A pixel \( u_i \) of a 1-D space-discrete signal \( u \) is called a local extremum whenever \( (u_i - u_{i-1}) (u_{i+1} - u_i) \leq 0 \). For example, in the sequence of five pixels \( u_1 > u_2 = u_3 = u_4 > u_5 \), pixels \( u_2 \) and \( u_3 \) are local minima while \( u_3 \) and \( u_4 \) are local maxima.
2.2 Well-Posedness Results

The following theorem contains our main result.

**Theorem 2.1 (Well-Posedness)** For our Problem, assume that the equality $f_{k+1} - 2f_k + f_{k-1} = 0$ does not hold for any pixel $f_k$ which is not a local maximum or minimum of $f$. Then the following are true:

(i) **Existence and uniqueness:** The problem has a unique solution for all $t \geq 0$.

(ii) **Maximum–minimum principle:** If there are real bounds $a, b$ such that $a < f_k < b$ holds for all $k$, then $a < u_k(t) < b$ holds for all $k$ and all $t \geq 0$.

(iii) **$l_{\infty}$-stability:** There exists a $\delta > 0$ such that for any initial signal $\bar{f}$ with $\|\bar{f} - f\|_{\infty} < \delta$ the corresponding solution $\bar{u}$ satisfies the inequality

$$
\|\bar{u}(t) - u(t)\|_{\infty} < \|\bar{f} - f\|_{\infty}
$$

for all $t > 0$. The solution therefore depends $l_{\infty}$-continuously on the initial conditions within a neighbourhood of $f$.

(iv) **Total variation preservation:** If the total variation of $f$ is finite, then the total variation of $u$ at any time $t \geq 0$ equals that of $f$.

(v) **Steady state:** For $t \to \infty$, the signal $u$ converges to a piecewise constant signal. The jumps in this signal are located at the steepest slope positions of the original signal.

All statements of this theorem follow from an explicit analytical solution of the Problem that will be described in the following proposition.

**Proposition 2.2 (Analytical Solution)** For our Problem, let the segment $(f_1, \ldots, f_m)$ be strictly decreasing and concave in all pixels. Assume that the leading pixel $f_1$ is either a local maximum or a neighbour to a convex pixel $f_0 > f_1$. Then the following hold for all $t \geq 0$:

(i) If $f_1$ is a local maximum of $f$, $u_1(t)$ is a local maximum of $u(t)$.

(ii) If $f_1$ is neighbour to a convex pixel $f_0 > f_1$, then $u_1(t)$ also has a convex neighbour pixel $u_0(t) > u_1(t)$.

(iii) The segment $(u_1, \ldots, u_m)$ remains strictly decreasing and concave in all pixels. The grey values of all pixels at time $t$ are given by

$$
u_k(t) = C \cdot \left( 1 + (-1)^k e^{-2t} - e^{-t} \sum_{j=0}^{k-2} \frac{j!}{j!} (1 + (-1)^{k-j}) \right)
$$

$$
+ e^{-t} \sum_{j=0}^{k-2} \frac{j!}{j!} f_{k-j} - (-1)^k f_1 e^{-t} \left( e^{-t} - \sum_{j=0}^{k-2} \frac{(-t)^j}{j!} \right)
$$

for $k = \ldots, m$, where $C = f_1(0)$ if $f_1$ is a local maximum of $f$, and $C = \frac{1}{2}(f_0(0) + f_1(0))$ otherwise.
(iv) At no time \( t \geq 0 \), the equation \( 2u_i(t) = u_{i+1}(t) + u_{i-1}(t) \) becomes true for any \( i \in \{1, \ldots, m\} \).

Analogous statements hold for increasing concave and for convex signal segments.

**Proof.** We show in steps 1.–3. that the claimed evolution equations hold as long as the initial monotonicity and convexity properties of the signal segment prevail. Step 4. then completes the proof by demonstrating that the evolution equations preserve exactly these monotonicity and convexity requirements.

1. From (2) it is clear that any pixel \( u_i \) which is extremal at time \( t \) has \( \dot{u}_i(t) = 0 \) and therefore does not move. Particularly, if \( f_1 \) is a local maximum of \( f \), then \( u_1(t) \) remains constant as long as it continues to be a maximum.

2. If \( u_0 > u_1, u_0 \) is convex and \( u_1 \) concave for \( t \in [0, T) \). Then we have for these pixels

\[
\begin{align*}
\dot{u}_0 &= u_1 - u_0, \\
\dot{u}_1 &= u_0 - u_1, \\
\end{align*}
\]  

which by the substitutions \( y := \frac{1}{2}(u_0 + u_1) \) and \( v := u_1 - u_0 \) becomes

\[
\begin{align*}
\dot{y} &= 0, \\
\dot{v} &= -2v .
\end{align*}
\]

This system of linear ordinary differential equations (ODEs) has the solution \( y(t) = y(0) = C \) and \( v(t) = v(0) \exp(-2t) \). Backsubstitution gives

\[
\begin{align*}
u_0(t) &= C \cdot (1 - e^{-2t}) + f_0 e^{-2t} , \\
u_1(t) &= C \cdot (1 - e^{-2t}) + f_1 e^{-2t} .
\end{align*}
\]  

This explicit solution is valid as long as the convexity and monotonicity properties of \( u_0 \) and \( u_1 \) do not change.

3. Assume the monotonicity and convexity conditions required by the proposition for the initial signal hold for \( u(t) \) for all \( t \in [0, T) \). Then we have in all cases, defining \( C \) as in the proposition, the system of ODEs

\[
\begin{align*}
\dot{u}_k &= -2(u_1 - C) , \\
\dot{u}_k &= u_{k-1} - u_k , \quad k = 2, \ldots, m
\end{align*}
\]

for \( t \in [0, T) \). We substitute further \( v_k := u_k - C \) for \( k = 1, \ldots, m \) as well as \( w_1 := v_1 \) and \( w_k := v_k + (-1)^k v_1 \) for \( k = 2, \ldots, m \). This leads to the system

\[
\begin{align*}
\dot{w}_1 &= -2w_1 , \\
\dot{w}_2 &= -w_2 , \\
\dot{w}_k &= w_{k-1} - w_k , \quad k = 3, \ldots, m
\end{align*}
\]
This system of linear ODEs has the unique solution
\[ w_1(t) = w_1(0)e^{-2t}, \]
\[ w_k(t) = e^{-t} \sum_{j=0}^{k-2} \frac{t^j}{j!} u_{k-j}(0), \quad k = 2, \ldots, m \]
which after reverse substitution yields (4) for all \( t \in [0, T] \).

4. Note that (5) and (7) are systems of linear ODEs which have the unique explicit solutions
(6) and (4) for all \( t > 0 \). As long as the initial monotonicity and convexity conditions are satisfied, the solutions of (2) coincide with those of the linear ODE systems.

We prove therefore that the solution (4) fulfills the monotonicity condition
\[ u_k(t) - u_{k-1}(t) < 0, \quad k = 2, \ldots, m, \]
and the concavity conditions
\[ u_{k+1}(t) - 2u_k(t) + u_{k-1}(t) < 0, \quad k = 1, \ldots, m, \]
for all \( t > 0 \) if they are valid for \( t = 0 \). To see this, we calculate first

\[
u_k(t) - u_{k-1}(t) = e^{-t} \sum_{j=0}^{k-2} \frac{t^j}{j!} (f_{k-j} - f_{k-1-j})
+ 2e^{-t}(-1)^{k-1} \left( e^{-t} - \sum_{j=0}^{k-2} \frac{(-t)^j}{j!} \right) (f_1 - C). \]

By hypothesis, \( f_{k-j} - f_{k-1-j} \) and \( f_1 - C \) are negative. Further, \( \exp(-t) - \sum_{j=0}^{k-2} (-t)^j / j! \) is the error of the (alternating) Taylor series of \( \exp(-t) \), thus having the same sign \((-1)^{k-1}\) as the first neglected member. Consequently, monotonicity is preserved by (4) for all \( t > 0 \).

Second, we have for \( k = 2, \ldots, m - 1 \)

\[
u_{k+1}(t) - 2u_k(t) + u_{k-1}(t) = e^{-t} \sum_{j=0}^{k-2} \frac{t^j}{j!} (f_{k-j+1} - 2f_{k-j} + f_{k-j-1})
+ 4e^{-t}(-1)^k \left( e^{-t} - \sum_{j=0}^{k-2} \frac{(-t)^j}{j!} \right) (f_1 - C)
+ e^{-t} \frac{t^{k-1}}{(k-1)!} (f_2 - 3f_1 + 2C)
\]

which is seen to be negative by similar reasoning as above.

Concavity at \( u_m(t) \) follows in nearly the same way. By extending (4) to \( k = m + 1 \), one obtains not necessarily the true evolution of \( u_{m+1} \) since that pixel is not assumed to be included in the concave segment. However, the true trajectory of \( u_{m+1} \) can only lie below or on that predicted by (4).

Third, if \( f_1 \) is a maximum of \( f \), then \( u_1(t) \) remains one for all \( t > 0 \) which also ensures concavity at \( u_1 \). If \( f_1 \) has a convex neighbour pixel \( f_0 > f_1 \), we have instead

\[ u_2(t) = 2u_1(t) + u_0(t) = e^{-t} (f_2 - 2f_1 + f_0) + 4e^{-t} (1 - e^{-t})(f_1 - C) < 0 \]
which is again negative for all $t > 0$.

Finally, we remark that the solution (6) ensures $u_0(t) > u_1(t)$ for all $t > 0$ if it holds for $t = 0$. That convexity at $u_0$ is preserved can be established by analogous reasoning as for the concavity at $u_1$.

Since the solutions from the linear systems guarantee preservation of all monotonicity and convexity properties which initially hold for the considered segment, these solutions are the solutions of (2) for all $t > 0$.

**Proof of Theorem 2.1.** Existence and uniqueness of the solution follow from the proof of Proposition 2.2.

The maximum–minimum principle and preservation of total variation follow from the fact that extrema do not move, and monotonicity is preserved for all $t > 0$. Note that by the specification of our Problem, each non-extremal pixel in the signal belongs to a segment enclosed by two extrema.

For the $l_{\infty}$-stability, note that for each admissible initial signal $f$, there exists a lower bound $\gamma > 0$ for all values of $|f_{k-1} - f_k|$ which are not zero, and a lower bound $\eta > 0$ for all values of $|f_{k+1} - 2f_k + f_{k+1}|$ for pixels $k$ which are not local extrema of $f$. Let a signal $\tilde{f}$ with $\|f - \tilde{f}\|_{\infty} =: d < \min\{\gamma/2, \eta/4\}$ be given. One easily checks that then the monotonicity and convexity properties of all strictly monotone segments of $f$ are preserved in $\tilde{f}$. Moreover, isolated extremal pixels of $f$ will be such in $\tilde{f}$. Only chains of equal pixels $f_k = \ldots = f_{k+l}$ may break up in $\tilde{f}$, but in this case the corresponding chain $\tilde{f}_k, \ldots, \tilde{f}_{k+l}$ contains at least one extremum or one inflection pair. To sum up, each monotone and convex/concave segment in $f$ with one of the pixels $f_k = \ldots = f_{k+l}$ as leading pixel is turned into a segment of equal character in $\tilde{f}$ whose leading pixel is one of $\tilde{f}_k, \ldots, \tilde{f}_{k+l}$.

Let us choose without loss of generality a pixel $\tilde{f}_k$ within a decreasing and concave segment as in the proposition. By virtue of $|C - \tilde{C}| \leq d$ and $|\tilde{f}_j - f_j| \leq d$, we can estimate the difference of the explicit solutions (4) for $\tilde{u}_k$ and $u_k$:

$$|\tilde{u}_k(t) - u_k(t)| \leq d \cdot \left(1 - e^{-t \sum_{j=0}^{k-2} \frac{\tilde{U}_j}{j!}} + (-1)^k e^{-t} \left(e^{-t} - e^{-t \sum_{j=0}^{k-2} \frac{(-t)_j}{j!}}\right)\right)$$

$$+ e^{-t \sum_{j=0}^{k-2} \frac{\tilde{U}_j}{j!}} + e^{-t} \cdot \left|e^{-t - \sum_{j=0}^{k-2} \frac{(-t)_j}{j!}}\right|$$

$$\leq d$$

which proves the $l_{\infty}$-stability statement with $\delta := \min\{\gamma/2, \eta/4\}$.

Finally, the convergence to a steady state is obvious from the exponential decay of all summands but $C$ in (4).}

The fact that the signal reaches its steady state only for $t \to \infty$ stands in contrast to the behaviour of space-continuous dilation and erosion where extrema propagate in space with constant speed. In our semidiscrete setting, non-extremal pixels only asymptotically approach their limit values. This can be seen as a blurring, or approximation error, which is the price for the performed spatial discretisation.
2.3 Signals With Linear Pixels

Our well-posedness statements in the previous section explicitly exclude signals which contain pixels \( f_k \) with \( f_{k+1} - 2f_k + f_{k-1} = 0 \). For brevity, we shall call such pixels non-extremal linear pixels.

As we are going to demonstrate, no uniqueness and continuous dependence on initial condition holds for initial signals \( f \) with non-extremal linear pixels. First of all, condition (III) allows the right-sided time derivative of such a pixel at \( t = 0 \) to deviate from the right-hand side of equation (2). Without this relaxation, the condition would in many cases stand in contradiction to the uniquely determined evolution of non-linear surrounding pixels, preventing the existence of solutions.

While the evolution of the neighbours of a non-extremal linear pixel will often force it to become convex or concave, this process is never unique: Instead, one can label each non-extremal linear pixel at \( t = 0 \) arbitrarily to impose either convexity or concavity. Whatever labelling is chosen, it leads to a consistent evolution for \( t > 0 \) in which no linear pixels reappear. This is precised in the following proposition.

**Proposition 2.3 (Forking Solutions at Linear Pixels)** Let an initial signal \( f = (\ldots, f_1, f_2, f_3, \ldots) \) be given. Let for each non-extremal linear pixel \( f_k \) a sign \( \sigma_k \in \{+1, -1\} \) be chosen. Then there exists a unique solution of (2) in the sense of (I)-(III) with initial conditions (3) such that for each non-extremal linear pixel \( f_k \), the right-sided derivative of \( u_k \) at \( t = 0 \) is

\[
\dot{u}_k^+(0) = \begin{cases} 
\max(u_{i+1} - u_i, u_i - u_{i-1}, 0), & \sigma_k = +1, \\
\min(u_{i+1} - u_i, u_i - u_{i-1}, 0), & \sigma_k = -1.
\end{cases}
\]

For none of these solutions, non-extremal linear pixels exist in \( u(t) \) for \( t > 0 \).

**Proof.** Assume we are given a fixed choice of the \( \sigma_k \). Interpreting non-extremal linear pixels with \( \sigma_k = +1 \) as convex and such with \( \sigma_k = -1 \) as concave, we have again a segmentation of the entire signal into concave and convex regions as for signals without non-extremal linear pixels. For each of these regions, we proceed as in the proof of Proposition 2.2 by rewriting the evolution into a system of linear ODEs which has verbatim the same analytical solutions as before. Inspecting the proof of Proposition 2.2, one finds first that the proof of monotonicity preservation suffers no change at all. In the concavity preservation proof one finds that some of the summands of type \( (f_{j-1} - 2f_j + f_{j+1}) \) in the inequalities used there for \( u_{k+1}(t) - 2u_k(t) + u_{k+1}(t) \) now vanish. However, for the entire right-hand side to vanish it is necessary that \( (f_1 - C) \) and all \( (f_{j-1} - 2f_j + f_{j+1}) \) for \( j = 1, \ldots, k \) are zero. This is true if and only if \( f_k \) itself is extremal. Consequently, under the evolution of the linear system, non-extremal pixels which are linear at \( t = 0 \) become convex if \( \sigma_k = +1 \) or concave if \( \sigma_k = -1 \) for any positive \( t \). □
3 Time Discretisation

3.1 Explicit Time Discretisation

In the following we discuss time discretisations of our time-continuous system. We denote the time step by $\tau > 0$. A straightforward explicit time discretisation of our Problem then reads as follows:

**Time-Discrete Problem.** Let $(\ldots,u_{0}^{l},u_{1}^{l},u_{2}^{l},\ldots), \ l = 0,1,2,\ldots$ be a series of bounded real-valued signals which satisfy the equations

$$
\frac{u_{i}^{l+1} - u_{i}^{l}}{\tau} = \begin{cases} 
\max(u_{i+1}^{l} - u_{i}^{l}, u_{i-1}^{l} - u_{i}^{l},0), & 2u_{i}^{l} > u_{i+1}^{l} + u_{i-1}^{l}, \\
\min(u_{i+1}^{l} - u_{i}^{l}, u_{i-1}^{l} - u_{i}^{l},0), & 2u_{i}^{l} < u_{i+1}^{l} + u_{i-1}^{l}, \\
0, & 2u_{i}^{l} = u_{i+1}^{l} + u_{i-1}^{l}
\end{cases}
$$

with the initial conditions

$$
u_{i}^{0} = f_{i}.
$$

Assume further that the signal is either of infinite length or finite with reflecting boundary conditions.

**Theorem 3.1 (Time-Discrete Well-Posedness)** Assume that in the Time-Discrete Problem the equality $f_{k+1} - 2f_{k} + f_{k-1} = 0$ does not hold for any pixel $f_{k}$ which is not a local maximum or minimum of $f$. Assume further that $\tau < 1/2$. Then the statements of Theorem 2.1 are valid for the solution of the Time-Discrete Problem if only $u_{k}(t)$ for $t > 0$ is replaced everywhere by $u_{k}^{l}$ with $l = 0,1,2,\ldots$

The existence and uniqueness of the solution of the Time-Discrete Problem for $l = 0,1,2,\ldots$ is obvious. Maximum-minimum principle, $l_{\infty}$-stability, total variation preservation and the steady state property are immediate consequences of the following proposition. It states that for $\tau < 1/2$ all qualitative properties of the time-continuous solution transfer to the time-discrete case.

**Proposition 3.2 (Time-Discrete Solution)** Let $u_{i}^{l}$ be the value of pixel $i$ in time step $l$ of the solution of our Time-Discrete Problem with time step size $\tau < 1/2$. Then the following hold for all $l = 0,1,2,\ldots$:

(i) If $u_{i}^{l}$ is a local maximum of $u^{l}$, then $u_{i}^{l+1}$ is a local maximum of $u^{l+1}$.

(ii) If $u_{i}^{l}$ is a concave pixel neighbouring to a convex pixel $u_{0}^{l} > u_{i}^{l}$, then $u_{i}^{l+1}$ is again concave and has a convex neighbour pixel $u_{0}^{l+1} > u_{i}^{l+1}$.

(iii) If the segment $(u_{1}^{l}, \ldots,u_{m}^{l})$ is strictly decreasing and concave in all pixels, and $u_{i}^{l}$ is either a local maximum of $u^{l}$ or neighbours to a convex pixel $u_{0}^{l} > u_{i}^{l}$, then the segment $(u_{i}^{l+1}, \ldots,u_{m}^{l+1})$ is strictly decreasing.

(iv) Under the same assumptions as in (iii), the segment $(u_{1}^{l+1}, \ldots,u_{m}^{l+1})$ is strictly concave in all pixels.

(v) If $2u_{i}^{l} = u_{i+1}^{l} + u_{i-1}^{l}$ holds for no pixel $i$, then $2u_{i}^{l+1} = u_{i+1}^{l+1} + u_{i-1}^{l+1}$ also holds for no pixel $i$.
(vi) Under the assumptions of (iii), all pixels in the range $i \in \{1, \ldots, m\}$ have the same limit
\[ \lim_{t \to \infty} u_i^l = C \] with $C := u_i^1$ if $u_i^1$ is a local maximum, or $C := \frac{1}{2}(u_0^1 + u_1^1)$ if it neighbours
the convex pixel $u_i^1$.

Analogous statements hold for local minima, for increasing concave and for convex signal segments.

**Proof.** Assume first that $u_i^l$ is a local maximum of $u^l$. From the evolution equation (9) it
is clear that $u_{j+1}^{l+1} \leq u_j^l + \tau(u_j^l - u_{j-1}^l)$ for $j = 0, 2$. For $\tau < 1/2$ this entails $u_{j+1}^{l+1} \leq u_j^l$, thus (i).

If instead $u_i^l$ is a concave neighbour of a convex pixel $u_i^0 > u_i^1$, then we have $u_i^{l+1} = u_i^1 = u_i^0 + \tau(u_i^0 - u_i^1)$ and $u_{i+1}^{l+1} = \tau(u_i^1 - u_{i+1}^1)$. Obviously, $u_i^{l+1} > u_i^0$ holds if and only if
$\tau < 1/2$. For concavity, note that $u_i^{l+1} \leq u_0^1 + \tau(u_0^1 - u_i^1)$ and therefore $u_i^{l+1} - 2u_i^{l+1} + u_{i+1}^{l+1} \leq (1 - \tau)(u_0^1 - 2u_i^1 + u_{i+1}^1) + 2\tau(u_i^1 - u_0^1)$. The right-hand side is certainly negative for $\tau \leq 1/2$.

An analogous argument secures convexity at pixel 0 which completes the proof of (ii).

In both cases we have $u_i^{l+1} \geq u_i^l$. Under the assumptions of (iii), (iv) we then have $u_i^{l+1} = u_i^k + \tau(u_k^1 - u_i^l)$ for $k = 2, \ldots, m$. If $\tau < 1$, it follows that $u_i^l < u_i^{l+1} \leq u_i^{l+1}$ for $k = 2, \ldots, m$ which together with $u_i^{l+1} \geq u_i^l$ implies that $u_i^{l+1} > u_i^k$ for $k = 2, \ldots, m$ and thereore (iii).

For the concavity condition we compute
\[ u_{k+1}^{l+1} - 2u_{k}^{l+1} + u_{k-1}^{l+1} = (1 - \tau)(u_{k-1}^{l+1} - 2u_k^l + u_{k+1}^l) \]
for $k = 3, \ldots, m - 1$. The right-hand side is certainly negative for $\tau \leq 1$ which secures concavity in the pixels $k = 3, \ldots, m - 1$. Concavity in pixel $m$ for $\tau \leq 1$ follows from essentially the same argument. However, the equation is now replaced by an inequality since for pixel $m + 1$ we know only that $u_m^{l+1} \leq u_m^l + \tau(u_m^l - u_m^{l+1})$. If $u_i^1$ is a local maximum and therefore $u_i^{l+1} = u_i^1$, we find for pixel 2 that $u_i^{l+1} - 2u_i^{l+1} + u_{i+1}^{l+1} = (1 - \tau)(u_i^1 - 2u_2^l + u_3^l) + \tau(u_2^l - u_i^1)$ which again secures concavity for $\tau \leq 1$. As was proven above, concavity in pixel 1 is preserved for $\tau \leq 1/2$ such that (iv) is proven.

Under the hypothesis of (v), the evolution of all pixels in the signal is described by statements (i)-(iv) or their obvious analogs for increasing and convex segments. The claim of (v) then is obvious.

Finally, addition of the equalities $C - u_i^{l+1} = (1 - 2\tau)(C - u_i^l)$ and $u_i^{l+1} - u_i^{l+1} = (1 - \tau)(u_{i-1}^l - u_i^l)$ for $i = 2, \ldots, m$ implies that
\[ C - u_k^{l+1} = (1 - \tau)(C - u_k^l) < (1 - \tau)(C - u_k^l) \]
for all $k = 1, \ldots, m$. By induction, we have
\[ C - u_k^{l+1} \leq (1 - \tau)^l(C - u_k^l) \]
where the right-hand side tends to zero for $l' \to \infty$. Together with the monotonicity preservation for $\tau < 1/2$, statement (vi) follows. \hfill \square

We remark that in the presence of non-extremal linear pixels, uniqueness fails in a similar way as in the semidiscrete setting.
3.2 Modified Explicit Time Discretisation

A closer look at the proof of Proposition 3.2 reveals that the limitation \( \tau < 1/2 \) is made necessary only by the situation of case (ii) of the proposition, i.e. a concave pixel following a convex one within a decreasing segment, or symmetrical situations. In the absence of such a configuration, the statements hold even for all \( \tau < 1 \).

By a small adaptation of the time-discrete evolution rule we can therefore obtain a scheme which satisfies well-posedness properties for step size up to 1.

**Modified Time-Discrete Problem.** Let \((\ldots, u_0^l, u_1^l, u_2^l, \ldots), l = 0, 1, 2, \ldots\) be a series of bounded real-valued signals which is generated by the equations

\[
\frac{u_i^l - u_i^l}{\tau} = \begin{cases} 
\max(u_{i+1}^l - u_i^l, u_{i-1}^l - u_i^l, 0), & 2u_i^l > u_{i+1}^l + u_{i-1}^l, \\
\min(u_{i+1}^l - u_i^l, u_{i-1}^l - u_i^l, 0), & 2u_i^l < u_{i+1}^l + u_{i-1}^l, \\
0, & 2u_i^l = u_{i+1}^l + u_{i-1}^l,
\end{cases}
\]

(11)

\[
u_i^{l+1} = \begin{cases} 
\frac{1}{2}(u_{i+1}^l + u_i^l), & (u_{i+1}^l - u_i^l)(u_{i+1}^l - u_i^l) < 0, \\
\frac{1}{2}(u_{i-1}^l + u_i^l), & (u_{i-1}^l - u_i^l)(u_{i-1}^l - u_i^l) < 0, \\
u_i^l, & \text{else}
\end{cases}
\]

(12)

with the initial conditions (10) and boundary conditions as in the previous Time-Discrete Problem.

Note that the case distinction on the right-hand side of (12) is sound since the case (ii) of Proposition 3.2 cannot occur simultaneously on both sides of the same pixel. Further, it is clear from Proposition 3.2 that the Modified Time-Discrete Problem is identical with the Time-Discrete Problem for \( \tau < 1/2 \).

For the Modified Time-Discrete Problem, the well-posedness statements of Theorem 3.1 hold for all \( \tau \leq 1 \). Since the proof contains only slight modifications compared to the previous one, we do not repeat it here but state only the suitably modified version of Proposition 3.2. The main modification is that extremal linear pixels can now arise during the evolution.

**Proposition 3.3 (Modified Time-Discrete Solution)** Let \( u_i^l \) be the value of pixel \( i \) in time step \( l \) of the solution of our Modified Time-Discrete Problem with time step size \( \tau \leq 1 \). Then the following hold for all \( l = 0, 1, 2, \ldots \):

(i) If \( u_i^l \) is a local maximum of \( u_i^l \), then \( u_i^{l+1} \) is a local maximum of \( u_i^{l+1} \).

(ii) If \( u_i^l \) is a concave pixel neighbouring to a convex pixel \( u_0^l > u_i^l \), then \( u_i^{l+1} \) is again concave and has a convex neighbour pixel \( u_{i+1}^{l+1} \geq u_i^{l+1} \).

(iii) If the segment \((u_1^l, \ldots, u_m^l)\) is strictly decreasing and concave in all pixels, and \( u_i^l \) is either a local maximum of \( u_i^l \) or neighbours to a convex pixel \( u_0^l > u_i^l \), then the segment \((u_2^{l+1}, \ldots, u_m^{l+1})\) is strictly decreasing, and \( u_i^{l+1} \geq u_2^{l+1} \).

(iv) Under the same assumptions as in (iii), the segment \((u_2^{l+1}, \ldots, u_m^{l+1})\) is strictly concave in all pixels. Pixel \( u_i^{l+1} \) is strictly concave except if \( u_i^l \) is a local maximum, and \( \tau = 1 \).

(v) If \( 2u_i^l = u_i^{l+1} + u_i^{l-1} \) holds for no pixel \( i \) for which \( u_i^{l-1} = u_i^l = u_i^{l+1} \) does not hold, then \( 2u_i^{l+1} = u_i^{l+1} + u_i^{l-1} \) also holds for no pixel \( i \) for which \( u_i^{l-1} = u_i^{l+1} = u_i^{l+1} \) does not hold.
(vi) Under the assumptions of (iii), all pixels in the range $i \in \{1, \ldots, m\}$ have the same limit
\[ \lim_{t \to \infty} u_i^t = C \] with $C := u_i^1$ if $u_i^1$ is a local maximum, or $C := \frac{1}{2}(u_0^1 + u_i^1)$ if it neighbours
to the convex pixel $u_0^1$.

Analogous statements hold for local minima, for increasing concave and for convex signal segments.

The special case $\tau = 1$ deserves a closer consideration. Straightforward calculations reveal
that a pixel neighbouring to a local maximum attains the same value as the maximum in the
next time step. Moreover, a pair of a convex pixel followed by a concave one in a decreasing
segment aligns to equal values within one time step, turning both pixels into discrete local
extrema. These facts give rise to the following corollary.

**Corollary 3.4** Consider the Modified Time-Discrete Problem with $\tau = 1$. If at time step
$k = 1$ the segment $(u_1^1, u_2^1, \ldots, u_m^1) \) has the properties required in Prop. 3.3, (iii), then each
pixel $u_k$ of this segment becomes constant after not more than $k$ time steps.

In the case of finite-length signals, or infinite signals in which there exists a finite upper
bound to the length of monotonic concave or convex segments, this corollary implies that the
steady state is reached in finite time.

Since the modified time-discrete filter propagates grey-values in $x$ direction one pixel per
time step, it turns out to reflect particularly well the behaviour of dilations and erosions on a
continuous domain. In the light of our remark at the end of Subsection 2.2, it can be said that
the approximation error introduced by spatial discretisation has been compensated exactly
by the time discretisation.

A further view on the modified scheme is that it can be related to *locally analytic schemes*. Schemes of this type compose analytic solutions for certain dynamical systems in a neighbour-
hood of each location into approximate solutions on the entire grid. Examples are the
numerical scheme for 1-D total variation (TV) flow introduced in [15, Sec. 4.2] which relies on
analytic solutions for TV flow on two pixels, or its 2-D extension set forth in [21] which is
based on four-pixel systems.

Our modified time-discrete shock filtering scheme can be understood as a locally analytic
scheme built on the analytical solutions (for $0 \leq \tau \leq 1$) of space-continuous dilation and
erosion filters where the signal is linearly interpolated between subsequent pixels.

### 4 Equivalence to Local Mode Filtering

Now that we have derived well-posedness properties for semidiscrete and fully discrete shock
filters, let us also establish an equivalence result between shock filters and a class of discrete
filters based on local signal statistics, namely so-called mode filters.

Discrete filters exploiting local signal statistics evaluate signal values from a sliding window
neighbourhood to determine a new value for each pixel. Commonly used representatives of
this class are box-average filters, median filters and generally M-smoothers, but also discrete
dilation and erosion.

One statistical parameter of the local greyvalue distribution that is not used in one of the
aforementioned filters is the *mode*. The mode of a continuous distribution is defined as its
most probable value. Analogous as above, determining the mode of the grey-values within a sliding window or structuring element constitutes a local mode filter for images [3, 16, 5].

However, applying this procedure to spatially discretised signals faces a problem because the distribution is now given only by finitely many values. Defining the mode simply as the most frequent value is not helpful since in generic cases there are no duplicates among the values.

Instead, we combine a polynomial approximation with local histogram properties to find the mode value within a sliding window containing three pixels.

Assume we have a one-dimensional discrete signal \( \ldots, u_0, u_1, u_2, \ldots \). By our sliding window we select three subsequent values \( u_{i-1}^l, u_i^l, u_{i+1}^l \). The value at pixel \( i \) of the signal filtered by our local mode filter should be the mode value of \( u_{i-1}^l, u_i^l, u_{i+1}^l \).

To begin with, we interpolate by a quadratic polynomial through the three points \((i - 1, u_{i-1}^l), (i, u_i^l), (i + 1, u_{i+1}^l)\). Translating, for simplicity, spatial coordinates by \(-i\), we therefore want the quadratic polynomial

\[
p(z) = az^2 + bz + c
\]
to satisfy the conditions

\[
p(-1) = u_{i-1}^l, \quad p(0) = u_i^l, \quad p(1) = u_{i+1}^l.
\]

This gives the system of three equations

\[
\begin{align*}
a - b + c &= u_{i-1}^l, \\
c &= u_i^l, \\
a + b + c &= u_{i+1}^l,
\end{align*}
\]

with the solution

\[
a = \frac{u_{i-1}^l - 2u_i^l + u_{i+1}^l}{2}, \quad b = \frac{u_{i+1}^l - u_{i-1}^l}{2}, \quad c = u_i^l.
\]

Having determined \( p(z) \), we are now interested in the location of the mode of its values. First, if \( a = 0 \), \( p \) is a linear polynomial whose values are uniformly distributed. In this case, we have our local mode filter not change the value of pixel \( i \).

If \( a \neq 0 \), the density of the distribution of values of \( p \) attains its maximum at the (uniquely determined) stationary value of \( p \). The extremum of \( p(z) \) is located at

\[
- \frac{b}{2a} = \frac{u_{i-1}^l - u_{i+1}^l}{u_{i-1}^l - 2u_i^l + u_{i+1}^l} =: \varepsilon.
\]

However, whenever \( \varepsilon \not\in \{-1, 0, +1\} \), the extremal value \( p(\varepsilon) \) will lie outside the interval \( I_i := [\min(u_{i-1}^l, u_i^l, u_{i+1}^l), \max(u_{i-1}^l, u_i^l, u_{i+1}^l)] \); choosing \( p(\varepsilon) \) as the value of the mode filter therefore results in over- and undershoots.

This leads us to stabilise our filtering procedure by projecting \( p(\varepsilon) \) to the interval \( I_i \), i.e. choosing as the new value of pixel \( i \) the value from the interval \( I_i \) which is closest to \( p(\varepsilon) \). We will call this procedure stabilised discrete local mode filtering. Clearly, for \( p \) not linear this filter will always return as its value one of the end points of \( I_i \), namely \( \min(u_{i-1}^l, u_i^l, u_{i+1}^l) \) if \( a < 0 \), or \( \max(u_{i-1}^l, u_i^l, u_{i+1}^l) \) if \( a > 0 \).

We have therefore arrived at the following equivalence result.
Proposition 4.1 Stabilised discrete local mode filtering of a 1-D discrete signal \((\ldots, f_0, f_1, f_2, \ldots)\) obeys the equation

\[
u_i = \begin{cases} 
\min(u_{i-1}^l, u_i^l, u_{i+1}^l), & \text{if } \max(u_{i-1}^l, u_i^l, u_{i+1}^l) < 0, \\
\max(u_{i-1}^l, u_i^l, u_{i+1}^l), & \text{if } \max(u_{i-1}^l, u_i^l, u_{i+1}^l) > 0, \\
u_i^l, & \text{if } \max(u_{i-1}^l, u_i^l, u_{i+1}^l) = 0.
\end{cases}
\]

Consequently, it is equivalent to one step of time-discrete shock filtering as described in Subsection 3.1 with time step size \(\tau = 1\).

Since the time step size \(\tau = 1\) exceeds the limit given in Theorem 3.1, the well-posedness properties do not transfer fully to stabilised discrete local mode filtering. However, a modification along the line of Section 3.2 could also be applied to local mode filtering to obtain a well-posed process.

5 Conclusions

Theoretical foundation for PDE-based shock filtering has long been considered to be a hopelessly difficult problem. In this paper we have shown that it is possible to obtain both an analytical solution and well-posedness by considering the space-discrete case where the partial differential equation becomes a dynamical system of ordinary differential equations (ODEs). Corresponding results can also be established in the fully discrete case when an explicit time discretisation is applied to this ODE system. Last but not least, we were able to derive equivalence results between shock filtering and local mode filtering.

Since local mode filtering also satisfies equivalence properties to robust estimation and mean shift analysis [17], it becomes clear that shock filtering is much more than a simple image enhancement method: It is a theoretically well-founded class of methods that can be justified in many different ways.

We are convinced that the basic idea in our paper, namely to establish well-posedness results for difficult PDEs in image analysis by considering the semidiscrete case, is also useful in a number of other important PDEs. While this has already been demonstrated for nonlinear diffusion filtering [18, 12], we plan to investigate a number of other PDEs in this manner, both in the one- and the higher-dimensional case. This should give important theoretical insights into the dynamics of these experimentally well-performing nonlinear processes.

References


