Revisiting the Relationship Between Adaptive Smoothing and Anisotropic Diffusion With Modified Filters

Bumsub Ham, Student Member, IEEE, Dongbo Min, Member, IEEE, and Kwanghoon Sohn, Member, IEEE

Abstract—Anisotropic diffusion has been known to be closely related to adaptive smoothing and discretized in a similar manner. This paper revisits a fundamental relationship between two approaches. It is shown that adaptive smoothing and anisotropic diffusion have different theoretical backgrounds by exploring their characteristics with the perspective of normalization, evolution step size, and energy flow. Based on this principle, adaptive smoothing is derived from a second order partial differential equation (PDE), not a conventional anisotropic diffusion, via the coupling of Fick’s law with a generalized continuity equation where a “source” or “sink” exists, which has not been extensively exploited. We show that the source or sink is closely related to the asymmetry of energy flow as well as the normalization term of adaptive smoothing. It enables us to analyze behaviors of adaptive smoothing, such as the maximum principle and stability with a perspective of a PDE. Ultimately, this relationship provides new insights into application-specific filtering algorithm design. By modeling the source or sink in the PDE, we introduce two specific diffusion filters, the robust anisotropic diffusion and the robust coherence enhancing diffusion, as novel instantiations which are more robust against the outliers than the conventional filters.

Index Terms—Anisotropic diffusion, adaptive smoothing, coherence enhancing diffusion, energy flow, generalized continuity equation, normalization.

I. INTRODUCTION

In low-level vision problems, there is a need to smooth images, while preserving universal features such as edges or boundaries, in order to find structures embedded in images [1]. Linear smoothing averages all pixels evenly without incorporating the local topology, leading to blurred features. Over the last two decades, there have been many advances in nonlinear smoothing in which prior knowledge is leveraged for grouping with similar pixels only. There having been many types of nonlinear smoothing [2], [3], partial differential equation (PDE) based smoothing and kernel based smoothing have been widely used. The typical examples of the PDE based smoothing are anisotropic diffusion [4] and the total variation diffusion [5], [6], which are also related to the wavelet shrinkage and morphology [6], [7]. The most representative examples of the kernel based smoothing are adaptive smoothing [8], [9], the bilateral filter [10], the mean-shift filter [11], and the non-local filter [12]. Nonlinear smoothing has been successfully applied to the image denoising [13], [14], segmentation [15], structure decomposition [1], optical flow estimation [16], and manifold smoothing [17].

Many researchers have made efforts to investigate a fundamental relationship between the PDE based smoothing and the kernel based smoothing [8], [12], [13], [17]–[23]. Saint-Marc et al. showed that adaptive smoothing is equivalent to anisotropic diffusion [8]. Barash also derived a relationship between adaptive smoothing and anisotropic diffusion, and showed that Saint-Marc’s results are not consistent [19]. He also verified that the bilateral filter [10] becomes a generalized formulation of adaptive smoothing by introducing 5-D pixels. Simoncelli and Hany generalized a steerable filter, as a type of adaptive filter, such that the orientation and magnitude of local structures can be captured and analyzed together [24]. Buades et al. showed an asymptotic behavior of neighborhood filters as the size of the neighborhood shrinks to zero, and proved that these filters are asymptotically equivalent to anisotropic diffusion [13]. Singer et al. viewed the non-local filter as a diffusion process, and analyzed a relationship between the non-local filter and the random walk theory [20]. Elad showed how the bilateral filter is improved and extended upon for handling more sophisticated reconstruction problems [21]. In [22], it was shown that the bilateral filter is the particular case of the mean-shift filter and can be obtained by fixing the spatial kernel of the mean-shift filter at each iteration. Motivated by these works, Paris and Durand casted the bilateral filter into a signal processing framework [25]. The intensity range was quantized and sampled into a small set of channels, which is similar to the channel smoothing [18], so that the computational efficiency could be dramatically improved. Recently, Sevilla-Lara and Learned-Miller extended the channels from an intensity range to an arbitrary feature space, enabling the channel smoothing to be applicable to high-level vision fields such as tracking [26].

In this paper, a traditional relationship between adaptive smoothing and anisotropic diffusion is revisited. Reinterpreting two approaches in terms of a normalization, an evolution step size and an energy flow, we show that adaptive smoothing...
is equivalent to anisotropic diffusion only when special constraints are imposed. Specifically, the energy flow of adaptive smoothing is asymmetric, whereas that of anisotropic diffusion is always symmetric. Considering an asymmetric energy flow, we derive adaptive smoothing from a second order PDE, not a conventional anisotropic diffusion, via the coupling of Fick’s law with a generalized continuity equation where a source or sink exists, which has not been extensively exploited. Namely, the equivalence between adaptive smoothing and the second order PDE with the source or sink is explicitly investigated. Based on this fact, it is shown that the normalization term used in adaptive smoothing, a fundamental form of the weighted average filter [19], comes from the source or sink in the generalized continuity equation. We also show that adaptive smoothing satisfies a maximum principle and is always stable with a perspective of a PDE. Furthermore, the proposed PDE gives us new diffusion filters such as the robust anisotropic diffusion (RAD) and the robust coherence enhancing diffusion (RCED).

The significance of our work is as follows: First, we distinguish adaptive smoothing from anisotropic diffusion with the perspective of an energy flow, providing new insights into application-specific filtering algorithm design. A symmetric energy flow of anisotropic diffusion implies that the diffusion process conserves the total energy of an initial image. Thus, anisotropic diffusion should be differentiated from adaptive smoothing although they show similar behavior. For instance, Gilboa and Osher [14] proposed a non-local diffusion filter (PDE based smoothing) which is a corresponding counterpart of the non-local filter (kernel based smoothing) [12]. They showed that the proposed diffusion filter is superior to the conventional non-local filter in some applications such as image denoising and supervised image segmentation, since the symmetric energy flow does not tend to blur rare and singular regions [14]. Recently, Aubry et al. proposed a variant of the bilateral filter in which the normalization is removed [27]. This unnormalized version has a weaker effect when the sum of weights become smaller, which leads to generating slightly softer images, thus preventing halos at strong edges. Second, the behavior of a weighted average filter can be analyzed with the viewpoint of a PDE, since the normalization term used in the weighted average filter comes from the source or sink in the generalized continuity equation. Third, a new filter can be designed by properly modeling the source or sink in the proposed PDE according to specific applications. One feasible example is the RAD which is more robust against various outliers such as salt-and-pepper noise, Gaussian noise, and their mixture [28]. In this paper, as an extension of the RAD, the RCED is examined as well.

The paper is organized as follows: Section II briefly summarizes adaptive smoothing and anisotropic diffusion followed by traditional relationship between them [8], [19], [22]. Then, adaptive smoothing is derived from a second order PDE and its behavior is analyzed with the view point of a PDE in Section III. In Section IV, the RAD and the RCED are introduced from the proposed PDE. Finally, Section V concludes the paper with a discussion.

II. ADAPTIVE SMOOTHING AND ANISOTROPIC DIFFUSION

A. Adaptive Smoothing

The adaptive smoothing aims to regularize an image while preserving features. The image is repeatedly convolved with a kernel weighted by a measure of the discontinuity [8]. Let \( I^{(t)}(p) \) denote an intensity value of \( p = (x, y) \) at the \( t^{th} \) iteration. A signal, filtered by adaptive smoothing, is defined as follows:

\[
I^{(t+1)}(p) = \frac{1}{\chi^{(t)}(p)} \sum_{q \in \mathcal{N}} I^{(t)}(q) g_s(d^{(t)}(q, p))
\]

with

\[
\chi^{(t)}(p) = \sum_{q \in \mathcal{N}} g_s(d^{(t)}(q, p))
\]

where \( g_s(d^{(t)}(q, p)) \) is a monotonically decreasing function according to the distance\( d^{(t)}(q, p) = |I^{(t)}(q) - I^{(t)}(p)| \) which discriminates the relative importance between points. \( \mathcal{N} \) is the set of neighboring pixels to the center node \( p \) as shown in Fig. 1(a). Note that the center node is also included in \( \mathcal{N} \).

B. Anisotropic Diffusion

The heat equation, or the diffusion, is a fundamental PDE that models the distribution of heat or temperature on a given domain over time. Perona and Malik applied this physics model to image processing, especially for edge preserving smoothing, with scale space theory [4]. They introduced a time and spatially varying diffusivity function into the diffusion model, which results in anisotropic diffusion, as follows:

\[
\partial_t I(p) = \nabla \cdot [c^{(t)}(p) \nabla I^{(t)}(p)]
\]

where \( t \) denotes the time. \( \nabla \) and \( \nabla \cdot \) denote the gradient and divergence operator, respectively. \( c^{(t)}(p) \) defined as in (4) is a thermal diffusivity function satisfying \( g_d(x) \to 0 \) as \( x \to \infty \)

\[
c^{(t)}(p) = g_d(\| \nabla I^{(t)}(p) \|).
\]

The 1-D counterpart of anisotropic diffusion as in (3) is discretized by an explicit finite difference method (FDM) as follows [29]

\[
\partial_t I(x) = \partial_x [c^{(t)}(x) \partial_x I^{(t)}(x)]
\]

\[
\approx \frac{1}{2} \left( c^{(t)}(x-1) + c^{(t)}(x) \right) \left( I^{(t)}(x-1) - I^{(t)}(x) \right)
\]

\[
+ \frac{1}{2} \left( c^{(t)}(x+1) + c^{(t)}(x) \right) \left( I^{(t)}(x+1) - I^{(t)}(x) \right)
\]

\[
(5)
\]
where
\[
\frac{1}{2} (c^{(i)}(x-1) + c^{(i)}(x)) \\
\approx \frac{1}{2} \left[ g_d(I^{(i)}(x-1) - I^{(i)}(x)) + g_d(I^{(i)}(x-1) - I^{(i)}(x)) \right] \\
= g_d(I^{(i)}(x-1) - I^{(i)}(x)),
\]
\[
\frac{1}{2} (c^{(i)}(x) + c^{(i)}(x+1)) \\
\approx \frac{1}{2} \left[ g_d(I^{(i)}(x+1) - I^{(i)}(x)) + g_d(I^{(i)}(x+1) - I^{(i)}(x)) \right] \\
= g_d(I^{(i)}(x+1) - I^{(i)}(x)).
\] (6)

Note that the first and second terms are approximated by the backward and forward differences, respectively [30]. Then, the 1D anisotropic diffusion is discretized by an explicit FDM with a forward Euler approximation as follows:
\[
[I^{(i+1)}(x) - I^{(i)}(x)]/\tau \\
= g_d(I^{(i)}(x+1) - I^{(i)}(x)) [I^{(i)}(x-1) - I^{(i)}(x)] \\
+ g_d(I^{(i)}(x+1) - I^{(i)}(x)) [I^{(i)}(x+1) - I^{(i)}(x)] \tag{7}
\]
where \( \tau \) is an evolution step size.

Similarly, the 2D anisotropic diffusion as in (3) is discretized as follows:
\[
I^{(i+1)}(p) = I^{(i)}(p) + \tau \sum_{q \in \mathcal{N}_4} g_d(I^{(i)}(q) - I^{(i)}(p)) [I^{(i)}(q) - I^{(i)}(p)] \tag{8}
\]
where \( \mathcal{N}_4 \) represents the 4-neighborhood of the center node \( p \), as shown in Fig. 1(b).

C. Traditional Relationship Between Adaptive Smoothing and Anisotropic Diffusion

We review the traditional relationship between adaptive smoothing and anisotropic diffusion. We assume that, without loss of generality, the functions \( g_d(\cdot) \) in (1) and \( g_d(\cdot) \) in (4) is identical in that they play the same role, i.e., preventing the diffusion across different features. From here on, we hence denote these functions as \( g(\cdot) \). The general relationship between two functions \( c^{(i)}(\cdot) \) and \( d^{(i)}(\cdot) \) can then be derived as follows:
\[
\frac{1}{2} (c^{(i)}(q) + c^{(i)}(p)) \\
\approx g([I^{(i)}(q) - I^{(i)}(p)]) = g(d^{(i)}(q, p)). \tag{9}
\]

Saint-Marc et al. formulated the 1-D case of adaptive smoothing in (1) as follows [8].
\[
I^{(i+1)}(x) = c^{(i)}(x-1)I^{(i)}(x-1) \\
+ c^{(i)}(x)I^{(i)}(x) + c^{(i)}(x+1)I^{(i)}(x+1) \tag{10}
\]
with
\[
c^{(i)}(x-1) + c^{(i)}(x) + c^{(i)}(x+1) = 1. \tag{11}
\]

After plugging (11) into (10) and rearranging the equation, the following equation can be derived:
\[
I^{(i+1)}(x) - I^{(i)}(x) \\
= c^{(i)}(x-1)[I^{(i)}(x-1) - I^{(i)}(x)] \\
+ c^{(i)}(x)[I^{(i)}(x) - I^{(i)}(x)] \\
+ c^{(i)}(x+1)[I^{(i)}(x+1) - I^{(i)}(x)]. \tag{12}
\]

It is similar to the 1-D discrete implementation of anisotropic diffusion in (7). Later, Barash showed that this is an inconsistent approximation of anisotropic diffusion in (7), since an extra term remains when the terms \( c^{(i)}(x+1) \), \( c^{(i)}(x-1) \) and \( I^{(i)}(x+1) \), \( I^{(i)}(x-1) \) are expanded with respect to \( c^{(i)}(x) \) and \( I^{(i)}(x) \), respectively, by using a Taylor series [19]. (See for more details in appendix of [19].) In order to address the inconsistency problem, Barash re-formulated the 1-D adaptive smoothing of (1) as follows [19]:
\[
I^{(i+1)}(x) = \frac{c^{(i)}(x-1)I^{(i)}(x-1) + c^{(i)}(x)I^{(i)}(x-1) + c^{(i)}(x+1)I^{(i)}(x+1)}{2} \\
+ \frac{c^{(i)}(x)I^{(i)}(x) + c^{(i)}(x+1)I^{(i)}(x+1)}{2} \tag{13}
\]
with
\[
\frac{c^{(i)}(x-1) + c^{(i)}(x) + c^{(i)}(x+1)}{2} = 1. \tag{14}
\]

That is
\[
g([I^{(i)}(x-1) - I^{(i)}(x)]) + g(0) \\
+ g([I^{(i)}(x+1) - I^{(i)}(x)]) = \chi^{(i)}(x) = 1. \tag{15}
\]

After similar manipulation to (12), we can derive the following equation:
\[
I^{(i+1)}(x) - I^{(i)}(x) \\
= \frac{c^{(i)}(x-1) + c^{(i)}(x)}{2} [I^{(i)}(x-1) - I^{(i)}(x)] \\
+ \frac{c^{(i)}(x+1) + c^{(i)}(x)}{2} [I^{(i)}(x+1) - I^{(i)}(x)]. \tag{16}
\]

Obviously, (16) can be referred to as the 1-D discrete implementation of anisotropic diffusion as in (7), [19]. However, this result is validated only when (14) or (15) is satisfied, i.e., the sum of weights is equal to 1, since the normalization used in adaptive smoothing of (1) is not considered in (13), which will be explained in the next section.

III. DERIVATION OF ADAPTIVE SMOOTHING FROM A SECOND ORDER PDE

A. Problem Statement

In this section, we show that adaptive smoothing is not equivalent to anisotropic diffusion by exploring the characteristics of two approaches with the perspective of a normalization, an evolution step size, and an energy flow.

1) Normalization: The metric \( d^{(i)}(q, p) \) in (1) is generally defined by an intensity similarity between two pixels, and meets following conditions:
\[
d^{(i)}(q, p) \geq 0 \text{ (non-negativity)} \tag{17}
\]
\[
d^{(i)}(q, p) = 0 \text{ if and only if } q = p \text{ (identity)} \tag{18}
\]
\[
d^{(i)}(q, p) = d^{(i)}(p, q) \text{ (symmetry)} \tag{19}
\]
\[
d^{(i)}(q, p) \leq d^{(i)}(q, r) + d^{(i)}(r, p) \text{ (triangle inequality)} \tag{20}
\]
Since the weight function \( g(d^{(i)}(q, p)) \) is calculated by the distance metric \( d^{(i)}(q, p) \) whose value is always positive, the sum of weights in (14) or (15) is spatially varying according to the characteristics of the distance metric \( d^{(i)}(q, p) \), not being fixed to 1.

**Proposition 1:** The adaptive smoothing is equivalent to anisotropic diffusion only when the sum of weights \( \chi \) is equal to 1.

**Proof:** Let us consider the following case:

\[
\frac{e^{(i)}(x - 1) + e^{(i)}(x)}{2} + \frac{e^{(i)}(x + 1) + e^{(i)}(x)}{2} = \chi^{(i)}(x) \tag{21}
\]

where \( \chi^{(i)}(x) \) is a normalization factor, and is an arbitrary constant that satisfies \( \chi^{(i)}(x) > 0 \).

After the same manipulation as (16), the following equation is derived:

\[
I^{(i+1)}(x) - I^{(i)}(x) = \frac{e^{(i)}(x - 1) + e^{(i)}(x)}{2} \left[ I^{(i)}(x - 1) - I^{(i)}(x) \right] \\
+ \frac{e^{(i)}(x + 1) + e^{(i)}(x)}{2} \left[ I^{(i)}(x + 1) - I^{(i)}(x) \right] \\
+ (\chi^{(i)}(x) - 1)I^{(i)}(x). \tag{22}
\]

Thus, adaptive smoothing is equivalent to anisotropic diffusion only when the sum of weights \( \chi^{(i)}(x) \) is equal to 1.

2) **Evolution Step Size:** It is assumed that the evolution step size of anisotropic diffusion in (16) is 1, making it unstable. In general, when 1-D anisotropic diffusion is discretized by an explicit FDM, the evolution step size should be smaller than 0.5 (0.25 in 2-D case) in order to ensure its numerical stability [31]. Fig. 2 shows images filtered by (b)–(e) anisotropic diffusion where the evolution step size varies from 0.10 to 3.0, and (f) adaptive smoothing, respectively. We found that the result diverges, when the evolution step size of anisotropic diffusion is larger than 0.25, i.e., the filtered results become noisy when the evolution step size is set to 1.5 or 3.0.

Recently, it was shown that each iteration of adaptive smoothing can be referred to as one step of anisotropic diffusion [23]. This relationship, however, still shares the same problems as described above, i.e., anisotropic diffusion derived from adaptive smoothing in [23] is unstable since the evolution step size of anisotropic diffusion is assumed to become 1.

3) **Energy Flow:** The energy of anisotropic diffusion in (8) is exchanged in a symmetric manner, that is, the energy flow between two nodes \( p \) and \( q \) is determined by the center node itself, as shown in Fig. 3(a). In contrast, the energy of adaptive smoothing is exchanged in an asymmetric manner due to the normalization [14], [32], and the energy flow is determined by their neighborhood nodes as well as the current node, as plotted by the dotted arrows in Fig. 3(b). Fig. 4 shows the normalized mean value of the results filtered by anisotropic diffusion and adaptive smoothing according to iteration with an initial condition as in Fig. 2(a). The anisotropic diffusion preserves the mean of an initial image regardless of time, whereas adaptive smoothing does not. However, (16) derived by Barash [19] does not reflect the asymmetric energy flow of adaptive smoothing, i.e., the flow of (16) is always symmetric although it is derived from adaptive smoothing, leading to the conclusion that the equivalence between adaptive smoothing and anisotropic diffusion is not valid.

**B. Derivation of Adaptive Smoothing from a Second Order PDE**

In this section, we first examine the origin of anisotropic diffusion, and re-derive adaptive smoothing from a second order PDE, considering the asymmetric energy flow.

The anisotropic diffusion is derived via the coupling of Fick’s law with the continuity equation. Fick’s law states that a concentration gradient causes a diffusion flux that aims to compensate for this concentration field as follows [33]:

\[
J(p) = -e^{(i)}(p)\nabla I(p). \tag{23}
\]
The generalized continuity equation is then expressed by:

$$\hat{\partial}_t I(p) = -\nabla \cdot J(p) + s$$

(24)

where $s$ is a function that describes the generation or removal of $I$, so-called source or sink. Plugging Fick’s law into the general continuity equation and setting $s$ to 0, we derive anisotropic diffusion as in (3). The anisotropic diffusion is hence adiabatic as shown in Fig. 4 since the source or sink in the continuity equation is eliminated, making the energy flow symmetric, as shown in Fig. 3(a).

As mentioned in previous section, the energy flow in adaptive smoothing is asymmetric, which enables us to classify the energy flow in adaptive smoothing into three cases as shown in Fig. 5.

1) First, the energy flow of adaptive smoothing is exactly the same as that of anisotropic diffusion as shown in Fig. 5(a), so there is no additional flow in adaptive smoothing.

2) Second, the energy flow of adaptive smoothing can be smaller or larger than that of anisotropic diffusion as shown in Fig. 5(b) and (c), corresponding that the sum of weights $\chi$ in (1) is larger or smaller than 1, respectively. It also implies that an additional flow exists in adaptive smoothing. Specifically, $p$ in Fig. 5(b) and (c) can be considered as the sink and source, respectively.

It leads to the conclusion that $s$ in (24) exists in adaptive smoothing as in (25).

$$\hat{\partial}_t I(p) = \nabla \cdot [c^{(t)}(p)\nabla I(p)] + s(p),$$

(25)

where $s(p)$ is a spatially-varying function. It is worthy of noting that one can design a new filter by appropriately modeling $s(p)$ according to specific applications [28].

Then, what function should be given as $s(p)$ in adaptive smoothing? By considering the energy flow in adaptive smoothing as described in Fig. 5, $s(p)$ should meet the following criteria.

1) First, it should scale the magnitude of the original flux $\nabla \cdot [c^{(t)}(p)\nabla I(p)]$ only while preserving its direction. It is worthy of noting that a rotation field can create an asymmetric flow as well, but it changes the direction/angle as well as the magnitude of a flow, which does not correspond to adaptive smoothing.

2) Second, its scaling strength depends on not only the current pixel but its neighboring pixels. When the current pixel is similar to the neighboring pixels, i.e., $\chi^{(t)}(p)$ is high, the flux decreases as shown in Fig. 5(b), corresponding that a scaling strength becomes smaller than 1 when the scaling strength is assumed to be 1 in case of Fig. 5(a), and vice versa. In summary, the scaling strength is inversely proportional to the sum of weights $\chi^{(t)}(p)$.

The following proposition can then be derived.

Proposition 2: In adaptive smoothing, $s(p)$ is a function of the original flux $\nabla \cdot [c^{(t)}(p)\nabla I(p)]$ and the sum of weights $\chi^{(t)}(p)$.

Proof: Equation (25) can be re-formulated by introducing a new function $\kappa$.

$$\hat{\partial}_t I(p) = \kappa(p)\nabla \cdot [c^{(t)}(p)\nabla I(p)]$$

(26)

where

$$\kappa(p) = \frac{\nabla \cdot [c^{(t)}(p)\nabla I(p)] + s(p)}{\nabla \cdot [c^{(t)}(p)\nabla I(p)]}.$$  

(27)
In adaptive smoothing, the scaling strength of $\kappa(p)$ is equal to the reciprocal of the sum of weights $\chi^{(t)}(p)$ by the second criterion. By arranging (27) with respective to $s(p)$, the following equation is derived:

$$s(p) = (\kappa(p) - 1) \nabla \cdot [e^{(t)}(p) \nabla I(p)] = \left(1 - \frac{1}{\chi^{(t)}(p) - 1}\right) \nabla \cdot [e^{(t)}(p) \nabla I(p)].$$

(28)

It is worth noting that the scaling factor $\kappa(p)$ in (26) is always larger than 0 since it should adjust the magnitude of flux only as in the first criterion.

**Corollary 1:** $\kappa(p)$ is a source or sink in the log diffusion equation.

**Proof:** Since $\kappa(p) > 0$, without loss of generality, we analyze (26) in the log domain by assuming that it becomes an equilibrium state as time $t$ goes infinite

$$ln[\tilde{c}_t I(p)] = \ln \left[\nabla \cdot [e^{(t)}(p) \nabla I(p)]\right] + ln[\kappa(p)].$$

(29)

By comparing (29) with (25), one can notice that (29) is anisotropic diffusion in the log domain. $ln[\kappa(p)]$ is hence called the source or sink in the log diffusion equation, which can be defined by

$$ln[\kappa(p)] = \ln \left[1 + \frac{s(p)}{\nabla \cdot [e^{(t)}(p) \nabla I(p)]}\right].$$

(30)

When $s(p)$ becomes 0 as in Fig. 5(a), i.e., $p$ is neither a source nor sink, it corresponds $ln[\kappa(p)] = 0$. In contrast, when $p$ is a sink ($s(p) < 0$) in Fig. 5(b) or a source ($s(p) > 0$) in Fig. 5(c), it corresponds to $ln[\kappa(p)] < 0$ or $ln[\kappa(p)] > 0$, respectively. Table I summarizes three cases of Fig. 5.

Equation (26) is then re-written by incorporating $\chi^{(t)}(p)$, which leads to the second order PDE as follows:

$$\chi^{(t)}(p) \tilde{c}_t I(p) = \nabla \cdot [e^{(t)}(p) \nabla I(p)].$$

(31)

Note that when $\chi^{(t)}(p)$ is set to 1, i.e., $s(p) = 0$, as in Fig. 5(a), adaptive smoothing in (31) becomes anisotropic diffusion in (3), and this exactly coincides with the constraint (14) or (15), i.e., sum of weights becomes 1. Therefore, the proposition 1 is supported once more.

Equation (31) is then discretized by an explicit FDM with a forward Euler approximation as follows:

$$I^{(t+1)}(x) = I^{(t)}(x) + \frac{\tau}{\sum_{q \in N_4} g \left(||I^{(t)}(q) - I^{(t)}(p)||\right)} \left[I^{(t)}(q) - I^{(t)}(p)\right]$$

$$+ \frac{\tau}{\sum_{q \in N_4} g \left(||I^{(t)}(q) - I^{(t)}(p)||\right)} g(0) + \frac{\tau}{\sum_{q \in N_4} g \left(||I^{(t)}(q) - I^{(t)}(p)||\right)} g(0)$$

$$= \frac{\alpha}{2} \frac{e^{(t)}(x-1) + e^{(t)}(x)}{\chi^{(t)}(x)} + \frac{\beta}{2} \frac{e^{(t)}(x+1) + e^{(t)}(x)}{\chi^{(t)}(x)} + \gamma$$

(32)

**Proposition 3:** The adaptive smoothing in (1) is equivalent to (32) when the evolution step size $\tau$ is 1.

**Proof:** Let us re-formulate the 1-D case of adaptive smoothing in (1) in order to link it with the second order PDE, considering the sum of weights $\chi^{(t)}$. Note that the normalization is considered in (33), different from (13)

$$I^{(t+1)}(x) = \alpha I^{(t)}(x-1) + \beta I^{(t)}(x) + \gamma I^{(t)}(x+1)$$

(33)

where

$$\alpha = \frac{e^{(t)}(x-1) + e^{(t)}(x)}{2\chi^{(t)}(x)},$$

$$\beta = \frac{e^{(t)}(x)}{\chi^{(t)}(x)},$$

$$\gamma = \frac{e^{(t)}(x+1) + e^{(t)}(x)}{2\chi^{(t)}(x)}.$$ (34)

Note that $\alpha$, $\beta$, and $\gamma$ are derived by using (6). In contrast to (22), the following equation is always satisfied regardless of $\chi^{(t)}(x)$ since the normalization is embedded in $\alpha$, $\beta$, and $\gamma$:

$$\alpha + \beta + \gamma = 1.$$ (36)

By substituting $\alpha$, $\beta$, and $\gamma$ in (33) with (34), we derive the following equation:

$$I^{(t+1)}(x) - I^{(t)}(x)$$

$$= \frac{e^{(t)}(x-1) + e^{(t)}(x)}{2\chi^{(t)}(x)} \left(I^{(t)}(x-1) - I^{(t)}(x)\right)$$

$$+ \frac{e^{(t)}(x+1) + e^{(t)}(x)}{2\chi^{(t)}(x)} \left(I^{(t)}(x+1) - I^{(t)}(x)\right)$$

$$= \frac{\left(I^{(t)}(x-1) - I^{(t)}(x)\right)}{g \left(||I^{(t)}(x-1) - I^{(t)}(x)||\right)} + \frac{\left(I^{(t)}(x+1) - I^{(t)}(x)\right)}{g \left(||I^{(t)}(x+1) - I^{(t)}(x)||\right)}$$

(37)

This represents the implementation of a second order PDE in (32) with the evolution step size $\tau$ being 1, meaning that adaptive smoothing in (33) is linked with the second order PDE with the source or sink.

**Remark 1:** The normalization term of a weighted average filter such as adaptive smoothing [8] generates an asymmetric energy flow, and comes from the generalized continuity equation in which the source or sink exists.

**Corollary 2:** All $p$'s in adaptive smoothing are sink if $g(0) = 1$. 

| Table 1: Relation Between $s(p)$ and $\kappa(p)$ |
|-------------------|-------------------|-------------------|
| Fig. 5(a) | Fig. 5(b) | Fig. 5(c) |
| $\chi^{(t)}(p)$ | 1 | 1 | 1 |
| $\kappa(p)$ | 1 | 0 | 0 |
| $s(p)$ | 0 | 1 | 0 |
| $\ln[s(p)]$ | 0 | 0 | 0 |
| Anisotropic Diffusion | O | O | O |

Adaptive Smoothing | O | O | O |
Proof: If \( g(0) = 1 \), \( \kappa(p) \) is always smaller than 1 as in Fig. 5(b), which makes \( \ln[\kappa(p)] < 0 \) (\( s(p) < 0 \)). Therefore, as shown in Fig. 4, the normalized mean value of the results filtered by adaptive smoothing, where \( g(0) \) is set to 1, monotonically decreases.

C. Behavior of Adaptive Smoothing

In this section, we will examine the behavior of adaptive smoothing such as the maximum principle and stability condition within the framework of a PDE.

1) Maximum Principle: We verify that (32), which was proven to be equivalent to adaptive smoothing, satisfies the maximum principle, i.e., no new maxima and minima appear as an image is filtered. Although anisotropic diffusion in (8) also satisfies the maximum principle [4], anisotropic diffusion and adaptive smoothing have a different theoretical origin, as mentioned in section III-B.

Proposition 4: The adaptive smoothing satisfies the maximum principle.

Proof: When the 4-neighborhood is used, the maximum and minimum values among the center node \( p \) and \( N_4 \) are defined by

\[
M^{(t)}(p) = \max \{ I^{(t)}(p), I^{(t)}(q) \}_{q \in N_4} \tag{38}
\]

\[
m^{(t)}(p) = \min \{ I^{(t)}(p), I^{(t)}(q) \}_{q \in N_4} \tag{39}
\]

Equation (32) can be modified as

\[
I^{(t+1)}(p) = I^{(t)}(p) \left( 1 - \tau \frac{\sum_{q \in N_4} g(I^{(t)}(q) - I^{(t)}(p)) \chi^{(t)}(q)}{g(0) + \sum_{q \in N_4} g(I^{(t)}(q) - I^{(t)}(p))} \right) + \tau \frac{\sum_{q \in N_4} g[I^{(t)}(q) - I^{(t)}(p)] I^{(t)}(q) g(0) + \sum_{q \in N_4} g(I^{(t)}(q) - I^{(t)}(p))}}{g(0) + \sum_{q \in N_4} g(I^{(t)}(q) - I^{(t)}(p))} \]

\[
\leq M^{(t)}(p) \left( 1 - \tau \frac{\sum_{q \in N_4} g[I^{(t)}(q) - I^{(t)}(p)] \chi^{(t)}(q)}{g(0) + \sum_{q \in N_4} g[I^{(t)}(q) - I^{(t)}(p)]} \right) + \tau \frac{\sum_{q \in N_4} g[I^{(t)}(q) - I^{(t)}(p)] m^{(t)}(p) g(0) + \sum_{q \in N_4} g[I^{(t)}(q) - I^{(t)}(p)]}{g(0) + \sum_{q \in N_4} g[I^{(t)}(q) - I^{(t)}(p)]} \]

\[
m^{(t)}(p) \leq I^{(t+1)}(p) \leq M^{(t)}(p). \tag{40}
\]

Similarly

\[
I^{(t+1)}(p) \geq m^{(t)}(p) \left( 1 - \tau \frac{\sum_{q \in N_4} g[I^{(t)}(q) - I^{(t)}(p)] \chi^{(t)}(q)}{g(0) + \sum_{q \in N_4} g[I^{(t)}(q) - I^{(t)}(p)]} \right) + \tau \frac{\sum_{q \in N_4} g[I^{(t)}(q) - I^{(t)}(p)] m^{(t)}(p) g(0) + \sum_{q \in N_4} g[I^{(t)}(q) - I^{(t)}(p)]}{g(0) + \sum_{q \in N_4} g[I^{(t)}(q) - I^{(t)}(p)]} \]

\[
m^{(t)}(p) \leq I^{(t+1)}(p) \leq M^{(t)}(p). \tag{41}
\]

Thus

\[
m^{(t)}(p) \leq I^{(t+1)}(p) \leq M^{(t)}(p). \tag{42}
\]

2) Stability: It is somewhat intuitive that the normalization term prevents filtered images from diverging. It is related to the maximum norm stability in graph theory [17]. To the best of our knowledge, there have been no studies exploring this observation in the viewpoint of a PDE.

Proposition 5: The adaptive smoothing is always stable.

Proof: The stability condition of (32) is \( 0 \leq \tau \leq 5/4 \), since the weight of the center node \( p \) should be between 0 and 1 in order to ensure the stability as follows:

\[
\tau < \max \left( \frac{g(0) + \sum_{q \in N_4} g[I^{(t)}(q) - I^{(t)}(p)]}{\sum_{q \in N_4} g[I^{(t)}(q) - I^{(t)}(p)]} \right) = \max \left( 1 + \frac{g(0)}{\sum_{q \in N_4} g[I^{(t)}(q) - I^{(t)}(p)]} \right) = 5/4.
\]

(43)

We showed that adaptive smoothing is a discrete approximation of (31) with the evolution step size being 1. Since the evolution step size in adaptive smoothing is always smaller than 5/4 regardless of the function \( g(\cdot) \), adaptive smoothing is always stable.

IV. INSTANTIATIONS OF THE PROPOSED PDE: ROBUST DIFFUSION

In this section, two specific diffusion methods are instantiated by leveraging the asymmetric energy flow in the diffusion. First, the RAD [28] is derived by differently modeling the source or sink in the proposed PDE. Based on this observation, the RCED is further proposed, which preserves universal features and enhances coherence structures better than the conventional one [34], [36], [37].

A. RAD

The RAD regularizes the image with an assumption that the adiabatic process such as the diffusion is not suitable in handling the outliers, e.g., impulsive noise [28]. Namely, an additional flux exists in the RAD so that the source or sink \( s(p) \) in (25) is not 0, as opposed to anisotropic diffusion [4]. The additional flux plays a role in such a way that the outlier signal is compensated by adaptively changing the amount of flux according to the local topology of the neighborhood, which results in reducing the influence of outliers significantly.

Ham et al. modeled \( \kappa(p) \) as follows [28]

\[
\kappa(p) = \frac{1}{\chi^{(t)}(p) - g(0)}. \tag{44}
\]

The quantity of \( \chi^{(t)}(p) - g(0) \) is an indicator of the outlier, e.g., when this quantity is small, it implies that the center node is likely to be an outlier.

Then, the RAD is defined as follows:

\[
[\chi^{(t)}(p) - g(0)]_{\partial I(p)} = \nabla \cdot [\varepsilon^{(t)}(p) \nabla I(p)]. \tag{45}
\]

Note that anisotropic diffusion [4] in (3), the adaptive smoothing [8] in (31) and the RAD in (45) are all the special case of the PDE of (26).

Fig. 6 shows an example of (b) anisotropic diffusion [4] and (c) the RAD with (a) a degraded image. The initial Cat image
[10] was degraded by the Gaussian noise with a standard deviation 0.1 and the impulse noise with a density of 0.05. All parameters were set equal in both methods: \( g(\cdot) \) as the Gaussian kernel with an amplitude 1 and a fixed standard deviation 0.01, an evolution step size \( \tau \) of 0.25, and the number of iteration \( t \) of 500. It demonstrates that the RAD can handle the mixture noise very well, in contrast to the conventional anisotropic diffusion. Please refer to [28] for more results and intensive analysis of the RAD.

B. RCED

In this section, we further propose the RCED by incorporating additional fluxes into coherence enhancing diffusion in a similar manner to the RAD, making the proposed diffusion more robust against outliers as well as better enhance coherence structures.

1) RCED: Over the last two decades, there have been many studies on analyzing and enhancing the flow-like structure in the field of image processing [34]–[37]. It has been usually done by a well-established tool from texture analysis based on the structure tensor (second moment matrix), the eigenvalues and eigenvectors of which provide us with all required information for speculating the structure embedded in the image [39].

First, let us define the structure tensor as follows:

\[
J_\rho = K_\rho(p) \ast \left[ \nabla I_\sigma^{(1)}(p) \nabla I_\sigma^{(1)}(p)^T \right] = \lambda_+ \theta_+^T + \lambda_- \theta_-^T.
\]  

(46)

where

\[
K_\sigma(p) = \frac{1}{2\pi\sigma^2} \exp \left( -\frac{||p||^2}{2\sigma^2} \right)
\]  

(47)

\[
\nabla I_\sigma^{(1)}(p) = K_\sigma(p) \ast \nabla I^{(1)}(p).
\]  

(48)

This matrix is symmetric and positive definite, thus has two eigenvalues \( \lambda_+, \lambda_- \) and corresponding eigenvectors \( \theta_+, \theta_- \) which are tangential and orthogonal to \( \nabla I_\sigma^{(1)}(p) \), respectively.

The eigenvalues of \( J_\rho \) are

\[
\lambda_\pm = \frac{1}{2} (j_{11} + j_{22} \pm \Delta)
\]  

(49)

where

\[
\Delta = \sqrt{(j_{11} - j_{22})^2 + 4j_{12}^2}.
\]  

(50)

Note that in a scalar image, the eigenvalues and eigenvectors of \( J_0 \) are

\[
\lambda_+ = \|\nabla I\|^2, \quad \lambda_- = 0
\]  

(51)

and

\[
\theta_+ = \frac{\nabla I}{\|\nabla I\|}, \quad \theta_- = \frac{\nabla I_\perp}{\|\nabla I\|}.
\]  

(52)

Directly employing the structure tensor \( J_\rho \) as the diffusion tensor will lead to fast diffusion across the edge and slow diffusion along the edge, which is opposite to our intention [40]. For enhancing coherence within the flow-like structure, a regularization should act mainly along the flow direction. Also, the smoothing should increase according to the strength of its orientation which can be measured by some metric, e.g., \((\lambda_+ - \lambda_-)^2\) becomes large for strongly differing eigenvalues, and tends to zero for isotropic structures. Therefore, the diffusion tensor is constructed as in (53) with the same
Fig. 8. Filtering results for corrupted images of Fig. 7(b). (a) Coherence enhancing diffusion (CED) [34]. (b) Anisotropic Kuwahara filter (AKF) [36]. (c) Coherence enhancing shock filter (CES) [37]. (d) Robust coherence enhancing diffusion (RCED). All parameters are fixed during experiments. In the CED and the RCED, $\rho = 5$, $\sigma = 0.7$, $\omega = 0.01$, and $C = 0.001$. The number of iterations and the evolution step size were set to 100 and 0.2, respectively. In the AKF and the CES, the parameters are set to default values used in [36] and [37], respectively. Note that the results of both methods were produced using software provided by the authors [38].

eigenvectors as the structure tensor $J_{\rho}$ [41], [42]

$$D^{(1)}(p) = \lambda_1 \theta_+ \theta_+^T + \lambda_2 \theta_- \theta_-^T. \quad (53)$$

The diffusion tensor $D^{(1)}(p) = \begin{pmatrix} d_{11} & d_{12} \\ d_{12} & d_{22} \end{pmatrix}$ is a $2 \times 2$ symmetric and positive definite matrix with two positive eigenvalues $\lambda_1$, $\lambda_2$ and two corresponding eigenvectors $\theta_+, \theta_-$. Each component can be calculated as follows:

$$d_{11} = \frac{1}{2} \left[ \lambda_1 + \lambda_2 - \frac{(\lambda_2 - \lambda_1)(j_{11} - j_{22})}{\Delta} \right],$$
$$d_{12} = \frac{(\lambda_1 - \lambda_2) j_{12}}{\Delta},$$
$$d_{22} = \frac{1}{2} \left[ \lambda_1 + \lambda_2 + \frac{(\lambda_2 - \lambda_1)(j_{11} - j_{22})}{\Delta} \right].$$

where the eigenvalues are

$$\lambda_1 = \omega \quad \text{if} \quad \lambda_+ = \lambda_-$$
$$\lambda_2 = \begin{cases} \omega & \text{if} \quad (\lambda_+ - \lambda_-)^2 \gg C \\ \omega + (1 - \omega) \exp \left\{ -\frac{C}{(\lambda_+ - \lambda_-)^2} \right\} & \text{else} \end{cases} \quad (54)$$

$\omega \in (0, 1)$ represents the regularization parameter which keeps the diffusion tensor positive definite [34]. $C > 0$ serves as a threshold parameter: $\lambda_2 \approx 1$ for $(\lambda_+ - \lambda_-)^2 \gg C$, and $\lambda_2 \approx \omega$ for $(\lambda_+ - \lambda_-)^2 \ll C$.

Then, the RCED is defined as follows:

$$\tilde{\partial}_t I(p) = \kappa(p) \nabla \cdot [D^{(1)}(p) \nabla I(p)]. \quad (55)$$

Similar to RAD, the source or sink in (56) is modeled as

$$\kappa(p) = \frac{1}{\kappa^{(0)}(p) - g(0)}. \quad (56)$$
Note that $\kappa(p)$ is an isotropic since $\kappa(p)$ which is an indicator of the existence of the outliers, is irrelevant to an orientation of the edges. In other words, it is assumed that a probability of being corrupted by the outliers is independent of the orientation of the edges.

2) Experimental Results: To verify the performance, we compared the proposed method with the conventional coherence enhancing diffusion (CED) [34] and other related methods such as anisotropic Kuwahara filter (AKF) [36] and the coherence enhancing shock filter (CES) [37]. These were applied to original images and images degraded by the mixture noises, i.e., the Gaussian noise with a standard deviation of 0.1 and the impulsive noise with a density of 0.1, as shown in Fig. 7. All the parameters were fixed during experiments. In the CED and the RCED, $\rho = 5$, $\sigma = 0.7$, $\omega = 0.01$ and $C = 0.001$. The number of iteration and the evolution step size were set to 100 and 0.2, respectively. The Gaussian kernel with a standard deviation 0.01 was used as $g(\cdot)$. In the AKF and the CES, the parameters were set to default values used in [36] and [37], respectively. Note that the filtering results of the two methods were produced by authors-provided softwares [38].

Fig. 8(a) shows that the Gaussian noise can be effectively handled by the CED [34], but impulsive noise still exists even after long evolution. Note that the vector $\nabla I_0$ and the matrix $J_0$ are regularized by the Gaussian kernel $K$ in constructing the diffusion tensor so some outliers have been eliminated before the diffusion process. The AKF is anisotropic counterpart of the weighted Kuwahara filter [43], thus artifacts which exist in the Kuwahara filter are avoided while directional image features are better preserved and emphasized. Fig. 8(b) shows that the AKF is robust against the outliers [36], but the region boundaries are distorted. The CES is not robust against the outliers, and even the outliers are sharpened and enhanced since the shock filter is embedded in the CES. In contrast, the RCED
handles the impulsive noise as well as the Gaussian noise very well. Furthermore, it can preserve singular features better than other methods, e.g., mandrill’s eye.

The same experiments were conducted with the noise-free image as shown in Fig. 9. It also demonstrates the proposed method is more capable of enhancing the flow-like structures than other methods. Although the CED enhances the coherence structure well, it also blurs some important features. Meanwhile, the diffusion velocity of the proposed method automatically decreases when the features begin to be flattened, i.e., $|\chi(0)(p) - g(0)|$, an indicator of the existence of the outliers, increases, thus leading to preserving important features while enhancing the coherence structures. The CES sharpens and enhances the coherence structure well, but the additional procedure such as shock process is needed. We also found that the coherence enhancing capability of the RCED is better than that of the AKF (see mandrill’s fur).

V. CONCLUSION

This paper differentiated adaptive smoothing from anisotropic diffusion in the viewpoint of a normalization, an evolution step size, and an energy flow. While anisotropic diffusion has a symmetric flow since the diffusion is theoretically an adiabatic process, adaptive smoothing has an asymmetric energy flow. Based on this principle, adaptive smoothing was drawn from the generalized second order PDE where the source or sink exists. It provides new insights into application-specific filtering algorithm design such as the non-local diffusion [14] and the unnormalized bilateral filter [27]. Also, the behavior of adaptive smoothing such as the maximum principle and stability has been examined with the perspective of a PDE by leveraging that the source or sink is closely related to the normalization term of adaptive smoothing. Furthermore, new diffusion filters such as the RAD and the RCED have been designed by properly modeling the source or sink, thus generating the asymmetric diffusion flow which is more robust against the outliers.

REFERENCES


[38] [Online]. Available: http://www.kyprianidis.com/projects.html


Bumsun Ham (S’09) received the B.S. degree in electrical and electronic engineering from Yonsei University, Seoul, Korea, in 2008, where he is currently pursuing the joint M.S. and Ph.D. degree at Yonsei University.

His current research interests include variational methods and PDEs to solve 3-D computer vision and image processing problems, particularly stereo vision, 3-D modeling, super-resolution, and HDR imaging.

Bumsub Ham (S’09) received the B.S. degree in electrical and electronic engineering from Yonsei University, Seoul, Korea, in 2008, where he is currently pursuing the joint M.S. and Ph.D. degree at Yonsei University.

His current research interests include variational methods and PDEs to solve 3-D computer vision and image processing problems, particularly stereo vision, 3-D modeling, super-resolution, and HDR imaging.

Dongbo Min (M’09) received the B.S., M.S., and Ph.D. degrees in electrical and electronic engineering from Yonsei University, Seoul, Korea, in 2003, 2005, and 2009, respectively.

He was a Post-doctoral Researcher with Mitsubishi Electric Research Laboratories (MERL) from June 2009 to June 2010. He is now with the Advanced Digital Sciences Center (ADSC), which was jointly founded by the University of Illinois at Urbana-Champaign (UIUC) and the Agency for Science, Technology, and Research (A*STAR), a Singapore government agency. His research interests include 3-D computer vision, video processing, 3-D modeling, and hybrid sensor systems.

Kwanghoon Sohn (M’92) received the B.E. degree in electronic engineering from Yonsei University, Seoul, Korea, in 1983, the MSEE degree in electrical engineering from University of Minnesota in 1985, and the Ph.D. degree in electrical and computer engineering from North Carolina State University in 1992.

He was a Senior Member of the Research Staff with the Satellite Communication Division at the Electronics and Telecommunications Research Institute, Daejeon, Korea, from 1992 to 1993 and a Post-doctoral Fellow at the MRI Center, Medical School of Georgetown University in 1994. He was a Visiting Professor at Nanyang Technological University from 2002 to 2003. He is currently a Professor in the School of Electrical and Electronic Engineering, Yonsei University. His research interests include 3-D image processing, computer vision and image communication.

Dr. Sohn is a member of SPIE.