Mumford-Shah Regularizer with Contextual Feedback [*]

Erkut Erdem [**], **Sibel Tari**[,***,†]

Department of Computer Engineering, Middle East Technical University

e-mail: `erkut@ceng.metu.edu.tr`, `stari@metu.edu.tr`

Abstract We present a simple and robust feature preserving image regularization by letting local region measures to modulate the diffusivity. The purpose of this modulation is to disambiguate low level cues in early vision. We interpret the Ambrosio-Tortorelli approximation of the Mumford-Shah model as a system with modulatory feedback and utilize this interpretation to integrate high level information into the regularization process. The method does not require any prior model or learning; the high level information is extracted from local regions and fed back to the regularization step. An important characteristic of the method is that both negative and positive feedback can be simultaneously used without creating oscillations. Experiments performed with both gray and color natural images demonstrate the potential of the method under difficult noise types, non-uniform contrast, existence of multi-scale patterns and textures.

Key words variational and PDE methods, feature preserving diffusion, structure preserving diffusion, disambiguation in low level vision.

1 Introduction

The prevalent view in computer vision since Marr [53] is to assume early vision as a data-driven, bottom-up process. As the raw image data is processed in a feedforward fashion, a sequence of processes transforms low-level cues into larger perceptual units. Over the years, a variety of techniques for low-level vision has been developed using this paradigm. Most of these techniques cannot easily cope with ambiguities since information flow is in one direction and low-level cues are always

presumed to be reliable. Hence, any misinterpretation in the early stages greatly affects the later ones.

For example, consider edge detection from a noisy image. When a low-pass filter denoising process is applied to the image, the process not only removes the noise but also softens intensity discontinuities making the edges difficult to detect. If the cut-off frequency of the low-pass filter is set to a large value, then one cannot eliminate noise. The noise, which is passed to the second level in the hierarchy, is enhanced more than the signal itself [39] by the edge detection process. About twenty years ago, the dilemma in the one way information flow between smoothing and edge detection, prompted eminent researchers including Mumford and Shah to propose methods that combine smoothing and edge detection [14, 41, 56, 59].

Mumford and Shah [56] formulated image segmentation process as a functional minimization via which a piecewise smooth approximation of a given image and an edge set are to be recovered simultaneously. It aims to decompose an image into cartoon and noise components. The Mumford-Shah (MS) model is:

$$E_{MS}(u, \Gamma) = \beta \int_R (u - g)^2 dx + \alpha \int_{R \setminus \Gamma} |\nabla u|^2 dx$$
$$+ length(\Gamma) \quad (1)$$

where

– $R \subset \Re^2$ is connected, bounded, open subset representing the image domain,

– g is an image defined on R,

– $\Gamma \subset R$ is the edge set segmenting R,

– u is the piecewise smooth approximation of g,

– α, β are the scale space parameters of the model.

The first term in $E_{MS}(u, \Gamma)$ is the data fidelity term, which forces the solution u to be close to the original image g. The other two terms are regularization terms, which encode a priori information about the solution and give preference to piecewise smooth images with simple edge sets [55].

The unknown edge set Γ of a lower dimension makes the minimization difficult. A convenient approximation is suggested by Ambrosio and Tortorelli [4] following the Γ convergence framework [15]. The basic idea is to introduce a smooth edge indicator function v which is more convenient than the original edge indicator represented by the characteristic function $1 - \chi_\Gamma$. The function v depends on a parameter ρ, and as $\rho \to 0$, $v \to 1 - \chi_\Gamma$. That is, $v(x) \approx 0$ if $x \in \Gamma$ and $v(x) \approx 1$ otherwise. The cardinality of the edge set Γ can be approximated by $\frac{1}{2} \left(\rho |\nabla v|^2 + \frac{(1-v)^2}{\rho} \right)$. The new functional is as follows:

$$E_{AT}(u, v) = \int_R \left(\beta(u - g)^2 + \alpha(v^2 |\nabla u|^2) + \frac{1}{2} \left(\rho |\nabla v|^2 + \frac{(1 - v)^2}{\rho} \right) \right) dx \quad (2)$$

The most appealing property of the Ambrosio-Tortorelli (AT) approximation is that one can apply gradient descent to obtain the condition for minima in the form of

a system of coupled PDEs:

$$\frac{\partial u}{\partial t} = \nabla \cdot (v^2 \nabla u) - \frac{\beta}{\alpha}(u - g); \quad \left.\frac{\partial u}{\partial n}\right|_{\partial R} = 0 \quad (3)$$

$$\frac{\partial v}{\partial t} = \nabla^2 v - \frac{2\alpha|\nabla u|^2 v}{\rho} - \frac{(v-1)}{\rho^2}; \quad \left.\frac{\partial v}{\partial n}\right|_{\partial R} = 0 \quad (4)$$

where ∂R denotes the boundary of R and n denotes the direction normal to ∂R.

The coupled equations can be simultaneously solved for u and v using standard numerical discretization techniques such as Finite Differences. When these equations are discretized using a modified explicit scheme (Appendix A), the iterations converge in the sense that the rate of change is smaller than a threshold. In each iteration, only one variable is updated while the other variable is kept fixed.

Keeping v fixed, (3) minimizes a convex quadratic functional given by

$$\int_R \left(\alpha v^2 |\nabla u|^2 + \beta(u - g)^2\right)dx \quad (5)$$

While the bias term in (3) or equivalently in (5) forces u to be close to the original image g, the first term acts as an edge preserving regularizer. It smoothes the image with a smoothing radius proportional to the values of v^2 and $\frac{\alpha}{\beta}$. If there is an edge ($v \approx 0$), no smoothing (diffusion) is carried out.

As $v \to 0$, the smoothness constraint in the piecewise smooth model is switched off. It is possible to interpret v^2 as an analog form of the line process introduced by Geman and Geman [41]. As shown by Bar, Kiryati and Sochen [7] and Teboul et al. [78], the AT approximation of the MS functional defines an extended line process regularization where the regularizer has an additional constraint introduced by the term $\rho|\nabla v|^2$. This term mildly forces some spatial organization by demanding the edges to be smooth. However, the reconstruction results are still affected by the existing noise or texture (Fig. 1).

Over the years, a variety of modifications were proposed for the AT model. Shah [72] considered replacing the quadratic cost functions in both the data fidelity and the regularizer with L^1-functions. Erdem, Erdem and Tari [36] considered incorporating prior shape information into the AT by introducing additional terms that also make the minimization difficult. Esedoglu and Shen [38] considered incorporating higher order geometric terms in the length functional in order to improve the regularity of the edge sets found by the MS. An extreme modification to the AT has been proposed by Tari, Shah and Pien [77], by using large ρ values and interpreting v as a smoothed distance function from which a skeletal analysis is performed. Recently, Aslan and Tari [5] considered letting $\rho \to \infty$. Whereas the latter method can only handle silhouettes, the former one can handle piecewise smooth images.

In this paper, we propose a simple yet very effective modification to the AT that will convert it to a regularizer with much better feature preserving capa-
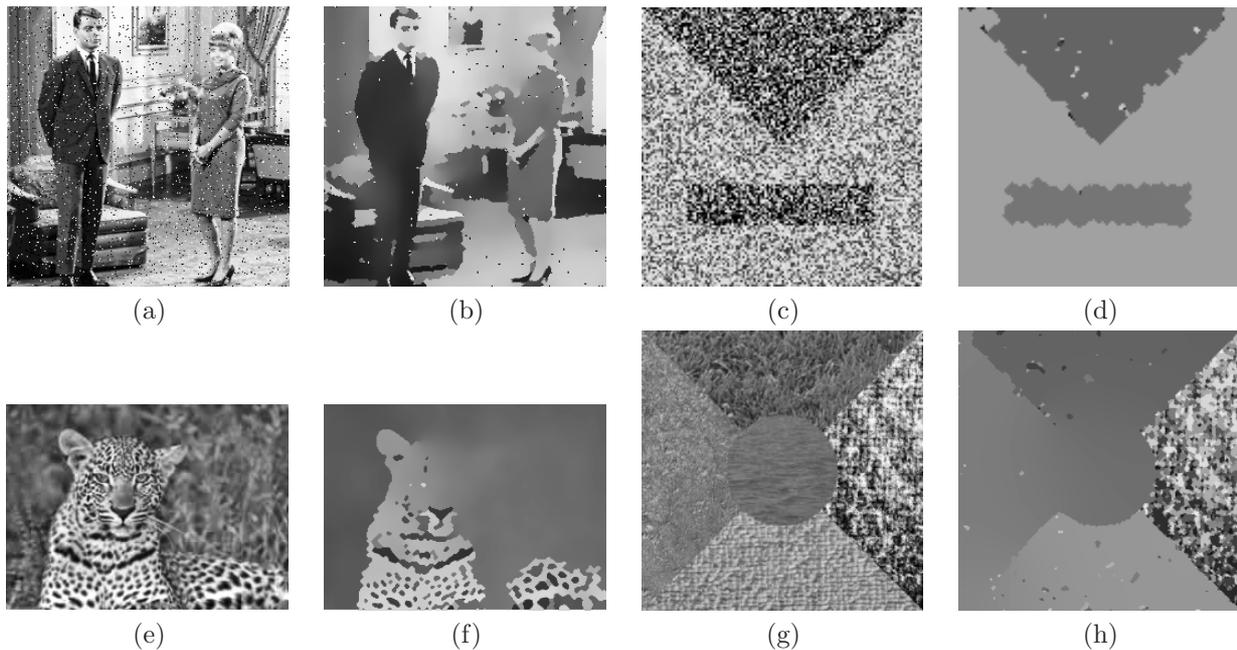
Fig. 1 Images which cannot be handled by the AT model. (a)-(b) *couple* image corrupted with 5% salt and pepper noise and its reconstruction using AT. (c)-(d) A noisy test image –70% of the pixels are degraded with uniform noise– and its reconstruction using AT. (e)-(f) An animal image and its reconstruction using AT. (g)-(h) A texture mosaic image and its reconstruction using AT.

bilities, without resorting to additional energies, prior models and difficult to minimize cost functions. The new model can cope with difficult noise types such as impulse noise and preserve structures even in highly textured situations. Interesting non-linear and non-local behaviors arise, while the computations are carried out in a locally linear fashion.

Key to our approach is the link between regularization and diffusion filters [13,55,57,68,71]. After casting the AT as a biased diffusion filter, we use local measures to steer the diffusion. The local measures are computed from a collection of local neighborhoods that are not necessarily centered on the pixel to be regularized. They are referred as *contextual feedback measures* or simply *feedback measures* throughout the paper.

1.1 Related Work

Contextual influences have been explored within both computer vision and neuroscience community, e.g. [1,9, 10,33–35,44,50,52,69,73,75,76,79,83,85]. The notion of context differs and typically means global image statistics, nearby objects, nearby pixels or statistics of features [74,83]. In our work, context is a collection of nearby local neighborhoods or nearby local image patches and contextual measures are aggregate quantities that are computed from these collections. Our work differs from the above efforts in the sense that we use a PDE framework as the computational platform and we focus on the MS regularization.

There are works which use the MS regularizer or its modification [3,72] for restoration in the presence of im-

pulse noise or textures [7,8,72,78]. In [7,8], Bar, Kiryati and Sochen present a very promising approach. However, the success of their method stems mostly from the use of robust data fidelity by replacing the L^2-norm with the L^1. Similarly, in [72], Shah uses the L^1-norm for both the data fidelity and the regularizer. This generates shocks in u and thus object boundaries are recovered as actual discontinuities without being much affected by noise or texture. In fact, the use of non-smooth cost functions such as the L^1-norm for the data fidelity term in order to deal with outliers is well motivated both theoretically and experimentally (e.g. [17,30,31,58,72]). Teboul et al. [78] present a modification of (2), by replacing the quadratic cost $|\nabla v|^2$ with the L^1-cost, which leads to singular diffusivity. Numerical difficulties are the major drawback of singular diffusivities [21]. The cost function choice in [78] also leads to directional smoothing. As explored by Weickert [82], directional smoothing can offer significant feature preserving capabilities. However, the models get complicated and the numerics is not as simple as in the case of isotropic diffusion. Our effort can be compared to that of Weickert [82] in the sense that we also consider gray values in a local image region to steer diffusion. Our contextual feedback measures are more general and higher level. Moreover, we remain in the isotropic setting.

Incorporation of high level knowledge to segmentation has been used in practice (e.g. [18,19,24,28,29,36, 45,47,49,51,60,62,65,80,84]). On the other hand, use of priors in filtering is much less explored [64,67,86]. Our approach differs from the latter works in the sense that we do not employ any prior model or learning.

One of our contextual feedback measures (edge continuity - Section 2.2) is closely related to spatial organization energy of Black and Rangarajan [11]. Rather than introducing an additional energy term, we use spatial organization to provide feedback that steers diffusion.

Our motivation in defining the texture edges feedback measure in Section 2.3 is to eliminate texture and preserve structure. In this respect, our texture edges feedback strategy is related to the work of Aleman-Flores, Alvarez and Caselles [2]. Whereas they modify Perona-Malik so that the diffusivity is a function of texture gradient computed from Gabor feature space, we use a measure computed from local patches as a modulatory feedback.

We utilize the concept of spatially adaptive image processing [32] which has also been investigated under the computational frameworks of image algebra [63] and mathematical morphology [70]. Our selection of neighborhood in the computation of directional consistency feedback measure in Section 2.1 is adaptive in the sense defined in [32].

Our approach can also be related to the one in Gilboa et al. [42] in the sense that both of the methods consider contextual knowledge apart from the other variational denoising frameworks that are based on the local features of the image such as derivatives. While [42] is based

on a non-local convex regularization term that depends on the contextual similarity suggested in [20], we remain in the local and simple framework of the AT.

2 Proposed Method

In the proposed method, as in the AT approximation, there are two coupled processes: the image process u and the edge process v. The edge process v evolves according to the same dynamics as given in (4). On the other hand, the image process u evolves according to

$$\frac{\partial u}{\partial t} = \nabla \cdot ((cv)^2 \nabla u) - \frac{\beta}{\alpha}(u - g); \quad \frac{\partial u}{\partial n}\bigg|_{\partial R} = 0 \quad (6)$$

$$cv = \phi v + (1 - \phi)V$$

$V \in \{0, 1\}$ and ϕ is a measure which takes values in $[0, 1]$ and it depends on u and/or v in a collection of neighborhoods. Therefore ϕ is at a higher place, in the visual hierarchy, than u and v. We discuss the role played by the multiplier c, by considering an abstract view (Fig. 2) of the discrete approximations of the coupled PDEs which are given by (24) and (25) in Appendix A.

For any image point i, the value of u_i^{t+1} is computed using u_i^t, the neighborhood of u_i^t, g_i, and the feedback which depends on ϕ_i. Notice that the direct connection from v to u in AT is replaced by an indirect path; first rising higher up in the visual hierarchy and coming back to steer the diffusion of u.
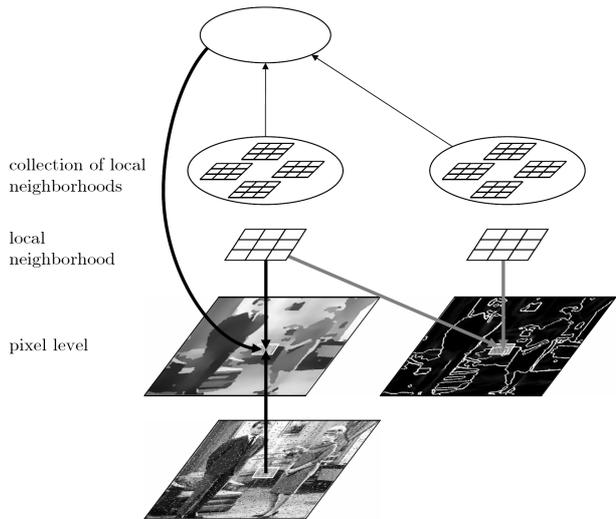

Fig. 2 Proposed architecture. Aggregate quantities computed from a collection of neighborhoods are utilized to steer the evolution of the processes u and v. See text.

Suppose that we wish to eliminate an accidentally occurring feature such as a high gradient due to noise. We can use edge gradient direction consistency (as in the edge linking step of traditional boundary detectors) as a feedback measure to modulate the diffusivity in (6) as

$$(cv)_i^2 = (\phi_i v_i + (1 - \phi_i) \, 1)^2 \quad (7)$$

Notice that the higher the value of ϕ_i, the higher the likelihood that the high gradient is meaningful. Thus, the lower the value of ϕ_i, the higher the diffusivity shifts to the maximum value 1. We call this type of modulation *negative feedback*.

Suppose that we wish to prevent an accidental elimination of a feature of interest, e.g. preserving a fine texture or encouraging edge formation. One can define ϕ_i such that it is low for meaningful occurrences, and let it

modulate the diffusivity in (6) as

$$(cv)_i^2 = (\phi_i v_i + (1 - \phi_i) \, 0)^2 \qquad (8)$$

Notice that the lower the value of ϕ_i, the higher the diffusivity shifts to the minimum value 0 in order to preserve the meaningful feature. We call this type of modulation *positive feedback*.

A computational advantage of the proposed approach is that both negative and positive feedbacks are implemented as regularization. This provides robustness. Another source of robustness is the separate evolution of the edge indicator v_i from its modulated form $(cv)_i$, even though, they eventually seem to converge to each other. The separate evolution of these variables prevents oscillations in the edge indicator, and it allows us to effectively use both positive and negative feedback. Finally, notice that c_i, as indicated by (7) and (8), is not explicitly computed. We can think of c_i as an indirect measure of feedback.

In the following subsections, we present different choices of the contextual feedback measure ϕ, which are based on directional consistency of the edges (ϕ^{dc}), edge continuity (ϕ^h), texture edges (ϕ^{te}), and local scale (ϕ^{ls}), respectively. The ones presented in Section 2.1 and Section 2.3 are to be interpreted as negative feedback whereas the ones presented in Section 2.2 and Section 2.4 are to be interpreted as positive feedback.

Since the feedback measure induces a multiplier for the diffusivity function, it is possible to consider a coalition of contextual feedback measures by taking the multiplier c as the product of individual c values.

2.1 Directional Consistency: ϕ^{dc}

Traditionally, segmentation is defined as a sequential bottom-up process composed of the following three steps: smoothing, edge detection and edge linking. The purpose of the edge linking step is to force global consistency to locally detected edges in order to come up with a coherent edge set. In this step, the edge pixels that are detected based on the magnitude of image gradients are linked to give a connected edge set if their gradient directions are in agreement. The unlinked edge pixels are discarded. Interestingly, this last step is what the MS model or its AT approximation lacks. We induce such an effect in our diffusion model by increasing the relative persistence of the edge pixels, which are consistent with their neighbors, by increasing the diffusivity at inconsistent ones. We consider a feedback measure ϕ_i^{dc} such that $\phi_i^{dc} \to 1$ on the preferred configurations and $\phi_i^{dc} \to 0$ on the incoherent configurations and let c_i has the following form:

$$c_i = 1 + (1 - \phi_i^{dc})\frac{1 - v_i}{v_i} \qquad (9)$$

Firstly, notice that c_i increases in proportion to the image gradient $|\nabla u_i|$, which is proportional to $\frac{1 - v_i}{v_i}$ (see (26) in Appendix B). Secondly, notice that the overall

diffusivity coefficient $(cv)_i^2$ can be estimated as follows, without explicitly computing the variable c_i:

$$(cv)_i^2 = \left(\phi_i^{dc} v_i + \left(1 - \phi_i^{dc}\right) 1\right)^2 \qquad (10)$$

The value of the diffusivity is bounded by 1. It attains its maximum as $\phi_i^{dc} \to 0$ and decays to a value determined by the edge indicator v_i as $\phi_i^{dc} \to 1$. In the experiments, we compute ϕ_i^{dc} as

$$\phi_i^{dc} = exp\left(\varepsilon^{dc}\left[\frac{1}{|\eta_s|}\sum_{j\in\eta_s}\nabla u_i \cdot \nabla u_j - 1\right]\right) \qquad (11)$$

In (11), η_s represents the neighborhood of pixel i having s neighbors and is defined as $\pm s$ pixels along the orthogonal edge direction $\nabla u_i^\perp$. The parameter ε^{dc} is a scalar, which determines the decay rate of the ϕ_i^{dc} function. If the neighboring pixels are coherent (having similar edge directions), then the average angle between ∇u_i and ∇u_j's is close to 0, yielding $\phi_i^{dc} \to 1$. Therefore, when ϕ_i^{dc} is high, occurrence of a low level feature (high gradient) is not accidental and we can rely on our edgeness measure. On the other hand, as ϕ_i^{dc} approaches to 0, an occurrence of the same low level feature is very likely to be accidental and we warp the diffusivity value towards 1.

The importance of directional consistency is best observed if the input image contains impulse noise. Our method can be related to the ones in [23, 27] in the sense that some regularization is performed on the noisy image points. While [23, 27] use hard decisions on noisy image points and smooth out only these points, our method uses a soft decision strategy by means of the local region measure ϕ^{dc} to define the new diffusivity $(cv)^2$ for each image point.

2.2 Edge Continuity: ϕ^h

In boundary detection methods, the principle of edge continuity is used to eliminate streaking or breaking up of an edge contour due to noise or changing contrast. This procedure is commonly referred as *hysteresis* due to successful application of threshold retardation in the Canny edge detector [22]. In our diffusion model, we lower the diffusivity at pixels that correspond to the broken parts of boundary segments to favor edge formation. There may be various choices for the selection of the feedback measure ϕ^h. The important point is to decrease the modified diffusivity $(cv)_i^2$ if the neighboring site encourages formation of an edge, i.e., having a low v value. As shown in Appendix B, there is a reciprocal relationship between v_i and $|\nabla u_i|$ given by $v_i \approx \frac{1}{1+2\alpha\rho|\nabla u_i|^2}$. A decreasing diffusivity can be achieved by increasing the estimate of the image gradient, which is used in estimating the diffusivity. Therefore, a natural choice is to add an offset $h_i \in [0, 1]$ indicating a support in favor of edge formation to the gradient term in the diffusivity

estimate:

$$(cv)_i^2 = \left(\frac{1}{1 + h_i + 2\alpha\rho|\nabla u_i|^2} \right)^2$$
$$= \left(\frac{1}{1 + 2\alpha\rho(|\nabla u_i|^2 + h_i^*)} \right)^2 \qquad (12)$$

We can rewrite

$$(cv)_i^2 = \left(\phi_i^h v_i + \left(1 - \phi_i^h\right) 0 \right)^2 \qquad (13)$$

by letting

$$\phi_i^h = \frac{1}{1 + h_i v_i} \qquad (14)$$

There is a subtle difference between ϕ^{dc} discussed in the previous subsection and ϕ^h. Whereas ϕ^{dc} is high for the non-accidental occurrences of certain low level features e.g., the gradient, ϕ^h is high for the accidental occurrences of the same feature. As ϕ^h approaches to 0, we can rely on a non-accidental occurrence of the feature of interest and warp our diffusivity towards 0. Therefore, the modified edge indicator cv is a linear combination of v, and the maximum edgeness value which is 0.

In the discrete implementation of (6), diffusivities are estimated at mid-grid points. Hence, h_i should be computed as a support from a suitably chosen neighbor. For example, a modified diffusivity $(cv)_{i+0.5,j}^2$ at a mid point between (i, j) and $(i+1, j)$ may receive a support in the form of either $(1 - v_{i+0.5,j-1})$ or $(1 - v_{i+0.5,j+1})$. Notice, that the lower the value of edge indicator at a neighboring site, the higher the support it provides.

Adding spatial organization to energies defining regularization with line process has been previously proposed by Black and Rangarajan [11]. In [13], Black et al. derived the necessary update equations. If we let $\frac{v_k^2 + v_l^2}{2}$ define a line process between site k and site l, then our development becomes equivalent to that of Black et al. Thus, solving new coupled equations are qualitatively equivalent to modifying the MS model with an additive term favoring unbroken contours as in Black and Rangarajan [11] if the parameter h is employed within each iteration.

2.3 Texture Edges: ϕ^{te}

The MS model relies on the image gradient to detect edges. For textured images, a piecewise smoothness assumption fails and large image gradients may result from a texture inside a region rather than from a region boundary. Hence, the MS model may not capture true object boundaries (Fig. 1(e)-(h)).

In order to capture the true object boundaries of a textured image, we need more higher level features than the local derivatives. In the feed-forward step, we estimate the contextual feedback measure ϕ^{te} using the probability map of the texture edges of the source image. For computational concerns, this probability map is extracted only once, at the beginning, using a simple patch-based technique proposed by Wolf et al. [84]. The probability of an image point being near a texture edge

is computed using a non-parametric test of distributions called Wilcoxon Mann-Whitney test.

For each point i, four distributions are sampled: D_i^{up}, D_i^{down}, D_i^{left} and D_i^{right}. These distributions correspond to the contextual similarities between the patch centered at the point i and the patches which are Δx pixels to the related direction. The similarities are simply computed by using the Euclidean distance between the $n \times n$ patches. If the point i lies on near a texture edge, then sampled distributions D_i^{up} and D_i^{down}, or D_i^{left} and D_i^{right} should be different from each other. To test this hypothesis, the Wilcoxon Mann-Whitney test is employed. As in [84], the final probability is taken as the minimum of the p-value's p_i^1 and p_i^2 returned by this test for the distributions D_i^{up}-D_i^{down} and D_i^{left}-D_i^{right}. We use this probability value to define the function ϕ_i^{te} as

$$\phi_i^{te} = exp\left(-\varepsilon^{te}\left(min(p_i^1, p_i^2)\right)\right), \qquad (15)$$

where ε^{te} is again a scalar parameter that determines the decay rate. For the points that do not correspond to the texture edges, the estimated p-values are high, yielding $\phi_i^{te} \to 0$ and $(cv)_i \to 1$.

This case is exactly equivalent to the case discussed under directional consistency. Just like ϕ^{dc}, a low value of ϕ^{te} indicates an accidental occurrence of a feature of interest (again, it is high gradient). Therefore, we should warp our diffusivity estimate towards a maximum value of 1 in order to diffuse more at those locations where

the existence of a local feature is not supported by the context. The modulated diffusivity is given by

$$(cv)_i^2 = \left(\phi_i^{te} v_i + \left(1 - \phi_i^{te}\right) 1\right)^2 \qquad (16)$$

2.4 Local Scale: ϕ^{ls}

The MS model decomposes an image into two components, cartoon and noise. During denoising, some important features of the image like textures or fine details are also smoothed out; since they are treated as noise by the model. This behavior is more apparent in the models with more robust norms, e.g. [66,72]. In [43], a spatially varying fidelity term is proposed for the Rudin-Osher-Fatemi (ROF) model [66] to partly preserve local image structures. The term is based on local variance measures and determines the level of denoising.

Our framework can be also used to devise a texture preserving denoising model by lowering the diffusivity at the pixels inside textured regions. The contextual feedback measure ϕ^{ls} determines whether a point lies on a textured region or not, and controls the level of smoothing accordingly.

We use robust statistics following [12] to determine the textured regions. For an image point i, the local scale is specified by

$$\sigma_i = median_{j \in \Omega_i}(|\nabla u_j - median_{j \in \Omega_i}(|\nabla u_j|)) \qquad (17)$$

where higher values of σ_i means that the point i lies on a textured image region. In deriving σ_i, we use an $n \times n$ pixel patch around the image point i which is represented by Ω_i in (17). Considering σ_i, we define the new contextual feedback measure ϕ_i^{ls} as follows:

$$\phi_i^{ls} = exp\left(-\varepsilon^{ls} \sigma_i\right) \qquad (18)$$

For the pixel inside a textured region, the value of σ_i is high, making $\phi_i^{ls} \to 0$. Just as ϕ_i^h in Section 2.2, ϕ_i^{ls} is low for the non-accidental occurrences of high gradient. When ϕ_i^{ls} is low, we should warp our diffusivity estimate towards minimum level of 0. That is, we compute the modulated edge indicator $(cv)_i$ as the convex combination of 0 and v_i

$$(cv)_i^2 = \left(\phi_i^{ls} v_i + \left(1 - \phi_i^{ls}\right) 0\right)^2 \qquad (19)$$

In the experiments, the contextual feedback measure ϕ^{ls} is estimated at 0th, 10th and 20th iterations.

3 Multi-Channel Images

The straightforward way to smooth a multi-channel image is to diffuse each channel independently. For an m-channel image $g = (g^1, \dots, g^m)$, this yields the following PDEs:

$$\frac{\partial u^i}{\partial t} = \nabla \cdot ((c^i v^i)^2 \nabla u^i) - \frac{\beta}{\alpha}(u^i - g^i); \quad \left.\frac{\partial u^i}{\partial n}\right|_{\partial R} = 0$$
$$(i = 1, \dots, m) \qquad (20)$$

$$\frac{\partial v^i}{\partial t} = \nabla^2 v^i - \frac{2\alpha|\nabla u^i|^2 v^i}{\rho} - \frac{(v^i - 1)}{\rho^2}; \quad \left.\frac{\partial v^i}{\partial n}\right|_{\partial R} = 0$$
$$(i = 1, \dots, m) \qquad (21)$$

Notice that different edge points would be created in different channels. To overcome this inconsistency, we can diffuse each channel with a common edge indicator function as proposed in [61] using the following PDEs [1]:

$$\frac{\partial u^i}{\partial t} = \nabla \cdot ((c^i v)^2 \nabla u^i) - \frac{\beta}{\alpha}(u^i - g^i); \quad \left.\frac{\partial u^i}{\partial n}\right|_{\partial R} = 0$$
$$(i = 1, \dots, m) \qquad (22)$$

$$\frac{\partial v}{\partial t} = \nabla^2 v - \sum_{i=1}^{m}\left(\frac{2\alpha|\nabla u^i|^2 v}{\rho}\right) - \frac{(v - 1)}{\rho^2};$$
$$\left.\frac{\partial v}{\partial n}\right|_{\partial R} = 0 \qquad (23)$$

Contextual feedback measures can be estimated in various ways. First, each ϕ^i and consequentially $(cv)^i$ may be computed directly from the ith channel of the image. Second, a common feedback measure may be estimated. One possibility is to estimate a common measure from a weighted average of the multi-channel data. For color images, the average readily corresponds to the intensity image. Another possibility in estimating a common feedback measure is to consider individual channels to form a statistical measure. When the latter strategy is adopted in the experiments, the median of the individually computed measures is used for the directional consistency and the local scale. On the other hand, for

---

[1] An alternative color image segmentation method was also proposed in [16] that considers images as manifolds within the AT regularization framework.

the texture edges, the measure is estimated by considering the distribution of contextual similarities computed from vector-valued image patches.

4 Experimental Results

In all experiments, discrete approximations of the coupled equations are used (Appendix A (24) and (25)). The homogeneous Neumann boundary condition is assumed along the image boundary. The results are obtained by using a convergence tolerance of $\epsilon = 5 \times 10^{-4}$ unless stated otherwise.

The importance of the directional consistency of the edges is best observed if the image contains impulse noise. The postprocessing of the noisy image shown in Fig. 3(a) is illustrated in Fig. 3(b)-(f). Fig. 3(b) and (c) illustrate smoothing obtained using the AT with different values of smoothing radius, $\frac{\alpha}{\beta}$. The result in Fig. 3(b) is obtained with $\alpha = 0.75$, $\beta = 0.005$, $\rho = 0.01$. When we increase the smoothing radius by choosing $\beta = 0.001$, diffusion is so strong that we even lose the upper body part of the woman (Fig. 3(c)). Yet, the noise is still present. If we use a regularization term which uses contextual feedback by means of directional consistency of the edges, as discussed in Section 2.1, the image is denoised without blurring (Fig. 3(d) and (e)). The perceptual difference between Fig. 3(d) and (e) is in the sharpness level. The result in Fig. 3(d) is obtained with the segmentation parameters specified $\alpha = 1, \beta = 0.01, \rho = 0.01$ and the contextual feedback parameters $s = 2$ and

$\varepsilon^{dc} = 0.25$ (these are the default values for the directional consistency parameters). For the result given in Fig. 3(e), we use the same parameters except $\varepsilon^{dc} = 0.1$. The variable ε^{dc} determines the decay rate of the coherency function used in the segmentation process and therefore specifies the level of sharpness. For large ε^{dc} value, the decay rate is high and the edges are more smoothed out depending on the contextual feedback. Hence, as observed, the resulting image is smoother. On the other hand, for small ε^{dc} values, we get sharper results. The result given in Fig. 3(e) is also comparable to that is obtained by Shah's modification [72] (Fig. 3(f)), which uses the L^1-norm for both the data fidelity and the regularizer (for the experiments a half-quadratic approximation of Shah's modified energy proposed in [48] is used). While both results are satisfactory, the reconstruction using the model due to Shah is blockier. The used robust norms attract the image towards the cartoon limit and catches unintuitive regions such as the one at the man's right shoulder and the ones on the floor. Better result can be obtained by *merely* replacing the data fidelity term in the AT model with the L^1-norm as in [17,58].

Increasing the value of α while keeping $\frac{\alpha}{\beta}$ fixed means decreasing the penalty of the length term, yielding more detailed reconstruction. In Fig. 4, the proposed modification is tested with again the image given in Fig. 3(a), however, forcing the reconstruction to be more detailed by the proper choice of parameters. Fig. 4(a) demon-
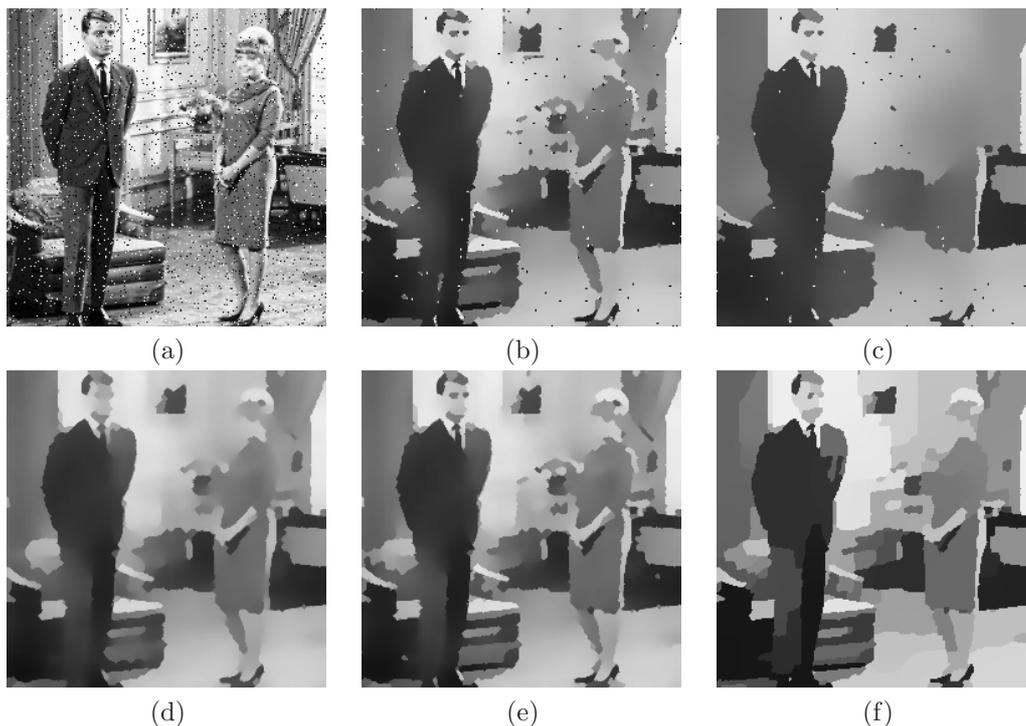
Fig. 3 Considering directional consistency of the edges ϕ^{dc} as the contextual feedback measure eliminates impulse noise. (a) Source image corrupted with 5% salt and pepper noise. (b)-(c) Reconstructions using AT with two different smoothing radii. Notice that the noise is still present even when we lose the upper body part of the woman. (d)-(e) Reconstructions with directional consistency with two different sharpness levels. Notice that at comparable scales noise is completely eliminated. (f) L^1-reconstruction catches unintuitive regions.
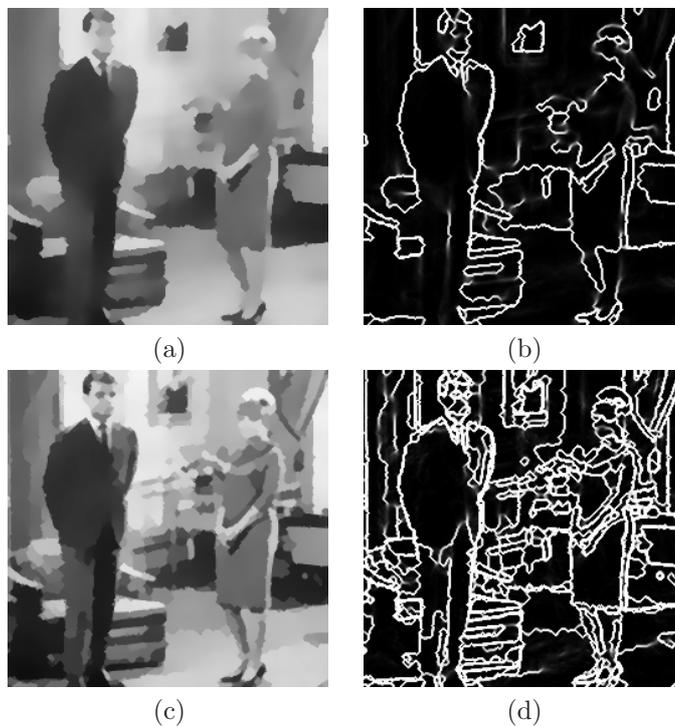
Fig. 4 u and $1-v$ functions computed with $\alpha = 1$, $\beta = 0.01$ and $\alpha = 4$, $\beta = 0.04$ respectively. Even in detailed reconstruction, directional consistency feedback is sufficient to remove noise completely.

strates the outcome of the proposed modification presented in the first experiment. The result displayed in Fig. 4(c) is obtained by using the same parameters except $\alpha = 4$, $\beta = 0.04$. The corresponding edge indicator functions are also shown in Fig. 4(b) and (d) respectively. As they demonstrate, even the detailed reconstruction with $\alpha = 4$ is noise free.

The example presented in Fig. 5 illustrates the effect of edge continuity as described in Section 2.2. The reconstructions of the source image shown in Fig. 5(a) , which is taken from [13], are presented in Fig. 5(b) and (c) together with the corresponding edge indicator functions. Fig. 5(b) illustrates the outcome of AT whereas Fig. 5(c) illustrates the result obtained by considering edge continuity. Both results are obtained by using the same scale space parameters, $\alpha = 1$, $\beta = 0.01$, $\rho = 0.01$. As it can be clearly seen from the enlarged indicator functions given in Fig. 5(d), the modified scheme eliminates broken contours.

In Fig. 6, we demonstrate the results obtained with a regularization considering the coalition of edge continuity and directional consistency of the edges as the contextual feedback measure via the product of individual c functions. Recall that a combination can be achieved by multiplying individual c values. In the case of directional consistency, the multiplier c^{dc} is given by (9). In the case of edge continuity, the multiplier c^h is equal to ϕ^h as indicated by (13). The reconstruction results of the source image corrupted with 10% salt and pep-

per noise (Fig. 6(a)) are given in Fig. 6(b)-(d). They are obtained using the scale space parameters specified in the previous experiment. Fig. 6(b) is the result obtained with edge continuity. As it can be clearly seen, the noise is not eliminated. Fig. 6(c) is obtained with the modification which considers directional consistency. Finally, Fig. 6(d) is the outcome of the combined framework that considers the coalition of edge continuity and directional consistency of the edges, which is not only noise free but also has stronger edges.

In Fig. 7, the framework considering the coalition of directional consistency of the edges and edge continuity is tested on a noisier image (Fig. 7(a)). Fig. 7(b)-(c) are the outcomes of the AT approximation considering $\alpha = 2.5$, $\beta = 0.0001$, $\rho = 0.01$. Fig. 7(b) is obtained with a convergence tolerance of $\epsilon = 4.8 \times 10^{-4}$ whereas Fig. 7(c) is obtained with $\epsilon = 1.6 \times 10^{-4}$. The effect of decreasing β dramatically, is to shut off the data fidelity term and to increase the relative importance of the length term. Notice that the relative increase (α/β) in the weight of the second term of the AT model is less than the relative increase $(1/\beta)$ in the weight of the third term. Our goal is to compare the effect of the length term –which is a part of the original model– to the effect of ϕ^{dc}. As it can be seen in Fig. 7(c), eliminating the noise entirely results smoothing out the whole rectangular region. This is mainly due to the perceived difference in the contrast. Fig. 7(d)-(f) are the outcomes of the proposed framework that are produced by using
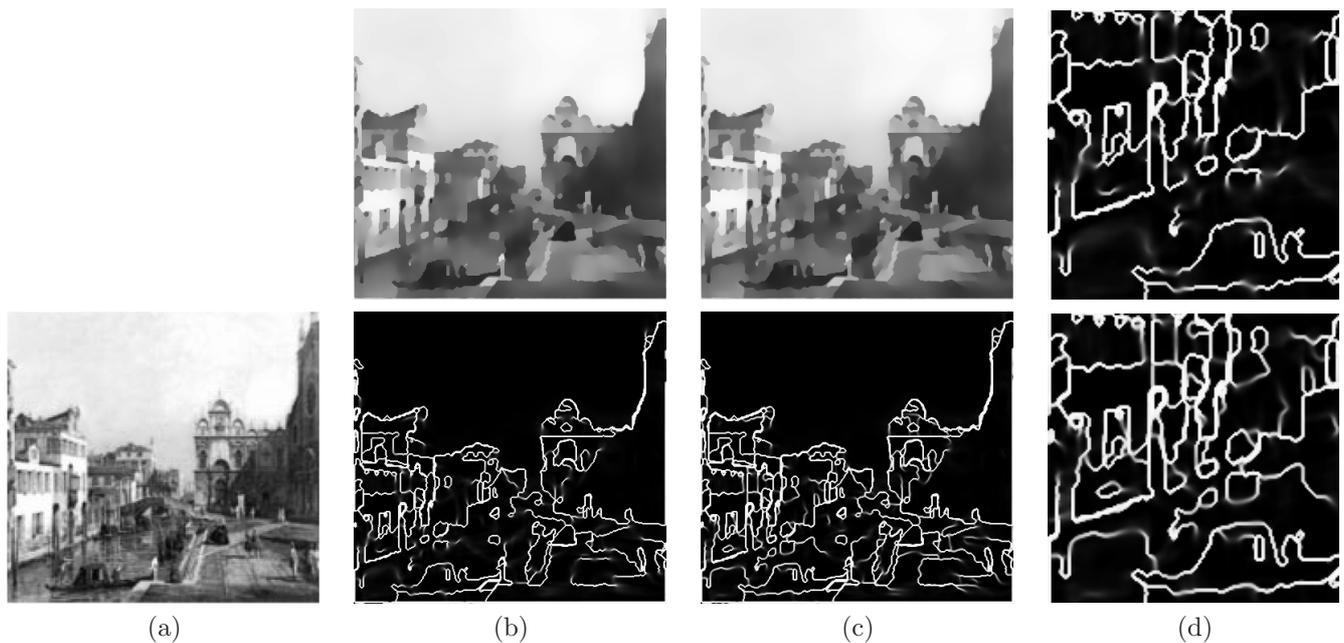
Fig. 5 Considering edge continuity ϕ^h as the contextual feedback measure eliminates broken contours. (a) Source image. (b) Reconstruction using AT (u and $1 - v$). (c) Reconstruction with the new method forcing edge continuity (u and $1 - v$) (d) Details from the edge indicator functions given in (b) and (c) respectively.
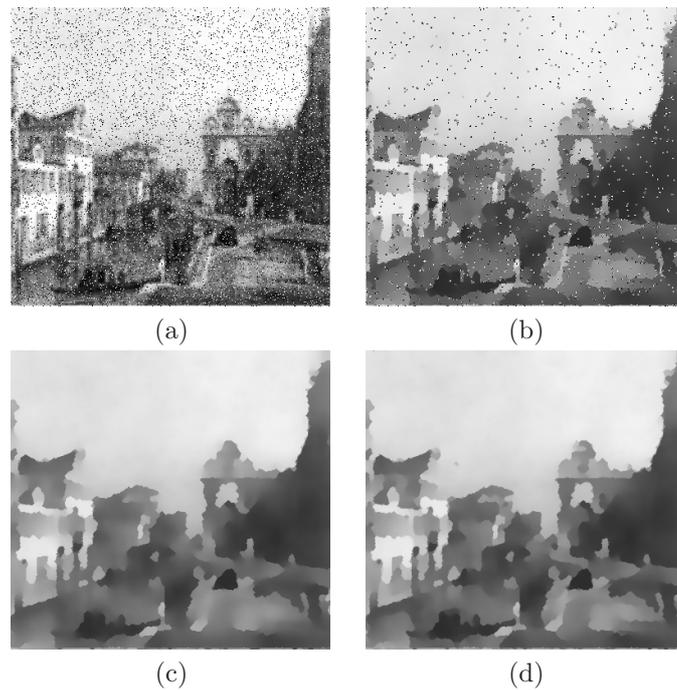

Fig. 6 Considering a framework that uses the coalition of ϕ^{dc} and ϕ^h as the contextual feedback measure eliminates both noise and the broken contours. (a) Source image corrupted with 10% salt and pepper noise. (b) Reconstruction with edge continuity. (c) Reconstruction with directional consistency of the edges. (d) Reconstruction using both edge continuity and directional consistency of the edges.
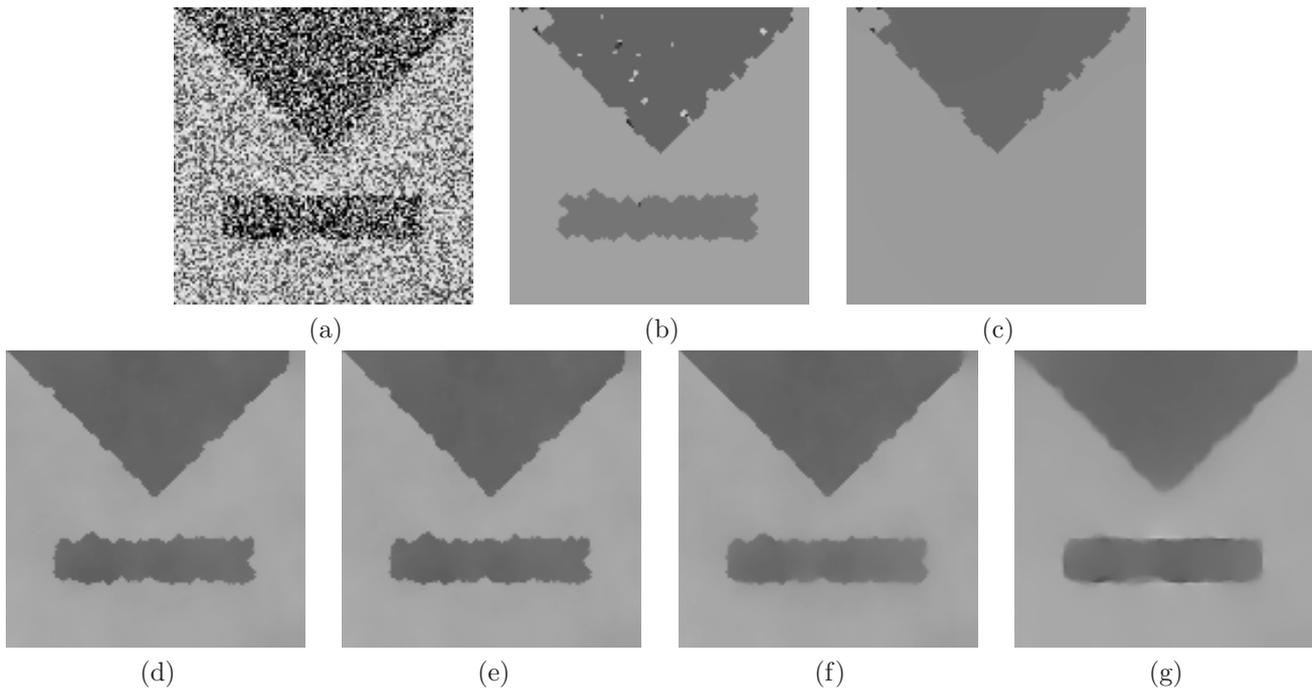
Fig. 7 A difficult denoising case. (a) Source image, 70% of the pixels are degraded with uniform noise. (b)-(c) Reconstructions using AT with different choices of convergence tolerance. (d) Reconstruction considering directional consistency. (e) Reconstruction using the coalition of ϕ^{dc} and ϕ^h. (f) Reconstruction considering directional consistency with a larger neighborhood size. (g) Reconstruction using the ROF model.

the scale space parameters $\alpha = 1$, $\beta = 0.01$, $\rho = 0.01$. Fig. 7(d) is obtained by using directional consistency with $\epsilon = 2.5 \times 10^{-4}$. Fig. 7(e) is the outcome of the coalition of ϕ^{dc} and ϕ^h with $\epsilon = 1.25 \times 10^{-4}$. Notice that the results given in Fig. 7(d) and (e) are visually similar. This is due to the fact that the contrast is almost constant in the image. Hence, broken lines do not occur. Fig. 7(f) is obtained by using directional consistency with a larger neighborhood size, $s = 10$, and a convergence tolerance of $\epsilon = 1.6 \times 10^{-4}$. The perceptual difference between Fig. 7(d) and (f) is clearly evident. Imposing coherency in a large neighborhood produces smoother object boundaries, however, at the expense of losing sharpness. Both reconstructions are qualitatively comparable to the one obtained by the ROF model[66]

(Fig. 7(g)). Note that the control parameters of ROF were tuned to eliminate the noise and to capture the shape boundaries.

Fig. 8-10 demonstrate the use of texture edges measure ϕ^{te} described in Section 2.3 for smoothing textured images. We observe that this measure, if used together with the other feedback measures, captures the actual object boundaries. Note that in our formulation we do not decompose the original image into structure, texture and noise components as in [6,81], we only retain the structure. We do so without using complicated norms.

Fig. 8 and Fig. 9 includes two texture mosaic images [46] as the source images. Fig. 8(b) and Fig. 9(b) are the outcomes of our proposed method. Fig. 8(b) is obtained by considering ϕ^{te} as the only contextual feedback
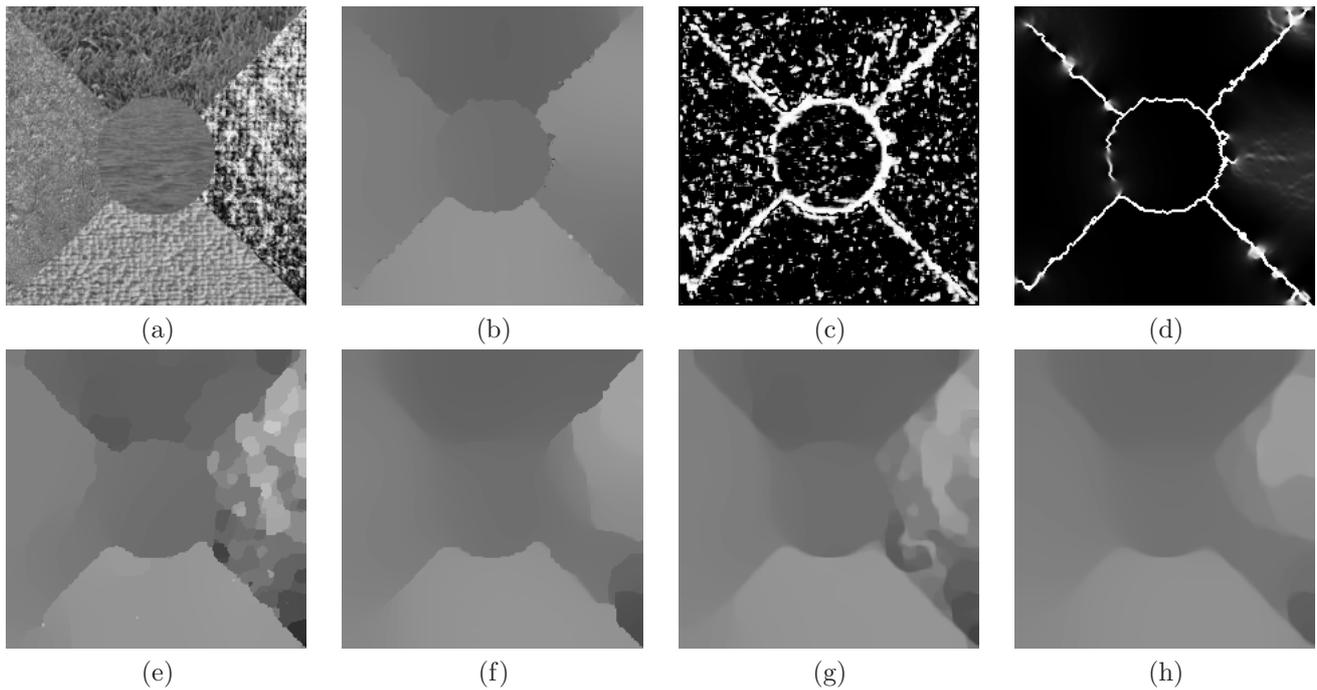
Fig. 8 Actual region boundaries are captured by considering ϕ^{te} as the feedback measure. (a) Source image. (b) Reconstruction result with contextual feedback. (c) Texture edges measure ϕ^{te}. (d) Final edge indicator function. (e)-(f) Reconstructions using Shah's modification with different scale parameters. (g)-(h) Reconstructions using the ROF model with different scale parameters.
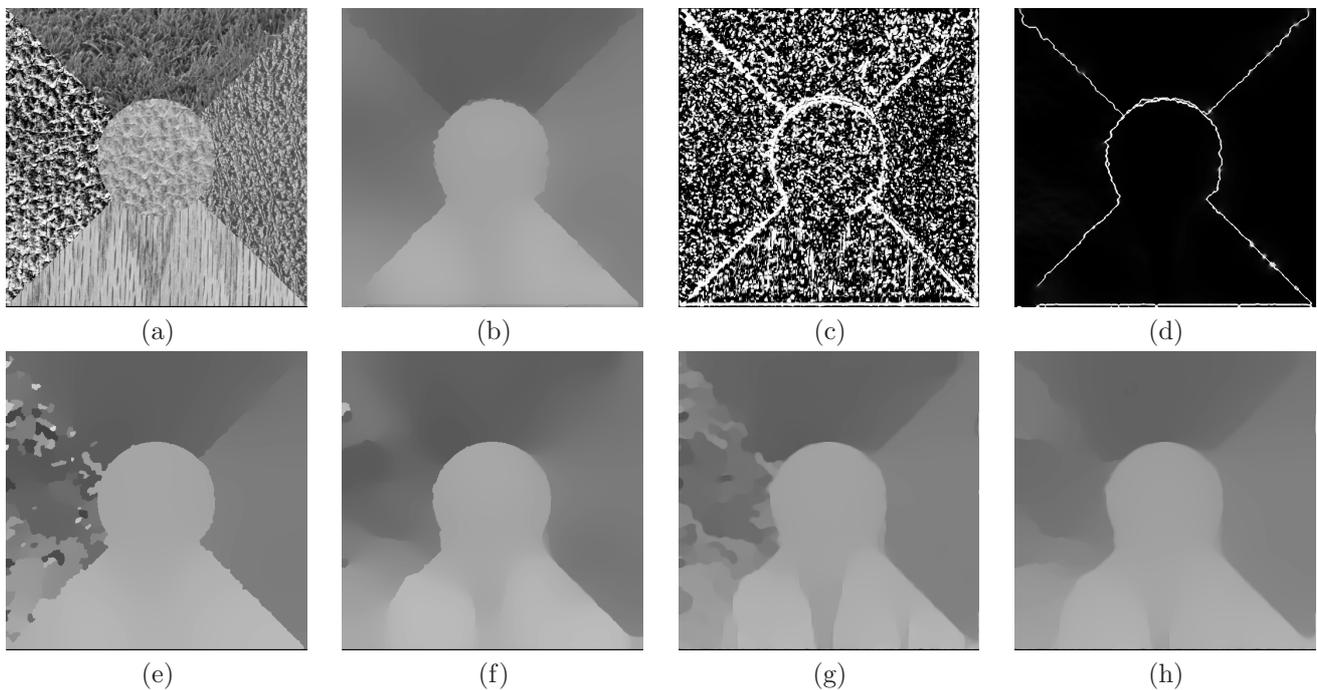

Fig. 9 Coalition of three feedback measures. (a) Source image. (b) Reconstruction result with contextual feedback. (c) Textures edges measure ϕ^{te}. (d) Final edge indicator function. (e)-(f) Reconstructions using Shah's modification with different scale parameters. (g)-(h) Reconstructions using the ROF model with different scale parameters.
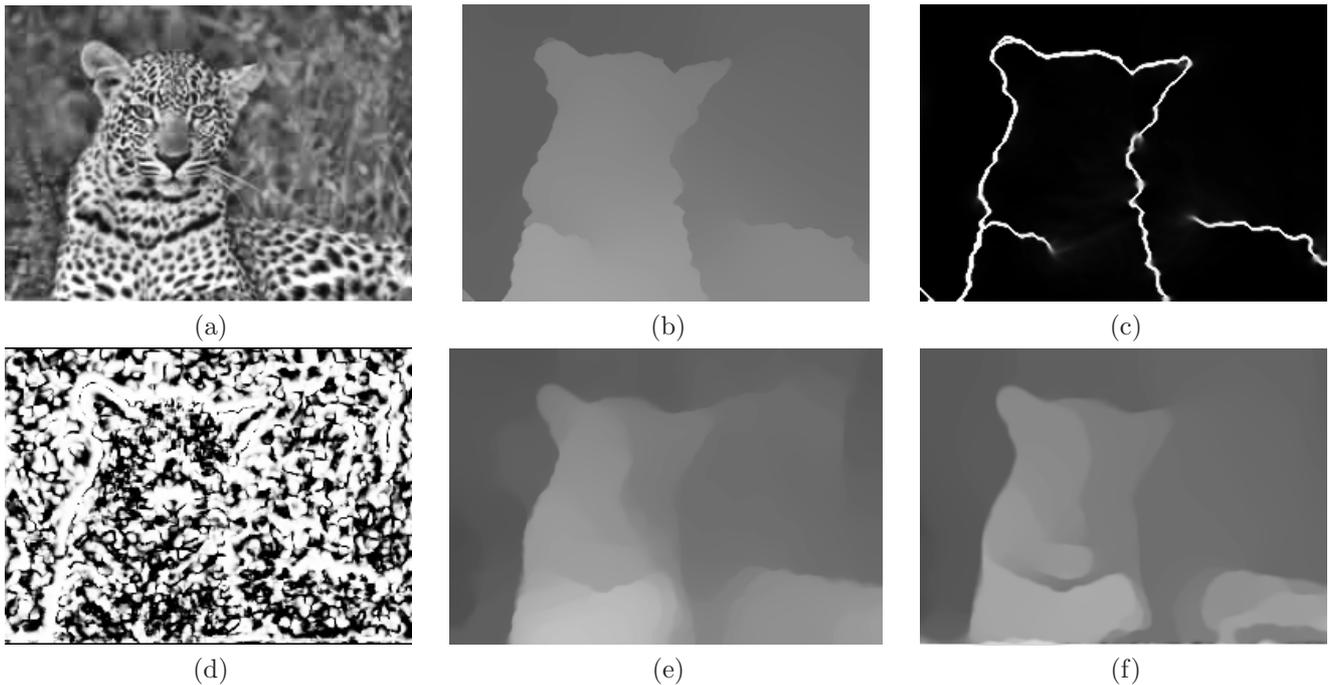
Fig. 10 Coalition of three feedback measures. (a) Source image. (b) Reconstruction result with contextual feedback (c) Final edge indicator function. (d) Texture edges measure ϕ^{te}. (e) Reconstruction using Shah's modification. (f) Reconstruction using the ROF model.

measure with the segmentation parameters $\alpha = 100$, $\beta = 0.0075$, $\rho = 0.01$ and feedback parameters $n = 5$, $\Delta x = 15$ and $\varepsilon^{te} = 1000$ (these are the default values for the texture edges). On the other hand, Fig. 9(b) is obtained by considering ϕ^{te} together with ϕ^{dc} and ϕ^{h} measures by using the same parameters except $\alpha = 300$, $\beta = 0.005$, $\rho = 0.001$ and $\varepsilon^{te} = 100$. For each texture mosaic image, the texture edges measure and the resulting edge indicator function are also given. As it can be clearly seen in Fig. 8(c)-(d) and Fig. 9(c)-(d), while the estimated texture edges measures computed using the method proposed in [84] are noisy, the resulting edge indicator functions succeed to capture the actual object boundaries. The results of Shah's modification and the ROF model

with different scale parameters are also provided in Fig. 8 and Fig. 9 for comparison.

In Fig. 10, the framework considering the coalition of directional consistency, edge continuity and texture edges is tested with a natural animal image. The source image Fig. 10(a) is taken from [40]. Fig. 10(b) is the outcome of the combined framework with the segmentation parameters $\alpha = 200$, $\beta = 0.05$, $\rho = 0.001$ and the default contextual feedback parameters except $\varepsilon^{te} = 20$.

As discussed in Section 2.4, the MS model does not distinguish textures and fine details from noise. Hence, they are smoothed out during denoising. Fig. 11 depicts how contextual feedback measure based on local scale affects the reconstruction results. For the source image given in Fig. 11(a), the outcomes of the AT approxi-

Fig. 11 Texture preserving denoising with local scale measure ϕ^{ls}. (a) Source image. (b)-(c) Reconstructions using AT with different choices of smoothing levels. (d) Reconstruction result with contextual feedback. Notice that texture in the fabric is preserved. (e) Local scale measure ϕ^{ls}.

mation are presented in Fig. 11(b)-(c). Fig. 11(b) is obtained by using the parameters $\alpha = 10$, $\beta = 2.5$, $\rho = 0.001$ with a convergence tolerance of $\epsilon = 1 \times 10^{-4}$ while Fig. 11(c) is obtained using $\alpha = 15$, $\beta = 1$, $\rho = 0.001$. Fig. 11(d) is the outcome of the proposed method that is estimated using $\alpha = 10$, $\beta = 0.1$, $\rho = 0.001$, $\varepsilon^{ls} = 0.25$, $n = 25$ and the local scale measure depicted in Fig. 11(e) as the contextual feedback measure. The L^2-norms of the differences between the original image and the reconstructions presented in Fig. 11(b)-(d) are 6.0954×10^3, 7.6363×10^3 and 6.0735×10^3, respectively. Although the amount of signal eliminated from the source image in Fig. 11(b) and (d) is approximately the same and the

reconstructions are in a visually similar scale, Fig. 11(b) is not entirely smooth and contains noisy pixels whereas the source image is denoised while preserving textures in Fig. 11(d). Increasing the level of smoothing in the AT model results noise-free results as presented in Fig. 11(c), however the textured regions are also smoothed out during the process. Fig. 12 illustrates the results of two more texture preserving denoising experiments. Fig. 12(c) and (d) are obtained using the parameters $\alpha = 20$, $\beta = 0.1$, $\rho = 0.001$, $\varepsilon^{ls} = 0.125$, $n = 15$ and $\alpha = 4$, $\beta = 0.1$, $\rho = 0.001$, $\varepsilon^{ls} = 0.125$, $n = 25$ respectively. Fig. 12(b) is taken from *Berkeley Segmentation Dataset* [54]. Note that the basic idea is just to lower the diffusion at tex-

tured locations. The measure ϕ^{ls} does not make a distinction between noise and texture.

Finally, we apply our method to color images as described in Section 3. We use the RGB channels as our multi-channel data. For Fig. 13, we have repeated the texture preserving denoising experiment using now colored versions of Fig. 11(a) and Fig. 12(b). The denoising results presented in Fig. 13(b) are obtained by diffusing each channel separately with a common feedback measure estimated from the intensity image.

Fig. 14 illustrates the results of using different strategies for color image smoothing of the source image presented in Fig. 14(a) (image taken from *Berkeley Segmentation Dataset* [54]). All smoothing results given in Fig. 14(b)-(d) are obtained by considering the coalition of directional consistency, edge continuity and texture edges with the segmentation parameters $\alpha = 100$, $\beta = 0.1$, $\rho = 0.001$ and the default contextual feedback parameters except $\varepsilon_{dc} = 0.05$. Fig. 14(b) is the result obtained by diffusing each channel separately by using the feedback measures estimated from corresponding multi-channel data. Fig. 14(c) is obtained by again diffusing each channel separately but with a common feedback measure estimated from the intensity image. Finally, Fig. 14(d) is the outcome of the color image smoothing with a common edge strength function and a feedback measure estimated from intensity image.

In Fig. 15, we demonstrate a case where using contextual feedback measures estimated from the intensity

image yields an inaccurate smoothing result. The reconstructions of the color image shown in Fig. 15(a), which is reproduced from [26], are given in Fig. 15(c) and (e) together with the corresponding edge indicator functions provided in Fig. 15(d) and (f). All smoothing results are obtained by considering the coalition of directional consistency, edge continuity and texture edges with the segmentation parameters $\alpha = 100$, $\beta = 0.1$, $\rho = 0.001$ and the default contextual feedback parameters except $\varepsilon^{dc} = 0.05$. Fig. 15(c) is the outcome of the color image smoothing with a common edge strength function and a feedback measure estimated from intensity image. Since transforming the color image into intensity image makes the upper and the lower objects disappear (Fig. 15(b)), these objects are smoothed out during processing and the reconstruction fails to capture the actual object boundaries. On the other hand, when the feedback measures are estimated by considering each channel simultaneously, we get a fairly good result (Fig. 15(e)).

Fig. 16 shows some typical smoothing results of natural color images obtained with the proposed method. In all the cases, textured regions are smoothed out and the structures are preserved.

5 Summary and Discussion

Diffusivity modulated by the context is the key mechanism behind the proposed method. The modulated diffusivity $(cv)^2$ is a square of a convex combination of the edge indicator v and a fixed value which is either 1
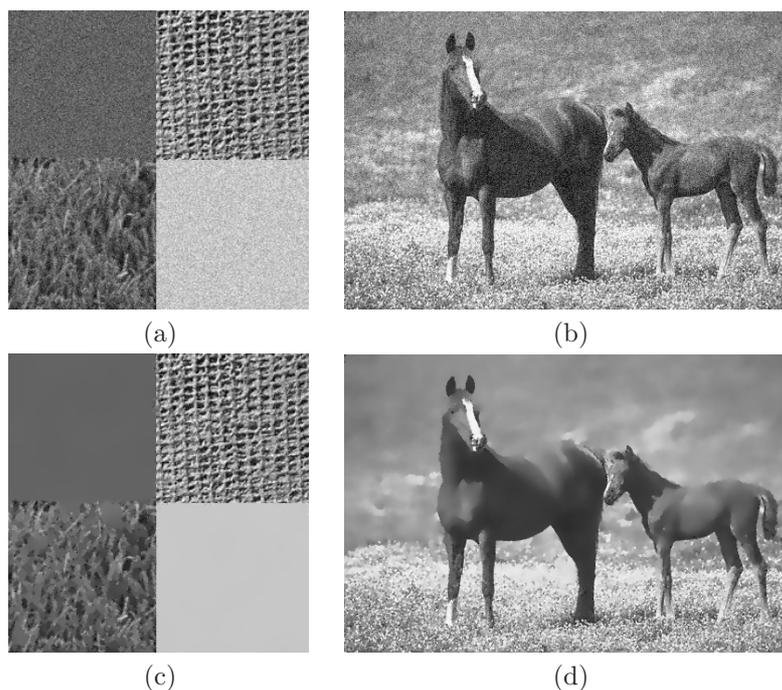
Fig. 12 Some more experiments on texture preserving denoising. Notice the difference between non-textured and textured regions. (a)-(b) Source images. (c)-(d) Smoothed images.
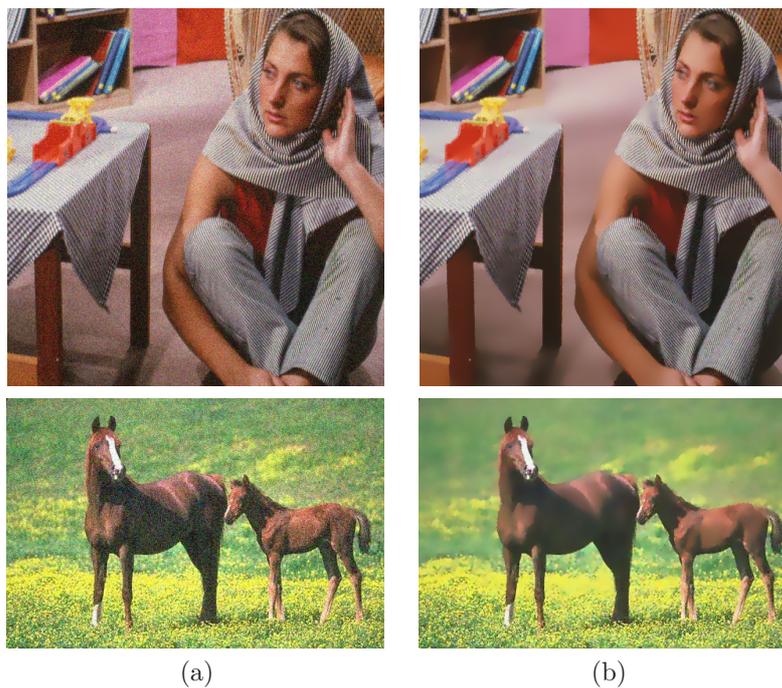

Fig. 13 Texture preserving denoising on a color image. Notice the difference between non-textured and textured regions. (a) Source images. (b) Smoothed images.

(negative feedback) and 0 (positive feedback). Four different possibilities for contextual feedback measure are presented, in order to perform several seemingly different tasks: elimination of unintuitive edges due to noise or texture, preserving texture, catching texture boundaries, and enhancing weak edges. Experimental results
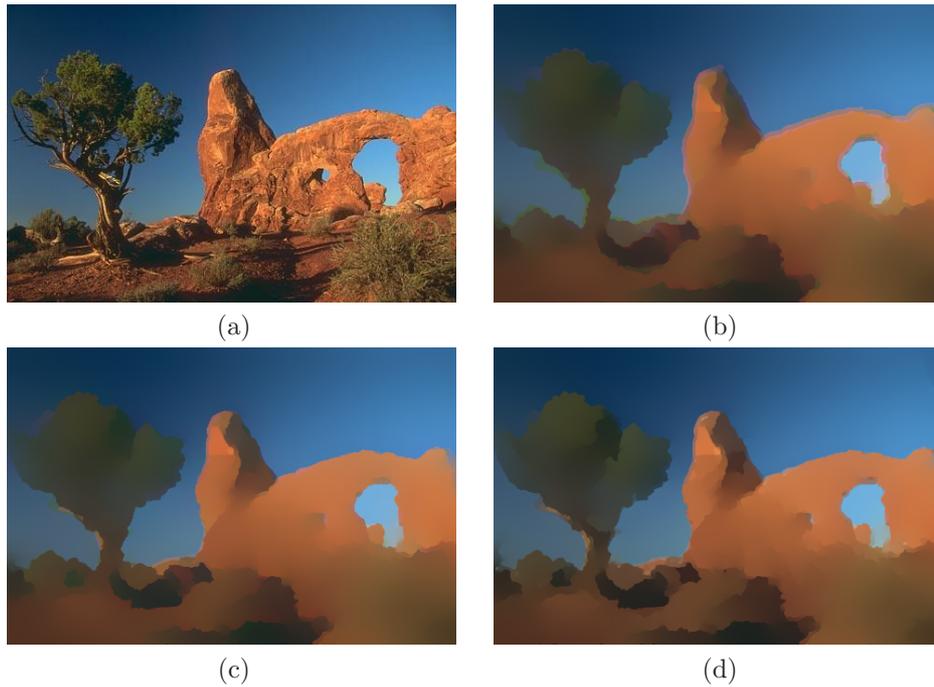
(a)   (b)

(c)   (d)

Fig. 14 Color image smoothing. (a) Source Image. (b) Result of channel-by-channel smoothing. (c) Result of channel-by-channel smoothing with a common feedback measure estimated from intensity image. (d) Color image smoothing with a common edge strength function and a feedback measure estimated from intensity image.
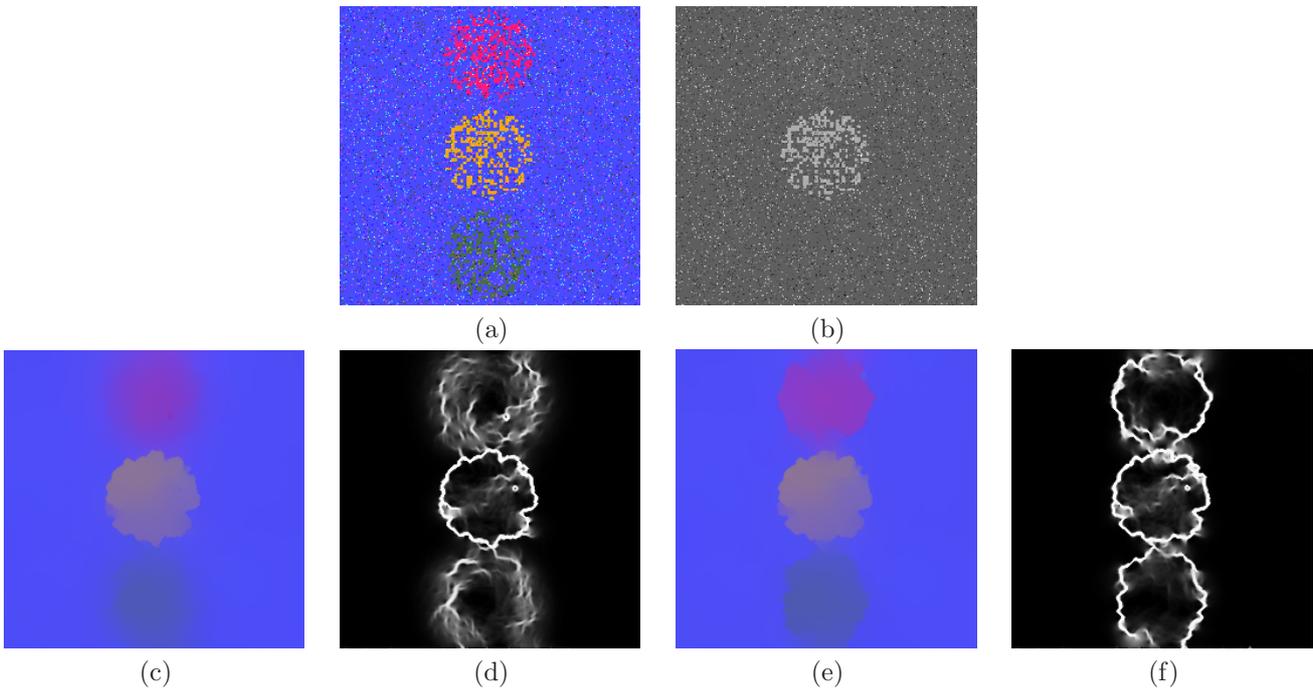

(a)   (b)

(c)   (d)   (e)   (f)

Fig. 15 Coalition of measures in color image smoothing. (a) Source Image. (b) Corresponding intensity image. (c)-(d) Reconstruction result of color image smoothing with a common edge strength function and a feedback measure estimated from the intensity image (u and 1-v). (e)-(f) Reconstruction result of color image smoothing with a common edge strength function and a feedback measure estimated from all channels (u and 1-v).
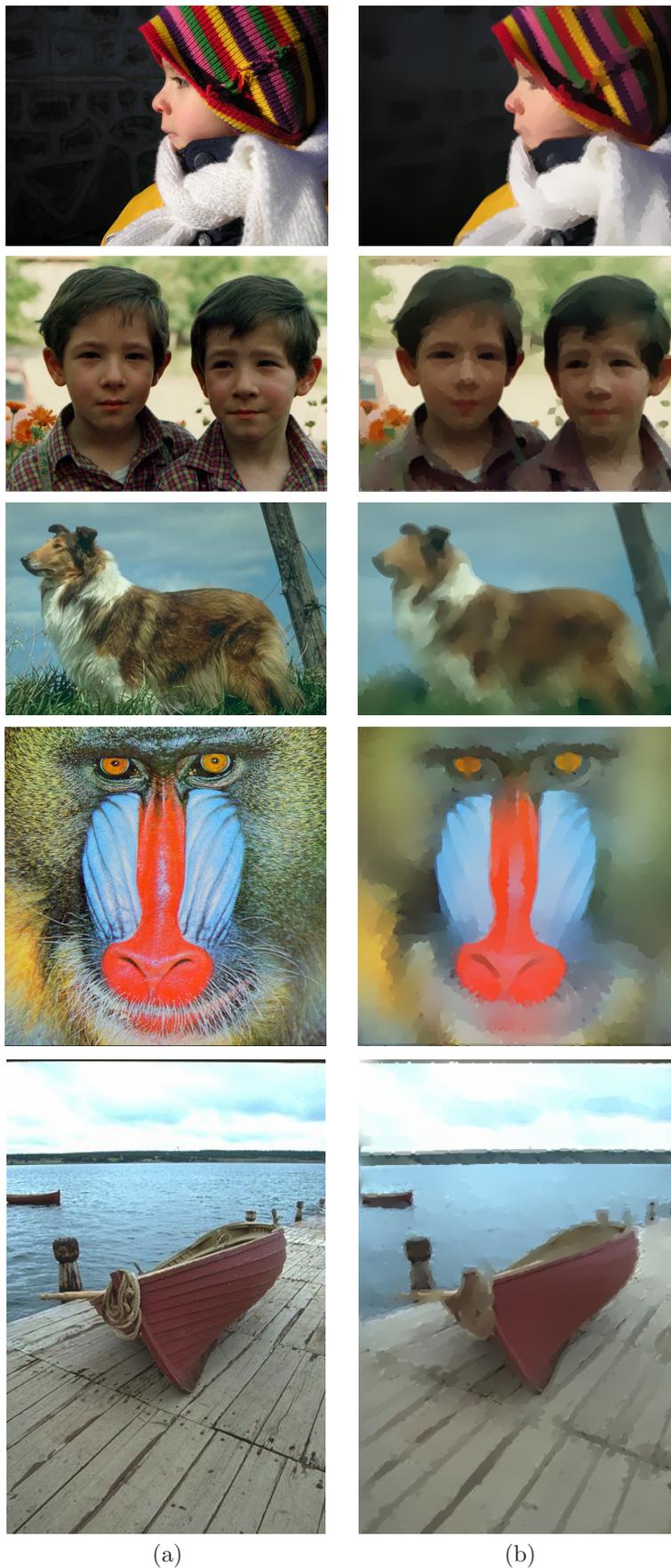
Fig. 16 Smoothing of natural color images. (a) Source images *Devin*, *twins*, *dog*, *baboon* and *boat*. (b) Smoothing results.

demonstrate that the proposed regularization is able to remove difficult noise types, capture texture boundaries and produce almost segmentation like results without using singular diffusivities that arose from non-quadratic cost functions.

In the same computational frame, the measures can be combined and the list can be extended. As an example, continuous measures by Kokkinos, Evangelopoulos and Maragos [49] that allow probabilistic discrimination between edges, textured and smooth areas can be directly utilized in our framework. Moreover these type of measures can be used as weights in combining the effects of various feedback measures.

In the proposed approach, both negative and positive feedback are implemented as regularization. This may have the disadvantage that if a feature does not exist in the data (e.g. illusory contour), it will not emerge. There may be two solutions. Firstly, an occasional random noise may be added in order to generate spurious edges. If these edge hypothesis do not get sufficient support from the context, they will be eliminated during the regularization process. Secondly, the information rising up in the hierarchy may get connected to a database object whose shape is known and impose top-down constraints to the regularization process. Both solutions are currently being investigated by the authors.

Appendix A

We first discretize the coupled system w.r.t. spatial variables. The discrete approximation of (6) and (4) used in our computations are as follows:

$$\frac{d}{dt}u_{i,j}(t) = \nabla \cdot (d_{i,j}(t)\nabla u_{i,j}(t))$$
$$- \frac{\beta}{\alpha}(u_{i,j}(t) - g_{i,j})$$
$$\frac{d}{dt}v_{i,j}(t) = \nabla^2 v_{i,j}(t) - \frac{2\alpha|\nabla u_{i,j}(t)|^2 v_{i,j}(t)}{\rho}$$
$$- \frac{(v_{i,j}(t) - 1)}{\rho^2}$$

where $d_{i,j}(t) = (v_{i,j}(t))^2$ for AT and $d_{i,j}(t) = ((cv)_{i,j}(t))^2$ for the proposed method.

We discretize the time derivative using forward difference. Regularization terms on the right hand side of each equation are evaluated at time t. Bias terms in the right hand side of each equation are evaluated at time $t+1$. Hence the scheme is neither fully explicit nor fully implicit. However, as in an explicit scheme, we can evaluate u_{t+1} and v_{t+1} explicitly using forward recursion.

Let h be the space step, Δt be the time step, and $(x_i, y_i) = (ih, jh)$ be the space coordinates. The finite differences are

$$\Delta^x_+ u_{i,j} = u_{i+1,j} - u_{i,j}, \qquad \Delta^x_- u_{i,j} = u_{i,j} - u_{i-1,j}$$
$$\Delta^y_+ u_{i,j} = u_{i,j+1} - u_{i,j}, \qquad \Delta^y_- u_{i,j} = u_{i,j} - u_{i,j-1}$$

The values at mid points are estimated using the averages

$$d_{i\pm\frac{1}{2},j} = \frac{d_{i\pm1,j} + d_{i,j}}{2}, \qquad d_{i,j\pm\frac{1}{2}} = \frac{d_{i,j\pm1} + d_{i,j}}{2}$$

$$\frac{u_{i,j}^{t+1} - u_{i,j}^{t}}{\Delta t} = \frac{1}{h^2}\left(d_{i+\frac{1}{2},j}^{t}\Delta_+^x u_{i,j}^t - d_{i-\frac{1}{2},j}^{t}\Delta_-^x u_{i,j}^t\right.$$
$$\left. + \ d_{i,j+\frac{1}{2}}^{t}\Delta_+^y u_{i,j}^t - d_{i,j-\frac{1}{2}}^{n}\Delta_-^y u_{i,j}^t\right)$$
$$- \ \frac{\beta}{\alpha}\left(u_{i,j}^{t+1} - g_{i,j}\right)$$

$$\frac{v_{i,j}^{t+1} - v_{i,j}^{t}}{\Delta t} = \frac{1}{h^2}\left(\Delta_+^x v_{i,j}^t - \Delta_-^x v_{i,j}^t + \Delta_+^y v_{i,j}^t - \Delta_-^y v_{i,j}^t\right)$$
$$- \ \frac{2\alpha}{\rho}\left[\frac{(u_{i,j+1}^{t+1} - u_{i,j-1}^{t+1}) + (u_{i+1,j}^{t+1} - u_{i-1,j}^{t+1})}{(2h)^2}\right]v_{i,j}^{t+1}$$
$$- \ \frac{(v_{i,j}^{t+1} - 1)}{\rho^2}$$

$$u_{i,j}^{t+1} = \frac{1}{1+\frac{\Delta t\beta}{\alpha}}\left[u_{i,j}^t + \frac{\Delta t\beta}{\alpha}g_{i,j}\right.$$
$$+ \ \frac{\Delta t}{h^2}\left(d_{i+\frac{1}{2},j}^{t}\Delta_+^x u_{i,j}^t - d_{i-\frac{1}{2},j}^{t}\Delta_-^x u_{i,j}^t\right.$$
$$\left.\left. + \ d_{i,j+\frac{1}{2}}^{t}\Delta_+^y u_{i,j}^t - d_{i,j-\frac{1}{2}}^{n}\Delta_-^y u_{i,j}^t\right)\right] \quad (24)$$

$$v_{i,j}^{t+1} = \left[1 + \Delta t\left(\frac{2\alpha\rho A + 1}{\rho^2}\right)\right]^{-1}\cdot$$
$$\left[v_{i,j}^t + \frac{\Delta t}{\rho^2}\right.$$
$$\left. + \frac{\Delta t}{h^2}\left(\Delta_+^x v_{i,j}^t - \Delta_-^x v_{i,j}^t + \Delta_+^y v_{i,j}^t - \Delta_-^y v_{i,j}^t\right)\right]$$
$$(25)$$

where $A = \left[\dfrac{(u_{i,j+1}^{t+1} - u_{i,j-1}^{t+1})^2 + (u_{i+1,j}^{t+1} - u_{i-1,j}^{t+1})^2}{(2h)^2}\right].$

Iterations stop if $|u^{t+1} - u^t| < \epsilon|u^t|$.

Appendix B

Keeping u fixed, (4) minimizes a convex quadratic functional given by

$$\int_R \left(\rho|\nabla v|^2 + \frac{1 + 2\alpha\rho|\nabla u|^2}{\rho}\left(\frac{1}{1 + 2\alpha\rho|\nabla u|^2} - v\right)^2\right)dx$$

The reciprocal relationship between v and $|\nabla u|^2$ can be best observed in the above functional. Clearly, it asserts that the function v is nothing but a smoothing of $\frac{1}{1 + 2\alpha\rho|\nabla u|^2}$ with a blurring radius proportional to ρ and reciprocal to $|\nabla u|$. Ignoring the smoothing, by letting $\rho \rightarrow 0$ [25,72],

$$v \approx \frac{1}{1 + 2\alpha\rho|\nabla u|^2} \qquad (26)$$

References

1. Symposium on the Role of Context in Recognition, European Conference on Visual Perception, August 2005.

2. M. Aleman-Flores, L. Alvarez, and V. Caselles. Texture-oriented anisotropic filtering and geodesic active contours in breast tumor ultrasound segmentation. *J. Math. Imaging Vis.*, 2007.

3. R. Alicandro, A. Braides, and J. Shah. Free-discontinuity problems via functionals involving the L^1-norm of the

gradient and their approximation. *Interfaces and Free Boundaries*, 1(1):17–37, 1999.

4. L. Ambrosio and V. Tortorelli. On the approximation of functionals depending on jumps by elliptic functionals via Γ-convergence. *Commun. Pure Appl. Math.*, 43(8):999–1036, 1990.

5. C. Aslan and S. Tari. An axis-based representation for recognition. In *ICCV*, volume 2, pages 1339–1346, 2005.

6. J.-F. Aujol, G. Aubert, L. B.-F., and A. Chambolle. Image decomposition into a bounded variation component and an oscillating component. *J. Math. Imaging Vis.*, 22(1):71–88, 2005.

7. L. Bar, N. Kiryati, and N. Sochen. Image deblurring in the presence of impulsive noise. *Int. J. Comput. Vision*, 70(3):279–298, 2006.

8. L. Bar, N. Sochen, and N. Kiryati. Image deblurring in the presence of salt-and-pepper noise. In *Scale-Space*, pages 107–118, 2005.

9. M. Bar. Visual objects in context. *Nature Reviews: Neuroscience*, 5:617–629, August 2004.

10. P. Bayerl and H. Neumann. Disambiguating visual motion through contextual feedback modulation. *Neural Computation*, 16:2041–2066, 2004.

11. M. J. Black and A. Rangarajan. On the unification of line processes, outlier rejection, and robust statistics with applications in early vision. *Int. J. Comput. Vision*, 19(1):57–91, 1996.

12. M. J. Black and G. Sapiro. Edges as outliers: Anisotropic smoothing using local image statistics. In *Scale-Space*, pages 259 – 270, 1999.

13. M. J. Black, G. Sapiro, D. H. Marimont, and D. Heeger. Robust anisotropic diffusion. *IEEE Trans. Image Processing*, 7(3):421–432, March 1998.

14. A. Blake and A. Zisserman. *Visual Reconstruction.* The MIT Press, Cambridge, MA, 1987.

15. A. Braides. *Approximation of Free-Discontinuity Problems.* Lecture Notes in Mathematics, Vol. 1694. Springer-Verlag, 1998.

16. A. Brook, R. Kimmel, and N. A. Sochen. Variational restoration and edge detection for color images. *J. Math. Imaging Vis.*, 18(3):247–268, 2003.

17. T. Brox, A. Bruhn, N. Papenberg, and J. Weickert. High accuracy optical flow estimation based on a theory for warping. In *ECCV*, volume 4, pages 25–36, 2004.

18. T. Brox and D. Cremers. On the statistical interpretation of the piecewise smooth Mumford-Shah functional. In *SSVM*, pages 203–213, 2007.

19. T. Brox and J. Weickert. A TV flow based local scale estimate and its application to texture discrimination. *Journal of Visual Communication and Image Representation*, 17(5):1053–1073, 2006.

20. A. Buades, B. Coll, and J.-M. Morel. A non-local algorithm for image denoising. In *CVPR*, volume 2, pages 60–65, 2005.

21. B. Burgeth, J. Weickert, and S. Tari. Minimally stochastic schemes for singular diffusion equations. In X.-C. Tai, K.-A. Lie, T. F. Chan, and S. Osher, editors, *Image Processing Based on Partial Differential Equations*, Mathematics and Visualization, pages 325–339. Springer Berlin Heidelberg, 2006.

22. J. Canny. A computational approach to edge detection. *IEEE Trans. Pattern Anal. Mach. Intell.*, 8(6):679–698, 1986.

23. R.H. Chan, C.-W. Ho, and M. Nikolova. Salt-and-pepper noise removal by median-type noise detectors and detail-preserving regularization. *IEEE Trans. Image Processing*, 14(10):1479–1485, 2005.

24. T. Chan, S. Esedoglu, and K. Ni. Histogram based segmentation using Wasserstein distances. In *SSVM*, pages 697–708, 2007.

25. T. Chan and L. Vese. Variational image restoration and segmentation models and approximations. *UCLA CAM-report 97-47*, September, 1997.

26. T. F. Chan, B.Y. Sandberg, and L.A. Vese. Active contours without edges for vector-valued images. *Journal of Visual Communication and Image Representation*, 11(2):130–141, 2000.

27. S. Chen and X. Yang. A variational method with a noise detector for impulse noise removal. In *SSVM*, pages 442–450, 2007.

28. Y. Chen, H. D. Tagare, S. Thiruvenkadam, F. Huang, D. Wilson, K. S. Gopinath, R. W. Briggs, and E. A. Geiser. Using prior shapes in geometric active contours in a variational framework. *Int. J. Comput. Vision*, 50(3):315–328, 2002.

29. D. Cremers, F. Tischhäuser, J. Weickert, and C. Schnörr. Diffusion snakes: Introducing statistical shape knowledge into the Mumford-Shah functional. *Int. J. Comput. Vision*, 50(3):295–313, 2002.

30. J. Darbon and M. Sigelle. Image restoration with discrete constrained total variation part I: Fast and exact optimization. *J. Math. Imaging Vis.*, 26(3):261–276, 2006.

31. J. Darbon and M. Sigelle. Image restoration with discrete constrained total variation part II: Levelable functions,

convex piors and non-convex cases. *J. Math. Imaging Vis.*, 26(3):277–291, 2006.

32. J. Debayle and J.C. Pinoli. General adaptive neighborhood image processing: Part I. *J. Math. Imaging Vis.*, 25:245–266, 2006.

33. G. Deng. Symbol mapping and context filtering for lossless image compression. In *ICIP*, volume 1, pages 526–529, 1998.

34. A. Desolneux, L. Moisan, and J.-M. Morel. Meaningful alignments. *Int. J. Comput. Vision*, 40(1):7–23, 2000.

35. A. Desolneux, L. Moisan, and J.-M. Morel. *From Gestalt Theory to Image Analysis: A Probabilistic Approach.*, volume 35 of *Interdisciplinary Applied Mathematics.* Springer-Verlag, New York, NY, 2007.

36. E. Erdem, A. Erdem, and S. Tari. Edge strength functions as shape priors in image segmentation. In *EMM-CVPR*, pages 490–502, 2005.

37. E. Erdem, A. Sancar Yilmaz, and S. Tari. Mumford-Shah regularizer with spatial coherence. In *SSVM*, pages 545–555, 2007.

38. S. Esedoglu and J. Shen. Digital image inpainting by the Mumford - Shah - Euler image model. *European J. Appl. Math.*, 13:353–370, 2002.

39. O. Faugeras. *Three-Dimensional Computer Vision.*, chapter 4.2. The MIT Press, 1993.

40. M. Galun, E. Sharon, R. Basri, and A. Brandt. Texture segmentation by multiscale aggregation of filter responses and shape elements. In *ICCV*, volume 1, pages 716–723, 2003.

41. S. Geman and D. Geman. Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images.

IEEE Trans. Pattern Anal. Mach. Intell., 6:721–639, 1984.

42. G. Gilboa, J. Darbon, S. Osher, and T. Chan. Nonlocal convex functionals for image regularization. *UCLA CAM-report 06-57*, 2006.

43. G. Gilboa, N. Sochen, and Y. Y. Zeevi. Variational denoising of partly textured images by spatially varying constraints. *IEEE Trans. Image Processing*, 15(8):2281–2289, August 2006.

44. R. M. Haralick and J. S. J. Lee. Context dependent edge detection. In *CVPR*, pages 223–228, 1988.

45. M. Heiler and C. Schnörr. Natural image statistics for natural image segmentation. *Int. J. Comput. Vision*, 63(1):5–19, 2005.

46. T. Hofmann, J. Puzicha, and J. M. Buhmann. An optimization approach to unsupervised hierarchical texture segmentation. In *ICIP*, volume 3, pages 213–216, 1997.

47. B.-W. Hong, E. Prados, S. Soatto, and L. Vese. Shape representation based on integral kernels: Application to image matching and segmentation. In *CVPR*, volume 1, pages 833– 840, 2006.

48. J. P. Kaufhold. *Energy Formulations of Medical Image Segmentations.* PhD thesis, Boston University College of Engineering, Department of Biomedical Engineering, Boston, MA, 2000.

49. I. Kokkinos, G. Evangelopoulos, and P. Maragos. Texture analysis and segmentation using modulation features, generative models and weighted curve evolution. *IEEE Trans. Pattern Anal. Mach. Intell.*, to appear.

50. G. C. Lai and R. J. P. de Figueiredo. Image interpretation using contextual feedback. In *ICIP*, volume 2, page 2623, 1995.

51. M. E. Leventon, E. L. Grimson, and O. D. Faugeras. Statistical shape influence in geodesic active contours. In *CVPR*, pages 1316–1323, 2000.

52. P. Lombardi, V. Cantoni, and B. Zavidovique. Context in robotic vision: Control for real-time adaptation. In *ICINCO*, volume 3, pages 135–142, 2004.

53. D. Marr. *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information.* W. H. Freeman, San Francisco, 1982.

54. D. Martin, C. Fowlkes, D. Tal, and J. Malik. A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics. In *ICCV*, volume 2, pages 416–423, 2001.

55. J.-M. Morel and S. Solimini. *Variational Methods in Image Segmentation.* Birkhauser, 1995.

56. D. Mumford and J. Shah. Optimal approximations by piecewise smooth functions and associated variational problems. *Commun. Pure Appl. Math.*, 42(5):577–685, 1989.

57. M. Nielsen, L. Florack, and R. Deriche. Regularization, scale-space and edge detection filters. *J. Math. Imaging Vis.*, 7(4):291–307, 1997.

58. M. Nikolova. A variational approach to remove outliers and impulse noise. *J. Math. Imaging Vis.*, 20(1-2):99–120, 2004.

59. P. Perona and J. Malik. Scale-space and edge detection using anisotropic diffusion. *IEEE Trans. Pattern Anal. Mach. Intell.*, 12:629–639, 1990.

60. H. H. Pien, M. Desai, and J. Shah. Segmentation of MR images using curve evolution and prior information. *IJPRAI*, 11(8):1233–1245, 1997.

61. H. H. Pien and J. M. Gauch. Variational segmentation of multi-channel MRI images. In *ICIP*, volume 3, pages 508–512, 1994.

62. T. Riklin-Raviv, Kiryati, and N. A. Sochen. Unlevelsets: Geometry and prior-based segmentation. In *ECCV*, volume 4, pages 50–61, 2004.

63. G.X. Ritter and J.N. Wilson. *Handbook of Computer Vision Algorithms in Image Algebra,*. CRC Press, 1996.

64. S. Roth and M. J. Black. Fields of experts: A framework for learning image priors. In *CVPR*, volume 2, pages 860– 867, 2005.

65. M. Rousson, T. Brox, and R. Deriche. Active unsupervised texture segmentation on a diffusion based feature space. In *CVPR*, volume 2, pages 699–704, 2003.

66. L. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. *Phys. D.*, 60:259–268, 1992.

67. H. Scharr, M. J. Black, and H. W. Haussecker. Image statistics and anisotropic diffusion. In *ICCV*, volume 2, pages 840– 847, 2003.

68. O. Scherzer and J. Weickert. Relations between regularization and diffusion filtering. *J. Math. Imaging Vis.*, 12(1):43–63, 2000.

69. O. Schwartz, A. Hsu, and P. Dayan. Space and time in visual context. *Nature Reviews: Neuroscience*, 8:522–535, July 2007.

70. J. Serra. *Image Analysis and Mathematical Morphology.* Academic Press, 1982.

71. J. Shah. Segmentation by nonlinear diffusion. In *CVPR*, pages 202–207, 1991.

72. J. Shah. A common framework for curve evolution, segmentation and anisotropic diffusion. In *CVPR*, pages 136–142, 1996.

73. C. Sminchisescu, A. Kanaujia, Z. Li, and D. Metaxas. Conditional models for contextual human motion recognition. In *ICCV*, volume 2, pages 1808–1815, 2005.

74. M. Sonka, V. Hlavac, and R. Boyle. *Image Processing, Analysis and Machine Vision.* Chapman and Hall, 1993.

75. A. Srivastava, A. B. Lee, E. P Simoncelli, and S.-C. Zhu. On advances in statistical modeling of natural images. *J. Math. Imaging Vis.*, 18(1):17–33, 2003.

76. Q. Tang, N. Sang, and T. Zhang. Contour detection based on contextual influences. *Image and Vision Computing*, 25:1282–1290, 2007.

77. S. Tari, J. Shah, and H. Pien. Extraction of shape skeletons from grayscale images. *CVIU*, 66(2):133–146, 1997.

78. S. Teboul, L. Blanc-Fraud, G. Aubert, and M. Barlaud. Variational approach for edge preserving regularization using coupled PDE's. *IEEE Trans. Image Processing*, 7(3):387–397, March 1998.

79. A. Torralba and A. Oliva. Statistics of natural image categories. *Network-Computation in Neural Systems*, 14(3):391–412, 2003.

80. A. Tsai, A. Yezzi, W. Wells, C. Tempany, D. Tucker, A. Fan, E. Grimson, and A. Willsky. A shape-based approach to the segmentation of medical imagery using level sets. *IEEE Trans. Med. Imaging*, 22(2):137–154, 2003.

81. L. A. Vese and S. J. Osher. Image denoising and decomposition with total variation minimization and oscillatory functions. *J. Math. Imaging Vis.*, 20(1-2):7–18, 2004.

82. J. Weickert. Coherence-enhancing diffusion filtering. *Int. J. Comput. Vision*, 31(2-3):111–127, 1999.

83. L. Wolf and S. Bileschi. A critical view of context. *Int. J. Comput. Vision*, 69(2):251–261, 2006.

84. L. Wolf, X. Huang, I. Martin, and D. Metaxas. Patch-based texture edges and segmentation. In *ECCV*, volume 2, pages 481–493, 2006.

85. S. Zheng, Z. Tu, and A. Yuille. Detecting object boundaries using low-, mid-, and high-level information. In *CVPR*, pages 1–8, 2007.

86. S. C. Zhu and D. Mumford. Prior learning and Gibbs reaction-diffusion. *IEEE Trans. Pattern Anal. Mach. Intell.*, 19(11):1236–1250, 1997.