

Density Driven Diffusion

Freddie Åström^{1,2}, Vasileios Zografos¹, and Michael Felsberg^{1,2}

¹ Computer Vision Laboratory, Linköping University, Sweden

² Center for Medical Image Science and Visualization (CMIV), Linköping University
{freddie.astrom, vasileios.zografos, michael.felsberg}@liu.se

Abstract. In this work we derive a novel density driven diffusion scheme for image enhancement. Our approach, called D3, is a semi-local method that uses an initial structure-preserving oversegmentation step of the input image. Because of this, each segment will approximately conform to a homogeneous region in the image, allowing us to easily estimate parameters of the underlying stochastic process thus achieving adaptive non-linear filtering. Our method is capable of producing competitive results when compared to state-of-the-art methods such as non-local means, BM3D and tensor driven diffusion on both color and grayscale images.

1 Introduction

Image denoising is still an open problem in computer vision. The most common approach is to assume a noise distribution, often modeled as an additive component of the observed image. Thus making the goal of image denoising the reduction of the noise variance embedded in the image signal, such that the noise free image can be recovered.

Estimation of the noise distribution from a noisy signal is an ill-posed problem and the challenge is to preserve lines and edges for achieving robust filtering. In image processing, by the principles of ergodicity and stationarity, it is commonly accepted that an image pixel value can be assumed to belong to the same stochastic process as its local neighborhood. Within this statistical interpretation of image representation, two coarse subgroups of denoising algorithms can be identified; local and semi-local methods. One of the currently most notable semi-local denoising algorithms is the non-local means (NLM) [1]. The basic approach of NLM is to compute averages of similar image patches in a neighborhood. The rationale being that with more similar patches the better the image structure is represented. The mean shift method (MS) [2] also belongs to this category. It estimates kernel density functions in feature space and using mode locations of the densities they achieve a denoising effect by applying spatial and range filters. Similarly to NLM, the image denoising method BM3D [3, 4] locates patches in the image by grouping similar image features. This produces a sparse 3D representation on which a weighted filtering can be performed. Another recent approach is the extended anisotropic diffusion (EAD) [5], which is a local method that estimates the orientation of image structures using a diffusion tensor, thus

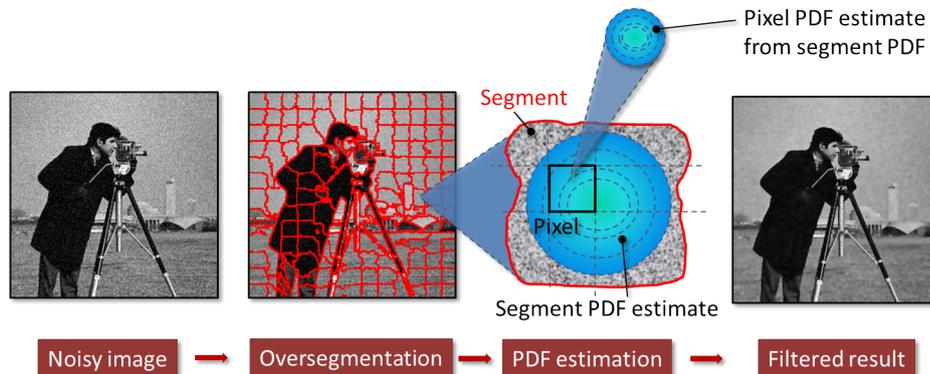


Fig. 1. Overview of the three main components of our approach. First the noisy image is oversegmented into multiple homogeneous regions (segments). Then local PDFs are estimated using the segments’ information. The estimated PDFs drive the filtering process, resulting in the final denoised image.

achieving adaptive filtering. The difference between EAD and anisotropic diffusion as defined by Weickert [6] is that EAD models a non-symmetric tensor in addition to the standard diffusion tensor.

Under the statistical motivation and in contrast to the other methods we propose a novel Density Driven Diffusion scheme (D3) which incorporates probability density estimates derived from local information, such as texture or intensities to achieve robust image filtering. The proposed method consists of three components. The first part involves generating a structure preserving segmentation map by applying an oversegmentation process to each image. Such a map allows for simple, unimodal density models to be easily estimated from the homogeneous information that will be contained in each segment. The second part involves extracting density functions from the segmentation map and the third part minimises the proposed energy functional, thus achieving a semi-local non-linear adaptive filtering scheme. Therefore, our main contribution is to incorporate density information into an energy functional, resulting in an adaptive filtering scheme based on a stochastic image representation.

2 Method Description

Our proposed method D3, consists of three basic steps. First we generate a *segmentation map* from a given noisy image. That is, we **oversegment** the image, obtaining a number of segments (image regions) that ideally exhibit two important properties: i) they are homogeneous (e.g. in terms of colour, texture, intensity and so on) and ii) they obey image boundaries, i.e. they align themselves along strong discontinuity boundaries in the image, instead of crossing them.

The segmentation map is utilised in the next step, which is the **estimation** of a number of density functions (PDFs), each describing the distribution of pixel

values in each segment. We finally exploit the estimated PDFs in order to drive the **filtering** process in the third and final step, and thereby obtain the denoised image. Such a scheme takes advantage of the structure preserving properties of the oversegmentation and thus filters more heavily in regions of near-uniform colour/intensity (i.e. *within* each segment), while reducing the amount of filtering near edges (i.e. *between* different segments). As a result, our method is more powerful than the simple linear diffusion schemes.

In the following sections, we describe in more detail the three main components of our approach, starting with the oversegmentation step. We follow with the density-driven filtering scheme, and conclude with a discussion of the PDF estimation process. An overview of the method is illustrated in Fig. 1.

2.1 Oversegmentation

The first step of our method involves dividing (oversegmenting) the image into distinct, non-overlapping regions, each containing a number of pixels with some consistent and perceptually meaningful set of properties (e.g. colour, texture or intensity). These regions form the segmentation map, which in turn is used for the estimation of the local PDFs necessary for the filtering stage. Since as we will discuss later, we estimate a simple, unimodal PDF in each region/segment, we assume that each segment spans approximately a homogeneous intensity/colour image region and that the pixel values are i.i.d. across different segments. In other words, the segments do not span strong discontinuity boundaries. Examples from such segmentation maps can be seen in Fig. 3. Obviously, not every oversegmentation method can produce segmentation maps with the above properties. There is a line of approaches called “superpixel” segmentation methods that can generate such maps. In particular, we have used the “relaxation labelling” method by [7], which adapts the size and shape of each segment (superpixel) according to the image structure, always favouring map configurations where each segment expands to fill a homogeneous image region, while avoiding crossing edges.

The advantage of using a segmentation map to drive the filtering process, as opposed to a uniform, local PDF estimation scheme can be seen in Fig. 2. In Fig. 2(a) we can see the noisy synthetic input image. Fig. 2(c) shows the segmentation map produced by [7]. Compare this with the randomized uniform grid in Fig. 2(f). In Fig. 2(d) and (g) we can see the “probability maps” estimated from the segmentations in Fig. 2(c) and (f) respectively. A probability map illustrates the diffusion strength as determined by the estimated PDFs. Red indicates strong diffusion and that should occur in homogeneous image regions. Blue indicates very weak diffusion and that should be along strong image edges. Note that the probability map in Fig. 2(c) shows a more desirable diffusion behaviour leading to the very good denoising result in Fig. 2(e). This is because the unimodal local PDFs that we have estimated in each segment from Fig. 2(c) are good approximations to the true pixel distributions. Contrast this with the probability map Fig. 2(g) estimated from the non-structure preserving segmentation in Fig. 2(f). We see that because of the poor estimation of the local PDFs in that case, the probability map is noisy, with artifacts and does not reduce

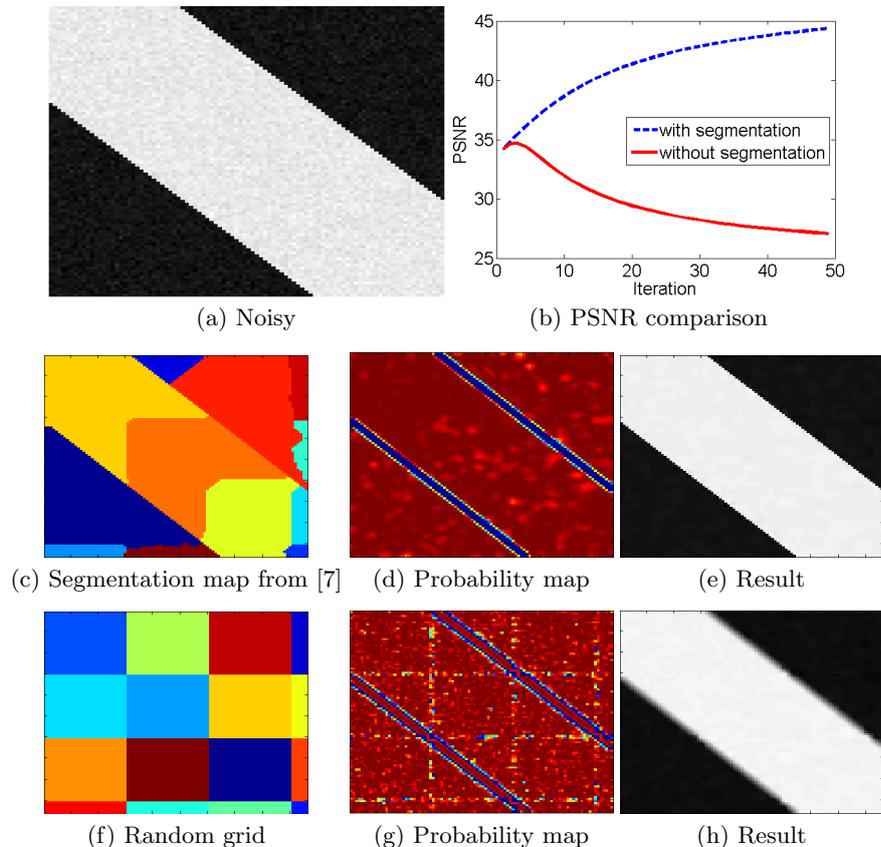


Fig. 2. Example of applying the proposed filtering method on an synthetic image driven by an accurate segmentation map, compared to a map initialized with random labels. See section 2.1 for more details.

the filtering strength sufficiently along edges. As expected the resulting image in Fig. 2(h) contains the same artifacts, and also most of the edge information has been lost. We can further see the quantitative difference between the two examples in terms of the PSNR plots in Fig. 2(b).

2.2 Novel Filtering Scheme

In the motivation of the proposed filtering scheme, originally derived in [8], we assume that content in a local neighborhood adheres to the same stochastic process described by the random variable U . Then let $m(u)$ be the probability of a sample $u(x, y)$ belonging to this process. The PDF associated with the cumulative distribution function (CDF) $m(u)$ is then the non-negative function $m'(u)$. Thus we define a functional $E(u)$ such that

$$E(u) = \int_{\Omega} (u - u^0)^2 d\mathbf{x} + \lambda \int_{\Omega} |\nabla m(u)|^2 d\mathbf{x} \quad (1)$$

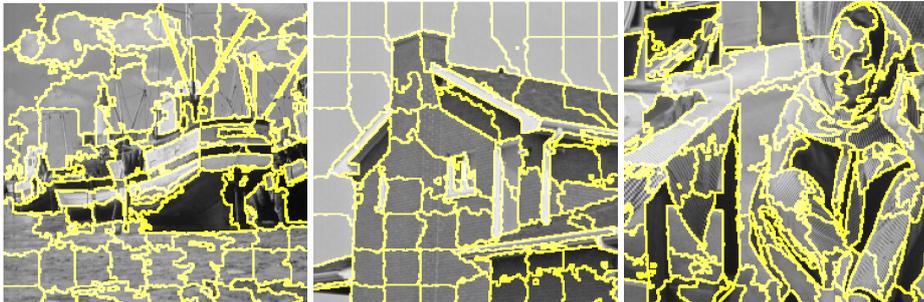


Fig. 3. Typical segmentation maps produced by [7]. Notice how the segments are generally aligned with strong discontinuity boundaries in the image, while they adjust their shape and size to better approximate the homogeneous image regions.

where $m(u) : \mathbb{R} \rightarrow \mathbb{R}$ and $u : \mathbb{R} \rightarrow \mathbb{R}$. We require that $m(u) \in \mathcal{C}^3(\Omega)$ and $\lambda > 0$ is a parameter determining the influence of the smoothness term. The gradient operator ∇ is defined as $\nabla = \partial_{\mathbf{x}}$. The interpretation of the smoothness term is that by *minimizing the gradient of the density function, its variance is reduced, i.e. the distribution is sharpened*. In terms of noise estimation this implies that the samples belonging to the new random variable describing the sharpened distribution \tilde{U} more accurately reflects the true value distribution of the signal.

In order to minimize the functional $E(u)$ the variational derivative of the smoothness term is computed using the Gâteaux derivative

$$\langle \partial R, v \rangle = \lim_{\varepsilon \rightarrow 0} \frac{|\nabla m(u + \varepsilon v)|^2 - |\nabla m(u)|^2}{\varepsilon} \quad (2)$$

where $v \in \mathcal{C}^1(\Omega)$ is an arbitrary function such that $\partial_n v|_{\partial\Omega} = 0$. Using the chain rule $\nabla m(u) = m'(u)\nabla u$ we obtain with Green's identity and Neumann boundary conditions the Euler-Lagrange (E-L) equation

$$\begin{cases} u - u^0 - \lambda(\operatorname{div}(m'(u)^2 \nabla u) + m'(u)^2 \Delta u) = 0 & \text{in } \Omega \\ m'(u)^2 \nabla u \cdot \mathbf{n} = 0 & \text{on } \partial\Omega \end{cases} \quad (3)$$

Since $m'(u)^2 \geq 0$ it is guaranteed that a solution of (3) exists. If and only if m is a globally linear function, (3) becomes identical to linear diffusion, that is, only the Laplacian in (3) remains. Numerically we solve the E-L equation as an initial value problem (IVP) to obtain the diffusion equation.

2.3 Estimating Density Functions

For every segment $s \in \mathcal{S}$, where \mathcal{S} is the set of segments, found by the oversegmentation step, we fit a Gaussian distribution with μ_s and σ_s^2 being the mean value and the variance of the value distribution. Due to the properties of the oversegmentation we can assume that the distributions in \mathcal{S} are independent.

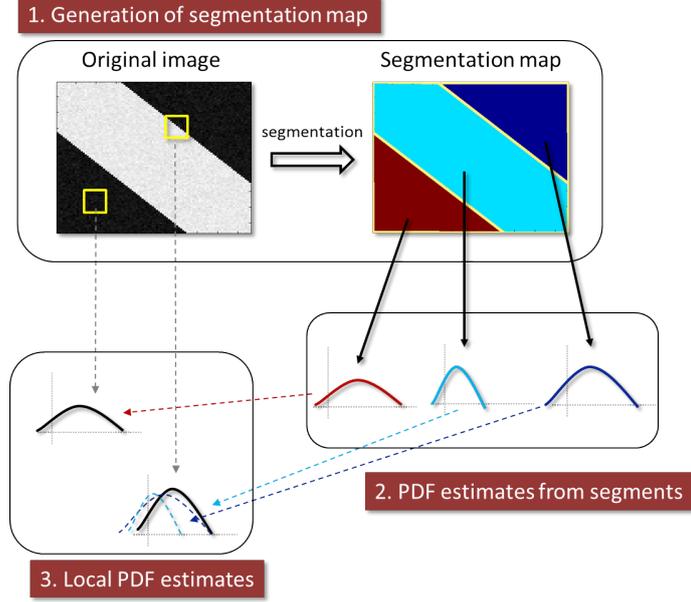


Fig. 4. Overview of the estimation of density functions in homogeneous regions giving rise to a single local density estimate. On the other hand, in textured regions they will give rise to mixture models. The number of mixtuers are given by the number of spanned segments, and since the distributions are independent inside each segment, the mixture fitting is trivially determined.

Therefore, let the derivative of the mapping function in the E-L equation (3) be the product of n distributions in a spatial neighborhood \mathcal{L} at a pixel with location $(x, y) \in \Omega$ then

$$m'(u) = \prod_{i \in \mathcal{L}} m'_i(u) \quad (4)$$

where $m'_i(u) = \mathcal{N}(u; \mu_i, \sigma_i^2)$. Fig. 4 shows this process on an synthetic image for two cases; first for a homogeneous region where a single PDF is estimated and second in the case when a mixture of two regions give rise to a corresponding mixture model. The corresponding CDF of $m'_s(u)$ is given by the error function

$$m_s(u) = \frac{1}{2} \left(1 + \operatorname{erf} \left(\frac{u - \mu_s}{\sigma_s \sqrt{2}} \right) \right) \quad (5)$$

Note that the interpretation of the CDF is that it is a non-linear function which adaptively selects a range of intensity values corresponding to the segments value distribution. This can be interpreted in contrast to linear diffusion where the value domain is considered equally since $m'(u)$ would be a constant function. The difference to non-linear diffusion as defined by Perona and Malik [9] is that $m'(u)$ is selected by the image value distribution rather than an ad-hoc function



Fig. 5. Standard grayscale test images used in the evaluation.

operating on the gradient domain. In practice we are never required to compute the actual CDF since it does not appear in the E-L equation (3).

3 Evaluation

In the next sections we first investigate the sensitivity of the parameter setting of the oversegmentation with respect to obtained error values. Then a qualitative and quantitative evaluation using the peak-signal-to-noise (PSNR) ratio and the structural similarity index [10] (SSIM) is presented for a number of standard grayscale test images and color images. The usage of PSNR is a common metric in the denoising literature to measure image quality. However the SSIM index has been shown to better correlate error values with the perceptual impression of the image quality [10].

3.1 Sensitivity Analysis

Since the filtering performance will depend on how accurately the estimated density functions represent the local image regions, we are interested in understanding how sensitive the parameter selection of the used segmentation method is. The oversegmentation is controlled by two parameters: the “clique size” which determines the initial size of the segments and a the “clique cost” which is a sensitivity parameter. Depending on the image content different values of these parameters will be required. Typically for images with large homogeneous regions larger segments and a more rigid sensitivity parameter suffice. If the scene is highly textured then smaller segments are preferable. If noise is present then the sensitivity parameter needs to be set not to model the noise as structure.

In Fig. 6 (left), by setting the sensitivity parameter to 0.3 the segment size parameter is plotted against the obtained error value for 9 different noise levels in the range 1 to 100. Here we used an 8-bit quantisation representation of the noise. In Fig. 6 (right) we set the segment size to a fixed value (10×10 pixels) and the sensitivity parameter is plotted against the corresponding SSIM value. In both cases, the SSIM values are averaged over all the grayscale images seen in Fig. 5 with the corresponding segmentation sizes and sensitivity coefficients. We can see in Fig. 6 that the selection of parameters for a large number of noise levels is not critical for the performance of the density driven diffusion. This illustrates that the filtering method is robust in terms of parameter selection.

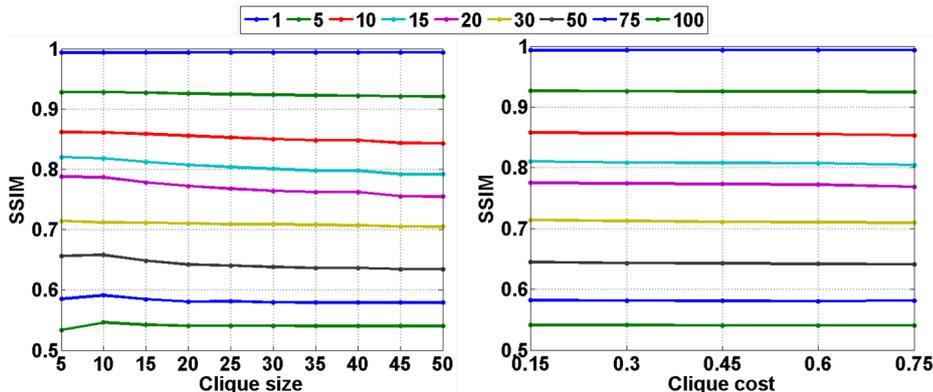


Fig. 6. Analysis of SSIM values for varying segmentation sizes for a fixed sensitivity value (left) and evaluation of the sensitivity parameter for a fixed segment size (right) as a function of different noise levels. See text for details. *Best viewed in color*

3.2 Denoising Results

We evaluate the method on a number of color images and a set of standard grayscale test images. For the considered test images the error values (given in Fig. 7 and 8) are comparable to the state-of-the-art denoising methods BM3D [3, 4], NLM [1] and EAD [5]. The implementation of each evaluated method was obtained from their respective authors. Since all methods except the proposed method, D3, rely on the image noise we follow [5] who estimated the noise according to the method presented in [11].

The obtained SSIM values for the grayscale images in Fig. 5 are seen in Fig. 7. The differences in SSIM value are on average not significant and the performance of respective method is correlated with the image content. Observe that BM3D [4] performs well in images with large homogeneous regions such as the House image. Color images indexed 8143 (owl), 87065 (lizard), 175043 (snake), 208001 (mushroom) available through the Berkeley segmentation dataset [12] were used in [5]. Thus we evaluate D3 also on these images. It has been shown in previous works that color image denoising in the RGB (red, green and blue) color space is suboptimal due to high correlation of the R,G and B color components [13, 5]. Our approach for applying D3 on color images is to again use the segmentation method of [7]. The produced segmentation map conforms to homogeneous color regions and it is used to estimate density functions using an opponent color transform derived in [14]. The density functions are now estimated in the transformed space based on the derived segmentation map. This approach is natural since the spatial location of the pixel has not changed.

The result of the color image evaluation is seen in Fig. 8. The poor error value of NLM and CBM3D [3] particularly seen in Fig. 9(a) is due to inaccurate estimation of the noise, thus illustrating that they heavily depend on the accuracy

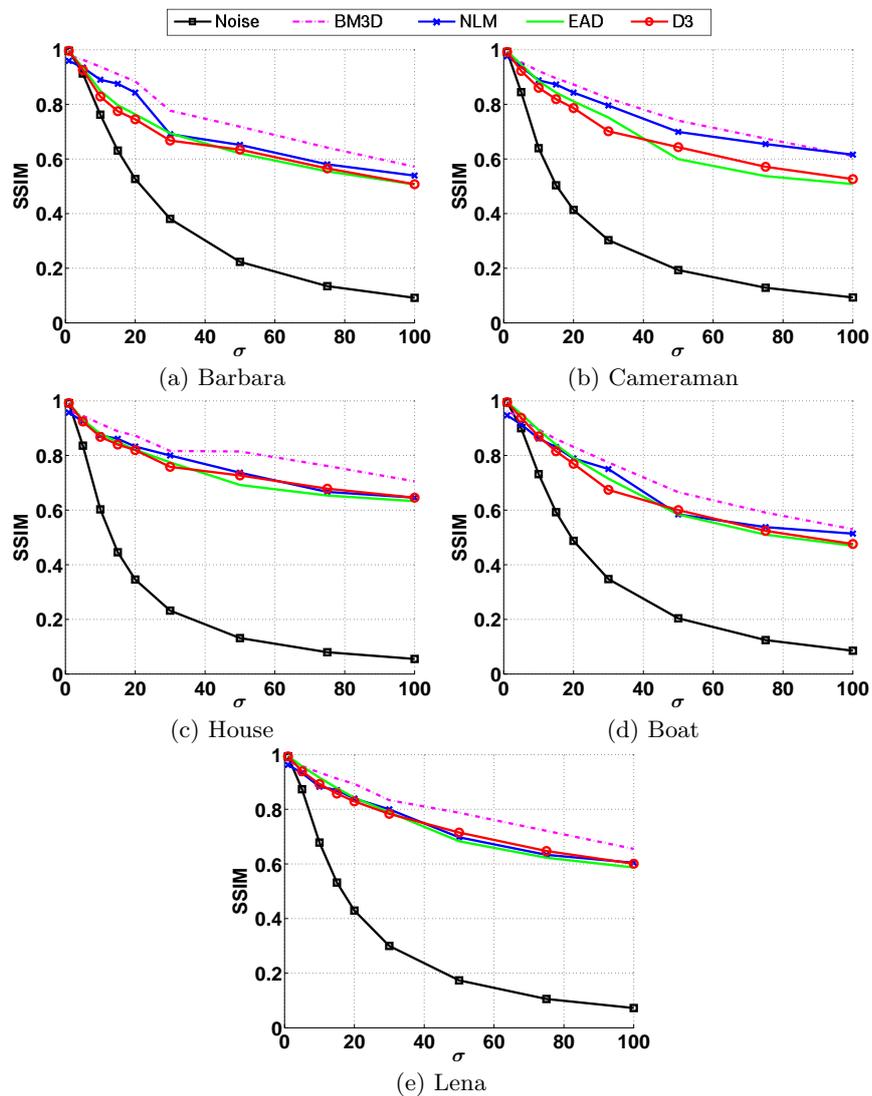


Fig. 7. Obtained SSIM values for standard grayscale test images. Where “Noise” corresponds to the image baseline noise. *Best viewed in color*

of any noise estimation technique to define their filtering parameters. The proposed method D3, compares well with the tensor-based EAD for the investigated color images. Here we point out that the EAD is a local tensor-based method estimating the local orientation, thus achieving adaptive filtering. In contrast, D3 is a semi-local method considering the density functions estimated from each segment of the segmentation map independently of the estimated noise. Example images of the denoising result can be seen in Fig. 9, indicated for each image is

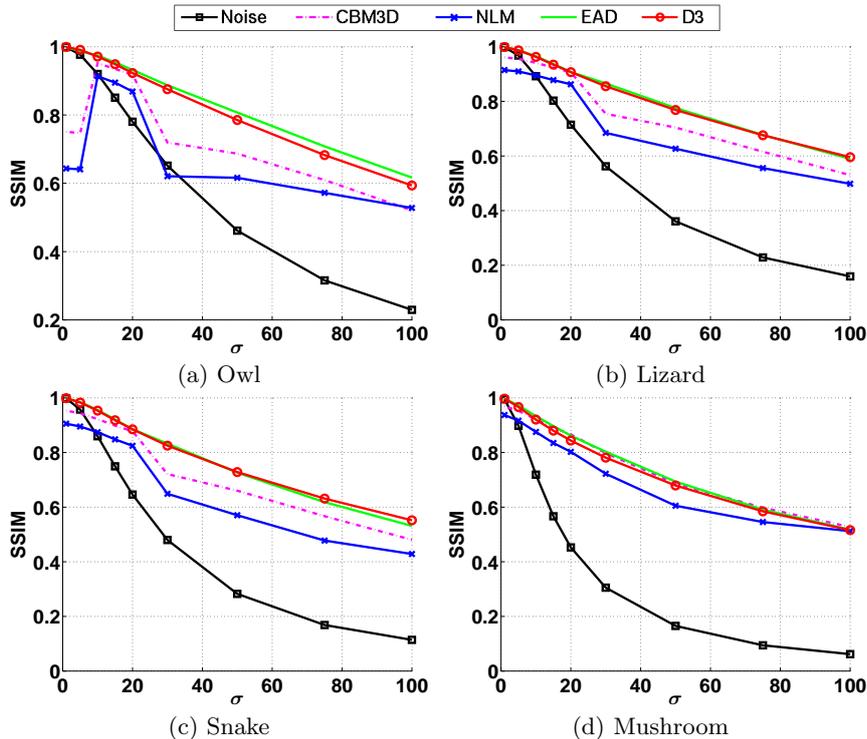


Fig. 8. Obtained SSIM results for color images. Where “Noise” corresponds to the image baseline noise. *Best viewed in color.*

the corresponding SSIM and PSNR values at noise level 30 in Fig. 8. For the images with many high-frequency components such as branches, leafs and feathers it is clear that the local method EAD performs well preserving these structures, particularly visible in the owl image. CBM3D and NLM have an advantage in images which contain many similar patches, i.e. in the mushroom image. The filtering result obtained by D3 appears crisp and no apparent oversmoothing is visible in the owl and the snake as is for the CBM3D and NLM. We have also tested the MS method by [2] but produced very poor results and for that reason is not illustrated here.

4 Conclusion

We have presented a novel method for image enhancement called Density Driven Diffusion (D3). D3 works by incorporating local density estimates into the diffusion process. These estimates are derived from a structure-preserving oversegmentation of the input image. The oversegmentation results into a number of homogeneous, edge-aligned segments, within which, density functions can be easily and independently estimated. The estimated densities drive the diffusion

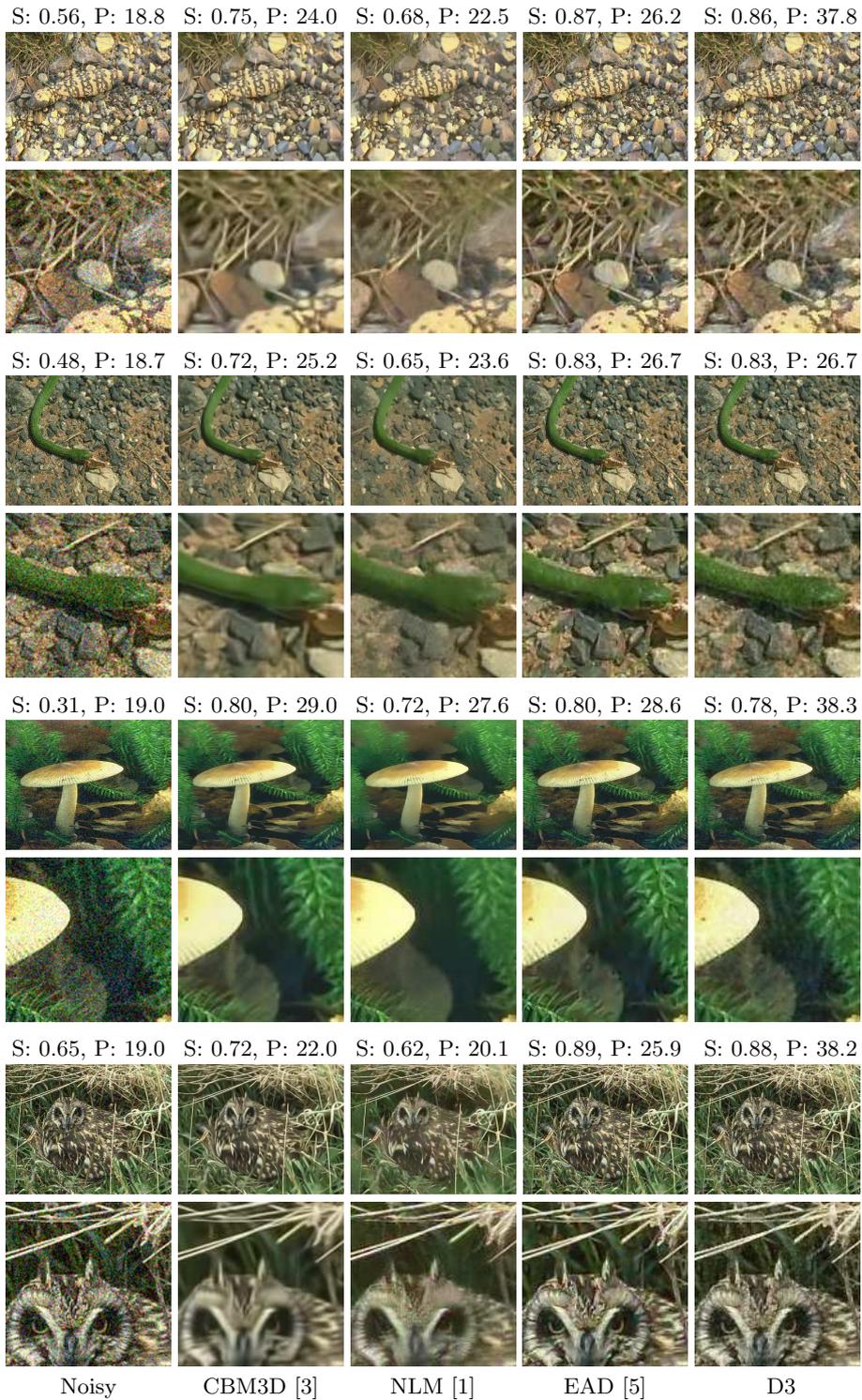


Fig. 9. Results of color images. The images in the first column have been corrupted by 30 standard deviation noise. *Best viewed in color.*

process such that it adheres to image boundaries resulting in a non-linear image enhancement scheme, comparable to state-of-the-art denoising methods. We have tested our approach on a number of images and compared with other methods in literature. We have found that D3 can produce very good results in line with the competing methods. Where D3 particularly excels however, is in the enhancement of colour images.

Acknowledgements

This research has received funding from the Swedish Research Council through grants for the projects Visualization-adaptive Iterative Denoising of Images and Extended Target Tracking (within the Linnaeus environment CADICS), from the ECs 7th Framework Programme (FP7/2007-2013), grant agreement 247947 (GARNICS), and the Vinnova project Online laboratory for medical image analysis.

References

1. Buades, A., Coll, B., Morel, J.M.: A non-local algorithm for image denoising. In: CVPR. Volume 2. (june 2005) 60 – 65 vol. 2
2. Comaniciu, D., Meer, P., Member, S.: Mean shift: A robust approach toward feature space analysis. PAMI **24** (2002) 603–619
3. Dabov, K., Foi, A., Katkovnik, V., Egiazarian, K.: Color image denoising via sparse 3d collaborative filtering with grouping constraint in luminance-chrominance space. In: ICIP. (2007)
4. Dabov, K., Foi, A., Katkovnik, V., Egiazarian, K.: Image denoising with block-matching and 3d filtering. In: SPIE. (2006)
5. Åström, F., Baravdish, G., Felsberg, M.: On Tensor-Based PDEs and their Corresponding Variational Formulations with Application to Color Image Denoising. In: ECCV. (2012)
6. Weickert, J.: Anisotropic Diffusion In Image Processing. ECMI Series, Teubner-Verlag, Stuttgart, Germany (1998)
7. Mester, R., Conrad, C., Guevara, A.: Multichannel segmentation using contour relaxation: fast super-pixels and temporal propagation. In: SCIA. (2011) 250–261
8. Åström, F., Michael, F., Baravdish, G., Lundström, C.: Targeted Iterative Filtering. In: SSVM’13. Volume 7893 of LNCS., Heidelberg, Springer (2013) 1–11
9. Perona, P., Malik, J.: Scale-space and edge detection using anisotropic diffusion. PAMI **12** (1990) 629–639
10. Wang, Z., Bovik, A., Sheikh, H., Simoncelli, E.: Image quality assessment: from error visibility to structural similarity. TIP **13**(4) (april 2004) 600 –612
11. Felsberg, M.: Autocorrelation-driven diffusion filtering. TIP **20**(7) (2011) 1797–1806
12. Martin, D., Fowlkes, C., Tal, D., Malik, J.: A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics. In: ICCV. Volume 2. (July 2001) 416–423
13. Åström, F., Michael, F., Lenz, R.: Color Persistent Anisotropic Diffusion of Images. In: Image Analysis. Volume 6688 of LNCS., Heidelberg, Springer (2011) 262–272
14. Lenz, R., Carmona, P.L.: Hierarchical s(3)-coding of rgb histograms. In: VISAPP. Volume 68. Springer (2010) 188–200