

Image Diffusion Using Saliency Bilateral Filter

Jun Xie^{1,3}, Pheng-Ann Heng^{2,3}, Simon S.M. Ho⁴, and Mubarak Shah¹

¹ Dept. of Computer Science, University of Central Florida, Orlando, USA

² Dept. of Computer Science, The Chinese University of Hong Kong, HK

³ Shun Hing Inst. of Advanced Engineering, The Chinese University of Hong Kong, HK

⁴ Dept. of Diagnostic Radiology & Organ Imaging, The Chinese University of Hong Kong, HK

Abstract. Image diffusion can smooth away noise and small-scale structures while retaining important features, thereby enhancing the performances of many image processing algorithms such as image compression, segmentation and recognition. In this paper, we present a novel diffusion algorithm for which the filtering kernels vary according to the perceptual saliency of boundaries in the input images. The boundary saliency is estimated through a saliency measure which is generally determined by curvature changes, intensity gradient and the interaction of neighboring vectors. The connection between filtering kernels and perceptual saliency makes it possible to remove small-scale structures and preserves significant boundaries adaptively. The effectiveness of the proposed approach is validated by experiments on various medical images including the color Chinese Visible Human data set and gray MRI brain images.

1 Introduction

To accentuate certain image features for subsequent analysis or display, image enhancement is usually performed by either suppressing the noise or increasing the image contrast. In the early development of image processing, linear filters were used primarily owing to their mathematical simplicity and ease of implementation. However, linear filters such as low-pass filtering [1] tend to blur away the sharp boundaries that help to differentiate large scale anatomical structures. Even in cases where linear filters do not obliterate boundaries, they tend to distort the fine structures of the image and thereby change subtle aspects of the anatomical shapes in question.

To alleviate the detrimental effects of linear filters, many efforts have been devoted in the literature for preserving the signal details while eliminating the noise. A powerful approach for denoising is to define a penalty functional that results in the recovered signal as close as possible to the measured signal while maintaining the smoothness except for the salient boundaries. As shown in [2], many well-known approaches, such as the Weighted Least Squares (WLS) [3], Robust Estimation technique (RE) [4] and Anisotropic diffusion (AD) [5], have a fundamental relationship. Similar to the Complex Diffusion [6], all these methods are based on the assumption that local relations between the samples dictate the final result and hence resort to an iterative algorithm.

Tomasi and Manduchi proposed an alternative non-iterative Bilateral Filter (BF) [7] which performs smoothing by using both domain and range neighborhoods. This non-linear approach does not involve the solution of partial differential equations and can be

implemented in a single iteration to generate results as good as iterative approaches [8]. Elad [2] shows that the bilateral filter can also be derived from the Bayesian approach using a novel penalty functional. Despite the relationship between the bilateral filter and AD/RE/WLS, the major difference of the former approach is that it considers all the relative neighborhoods in parallel while the latter focuses on the diffusion of connected samples. Although the bilateral filtering is capable of reducing the noise in an image by an order of magnitude while maintaining edges, it is a challenging task to choose an appropriate kernel, for balancing the trade-off between edge maintenance and noise removal. In [9], Ramanath and Snyder proposed a demosaicking approach using an adaptive bilateral filter kernel generated by linear interpolation.

In this paper, we propose to enhance the performance of the bilateral filter by replacing the constant kernels with a function decreasing with the boundary saliency. This leads to an adaptive filtering scheme which can average smooth regions with a broad filtering kernel and preserves strong boundaries with a sharp kernel. For each pixel in an image, we define a fixed set of “orientation elements” connecting the pixel to its neighboring pixels. The saliency of boundaries in the original images are evaluated through a novel saliency measure which is determined by the intensity gradient at an element, the interaction of neighboring elements and the curvature changes along a sequence of elements. The curvature gradient factor provides us a measure that can treat different types of curves equally. Another advantage of this novel measure is that it doesn’t depend on the closure criterion, which makes it more suitable for realistic applications. Using this edge feature, we propose the Saliency Bilateral Filter (SBF) which can be applied on both gray and color images for adaptive edge-preserving smoothing.

2 Method

The idea behind the bilateral filter is to combine domain and range filtering together, thereby enforcing both geometric and photometric locality. This filter generally smooths signals while preserving steps via a nonlinear combination of nearby sample values. For a 2D image f , the bilateral filter can be described as:

$$h(p) = \frac{\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(q) \Upsilon_d(p, q) \Upsilon_s(f(p), f(q)) dq}{\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \Upsilon_d(p, q) \Upsilon_s(f(p), f(q)) dq} \quad (1)$$

where p and q refer to space variables. Function $\Upsilon_d(p, q)$ is the closeness function which measures the geometric distance:

$$\Upsilon_d(p, q) = \exp\left\{\frac{-d^2(p - q)}{2\sigma_d^2}\right\} \quad (2)$$

and $\Upsilon_s(f(p), f(q))$ is the photometric similarity function:

$$\Upsilon_s(p, q) = \exp\left\{\frac{-[f(p) - f(q)]^2}{2\sigma_s^2}\right\}. \quad (3)$$

The diffusion effect of the bilateral filter is controlled by three parameters: the filtering window size N , the geometric spread σ_s , and the photometric spread σ_d . Usually,

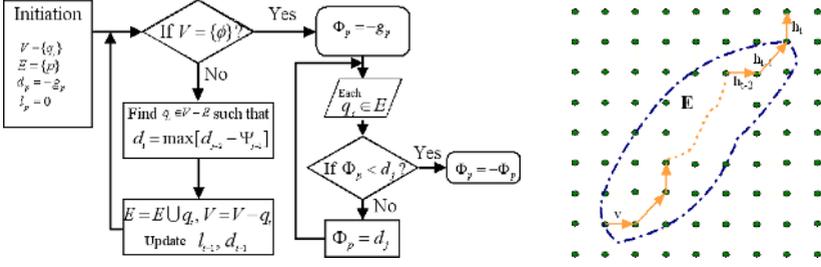


Fig. 1. (Left) Summary of the proposed searching algorithm to compute the novel saliency measure. (Right) Schematic explanation of an iteration of the searching algorithm.

the filtering window size is decided by the spatial kernel radius, while the photometric spread σ_s and geometric spread σ_d are chosen to achieve the desired amount of combination of sample values and the desired amount of low-pass filtering.

2.1 Curvature Gradient Saliency Measure

In this section, we propose to choose the filtering kernel in an adaptive scheme by using perceptual metrics so that the regions with different properties can be treated accordingly. For the measurement of the perceptual saliency of structures in the original image, we propose a novel saliency measure, called Curvature Gradient Saliency (CGS). For each orientation element v on the input image, the L level saliency $\Phi_L(v)$ is defined to be the maximum saliency over all curves of length L emanating from v . The saliency of a curve $\Gamma(s)$ of length l (s denotes arc length, $0 \leq s \leq l$) is defined as

$$S(\Gamma) = \int_0^l \rho^s g(s) C(s) ds \quad (4)$$

where ρ is a constant that determines how quickly the contribution of an element to its neighboring elements decays with distance, and $g(s)$ is the normalized intensity gradient of arc $\Gamma(s)$ defined as:

$$g(s) = \frac{G(s) - G_{min}}{G_{max} - G_{min}} \quad (5)$$

with $G(s)$ represents the intensity gradient of $\Gamma(s)$. Terms G_{max} and G_{min} are the maximum and minimum values of the gradient map respectively. The purpose of employing the normalized intensity gradient is to guide the expansion of saliency in a smoother manner so that elements with high intensity gradient can extend their saliency farther. Function $C(s)$ in Eq. (4) is the curvature gradient function defined as

$$C(s) = \exp(-|\nabla k(s)|) \quad (6)$$

where $k(s)$ refers to the curvature of arc $\Gamma(s)$.

This measure favors significant, long boundaries with constant curvature. Further, it calculates the smoothing factor in local areas providing an efficient way to restrain

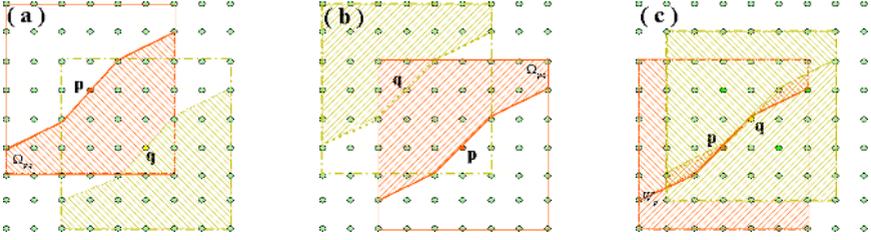


Fig. 2. Three cases of spatial relationship between the central pixel p and its neighbor q for photometric similarity computation, including two cases ((a) and (b)), where q is away from the salient edge across the central point p and case (c) when q is just located on the salient edge.

the spread of local noise. Unlike most existing methods [10, 11], this measure does not depend on the closure criterion. Since the boundaries are not always closed curves in natural and medical images, the independence of closure criterion makes our measure more genetic for realistic applications.

2.2 Computation of CGS

To compute this saliency measure in gradient maps of 2D images, we developed a local searching algorithm which is summarized in Fig. 1(left) and is briefly described as follows. To compute the L level saliency $S_L(v)$ of element p , we consider all the elements $V = \{h_i\}, i = 1, \dots, N$ in the $2L * 2L$ window centered on p . A set of elements is maintained $E = \{h_1, h_2, \dots, h_n\}$ to which we currently know the most salient curve from v as illustrated in Fig. 1(right). Initially, only element v is in set E . Then at each following iteration, we find an element h_t from set $V - E$ such that

$$d_t = \max_{h_j \in (V-E)} [d_{j-2} - \Psi_{j-1}] \tag{7}$$

where d_t is the saliency of the most salient curve from v to h_t and d_{j-2} is the saliency of the most salient curve from v to h_{j-2} , the previous second element before element h_j . It is obvious that all previous elements should have been in E . Ψ_{j-1} is the saliency of the fragment from h_t to h_{t-1} and is defined as

$$\Psi_{j-1} = (\rho^{j-1} g_{j-1} + \rho^j g_j) C_{j-1} \tag{8}$$

with

$$C_{j-1} = |K_{j-2,j-1} - K_{j-1,j}| \tag{9}$$

where $K_{j-1,j}$ is the curvature of fragment from h_{j-1} to h_j . Since h_t satisfies Eq. (7), it is guaranteed that all elements of the most salient curve from v to h_t must have been in set E . Hence, we put h_t into set E as a new element with a known most salient curve from v . After finding the most salient curve for each element in the window, we search for the element h_{max} with the maximum saliency value among the elements whose lengths l_j are equal to L . Then, we define its saliency value d_{max} as $S_L(v)$, the saliency of the central element v .

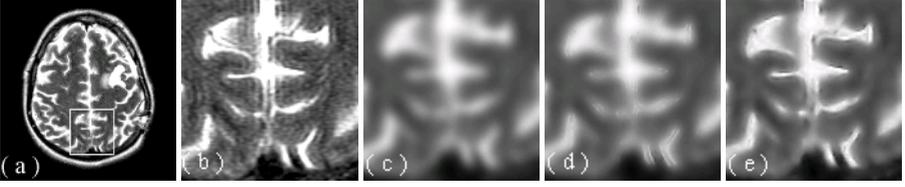


Fig. 3. The filtering results on a 2D gray MRI brain image. (a) The input MRI image. (b) An enlarged part. (c) The result of bilateral filter. (d) The result of adaptive bilateral filter using intensity gradient. (e) The result using saliency bilateral filtering.

2.3 Adaptive Filtering Based on CGS

After extracting the saliency features using the approach introduced above, each pixel p on the 2D image plane Ω has a saliency measure s_p which is defined as the saliency of the greatest boundary across it. Using the basic idea in Eq. (2), the geometric closeness function for SBF can be defined as:

$$\Upsilon_d(p, q) = \exp \left\{ -\frac{(1 + s_p)^2 \|p - q\|}{2} \right\}. \quad (10)$$

Further, we follow the region based filtering idea to measure the photometric similarity between two locations using region difference. The region filtering can improve the robustness of the filtering process instead of using the intensity of single pixel. For pixels on salient boundaries, there are two subregions with significantly different appearances besides the boundary. To compute the photometric similarity more accurately, we compare two pixels within the same side of the salient boundary so that the interference of other subregions can be avoided. As shown in Fig. 2, there are three relations of the locations of pixels p with its neighbor q . For both the cases (Fig. 2(a) and 2(b)) where q is not located on the salient edge, we compute the similarity using only the subregion Ω_{pq} containing q . When pixel q is located on the salient edge (Fig. 2(c)), we compare the similarity within the whole filtering window W_p . Thus, the photometric similarity function for our approach is defined as:

$$D(p, q) = \begin{cases} \frac{\iint_{\Omega_{pq}} (f(x) - f(y))^2 dx dy}{\iint_{\Omega_{pq}} dx dy}, & \text{Case 1 \& 2} \\ \frac{\iint_{W_p} (f(x) - f(y))^2 dx dy}{\iint_{W_p} dx dy}, & \text{Case 3} \end{cases} \quad (11)$$

The range similarity function using saliency measures is then obtained by:

$$\Upsilon_s(p, q) = \exp\left(\frac{-D(p, q)}{2\sigma_s^2}\right) \quad (12)$$

When σ_s is set at a larger value (e.g. 100), the range filtering kernel is very broad, and the range component of the filter has little effect for the domain component. In other words, all neighbors have the same weight so that the combined filter acts as a simple domain filter. When σ_s is small (e.g. 10), only pixels with very similar contexts will be averaged.

3 Results

In this section, we analyze the performance of the proposed approach on both synthetic and medical images. One of the constants for saliency computation is the maximum curve length L which should be chosen large enough to smooth away small-scale structures and noise, but not too large to corrupt significant features. Another constant is the interaction factor ρ , ranging from 0 to 1. A large value of ρ enables our algorithm to fill in large gaps between edge segments, while a small ρ can restrain the spread of local noise more effectively. For all experiments presented here, we set $L = 30$, $\rho = 0.8$ and $\sigma_s = 30$.

Figure 3 shows the results on a gray MRI brain image. The bilateral filter was applied with $\sigma_d = 3$ and $\sigma_s = 30$. From the result in Fig. 3(c), we observe that most of small-scale structures have been removed by the bilateral filter but important edges were also blurred. Figures 3(d) and 3(e) show the results using adaptive filtering scheme based on intensity gradient and saliency measure respectively. It is evident that using intensity gradient can improve the preserving of boundaries, and SBF maintains most of the significant boundaries in the original image. Following the same framework in [7], our algorithm can be easily extended to vector-valued images. Figure 4 shows an experiment on a color image (Fig. 4(a)). When using the RGB space (Fig. 4(c)), there are some extraneous colors on the filtering result (e.g., the blurred pink-purple area). This is because the RGB color models are not well suitable for describing colors in terms that are practical for human interpretation. Therefore we make use of the CIE-Lab color model [12] which is based on a large body of psychophysical data concerning color-matching experiments performed by human observers. Thus, the similarity measured in this space correlates strongly with the perception of color discrepancy. As seen

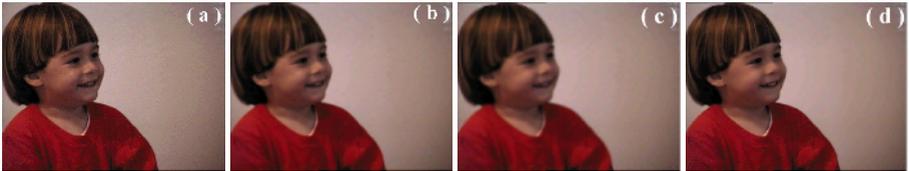


Fig. 4. An experiment on a natural color image. (a) The original color image. (b) The output of the bilateral filter using the CIE-Lab color space with $\sigma_d = 1$ and $\sigma_s = 50$. (c) The output of the adaptive bilateral filter using the RGB color space. (d) The output of the SBF approach using CIE-Lab color space.

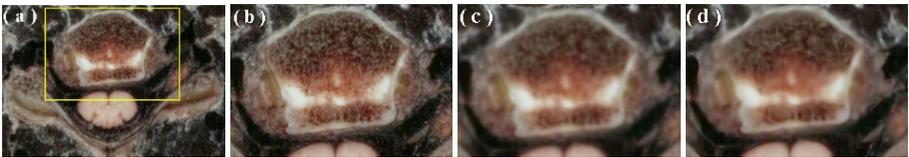


Fig. 5. A diffusion example on the color CVH images. (a) The original color image in the CVH dataset. (b) The patch selected as the test image. (c) The result of the bilateral filter. (d) The result of SBF.

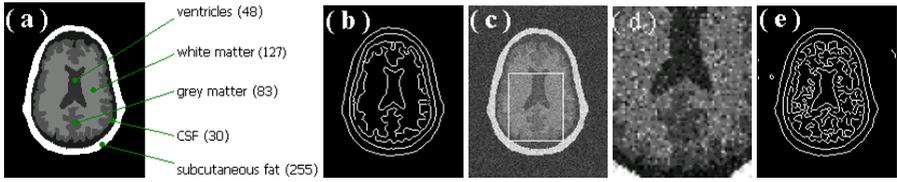


Fig. 6. The brain MR phantom. (a) The artificial phantom of the MR head tomogram as a reference image. (b) The canny edge map. (c) The noisy version of (a). (d) An enlarged view of the noisy image. (e) The canny edge map of the noisy image.

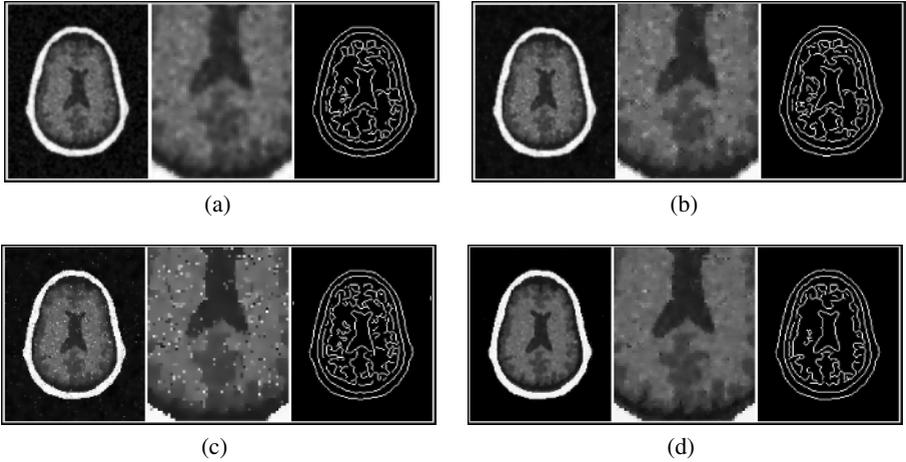


Fig. 7. The diffusion results of the noisy brain phantom image using different filters including (a) the bilateral filter with $\sigma_d = 1$ and $\sigma_s = 40$, (b) the Perona-Malik anisotropic filter 100 iterations with diffusivity constant $K = 10$ and time step of 0.1, (c) the complex decomposition using $\theta = \pi/15$, and (d) the proposed SBF. Each result shows the filtered image, an enlargement part of the result and the canny edge map.

above, by measuring the color distance in the CIE-Lab space, both the bilateral filter (Fig. 4(b)) and the SBF approach (Fig. 4(d)) have greatly corrected this undesirable effect. However, it is noticeable that our algorithm retains more details (e.g. the ear) than the bilateral filter.

Figure 5 illustrates one of our experiments on the color Chinese Visible Human (CVH) data set. The highest resolution CVH data set is about 1143 GB in size and consists of 18,200 cross-sectional color digital images (4064×2704 resolution, 24 bits per pixel) at 0.1mm intervals. Both the bilateral filter and our algorithm have successfully removed most of shading, while more perceptually important edges were preserved by the SBF filter. Both the methods were implemented using C code. For a color image in the CVH data set, the bilateral filter took 1.4 mins on a 2.8-GHz Pentium IV, and our algorithm needs 3.2 mins including both saliency computation and adaptive filtering.

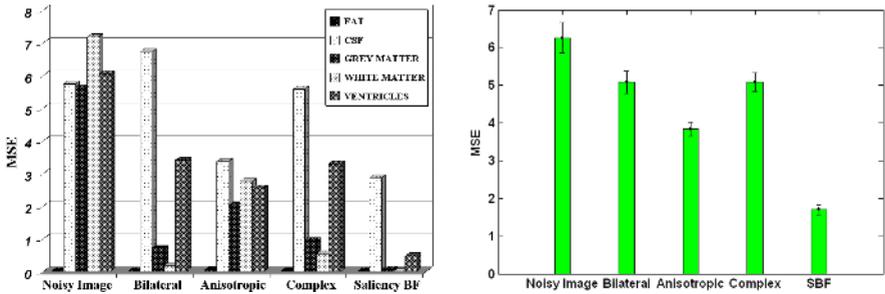


Fig. 8. (Left) The intensity difference of each region interior between brain phantom image with the noisy image and the filtering results of different approaches. (Right) The Mean-Squared Errors between the brain phantom image with the noisy image and the filtering results of different approaches.

To evaluate the performances of those algorithms quantitatively, we use the stepwise constant MR-head phantom in [13]. It contains five ideally segmented components as shown in Fig. 6(a). Each component is stepwise constant with a given intensity. The input image (Fig. 6(c)) is a noisy version of this reference image with a Gaussian noise of standard deviation $\sigma = 20$. Figure 7 shows the filtered results using four algorithms including the bilateral filter, anisotropic diffusion [5], complex diffusion [6] and the proposed SBF algorithm. All the parameters involved in the previous methods were selected by trial-and-error in order to get an optimal result from each method.

We compared the filtering results with the original phantom image. Figure 8(left) shows the intensity differences in five region interiors of the filtering results and the stepwise constant reference image. The average difference of interior intensity reflects the similarity of filtered images with the reference image. The statistics show that among the tested four algorithms, the SBF approach produced the closest result to the phantom image. For the whole frame of filtered images, the obtained Mean-Squared Errors are illustrated in Fig. 8(right), where the result of the SBF algorithm hold the lowest overall MSE for this experiment.

4 Conclusions

The paper presents a novel diffusion approach based on perceptual metrics. To accommodate regions with different features, the filtering kernels are adjusted based on the saliency of boundaries. The connection between the filtering kernel and perceptual saliency makes it possible to remove noise and preserve salient boundaries adaptively. Moreover, the calculation of the saliency provides an efficient way to accurately measure the photometric similarity. We have shown that the proposed approach can be extended for vector-valued images by combining the CIE-Lab color model and tested the proposed saliency bilateral filter on a variety of synthetic and medical images. The qualitative and quantitative results show convincing advantages in comparison with the bilateral filter and other popular approaches.

Acknowledgement

The work described in this paper was supported by a grant from the Research Grants Council of the Hong Kong Special Administrative Region (Project no. CUHK4461/05M) and CUHK Shun Hing Institute of Advanced Engineering

References

1. R. Deriche, "Fast algorithms for low-level vision," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 12, no. 1, pp. 78–87, 1990.
2. M. Elad, "On the origin of the bilateral filter and ways to improve it," *IEEE Trans. Image Processing*, vol. 11, no. 10, pp. 1141–1151, Oct. 2002.
3. R. L. Lagendijk, J. Biemond, and D. E. Boeke, "Regularized iterative image restoration with ringing reduction," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 36, no. 12, pp. 1874–1887, 1988.
4. M. J. Black and G. Sapiro, "Edges as outliers: Anisotropic smoothing using local image statistics," in *Scale-Space*, 1999, pp. 259–270.
5. P. Perona and J. Malik, "Scale-space and edge detection using anisotropic diffusion," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 12, no. 7, pp. 629–639, 1990.
6. G. Gilboa, N. A. Sochen, and Y. Y. Zeevi, "Image enhancement and denoising by complex diffusion processes," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 26, no. 8, pp. 1020–1036, 2004.
7. C. Tomasi and R. Manduchi, "Bilateral filtering for gray and color images," in *Proc. IEEE Int. Conf. on Computer Vision*, Washington DC, USA, 1998, pp. 839–846.
8. D. Barash, "A fundamental relationship between bilateral filtering, adaptive smoothing, and the nonlinear diffusion equation," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 24, no. 6, pp. 844–847, 2002.
9. R. Ramanath and W. Snyder, "Adaptive demosaicking," *J. Electron. Imaging*, vol. 12, no. 4, pp. 633–642, Oct. 2003.
10. L. Williams and D. Jacobs, "Stochastic completion fields: A neural model of illusory contour shape and salience," *Neural Computation*, vol. 9, pp. 849–870, 1997.
11. A. Desolneux, L. Moisan, and J. Morel, "A grouping principle and four applications," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 25, no. 4, pp. 508–513, Apr. 2003.
12. G. Wyszecki and W. Stiles, *Color Science: Concepts and Methods, Quantitative Data and Formulae*. New York, USA: John Wiley and Sons, 1982.
13. I. Bajla and I. Hollander, *Empirical Evaluation Methods in Computer Vision*. Stockholm, Sweden: World Scientific, 2001, ch. Task-based Evaluation of Image Filtering within A Class of Geometry-Driven-Diffusion Algorithms.