

Efficient High-Dimensional Filtering for Image and Video Processing

Eduardo S. L. Gastal¹, Manuel M. Oliveira¹ (Advisor)

¹Instituto de Informática – Universidade Federal do Rio Grande do Sul (UFRGS)

{eslgastal,oliveira}@inf.ufrgs.br

Abstract. *High-dimensional filters are a fundamental building block for several applications, having recently received considerable attention from the research community. Unfortunately, naive implementations of such an important class of filters are too slow for many practical uses. This dissertation describes three novel approaches to efficiently perform high-dimensional filtering with linear cost in both the number of pixels and in the dimensionality of the space in which the filters operate. Our filters address the main limitations of previous techniques, in addition to providing the fastest performance (both on CPU and GPU) for a variety of real-world applications.*

1. Introduction

Filtering is arguably the single most important operation in image and video processing. In particular, *high-dimensional filters* are a fundamental building block for several applications, including upsampling [Kopf et al. 2007], stylization [Winnemöller et al. 2006], spatio-temporal filtering [Lang et al. 2012], denoising [Buades et al. 2005], detail manipulation [Gastal and Oliveira 2012], tone mapping [Fattal 2009], filtering of photon maps for rendering [Bauszat et al. 2011], and recoloring [Gastal and Oliveira 2011].

High-dimensional filters can be classified as *Euclidean* or *geodesic*, according to how they compute the distances between samples. The main difference between the two groups is the filter behavior near strong discontinuities (commonly called edges) in the signal. In general, Euclidean filters allow for samples belonging to different sides of a discontinuity to be combined, while geodesic filters do not. Thus, each filter type provides best results for different applications.

Due to their wide applicability, several high-dimensional and related filters have been proposed—the most popular one being the bilateral filter [Smith and Brady 1997]. Unfortunately, naive implementations of such an important class of filters are too slow for many practical uses, specially in light of the ever increasing resolution of digitally captured images, and the fact that some applications may require working with spaces containing hundreds of dimensions. As a result, several techniques have been developed that either try to accelerate high-dimensional filters, or introduce alternative ways of performing similar filtering operations on images and videos. While they clearly improve performance, these solutions still face limitations such as not being sufficiently fast for real-time applications [Adams et al. 2010], restricting filtering to certain scales [Fattal 2009], and only handling grayscale images [Yang et al. 2009], among others.

2. Contributions of this Dissertation

This dissertation describes three novel approaches to efficiently perform high-dimensional filtering, capable of achieving real-time performance: the **domain transform** for geodesic

response, the **adaptive manifolds** for Euclidean response, and a mathematical formulation for recursive filtering of **non-uniformly sampled signals**. With these frameworks, we propose several filters that address the main limitations of previous techniques, in addition to providing the fastest performance (both on CPU and GPU) for a variety of real-world applications (Figures 1 to 3). This efficiency comes from their *linear cost* in both the number of pixels and in the dimensionality of the space in which the filters operate. The contributions of this work include:

- The *first geodesic high-dimensional filter* that simultaneously exhibits the following properties: it supports a continuum of scales; its processing time is linear in the number of pixels and independent of the filter parameters, allowing for real-time computations; it correctly handles color images; and it offers control over the kernel's shape.
- The *first Euclidean high-dimensional filter* with linear complexity in both the number of pixels and in the dimensionality of the space in which the filter operates. We also give the first demonstration of a single-pass *hybrid Euclidean-geodesic filter*.
- A mathematical formulation for applying recursive digital filters to non-uniformly sampled signals, whose time complexity is linear in both the number of samples being filtered and in the order of the filter. This formulation enables, for the first time, *geodesic edge-aware* evaluation of *arbitrary* recursive infinite impulse response filters (not only low-pass), allowing for practically unlimited control over the shape of the filtering kernel.

2.1. Impact in the Area

This thesis received the inaugural ACM SIGGRAPH *Outstanding Doctoral Dissertation Award* “in recognition of the potential impact and the notable contributions” (<http://s2016.siggraph.org/content/acm-siggraph-awards>). This is the most prestigious award for a dissertation in the field of Computer Graphics and Interactive Techniques. Furthermore, our research has resulted in two publications at ACM SIGGRAPH (2011 and 2012), and a publication at Eurographics (2015). It is important to note that ACM SIGGRAPH is the world's biggest and most important conference of Computer Graphics, whose proceedings are published as a special edition of *ACM Transactions on Graphics*, the journal with the highest impact index in the area. Similarly, Eurographics is the most important computer graphics conference in Europe, and its proceedings are published as a special edition of the *Computer Graphics Forum* journal.

Our filters and algorithms have been well received by the international community. Our latest publication [Gastal and Oliveira 2015] received the Eurographics 2015 *Best Paper Award Honorable Mention*, conceded by the European Association for Computer Graphics. Furthermore, the papers describing our filters, despite of being recently published, have already received a few hundred citations, according to Google Scholar. In particular, our filters have already been put to use by researchers at Adobe, NVIDIA, Google, and Disney, in applications such as color grading [Bonneel et al. 2013], enforcing temporal consistency in video effects [Lang et al. 2012], synthetic defocus for smartphone cameras [Barron et al. 2015], and viewfinder editing for digital cameras [Baek et al. 2013]. Perhaps a better indication of this interest in our algorithms is the fact that third-party implementations of the Domain-Transform and Adaptive-Manifold filters have been included in [OpenCV 2016], a widely used library in computer vision and image processing.

In conclusion, the filters introduced by this dissertation provide a valuable tool for

the image and video processing, computer graphics, computer vision, and computational photography communities.

3. Overview of our High-Dimensional Filters

Due to space restrictions, this article cannot present all the details of the developed filters and algorithms. The complete text of the dissertation is available at <http://hdl.handle.net/10183/118258>, and the homepages of our three publications [Gastal and Oliveira 2011, Gastal and Oliveira 2012, Gastal and Oliveira 2015] may be found at <http://inf.ufrgs.br/~eslgastal>. We suggest that the readers watch the videos available in these homepages for more intuitive descriptions and examples of applications of the developed techniques.

Geodesic Filtering In [Gastal and Oliveira 2011], we introduced the *Domain Transform*, a state-of-the-art technique for geodesic filtering of images and videos. The innovation behind this method is the use of a dimensionality reduction strategy for efficient implementation of high-dimensional filters. Its inner workings are based on the interpretation of images as manifolds embedded in high-dimensional spaces. For instance, an RGB color image is a 2D manifold embedded in 5D space: x, y, r, g, b ; and a manifold is simply the generalization of a surface to high-dimensional spaces. Our idea is to map curves on this surface to a lower-dimensional space, in a way that the geodesic distances between points on the curve is preserved. For this, we derived an isometric (distance-preserving) domain transformation $t(x)$, which maps image-space pixel positions to a new space which preserves the original distances measured in high-dimensional space:

$$t(x) = \int_0^x 1 + \frac{\sigma_s}{\sigma_r} \sum_{k=1}^{d_{\mathcal{R}}} \left| \frac{d}{du} f_{[k]}(u) \right| du. \quad (1)$$

In [Gastal and Oliveira 2011], we present all the necessary steps for the derivation of this formula, along with mathematical proofs of its properties. The variables σ_s and σ_r are parameters of the filter, while $f(u)$ represents the image being filtered. Equation (1) allowed us to demonstrate real-time edge-preserving color filtering of high-definition videos. Furthermore, ours is the first geodesic filtering technique capable of working on color images at arbitrary scales in real time, without resorting to subsampling or quantization. Figure 1 illustrates the flexibility of our domain transform by using it to obtain various effects from the original photograph in Figure 1(a).

Euclidean Filtering In [Gastal and Oliveira 2012], we introduced the *Adaptive Manifolds*, a state-of-the-art technique for Euclidean filtering of images and videos. The central idea of this technique consists in subdividing the high-dimensional space into low-dimensional manifolds which adapt to the signal being filtered. Filtering is then performed independently and efficiently *on top of* each low-dimensional manifold, and the results interpolated to approximate the desired filter. Mathematically, a naive *quadratic-cost* (thus inefficient) Euclidean filter is given by $g_i = \sum_{p_j \in \mathcal{S}} \phi_{\Sigma}(\hat{p}_i - \hat{p}_j) f_j$, where ϕ is a Gaussian function, f_j are the colors of the pixels in the input image, g_i are the colors of the output pixels, and \hat{p}_i are the high-dimensional positions of the pixels. We accelerate this filter by evaluating its response on a reduced set of sampling points, and using these values to interpolate the filter's response at all input pixels. Our filter is thus defined by:

$$g_i = \sum_{k=1}^K w_{ki} \sum_{p_j \in \mathcal{S}} \phi_{\Sigma_S}(\hat{p}_i - \hat{p}_j) \phi_{\Sigma_{\mathcal{R}/2}}(\eta_{ki} - f_j) f_j,$$

where w_{ki} are weights derived from the Gauss-Hermite quadrature. With this formulation, we present a filter which is faster and requires less memory than previous approaches, being the first high-dimensional Euclidean filter with *linear cost* in both the number of pixels and in the dimensionality of the space. In [Gastal and Oliveira 2012], we present a derivation for the equations that define our method, providing a theoretical justification for the technique and for its properties. Furthermore, the flexibility of our approach allows for the first demonstration of a hybrid Euclidean-geodesic filter that runs in a single pass. Figure 2 shows examples of Euclidean-filtering applications using the adaptive manifolds.

Non-Uniform Filtering More recently [Gastal and Oliveira 2015], we introduced a mathematical formulation for applying recursive digital filters to *non-uniformly sampled signals*. This formulation enables, for the first time, geodesic edge-aware evaluation of *arbitrary* recursive infinite impulse response digital filters (not only low-pass), which allows practically unlimited control over the shape of the filtering kernel. For this, we decompose a P -th order filter into a set of 1st-order ones, and we then derive the equations for the individual 1st-order filters in a non-uniform domain. Traditionally, a 1st-order recursive filter (with standard uniform sampling) is defined by the equation $g[k] = a f[k] + b g[k-1]$, where a and b are the coefficients of the filter, $f[k]$ are the input samples, and $g[k]$ the output ones. In [Gastal and Oliveira 2015], we extend this equation to support non-uniform sampling, which results in the following expression:

$$g[k] = a f[k] + b^{\Delta t_k} g[k-1] + \left(\frac{b^{\Delta t_k} - 1}{r_0 \Delta t_k} - r_1 b \right) f[k] - \left(\frac{b^{\Delta t_k} - 1}{r_0 \Delta t_k} - r_1 b^{\Delta t_k} \right) f[k-1],$$

where r_0 and r_1 are constants, and $\{\Delta t_k\}$ is the set containing the non-uniform distances between samples of the signal. This formulation allows for greater control over the desired responses of the filters being used, which can then be adapted to specific applications. As examples, we provide the first demonstrations of a *variety of high-dimensional geodesic filters*, including Gaussian, Laplacian of the Gaussian, and low/high/band-pass Butterworth and Cauer filters, among others. Figure 3 illustrates some of the possible detail-enhancement effects generated by such responses. By allowing the use of arbitrary filters, our method enables the creation of a new rich variety of effects for image and video applications, many of which were not previously within the reach of previous techniques.

References

- Adams, A., Baek, J., and Davis, M. A. (2010). Fast high-dimensional filtering using the permutohedral lattice. *Computer Graphics Forum*, 29(2):753–762.
- Baek, J., Pajak, D., Kim, K., Pulli, K., and Levoy, M. (2013). Wysiwyg computational photography via viewfinder editing. *ACM TOG*, 32(6):198:1–198:10.
- Barron, J. T., Adams, A., Shih, Y., and Hernández, C. (2015). Fast Bilateral-Space Stereo for Synthetic Defocus. In *IEEE CVPR*, pages 4466–4474.
- Bauszat, P., Eisemann, M., and Magnor, M. (2011). Guided image filtering for interactive high-quality global illumination. *Computer Graphics Forum*, 30(4):1361–1368.
- Bonneel, N., Sunkavalli, K., Paris, S., and Pfister, H. (2013). Example-based video color grading. *ACM TOG*, 32(4):39:1–39:12.
- Buades, A., Coll, B., and Morel, J. (2005). A non-local algorithm for image denoising. In *IEEE CVPR*, volume 2, pages 60–65.

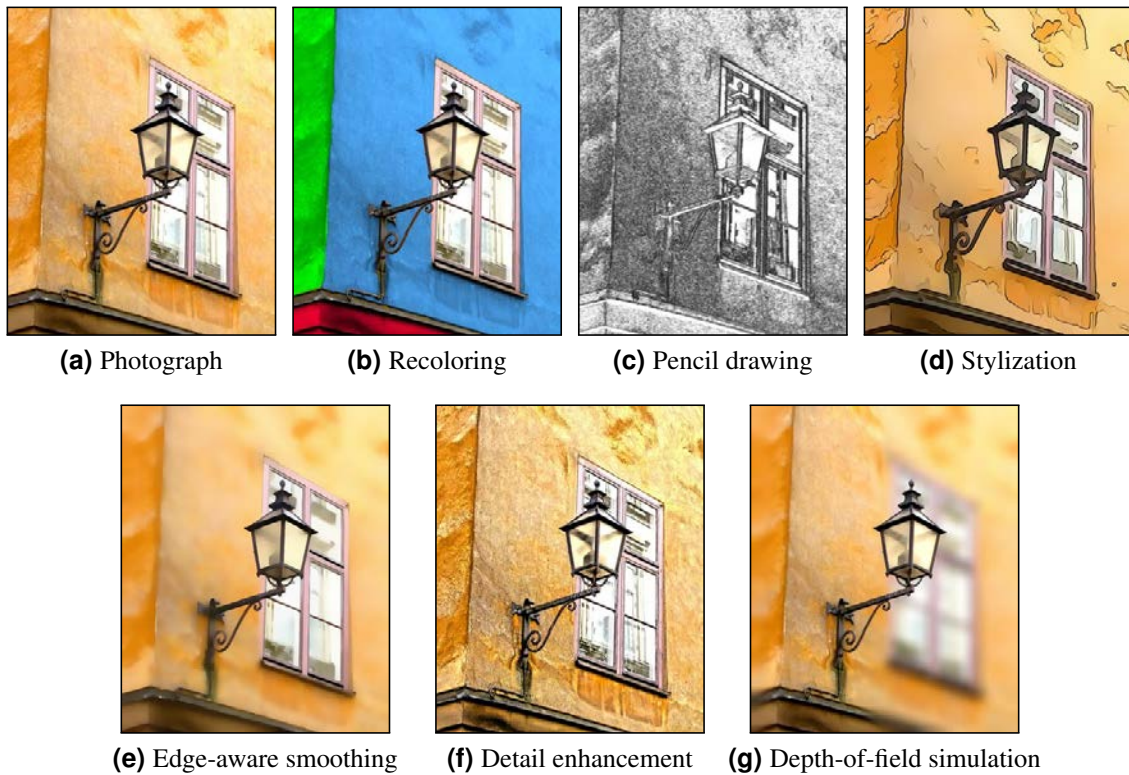


Figure 1. A variety of effects illustrating the versatility of our *domain transform* [Gastal and Oliveira 2011] when used to filter the photograph in (a).

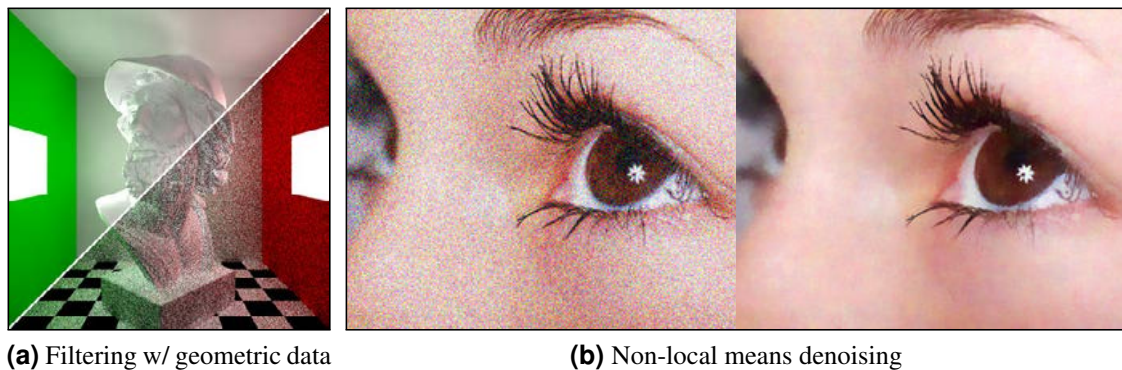


Figure 2. Examples of filtering results produced with our adaptive-manifold filter [Gastal and Oliveira 2012]. (a) Filtering (performed in 8 dimensions) of a noisy undersampled image generated using path tracing. The split rendition compares the input (bottom right) and the filtered result (top left). Note the smoothness of the shading. (b) Denoising of natural images using non-local-means (performed in 27 dimensions). Notice the noise reduction while retaining fine details.

Fattal, R. (2009). Edge-avoiding wavelets and their applications. *ACM TOG*, 28(3):22.

Gastal, E. S. L. and Oliveira, M. M. (2011). Domain transform for edge-aware image and video processing. *ACM TOG*, 30(4):69:1–69:12. Proc. of SIGGRAPH 2011.

Gastal, E. S. L. and Oliveira, M. M. (2012). Adaptive manifolds for real-time high-dimensional filtering. *ACM TOG*, 31(4):33:1–33:13. Proc. of SIGGRAPH 2012.

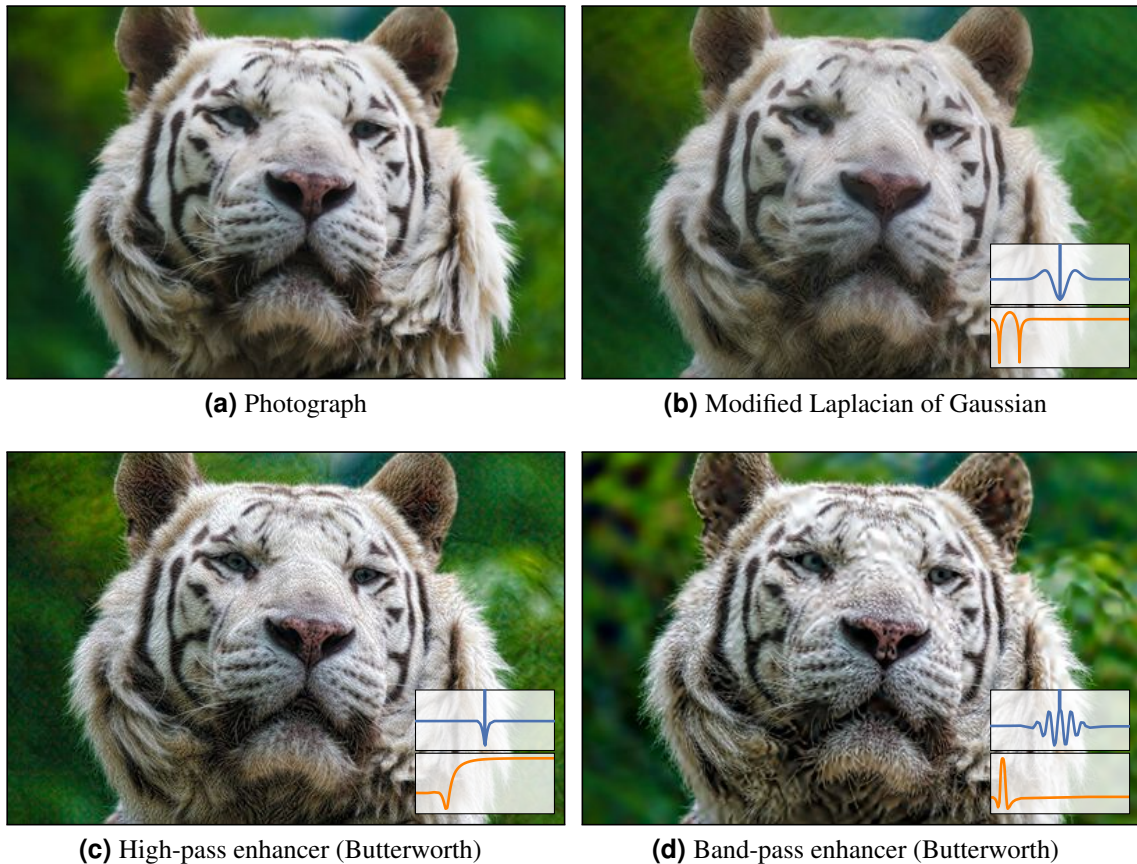


Figure 3. A variety of high-order recursive filters applied by our method for non-uniform filtering [Gastal and Oliveira 2015] to the photograph in (a). The resulting filters preserve the image structure and do not introduce visual artifacts such as halos around objects. The graphs in the insets show the filter's impulse response in blue, and its frequency response (Bode magnitude plot) in orange.

- Gastal, E. S. L. and Oliveira, M. M. (2015). High-order recursive filtering of non-uniformly sampled signals for image and video processing. *Computer Graphics Forum*, 34(2):81–93. Proc. of Eurographics 2015.
- Kopf, J., Cohen, M. F., Lischinski, D., and Uyttendaele, M. (2007). Joint bilateral upsampling. *ACM TOG*, 26:96:1–96:5.
- Lang, M., Wang, O., Aydin, T., Smolic, A., and Gross, M. (2012). Practical temporal consistency for image-based graphics applications. *ACM TOG*, 31(4):34:1–34:8.
- OpenCV (2016). Open Source Computer Vision library 3.0. <http://opencv.org>.
- Smith, S. M. and Brady, J. M. (1997). SUSAN – a new approach to low level image processing. *International Journal of Computer Vision*, 23(1):45–78.
- Winnemöller, H., Olsen, S. C., and Gooch, B. (2006). Real-time video abstraction. *ACM TOG*, 25(3):1226.
- Yang, Q., Tan, K. H., and Ahuja, N. (2009). Real-time O(1) bilateral filtering. In *CVPR*, pages 557–564.