Image Denoising by Regularization on Characteristic Graphs

Thomas J. Asaki* and Pavlo Cherepanov+ and Matthew Sottile† and Kevin R. Vixie#

* *Computational Physics and Methods, Los Alamos National Laboratory, MS D413, Los Alamos, NM 87545 and Center for Data-Driven Modeling and Analysis, New Mexico Consortium, Los Alamos, NM 87545*

+ *Department of Mathematics and Statistics, MSC03 2150, 1 University of New Mexico, Albuquerque, New Mexico, 87131-0001*

† *Department of Computer Science, 1202 University of Oregon, 120 Deschutes Hall, Eugene, OR 97403-1202*

Mathematical Modeling and Analysis, Los Alamos National Laboratory, MS B284, Los Alamos, NM 87545 and Center for Data-Driven Modeling and Analysis, New Mexico Consortium, Los Alamos, NM 87545

Abstract

We present the concept of characteristic graph representations of images. Two such non-parametric graph constructions are detailed. Their use in image denoising is demonstrated within a regularization framework. The results are compared with those of more traditional approaches of total variation and Tikhonov regularization. We show that in some denoising scenarios our methods perform more favorably in preserving intensity levels and object boundary shape details.

Key Words: Image Restoration, Denoising, Non-Parametric, Graph Regularization.

1 Introduction

We consider an interesting and useful modification to the practice of image denoising by functional minimization methods. A standard formulation is to impose a regularization penalty weighted with respect to a data fidelity measure. The discrete realization is the problem

$$u* = \arg\min_u \left(\|u - f\|_q^q + \alpha \|\nabla u\|_p^p \right). \tag{1}$$

$f \in \mathbb{R}^n$ is a noisy image of n pixels, $1 \leq p, q < \infty$ (typically), u and $u*$ are candidate and optimal denoised images, respectively. The data fidelity and regularization terms are weighted by a scalar α.

The popular L^2 realization of Eq. 1 ($p = 2, q = 2$) was first implemented in the work of Tikhonov (See [18]). The regularization enforces smoothness in u and is equivalent to a finite-time application of the heat equation to the noisy image. The net effect is to reduce all high-frequency content such as noise. One undesirable feature is that sharp intensity boundaries are smoothed.

The total variation (TV) realization Eq. 1 ($p = 1$, $q = 2$), or ROF model [16], gained rapid popularity from its ability to produce denoised images that retain sharp intensity boundaries. While successful for visual presentation, there are known side effects [17]. First, absolute intensity levels are altered. This is most obvious for images of piecewise constant intensity. Second, intensity regions of smooth variation become stair-stepped. Third, shapes in images are altered by elimination of high boundary curvature, reducing object perimeter.

While Eq. 1 has been highly successful and ubiquitously applied in image processing, the side effects arising from the regularization and choice of p are problematic for many applications. Several approaches have been taken in order to address these issues. Additional regularization based on object boundary lengths was proposed

Correspondence to: <asaki@lanl.gov>, Los Alamos National Laboraotory Document LA-UR-08-0116.

by Mumford and Shah [13]. This approach requires an additional functional weight parameter and assumes objects constrained by a smooth and minimal perimeter.

Another approach proposed a variable p. See [3] and the developments in [10, 8, 4]. In this approach p is made a function of the gradient of the image. While this adds to the analytic complexity of the functional, the stair-stepping is reduced.

The L^1TV approach, introduced in its continuous version by Chan and Esedoḡlu in [6] after its initial discrete study by Alliney [2] and Nikolova [15], does a much better job of contrast preservation. This particular approach has generated continued interest; see for example [1, 12, 19].

Recently, Chartrand [7] has shown that solving the nonconvex optimization problem for $0 < p < 1$ yields improved results for shape and intensity preservation relative to TV. These results hold for images that can be represented sparsely in the gradient, that is, piecewise constant intensities.

There are many other methods which can be found in the literature, each of which attempts to improve on the initial inspiration of the Rudin-Osher-Fatemi method. Most require some input in the form of parameters that, in effect, chooses what the noise level is. This is not unreasonable since this can be chosen from knowledge of the noise level in the image when this information is available. But it seems to us a useful endeavor to see what can be done without using such parameters.

We introduce the concept of non-parametric characteristic graphs and introduce two realizations. Graphs are computed without assumptions on image content. We illustrate their use for simple image denoising and compare our results with standard methods. The characteristic graph approach imposes no additional parameters on the denoising process.

2 Approach

Consider the discretized image gradient of Eq. 1. Let $u_{i,j}$ be the image intensity at image pixel location (i, j). The standard forward difference gradient approximation is

$$\nabla u_{i,j} \equiv \sqrt{(u_{i+1,j} - u_{i,j})^2 + (u_{i,j+1} - u_{i,j})^2}. \tag{2}$$

The first (second) difference term in Eq. 2 is the vertical (horizontal) gradient. We impose the Neumann boundary condition $\nabla u_{i,j} = 0$ if either i or j causes $u_{i,j}$ to fall outside of the image.

This gradient can be represented as an operation on the graph $G_0 = \{V_0, E_0\}$, the non-directed graph representation of u with vertex set V_0 containing all image pixels and edge set E_0 containing all pixel pairs that share a pixel edge (*not a graph edge*). This graph has been called the grid graph [5] and is recognized as that of a fully connected Von Neumann neighborhood (see [14]). In addition we consider the graph edges to have weights given by the absolute value of the image intensity difference associated with the vertex pair connected by the edge. Quite simply, G_0 is the fully-connected nearest (geographic) neighbor graph in the 2-d image space. Figure 1(b) illustrates G_0 for a small sample image.

The denoising problem is now represented as

$$u* = \arg\min_u \left(\|u - f\|_q^q + \alpha \|\nabla_G u\|_p^p \right). \tag{3}$$

where the gradient operator is applied to a graph G with weights W. The fundamental graph, $G = G_0$, is the one typically applied. We will consider subgraphs $G = \{V = V_0, E \subset E_0\} \subset G_0$ which we consider to be *characteristic* of images for denoising purposes. In the next section we will detail two such graphs.

The optimization problem described in 3 is nice: it is convex for $(1 \leq p, q \leq \infty)$ and strictly convex for $(1 < q < \infty)$. Though the regularization term suffers from a degeneracy yielding a null space with dimension equaling the number of components in the graph G, the data fidelity term keeps the functional coercive. Our choice of $q = 2$ keeps everything strictly convex.

In practice, the gradient computation is carried out over the fundamental graph G_0 with imposed zero edge weights that effectively define the subgraph of interest. More precisely,

$$\|\nabla_G u\|_p^p = \sum_{w \in W} |w|^p. \tag{4}$$

Note in particular that we are using the anisotropic version* of equation 3 here: this is the only version that makes sense on the graph – we do not have differences in the x and y directions at every node in the reduced graph G unless the graph is the complete graph G_0.

We utilize the Lagged-Diffusivity method for solving Eq. 3 together with an application of the discrepancy principle for choosing the constant α [20]. We choose this approach for its general applicability to $p > 0$, $q > 0$, though we limit our presentation to the cases $p \in \{1, 2\}$ and $q = 2$. Lezoray, et al. [11] present a general method for solving Eq. 3 on graphs of arbitrary topology. They also present a variety of results for color image denoising using the fundamental graph G_0. Their focus is on computational aspects of very general discrete graph-based denoising. Our main objective is the construction and application of characteristic graphs that are natural choices for image processing applications.

3 Characteristic Graphs

In this section we present two new characteristic image graphs, detail how they are constructed, and briefly discuss our expectations in denoising applications. An example of each graph is shown in Figure 1 as computed from a simple example image.

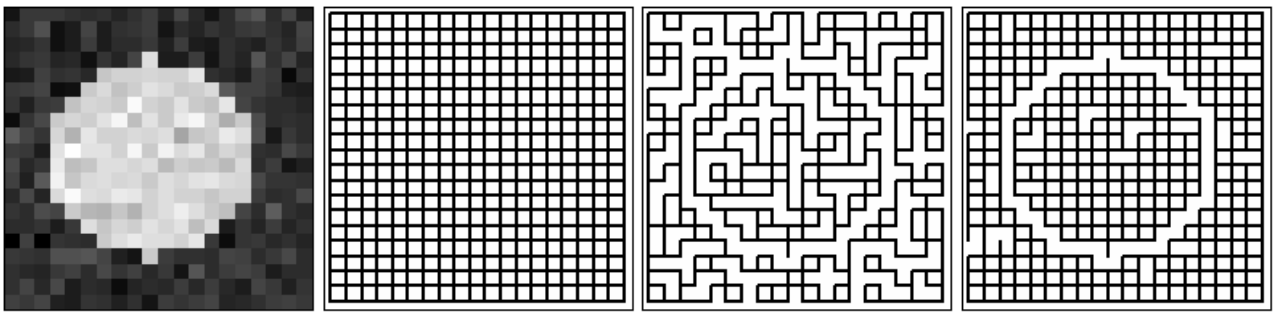

Figure 1: Left to right: A simple noisy test image; and the corresponding three characteristic graphs: the full graph or grid graph G_0, the truncated Kruskal algorithm graph with next-nearest neighbor edge extension K_1, and the vertex inclusion graph, K_2.

In particular, our subgraphs will have the following properties.

1. A graph should preserve spatial information. Distinct image regions should only be able to influence each other through neighbor pixel paths through the graph. By choosing graphs as subsets of G_0 this requirement is automatically met.

2. A graph should naturally detect image intensity content. In image regions of smooth intensity variation, the graph should be highly connected. Conversely, graphs should have few or no connections across edges of large weight.

3. Graph construction should be non-parametric utilizing *only* the image intensity information without relying on user judgement or prior information.

*IN the case of $G = G_0$, the isotropic version is given by $\sum_i ((u_x^i)2 + (u_y^i)2)^{p/2}$ instead of the anisotropic analog given by $\sum_i (|u_x^i|^p + |u_y^i|^p$

3.1 K_1: Truncated-Kruskal Graph

We introduce a graph K_1 based on Kruskal's algorithm for constructing a minimum spanning tree of an arbitrary graph[9]. The tree construction is prematurely terminated when all vertices are included and then additional edges are included to improve graph density.

1. Begin with an empty graph $V = \{\}$ and $E = \{\}$.
2. Add to E the edge of minimum weight in E_0 that does not create a cycle in $\{V, E\}$. Add the corresponding vertices to V that are not already in V.
3. Repeat step 2 until $V = V_0$.
4. Perform step 2 one more time. If a new edge is added set w equal to the edge weight associated with this edge and then remove the edge from E. If a new edge was not added set $w = \infty$.
5. To each vertex of degree one add to E the associated edge from E_0 that (a) has smallest weight, (b) is not already in E, and (c) if the edge weight is less than w.

The first four steps of the algorithm retain the time complexity of Kruskal's algorithm, $O(|E|\ log\ |V|)$. The last step requires a $O(|E|\ log\ |E|)$ sorting step to determine the minimum weight edge connected to each vertex. Due to the special structure of the grid graph derived from the image, we know that $|E|$ is bounded from above by $4|V|$, so this complexity is equivalent to that of Kruskal's algorithm.

The characteristic graph K_1 will attempt to prevent graph connections across pixel neighbors of significantly different intensities. The density may be low relative to the grid graph G_0, and the connectivity even in regions of like intensity may be circuitous or lacking. The third image in Figure 1 illustrates these properties. Note the small 4-pixel disconnected subgraph near the upper left corner. Note also the somewhat circuitous graph paths one must often take to traverse from one pixel to its neighbor even if the intensities are similar. The best graph property is the absence of any graph edges that would connect object regions to background regions.

3.2 K_2: Vertex Inclusion Graph

We also introduce the Vertex Inclusion Graph K_2 in order to address the potential drawbacks of K_1, low density and significant dis-connectivity. K_2 is constructed by including all edges of weight less than or equal to a cutoff value non-parametrically determined from the Kruskal construction.

1. Begin with an empty graph $V = \{\}$ and $E = \{\}$.
2. Add the edge of smallest weight from E_0 to E that is not already in E, and the corresponding vertices from V_0 to V that are not already in V.
3. Repeat step 2 until $V = V_0$.
4. Add all edges from E_0 to E that are of equal weight to the largest edge weight in E.

Graph K_1, like graph K_2 can be disconnected across image intensity discontinuities. Also, we expect a significant number of graph edges in regions where the intensity variations are governed by noise. These properties are evident in the fourth image of Figure 1. Note both the absence of edges that cross the object-background boundary and the high graph density relative to K_1. For very noisy images or images with salt and pepper noise, K_2 can produce a nearly fully connected graph $K_2 \sim G_0$. In these cases, there is no expected advantage to using K_2. Like the construction of K_1, this approach has logarithmic time complexity of $O(|E|\ log\ |E|)$ due to sorting the edges by edge weight.

3.3 Discussion on Parametrization

These graphs are considered to be *characteristic* representations of images because they satisfy, in our estimation, the criteria given at the beginning of this section. One might argue that these graph constructions contain hidden parameters. For example, why did we choose to add edges only to leaf vertices in forming K_1? And, why did we specify only second nearest neighbors instead of, say, third nearest? Parameters such as these are of a different nature than more intrusive and application specific parameters such as edge weight thresholding, graph density thresholding, and local graph construction based on prior assumptions. We contend that our choices are related primarily to the desired structure of the graphs independent of the consituent intensity distributions of the images from which they are derived. Our graphs are intended for very general application because they capture spatial and intensity connectivity for any type of image.

4 Test Images

We consider four test images chosen to test a variety of denoising situations. These reference images are shown in Figure 2. Each image is 200 by 200 pixels.

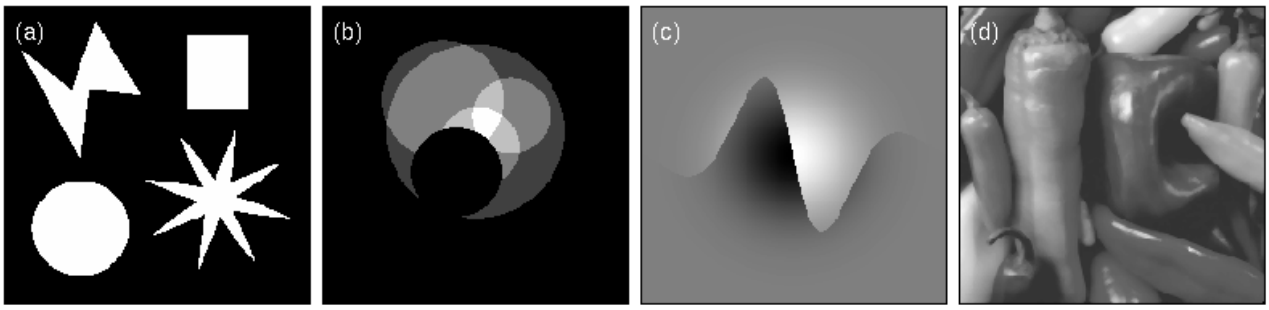

Figure 2: The four test images (before addition of noise): (a) simple shapes delineated by binary intensity boundaries, (b) multiple circle object of five distinct, constant, and equally-spaced intensities, (c) a Gaussian bump with a sign change across a modulated sinusoid boundary, and (d) a natural image of a collection of peppers.

Figure 2(a) shows test image T_1. The pixel intensity is unity inside each of four objects and is zero on the background. The four objects have structure on a variety of length scales from a single pixel to several tens of pixels. They were chosen to test denoising algorithms on both intensity and object shape preservation.

Figure 2(b) shows test image T_2. The pixel intensities belong to the discrete set $\{0.00, 0.25, 0.50, 0.75, 1.00\}$. The object of interest is a combination of circular individuals. This image was also chosen to test denoising algorithms on both intensity and shape preservation.

Figure 2(c) shows test image T_3. This image has smooth intensity variation everywhere except along a curved boundary. The intensity jump varies from near zero at the left and right image edges to a value of 1 at the image center. The upper (lower) portion of the image is a positive (negative) Gaussian bump. The curved boundary is a sine function modulated by an exponential decay away from the image center. This image was chosen to test denoising algorithms on ability to handle images with both continuous and discontinuous intensity variations.

Figure 2(d) shows test image T_4. This natural image of a collection of peppers was chosen to test denoising algorithms on ability to handle simple realistic images.

5 Results

We now present some specific results that illustrate the denoising concepts of intensity level and object shape preservation.

5.1 TO1

Figure 3 shows TV ($p = 1$) denoising results for the binary test image T_1. The upper left image (a) shows T_1 with additive Gaussian noise of $\sigma^2 = 0.10$. The remaining images show the results of denoising via three characteristic graphs: (b) the full graph G_0, the standard TV approach, (c) K_1 and (d) K_2. All three methods produce visually pleasing results. Noise is significantly reduced and the general shapes are visually preserved. Differences between the various results are more readily seen in Figure 4 which shows the image subregions of the lower part of the square object and upper part of the star object. The G_0-denoised image is now seen to have somewhat fuzzy boundaries. This is a consequence of graph connections across object boundaries combined with noise. A larger value of α would sharpen the boundaries at the cost of reducing image contrast. The K_1 and K_2-denoised images have more well-defined object boundaries. This is due to limited or few graph connections across the object boundary. The K_1 recovered image does show some contrast reduction along the arms of the star. This is likely due to low graph density naturally expected in thin regions of object or background.

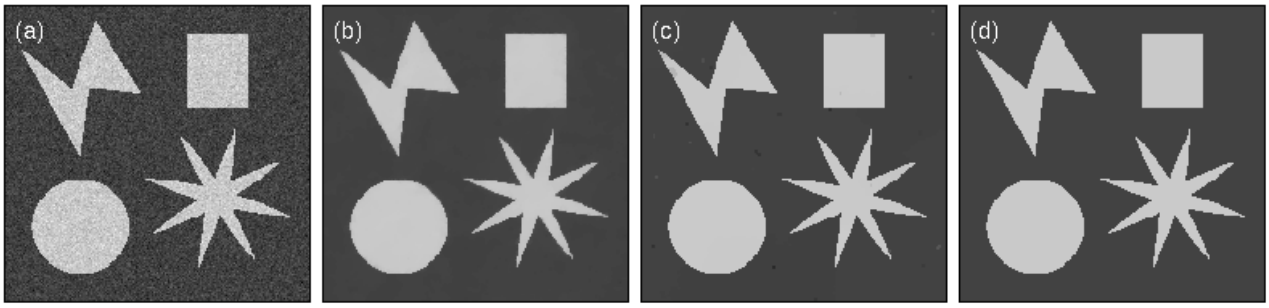

Figure 3: TV denoising example for test object T_1. The test image (a) is the clean image of Figure 2(a) corrupted by a $\sigma^2 = 0.10$ additive Gaussian process (note the grayscale change). Recovered images are based on three graphs: (b) G_0, (c) K_1 and (d) K_2.

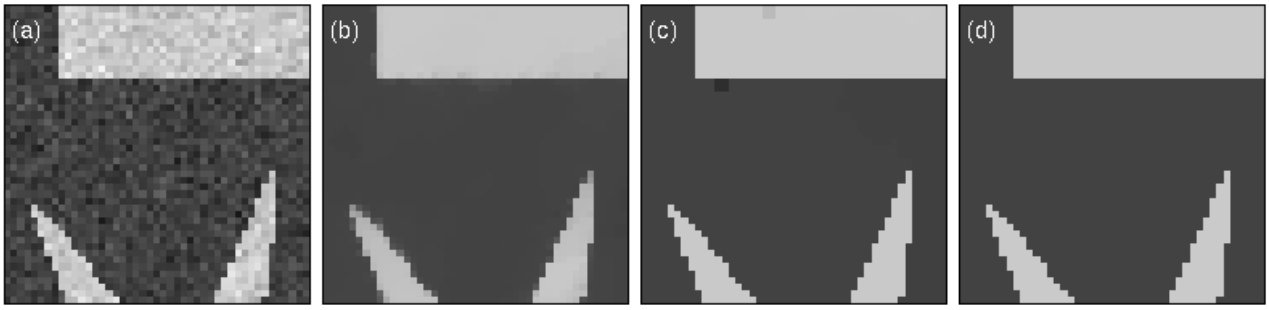

Figure 4: TV denoising example for test object T_1. The images are zoomed in representations of those in Figure 3 to a region focused near the top of the star shape.

K_1 and K_2 denoising performs better than G_0 at recovering object shape details. Significant loss of contrast

with G_0 in regions of high boundary curvature prevents the usual method from performing well. The reduced graph methods maintain reasonable contrast even for object regions of single pixel width.

Intensity level preservation is also improved with the use of K_1 and K_2 over G_0. For images I and T of the same size and pixel index i, let $I^{T,v} \equiv mean\{I_i \mid T_i = v\}$. Also, let nT indicate a noisy version of image T and let T^G indicate an image T denoised by via graph G. Table 1 shows an intensity comparison for T_1 denoising examples. The standard TV denoising reduces the binary intensity contrast by 3.1% while the characteristic graph methods show essentially no contrast reduction.

I	$I^{(T_1,0)}$	$I^{(T_1,1)}$
nT_1	0.001	1.000
$nT_1^{G_0}$	0.008	0.977
$nT_1^{K_1}$	0.002	1.001
$nT_1^{K_2}$	0.001	1.001

Table 1: Intensity preservation comparison for T_1 denoising examples.

5.2 TO2

Figure 5 shows TV ($p = 1$) denoising results for the piecewise constant intensity test image T_2. The upper left image (a) shows T_2 with additive Gaussian noise of $\sigma^2 = 0.05$. The remaining images show the results of denoising via three characteristic graphs: (b) the full graph G_0, the standard approach, (c) K_1 and (d) K_2. Again, all three methods produce visually pleasing results. Noise is significantly reduced and the general shapes are visually preserved. Differences between the various results are more readily seen in Figure 6 which shows the image subregions of the lower left arm of the object. Similar to the T_1 example, the G_0-denoised image is now seen to have somewhat fuzzy boundaries. The K_1 and K_2-denoised images have more well-defined object boundaries. Note that the noise level in this case is a larger in relation to some of the intensity discontinuities than for the T_1 examples.

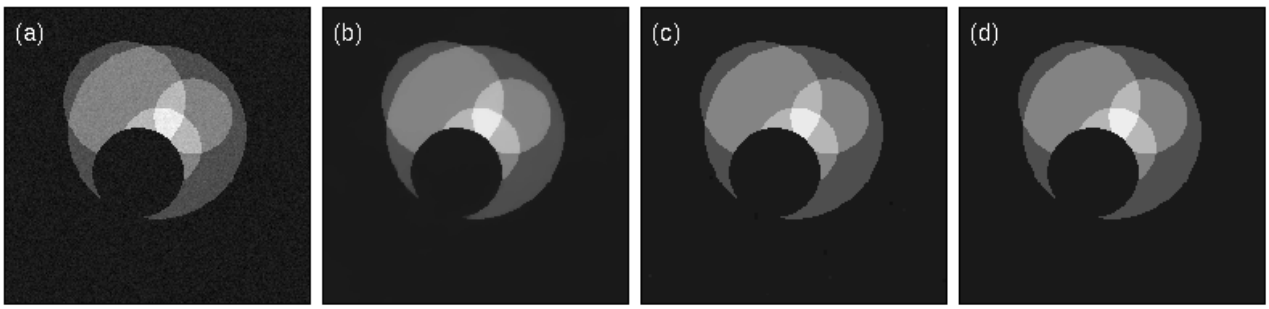

Figure 5: TV denoising example for test object T_2. The test image (a) is the clean image of Figure 2(b) corrupted by a $\sigma^2 = 0.05$ additive Gaussian process (note the grayscale change). Recovered images are based on three graphs: (b) G_0, (c) K_1 and (d) K_2.

Again we find that K_1 and K_2 denoising performs better than G_0 at recovering object shape details. Significant loss of contrast with G_0 in regions of high boundary curvature prevents the usual method from performing well. The reduced graph methods maintain reasonable contrast even for object regions of single pixel width.

Intensity level preservation is also improved with the use of K_1 and K_2 over G_0. Table 2 shows an intensity comparison for T_2 denoising examples. K_1 and K_2 denoising perform better than the standard G_0 approach.

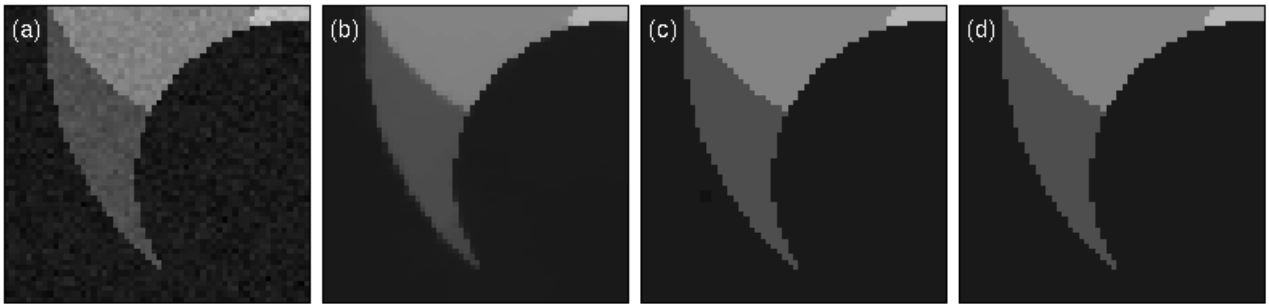

Figure 6: TV denoising example for test object T_2. The images are zoomed in representations of those in Figure 5 to a region focused on the left side of the object.

I	$I^{(T_1,0)}$	$I^{(T_1,0.25)}$	$I^{(T_1,0.50)}$	$I^{(T_1,0.75)}$	$I^{(T_1,1)}$
nT_2	0.000	0.249	0.500	0.751	1.000
$nT_2^{G_0}$	0.001	0.246	0.498	0.743	0.984
$nT_2^{K_1}$	0.000	0.249	0.502	0.744	0.992
$nT_2^{K_2}$	0.000	0.249	0.501	0.750	1.000

Table 2: Intensity preservation comparison for T_1 denoising examples.

5.3 TO3

Figure 7 shows $p = 2$ denoising results for test image T_3. The upper left image (a) shows T_3 with additive Gaussian noise of $\sigma^2 = 0.03$. The remaining images show the results of denoising via three characteristic graphs: (b) the full graph G_0 which represents the standard approach, (c) K_1 and (d) K_2. Again, all three methods produce visually pleasing results. Noise is significantly reduced and the general shapes are visually preserved. However, even at this magnification, differences in the three solutions are readily apparent to the eye. G_0 produces a relatively noisy image with a smoothed, although narrow, intensity transition region along the sinusoid boundary. K_1 preserves the boundary very well, but suffers from a stair-stepping effect associated with locally disconnected graph regions. K_2 performs the best of the three with good noise reduction and intensity discontinuity preservation. These observations are seen even more clearly in Figure 8.

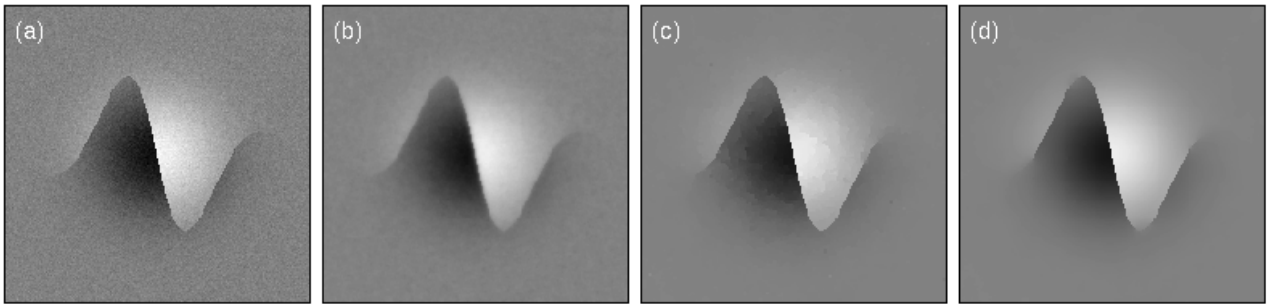

Figure 7: $p = 2$ denoising example for test object T_3. The test image (a) is the clean image of Figure 2(c) corrupted by a $\sigma^2 = 0.03$ additive Gaussian process (note the grayscale change). Recovered images are based on three graphs: (b) G_0, (c) K_1 and (d) K_2.

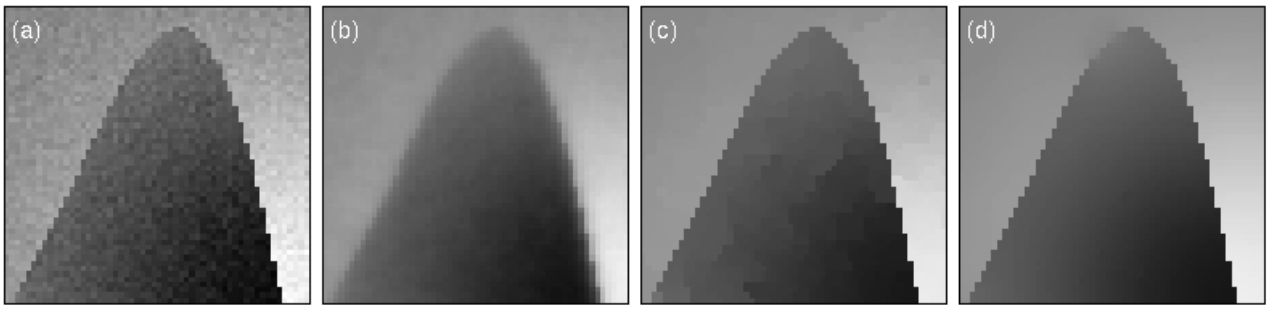

Figure 8: TV denoising example for test object T_3. The images are zoomed in representations of those in Figure 7 to a region focused on the upper left intensity discontinuity.

5.4 TO4

Finally, we consider the natural image T_4 denoised using $p = 2$. The results are shown in Figure 9. The upper left image (a) shows T_4 with additive Gaussian noise of $\sigma^2 = 0.05$. The remaining images show the results of denoising via three characteristic graphs: (b) the full graph G_0 which represents the standard approach, (c) K_1 and (d) K_2. The results are markedly different from the previous test cases. The G_0 and K_2 results are essentially identical because, in this case K_2 differs from G_0 only by the removal of two graph edges. This can happen when one or more pixels are largely isolated (in intensity) from all of its G_0 graph neighbors. In order to include these isolates, the K_2 graph becomes very dense. The K_1 result clearly shows graph-region intensity stepping.

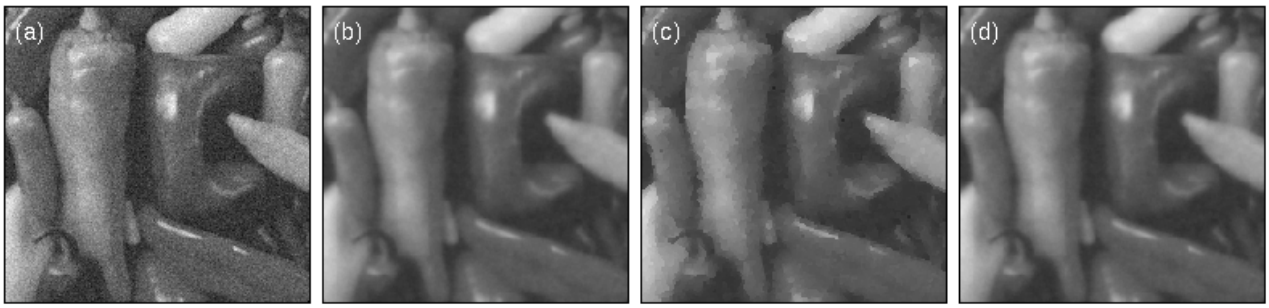

Figure 9: $p = 2$ denoising example for test object T_4. The test image (a) is the clean image of Figure 2(d) corrupted by a $\sigma^2 = 0.05$ additive Gaussian process (note the grayscale change). Recovered images are based on three graphs: (b) G_0, (c) K_1 and (d) K_2.

6 Discussion

The characteristic graphs K_1 and K_2 are especially useful for denoising image of piecewise smooth content. They help to preserve intensity levels because intensity jumps are not penalized if they large relative to the characteristic weight determined by the truncated Kruskal algorithm. They also help to preserve object boundary shape details for much the same reason. Intensity jump boundaries have zero penalty regardless of length or curvature. Object details can be preserved down to the pixel level. Even in the non-ideal case, where noise has allowed the Kruskal algorithm to build links across intensity discontinuities, these links are sparse relative to a full graph implementation and still serve to reduce distortion effects.

Our characteristic graphs will fail to produce significantly improved results in some cases. If images are very noisy, are corrupted by salt and pepper noise, or contain intensity isolated pixels, then we expect their characteristic graphs to be nearly as dense as the full graph. This simply means that the image information needed to construct a good subgraph is lacking. We note that our results are not expected to be inferior to methods that employ a full graph.

The use of characteristic image subgraphs is general to PDE-based methods for image processing. We have defined here example graphs and demonstrated their potential using some simple denoising examples. Clear extensions are applications to segmentation and texture extraction.

We have focused exclusively on non-parametric modifications to Equation 1. Other important graphs are certainly possible that make use of (problem-dependent) parameters. Some examples include graphs that achieve a certain density or connectivity, or user-defined graph construction cutoff values. We also note that there are certain possible parametric modifications to the denoising procedure. For example, one is tempted to employ a variable α procedure in which the characteristic graph is updated as α is incrementally raised from some low value to a final value. This may help to reduce graph connectivity across boundaries significantly obscured by noise.

7 Conclusion

We have defined two characteristic image graphs, the truncated Kruskal graph K_1 and the vertex inclusion graph K_2. We have demonstrated their use by application to three simple test images and one natural test image. Comparisons were made with standard TV and L^2 denoising methods. Results show that denoising by these graphs improves intensity level and object shape preservation in images with piecewise smooth content. We point out several areas for future investigation including applications beyond denoising and parametric extensions.

8 Acknowledgements

This research was performed, in part, under the auspices of the Department of Energy under contract W-7405-ENG-36. It was funded at Los Alamos National Laboratory through the 2007 Data Driven Modeling and Analysis summer school jointly funded by the T-07 summer student program and the New Mexico Consortium.

References

[1] W. K. ALLARD, *Total variation regularization for image denoising; I. Geometric Theory.*, To appear in SIAM Journal on Mathematical Analysis, (2007).

[2] S. ALLINEY, *A property of the minimum vectors of a regularizing functional defined by means of the absolute norm*, IEEE Trans. Signal Process., 45 (1997), pp. 913–917.

[3] P. BLOMGREN, T. F. CHAN, P. MULET, AND C. WONG, *Total variation image restoration: Numerical methods and extensions*, in Proceedings of the 1997 IEEE International Conference on Image Processing, vol. III, 1997, pp. 384–387. CAM report 97-50.

[4] E. M. BOLLT, R. CHARTRAND, S. ESEDOḠLU, P. SCHULTZ, AND K. R. VIXIE, *Existence and non-uniqueness for $\int |\nabla u|^{p(|\nabla u|)}$ regularized image functionals*, Advances in Computational Mathematics, Special issue, "Mathematical Methods for Image Processing" (2008).

[5] T. CHAN, S. OSHER, AND J. SHEN, *The digital TV filter and nonlinear denoising*, IEEE Transactions on Image Processing, 10 (2001).

[6] T. F. CHAN AND S. ESEDOḠLU, *Aspects of total variation regularized L^1 function approximation*, SIAM J. Appl. Math., 65 (2005), pp. 1817–1837.

[7] R. CHARTRAND, *Nonconvex regularization for shape preservation*, in IEEE International Conference on Image Processing (ICIP), 2007.

[8] Y. CHEN, S. LEVINE, AND M. RAO, *Variable exponent, linear growth functionals in image restoration*, SIAM Journal of Applied Mathematics, 66 (2006), pp. 1383–1406.

[9] T. H. CORMEN, C. E. LEISERSON, R. L. RIVEST, AND C. STEIN, *Introduction to Algorithms*, The MIT Press, second ed., 2002.

[10] S. E. LEVINE, *An adaptive variational model for image decomposition*, in Energy Minimization Methods in Computer Vision and Pattern Recognition, no. 3757 in Lecture Notes in Computer Science, Springer, 2005, pp. 382–397.

[11] O. LEZORAY, A. ELMOATAZ, AND S. BOUGLEUX, *Graph regularization for color and image processing*, Compter Vision and Image Understanding, 107 (2007).

[12] S. P. MORGAN AND K. R. VIXIE, *L^1TV computes the flat norm for boundaries*, Abstract and Applied Analysis, 2007 (2007), pp. Article ID 45153, 14 pages. doi:10.1155/2007/45153.

[13] D. MUMFORD AND J. SHAH, *Optimal approximations by piecewise smooth functions and associated variational problems*, Communications on Pure and Applied Mathematics, 42 (1989).

[14] J. V. NEUMANN AND A. W. BURKS, *Theory of self-reproducing automata*, University of Illinois Press, 1966.

[15] M. NIKOLOVA, *Minimizers of cost-functions involving nonsmooth data-fidelity terms*, SIAM J. Numer. Anal., 40 (2003), pp. 965–994.

[16] L. RUDIN, S. OSHER, AND D. FATEMI, *Nonlinear total variation based noise removal algorithms*, Physica D, 60 (1992).

[17] D. STRONG AND T. CHAN, *Edge-preserving and scale-dependent properties of total variation regularization*, Inverse Problems, 19 (2003).

[18] A. N. TIKHONOV AND V. Y. ARSENIN, *Solutions of ill-posed problems*, Winston, 1977.

[19] K. R. VIXIE, *Some properties of minimizers for the Chan-Esedoḡlu L^1TV functional*, arXiv.org, (2007).

[20] C. VOGEL, *Computational methods for inverse problems*, SIAM, 2002.