RECOVERING FILAMENTARY OBJECTS IN SEVERELY DEGRADED BINARY IMAGES USING BEAMLET-DECORATED PARTITIONING

Xiaoming Huo

School of ISyE,
Georgia Institute of Technology
Atlanta, GA 30332
xiaoming@isye.gatech.edu

David Donoho

Department of Statistics
Stanford University
Stanford, CA 94305
donoho@stat.stanford.edu

ABSTRACT

We consider the problem of recovering a binary image consisting of many filaments or linear fragments in the presence of severe binary noise. Our approach exploits beamlets—a dyadically organized, multiscale system of line segments—and associated fast algorithms for beamlet analysis and complexity-penalized model fitting. Simulation results demonstrate the effectiveness of the method.

1. INTRODUCTION

Real-world imaging often presents problems where the object of interest consists of many filaments and linear fragments – think of road networks or river networks seen from above, or boundaries of objects in computer vision.

In this paper, we consider a model for problems of this kind, where the object of interest is a binary image containing many linear or curvilinear fragments and there is severe binary noise degrading the image. One can think of this problem as arising when we must process the output of a pixelwise detector, where each pixel-level detector conducts a hypothesis test for presence or absence of a feature. For example, this could arise in hyperspectral imaging where the pixel-level computation involves analyzing a spectrum of reflected light at that pixel and testing for spectral similarity to a known substance, but the pixel-level test has a rather high error rate. Other examples include processing noisy infrared imagery of a scene by an edge detector, setting a threshold on the amplitude of the edge detector, and obtaining a binary image of pixel-level 'edge-detections'. For helpful literature, see for example [7, 8].

State-of-the-art methods for image processing in such settings generally have two components: (1) local feature identification and (2) global selection and fusion. Often, the underlying features are assumed to be obvious at pixel level, so that it is not difficult to obtain an effective local feature detector, and the problem of integration and fusion is correspondingly easy. We study a much harder case here, where the image is so severely noisy that when operating at pixel level, it is not possible to tell if it is a part of a larger-scale linear or curvilinear feature or not; and attempts to fuse local decisions into global recognition are doomed. We develop in this paper a multiscale approach which automatically operates at the smallest scale where the underlying features are extractible in a statistically reliable manner, and which can therefore detect features in far noisier environments than possible using the current state of the art.

Our method is based on beamlet analysis, as defined and delineated in [5]. Beamlets provide a multiscale dictionary of line segments having a range of orientations, locations, and scales, obeying a natural tree hierarchy and having a natural multiscale graph structure. Beamlet analysis is the operation of integrating an image along each of these line segments, and is best viewed, on first acquaintance, as a kind of multiscale Radon or Hough transform. Beamlet analysis is essentially comparing the image with a rich set of linear features at all scales, orientations, and locations and storing the results in a well-organized data structure. Our method exploits this analysis by further processing the data structure, looking for beamlet coefficients which are unusually large given the number of pixels involved and the ambient noise level. The method returns a collection of beamlets, which can be viewed as providing a reconstructed image extracted from the noisy image.

The reconstruction is in fact the solution of a formal optimization problem seeking the minimum of a complexity-penalized likelihood; informally, we seek a faithful description of the data with low complexity (i.e. using few beamlets). The optimization problem is framed using the natural tree hierarchy of the beamlet system and, by applying ideas from dynamic programming, can be solved rapidly using a tree-pruning algorithm.

The method intrinsically carries out an adaptive multiresolution analysis, since the beamlets used in the analysis occur at all possible scales and locations and the beamlets used in the reconstruction are selected by a statistical principle guaranteeing that at each location, the beamlets (if any) which are extracted, have the smallest scale where the results are statistically reliable.

This is a part of a bigger effort to analyze the potential of using beamlets, and their tree and graph structure to derive reliable methods for solving problems detection and extraction of curvilinear objects at very challenging noise levels. In this project, a strategy that is based on multiresolution analysis helps to solve a problem which is otherwise seemingly hopeless. We have here restricted ourselves to the analysis of binary images. Similar approaches can be derived for more complex imagery, with Synthetic Aperture Radar images and color digital images being promising areas for attention.

A related idea for linear feature extraction in a different (Gaussian) noise model was described in [6], together with some illustrations. The statistical modeling approach and the idea of minimizing a penalized distortion function has a precursor in [3], where the idea was to extract regions and their boundaries rather than filaments and curvilinear fragments, and the noise model was Gaussian rather than Bernoulli. Novel features in this work include the statistical model, and the choice of penalty parameter by the principle of controlled false alarm rate in each scale.

In the simulations, when a high proportion of the pixels in a binary image are contaminated (switching from 0's to 1's or vice versa), the model based method can still reliably extract most of the underlying linear features.

In this paper, in Section 2 the beamlet-based statistical models are described. In Section 3 the penalization is determined so that the false alarm rates are controlled. In Section 4, two representative simulations are reported. In Section 5, the advantages and limitations of this method are discussed. Finally, we conclude in Section 6.

2. BEAMLET BASED MODEL

The design of the *beamlet dictionary* is provided in detail in [5], where relations to earlier work, for example with multiscale Radon transforms, was described; the reader should consult that work for the details we omit here. There are three steps in designing the beamlet dictionary: (1) partitioning the image domain into dyadic squares at all scales, (2) marking the boundaries of squares with equispaced vertices at pixel spacing and (3) defining beamlets as line segments connecting pairs of vertices within a dyadic square. Here we consider a squared image with dyadic length (the number of pixels on each row or in each column is an integral power of 2). Within a square, a *digital beamlet* is a set of pixels that are aligned with a line segment which begins and ends at the boundary of this square. Given a line segment, the selection of the pixels are determined by following Bresenham's algorithm [1]. More details on the digital beamlets are provided in a companion paper [2].

A Recursive Dyadic Partition (RDP) of the image domain is any partition arrived at by starting with the whole image domain and recursively decomposing any part of an existing RDP into dyadic squares based on the dyadic subdivision principle of replacing a square by four similar squares of half the size; for fuller discussion and pictures, see [5]. The idea will be familiar as *quadtree decomposition* to many readers. A beamlet-decorated RDP (BD-RDP) is an RDP in which some of the terminal nodes of the partition are decorated by beamlets associated with those squares.

Our statistical model specifies that the pixels in the image have a statistical distribution based on Bernoulli trials, in which the probability of "1" differs from pixel to pixel. Given a beamlet-decorated RDP, the model states in an undecorated square, we have a statistically homogeneous i.i.d. Bernoulli with parameter $1/2 - \epsilon$; in a decorated square, we have such a noise in all pixels *except* along a single digital beamlet associated with that square, where the Bernoulli parameter is $1/2 + \epsilon$. Hence, on pixels belonging to the beamlet decoration, there is an elevated probability of "1". Note that in our model, no pair of beamlets overlaps. The beamlets can have various lengths, depending on the size of the subsquares that they reside on.

Each BD-RDP defines a statistical model, so there is a huge number of possible models for binary data. Our goal will be to fit to a given binary image a good model from the collection of all BD-RDP's. In short, we have converted the problem of recovering the underlying filaments and linear fragments into a model selection problem.

We propose a *Complexity-Penalized Distortion* (CPD) approach for estimation in the above setting. Suppose $\{S \in \mathcal{P}\}$ is the collection of subsquares making up a RDP $\mathcal{P}$, and that $\mathcal{P}'$ is the collection of decorated subsquares. For such a decorated square $S' \in \mathcal{P}'$, let $b_{S'}$ denote the digital beamlet that decorates S'; for convenience, we set $b_S = \emptyset$ if $S \notin \mathcal{P}'$. A statistical model (denoted by m) can be written as $m = \{(b_{S'}, S' \in \mathcal{P}'), \mathcal{P}\}$. The complexity penalized distortion is

$$\text{CPD}(m, \lambda) = \|y - 1_m\|_0 + \sum_{S \in \mathcal{P}'} \lambda(S), \qquad (1)$$

where y stands for the observed image, 1_m is the indicator of the pixels in the model (i.e. pixels in the decorations), $\|\cdot\|_0$ measures the number of nonzero entries, and $\lambda(S)$ is the penalty associated with square S. Note that binary images are considered, so the ℓ_0 norm of two binary images is sufficient to measure the distortion.

The second term in the above equation is the penalty term, which is intentionally decomposed as functions of subsquares. It is reasonable to assume that $\lambda(S)$ is only a function of the size of the subsquare S, and also of the fact that it is decorated. Our choices of λ's are determined by the size of the subsquares, which are also the scales of the corresponding beamlets. Note that this realizes a scale-dependent penalization, which is in some ways similar to scale-dependent thresholding in wavelet analysis. The model that minimizes the function CPD(m, λ) is our Complexity-Penalized Distortion estimate.

In a binary image, if the distortion of a particular pixel follows the Bernoulli distribution (it has the chance p to switch to another value and the chance $1-p$ to stay at the same value), the term $\|y - 1_m\|_0$ is proportional to the logarithm of the likelihood function given that m is the true underlying model. Hence the CPD is in fact a complexity-penalized likelihood criterion; with suitable reinterpretation it can be viewed as a kind of maximum a posteriori estimator under a certain random beamlet model.

Since the penalty function can be decomposed as functions of subsquares of a RDP, the penalty is additive according to the quadtree structure given by RDPs. So an algorithm that has the flavor of the *Best Orthonormal Basis* can be deployed. Since the algorithm is a bottom-up tree pruning algorithm, it will be fast. Its order of complexity is roughly equal to the size of the binary image.

The function λ is chosen based on controlling the *false alarm rate*. This will be addressed in the next section.

3. CHOICE OF PARAMETER

Suppose s is a subsquare with side $n(s)$. Let $\lambda(s) = 0$, if there is no beamlet in s. Let y_s denote the subimage that resides on the subsquare s. Recall that we want to choose a function λ which is a function of the size, $n(s)$, of the subsquare. In this fixed subsquare, the probability of false alarm is given by

$$\text{f. a. r.} = P\{\exists b \subset s : \|y_s - 1_b\|_0 + \lambda(n(s)) < \|y_s\|_0\},$$

where b is a beamlet in s. Moreover, we have

$$\begin{aligned} \text{f. a. r.} &= P\{\exists b \subset s : \lambda(n(s)) < 2(\#\text{ones on } b) - \|b\|_0\} \\ &= P\{\lambda(n(s)) < \max_{b \subset s}[2(\#\text{ones on } b) - \|b\|_0]\}. \end{aligned}$$

When the switching probability p of each pixel is given, in a fixed subsquare s, the distribution of the statistic $\max_{b\subset s}[2(\#\text{ones on } b) - \|b\|_0]$ can be derived. Hence a value of $\lambda(n(s))$ can be chosen so that the false alarm rate is controlled. We determined the values of λ via simulations. More details on the analysis of the distribution of this maximum will be reported later.

Since a beamlet at scale n—which is the sidelength of the subsquare this beamlet residing on—can be decomposed as two or three smaller beamlets at scale $n/2$, it is desirable for the function λ to satisfy the condition

$$2\lambda(n/2) > \lambda(n),$$

so that the optimal solution always chooses a coarser scale elements, instead of its superposition. The function λ that is chosen based on the controlled false alarm rate satisfy the above inequality. As a matter of fact, we observe the following: for $p < 1/2$ and $n > 4$,

$$\lambda(n) \propto c \log_2(n),$$

where c is a constant that depends on p. A verification of this result will be studied later.

4. SIMULATIONS

The results of two simulations are shown in Figure 1 and Figure 2. In both figures, the top are the original images, which are 128 by 128 binary images. In the distorted images (the middle ones), 20% of the pixels switch values (from 0 to 1, or from 1 to 0). The bottom images are the estimates that are from the beamlet model based approach. We can see that the estimates apparently preserve a significant amount of the original features. There are a few mis-detections. But compared to the distorted images, the false identification is well-controlled.

5. DISCUSSION

There are some desired properties of this approach. First since the objective function is additive in a quadtree structure, the solution can be efficiently computed through a tree-pruning algorithm. So this approach has a fast algorithm. Note that this is different from many methods that relies on a Markovian Random Field model, which is usually expensive to find the optimal solution. Second, in the contaminated binary image model that we study, it performs very well in the sense that most of the distortion has been suppressed. Third, this method automatically carries out a multiresolution analysis.

This approach is not perfect. For example, from the simulation results, the method fails at the connecting points of multiple linear features. This is due to the non-overlapping property in the BD-RDP model. Also due to the structure of RDP and the criterion of controlling false alarm rate, if the “true” statistical model contains fine scale beamlets, then in our approach, the fine scale components will be lost. However as stated earlier, our method automatically starts working at the finest scale that the problem is solvable.

6. CONCLUSION

A method that is based on multiscale feature extraction is proposed to identify linear features in a severely distorted binary image. Simulation demonstrates the effectiveness of this approach. The proposed method has low computational complexity. The fundamental idea that is embedded in the algorithmic approach could be used to generate new and efficient methods in many other situations where the Signal to Noise Ratio is extremely low.

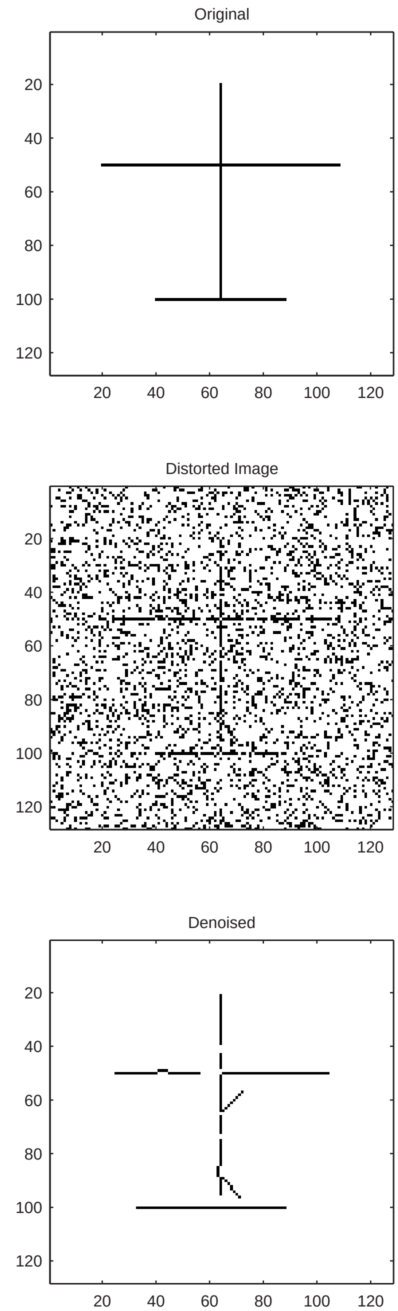

Figure 1: An example of denoising.

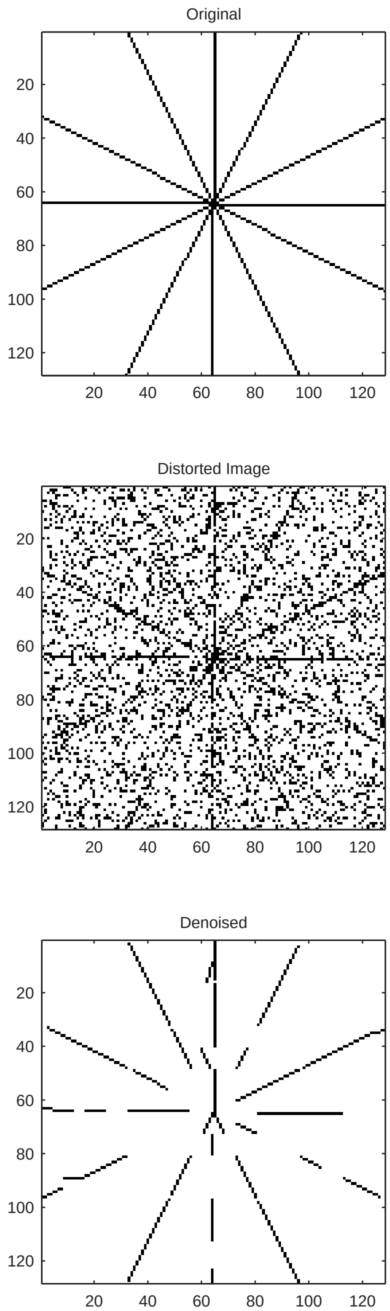

Figure 2: The second example of denoising: a star pattern.

7. REFERENCES

[1] J. Bresenham. Algorithm for computer control of a digital plotter. *IBM Systems Journal*, 4(1):25–30, Jan. 1965.

[2] J. Chen and X. Huo. Beamlet coder: A tree-based, hierarchical contour representation and coding method. Submitted to ICASSP 2002, Sep. 2001.

[3] D.L. Donoho. Wedgelets: Nearly minimax estimation of edges. *Annals of Statistics*, 27(3):859–897, 1999.

[4] D. Donoho and X. Huo. Beamlet pyramids: A new form of multiresolution analysis, suited for extracting lines, curves, and objects from very noisy image data. In *Proceedings of SPIE*, volume 4119, July 2000.

[5] D.L. Donoho and X. Huo. Beamlets and multiscale image analysis. In *Multiscale and Multiresolution Methods*, Springer Lecture Notes in Computational Science and Engineering, T.J. Barth, T. Chan, and R. Haimes (Eds.), Vol. 20, November, 2001.

[6] X. Huo. Some Examples of Untraditional Statistical Computing. In *Proceedings of Joint Statistical Meeting*, August 2001.

[7] A. Katartzis, H. Sahli, V. Pizurica, and J. Cornelis. A model-based approach to the automatic extraction of linear features from airborne images. *IEEE Trans. on Geoscience and Remote Sensing*, 39(9):2073–2079, Sep. 2001.

[8] F. Tupin, H. Maître, J.F. Mangin, J.M. Nicolas, and E. Pechersky. Detection of linear features in SAR images: Application to road network extraction. *IEEE Trans. on Geoscience and Remote Sensing*, 36(2):434–453, Mar. 1998.