A GEOMETRIC FRAMEWORK FOR CHANNEL NETWORK EXTRACTION FROM LiDAR: NONLINEAR DIFFUSION AND GEODESIC PATHS

By

Paola Passalacqua

Tien Do Trung

Efi Foufoula-Georgiou

Guillermo Sapiro

and

William E. Dietrich

IMA Preprint Series # 2234

(February 2009)

INSTITUTE FOR MATHEMATICS AND ITS APPLICATIONS

UNIVERSITY OF MINNESOTA

400 Lind Hall

207 Church Street S.E.

Minneapolis, Minnesota 55455–0436

Phone: 612-624-6066 Fax: 612-626-7370

URL: http://www.ima.umn.edu

JOURNAL OF GEOPHYSICAL RESEARCH, VOL. ???, XXXX, DOI:10.1029/,

A geometric framework for channel network extraction from LiDAR: nonlinear diffusion and geodesic paths

Paola Passalacqua,[1] Tien Do Trung,[2] Efi Foufoula-Georgiou,[1] Guillermo Sapiro,[3] and William E. Dietrich[4]

E. Foufoula-Georgiou and P. Passalacqua, National Center for Earth-Surface Dynamics and St. Anthony Falls Laboratory, Department of Civil Engineering, University of Minnesota, 2 3rd Avenue SE, Minneapolis, MN 55414, USA. (efi@umn.edu)

T. Do Trung, Département de Mathématiques, École Normale Supérieure de Cachan, Bâtiment Laplace 1er étage, 61 Avenue du Président Wilson 94235 Cachan Cedex, France.

G. Sapiro, Department of Electrical and Computer Engineering, University of Minnesota, 200 Union Street SE, Minneapolis, MN 55455, USA.

W. E. Dietrich, Department of Earth and Planetary Science, University of California, Berkeley, McCone Hall, Berkeley, CA 94720, USA.

Abstract.

A geometric framework for the automatic extraction of channels and channel networks from high resolution digital elevation data is introduced in this paper. The proposed approach incorporates nonlinear diffusion for the pre-processing of the data, both to remove noise and to enhance features that are critical to the network extraction. Following this pre-processing, channels are defined as curves of minimal effort, or geodesics, where the effort is measured based on fundamental geomorphological characteristics such as flow accumulation area and iso-height contours curvature. The merits of the proposed methodology, and especially the computational efficiency and accurate localization of the extracted channels, are demonstrated using LiDAR data of two river basins.

1. Introduction

The development of channel network extraction methodologies from Digital Elevation Maps (DEMs) has been the subject of research for a long time [e.g., *Montgomery and Dietrich*, 1988; *Tarboton et al.*, 1991; *Montgomery and Foufoula-Georgiou*, 1993; *Costa-Cabral and Burges*, 1994; *Giannoni et al.*, 2005; *Hancock and Evans*, 2006]. Most of these extraction methodologies have been developed on 30 m or 90 m DEMs, resolutions commonly available until recently. At these resolutions, channel initiation can only be inferred indirectly, using a threshold on drainage area and slope or a threshold on curvature [e.g., *Dietrich et al.*, 1988, 1992, 1993; *Tarboton et al.*, 1988; *Howard*, 1994a, b; *Rodriguez-Iturbe and Rinaldo*, 1997; *Heine et al.*, 2004]. Moreover, these methodologies are sensitive to the resolution of the initial DEM [e.g., *Helmlinger et al.*, 2004]. The availability of high resolution (1 to 3 m data spacing) elevation data offers an opportunity to directly and objectively extract channels and other features of geomorphologic/hydrologic interest, but requires the development of new computational tools.

Recently *Lashermes et al.* [2007] proposed a wavelet-based filtering methodology to compute the Laplacian and slope-direction-change across scales, and exploited the statistical signature of these features for extracting channel networks from LiDAR data. It was found that a sharp deviation in the positive tails of the probability distribution of the Laplacian from a Gaussian distribution defines a critical threshold which delineates the channelized valleys of the terrain. Within those valleys, a maximum slope-direction-change algorithm was used to direct the forward tracing of the channel centerline (being this forward tracing very sensible to noise, errors in the computations, and missing data).

The method was applied to a watershed, within the South Fork Eel River (California) using ALSM data acquired by NCALM (data set available at data distribution archive http://www.ncalm.org/). This wavelet-based methodology of channel extraction from LiDAR data presents a real advantage over prior methodologies developed for lower resolution DEMs, allowing the multi-scale analysis of the elevation data and the extraction of the corresponding channel network.

In this paper, a geometric framework which significantly advances the accurate and automatic extraction of channel networks from LiDAR data is developed. The first component of the framework is the use of nonlinear geometric filtering (via partial differential equations), instead of linear filtering via wavelets, which naturally adapts to a given landscape and facilitates the enhancement of features for further processing. Early uses of nonlinear partial differential equations for digital elevation maps appear in *Braunmandl et al.* [2003] for scale space generation and in *Almansa et al.* [2002]; *Solé et al.* [2002] for interpolation. The form of this filtering is such that it behaves as linear diffusion at low elevation gradients, while it arrests diffusion as the gradients become large. The second key component of the proposed framework, is the novel formulation of the channel network extraction problem as a geodesic energy minimization problem with a cost-function which is geomorphologically informed, i.e., it is defined in terms of local attributes of the landscape such as upstream drainage area and iso-height contours curvature.

The remainder of this paper is organized as follows. Section 2 gives a brief mathematical background on nonlinear diffusion, geometric filtering, geodesics, and energy minimization principles. In Section 3 these techniques are applied to the problem of channel network extraction and demonstrated in two real basins. Finally Section 4 presents conclusions and challenges for future research.

2. Mathematical background on the proposed methodology

This section presents the basic mathematical background that provides the foundation for the channel networks extraction geometric framework introduced in this paper. First, the notion of nonlinear anisotropic filtering, to replace the linear isotropic one such as via wavelets, is introduced. Next, the framework of geodesic computations is presented. This leads to consider channel networks, computed on the nonlinearly filtered data, as paths of minimal effort. Such paths spatially integrate geomorphological local features that characterize the channel network, thereby providing a local-global approach to the detection of important geomorphological features.

2.1. Nonlinear diffusion and geometric filtering

Let us denote by $h_0(x, y) : \mathbb{R}^2 \to \mathbb{R}$ the provided DEM image, i.e., high resolution digital elevation data. Typical of any feature extraction methodology is the application of a *smoothing* filter on the original data $h_0(x, y)$ to remove noise and identify features as entities that persist over a range of scales. This operation of smoothing is also very important to make computations such as derivatives mathematically well-posed. A popular smoothing filter is the Gaussian kernel, which, when applied to $h_0(x, y)$, results in landscapes at coarser resolutions, i.e.,

$$h(x, y, t) = h_0(x, y) \star G(x, y; t) \tag{1}$$

where $\star$ denotes the convolution operation and $G(x, y; t)$ is a Gaussian kernel of standard deviation t (larger values of t result in coarser resolution landscapes), centered at location

(x, y):

$$G_{x,y,t}(u,v) = \frac{1}{2\pi t} exp\left[-\frac{(u-x)^2+(v-y)^2}{2t}\right] \tag{2}$$

As it was shown and exploited in *Lashermes et al.* [2007], the use of the classical Gaussian smoothing kernel naturally leads to a multiscale (scale-space in the computer vision community) efficient computation of local slopes and the Laplacian via wavelets, where wavelets were selected as the first and second derivatives of a Gaussian kernel (see *Burt and Adelson* [1983]; *Koenderink* [1984]; *Witkin* [1983] for early developments and the introduction of Gaussian filtering for multiscale image analysis).

The family of coarsened landscapes resulting from (1) may be seen as solutions of the linear heat or diffusion equation, e.g., [*Koenderink*, 1984], with the initial condition $h(x, y; 0) = h_0(x, y)$, i.e.,

$$\partial_t h(x,y,t) = \nabla \cdot (c\nabla h) = c\nabla^2 h \tag{3}$$

where c is the diffusion coefficient and ∇ is the gradient operator. Thus, processing the landscape with filters of increasing spatial scale, as done in *Lashermes et al.* [2007], is equivalent to applying an isotropic diffusion equation over time on the landscape with the spatial scale of the filter (variance) and the time of diffusion being related to each other (since derivatives are linear operations, filtering and then differentiating is equivalent to filtering with the corresponding derivatives of the filter). Once the scale of the Gaussian kernel is fixed, the time over which diffusion is applied on the original landscape is spatially uniform, i.e., the landscape is diffused with the same rate at all points and in all directions.

The choice of the Gaussian kernel as smoothing filter is motivated in part by two criteria defined by *Koenderink* [1984] as (1) *causality* and (2) *homogeneity/isotropy*. The *causality* guarantees that no spurious feature should be generated at coarser resolutions, since any feature at a coarse level of resolution must have a cause at a finer level of resolution. This guarantees denoising of the original data as the resolution is coarsened. The *homogeneity/isotropy* criterion requires the blurring to be space invariant. The Gaussian kernel thus satisfies the standard scale-space paradigm as stated by *Koenderink* [1984]. It is noted, however, that the Gaussian filtering is isotropic and does not respect the natural boundaries of the features and diffuses across boundaries throughout the landscape. This obviously degrades the spatial localization of these boundaries, especially at larger scales of smoothing. These boundaries represent, in the case of landscapes, important discontinuities such as crests and valleys. *Perona and Malik* [1990] reformulated the space-scale paradigm to address this issue. The new paradigm was reformulated to satisfy three criteria: (1) *causality*, as previously stated by *Koenderink* [1984], (2) *immediate localization*, which searches, at each resolution, sharp and meaningful region boundaries, and (3) *piecewise smoothing*, which indicates preferential smoothing (intraregion rather than interregion).

In the standard linear diffusion equation (3), the diffusion coefficient c is constant, that is, independent of the space location. An extension to the Gaussian filtering is obtained by choosing the diffusion coefficient c to be a suitable function of spatial location, such that the new space-scale paradigm criteria are satisfied. The modified diffusion equation can be written as

$$\partial_t h(x, y, t) = \nabla \cdot [c(x, y, t)\nabla h] = c(x, y, t)\Delta h + \nabla c \cdot \nabla h \tag{4}$$

Note that (4) reduces to the linear diffusion equation (3) if $c(x, y, t)$ is constant. If the location of a channel were known, then, in order to achieve denoising and edge enhancement, smoothing should preferentially happen in the region outside and within the channel, rather then across its boundary. This could be achieved by setting $c = 0$ at the channel boundaries and $c = 1$ everywhere else. However the channel location is not known in advance, and what can be computed instead is an estimate of it, or some geometric characteristic that defines it, thereby stopping, or at least reducing, diffusion across the channel boundary.

Let $\vec{E}(x, y, t)$ denote the vector-valued function representing an estimate of the channel's location. The diffusion coefficient can be chosen as a function of the magnitude of $\vec{E}(x, y, t)$, i.e.,

$$c = p(\|\vec{E}\|) \tag{5}$$

where $p(\cdot)$ has to be designed such that it ideally does not allow diffusion across boundaries. *Perona and Malik* [1990] have proposed a simple first estimate of the channel's location (or image edges in their original application), given by the gradient of the elevation $h(x, y; t)$ at the location (x, y) and time t, i.e.,

$$\vec{E}(x, y, t) = \nabla h(x, y, t) \tag{6}$$

This provides a local estimator of the edges/discontinuities within the nonlinear space-scale paradigm. Note that we could also use curvature or other higher order features to define the diffusion coefficient c, while the use of gradients is the most standard formulation and found to be sufficient for our application. The diffusion equation thus takes the following form:

$$\partial_t h(x, y, t) = \nabla \cdot [p(\|\nabla h\|)\nabla h] \tag{7}$$

Perona and Malik suggested the following as possible edge-stopping functions:

$$p(\nabla h) = \frac{1}{1 + (\|\nabla h\|/\lambda)^2} \quad (8)$$

or

$$p(\|\nabla h\|) = e^{-(\|\nabla h\|/\lambda)^2} \quad (9)$$

where λ is a constant. Such expressions of the edge-stopping function, when regularized, guarantee basic properties of the scale-space paradigm, while at the same time enhancing the discontinuities, thereby allowing their easier extraction (see *Alvarez et al.* [1992]; *Perona and Malik* [1990] for details).

The just introduced nonlinear diffusion equation will be used as a pre-processing step in the elevation data, to remove unwanted details and enhance the features that are relevant for channel network extraction. While many alternatives exist in the literature for nonlinear diffusion, we found this basic and most classical one to be sufficient to introduce the ideas and to obtain state-of-the-art results for the tested LiDAR elevation data.

2.2. Geodesics and energy minimization principles for network extraction

In this section, a new geometric methodology is proposed for extracting channels on the regularized DEM image obtained by applying the Perona-Malik filter to the initial high resolution elevation data set. Let us now denote by h the regularized LiDAR data and by Ω the terrain region described by these data. Let C be a curve restricted on Ω, $C \in \Omega$. Consider two fixed points a and b on the surface Ω such that the curve C passes through them. The cost of traveling on the curve C is given by the function $\psi(C) : \Omega \rightarrow \mathbb{R}^+$. The *geodesic distance* from point a to any other point $x \in \Omega$ is defined as:

$$d(a, x) := \min_{C \in \Omega} \int_a^x \psi(s) ds \tag{10}$$

where s is the standard Euclidean arc-length [*Do Carmo*, 1976]. The minimum is taken over all possible curves $C \in \Omega$ that start at a and end at x, and the cost of traveling is integrated on that curve. The *geodesic curve* is defined as the curve with the minimal cost, among all possible curves connecting the two points a and b. Thus it is the actual curve that achieves that minimum (not necessarily a unique curve):

$$g(a, b) := \arg \left(\min_{C \in \Omega} \int_a^b \psi(s) ds \right) \tag{11}$$

It is easy to see that such a curve is computed by gradient descent on the distance function $d(a, \cdot)$, backtracking from the 'downstream' point b. The geodesic is thus the integral curve of ∇d starting at b, and the gradient is intrinsically computed on the surface. An efficient computation of the distance function can be obtained in linear time [*Yatziv et al.*, 2006], by extending classical algorithms for computing distance functions on graphs; namely, Dijkstra and Dial algorithms [*Dial*, 1969; *Dijkstra*, 1959]. These algorithms are applicable to all diverse types of surface representations, from triangulated surfaces [*Kimmel*, 2003] to point cloud data as in LiDAR [*Memoli and Sapiro*, 2005]. These extensions are based on the fact that such a distance function satisfies a Hamilton-Jacobi geometric partial differential equation, $\|\nabla d\| = \psi$, where the gradient is intrinsic to the surface in the most general case. Additional information on these efficient computations can be found in *Helmsen et al.* [1996]; *Sethian* [1999]; *Tsitsiklis* [1995]; *Tsai et al.* [2002]; *Zhao* [2004].

Different selections of the cost function ψ will lead to different curves on the surface. *Natural geodesics* are, for example, obtained in the case of a constant cost function. For the problem of detecting channel networks, the cost function has to include topographic attributes which differentiate channels from the rest of the landscape. Such attributes are the surface curvature (positive curvature, or curvature above a threshold value, commonly indicates convergent topography correspondent to channelized areas, negative curvature indicates divergent topography correspondent to hillslopes), and the flow accumulation (larger values are expected along channelized paths). In the next section we propose such a geomorphologically meaningful cost function and demonstrate its performance for automatic channel network extraction from LiDAR data.

3. River network extraction

The objective of this section is to illustrate the concepts described above through their application on LiDAR data of the South Fork Eel River basin in Northern California. We use the ALSM data (2.6 m average bare earth data spacing, gridded to 1 m) acquired by NCALM (the data are available online at the data distribution archive http://www.ncalm.org/). We focus in particular on two sub-basins. One is a 2.8 km^2 mostly forested tributary that lies just north of the Angelo Coast Range Reserve, about 3 km downstream from the junction of the Ten Mile Creek and the South Fork Eel River. The second sub-basin is the Skunk Creek, a 0.54 km^2 tributary located just upstream of the Elder Creek. The two sub-basins are shown in Fig. 1 and Fig. 2 respectively.

3.1. Preprocessing: Regularization of high resolution digital elevation data through nonlinear filtering

We focus our analysis on a 300m by 300m portion of the first sub-basin, referred to as portion A (see Fig. 1). The landscape A has been processed with a Gaussian filter (isotropic linear diffusion) and the Perona-Malik filter (anisotropic nonlinear diffusion).

To allow comparison of the two filtered landscapes the time of forward diffusion (iteration steps) has been set to 50 iterations in both (in general, there is no exact mathematical correspondence between the corresponding diffusion times). This corresponds to a Gaussian spatial filter of approximate $\sigma = 7$m (scale of smoothing of the landscape of approximately $4\sigma = 28$m; see Table 1 of *Lashermes et al.* [2007]). As is apparent from the theory, no such unique and uniform equivalent spatial scale of smoothing can be assigned to the Perona-Malik nonlinearly filtered landscape as the effective smoothing scale varies locally at every point depending on the local gradient. Specifically, the effective spatial scale of smoothing is smaller close to the streams (where the gradient is large and the edge stopping function of equation (8) assigns a small diffusivity coefficient), and larger in areas of spatially homogeneous and small gradients. The Perona-Malik filter used in this analysis is that of equation (8) with parameter λ estimated from the 90% quantile of the pdf of the gradients, as also suggested in *Perona and Malik* [1990] (the selection of such a parameter can be made fully automatic also following the robust statistics approach in *Black el al.* [1998]).

Fig. 3(a) shows the original landscape at the resolution of 1m with 3m contours superimposed on it, as well as the computed gradients and curvatures (using simple first and second order differentiation). Figs. 3(b) and 3(c) show the filtered landscapes with the Gaussian filter and Perona-Malik filter, respectively, using for both 50 iterations as the stopping time of the forward diffusion as explained above. The curvature reported here in all cases is the (geometric) curvature of the iso-height contours, $\kappa = \nabla \cdot (\nabla h / \|\nabla h\|)$, computed by finite differences. The advantages in using the curvature instead of the Laplacian will be addressed later in this section.

Several observations can be made from these figures. First, it is easily seen from Fig. 3(b) that the Gaussian filter smoothes the contours along the channels much more than the Perona-Malik filter. This is expected from the theoretical properties of the Perona-Malik filter which deforms the landscape much less along the edges. In fact, the Perona-Malik filter achieves a limited deformation of contours along the edges such that it encourages the localization of these features. It is also observed that the areas of the landscape over which the curvature is positive (along the channelized areas) are much broader in the Gaussian filtered landscape than in the Perona-Malik landscape. This is also expected from the basic properties of the two filters. One can argue that the Gaussian filtering (isotropic diffusion) could be stopped earlier, i.e, smaller spatial scale of filtering, to result in better localization of the channelized valleys. However, as it will be seen later, such a smaller-scale filtering would not adequately eliminate the isolated high curvature areas that are not pertinent to channel extraction. Furthermore, nonlinear diffusion is enhancing the discontinuities (acting in those regions as backward diffusion as shown in *Perona and Malik* [1990]), which is critical for facilitating the automatic channel network extraction.

Fig. 4 shows the pdfs of the curvatures of the original data and the filtered landscapes as well as the quantile-quantile plots of those curvatures. As discussed in *Lashermes et al.* [2007] for the Laplacian, the sudden change in the statistical signature of the landscape, depicted by the (positive) curvature at which the pdf deviates from a Gaussian pdf, marks the transition from hillslopes to valleys. It is interesting to observe that although the actual value of the threshold curvature is different for the original image and the two filtered images, as expected, the quantile at which this transition occurs is scale- and filter-independent and as reported in *Lashermes et al.* [2007] for the Laplacian, corresponds to the standard normal deviate of $z = 1$ (approximately the 84^{th} quantile of the pdf of curvatures). The right panels of Fig. 4 depict the pixels at which the curvature was greater than the threshold curvature identified from the corresponding pdfs; white pixels correspond to pixels with curvature greater than the threshold value while black pixels correspond to pixels with curvature smaller than the threshold value. Several observations can be made. First, the above-threshold-curvature pixels in the original high resolution data depict the channelized part of the landscape but at the same time one sees several isolated small areas which are strongly convergent due to inaccurate computation of second order differences from the original noisy elevation data. Second, the above-threshold-curvature pixels on the Gaussian filtered landscape eliminate the noise and nicely depict the valleys or channelized areas only; however, the corridors of the convergent areas are too wide due to the smoothing of the landscape which has been done at the scale of approximately 27m throughout the landscape. The above-threshold-curvature pixels in the Perona-Malik filtered landscape (shown in Fig. 4(c)), not only eliminate the noise but also depict in a much sharper way the channelized valleys. Again, one could argue that by using a smaller scale Gaussian filter, sharper delineation of the channelized valleys would result. While this is true, the smaller scale of smoothing would not eliminate the isolated small convergent areas which are not part of the channel network. This is demonstrated in Fig. 5 which displays the above-threshold-curvature pixels for three standard deviations of the Gaussian filter: $\sigma = 2m$ (landscape smoothing scale $a = 8.9$m); $\sigma = 4$m (landscape smoothing scale $a = 17.8$m); $\sigma = 6$m (landscape smoothing scale $a = 26.7$m). It is noted by comparing Fig. 4(c) right panel and Fig. 5, that the Perona-Malik localization of the channelized valleys (measured by the width of the white corridors) is comparable to the localization provided by the Gaussian filter at scale of approximately 9m (σ = 2m). However, at this small scale of smoothing, the Gaussian filtering results in many more isolated high curvature areas as can be seen in Fig. 5 left panel. Thus we conclude overall, that the Perona-Malik filter is a more efficient filter to use for pre-processing of the raw data (to produce what is called 'regularized data') on which further operations for automatic channel extraction can be performed. It is also worth pointing out the advantage of using the (geometric) curvature κ instead of the Laplacian. This can be seen by comparing Fig. 4(c) right panel to Fig. 6. The figures show the skeletons of pixels above-threshold-curvature obtained on the Gaussian filtered data (scale $\sigma = 7$m) using curvature (Fig. 4(c)) and Laplacian (Fig. 6). Note how sharper and well defined is the skeleton obtained using the curvature.

Before demonstrating in the next section the geodesic energy minimization approach for the automatic extraction of the whole river network, we note that one could further process the regularized data to eliminate even more the occasional isolated convergent pixels seen in Fig. 4(c). This is an optional further operation which can be easily done via a contributing-area-threshold, where the threshold used is arbitrary but small enough not to interfere with channel initiation. For example, Fig. 7 shows the skeleton of Fig. 4(c) after applying an additional contributing area threshold of $A \leq 1000\text{m}^2$. It is observed that this further operation not only removes isolated convergent areas, but also further narrows the width of the likely channelized valleys providing thus a better pre-processed data on which the geodesic optimization will be performed.

3.2. Automatic extraction of channel paths from the regularized data

In this section we focus on the regularized data set obtained through nonlinear filtering and illustrate how the concepts of geodesics and energy minimization described earlier allow a fast and efficient extraction of the river network. We applied the above pre-processing through Perona-Malik filtering to the 2.8km^2 sub-basin and to the Skunk Creek basin (previously shown in Fig. 1 and Fig. 2 respectively). For the Skunk Creek basin we had available a hand-drawn channel network map (field survey done by Joel Scheingross and Eric Winchell, University of California, Berkeley), which is shown in Fig. 8. The first step of the extraction procedure is the creation of the skeleton obtained by thresholding the curvature and the contributing area, as discussed in the previous section. The threshold curvature was easily identified by a clear change in the statistical behavior of the curvature, while the threshold area was set to a low value of 1000m^2. The extracted skeleton for the Skunk river basin is shown in Fig. 9.

Several observations can be made by comparing Fig. 9 with the hand-drawn network shown in Fig. 8. First, in Fig. 8 one observes that most of the left-side channels are labeled as 'poorly developed' and indeed the extracted skeleton depicts this topography by a series of interrupted areas of high curvature. Second, at the points that the hand-drawn channels terminate, our algorithm depicts a substantial interruption in the channelized valley. It is observed therefore, that the pre-processing allows one to investigate more closely the richness of the landscape form, something not possible with other current algorithms. From the skeleton of Fig. 9, we detected the river network outlet, as the point with the maximum flow accumulation area, computed, for example, using the Dinf algorithm developed by *Tarboton* [1997]. We also detected the 'end points' of the network as the *farthest* points from the outlet on each branch, still belonging to the skeleton. For this application we ignored interruptions if these were smaller than approximately 20-30m, but this can be a user-specified value.

Once the outlet and end points were detected, they were automatically connected with geodesic curves through an appropriately chosen *cost function*. This *cost function* was chosen to give penalty for selecting paths along which the drainage area does not have the largest flow accumulation and along which the curvature is not maximum compared to the surrounding points. The chosen form of the cost function ψ used in (11) is the following (α and δ are parameters):

$$\psi = \frac{1}{(\alpha \cdot A + \delta \cdot \kappa)} \tag{12}$$

where A is the contributing area, *Tarboton* [1997], and κ is the curvature (of iso-height contours for our examples). In the steep relief landscapes we considered herein, the parameters α and δ were set to 1, giving the same weight to both contributing area and curvature. The selection of other parameter values that give more (or less) penalty to contributing area vs. curvature is a topic of future investigation, as the optimal values might depend on different terrain characteristics.

Applying the above described algorithm to the regularized and thresholded data sets of the two basins, the channel networks were automatically extracted very efficiently. Fig. 10 shows the extracted network obtained from the Northern subbasin in the South Fork Eel River basin. Although for this basin no field data are available to test the performance of the algorithm, we only show it here for demonstration purposes and point out that the whole channel network extraction processing takes less than 10 mins if run on a laptop computer. Fig. 11 shows the extracted channel network obtained for the Skunk Creek.

As discussed before, this is a challenging basin for automatic river network extraction due to many interruptions due to landslides and debris flows. Nevertheless, the automatically extracted river network shown in Fig. 11 compares well with the field-monitored river network. Recall that the only information that was externally provided was the threshold area of 1000m^2 and the values of the parameters α and δ (which were set to 1). Point by point comparison with the real network is not attempted as further work is needed anyway to more accurately register all real channels by surveying.

To demonstrate the different usage of the contributing area threshold used in our pre-processing and the channel initiation area threshold used in the available channel network extraction algorithms, the network obtained through our proposed methodology is plotted in red on 3m contours and is compared with the one obtained using *Taudem* [*Tarboton*, 2002], plotted in black, both extracted using a value of $A = 1000$ m^2. As it can be observed, our methodology prevents the detection of many channels not present in the hand-drawn map. Also as discussed earlier, our algorithm allows the detection of channel disruptions (see Fig. 7 and Fig. 9) which are depicted in the skeleton and can be kept before the geodesic optimization is performed. In the application presented here, we traced continuous channels to the farthest end points detected, but the user can decide to keep some of the disruptions shown in the skeleton, if they correspond to actual channel interruptions in the field. This is not possible using a global extraction algorithm, as channels are always traced continuously and channels disruptions are not detected.

4. Conclusions

In this paper we introduced a geometric framework for the extraction of channel networks from LiDAR data. The proposed approach includes two main components: the pre-processing of the data via nonlinear diffusion, to reduce noise and enhance features that are relevant to the network extraction, and the computation of channel networks in the filtered data via geodesic curves that incorporate geomorphological knowledge such as contributing area and (geometric) curvature. Even though a complete validation of the extraction methodology still needs to be performed through comparison with several field-mapped real river networks, the methodology presented in this paper has been demonstrated to be computationally efficient and able to detect, not only channels, but also the presence of channel disruptions.

This work, which introduces the idea of approaching geomorphological analysis as a geometric task, opens the door to many problems in the automatic extraction of information from LiDAR data. For the particular case of channel networks, it is important to study the possible benefits of using other nonlinear equations for pre-processing and the introduction of additional features in the geodesic penalty function. Similarly, the exploitation for geomorphological analysis of other models which are popular in the partial differential equations and variational formulations in image processing community, such as the Mumford-Shah functional [*Mumford and Shah*, 1989], is of great interest. For example, the channel networks can be considered as discontinuity fields and outliers, and as such automatically computed by such an approach [*Sapiro*, 2001]. Beyond this, channel networks are just one of the many important structures in landscapes, and the exploration of the geometric approach here initiated for the extraction of other geomorphic features, such as landslides, debris flow regions, ravines, etc., is a subject of future research.

Acknowledgments. We thank Prof. Jean-Michel Morel for suggesting to incorporate the contributing area flow in the geodesic penalty function, which turned out to be critical in order to achieve the high quality results here reported. We double thank him for initiating the visit of T. Do Trung to the University of Minnesota to collaborate in this project. This work has been supported by an NSF CDI grant to E.F.G., G.S. (EAR-0835789) as well as by the STC program of the NSF via the National Center for Earth-surface Dynamics (NCED) under agreement Number EAR-0120914NGA. Other support to G. S. by NSF, ONR, DARPA, and ARO is acknowledged gratefully. Computer resources were provided by the Minnesota Supercomputing Institute, Digital Technology Center, at the University of Minnesota. Joel Scheingross and Eric Winchell are thanked for providing a difficult to make channel map of Skunk Creek. We thank David Olsen for his expert help with the manuscript preparation.

References

Almansa, A., F. Cao, and B. Rouge, Interpolation of digital elevation models via partial differential equations, *IEEE Trans. on Geoscience and Remote Sensing, 40(2)*, 314325, 2002.

Alvarez, L., P. L. Lions, and J. M. Morel, Image selective smoothing and edge detection by nonlinear diffusion, *SIAM J. Numer. Anal.*, *29*, 845–866, 1992.

Black, M., G. Sapiro, D. Marimont, and D. Heeger, Robust anisotropic diffusion, *IEEE Trans. Image Processing 7:3*, 421–432, 1998.

Braunmandl, A., T. Canarius, and H. P. Helfrich, Diffusion methods for form generalization, in *Dynamics of Multiscale Earth Systems*, H. J. Neugebauer and C. Simmer (Eds.), Lecture Notes in Earth Sciences, 97, Springer-Verlag, Germany, 2003.

Burt, P. and E. Adelson, The Laplacian pyramid as a compact image code, *IEEE Transactions on Communication*, *COM-31*, 532–540, 1983.

Costa-Cabral, M. C., and S. J. Burges, Digital elevation model networks (DEMON): A model of flow over hillslopes for computation of contributing and dispersal areas, *Water Resources Research*, *30*(6), 1681–1692, 1994.

Dial, R. B., Algorithm 360: Shortest-path forest with topological ordering, *Communications of the ACM*, *12*,1969.

Dietrich, W. E., D. R. Montgomery, S. L. Reneau, and P. Jordan, The use of hillslope convexity to calculate diffusion coefficients for a slope dependent transport law, *EOS Trans. AGU*, *69* (16), 316–317, 1988.

Dietrich, W. E., C. J. Wilson, D. R. Montgomery, J. McKean, and R. Bauer, Erosion thresholds and land surface morphology, *The Journal of Geology*, *20*, 675–679, 1992.

Dietrich, W. E., C. J. Wilson, D. R. Montgomery, and J. McKean, Analysis of erosion thresholds, channel networks and landscape morphology using a digital terrain model, *The Journal of Geology*, *3*, 161–180, 1993.

Dijkstra, E., A note on two problems in connection with graphs, *Numerische Math.*, *1*,269–271,1959.

Do Carmo, M. P., Differential Geometry of Curves and Surfaces, *New Jersey, Prentice Hall*, 1976.

Giannoni, F., G. Roth, and R. Rudari, A procedure for drainage network identification from geomorphology and its application to the prediction of the hydrologic response, *Advances in Water Resources*, *28*(6), 567–581, 2005.

Hancock, G. R., and K. G. Evans, Channel head location and characteristics using digital elevation models, *Earth Surface Processes and Landforms*, *21*(7), 809–824, 2006.

Heine, R. A., C. L. Lant, and R. R. Sengupta, Development and comparison of approaches for automated mapping of stream channel networks, *Ann. Assoc. Am. Geogr.*, *94* (3), 477–490, 2004.

Helmlinger, K., P. Kumar, and E. Foufoula-Georgiou, On the use of DEM data for Hortonian and fractal analyses of channel networks, *Water Resour. Res.*, *94* (3), 477–490, 2004.

Helmsen, J., E. G. Puckett, P. Collela, and M. Dorr, Two new methods for simulating photolithography development in 3D, *Proc. SPIE microlithology*, IX, 253, 1996.

Howard, A. D., A detachment-limited model of drainage basin evolution, *Water Resour. Res.*, *29* (8), 2599–2613, 1993.

Howard, A. D., Badlands, in *Geomorphology of Desert Environments*, A. D. Abrahams and A. J. Parsons (Eds.), Chapman and Hall, New York, 213–242, 1994.

Kimmel, R., Numerical Geometry of Images: Theory, Algorithms, and Applications, *Springer, New York*, 209 pp, 2003.

Koenderink, J., The structure of images, *Biol. Cybern.*, *50*, 363–370, 1984.

Lashermes, B., E. Foufoula-Georgiou, and W. E. Dietrich, Channel network extraction from high resolution topography using wavelets, *Geophysical Research Letters*, *34*, L23S04, doi:10.1029/2007GL031140, 2007.

Memoli, F., and G. Sapiro, Distance functions and geodesics on submanifolds of R^d and point clouds, *SIAM Journal Applied Math.*, *65*, 1227–1260, 2005.

Montgomery, D., and W. E. Dietrich, Where do channels begin?, *Nature*, *336*, 232–234, doi:10.1038/336232a0, 1988.

Montgomery, D., and E. Foufoula-Georgiou, Channel network source representation for Digital Elevation Models, *Water Resources Research*, *29*(12), 3925–3934, 1993.

Mumford, D., and J. Shah, Optimal approximation by piecewise smooth functions and associated variational problems, *Communications on Pure and Applied Mathematics*, *XLII*, 577–685, 1989.

Perona, P., and J. Malik, Scale-space and edge detection using anisotropic diffusion, *IEEE-PAMI*, *12*, 629–639, 1990.

Rodriguez-Iturbe, I., and A. Rinaldo, Fractal River Basins. Chance and Self-Organization, *New York: Cambridge University Press*, 528 pp., 1997.

Sapiro, G., Geometric Partial Differential Equations and Image Processing, *New York: Cambridge University Press*, 412 pp., 2001.

Sethian, J., Level Set Methods and Fast Marching Methods, *Cambridge University Press, Cambridge, UK*, 400 pp, 1999.

Solé, A., V. Caselles, G. Sapiro, and F. Arándiga, Morse description and geometric encoding of digital elevation maps, *IEEE Trans. Image Processing*, *13:9*, 1245–1262, 2004.

Tarboton, D. G., R. L. Bras, and I. Rodriguez-Iturbe, The fractal nature of river networks, *Water Resources Research*, *24*(8), 1317–1322, 1988.

Tarboton, D. G., R. L. Bras, and I. Rodriguez-Iturbe, On the extraction of channel networks from digital elevation data, *Hydrological Processes*, *5*(1), 81–100, 1991.

Tarboton, D. G., A new method for the determination of flow directions and contributing areas in grid digital elevation models, *Water Resources Research*, *33*(2), 309–319, 1997.

Tarboton, D. G., Terrain Analysis Using Digital Elevation Models (Taudem), Utah Water Research Laboratory, Utah State University, http://www.engineering.usu.edu/dtarb, 2002.

Tsai, Y. R., L. T. Cheng, S. Osher, and H. K. Zhao, Fast sweeping algorithms for a class of Hamilton-Jacobi equations, *SIAM Journal on Numerical Analysis*, *41* (2), 673–694, 2002.

Tsitsiklis, J. N., Efficient algorithms for globally optimal trajectories, *IEEE transactions on Automatic Control*, *40*, 1528–1538, 1995.

Yatziv, L., A. Bartesaghi, and G. Sapiro, O(N) implementation of the fast marching algorithm, *Journal of Computational Physics*, *212*, 393–399, 2006.

Zhao, H. K., Fast sweeping method for Eikonal equations, *Mathematics of Computation*, *74*, 603–627, 2004.

Witkin, A. P., Scale-space filtering, *Int. Joint. Conf. Artificial Intelligence*, 1019–1021, 1983.

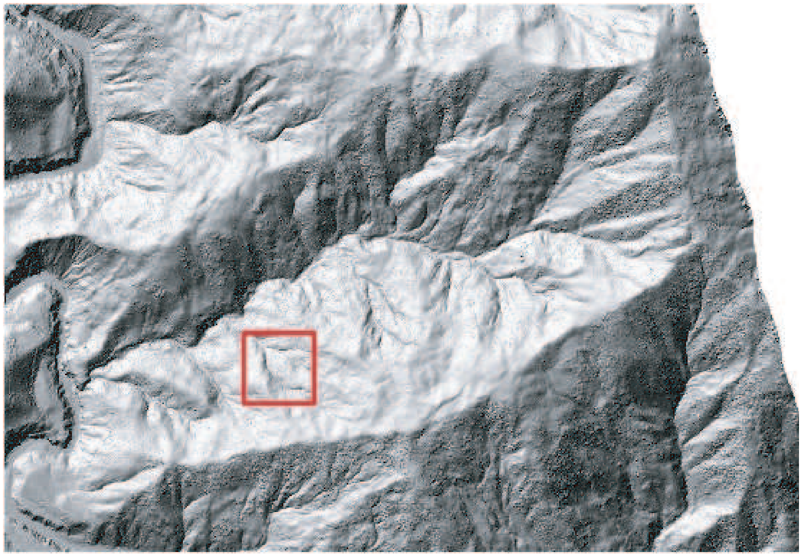

Figure 1. A 2.8 km^2 subbasin in the South Fork Eel River basin in Northern California. The square shows the 300m by 300m section used for illustration of the results.

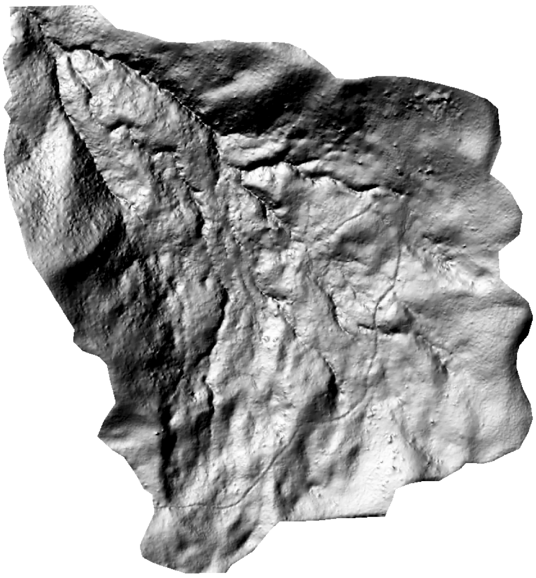

Figure 2. Skunk Creek, a 0.54 km^2 tributary located just upstream of Elder Creek, part of the South Fork Eel River in Northern California.

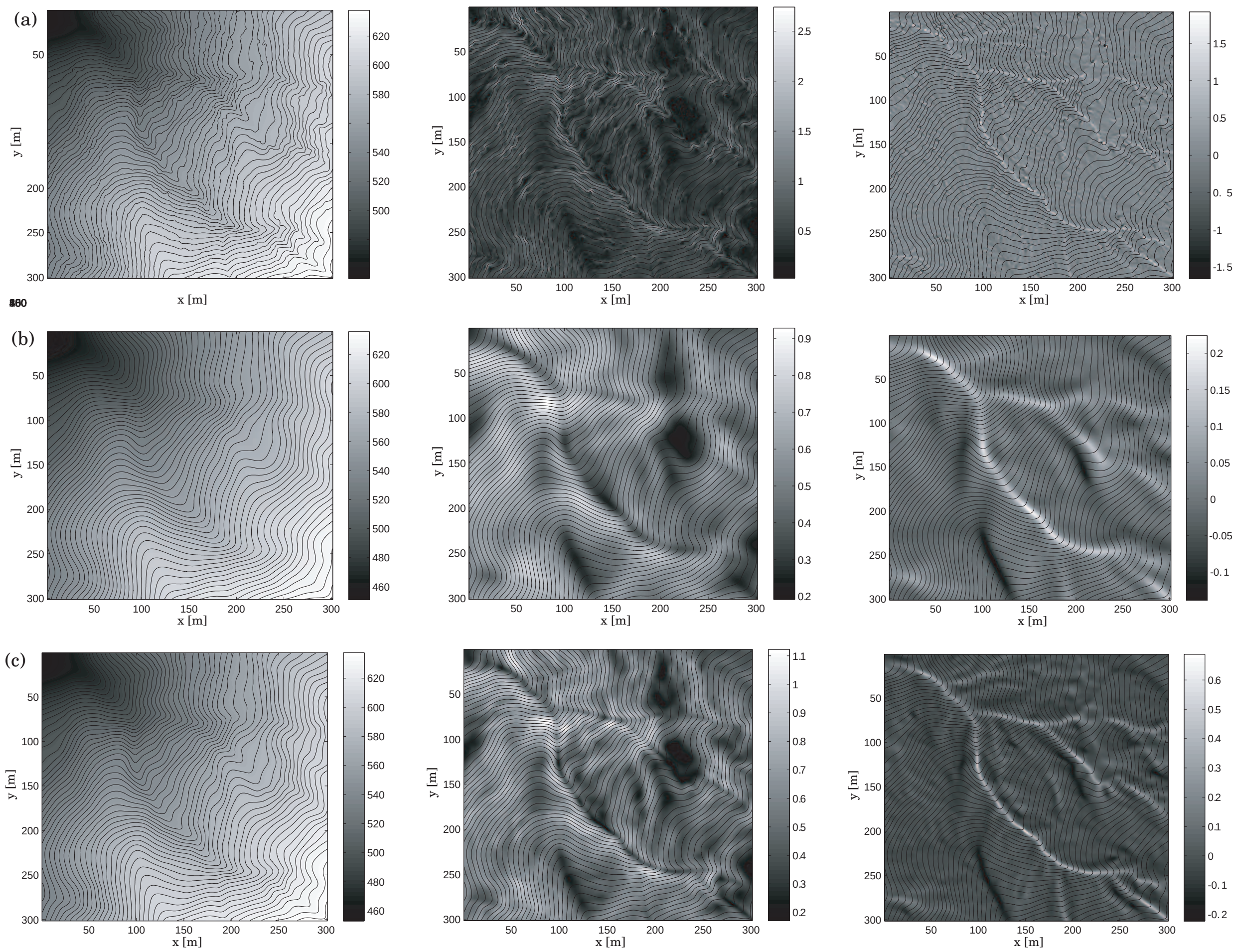

Figure 3. Comparison of the elevation (left), gradient (middle) and curvature (right) between the original data (a), the Gaussian filtered data (scale $\sigma = 7$m) (b) and the Perona-Malik filtered data (50 iterations) (c). In all plots, elevation contours at 3m spacing are superimposed. Notice the sharper localization of the channels in the Perona-Malik filtered LiDAR data.

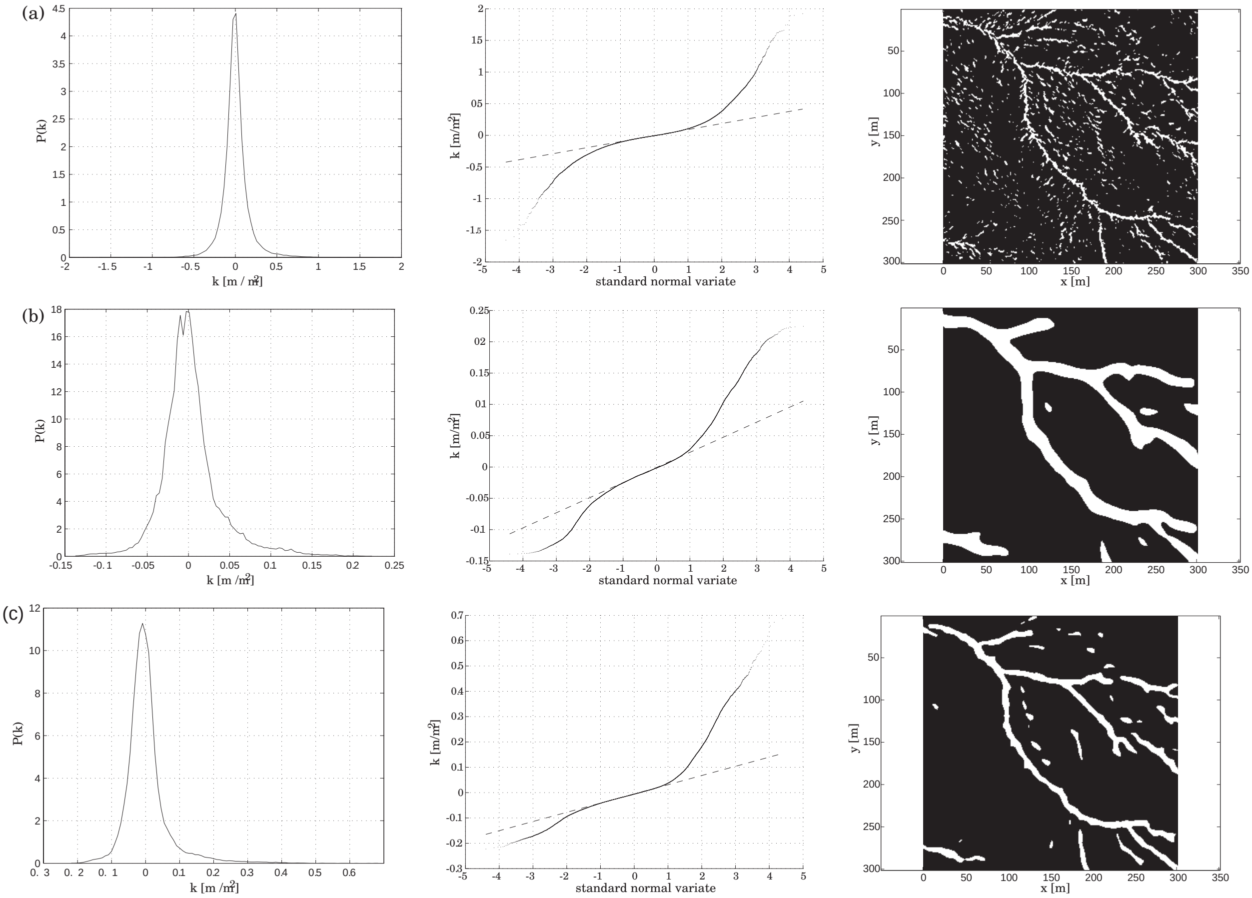

Figure 4. Comparison of the pdfs of curvature (left), q-q plots of curvature from which the threshold value is determined (middle), and skeleton of pixels with above-threshold-curvature for the original data (a), the Gaussian filtered data (scale $\sigma = 7$m) (b), and the Perona-Malik filtered data (50 iterations) (c). The Perona-Malik filter does the best in terms of accurately localizing the channelized valleys while reducing background noise (see text for more discussion).

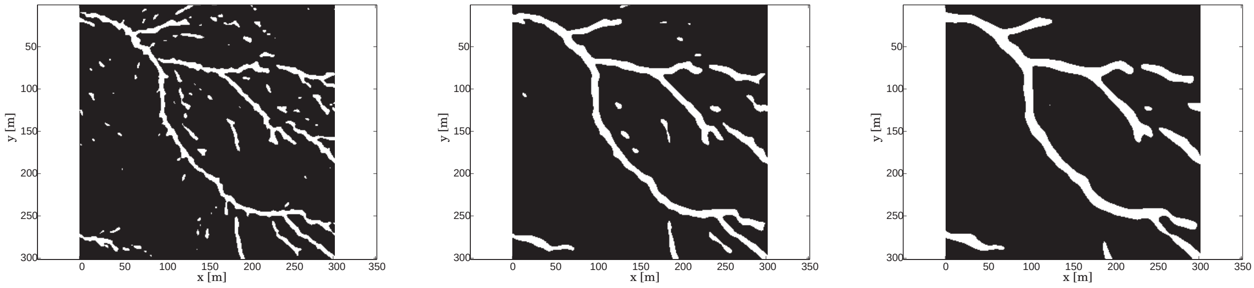

Figure 5. Comparison of the images obtained thresholding the curvature computed on the Gaussian filtered data with $\sigma = 2$m,4m,6m (landscape smoothing scales of 8.9m, 17.8m, and 26.7m) respectively. White pixels indicate pixels with above-threshold curvature.

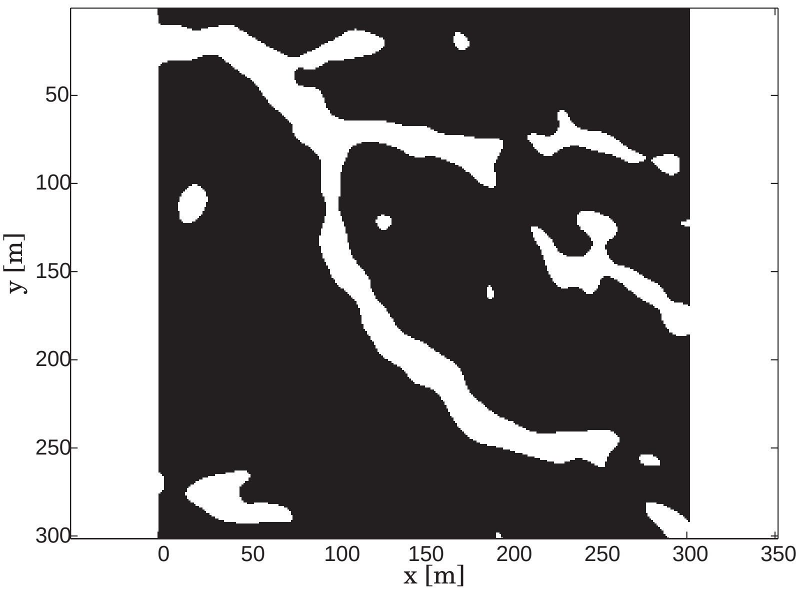

Figure 6. Skeleton of pixels above-threshold-curvature for the Gaussian filtered data using the Laplacian (scale $\sigma = 7$m.)

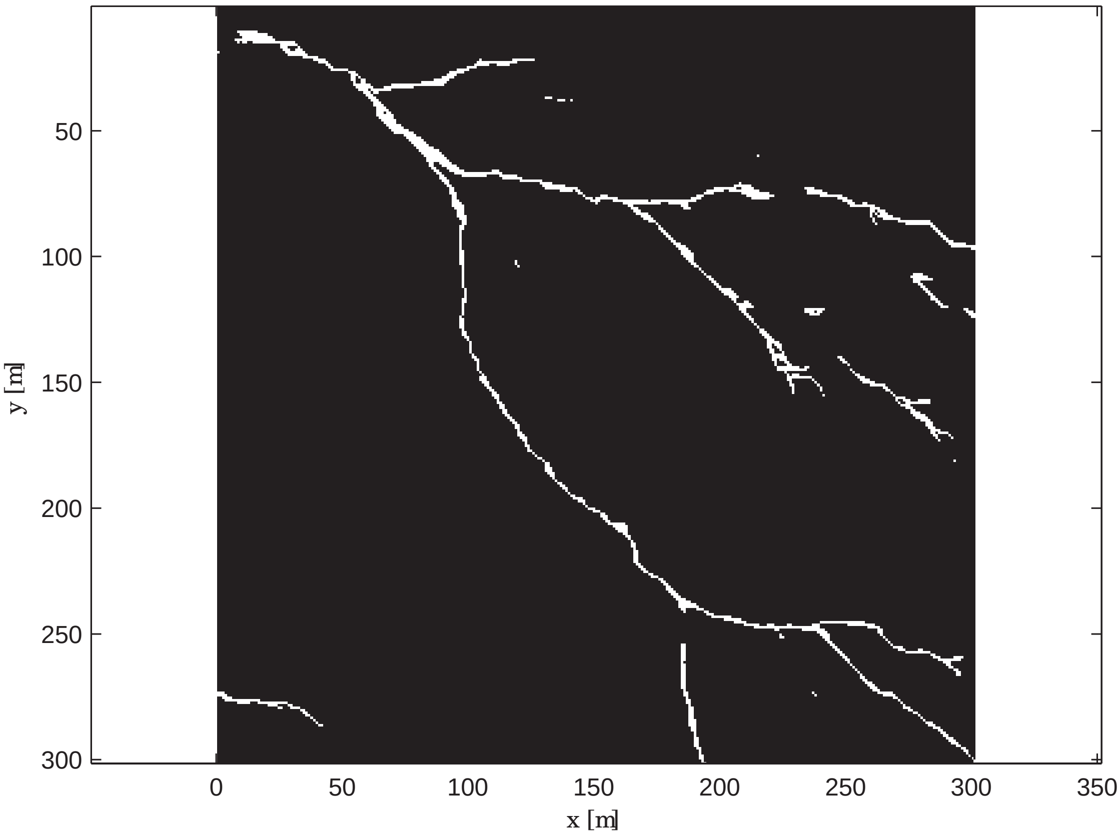

Figure 7. Skeleton obtained by thresholding curvature and contributing area for the portion of the subbasin shown in Fig. 1. Introducing the contributing area criterion eliminates all the isolated pixels which have a large curvature, but are not part of the channel network.

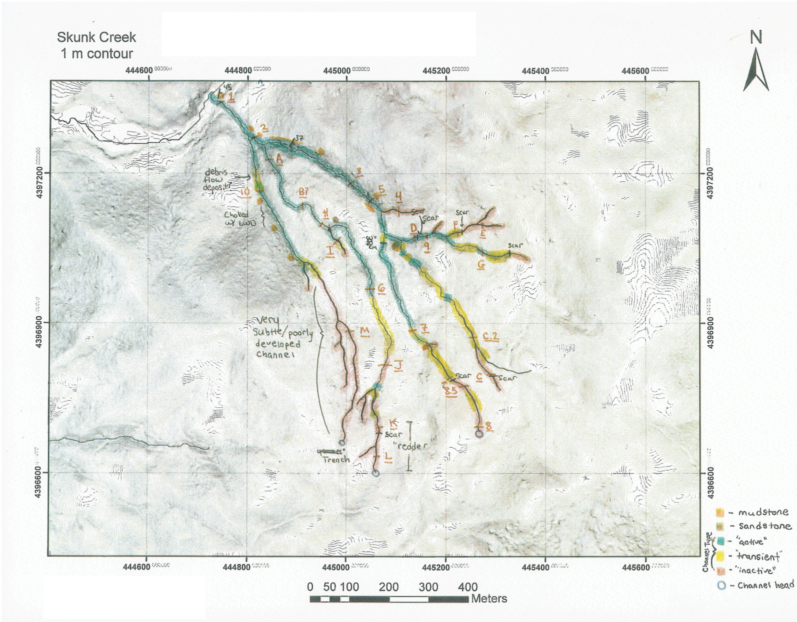

Figure 8. Hand drawn channel network map of the Skunk Creek from the field survey by Joel Scheingross and Eric Winchell (University of California, Berkeley).

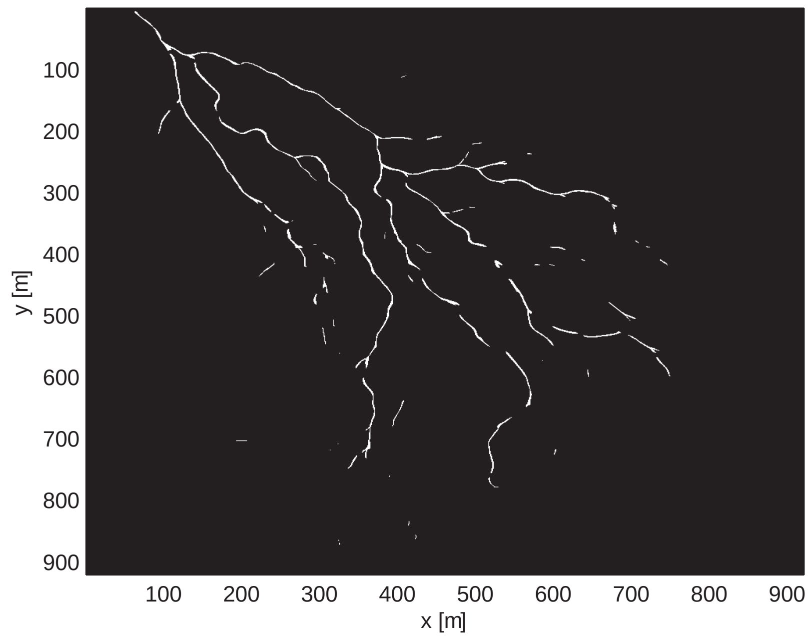

Figure 9. Skeleton obtained by thresholding curvature and contributing area for the Skunk Creek.

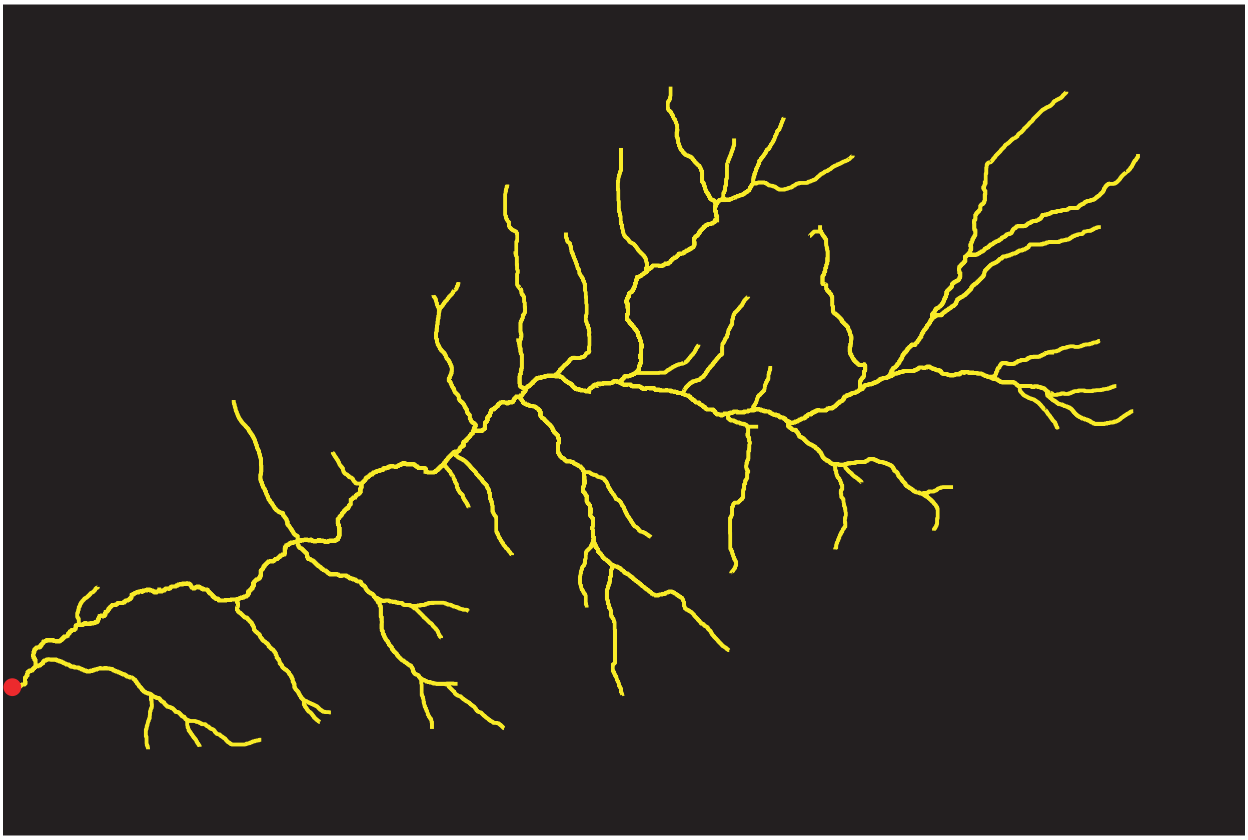

Figure 10. Automatically extracted river network for the sub-basin shown in Fig. 1, using the geodesic optimization on the Perona-Malik filtered landscape.

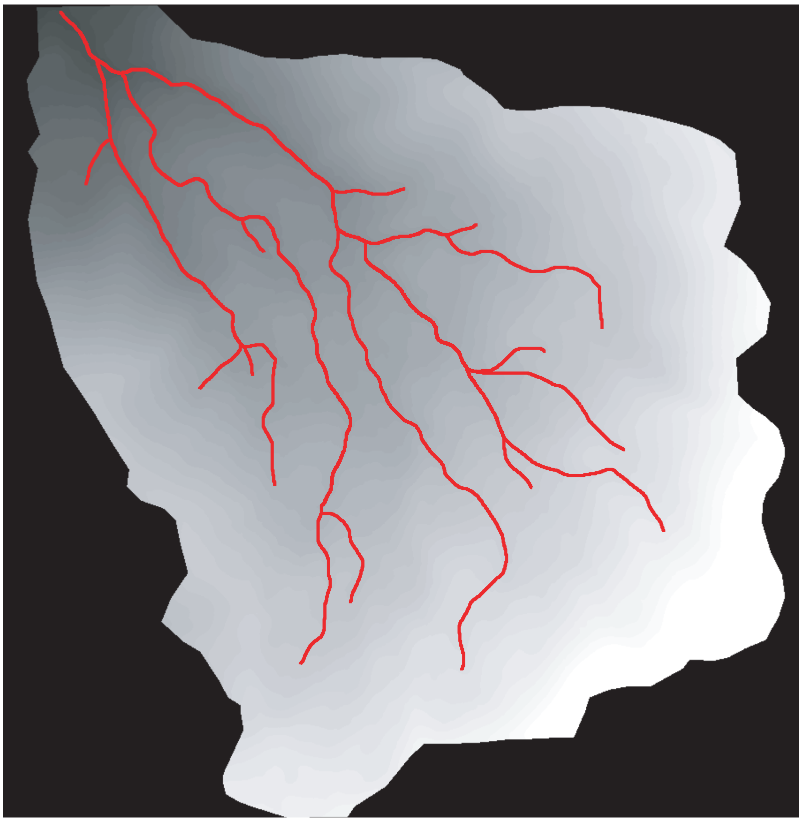

Figure 11. Automatically extracted river network for the Skunk Creek shown in Fig. 2, using the geodesic optimization on the Perona-Malik filtered landscape.

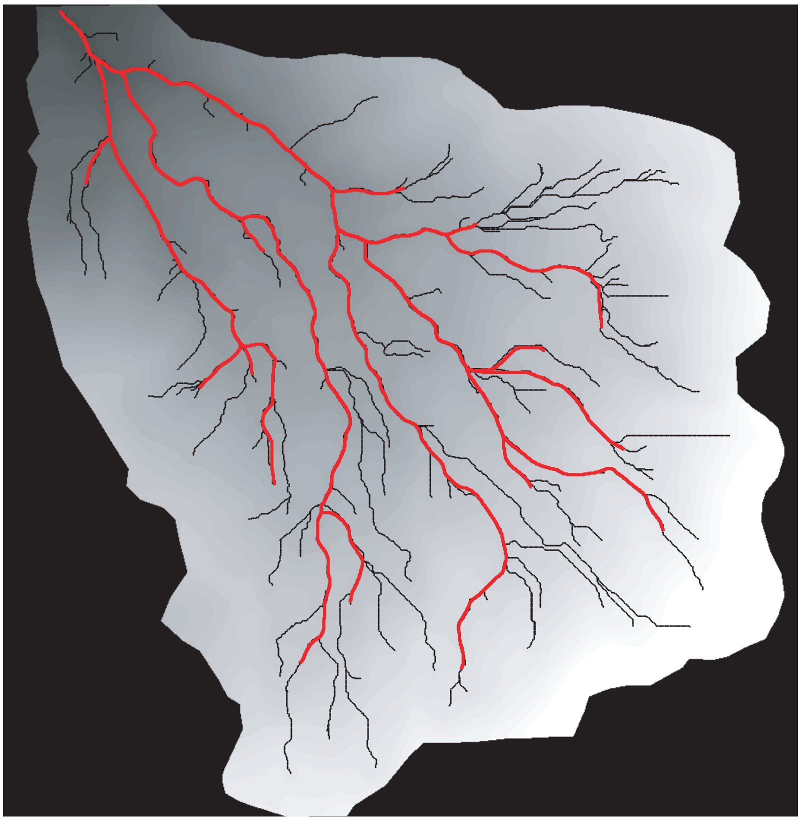

Figure 12. Automatically extracted river network for the Skunk Creek shown in Fig. 2, using the geodesic optimization on the Perona-Malik filtered landscape. The network is plotted in red and compared with the network obtained using *Taudem*, plotted in black, using the same threshold contributing area of 1000m^2.