TecLines: A MATLAB-Based Toolbox for Tectonic Lineament Analysis from Satellite Images and DEMs, Part 2: Line Segments Linking and Merging

Mehdi Rahnama 1,2,* and Richard Gloaguen 1,2

1 Remote Sensing Group, Institute of Geology, TU Bergakademie Freiberg (TUBAF), Bernhard-von-Cotta-Str. 2, Freiberg (Sachsen) 09599, Germany
2 Remote Sensing Group, Helmholtz Institute Freiberg of Resource Technology, Halsbruecker Str. 34, Freiberg (Sachsen) 09599, Germany; E-Mail: r.gloaguen@hzdr.de

* Author to whom correspondence should be addressed; E-Mail: meh.rahnama@gmail.com; Tel.: +49-373-139-4621; Fax: +49-373-139-3599.

External Editors: Andreas Vassilopoulos and Prasad S. Thenkabail

Received: 14 May 2014; in revised form: 11 September 2014 / Accepted: 29 October 2014 / Published: 18 November 2014

Abstract: Extraction and interpretation of tectonic lineaments is one of the routines for mapping large areas using remote sensing data. However, this is a subjective and time-consuming process. It is difficult to choose an optimal lineament extraction method in order to reduce subjectivity and obtain vectors similar to what an analyst would manually extract. The objective of this study is the implementation, evaluation and comparison of Hough transform, segment merging and polynomial fitting methods towards automated tectonic lineament mapping. For this purpose we developed a new MATLAB-based toolbox (TecLines). The proposed toolbox capabilities were validated using a synthetic Digital Elevation Model (DEM) and tested along in the Andarab fault zone (Afghanistan) where specific fault structures are known. In this study, we used filters in both frequency and spatial domains and the tensor voting framework to produce binary edge maps. We used the Hough transform to extract linear image discontinuities. We used B-spline as a polynomial curve fitting method to eliminate artificial line segments that are out of interest and to link discontinuous segments with similar trends. We performed statistical analyses in order to compare the final image discontinuities maps with existing references map.
Keywords: MATLAB; synthetic dataset; QuickBird2; image discontinuity; edge linking; hough transform; polynomial interpolation; tavares-phadilha; B-spline; statistical analysis

1. Introduction

In recent years, most lineament extraction methods are either based on visual image interpretation by an expert or (semi-) automatic detection by using remote sensing images [1–4]. The automatic methods have resulted in savings of time and improve the objectivity of lineament extraction process [5–12]. In general, automatic lineament extraction methods are based on edge detection techniques that enhance the pixels at the edges on an image, instead of directly extracting edge contours. The frequency and connectivity of the extracted edges by common edge detection methods are strongly affected by the type and spatial resolution of the source datasets, signal to noise ratio (SNR) and the parameters of the edge detection methods [5,9,13]. In other words, most edge detection methods results contain fragmented edges and should be ultimately interpreted visually. We discussed the different methods for detecting potential edge pixels in part 1 [14].

After edge detection, additional processing has to be performed in order to remove the false edge responses and to link the gaps between edges. These methods allow the linearization of edge pixels into continuous contours, using certain criteria such as closeness or some specific geometric properties such as straightness or curvature [15]. Edge linking methods can be classified into two categories [16–18]: (1) local processing methods [15,18–20], and (2) global processing methods [21–27]. In local processing methods, edge pixels are grouped to form edges by considering each pixel’s relationship to any neighboring edge pixels. This method is suitable to link edge pixels in situations where the shape of the edge is unknown. The global processing methods use all edge pixels. Pixels displaying similarities such as same edge geometry are used to find the best fit of a known shape. The global methods do not need to connect the edge pixels. However, these methods have inconveniences: small pieces of edges may be missing, or noise pixels may wrongly be handled as edge segments. The edge linking method that follows in this paper is based on the Hough transform (HT) method [28] and specifically focuses on linking edge pixels that represent tectonic linear features. The Hough transform method is a very robust technique for identifying and linking edge pixels that corresponds to linear features. The most important advantages of the technique are the relative insensitivity to noise, to object occlusion and to missing line parts (gaps) [28–31]. The Hough transform method has been widely used in order to increase the frequency and connectivity of detected lineaments [6,26,32–36], which is an efficient method for detecting straight lines [30,37,38].

However, after the line segments have been extracted using the Hough transform procedure the image discontinuities often need to be linked and/or merged [11]. This process is done by reducing the number of line segments having common characteristics (e.g., length, orientation), and building a new line with stronger characteristics. Several methods have been proposed to link non-overlapping or merging partially or fully overlapping line segments [39,40].

The main goal of this study is to develop a new MATLAB based toolbox (TecLines) for automatic linear image discontinuities mapping from satellite images and digital elevation models (DEM).
TecLines operates in two continues steps. In the first step, we implemented the edge detection procedure that consists of Butterworth band pass filter, edge detection method in the spatial domain and tensor voting framework technique. The first step has been explained in the first part of the TecLines paper series [14]. As described above, most of the edge detection methods are weak in confrontation with noisy images and the edges whose gray levels change slightly, and usually cannot differentiate the edge from a noisy area, which results in detecting false or discontinuous edges.

In this study, these problems are addressed using edge-linking methods. The goal of edge linking is to describe an edge as a linear segment of specified shape and estimate the missing edge pixels from the assumed equation of the curvilinear segment [14]. The specific objective of this study is to develop a procedure that consists of the integration between Hough transform (HT) method, Tavares-Phadilha algorithm [39], and B-spline polynomial curve fitting method [11,41] to extract the curvilinear image discontinuities with consideration of the object lengths and orientations as well as the distance between neighboring line segments.

In this second part of the TecLines paper series, the Hough transform method is described in detail, and its capabilities for automatic linear image discontinuities extraction, analysis and mapping are demonstrated. Also, we demonstrate how TecLines can be used to implement algorithms for linking or merging line segments, which are close to each other and have similar directions [39,40], and produce map of linear image discontinuities and their statistical analysis. This study also shows the implementation of the proposed methods on a synthetic image with known discontinuities and a high resolution satellite images (QUICKBIRD-2) from an active tectonic area: the Andarab fault zone in NE Afghanistan.

2. Data

In this paper, we demonstrate the performance of the TecLines for edge detection, where validation has been performed on a synthetic and a real dataset.

2.1. Synthetic Dataset

Applying the real dataset to detect the edges may obtain an irregular pattern, and hence the accuracy of the edge detection may decrease. Most of the performance evaluation of edge detection methods has been proposed according to the presence or absence of ground truth data [42–48], which are based on the characteristic of the images (i.e., real images, synthetic images) [42,47–49]. Several studies consider the binary edge maps generated by other edge detection methods (such as the Sobel method or the Canny method) to be optimal ground truth data [50,51]. However, this assumes that the binary edge maps generated by those methods are perfect. Moreover, different configurations or implementations of each method might produce different results [44,52]. In addition, a large number of approaches that rely on ground truth use simple synthetic images [53–56], because it is easy to specify the ground truth edge locations [42,43]. In these cases, the edge detection can be quantitatively evaluated based on the known ideal detection considered to be the ground truth [56]. The synthetic Digital Elevation Model (DEM) (Figure 1) used here is the result of landscape evolution algorithm created using set river incision and different uplift rates across tectonic faults. The lineaments and drainage system adapts to the evolving surface conditions. The resolution of the surface development model is independent of that of the
underlying continuum box [57]. The surface model follows the concept of the cellular automaton implemented on a regular Eulerian mesh. It incorporates an effective filling algorithm that guarantees flow direction in each cell computation of discharges and bedrock incision. This model was designed to be employed to synthetic topographies [57].

**Figure 1.** The synthetic Digital Elevation Model (DEM) that is the result of landscape evolution algorithm created using set river incision and different uplift rates across tectonic faults. The drainage system adapts to the evolving surface conditions.

![Synthetic Digital Elevation Model](image)

2.2. Real Dataset

**Study Area and Data**

We evaluated the performance of TecLines on a satellite image of the active Andarab fault zone, northeastern part of Afghanistan with a complex pattern of faults (Figure 2A). The tectonic activity in the northeast of Afghanistan is usually interpreted as a consequence of the collision between the Indian and Eurasian subcontinent [58,59]. The large east-west trending, right-lateral and strike-slip fault zones (for example, the Hari Rud, and Andarab fault systems) are key structures that have played an important role in the westward-extrusion of the Northern Afghan Platform [58–60]. The selected area has a long history of damaging earthquakes [60].

The Andarab fault is dextral and coincides with an approximately 150 km long, east-west-trending valley north of the intersection between the Paghman and Hari-Rud faults. Andarab fault lies in the north of Kabul in a transition region between the aseismic interior of Afghanistan and the eastern plate boundary. In the high valley of Darya, many evidences of recent tectonic movements were observed along the fault trace [58,61]. Wellman [62] also described ridges that have 25 m of dextral displacement. Several surveys were measured between 1880 and 1940 about tectonics of the study area but the tectonic activity is poorly known. Their re-measurement may reveal surface deformation associated with the tectonics of the region. However, as yet none of these early surveys have been re-measured.
Figure 2. (A) location of the study area in northeast Afghanistan; (B) panchromatic band of the Quickbird-2 (1 m spatial resolution) for 2 March 2006 of the study area.

The study area is approximately 10 km². In this study, we used panchromatic band of the Quickbird-2 (1 m spatial resolution) for 2 March 2006 (Figure 2B). This data is in UTM coordinate system, datum “WGS84” and zone 42 N.

3. Methodology

TecLines is a new MATLAB based framework that contains various functions for automatic detection and extraction of linear image discontinuities from satellite images and digital elevation models (DEM). Besides import and export functions that support the raster and vectors in standard file formats, TecLines provides functions for image filtering in the frequency and spatial domains to produce primary binary edge maps. Performing the computation of the Tensor voting framework produces final binary edge maps in TecLines. In addition, TecLines extracts line segments from final binary edge maps by employing standard Hough transform functions. A set of functions serves the grouping and merging line segments, which results in final linear image discontinuity maps. Comparing results with published/non-published linear image discontinuity maps and statistical analysis is also possible in TecLines. An overview of the edge linking procedure using TecLines toolbox can be found in Figure 3.

3.1. Hough Transform (HT)

In this study, we implemented the Hough transform (HT) method [28,29] by writing a set of MATLAB based functions. The Hough transform is a common and effective method for identifying the locations and orientations of straight lines in binary images. The main idea behind the Hough transform method is simple: in the image (x-y space), if all the points \((x_i, y_i)\) are located on a line \(L\), described by Equation (1), then the corresponding sine curves \(p\), described by Equation (2) form a bundle intersecting a point \(l(c, m)\) in the Hough parameter space (Figure 4).
\[ y_i = mx_i + c \]  
\[ c = y_i - mx_i \]  

**Figure 3.** Overview of the essential components of linear image discontinuities extraction and grouping using TecLines.

![Figure 3](image)

Since the parameter \( m \) in Equation (1) has infinite values for vertical lines, a different pair of parameters (Polar Coordinates), denoted \( p \) and \( \theta \), are used for the lines in the Hough transform. The Hough transform in image (x–y space) is defined as

\[ f(\theta, p) = \iint_D F(x,y) \delta(p - x \cos \theta - y \sin \theta) \, dx \, dy \]  

**Figure 4.** Lines through a point in the image space and \((m, c)\) space.

![Figure 4](image)
of the pixel \((x, y)\) is \(F(x, y)\), accordingly the value \(f(\theta, p)\) in the \((\theta, p)\) space is the accumulating gray intensity of the pixels of the corresponding line in the image space. Through accumulating array, the specific parameter \((\theta, p)\) is extracted with set the number of peak value in the accumulator array with respect to the candidate straight line. However, since the lengths of the lines in the \(x-y\) space are different, the peak value extraction in the \((\theta, p)\) space will cause false detection [63]. This problem is solved by normalizing gray intensity value in the \((\theta, p)\) space and setting a threshold for length of acceptable lines [64]. An overview of the Hough transform procedure can be found in Figure 5.

**Figure 5.** Simplified flow chart for Hough transform procedure.

3.2. **Grouping, Linking and Merging Line Segments**

The Hough transform result is a set of line segments with a wide range of length and orientation. We propose to link or merge segments with similar length and orientation in order to form continuous lines. In TecLines, we used two methods, an algorithm proposed by Tavares-Padilha [39] and a B-spline method [11,41] for linking non-overlapping line segments, or for merging line segments partially or completely overlapping. The main advantage of these methods is the consideration of the lengths, orientation and distance between neighboring segments as explicit weights for the grouping and the determination of the orientation and placement of the resulting line.

In this step, we implemented the Tavares-Padilha algorithm for generating an intermediate linear discontinuities map. In the first step, all line segments with azimuth less than a specified threshold are classified together in one class. The search area is defined using a maximum distance from line segments and two circle sectors with positive and negative directions. The radius of circle sectors is obtained by adding half the length of the line segment with the specified value that is selected by the user. The user can also define the opening angle of circle sectors. Then, the Tavares-Padilha algorithm is performed that involves five steps:
(1) Define point \((x_G, y_G)\) as a pair coordinates of the centroid by using the two segment endpoints (four points) and segment lengths:

\[
x_G = \frac{l_i(a_x + b_x) + l_j(c_x + d_x)}{2(l_i + l_j)}
\]

\[
y_G = \frac{l_i(a_y + b_y) + l_j(c_y + d_y)}{2(l_i + l_j)}
\]

where \(a = (a_x, a_y)\) and \(b = (b_x, b_y)\) are the endpoints of segment \(i\), and \(c = (c_x, c_y)\) and \(d = (d_x, d_y)\) are the endpoints of segment \(j\) and \(l_i\) and \(l_j\) are the lengths of segments \(i\) and \(j\), respectively (Figure 6).

The merged line will contain this centroid.

**Figure 6.** (A) Merging of two non-overlapping, (B) partially overlapping and (C) totally overlapping segments by Tavares-Padilha method. The red segments are merged to the green lines.

(2) The orientation of the merged line \((\theta_r)\) is defined as the weighted sum of the orientations of the given segments. If \(|\theta_i - \theta_j| \leq \frac{\pi}{2}\) then

\[
\theta_r = \frac{l_i\theta_i + l_j\theta_j}{l_i + l_j}
\]

else

\[
\theta_r = \frac{l_i\theta_i + l_j(\theta_j - \pi - \frac{\theta_i}{|\theta_j|})}{l_i + l_j}
\]

(3) \((X_G, Y_G)\) coordinate system is defined on the centroid \((x_G, y_G)\). The \(X_G\) axis is parallel to the direction \(\theta_r\) of the merged line.

(4) Coordinates for the endpoints a, b, c and d of both segments in the \((X_G, Y_G)\) coordinate system are determined:

\[
\delta_{X_G} = (\delta_y - y_G)\sin\theta_r + (\delta_x - x_G)\cos\theta_r
\]

\[
\delta_{Y_G} = (\delta_y - y_G)\cos\theta_r - (\delta_x - x_G)\sin\theta_r
\]
where \((\delta_{X_G}, \delta_{Y_G})\) are the coordinates of the point \(\delta\) in the \((X_G, Y_G)\) coordinate system. The endpoints coordinates in the new coordinate system are \(a_{(X,G)} = (a_{X_G}, a_{Y_G}), \ b_{(X,G)} = (b_{X_G}, b_{Y_G}), \ c_{(X,G)} = (c_{X_G}, c_{Y_G})\) and \(d_{(X,G)} = (d_{X_G}, d_{Y_G})\).

(5) The two orthogonal projections over the axis \(X_G\) of the four endpoints \(a, b, c\) and \(d\), which are farther apart, define the endpoints of the merged line [39].

In the final step we fit intermediate line segments with a polynomial curve by using B-spline method. We set the knots \((n + 1)\) as \(p_0, p_1, p_2, \ldots, p_n\) for all line segments \((n)\) within each group. The \(p_0\) and \(p_1\) knots are the starting and midpoints of the first segment, and \(p_{n-1}\) and \(p_n\) are the mid and end points of the last segment. The mid point of the other line segments is \(p_2, \ldots, p_{n-2}\) knots. The polynomial curve \(C(u)\) is called a spline of degree \(n\) with the knots \(p_i, \ldots, p_m\), where \(p_i \leq p_{i+1}\) and \(p_i < p_{i+n+1}\) for all possible \(i\), if \(C(u)\) is \(n - r\) times differentiable at any knot when \(p_i < p_{i+1} = \ldots = p_{i+r} < p_{i+r+1}\):

\[
C(u) = \sum p_i N_i^n(u)
\]

where the \(N_i^n(u)\) is a basis function that form a smooth curve with minimal support and certain continuity properties. Therefore, we defined B-splines \(N_i^n\) with knots \(p_i < p_{i+1}\) for all \(i\) by using Equation (5).

\[
N_i^n(u) = p_i^{n-1}N_i^{n-1}(u) + (1 - p_{i+1}^{n-1})N_{i+1}^{n-1}(u)
\]

where \(p_i^{n-1} = (u - p_i)/(p_{i+n} - p_i)\) is the local parameter with respect to support of \(N_i^{n-1}\) (Figure 7). From the definition above, \(N_i^n(u)\) is a piecewise continuous polynomial of degree \((n)\) that is positive in \((p_i - p_{i+n+1})\) and zero outside of \([p_i - p_{i+n+1}]\).

**Figure 7.** Approximation of line segments into a smooth curve using B-spline.

### 3.3. Accuracy Measurements

In order to evaluate the TecLines toolbox applicability, we considered two different analyses: qualitative and quantitative analyses. In the qualitative analysis, the results of the TecLines were compared with a reference map, which is based on manual extraction of the linear image discontinuities, by visual interpretation. The quantitative analysis is mainly for analyzing the directional, length and positional accuracy of the extracted image discontinuities. For directional accuracy assessment, rose diagram method is utilized to compare results with a reference map (ground truth). It should be noted that the discontinuities length weighted the rose diagrams, and thus, these orientations were influenced by longer discontinuities.
Length and positional accuracy were assessed by arithmetically finding the matching percentages of
the extracted discontinuities within a reference map specific buffer zone. Buffer size is initially based
on the resolution of the dataset used. If a discontinuity is located in the buffer zone of the reference
discontinuities, then it is considered matched. Therefore length accuracy (LA) is defined as the ratio
between the total length of all correctly extracted discontinuities (True positive: $TP$) and the total length
of all the discontinuities in the image reference dataset, which are visually interpreted as true
discontinuities ($TD$):

$$LA = \frac{TP}{TD} \times 100 \quad (12)$$

It is also necessary to measure the percentage of falsely extracted discontinuities (False positive: $FP$).
False positive is defined as the subtraction between $TP$, in terms of length, and all extracted
discontinuities ($AD$). False negative ($FN$) is the total lengths of the discontinuities in the reference map
that are not classified as discontinuities [65,66].

Accuracy requires that discontinuities should be detected as close as possible to their correct
positions. In a given image, the discontinuities positions and lengths can be varying according to
resolution and procedures. In this study, by comparing the discontinuities extracted using TecLines with
a reference map, the accuracy was computed as follows.

$$Ac = \frac{TP}{TP + FP + FN + \frac{TP}{TD}} \times 100 \quad (13)$$

where $Ac$ stands for accuracy. The value of $Ac$ for accurate edge linking methods should be close to
100%. In order to evaluate the performance of the proposed edge linking procedure a reference dataset
is required. The references dataset for both synthetic and QuickBird 2 images are determined based on
manual extraction of the linear image discontinuities (Figure 8).

**Figure 8.** (A) The reference discontinuity map for real dataset that is based on manual
extraction from panchromatic band of QuickBird-2; (B) The reference map of the synthetic
DEM consists in the digitized traces of the modeled discontinuities (black line).
4. Testing and Evaluating TecLines

4.1. Performance Evaluation of the TecLines on a Synthetic Digital Elevation Model (DEM)

Accuracy assessment is based on 21 known discontinuities in the synthetic image (Figure 8B). The total length of these features is 1021 m. The range of parameters to be used in edge linking procedure should be large enough to cover a wide range of extraction results.

To select the optimum set of parameters, we implemented 81 sets of combinations of mentioned parameters for Hough transform step. In the first step we extracted the peak value in the Hough domain (the \((\theta, p)\) space), where cutoff \((H(\%)) \in ((200, 300, 400))\) and the normalization factor \((T) \in ((0.01, 0.05, 0.1))\). The line segments extracted use the minimum distance, pixels, between two edges \((FG) \in ((1, 3, 5))\), and the minimum number of pixels that define the length of expected lines segments \((ML) \in ((3, 5, 10))\). Then, we selected the image discontinuities map as the best results using visual interpretation. In this study, the best results for Hough transform method are produced by: \(H(\%) = ((400)), T = ((0.01)), FG = ((5)),\) and \(ML = ((5)).\) The result of the Hough transform is shown in Figure 9A.

**Figure 9.** (A–C) the line segments extracted by Hough transform, Tavares-Padilha algorithm, and final resulting lineament map was obtained by B-spline method, respectively.

In the next step, the Tavares-Padilha algorithm and B-spline curve fitting method are used to fit polynomial curves between neighboring line segments. We used 81 set of combination parameters for generating intermediate discontinuities map (Figure 9B) using Tavares-Padilha algorithm, where the maximum azimuth difference \((MAD) \in ((10, 15, 20))\), the opening angle of circle segment \((OACS) \in ((10, 20, 30))\), the maximum distance from line segments \((MDLS) \in ((5, 15, 25))\), and the minimum length of accepted segment \((MLAS) \in ((1, 3, 5))\).

Finally, we performed B-spline method using 243 set of combination of the parameters, where the acceptable length for the fitted curve \((ALFC) \in ((10, 20, 30))\), the maximum distance between line segments \((MDBL) \in ((5, 10, 15))\), the maximum azimuth difference between line segments \((MADBL) \in ((5, 10, 15))\) and the minimum accepted length of the line segments \((MALL) \in ((1, 3, 5))\). The final discontinuity maps are shown in Figure 9C.
4.1.1. Qualitative Accuracy Assessment

Figure 9 visually shows the extracted discontinuities results using a synthetic DEM. As can be seen in Figure 9A, Hough transform found most of the references linear discontinuities (Figure 8B) but is affected by background textures. The Tavares-Padilha algorithm has relatively good responses on the specified discontinuities. In this step, small line segments are eliminated, and the numbers of the discontinuities are reduced (Figure 9B). The best result achieved uses B-spline method, and the number of remained line segments is near to the reference dataset (Figure 9C).

4.1.2. Quantitative Accuracy Assessment

In this section, we compared TecLines performance quantitatively on the synthetic DEM with a reference image discontinuities map (Figure 8B). For the directional analysis, rose diagrams for all of the discontinuities maps are produced (Figure 10). The result of the Hough transform (Figure 10A) has the major orientations of E-W and N-S. Also, the minor orientation is indicated in NE-SW. Figure 10B shows the E-W and N-S orientation for extracted discontinuities using Tavares-Padilha algorithm. The orientation of the final discontinuities map resulted by B-spline method (Figure 10C) is E-W which is quite similar to the rose diagram for the reference discontinuities (Figure 10D).

**Figure 10.** (A–C): Rose diagrams for discontinuities extracted by Hough transform, Tavares-Padilha algorithm and B-spline method, respectively. (D): Rose diagram for reference lineament map.

We evaluated the length and overall accuracies of the discontinuities summarized in Table 1. These results are indicated that B-spline method provides more consistent results than Tavares-Padilha
algorithm and Hough transform. There are 21 known lineaments in the synthetic dataset. The total length of these features is 1021 m. The number of extracted line segments by using Hough transform is 456 and the total length of them is 1763 m (Figure 9A). As seen in Figure 9A, 80% of the known discontinuities are correctly detected using Hough transform. The percentage of line segments that are extracted as false positive is around 55%. After Tavares-Padilha algorithm, the numbers of the discontinuities are reduced to 185 line segments with 1320 m total length. (Figure 9B). The length accuracy is improved to 85%, while false positive percentage is reduced to 30%.

**Table 1.** Quantitative measures obtained by Hough transform, Tavares-Padilha algorithm and B-spline method for synthetic dataset. True positive (TP) is the number of correctly extracted discontinuities. False positive (FP) is the number of line segments erroneously classified as discontinuities. False negative (FN) is the amount of line segments that were not classified as discontinuities.

<table>
<thead>
<tr>
<th>Method</th>
<th>TP (m)</th>
<th>FP (m)</th>
<th>FN (m)</th>
<th>Length Accuracy (Matching Percentages) (%)</th>
<th>Overall Accuracy (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hough Transform</td>
<td>817</td>
<td>946</td>
<td>204</td>
<td>80</td>
<td>60</td>
</tr>
<tr>
<td>Tavares-Padilha</td>
<td>868</td>
<td>452</td>
<td>153</td>
<td>85</td>
<td>72</td>
</tr>
<tr>
<td>B-spline</td>
<td>970</td>
<td>223</td>
<td>51</td>
<td>95</td>
<td>90</td>
</tr>
</tbody>
</table>

The number of remaining line segments after B-spline method drops to 43 segments. Their total length is 1193 m. Figure 9C shows that we detected 20 known linear features from the original 21 features. The accuracy assessment shows that 95% of the known discontinuities are detected as true positive and 5% are detected as false negative. The percentage of discontinuities that are false positives is about 20%. According to this result, the accuracy of the results after B-spline is increased to 90%.

4.2. Experimental Results and Accuracy Assessment Using Real Dataset

We used statistical tools (frequency, lengths of discontinuities histograms, and azimuth rose diagrams) in order to evaluate the performance of Hough transform and polynomial interpolation. We analyzed and compared the statistical relationship between the extracted discontinuities and a reference discontinuities map (Figure 9A) that is based on manual extraction method. In addition, we compared the final result of the TecLines with a discontinuities map that is extracted automatically using LINE module of PCI Geomatica [67].

To select the optimum set of parameters for Hough transform, we started with initial 81 sets of combinations of mentioned parameters, where cutoff \((H(\%))\) ∈ ((300, 500, 700)), the normalization factor \((T)\) ∈ ((0.01, 0.05, 0.1)), the minimum distance, pixels, between two edges \((FG)\) ∈ ((1, 3, 5)), and the minimum number of pixels that define the length of expected lines segments \((ML)\) ∈ ((3, 5, 10)) (Figure 11). The 243 initial discontinuity maps are generated from three input binary edge maps by using these sets of parameter combinations. Then, we compared the results and reference map using visual interpretation and selected three discontinuity maps, which have been more similar with reference dataset, as the best results. In this study, the best results for Hough transform method are produced by: \(H(\%) = ((500)), T = ((0.01)), FG = ((3)), and ML = ((5))\).
**Figure 11.** (A–C) The binary edge datasets that are produced by Sobel, LOG and Canny edge detection methods and tensor voting, respectively. (D–F): Hough domains from Sobel, LOG and Canny binary edge maps, respectively. Points on (D–F) images show peak values in matrix H.

HT-derived line segments from three edge map sources demonstrate in Figure 12. The corresponding statistics are summarized in Table 2.

**Table 2.** Statistics of extracted image discontinuities lengths, which are derived from the three binary edge map sources (Sobel, LOG, and Canny) using TecLines, reference discontinuity map (manual extraction) and extracted using LINE module of the PCI Geomatica.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>TecLines Toolbox</th>
<th>Hough Transform</th>
<th>Tavares-Padilha</th>
<th>B-Spline</th>
<th>Manually</th>
<th>PCI</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Sobel</td>
<td>LOG</td>
<td>Canny</td>
<td>Sobel</td>
<td>LOG</td>
<td>Canny</td>
</tr>
<tr>
<td>Mean (m)</td>
<td>32</td>
<td>35</td>
<td>22</td>
<td>98</td>
<td>89</td>
<td>96</td>
</tr>
<tr>
<td>St dev (m)</td>
<td>45</td>
<td>31</td>
<td>30</td>
<td>112</td>
<td>97</td>
<td>105</td>
</tr>
<tr>
<td>Sum (km)</td>
<td>75</td>
<td>84</td>
<td>88</td>
<td>56</td>
<td>43</td>
<td>58</td>
</tr>
<tr>
<td>Min (m)</td>
<td>5</td>
<td>2</td>
<td>2</td>
<td>10</td>
<td>14</td>
<td>9</td>
</tr>
<tr>
<td>Max (m)</td>
<td>353</td>
<td>256</td>
<td>281</td>
<td>540</td>
<td>372</td>
<td>511</td>
</tr>
<tr>
<td>Count</td>
<td>2324</td>
<td>2362</td>
<td>4043</td>
<td>1481</td>
<td>1298</td>
<td>2725</td>
</tr>
<tr>
<td>Range (m)</td>
<td>348</td>
<td>254</td>
<td>279</td>
<td>530</td>
<td>358</td>
<td>502</td>
</tr>
<tr>
<td>Median (m)</td>
<td>11</td>
<td>23</td>
<td>10</td>
<td>180</td>
<td>134</td>
<td>173</td>
</tr>
</tbody>
</table>
The second step concerns the identification, and concatenation of line segments to form one long discontinuity based on collinearity and proximity. We applied Tavares-Padilha algorithm using 81 sets of combination parameters for generating intermediate discontinuities maps. We selected the appropriate combination of the parameters by visual interpretation where the maximum distance from line segments (MDLS) \( \in ((10, 20, 30)) \), the maximum azimuth difference (MAD) \( \in ((10, 13, 15)) \), the radial distance (MAD) \( \in ((10, 15, 20)) \), the opening angle of circle segment (OACS) \( \in ((10, 20, 30)) \), and the minimum length of accepted segment (MLAS) \( \in ((1, 3, 5)) \). Acceptable results are obtained in this study when we set the maximum certain distance from the first line segment less or equal to 20 m, while the maximum azimuth difference between the segments is set to 13°. The appropriate radial distance is set to a distance of 20 m plus half the line segment length and the opening angle is set to 20°. The intermediate discontinuities maps after applying Tavares-Padilha algorithm are displayed in Figure 12. The resultant statistics are summarized in Table 2.

**Figure 12.** (A–C): Extracted line segments using HT for binary edge data sources from Sobel, LOG method and Canny methods, respectively; (D–F): Intermediate discontinuity map after applying Tavares-Padilha algorithm; (G–I): Final lineament maps from polynomial interpolation using B-spline method for Sobel, LOG and Canny edge data sources, respectively. (J): Extracted discontinuities using PCI.
In the next step for fitting polynomial curve between grouped line segments, we performed B-spline method using 243 sets of combination of the parameters, where the acceptable length for the fitted curve (ALFC) ∈ ((40, 50, 60)), the maximum distance between line segments (MDBL) ∈ ((15, 20, 25)), the maximum azimuth difference between line segments (MADBL) ∈ ((5, 10, 15)) and the minimum accepted length of the line segments (MALL) ∈ ((1, 3, 5)). The best result was achieved when we set the length of the extension to 40 m, the maximum distance to 20 m, the maximum azimuth difference to 15°, and minimum accepted length of the line segments to 5 m. The final discontinuities maps are displayed in Figure 12 and corresponding statistics are summarized in Table 2.

4.2.1. Qualitative Accuracy Assessment

As shown in Figure 12 and statistical analysis results summarized in Table 2, linear image discontinuities derived from the Sobel, LOG and Canny methods binary edge map sources are similar in the general shape and spatial distribution of the extracted discontinuities. The visual comparison of these results and a reference map shows the result of the Canny method is more acceptable than the results of the other two methods (Figure 12).

4.2.2. Quantitative Accuracy Assessment

We compared TecLines performance quantitatively on the real dataset with a reference discontinuities map, which is obtained by manual extraction. For the directional analysis, rose diagrams for all of the discontinuities maps are produced (Figure 13). The rose diagrams in Figure 13A–C show the similarity...
in major orientation azimuth for HT line segments from all three data sources. The N-S and E-W azimuthal direction are dominant in these rose diagrams, but NE-SW and NW-SE azimuthal directions for discontinuities derived automatically from the LOG data source are remarkable. However, these rose diagrams indicated that the frequency of the discontinuities in the major orientation is different. The highest frequencies in these orientation occurred in the results of the Canny method. The major orientations of the Tavares-Padilha algorithm for Sobel and Canny results (Figure 13D,F) are similar, which are N-S and E-W. The rose diagram for the LOG data source is indicated by a different pattern in the azimuthal orientations, and the azimuthal orientation extended in all directions. The distribution of azimuth directions for discontinuities extracted by polynomial interpolation method and a reference discontinuity map is similar. The major orientations in TecLines results are NE-SW and E-W azimuthal directions, but the frequency of discontinuities in that azimuthal direction in extracted discontinuities by TecLines are more than the frequency of discontinuities in manually and automatically (PCI) extracted discontinuities. In other words, TecLines is more capable than the visual interpretation method for discontinuities mapping, because TecLines is able to detect small discontinuities that cannot be recognized by visual discontinuities extracting and mapping.

**Figure 13.** (A–C): Rose diagram for extracted line segments by Hough transform from binary edge maps produced by Sobel, LOG and Canny methods, respectively; (D–F): Rose diagrams for intermediate discontinuities map extracted using Tavares-Padilha algorithm from three data sources (Sobel, LOG and Canny); (G–I): Rose diagrams for final discontinuities map extracted using B-spline method from three data sources (Sobel, LOG and Canny). (J) and (K): Rose diagrams for manually and automatically (PCI) extracted discontinuities, respectively.
The range of length of the final extracted discontinuities using Sobel, LOG and Canny methods are 695, 580 and 667 m, respectively. As can be seen in Figure 14, the most of the extracted discontinuities are between 200 and 400 m in all of the three methods. The discontinuities with length between 200 and 400 m are 52%, 65% and 68% of the extracted discontinuities using Sobel, LOG and Canny methods, respectively. The total number of the extracted discontinuities after applying LINE module is 85 with range in length from 158–1762 m, and with a geometric mean length of 379 m. Eighty percent of the discontinuities have lengths between 200 and 400 m.

Figure 15 shows the superimposition of linear image discontinuities extracted from Canny data sources and the reference discontinuities, which are extracted manually and automatically by LINE algorithm in PCI Geomatica software. There is a good agreement in orientation and position between extracted discontinuities using Canny method, the result of the LINE algorithm and the reference map.

We assessed the length and overall accuracy of the discontinuities and also evaluated the matching percentages of the results within a reference fault map specific 50 m buffer zone. The buffer size was initially based on the resolution of used dataset (QuickBird with 1 meter resolution). As seen in Table 3, the length accuracy (matching percentages) for Sobel, LOG, Canny and PCI approaches are 70%, 61%, 81% and 72%, respectively. These results are indicated that Canny method correctly extracted 81% of the known discontinuities and provides more consistent results than the other methods. However, the false positive achieved using PCI is the lowest but the overall accuracy of the Canny method is greater than other methods.
**Figure 14.** (A–C): Frequency of extracted discontinuities length by polynomial interpolation method from Sobel, LOG and Canny data sources, respectively; (D-F1): Frequency of length for automatically (PCI) extracted discontinuities; (D-F2): Enlarged image of (D-F1). (E-G1): Frequency of length for manually extracted discontinuities; (E-G2): Enlarge the image of (E-G1).
Figure 15. Superimposition of discontinuities extrapolated from Canny data sources (black lines) and the reference discontinuities, which are manually extracted (green lines), and automatically lineaments extracted by PCI Geomatica software (violet lines).

Table 3. Quantitative measures obtained by TecLines and PCI for panchromatic band of QuickBird-2. True positive (TP) is the number of correctly extracted discontinuities. False positive (FP) is the number of line segments erroneously classified as discontinuities. False negative (FN) is the amount of line segments that were not classified as discontinuities.

<table>
<thead>
<tr>
<th>Method</th>
<th>TP (km)</th>
<th>FP (km)</th>
<th>FN (km)</th>
<th>Length Accuracy (Matching Percentages) (%)</th>
<th>Overall Accuracy (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sobel</td>
<td>31</td>
<td>11</td>
<td>13</td>
<td>70</td>
<td>62</td>
</tr>
<tr>
<td>LOG</td>
<td>27</td>
<td>8</td>
<td>17</td>
<td>61</td>
<td>56</td>
</tr>
<tr>
<td>Canny</td>
<td>36</td>
<td>9</td>
<td>8</td>
<td>81</td>
<td>73</td>
</tr>
<tr>
<td>PCI</td>
<td>32</td>
<td>6</td>
<td>12</td>
<td>72</td>
<td>67</td>
</tr>
</tbody>
</table>

An analysis of the processing time is carried out on a MacBook Pro (Intel Core 2 Duo with 2.66 GHz) with 8 GB of RAM, and with an NVIDIA GeForce 320 M graphics card. Due to the large numbers of parameters used, which increases the number of iteration during image discontinuities extraction, the average computational costs for every individual step are analyzed. These parameters are set as a trade-off between the risk of selecting the wrong discontinuity on the image and the computational cost. The average computational costs for every individual step for TecLines (Canny method) and PCI Geomatica are shown in Figure 16 and summarized in Table 4. Note that the image discontinuity extraction procedure using TecLines consists of seven steps, whilst LINE module of the PCI Geomatica performed consists of three steps. Therefore, the total computational time for TecLines is more than PCI but in similar steps such as edge detection, performing time for the TecLines is less than PCI. In addition, the time necessary for a manual image discontinuities extraction from panchromatic band of the QuickBird 2 with 1 m resolution, and using ArcGIS 10 is about 30 min.
Figure 16. Comparison of run time between TecLines and PCI steps on the same datasets.

Table 4. Computational time for image discontinuities extraction using TecLines toolbox, and LINE module of the PCI Geomatica software.

<table>
<thead>
<tr>
<th>Step</th>
<th>Time (Sec)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>TecLines</td>
</tr>
<tr>
<td>Frequency filtering</td>
<td>15</td>
</tr>
<tr>
<td>Edge detection</td>
<td>35</td>
</tr>
<tr>
<td>Morphological filtering</td>
<td>20</td>
</tr>
<tr>
<td>Tensor voting framework</td>
<td>65</td>
</tr>
<tr>
<td>Hough transform</td>
<td>50</td>
</tr>
<tr>
<td>Grouping discontinuity</td>
<td>20</td>
</tr>
<tr>
<td>Linking discontinuity</td>
<td>35</td>
</tr>
</tbody>
</table>

5. Concluding Remarks

We developed a new MATLAB based toolbox in order to extract linear image discontinuities from satellite images and DEM. In this second paper of a two-paper series on TecLines, we propose an innovative approach which combines Hough transform method to extract line segments and polynomial curve fitting to linking or merging neighbor line segments and eliminate the small or isolated line segments, which have been applied in the part of the north-east of the Afghanistan (Andarab fault zone).

Extracted linear image discontinuities are correlative with the individual fault (Andarab fault) and the aggregate fault patterns of Hindukush-Pamir area. Comparing the experimental result with the similar automatic (LINE algorithm on PCI software) extraction method shows that the proposed toolbox produces a discontinuity map with higher overall accuracy and matching percentage than PCI when Canny method is used. Note that the PCI used the Canny method in the edge detection step. In addition, the experimental results showed that Canny method correctly extracted 81% of the references.
discontinuities and provides more consistent results than PCI with a 72% matching percentage. The overall accuracies achieved using Canny method and PCI are 73% and 67%, respectively. Note that TecLines may produce different results with different accuracies when using other dataset. These results prove the applicability of the presented toolbox (TecLines) for automatic extraction of linear image discontinuities from satellite images and DEM.

It should be considered that a high efficiency in the extraction of image discontinuities cannot be taken as a proxy for the efficient extraction and statistical analysis of tectonic lineaments. The mapping and interpretation of tectonic lineaments is a very complex research field that requires the integration of methods and approaches. This procedure includes the interpretation of geological, geomorphological and seismological data, which are obtained by multi-scale and multi-temporal analysis. Therefore, TecLines can be used as a preliminary tool for extracting linear image discontinuities prior to a more detailed investigation.

Furthermore, the quality and reliability of extracted linear image discontinuities by TecLines are improved in spite of noise and partial disconnection without any prior knowledge. A time-consuming task is avoided. The experimental results have shown the computational time of the presented toolbox with respect to the LINE module of the PCI Geomatica is acceptable. An observation reveals that time requirements for edge detection and linking steps, as essential components for the overall efficiency, are minimal in comparison to the time required for similar steps in other algorithms.

The proposed toolbox has the flexibility of allowing users to set parameters to meet application requirements. Also, users can write their own MATLAB functions in order to expand the toolbox capabilities. The proposed toolbox is available from the TecLines website [68]. This provides new opportunities for education and research in tectonic lineament analysis.

Acknowledgments

High resolution QuickBird image was freely obtained from the GeoEye Foundation. We would like to thank Louis Andreani for review and editing of this manuscript, and Sarah Schroeder for generating the synthetic digital elevation model.

Author Contributions

Mehdi Rahnama and Richard Gloaguen designed the research. Mehdi Rahnama wrote the manuscript and was responsible for the MATLAB programming, mathematical models and data analysis. Both authors contributed in editing and reviewing the manuscript.

Conflicts of Interest

The authors declare no conflict of interest.

References


© 2014 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution license (http://creativecommons.org/licenses/by/4.0/).