

Detection of Linear Structures in Remote-Sensed Images

Rui Gao and Walter F. Bischof

Computing Science,
University of Alberta
Edmonton, Alberta, CA
{rgao,wfb}@cs.ualberta.ca

Abstract. Over the past decades, considerable progress had been made in developing automatic image interpretation tools for remote sensing. There is, however, still a gap between the requirements of applications and system capabilities. Interpretation of noisy aerial images, especially in low resolution, is still difficult. We present a system aimed at detecting faint linear structures, such as pipelines and access roads, in aerial images. We introduce an orientation-weighted Hough transform for the detection of line segments and a Markov Random Field model for combining line segments into linear structures. Empirical results show that the proposed method yields good detection performance.

1 Introduction

Remote-sensed images provide us with accurate and frequently updated geographical information, which is used in map production and updating, in urban and regional planning, agriculture, forestry, mineral resources studies, and many other areas. The large amount of data necessitates the use of efficient, automated or computer-assisted interpretation methods. Over the past decades, there has been a tremendous effort to make these interpretation systems useful. Unfortunately, there is still a large gap between the requirements of practical applications and what is currently being achieved by automated methods in terms of completeness, correctness and reliability. To solve this problem, several solutions have been investigated. First, many successful systems focus on the design on semi-automated systems, where a human operator is in charge of image interpretation, and the computer acts as an assistant to the operator, taking over simple tasks and returning control to the operator whenever a difficulty in the interpretation process is encountered. Second, the image interpretation systems rely on many specialized modules, each concerned with one specific feature. The work presented here is concerned with one such module, which focuses on the detection of pipelines and associated access roads.

In the past, there have been many attempts to detect road networks in remote-sensed images (e.g., [14,16]). These systems make a number of assumptions about the appearance of roads: roads are elongated structures, road surfaces are usually homogenous, and there is adequate contrast between roads and adjacent

areas. All systems are faced with a number of difficulties, including a complex topology near crossings, bridges, ramps, etc., occlusion by ground objects such as vehicles, shadows, trees, etc., and inadequate contrast due to road texture, illumination conditions, weather conditions, and more. Despite these problems, a number of systems are quite successful and can be used in practical applications (for example, [16]).

The system presented here deals with a problem closely related to road detection, namely the detection of pipelines and associated access roads [6,7]. The detection of both structures is very similar to road detection discussed above, but has some important differences. First, pipelines and access roads are typically visible only as very faint lines (especially in low-resolution aerial images), but they typically extend over long distances. Hence, while the local evidence for these linear structures is usually very weak, the lines are typically straight and become visible through evidence accumulation over long distances. Second, in contrast to normal roads, these structures have a simple topology: there are, for example, no crossings, bridges, ramps and more.

Most linear feature detection techniques are based on two criteria, a local criterion, involving the use of local operators surrounding a target pixel, and a global criterion, incorporating large-scale knowledge about objects. Local operators evaluate local image information using edge or line detectors [11,4]. Local detection techniques are, however, insufficient, and global constraints must be introduced to identify these linear structures. The methods based on global criteria include, for example, dynamic programming [12], snake-based energy minimization [5], and tracking approaches [1,16]. Unfortunately, most of these techniques fail with occlusions and with noisy backgrounds, especially in low resolution images, as is the case in our application.

We present a method for detecting faint linear structures (such as pipelines) by combining a local analysis with a global interpretation. The local analysis is based on a combination of Gabor filters, orientation maps and a weighted Hough transform, and is shown to have a robust detection performance. The global interpretation relies on a Markov random field (MRF) model that includes prior and contextual knowledge to effectively improve the accuracy of linear structure detection. In the following sections, we first discuss the detection of linear structures (Section 2) and the combination of linear structures using global constraints (Section 3). Then we present experimental results (Section 4), and finally we discuss our system and present conclusions (Section 5).

2 Detection of Linear Structures

Figure 1 shows an aerial image of size 1000×1000 pixels with a resolution of 5 meters per pixel, containing a number of natural structures (hills, rivers, lakes) and several human-made structures. Clearly visible in the middle of the image is an S-shaped road, consisting of three straight segments connected by curves. In addition, several long, straight, but faint lines are visible, indicating the presence of pipelines and associated access roads. Existing road tracking systems have no

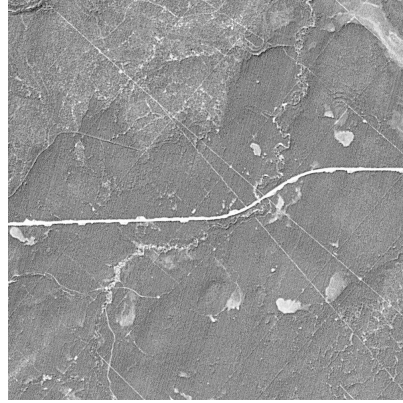


Fig. 1. Example of an aerial image

problem detecting the S-shaped road, but fail to detect the others. The work presented here is concerned with the detection of these linear structures.

Detection of linear structures proceeds in two steps. First, we construct a local orientation map using a bank of Gabor filters. Second, we use the orientation as weight for an orientation-weighted Hough transform to detect all linear structures.

2.1 Local Orientation Map

The input images are filtered with even Gabor filters

$$G(x, y) = \exp\left(-\frac{x'^2 + y'^2}{2\sigma^2}\right) \cos\frac{2\pi x'}{\lambda} \quad (1)$$

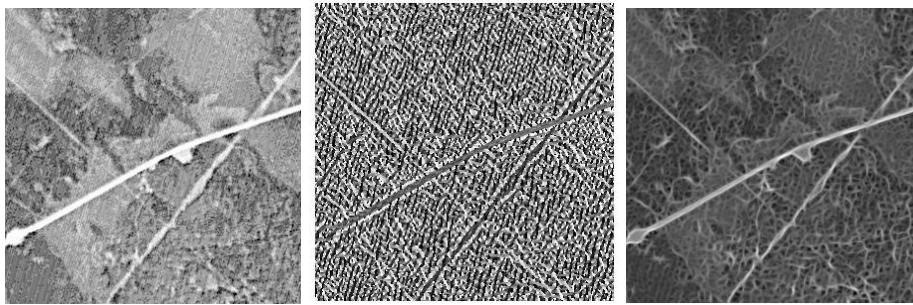


Fig. 2. Left: Aerial image of size 300×300 pixels with resolution of 5 meters per pixel. Middle: Orientation map with 10 orientations denoted by different grey levels. Right: Texture map representing the magnitude of the Gabor filter response.

with $x' = x \cos \theta + y \sin \theta$ and $y' = -x \sin \theta + y \cos \theta$, where λ represents the wavelength of the cosine carrier, σ defines the scale of Gaussian envelope, and θ is the filter orientation. The Gabor output for a line is maximal when θ matches the line orientation. We use a bank of Gabor filters with orientation θ uniformly distributed in the interval $[0, \pi)$.

The orientation map $o(x, y)$ is defined for each pixel (x, y) as the orientation θ of the Gabor filter with maximal response magnitude, and the texture map $g(x, y)$ is defined as the maximal response magnitude. The orientation map represents local image orientation whereas the texture map represents local texture [10]. Figure 2 shows an image of size 300×300 pixel and the corresponding orientation and texture maps. These two characteristics are used in the next step to extract linear segments in the images.

2.2 Orientation-Weighted Hough Transform

The conventional Hough transform [9,8] is used to detect lines by transforming the image space into Hough space, selecting maxima in Hough space, and using these maxima to identify lines in image space. The first step can be considered a pixel-to-curve transformation. Each straight line can be described by

$$\rho = x \cos \theta + y \sin \theta, \tag{2}$$

where ρ is the distance to the origin, θ is the angle between the normal of the line and x -axis, and (x, y) is the pixel position. Using Equation 2, any edge/line pixel in the image space is mapped to a sinusoidal curve in the Hough space, and collinear pixels should pass through the same peak.

In the conventional Hough transform, each pixel in image space votes equally for all possible straight lines through the pixel by increasing the corresponding accumulator cells in Hough space by one. Local maxima in Hough space correspond to line segments in image space. In this process, the detection of straight lines is susceptible to the presence of random noise and spurious line segments, which may generate false maxima in Hough space (see Figure 3b).

It is desirable to reduce the false contributions by introducing a preferential weighting scheme into the voting strategy [13], where the weight of each pixel is used as the value by which the accumulator cell in the Hough space is increased. We use the orientation map to assign weights in terms of how well the orientation in the orientation map matches the line orientation. For instance, if the orientation at a pixel is $o(x, y)$, the contributing weight to accumulator cell (θ, ρ) in Hough space should be large when θ is close to $o(x, y)$. The weight can thus be defined as follows:

$$W_{\theta, o(x, y)} = |\cos(o(x, y) - \theta)|, \tag{3}$$

where $o(x, y)$ is the pixel orientation, and θ is the orientation of the accumulator cell in Hough space. The introduction of this voting strategy can effectively reduce the presence of false alarms in line detection, as illustrated in Figure 3c.

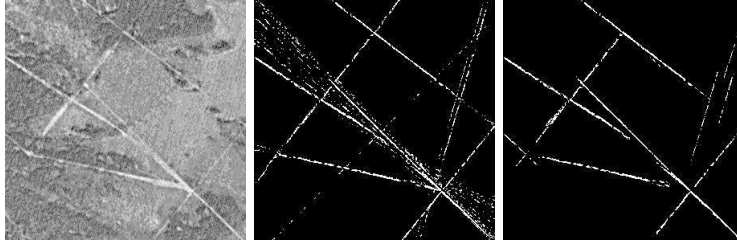


Fig. 3. Line detection results. (a) input image of size 300×300 pixels. (b) Lines detected using a conventional Hough transform. (c) Lines detected using our proposed orientation-weighted Hough transform.

Using the orientation-weighted Hough transform, we detect all possible line candidates. Some of these are false alarms while others are part of true linear structures. In the next Section, we discuss how to label these line candidates and combine them together to identify linear structures.

3 Combination of Linear Structures

The method introduced in the previous Section is able to identify linear structures, but they may be broken into many line segments, as seen in Figure 3. To complete the identification process, these line segments must be combined into linear structures. We begin by defining a graph of all line segments and define a Markov Random field over this graph. Then we discuss local line configurations and the definition of appropriate energy terms. The identification of linear structures can then be formulated as an energy minimization problem.

3.1 Graph Definition

The graph structure G consists of a set S of attributed nodes and an arc set E . The line candidates detected by the Hough transform are elements of S , and the attributes of a node S_k are the length l , the orientation θ , and the position (x_i, y_i) $i = 1, 2, \dots, l$ of the corresponding line segments. The textural property, denoted by g , corresponds to the average Gabor response of all pixels (x_i, y_i) in the line segment.

A neighborhood system is defined on the set S based on the spatial relations. A line segment S_i is regarded as the neighbor of the line segment S_j if and only if the end points of the two line segments are within a distance d_{max} and the orientation difference between the two line segments is at most θ_{max} . Let the neighborhood system on G be denoted by $n = n(S_1), n(S_2), \dots, n(S_N)$, where $n(S_i)$ is the set of all neighbors of node S_i . Then, the graph G is denoted by $G = \{S, E\}$, where $S = \{S_1, S_2, \dots, S_N\}$ and $E = \{S_i S_j | S_i \in n(S_j) \text{ and } S_j \in n(S_i)\}$. Some line candidates in S belong to the true linear structures, such as roads or pipelines, while others are false alarms, e.g. due to a noisy background. Let L denote the

set of labels on node set S , where L_i is the label value associated with node S_i . The label set L is the *random field* and the values of L_i are defined as:

$$L_i = \begin{cases} 1 & \text{if line candidate } S_i \text{ belongs to a true linear structure} \\ 0, & \text{if line candidate } S_i \text{ is a false alarm} \end{cases} \quad (4)$$

The label set $L = \{L_1, L_2, \dots, L_N\}$ takes a value in Ω , the set of all possible configurations. We use a Markov Random Field model to find the optimal configuration. L is a Markov Random Field on graph G with respect to the neighborhood system n if and only if the following conditions, the *Markov Conditions*, are met:

1. $P(L) > 0$ for all configurations, where $L \in \Omega$
2. $P(L_i|L_j, j \in S - S_i) = P(L_i|L_j, j \in n(S_i))$, i.e., the value of random variable L_i for node S_i depends only on the configuration of its neighborhood $n(S_i)$.

Here $P(L)$ is the joint probability and $P(L_i|L_j)$ the conditional probability.

3.2 Energy Function

We need to find the optimal configuration of $L = (L_1, L_2, \dots, L_N)$ that identifies true linear structures and describes their spatial relationships. Given observations $D = (D_1, D_2, \dots, D_N)$ resulting from the orientation-weighted Hough transform, the optimal configuration is defined as the one with the maximum *posterior probability* $P(L|D)$, which is evaluated by the maximum a posteriori estimation (MAP) criterion. According to the Bayes rule, the posterior probability $P(L|D)$ can be written as

$$P(L|D) = \frac{P(D|L)P(L)}{P(D)} \quad (5)$$

where $P(D)$ is a constant. The *conditional probability* $P(D|L)$ and the *prior probability* $P(L)$ stem from a priori knowledge or supervised learning. In our model, we introduce prior knowledge and contextual knowledge to estimate $P(D|L)$ and $P(L)$. The probability distributions $P(D|L)$ and $P(L)$ can be expressed as a MRF-Gibbs fields:

$$P(L) \propto \exp\left[-\sum_{c \in C} U_c(L)\right], \quad (6)$$

$$P(D|L) \propto \exp\left[-\sum_{i=1}^N U(D_i|L_i)\right] \quad (7)$$

where C denotes the clique set of graph G , $U(D_i|L_i)$ is called node potential, and $U_c(L)$ is called clique potential. The optimal labeling corresponds to the minimal energy.

Clique Potential. The clique potential represents contextual knowledge that expresses the conjunction rules for combining line segments. We are concerned with two kinds of linear structures, roads and pipelines. These have the following characteristics:

1. Linear structures are long. They are continuous over a long range, and end-points are rare.
2. The conjunctions of line segments can be curved.
3. Intersections are rare.
4. Lines belonging to the same linear structure have similar textural properties.

The characteristics listed above do not forbid crossroads, other conjunctions, or ends of the lines, but they are assigned a lower probability. As a consequence, a linear structure that can be a road or oil pipeline and can be modeled as a continuous succession of consistent line segments with low curvature and similar textural values. We can now express the energy function defined on a clique c as follows:

1. If all the line segments within clique c do not belong to any linear structure, i.e. they are all false alarms, then all these line segments are labeled 0. The energy is given a value of zero, defined as a stable state.

$$U_c(L) = 0. \quad (8)$$

2. If there is only one line segment S_i labeled 1, it could be the end of a road or pipeline. On the assumption that linear structures are long, a penalty is given for this case, except if the end point of the line segment is close to the image border, since the line segment could belong to a long linear structure that is not completely captured by the image.

$$U_c(L) = K_1, \quad (9)$$

3. If only two line segments S_i and S_j are labeled 1, there are two possible cases: S_i and S_j are two different linear structures or they belong to the same one. On the assumption that the end points of unrelated linear structures are far apart, the former case can be ignored. In the latter case, the energy function depends on how well S_i and S_j match. The energy term includes end point distance, the orientation difference and the texture of the two line segments. On the assumption that roads or pipelines have a low curvature, if two segments S_i and S_j are close to each other with a very small curve and their textures appear similar, we assign them a low energy (high probability); otherwise a high energy:

$$U_c(L) = K_2 \sin|\theta_i - \theta_j| + K_3 d + K_4 (g_i - g_j), \quad (10)$$

where θ is the line segment orientation, g is the textural value of a line segment, and d is the minimal distance between the endpoints of the line segments.

4. When there are more than two line segments labeled $L = 1$ in a clique c , we cannot use curvature or distance between line segments since complex conjunctions allow large curvatures. We calculate the energy based on the textural information, and add a conjunction penalty.

$$U_c(L) = K_4 \text{Var}(g_1 \dots g_n) + nK_5, \quad (11)$$

where n is the number of line segments involved in the conjunctions, and $\text{Var}(g_1 \dots g_n) = \sum_{i=1}^n (g_i - \mu)^2$ is the variance of the textural values of the n line segments, where μ is the mean value.

Node Potential. The node potential is used to evaluate how consistent the measurements are with the labels. A priori knowledge of (single) line segments can be summarized as follows:

1. Line segments are long, i.e. for line segments labeled 1, the longer ones should be given lower energy.
2. Line segments are continuous, with few gaps. Thus, the energy should be higher if line segments have many missing pixels.
3. A linear structure is consistent with respect to textural appearance, and the textural values of the collinear pixels should be similar.

The node potential can thus be defined as follows,

$$U(D_i|L_i) = K'_1 l_i + K'_2 \text{Var}(g) + K'_3 n', \quad (12)$$

where l_i is the length of the i th line segment, $\text{Var}(g)$ is the textural variance, and n' is number of missing pixels. In the MRF model, the energy terms (9)-(11) define the smoothness energy and represent contextual relations; the energy term (12) defines data energy related to the observed data.

3.3 Energy Minimization

The maximum a posterior estimate (MAP) configuration can be considered a global energy minimization problem. Energy minimization approaches have been used widely in image analysis and computer vision. They include, for example, iterated conditional models (ICM) [2], graph cuts [3], simulated annealing algorithm, and loopy belief propagation (LBP)[15]. To estimate the global minimal energy we use ICM. The ICM algorithm has no constraints on the energy form. Moreover, it can be very rapid in practice. The disadvantage is that the results are sensitive to the initial estimates. In our method, the initialization is defined using a length threshold. Line segment candidates that are longer than the threshold are labeled 1; the others are labeled 0.

4 Experimental Results

Experiments were performed on aerial photos with a resolution of 5 meters per pixel. The parameters of the Gabor filters were set to $\sigma=2$, $\lambda=5$ and a bank of 10 filters was used. The parameters of the clique potentials in the Markov Random Field model were $K_1 = 1$, $K_2 = 5$, $K_3 = 0.5$, and $K_4 = 0.001$ and $K_5 = 0.3$; the parameters of the node potentials were $K'_1 = -0.03$, $K'_2 = 0.001$

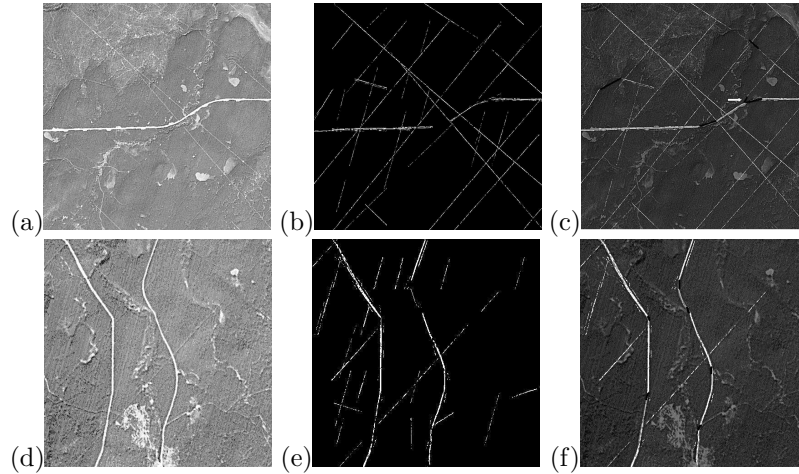


Fig. 4. Detection result. (a)(d) Input image of size 1000 by 1000 pixels.(b)(e) Line segment candidates detected on (a)(d) using Gabor filter and Hough transform. (c)(f) Line labeling result on (b)(e) using Markov Random Field model. Spatially related line segments are connected by black lines.

and $K'_3 = 0.01$. The results of applying the proposed Markov Random Field model-based method are shown in Fig. 4.

The proposed weighted Hough transform using orientation information has good performance in extracting fuzzy, thin lines, shown as in Fig. 4(b)(e). The proposed labeling method using MRF model can reduce false alarms effectively resulted from local analysis. In addition, the spatial meaning of the linear structures can be described correctly. For instance, two white, thick lines in the center of Fig. 4(c) are well connected. However, the connection in the center of Fig. 4(c) pointed by the white arrow is incorrect because of the complex conjunctions.

5 Discussion and Conclusions

In this paper, we proposed a system for the detection of linear features in aerial images. The local line detection is based on the local orientation determined using a set of Gabor filters. An orientation map and maximal Gabor responses are used to represent the spatial and textural distribution. A directionally-weighted Hough transform is used to extract lines using orientation information.

The local properties of the images is related to the global constraints by introducing a priori knowledge. A Markov random field model is built to discriminate true lines from false ones, and to identify the relations between true lines. The optimal interpretation is achieved by minimizing the energy using iterated conditional models. The method has been shown to be a powerful tool to detect fuzzy thin lines, and combine the desired linear structure networks

References

1. Arulampalam, S., Maskell, S., Gordon, N.: A tutorial on particle filters for online nonlinear/non-gaussian bayesian tracking. *IEEE Transactions on Signal Processing* 50(2), 174–188 (2002)
2. Besag, J.: On the statistical analysis of dirty pictures. *Journal of the Royal Statistical Society* 48(2), 259–301 (1986)
3. Boykov, Y., Veksler, O., Zabih, R.: Fast approximate energy minimization via graph cuts. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 23(11), 1222–1239 (2001)
4. Canny, J.: A computational approach to edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 8(6), 679–698 (1986)
5. Fua, P., Leclerc, Y.: Model driven edge detection. *Machine Vision Application* 3(2), 45–56 (1996)
6. Gao, R.: Detection of linear features in aerial images. Master’s thesis, University of Alberta (2009)
7. Gao, R., Bischof, W.F.: Bayesian tracking of linear structures in aerial images. In: *Conference on Computer and Robot Vision* (2009)
8. Gonzalez, R.C., Woods, R.E., Eddins, S.L.: *Digital Image Processing using Matlab*. Pearson, Upper Saddle River (2004)
9. Illingworth, J., Kittler, J.: A survey of the Hough transform. *Computer Vision, Graphics, and Image Processings* 44(1), 87–116 (1998)
10. Manthalkar, R., Biswas, P., Chatterji, B.: Rotation invariant texture classification using even symmetric gabor filters. *Pattern Recognition Letters* 24, 2061–2068 (2003)
11. Barzohar, M., Cooper, D.: Automatic finding of main roads in aerial images using geometric-stochastic models and estimation. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 18, 707–721 (1996)
12. Merlet, N., Zerubia, J.: New prospects in line detection by dynamic programming. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 18(4), 426–431 (1996)
13. Ruwwe, C., Zölzer, U., Duprat, O.: Hough transform with weighting edge-maps. *Visualization Imaging and Image Processing* (2005)
14. Vosselman, G., Knecht, J.: Road tracing by profile matching and kalman filtering. In: *Proc. Workshop Autom. Extraction Man-Made Objects From Aerial and Space Images*, pp. 265–274 (1995)
15. Yedidia, J., Freeman, W., Weiss, Y.: Generalized belief propagation. In: *Neural Information Processing Systems Conference*, pp. 689–695 (2000)
16. Zhou, J., Bischof, W., Caelli, T.: Road tracking in aerial images based on human-computer interaction and Bayesian filtering. *Photogrammetry and Remote Sensing* 6(2), 108–124 (2006)