8-6-2002

Linear feature detection using multiresolution wavelet filters

Samuel Peter Kozaitis

Somkait Udomhunsakul

Rufus H. Cofer

A. Agarawal

Shuwu Song

Linear feature detection using multiresolution wavelet filters

Samuel Peter Kozaitis
Somkait Udomhunsakul
Rufus H. Cofer
A. Agarawal
S.-W. Song

Linear feature detection using multiresolution wavelet filters

S. P. Kozaitis, S. Udomhunsakul, R. H. Cofer, A. Agarawal, and S-W. Song*
Florida Institute of Technology, ECE Dept., *CS Dept.,
150 W. University Blvd., Melbourne, FL 32901

ABSTRACT

We detected roads in aerial imagery based on multiresolution linear feature detection. Our method used the products of wavelet coefficients at several scales to identify and locate linear features. After detecting possible road pixels, we used a shortest-path algorithm to identify roads. The multiresolution approach effectively increased the size of the region we examined when looking for possible road pixels and reduced the effect of noise. We found that our approach leads to an effective method for detecting roads in aerial imagery

Keywords: dynamic programming, feature extraction, multiresolution, wavelet transform

1. INTRODUCTION

Many approaches to linear feature extraction involve the use of low-level vision methods using local neighborhoods followed by a symbolic approach using a global criterion. For example, to detect roads, several approaches use local neighborhoods to detect edges in an image, then a tracking method is used to find roads.[1-11] Often, methods use conventional or custom filters based on edge or line detection. Because the regional or global structure of a road network in usually not considered initially, errors in determining road pixels are unavoidable. A method that would use other than local information in the detection of road pixels or segments could form the basis of a particularly effective approach to road detection.

Multiresolution approaches effectively use larger neighborhoods than those at just one resolution or scale. Generally, these methods are based on knowledge that different characteristics of objects can be best detected in different scales. Features such as roads can be found at different resolutions then combined using a graph or rule-based method.[12-14] Using the results from different resolutions and knowledge of roads, hypotheses for road segments may be generated. Then, segments may be grouped into larger segments or networks. The wavelet transform provides a convenient method for relating information at different image resolutions.[15] For example, it has been shown that the evolution of wavelet local maxima across scales characterizes the local shape of structures.[16] Furthermore, a robust method has been developed to detect edges in noisy images using information across scales.[17] Such an approach uses relatively large neighborhoods for feature detection when compared to conventional approaches, and could be useful if applied to road detection.

We used a wavelet-based multiresolution approach to detect linear features in an image. We examined the response of wavelet coefficients across scales to generate spatial domain filters for linear feature extraction. In the next section we discuss our approach to linear feature extraction We also present an improved shortest-path algorithm for road detection. Finally, we present results on satellite imagery.

2. MULTIRESOLUTION APPROACH

2.1 Wavelet transform and multiresolution

The general idea behind the wavelet transform is to represent an arbitrary function as a superposition of functions of local support. Based on a mother wavelet, scaled and shifted versions of the mother wavelet can be summed to represent an arbitrary function. The wavelet transform maps a signal in the space domain into a scale-

 Battlespace Digitization and Network-Centric Warfare II, Raja Suresh, William E. Roper, Editors, Proceedings of SPIE Vol. 4741 (2002) © 2002 SPIE · 0277-786X/02/$15.00

translation domain using scaled and translated versions of a mother wavelet. The continuous one-dimensional wavelet transform of the function $f(x)$ with respect to a mother wavelet $w_{jk}(x)$ is

$$b_{jk} = \int f(x) w_{jk}(x) dx, \quad (1)$$

where b_{jk} are the collection of wavelet coefficients, and j and k indicate the scale and shift of the wavelet. A dyadic family of wavelets are often used so the wavelet is written as $w_{jk}(x) = 2^{-j/2} w_{jk}(2^{-j}x - k)$, where j = 1, 2, ... In this case, the wavelet is dilated by a factor of two at each increasing scale. In addition, the amplitude of the wavelet decreases by a factor of $\sqrt{2}$ at each increasing scale, so scaled wavelets have the same energy as the original mother wavelet.

In the wavelet transform, the input function $f(x)$ is being compared to a wavelet $w_{jk}(x)$ through a correlation or projection. A wavelet domain coefficient is computed for each particular scale and shift value; the wavelet coefficient is the correlation coefficient between $f(x)$ and $w_{jk}(x)$. Determining these coefficients is the wavelet transform. In other words, a function is approximated by a weighted sum of the scaled and shifted mother wavelet. When summed together the original signal is obtained. In addition, the wavelet transform is linear and superposition holds.

In our approach the wavelet coefficients are generated by a series of filters. Each filter represents a wavelet at a particular scale. In all scales, there may be relatively large wavelet coefficients in the vicinity of roads. However, the size and consistency of the wavelet coefficients across scales varies. A method is needed that allows wavelet coefficients to be characterized across scales. Then, the scales needed to detect features may be determined.

2.2 Linear feature extraction across scales

Mallat showed that the evolution across scales of the wavelet transform depends on the local Lipschitz regularity of the signal[15]. For dyadic wavelets, the uniformly Lipschitz exponent of a signal feature can be estimated from the absolute value of the wavelet transform across scales by,

$$|b_{2^j,k}| \leq K(2^j)^{\alpha}, \quad (2)$$

where $K > 0$, and α is the Lipschitz exponent. The estimate of α gives an indication of the decay of the wavelet transform maxima over a given range of scales. The wavelet maxima that propagate to the next scale may be found by looking at their value and position with respect to other maxima at the next scale.

The Lipschitz exponent has been used to discriminate a signal from white noise because the noise generates a negative exponent, and a signal may have a positive exponent.[16] However, factors such as clutter and the use of a limited number of scales decrease the effectiveness of using the Lipschitz exponent directly as a measure. Instead of calculating the Lipschitz exponent directly, we tracked the response of wavelet coefficients across scales to identify locations of features.

We compared wavelet transform maxima after rescaling the energy at adjacent scales and compared it with a threshold. This is equivalent to detecting features with Lipschitz exponents greater than some threshold value. For example, to detect linear features using j = 1 and 2, a product $C_{12}(k)$ is formed between a signal's nondownsampled wavelet transform coefficients of the two scales, $b_{1k}(k)$ and $b_{2k}(k)$. Then, the energy of the product $C_{12}(k)$ is rescaled so that it has the same energy as $b_{1k}(k)$ such that

$$C_{12}(k) = b_1(k) \cdot b_2(k) \left(\frac{\sum_k |b_1(k)|^2}{\sum_k |b_1(k) \cdot b_2(k)|^2} \right)^{1/2} . \quad (3)$$

A feature is identified at a position where $|C_{12}(k)| > |b_1(k)|$.

After detecting the presence of linear features as described above, we used an iterative process to find other features at the same scale with different magnitudes that may be present. After using Eq. (3), the feature locations are saved in a binary spatial mask, and the values of $C_{12}(k)$ and $b_1(k)$ are set to zero at the feature locations. The energy of $C_{12}(k)$ is again rescaled to that of $b_1(k)$ as in Eq. (3), and a new product is formed between $C_{12}(k)$ and $b_1(k)$. The energy of the new product is rescaled to that of $b_1(k)$ as before, features identified, their locations saved, and another product formed between the new $C_{12}(k)$ and $b_1(k)$. This process is repeated until the energy in $b_1(k)$ reaches some reference noise power in $b_1(k)$. Ultimately, we created a binary mask that indicates the location of detected features.

The individual wavelet transforms in the horizontal and vertical directions can be used to extend the 1-D feature extraction algorithm to images. The block diagram of the algorithm for images is shown in Fig. 1 using four scales of the wavelet transform. First, the wavelet transform is applied to the input image in both the horizontal and vertical directions as shown in Fig. 1(a). Then, the vertical and horizontal results are produced separately and labeled $W_{cv}(i,j)$ and $W_{ch}(i,j)$ respectively in Fig. 1(b). The binary feature location images $W_h(i,j)$ and $W_v(i,j)$ in the horizontal and vertical directions are determined from $W_{cv}(i,j)$ and $W_{ch}(i,j)$ by thresholding them with some reference value as indicated in Fig. 1(c). Features are found by forming the product of the binary masks and the wavelet coefficients from the smallest scale as indicated in Fig. 1(d). Finally, the modulus of the sum of these directional images,

$$M(x,y) = (|W_h(i,j) \cdot b_{1h}(i,j)|^2 + |W_v(i,j) \cdot b_{1v}(i,j)|^2)^{1/2} \quad (4)$$

represents the feature information in Fig. 1(e).

2.3 Implementation

According to Eq. (3) a product is formed between wavelet coefficients. Because feature locations may shift according to scale, this method could produce errors as more scales are considered. Therefore, we formed the product between $b_1(k)$ and shifted versions of $b_2(k)$, and used the maximum values at each location of $b_1(k)$ as the product between scales. For a feature to shift a maximum of s locations, and considering the product of n scales, this method uses the product,

$$C_{1n}(k) = \max\left\{b_{1,k}(k)b_{p,k+t}(k)\right\} \quad (5)$$

where $p = 2, 3, \ldots n$, $t = 0, \pm 1, \pm 2, \ldots \pm s$, and $(2s+1)^{n-1}$ signal multiplications are needed to determine the feature locations. The data in $C_{1n}(k)$ is the result of the iterative process similar to $C_{12}(k)$ in the previous section.

To reduce the number of computations, we approximated the direct multiplications between several scales by forming the products between two scales at a time. For example, using wavelet coefficients from four scales $b_1(k)$, $b_2(k)$, $b_3(k)$, and $b_4(k)$, we found features from the two largest scales $b_3(k)$, and $b_4(k)$. Then, the result was combined with $b_2(k)$. The result of that step was then combined with $b_1(k)$ to determine feature locations. For n scales, the number of signal multiplications is $(2s+1)(n-1)$.

3. LINKING

The result of the previous linear detection process assigns the probability to a pixel that it is part of a road. To create a road, a user selects a start and end pixel with a mouse, and the road is found between the two pixels. The path drawn between the start and end pixels is the shortest path, where shortest path is defined as the sum of pixel values.

Dynamic programming has been very successful for finding an optimal shortest path between two pixels in an image. We used a modified version of Dijkstra's algorithm which is a reverse dynamic programming algorithm.[18] It solves the problem of finding the shortest path by starting at the initial point and working toward a destination point. It works well for a general nonnegative weight digraph. Generally, the algorithm calculates the cost from the start pixel to every other pixel and chooses the path that has the shortest path as defined above.

We modified Dijkstra's algorithm to improve both accuracy and speed. First, the values of the linear feature detection portion were inverted so that bright pixels were assigned a low value, with the lowest value being set to one. Next, the image was soft thresholded. In other words, values above a threshold were set to infinity, and those below the threshold were retained. This approach creates regions where the algorithm will never consider; in effect, this process constrains the number of possible shortest paths. After thresholding, a check for connectivity takes place. If the shortest path between the start and end pixels is infinity, then connectivity does not exist, and the image is low-pass filtered. This process repeats until connectivity exists. Once connectivity exists, then the final shortest path is found. Note that after low-pass filtering, the filtered pixel value is only used if it is greater than the original pixel value, otherwise, the original value is used.

We showed some results using our method in Fig. 2. We used a discrete version of a mexican-hat wavelet on the original image shown in Fig. 2(a). The initial road pixel detection results from Sec. 2 are shown in Fig. 2(b). The image in Fig. 2(c) shows the region of the road-pixel result image where pixels were set to infinity (in white) for the shortest-path algorithm. The result in Fig. 2(d) shows the final result superimposed on the road-pixel result image.

4. CONCLUSION

We found that using a multiresolution approach to detect linear features was effective at determining possible road pixels. Using a shortest-path algorithm and restricting its possible solutions seemed to be form an effective method for finding road networks in images.

REFERENCES

1. R. Bajcsy, and M. Tavakoli, "Computer recognition of roads from satellite pictures," *IEEE Trans. Systems Man Cybernet.*, vol. 6, 623-637 (1976)
2. R. Nevatia, and K. R. Babu, "Linear feature extraction and description," *Computer Graphics and Image Proc.*, vol. 13, 257-269 (1980)
3. M. A. Fischler, J. M. Tenenbaum, and H.C. Wolf, "Detection of roads and linear structures in low-resolution aerial imagery using a multisource knowledge integration technique," *Computer Graphics and Image Proc.*, vol. 15, 201-223 (1981)
4. J. M. Canning, J. J. Kim, N. S. Netanyahu, and A. Rosenfeld, "Symbolic pixel labeling for curvilinear feature detection," *Pattern Recognition Lett.*, vol. 8, 299-310 (1988)
5. D. M. McKeown Jr., and J. L. Denlinger, "Cooperative methods for road tracking in aerial imagery, *in Proc. IEEE Comp. Soc. Conf. on Computer Vision and Pattern Recognition*, 662-672 (1988)
6. N. Merlet, and J. Zerubia, "New prospects in line detection by dynamic programming," *IEEE Trans. Pattern Anal. Machine Intel.*, vol. 18, 426-431 (1996)
7. D. Ziou, "Line detection using an optimal IIR filter," *Pattern Recognition*, vol. 24(6), pp. 465-478 (1991)
8. M. Petrou, "Optimal convolution filters and an algorithm for the detection of wide linear features," *IEE Proceedings-I*, vol. 140(5), (1993)

9. F. Tupin, H. Maitre, J. Mangin, J. Nicolas, and E. Pechersky, "Detection of linear features in SAR images: application to road network extraction," *IEEE Trans. Geosci. Remote Sensing*, vol. 36(2), 434-453 (1998)
10. D. Geman, and B. Jedynak, "An active testing model for tracking roads in satellite images," *IEEE Trans. Pattern Anal. Machine Intel.*, vol. 18, 1-14 (1996)
11. J. C. Trinder, and Y. Wang, "Automatic road extraction from aerial images," *Digital Signal Processing*, vol. 8, 215-224 (1998)
12. H. Mayer, "Automatic knowledge based extraction of objects of the real world from scanned maps," *Int. Archives of Photogrammetry and Remote Sensing*, vol. 30(3/2), 547-554 (1994)
13. A. Baumgartner, C. Steger, H. Mayer, W. Eckstein, and H. Ebner, "Automatic road extraction based on multi-scale, grouping, and context," *Photogrammetric Engineering & Remote Sensing*, vol. 65(7), 777-785 (1999)
14. C. Heipke, C. Stegar, and R. Multahmmer, "A hierarchical approach to automatic road extraction from aerial imagery," *Proc.* SPIE 2486, 222-231 (1995)
15. S. G. Mallat, "Multifrequency channel decompositions of images and wavelet models," *IEEE Trans. Acoustics Speech, and Sig. Proc*., vol. 37(12), 2091-2110 (1989)
16. S. G. Mallat, and S. Zhong "Characterization of signals from multiscale edges," *IEEE Trans. Pattern Anal. Machine Intel.*, vol. 14(7), 710-732 (1992)
17. Y. Lee, and S. P. Kozaitis, "Multiresolution gradient-based edge detection in noisy images using wavelet domain filters," *Optical Engineering*, vol. 39(9), 2405-2412 (2000)
18. T. K. Moon, and W. C. Stirling, "Mathematical Methods and Algorithms," Prentice-Hall:New Jeresy, 787-817 (2000)

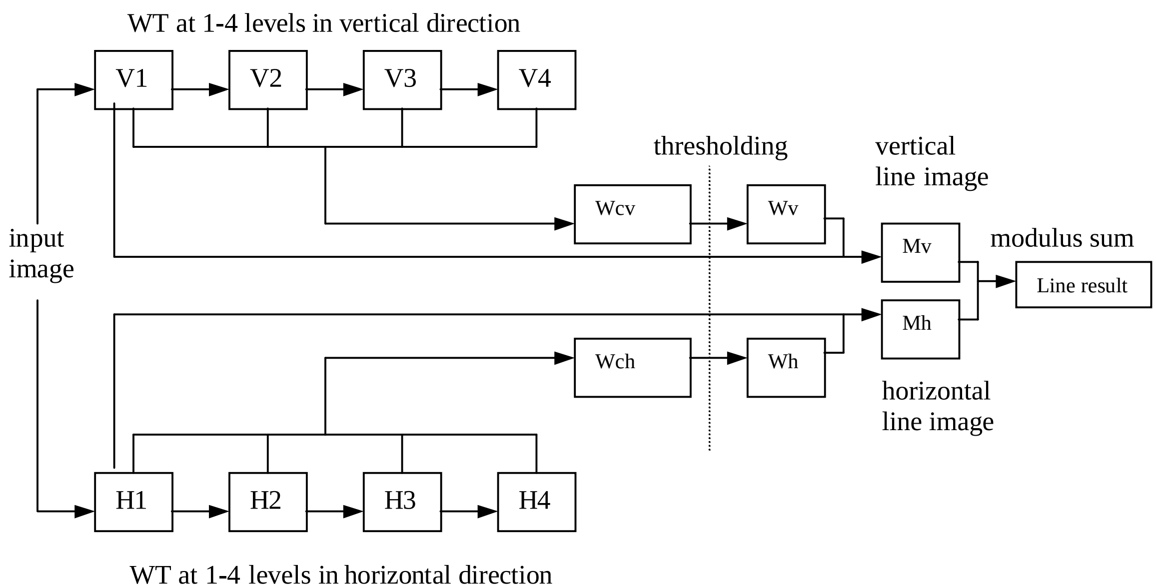
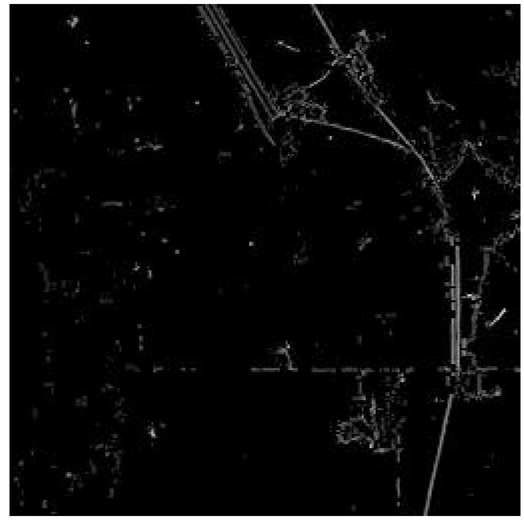


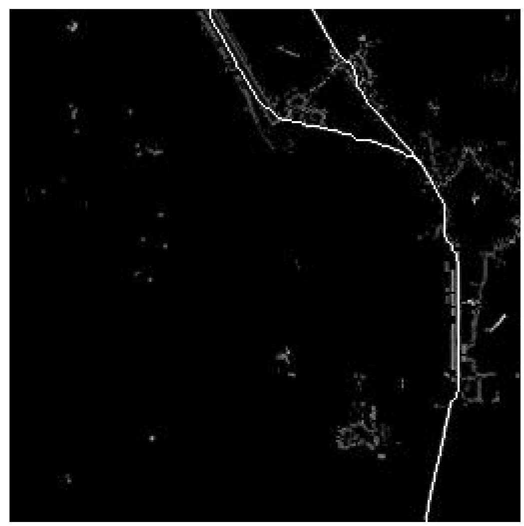
Figure 1 Block diagram of the image feature detection algorithm (a) wavelet transforms in horizontal and vertical directions for $j = 1, 2, 3, 4$ (b) results from combining different scale information (c) binary feature location (d) product of binary mask and wavelet coefficients at smallest scale (e) result, modulus of the sum of directional images.

(a)



(b)

(c)



(d)

Figure 2 Results of experiment (a) original image (b) image after road-pixel detection (c) region of road-pixel image set to infinite path (in white) (d) final result.