

FEATURE DETECTION IN IMAGES BY ADAPTIVE RANDOM SAMPLING

Ali Cafer Gurbuz, James H. McClellan and Waymond R. Scott, Jr.

Georgia Institute of Technology, Atlanta, GA USA

ABSTRACT

Random sample theory is an effective tool for detecting features in images. This paper presents an adaptive random sampling scheme that clusters random samples into candidate features. The required trial number is reduced by adaptive sampling, thereby reducing the run time of the algorithm. The proposed method quickly finds rough regions in the image that may include features using adaptive random sampling and re-estimates the features using the Hough Transform (HT) within the smaller regions. The proposed algorithm is tested on both simulated and experimental subsurface seismic and GPR images to search for linear features like pipes or tunnels. Faster results are obtained as compared to standard feature detection algorithms, such as the HT or its variants, while maintaining the similar performance level as the HT.

Index Terms— RANSAC, Hough Transform, Adaptive sampling, Fast line detection, Subsurface imaging,

1. INTRODUCTION

The Hough Transform (HT) [1], is the fundamental method used to detect parameterized shapes in images [2–5]. The HT uses a parameterized model of each feature to transform the feature in the original image space into a single mesh point in the parameter space. Although the HT is a robust algorithm for noisy images, it is not easily implementable because of its high computation time and large memory requirements.

Various methods have been proposed to decrease the computation requirements of the HT, the primary ones being the Probabilistic Hough Transform (PHT) [6], the Randomized Hough Transform (RHT) [7], and Line Detection using Random Sample Consensus (RANSAC) [8–10]. The RHT and PHT have less computation cost but decreased detection performance. For highly noisy images, PHT outperforms RHT [11], but neither algorithm works very well in detecting buried features in extremely noisy subsurface images.

Although RANSAC has been shown to perform line detection faster than the HT [9, 12] it can only be applied to binary images, not on gray-scale subsurface images. We have shown in [13] that using random sample theory, features can be detected in noisy gray-scale images more robustly and faster

than the HT without the degradation of detection performance suffered with the PHT and RHT algorithms; but information obtained from each random sample is not fully utilized. This paper uses the information obtained from each random sample to update the distribution from which the feature parameters are drawn to classify the features without *a priori* knowledge of the image statistics.

2. PROPOSED ALGORITHM

The basic idea of the algorithm is to first find rough areas or volumes in the image that may include features and then search only these rough regions with a more accurate algorithm like the HT. Reducing the search space of the HT decreases the computation time, while still maintaining detection performance at a level comparable to that of the HT. Rough regions are found using adaptive random sampling.

2.1. Algorithm Steps

A step by step description of the proposed two-stage algorithm follows:

Stage I (Candidate Model Selection)

(i) The feature to be detected is denoted by a set of parameters $P = (p_1, p_2, \dots, p_n)$, where each parameter p_i has limits defined by R_i . For example, a line in 2D is denoted by two parameters, (ρ, θ) , as $\rho = x \cos \theta + y \sin \theta$. Line parametrization in 3D and parameter ranges can be found in [14].

(ii) Generate a candidate feature by randomly sampling the current parameter distribution F . Thus, $(p_1, p_2, \dots, p_n) \sim F^k$ in the current iteration k . Here $P = (p_1, p_2, \dots, p_n)$ denotes the randomly selected feature and the parameters P make it possible to instantiate the feature in the image.

(iii) Sum the image within a σ distance of the selected feature. Call this candidate sum C_S . The summation regions in 2D and 3D images are illustrated in Fig.1. Call the total sum over the image T_S

(iv) Define a constant t , which depends on the summation result C_S from Step (iii), and compare it to the threshold t_0 defined in (1).

$$t = \frac{C_S}{T_S} \frac{1}{\Delta p_1 \dots \Delta p_n} \geq \frac{1}{R_1 \dots R_n} = t_0 \quad (1)$$

This work is supported under MURI by the U.S. Army Research Office under contract number DAAD19-02-1-0252.

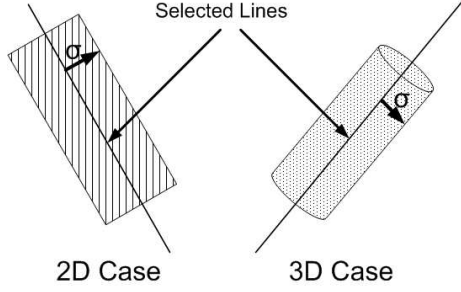


Fig. 1. Summing Regions in 2D and 3D images

Here Δp_i is the small parameter range around the selected parameter p_i that corresponds to the features summed within the σ distance of the selected feature in Step (iii). t_0 denotes the probability of a feature with a uniform distribution over the parameter space (R_1, \dots, R_n) . If $t > t_0$, the selected candidate feature is more likely than the default. Take the following steps depending on the cases:

Case 1 : If $t < t_0$ go to Step (v) and update the parameter distribution F . Note that the parameter distribution is only updated where the selected feature has lower probability than the threshold t_0 . This will reduce the probability of selecting another random sample close to this low probability feature region. Another reason why only low probability features are used to update the parameter distribution is to avoid missing weak features in the presence of a dominant feature. If the distribution were increased for one strong feature subsequent random samples would most probably be close to that feature while giving less chance to select other possible features. By only updating the parameter distribution for low probability features we avoid reselecting low probability regions and also give an equal chance for all possible features.

Case 2: If $t > t_0$ go to Step (vi) and classify the selected feature. In this case, the randomly selected feature has the potential to be a candidate feature. Use this feature in a clustering algorithm to classify it into possible feature classes.

(v) Update the parameter distribution as follows:

$$G = \exp\left\{-\frac{(p_1 - p_1)^2}{2\sigma_1^2} - \dots - \frac{(p_n - p_n)^2}{2\sigma_n^2}\right\} \quad (2)$$

$$F^{k+1} = (1 - G)F^k + Gt$$

$$F^{k+1} = \frac{F^{k+1}}{\sum F^{k+1}}$$

Here F^k is the parameter distribution at the k^{th} iteration. The variances used in G are given by $\sigma_i = c\Delta p_i$ and selection of Δp_i is given in Section 2.2. After updating the parameter distribution, increment k and go to Step (ii).

The initial distribution is selected as uniform over the parameter range if there is no *a priori* knowledge of the possible features.

(vi) Classify the selected candidate feature. Each class has parameter values μ_i^l and a class sum S^l for each parameter $i = 1, 2, \dots, n$ and class l . The clustering algorithm is implemented as follows:

Assign the first feature as one class with weight 1

Check if the selected feature falls in any of the current classes. It falls into a class if the mean of all selected parameters falls into the $(\mu_i - \Delta p_i, \mu_i + \Delta p_i)$ region of any current class. If it falls into a class and the selected feature has a higher summation value than the class, update the class parameters with the selected feature parameters, otherwise do nothing.

If the selected feature doesn't fall into any of the current classes a new class with $\mu_i^l = p_i$ and $S^l = C_S$ is formed.

Increase the iteration index k by 1 and go to Step (ii).

(vii) Stop after N trials

Stage II (Refine The Estimate) For each created class, define the parameter search space as $S_i = (\mu_i - \Delta p_i, \mu_i + \Delta p_i)$ and apply the HT only over S_i to find a better estimate of the features in that region. Note that this stage also enables multiple features within a region to be resolved and successfully detected. For example, if two lines are close to each other Stage I keeps only one of the features. To the extent that the HT can resolve multiple features in that region, all can be detected.

2.2. Selection of Algorithm Parameters

Selection of the parameter σ is important because it determines the number of trials, N , and the search space, ΔP , in Stage II. The relative size of σ with respect to the size of the image, s_i , is a crucial parameter. Selecting a very small σ/s_i ratio will increase N and reduce ΔP . In the limit $\sigma \rightarrow 0$, the algorithm approaches the RHT. Conversely, when the ratio σ/s_i increases, ΔP will increase, while the value of N will decrease. In the limit $\sigma \rightarrow s_i$, the number of trials in Stage I will be 1, making the candidate model selection stage useless and reducing the proposed method to the HT applied in Stage II to the whole image.

Once σ/s_i is selected, ΔP can be computed from σ/s_i as follows.

2D Line Detection	3D Line Detection	
$\Delta\theta = \arctan \frac{2\sigma}{s_i}$	$\Delta\theta = \arctan \frac{2\sigma}{s_i}$	(3)
$\Delta\rho = \sigma$	$\Delta\phi = \arctan \frac{2\sigma}{s_i}$	
	$\Delta u = \sigma$	
	$\Delta v = \sigma$	

Let N be the number of trials required to have at least one feature within ΔP vicinity of the true feature parameter with probability q . If the selected parameters fall beyond ΔP , the feature will not be able to be correctly detected since Stage II only searches for features within ΔP . If a non-adaptive

scheme is used and parameters are randomly chosen from a uniform distribution the number of trials for picking a random feature within ΔP with probability q will be

$$N = \frac{\log(1-q)}{\log\left(\frac{\prod_{i=1}^n R(p_i) - 2^n \prod_{i=1}^n \Delta p_i}{\prod_{i=1}^n R(p_i)}\right)} \quad (4)$$

In the adaptive selection of parameters, even though the distribution is reduced around the selected feature parameters for a low summation result, let's assume that part of the distribution is zeroed out. This will give a lower bound on the required trial number N and (4) gives an upper bound. The probability of getting at least one random selection in N trials within ΔP vicinity of the true feature parameter with probability q is equal to not getting any random selection in ΔP vicinity of the true feature parameter with probability $1-q$ in $N-1$ trials. Since an adaptive scheme is used and at each unsuccessful trial part of the distribution of size $\Delta p_1 \times \Delta p_2 \dots \Delta p_n$ is taken out, the probability $1-q$ can be given as

$$1-q = \prod_{j=1}^N \frac{\prod_{i=1}^n R_i - j \prod_{i=1}^n \Delta p_i}{\prod_{i=1}^n R_i} \quad (5)$$

To compare the minimum trial numbers for success probability q for adaptive and nonadaptive random schemes we take a 2D image of size 100×100 . For a summation width of $\sigma = 10$, and $q = 0.99$ using (3),(4) and (5), $N = 527$ trials are needed for nonadaptive sampling while adaptive sampling requires a minimum of $N = 64$ trials.

3. RESULTS

To illustrate how the proposed algorithm works, a 101×101 image with two linear structures having the parameter values $\rho_1 = 40, \theta_1 = 30^\circ$ and $\rho_2 = -20, \theta_2 = 150^\circ$ is created. The

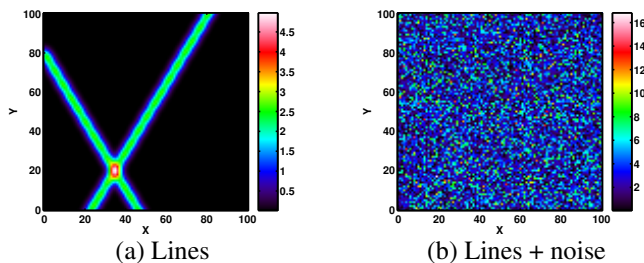


Fig. 2. (a) Two Linear features, (b) Two lines + 0-dB noise.

image is formed as sum of Gaussian distributions expressed by $f(i, j) = \sum_t G(i, j | x_t, y_t, \sigma_{x,y})$, where x_t, y_t are the true object positions in the image and $\sigma_{x,y}$ are the standard deviations on x and y axes. Zero-mean white Gaussian random

Table 1. True and Detected Target Parameters

2D Line Detection Results				
Targets	Target 1		Target 2	
Parameters	ρ	$\theta(^{\circ})$	ρ	$\theta(^{\circ})$
True Parameters	40	30	-20	150
Random Selection	34.61	28.04	-28.69	161.11
Refined by HT	39.35	30.70	-20.79	150.97

noise (WGN) is added to the image. The signal-to-noise ratio (SNR) is defined as the ratio of the maximum absolute value of the object signal to the average noise power. Figure 2 shows the synthetic linear structures with and without 0-dB noise. Note that the linear features cannot be visually seen and are masked by the noise.

The parameter σ is chosen as 10 with $N = 100$. The candidate feature parameters obtained from the Stage I of the algorithm are re-estimated by doing a HT within a vicinity $(\Delta\rho, \Delta\theta)$ of the parameters of the selected lines. The true parameters with the detected ones from stage I and stage II of the algorithm are listed in Table 1. The randomly selected and re-estimated lines are shown in Fig. 3.

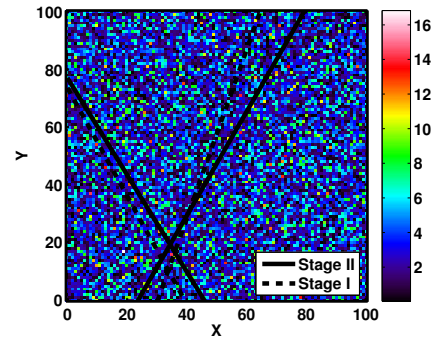


Fig. 3. Detected lines from Stage I (random selection) and Stage II (refined by the HT) of the proposed algorithm.

3.1. Adaptive vs Nonadaptive Random Sampling

The adaptive random sampling algorithm is compared to a nonadaptive version presented in [13]. Performance of the algorithms in varying SNR and average running times are compared. The proposed algorithm with $N = 100$ and $N = 200$ random samples is run 100 times on a 101×101 size image having one linear feature with random generations of white Gaussian noise added to the original signal at each SNR value. Nonadaptive feature detection algorithm uses 527 random samples. For each run, the detected line parameters with the run times are noted. The probability of detection (P_D) of

Table 2. Average Run Times of the Adaptive and Nonadaptive Algorithms

Adap($N = 100$)	Adap($N = 200$)	NonAdap($N = 527$)
0.72 s	1.44 s	1.95 s

each algorithm for varying SNR is shown in Fig.4. The detected line parameters are counted as a true detection if they fall within a $\pm\Delta p_i$ range of the true line parameters. Figure 4

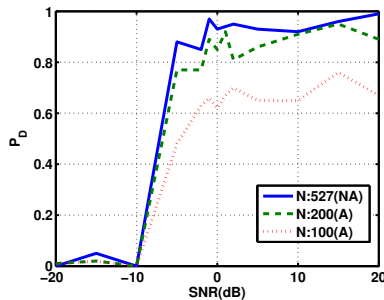


Fig. 4. Performance of the adaptive and nonadaptive algorithms in different SNR images

shows that while the nonadaptive algorithm and the adaptive algorithm with $N = 200$ have similar performances for all SNR values, the adaptive algorithm with $N = 100$ has worse performance. When the average run times of the algorithms (Table 2) are compared, the adaptive algorithm is the fastest.

3.2. Performance of the proposed method for varying SNR

This part compares the proposed algorithm (PA) and the nonadaptive feature detection algorithm (NA) from [13] with the Hough Transform (HT) and its variants RHT and PHT. The performance of the algorithms in varying SNR and the average times are compared. Two versions of the RHT and PHT algorithms with low (RHTL, PHTL) and high (RHTH, PHTH) trial number and data percentage are used for comparison. The PHT uses 50% and 5% of the data for PHTH and PHTL results; RHT uses 10^5 and 10^4 random point selections for RHTH and RHTL, respectively. In both RHT and PHT, the ratio of high and low trial numbers is kept the same at 10.

The probability of detection (P_D) for the algorithms is given in Fig. 5. Figure 5 shows that PA, NA and the HT have similar performance values, while PHTL and RHTL algorithms have much lower P_D . PA and NA algorithms are also more successful than RHT and PHT in detecting low SNR features. Here the advantage of the proposed method lies in its lower average run time. The average run times of the algorithms are given in Table 3 where it can be seen that random sample theory algorithms PA and NA have better performances and lower run times compared to PHT and RHT

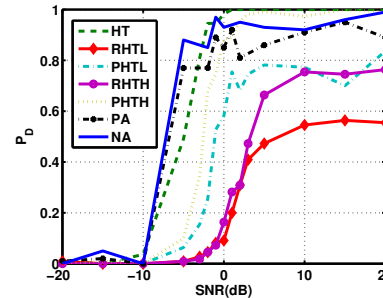


Fig. 5. Probability of detection for different algorithms with changing SNR

Table 3. Average Run Times of the Algorithms(secs)

PA	NA	HT	RHTH	RHTL	PHTH	PHTL
1.44	1.95	19.55	3.39	0.34	9.40	0.91

algorithms. Even though RHTL and PHTL have lower run times they also have worse performance relative to the PA. Random sample theory algorithms combines the fast running time of a random selection method with the best possible detection performance among the compared algorithms.

3.3. Experimental Data Results

An experimental system to collect seismic data was built to investigate the problem of detecting shallow tunnels [15–17]. In the experiments, a 10 cm diameter pipe which is a scale model of a tunnel was buried in the sandbox, which has a $1.8\text{ m} \times 1.8\text{ m}$ scanning area. The collected data from a seismic sensor was backprojected by a migration algorithm to form the subsurface images [15, 18].

The proposed algorithm and the HT have been applied to the surface energy image shown in Fig. 6. The two linear structures correspond to the buried pipe and the sloped bottom part of the sandbox tank. Detected features from Stage I and II of the proposed algorithm are shown in Fig. 6.

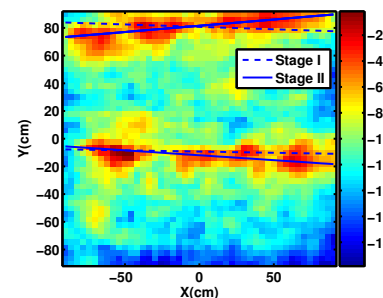


Fig. 6. Seismic surface energy image with the detected lines.

The detected parameters of the features in both the pro-

Table 4. Experimental Results

	PA		HT
	Stage I	Stage II	
ρ_1	80.72	81.24	78
$\theta_1, (^{\circ})$	1.53	1.66	1.64
ρ_2	-9.28	-11.91	-10
$\theta_2, (^{\circ})$	1.55	1.50	1.53
Time (s)	1.57		6.41

posed method and the HT along with the run times of the algorithms are summarized in Table 4.

The results show that the linear structures are detected successfully in a much shorter time than the Hough Transform.

4. CONCLUSIONS

An adaptive random sampling algorithm is introduced for detecting parameterized features. The random parameter samples are drawn from an updated distribution resulting in fewer required samples to detect a feature. Results from simulated and experimental data sets show that the proposed method has similar performance as the HT, and much lower computing time. The proposed algorithm also outperforms faster HT variant algorithms such as RHT and PHT. The proposed algorithm is well suited for applications such as detecting pipes, tunnels or other features in subsurface images from seismic or GPR sensors.

5. REFERENCES

- [1] P. V. C. Hough, "A Method and Means for Recognizing Complex Patterns," in *US Patent 3069654*, Dec. 1962.
- [2] G. Beylkin, "Discrete Radon Transform," *IEEE Trans. Acoustics, Speech, and Signal Processing*, vol. 35, pp. 162–172, 1987.
- [3] R. O. Duda and P. E. Hart, "Use of Hough transformation to detect lines and curves in pictures," in *Comm. ACM*, vol. 15, 1972, pp. 11–15.
- [4] D. H. Ballard, "Generalizing the Hough transform to detect arbitrary shapes," *Pattern Recognition*, vol. 13, pp. 111–122, 1981.
- [5] J. Illingworth and J. Kittler, "A survey of the Hough Transform," *Computer Vision, Graphics and Image Processing*, vol. 44, pp. 87–116, 1988.
- [6] N. Kiryati, Y. Eldar, and A. M. Bruckstein, "A probabilistic Hough transform," *Pattern Recognition*, vol. 24, pp. 303–316, 1991.
- [7] L. Xu and E. Oja, "Randomized Hough Transform (RHT): Basic mechanisms, algorithms, and computational complexities," *CVGIP: Image Understanding*, vol. 57, pp. 131–154, 1993.
- [8] M. A. Fischler and R. C. Bolles, "Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography," *Comm. ACM*, vol. 24, pp. 381–395, 1981.
- [9] R. Liu, Z. Ruan, and S. Wei, "Line detection algorithm based on random sample theory," in *SPIE Second International Conf. on Image and Graphics*, vol. 4875, 2002.
- [10] T. Chen and K. Chung, "A new randomized algorithm for detecting lines," *Real Time Imaging*, vol. 7, pp. 473–481, 2001.
- [11] N. Kiryati, H. Kalviainen, and S. Alaoutinen, "Randomized and probabilistic Hough transform: unified performance evaluation," *Pattern Recognition Letters*, vol. 21, pp. 1157–1164, 2000.
- [12] D. Chai and Q. Peng, "Image feature detection as robust model fitting," in *ACCV 7th Asian Conference on Computer Vision*, vol. 3852, 2006, pp. 673–682.
- [13] A. C. Gurbuz, J. H. McClellan, and W. R. Scott, Jr., "Feature detection in highly noisy images by random sample theory," in *15th International Conference on Digital Signal Processing*, July 2007.
- [14] P. Toft, "The Radon transform theory and implementation," Ph.D. dissertation, Technical University of Denmark, Lyngby, Denmark, 1996.
- [15] W. R. Scott, Jr., T. Counts, G. D. Larson, A. C. Gurbuz, and J. H. McClellan, "Combined ground penetrating radar and seismic system for detecting tunnels," in *IEEE Geoscience and Remote Sensing Symposium*, July 2006, pp. 1232–1235.
- [16] K. Kim and W. R. Scott, Jr., "Design of a resistively-loaded vee dipole for ultra-wideband ground-penetrating radar applications," *IEEE Trans. Antennas and Propagation*, vol. 53, 2005.
- [17] W. R. Scott, Jr., K. Kim, G. D. Larson, A. C. Gurbuz, and J. H. McClellan, "Combined seismic, radar, and induction sensor for landmine detection," in *Geoscience and Remote Sensing Symposium*, vol. 3, 2004, pp. 1613–1616.
- [18] A. C. Gurbuz, J. H. McClellan, W. R. Scott, Jr., and G. D. Larson, "Seismic Tunnel Imaging and Detection," in *ICIP 2006*, 8-11 Oct. 2006, pp. 3229–3232.