

# Histograms of Oriented Gradients for Landmine Detection in Ground-Penetrating Radar Data

Peter A. Torrione, Kenneth D. Morton, Jr., *Member, IEEE*, Rayn Sakaguchi, *Student Member, IEEE*, and Leslie M. Collins, *Senior Member, IEEE*

**Abstract**—Ground-penetrating radar (GPR) is a powerful and rapidly maturing technology for subsurface threat identification. However, sophisticated processing of GPR data is necessary to reduce false alarms due to naturally occurring subsurface clutter and soil distortions. Most currently fielded GPR-based landmine detection algorithms utilize feature extraction and statistical learning to develop robust classifiers capable of discriminating buried threats from inert subsurface structures. Analysis of these techniques indicates strong underlying similarities between efficient landmine detection algorithms and modern techniques for feature extraction in the computer vision literature. This paper explores the relationship between and application of one modern computer vision feature extraction technique, namely histogram of oriented gradients (HOG), to landmine detection in GPR data. The results presented indicate that HOG features provide a robust tool for target identification for both classification and prescreening and suggest that other techniques from computer vision might also be successfully applied to target detection in GPR data.

**Index Terms**—Computer vision, edge histogram descriptors, ground-penetrating radar (GPR), histogram of oriented gradients (HOG), random forest.

## I. INTRODUCTION

THE detection of buried explosives is of significant concern in both military and humanitarian applications. In the military sphere, recent events have illustrated the destructive efficacy of buried improvised and classical explosives (e.g., landmines), and both the humanitarian and military scope of these threats continues to grow [1]. Unlike classic landmines, which are produced en masse and usually contain at least small amounts of metal, modern homemade explosives are typically constructed in an *ad hoc* manner, and may contain no metal content whatsoever.

Electromagnetic induction (EMI) based sensing of buried threats is an efficient technique for subsurface threat detection, and improved modeling and related signal processing have illustrated the potential for EMI to detect very small amounts of metal [2]–[5]. However, the increase in the use of purely nonmetallic threats and the prevalence of metallic anthropic

clutter in conflict and post-conflict regions renders detection based solely on EMI sensing difficult.

Ground-penetrating radar (GPR) [6] is a highly complementary phenomenology to EMI modeling. Unlike EMI, which detects eddy currents induced in buried conductors, GPR measures reflections from changes in the electromagnetic properties of the subsurface environment. As a result, GPR sensing is suitable for finding nonmetallic threats as long as the electromagnetic properties of the buried object are sufficiently different from the surrounding soil. However, the sensitivity of GPR to changing subsurface properties results in additional responses from buried roots, rocks, pockets of moisture, and other subsurface changes. As a result, without proper additional processing, GPR detection of buried threats can be subject to high false alarm rates. A significant body of research has previously been undertaken to successfully detect buried threats in GPR data while reducing false alarm rates. This previous research can roughly be divided into three broad categories: model inversion, explicit hyperbola detection, and statistical feature-based techniques.

Physics- and model-inversion-based techniques for GPR processing are very popular in the research community due to strong underlying theoretical foundations (e.g., Maxwell's equations, Green's functions) and the wide variety of techniques available to solve the resulting propagation equations. Recent developments in the field of physics-based inversion include novel techniques for subsurface object shape and size estimation, even under rough surfaces using iterative optimization methods [7]. Significant work has also been investigated to automatically determine the properties of the subsurface, including layer identification and electromagnetic propagation properties [8]–[10]. Inversion techniques and their specific application in demining techniques have also been explored in [11] and [12]. Despite these advances, physics-based inversions of GPR signals are not widely utilized in fielded or real-time data collection systems for several reasons. First, many physics-based inversion models rely on accurate estimates (or computationally expensive iterative estimation) of the subsurface electromagnetic properties of the soil. These properties are difficult to estimate, and even when accurate large-scale estimates are available, the properties of most soils of interest can change drastically over small regions, making large-scale estimates irrelevant. Furthermore, despite recent advances, the computational expense associated with physics-based GPR data inversion is significant, and precludes most real-time operation.

Manuscript received October 2, 2012; revised February 6, 2013; accepted March 4, 2013. Date of publication May 21, 2013; date of current version December 17, 2013. This work was supported by the U.S. Army RDECOM CERDEC Night Vision and Electronic Sensors Directorate, via a Grant Administered by the Army Research Office under Grant W911NF-09-1-0487 and Grant W911NF-06-1-0357.

The authors are with the Department of Electrical and Computer Engineering, Duke University, Durham, NC 27708 USA (e-mail: pt@ee.duke.edu; kdm@ee.duke.edu; rayn.sakaguchi@duke.edu; lcollins@duke.edu).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TGRS.2013.2252016

Hyperbola-based landmine detection methods are another popular technique for object identification in GPR data. Hyperbola-based techniques can be motivated by considering a simplification of the general GPR propagation problem wherein objects of interest are modeled as point-scatterers, and the transmission medium is assumed to be homogeneous. Under these simplifying assumptions, the time of arrival for responses from buried objects constitute a hyperbola as the sensor moves across the object, and the parameters of the resulting hyperbola contain information about relevant object depth and soil properties. Typical approaches to hyperbola identification in GPR make use of modifications to the Hough transform [13] or alternative fitting techniques [14], to avoid computational complexity issues, account for uncertain numbers of reflectors, and mitigate variations due to unknown parameters. Additional modifications to account for refraction can also be considered [15]. However hyperbola-based techniques also face significant shortcomings for fielded operations. First, as discussed above, assumptions of homogeneous transmission media are routinely violated in soils of any complexity, which significantly obscures the theoretical hyperbolic responses. Also, at the sampling rates of interest, most buried threats are not well represented by point sources, and responses from nonmetal objects contain constructive and destructive interference which also obscure idealized hyperbolic responses.

Since responses from buried objects are highly variable, most effective fielded algorithms for landmine identification in GPR data use statistical feature extraction (often motivated implicitly or explicitly from the roughly hyperbolic nature of target responses) combined with machine learning techniques to develop algorithms to classify new responses as threats or nonthreats [16]–[22]. In our opinion, these techniques should not be seen as a rejection of the physics-based or hyperbolic modeling of target responses, but should be considered a natural extension of these techniques that take into account the noise and uncertainty inherent in realistic operational conditions. In actuality, many of the most successful techniques for fielded target identification in GPR data are based on exploiting spatial variations in target responses that are best understood by recourse to the underlying physics. For example, one of the first large-scale explorations of statistical learning for target detection in GPR data resulted in the application of hidden Markov models (HMMs) to detect and classify landmine responses based on a statistical representation of their characteristic hyperbolic shape [21]. Another equally efficient technique exploits the same rising/flat/falling structure as the HMM, but uses windows of fixed length for feature extraction [22]. These techniques have been shown to have very robust performance over very large data collections containing difficult GPR target populations [23].

The success of techniques such as in [21] and [22] illustrate the potential for statistical modeling and machine learning to provide robust results for object detection in GPR data. In a parallel vein, modern advances in the field of computer vision also make use of statistical object descriptors coupled with machine learning to develop accurate algorithms for both instance and category recognition (see [24] for recent summary and extensive bibliography). Notable examples include: 1) the

scale invariant feature transform (SIFT) and a related method; 2) speeded up robust features (SURF) [25], [26] which provide a low-dimensional representation of visual images for instance matching with images; 3) Viola–Jones’ cascade of simple features used for robust face detection [27]; and 4) histogram of oriented gradients (HOG) which has been successfully used for pedestrian detection [28].

In the context of subsurface threat detection, HOG features are particularly interesting since analysis of [22] and [28] illustrate that effective feature extraction techniques from two disparate image processing fields (MPEG encoding and pedestrian detection in images) share very similar underlying mathematical and theoretical structure. These features are also widely utilized in the field of computer vision and video processing, so are necessarily computationally inexpensive and efficient. This observation naturally leads to the question: can HOG and other tools from the computer vision literature be successfully brought to bear on the problem of landmine detection in GPR data?

The size and scope of modern computer vision techniques is enormous, so this paper focuses only on the specific case of HOG features. The rationale for an initial focus on HOG features is motivated by the similarity of HOG to preexisting feature extraction techniques [e.g., edge histogram descriptor (EHD)], and also to several aspects of HOG features which make them particularly well suited to landmine detection in GPR. For example, unlike SIFT, SURF, and others, HOG features are typically applied as a dense feature extraction technique. This enables HOG features to be extracted without previous interest point identification, which is an unsolved though interesting problem in the GPR literature [29]. Furthermore, unlike many modern feature descriptors, HOG features are fairly robust to moderate changes in object location within a HOG window. This enables the use of preexisting techniques for depth-binning in GPR classification. This paper shows how HOG features, coupled with off-the-shelf classification techniques, can enable significant improvements in target classification. Although this paper focuses on HOG features, we believe that, with additional research, other techniques from computer vision literature may have an important role to play in GPR-based subsurface threat detection.

Historically, for vehicle-mounted GPR processing to maintain speeds commensurate with operationally relevant rates of advance, computationally simple algorithms were utilized to rapidly process large volumes of GPR data and down-select locations of interest, alarms, which were then utilized in more advanced feature-based processing. This separation between prescreening and classification has always seemed artificial, as it could be advantageous to incorporate advanced statistical models into prescreening if feature extraction and classification could both be accomplished in a computationally tractable manner. This paper is the first to develop a technique for threat identification in vehicle-mounted GPR data that is fast enough and robust enough to operate as a prescreener (on every pixel) versus only at previously flagged locations of interest. The primary novel advances contained in this paper include the first application of techniques explicitly developed for feature extraction and object classification techniques from the

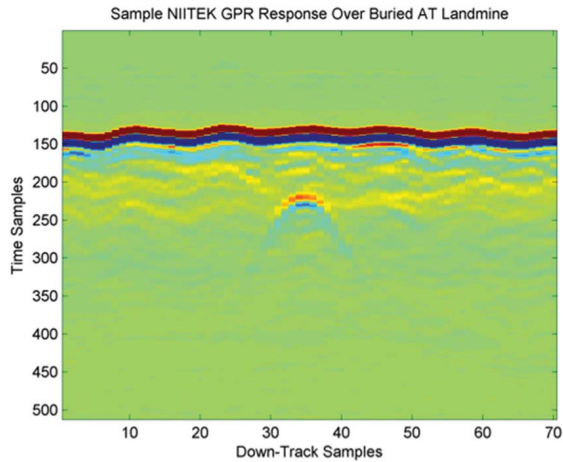


Fig. 1. Sample GPR data from a western test site, collected over a buried plastic antitank landmine at 3'' overburden. The dark wavy band near timesample 150 corresponds to the response at the air/ground interface. The target is visible near down-track sample 35, time sample 225.

computer vision literature to target classification in GPR data, the development of a series of preprocessing steps suitable for simple target extraction and classifier training across multiple feature extraction approaches (e.g., HOG, and EHD), and the development of a novel joint prescreening/classification algorithm for rapid pre-screening.

The remainder of this paper is organized as follows. Section II describes the ground penetrating radar used in this paper, and the GPR data under consideration. Section III describes the HOG features used in this paper. Section IV describes the application of HOG features to GPR data and how these features can be used for object classification as well as during pre-screening. Section V presents results from the HOG prescreener as well as HOG-based classification using cross validation on a large number of prescreener flagged objects of interest, and extends the preprocessing approaches described in Section IV for application to EHD features. Section VI presents conclusions, and describes several possible avenues for future work.

## II. GPR SYSTEM AND DATA UNDER CONSIDERATION

All of the data considered in this work were collected with a pulsed, time-domain, bi-static antenna, with a very wide bandwidth, and very low radar self-signature. A-scan's from each antenna are time-gated to sense approximately 1 m in air, using 512 temporal samples.

This work focusses on a vehicle-mounted array with 51 antenna pairs spaced approximately 5 cm apart. As the vehicle on which the antenna array is mounted moves (down-track) each of the 51 antenna pairs is sampled (approximately simultaneously) once every 5 cm of vehicle motion. This collection procedure creates a 3-D data set of size  $512 \times 51 \times n_{dt}$  samples where  $n_{dt}$  represents the number of down-track scans collected.

Fig. 1 illustrates sample data collected with the GPR at a western U.S. test site. The dark red band at time sample 150 represents the response from the air/ground interface.

Time samples before the interface correspond to the GPR signal propagating through air, and time samples after the interface represent propagation and reflection in the soil. A response from an antitank landmine buried at 3'' can be seen between time samples 200 and 250 and down-track samples 30–40. The characteristic hyperbolic response of the object is visible as the antenna moves across it. The goal of this research is to successfully identify responses such as the one in Fig. 1 and discriminate these from other similar responses from subsurface clutter such as rocks, roots, and other possible confounding objects.

For the purposes of applying computer vision approaches to GPR data, it will be convenient to treat a GPR image like the one in Fig. 1 as a 2-D matrix of data. Throughout the following sections, let  $I$  represent an image (or 2-D block of GPR data, in which case  $I(i, j)$  may represent the received field, often in volts, or units proportional to volts) of size  $[n_i, n_j]$ , and let  $I(i, j)$  represent the image intensity value (or amplitude of the GPR signal) at location  $i, j$ . For time-domain GPR data, typically the first dimension of  $I$  corresponds to propagation time, or (roughly) depth into the transmission media, and the second dimension may represent distance traveled (down-track) or channel (cross-track), e.g.  $[n_i, n_j] = [n_t, n_{dt}]$ . In the following, when HOG features are extracted from a point near an image boundary, pixels outside the bounds of the image are generated via symmetric padding, e.g.,  $I(-5, 1) = I(5, 1)$ .

## III. COMPUTER-VISION-BASED FEATURE EXTRACTION TECHNIQUES

A large number of very different techniques have previously been proposed for extracting low-dimensional representations of images and image patches (see [24] for detailed examples and bibliography). Many of these techniques were designed for the problem of instance recognition, or identifying the exact same object in two different images. For example, techniques such as SIFT [25] and SURF [26] were developed to enable rapid automatic matching between image patches which have undergone transformations between images, including rotations and scaling. Typically, the instance recognition problem is approached by extracting descriptors from a sparse set of interest points (e.g., [24], [25], [30]) and then matching descriptors across different images. In instance recognition, interest point extraction can significantly simplify processing by ignoring regions which are not identifiable, flagging regions which are interesting, and by reducing the overall computation time.

In contrast to instance recognition, the more challenging problem of object class recognition requires classifying distinct objects within an image as either belonging to or not belonging to a certain class. Category recognition is more complicated than instance recognition since robust classification requires modeling the wide statistical variability of objects within a class (e.g., chairs) compared to the instance recognition problem which only requires identifying changes due to transformations of a particular chair. For example, a commonly studied and challenging class recognition problem in computer vision is the robust

identification of pedestrians from vehicle-mounted video data (see [31], [32] for a recent review). In contrast to instance recognition, most algorithms for class recognition utilize feature descriptors extracted densely throughout the image, since it is difficult to *a priori* discriminate between informative and noninformative regions for the class-recognition problem.

The landmine detection problem in GPR can be seen as an example of object class recognition, where the system is presented with a series of responses from the subsurface and must identify which GPR images contain landmine responses. With this insight, a wide number of possible algorithms from the computer vision literature could be brought to bear on the landmine detection problem in GPR. To bound the problem and our current research, this paper focuses on the application of HOG features to target classification in GPR data, though as we specify in the conclusions, the computer vision literature provides ample opportunities for alternative feature extraction and object classification approaches.

This section outlines the basic implementation of HOG features, describes their application to GPR data, and discusses classification algorithms suitable for processing HOG features. Section IV presents details for applying these techniques to GPR data, and Section V presents results obtained from each of the techniques.

#### A. Histogram of Oriented Gradients

HOG features [28] provide a concise but powerful image representation for general object classification. They have found significant application in the field of pedestrian identification [28], where they have continued to provide one of the more robust feature extraction techniques even in recent analyses [32].

HOG features are based on gradient angle and magnitude distributions, and in visual data they are robust due to the gradient's natural invariance to slight changes in ambient lighting and color variations. Consider an image  $I$  and gradient estimation filters  $h_x = [-1, 0, 1]$ , and  $h_y = [-1, 0, 1]^T$ . Let  $g_x$  and  $g_y$  represent the gradient images generated by

$$g_x = I * h_x$$

and

$$g_y = I * h_y$$

where  $*$  represents convolution. The magnitude of the gradient at each pixel can be calculated as

$$G(i, j) = \sqrt{g_x(i, j)^2 + g_y(i, j)^2}$$

and the dominant gradient angle at each pixel can be estimated by

$$\theta(i, j) = \tan^{-1} \left( \frac{g_y(i, j)}{g_x(i, j)} \right). \quad (1)$$

For some applications, the sign of the gradients, and thus values of  $\theta$  between  $0^\circ$  and  $360^\circ$ , may be germane to the detection problem. For example, in detection problems where objects of interest can be assumed to always be darker than the background regions, gradients will always tend to flow toward background regions. In most applications, however, the sign

of the gradients  $g_x$  and  $g_y$  provides little or no discriminative power, so  $\theta(i, j)$  can be calculated in the range  $0$  and  $\pi$  radians (versus  $0$  to  $2\pi$ ).

HOG features can then be generated entirely on the basis of the gradient magnitude  $G$  and gradient angles  $\theta$  at each pixel. The primary insight provided by HOG feature extraction is that, while individual  $G(i, j)$  and  $\theta(i, j)$  are highly variable and subject to significant variations across nearby  $(i, j)$  locations, even for very similar images, the aggregate statistics of the spatial distribution of the gradient angles and magnitudes over small regions in similar images provide quite robust descriptors of those regions.

Consider  $n_\theta$  angle bins between  $0^\circ$  and  $180^\circ$  (or as discussed above, potentially, between  $0^\circ$  and  $360^\circ$ ). HOG descriptors encapsulate the local statistics of the gradient angles and magnitude by allowing each pixel to vote for a specific angle bin, with a vote magnitude proportional to the gradient magnitude at that pixel. Let the edges of the  $n_\theta$  bins correspond to  $n_\theta + 1$  edge values,  $\{\phi_k\} = \{180k/n_\theta\}$   $k = 0 \dots n_\theta$ , then a 3-D vote matrix of size  $n_x \times n_y \times n_\theta$  can be defined as

$$V(i, j, k) = G(i, j)\delta(\phi_{k-1} < \theta(i, j) \leq \phi_k), \quad k = 1 \dots n_\theta$$

where  $\delta(x)$  takes value 1 when the input argument is true, and 0 when the input argument is false.

Resulting individual pixel votes can be smoothed using bilinear interpolation and are then aggregated across subregions of the image referred to as cells, which can be rectangular or radial. Aggregating across a particular cell  $c$  is accomplished by

$$H_1(c, k) = \sum_{(i, j) \in c} V(i, j, k).$$

In many computer vision application areas, local changes in ambient lighting can have a significant effect on the magnitudes of feature responses, and this holds true for the gradient magnitudes used in HOG descriptors. Similarly, signal magnitudes in GPR data vary drastically as a function of propagation distance. As a result, normalizing HOG descriptors is fundamental to achieving robust performance. HOG descriptors can be normalized using groups of neighboring cells, also called blocks. Let  $N(c)$  represent the set of cells immediately surrounding the cell of interest; then the normalized HOG values can be calculated using

$$H(c, k) = H_1(c, k) / \left( \sum_{c_i \in N(c)} \sqrt{\|H_1(c_i)\|_2^2 + \epsilon^2} \right) \quad (2)$$

where  $H_1(c)$  represents the rasterized vector  $[H_1(c, 1), H_1(c, 2), \dots, H_1(c, n_\theta)]^T$ .

In [32], the authors utilize feature descriptors  $v$ , consisting of concatenations of  $H(c, k)$  over each  $N(c)$ , and normalize these vectors by their local cell regions. The resulting complete descriptor vector then consists of multiple repeated  $H(c, k)$ , each normalized with respect to a different set of cells. In this paper, for application to GPR processing that step was found to not significantly affect performance, so the complete feature vector is simply a rasterized vector of  $H(c, k)$ , as calculated

in (2). For more information about the HOG feature descriptors and a more detailed description of the HOG extraction process, the interested reader is referred to [28].

### B. Relationship to Edge Histogram Descriptors

EHD features [33] are a technique used in the MPEG codec to identify and contrast regions of varying texture, which were originally applied to landmine detection in GPR data in [22], and have continued to perform well over a wide range of operating conditions [23]. Like HOG features, EHD features are also based on describing an image through a histogram of gradients (edges). However there are several important distinctions to be made between the EHD and HOG techniques. First, EHD features rely on a global threshold on gradient magnitudes to determine whether a particular pixel's gradient has an opportunity to vote for its corresponding edge bin. These global thresholds can be difficult to set and tune, and require data normalization prior to feature extraction. Second, unlike the case with HOG features, where each pixel votes with strength proportional to the local (normalized) gradient, in EHD features the each pixel receives equal vote regardless of gradient magnitude. HOG features also introduce the notions of bins and cells (see above) which enable additional spatial information to be encoded in HOG features that is not present in EHD features (e.g., temporal as well as down-track information). Finally, the actual bins used in EHD are typically fixed to correspond to "diagonal," "anti-diagonal," "vertical," "horizontal," and "nonedge" orientations, whereas in HOG features the number and size of the individual bins is a parameter available for tuning. Extension of EHD features to arbitrary angle orientations requires the development of new edge extraction filters, which is not straightforward. This adds significant flexibility to HOG features for applications where the inter sample spacing in different orientations is not known *a priori*, and/or subtle shifts in gradient magnitude may be important for classification.

On a more practical level, the preprocessing steps developed in this paper for application with HOG features can significantly simplify algorithm development, training, and implementation for fielded application compared to current implementations of EHD features. For example, in [22] it was necessary to manually label the specific depth regions where target responses appeared in a very large body of GPR images to obtain prototype GPR segments for algorithm training. Manual labeling of GPR images is time consuming, error prone, and subject to variations between individual user preferences. In contrast, this paper shows how features can be reliably extracted from automatically flagged GPR depth segments (Section IV-B). This significantly simplifies algorithm training which often requires data from thousands landmine responses, and streamlines algorithm modifications as new data becomes available. Finally, [22] used a special-purpose classification algorithm based on a combination of self-organizing maps [34], and fuzzy K-NN classification. Results from this paper based on HOG features shows that both HOG and EHD features are more easily separated using standard off-the-shelf classification techniques such as random

forests [35], and can even be well separated with linear classifiers, such as partial-least-squares discriminant analysis (PLS-DA) [36]. Linear classification is particularly of interest, since extremely fast classification algorithms are required for techniques like HOG features to be applicable in GPR-based prescreening, and as the remainder of this paper illustrates, HOG features coupled with PLS-DA classifiers provide the first such example of a feature-based prescreening algorithm capable of operating in realtime on the large volumes of data generated by vehicle mounted GPR arrays.

## IV. APPLICATION OF HOG FEATURES TO GPR DATA

GPR-based object classification using HOG features requires incorporating additional information from our knowledge of GPR signals and propagation to understand how to best extract target and nontarget examples of subsurface responses, which subset of imagery at an alarm location to extract, proper data preprocessing, and combinations of down- and cross-track imagery.

### A. Spatial Locations for Feature Extraction

Developing a classification algorithm based on HOG features to detect possible threats in GPR data requires a large body of data from both target [e.g., landmines, other buried explosives] and nontarget responses. For the data considered here, it is relatively simple to identify target spatial locations since their GPS coordinates are known *a priori* and the GPR data is collected with accurate GPS locations.

Determining which nontarget responses to use during training offers more options. One approach to generate nontarget responses might be to randomly sample the remaining nontarget portions of the large volume of GPR data available. One could then undertake a bootstrapping approach to identify interesting regions of the subsurface for training (see [37] for an example of this process applied to video data). However, the vast majority of GPR data collected over testing lanes is relatively homogeneous, and the majority of randomly sampled points are therefore uninteresting from a machine learning perspective.

As discussed previously, prescreening algorithms are computationally inexpensive algorithms designed to achieve high probability of detection at moderate to high false alarm rates. False alarms from prescreening algorithms are typically due to surface or subsurface nontarget anomalies, rapid changes in sensor height, radar-ground coupling, or changes in the subsurface structure. As a result, prescreener false alarms provide a rich source of anomalous data likely to confuse classification algorithms, and these provide a good source of training data for machine learning algorithms. The remainder of this paper extracts nontarget data for training using spatial locations false alarms from a prescreener developed at and based on [38].

### B. Temporal Locations for Feature Extraction

Given a spatial location around a target or prescreener false alarm, one can easily extract a cube of data corresponding to



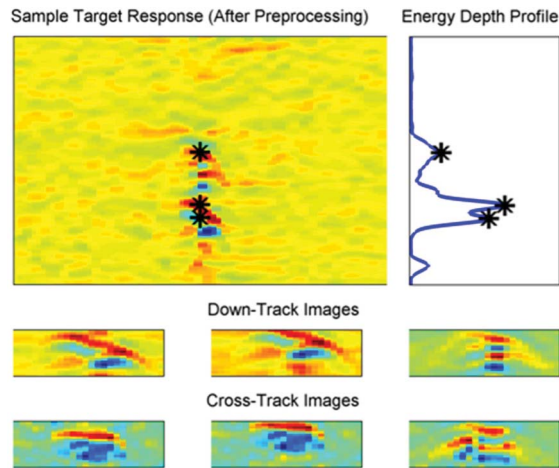


Fig. 2. Sample down-track GPR image and subimage extraction. The top left image shows a target response after time-gating to remove air/ground interface. The top right image shows a smoothed energy profile of the image on the left. Local maxima are marked with black stars. The bottom rows of images show subsets of data extracted around the three black star locations in the top left image in both the down- and cross-track orientations. In the bottom rows, the leftmost image corresponds to data extracted around the largest maxima peak in the top right, the middle image to the second highest peak, and the third image to the lowest magnitude peak. The six images in the bottom are enlarged to illustrate detail compared to the top left image.

that location. However, significant portions of a GPR response around a location of interest typically contain little to no information about the object (Fig. 1). As a result, it is typically necessary to extract regions of interest from within a typical B-scan (collection of A-scans) that are believed to contain information regarding the object under consideration. For the HOG features, additional preprocessing must be used to select different subwindows of the sensor data for further processing.

In the computer vision literature, the localization of an object or point of interest is often achieved using keypoint identification methods which seek to identify corner or blob-like regions [24]. However, based on some initial investigations, corners and blobs do not appear to be particularly well suited for interest point identification in GPR data [29]. Although GPR-specific keypoint identification is an interesting field for further study, it is beyond the scope of this paper. Instead, since subsurface objects typically result in larger magnitude GPR responses than the surrounding homogeneous soils (since responses from dielectric discontinuities cause reflections of significant magnitude), this paper uses a smoothed local signal energy metric to identify the three largest energy peaks in the central A-scan closest to an anomalous response as locations around which to extract HOG features. Each extracted window is of size  $54 \times 24$  image pixels, though this is later reduced to  $18 \times 24$  pixels (Section IV-C). The original window size of  $54 \times 24$  pixels was chosen based on visual analysis of multiple target responses. Responses from the targets of interest often span 50 or so samples temporally, and larger targets (e.g., buried shells) often have responses spanning 24 samples (about 1.2 m) spatially. Although for general application these parameters should be selected based on typical responses from the radar under consideration, investigations undertaken during algorithm development indicated

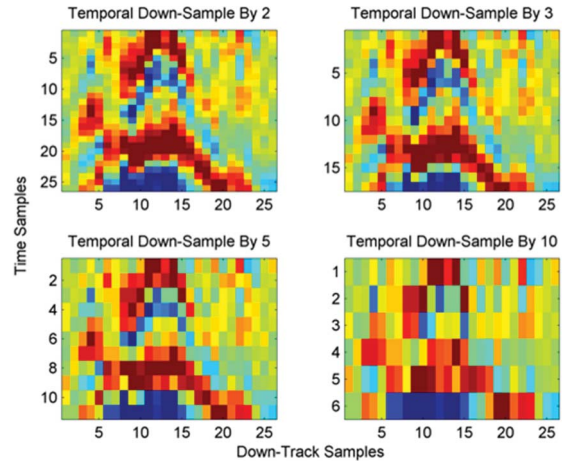


Fig. 3. Sample GPR images using different amounts of temporal down-sampling. The time axis is along the y-axis and spatial location is along the x-axis. Each image corresponds to a different amount of downsampling in the time axis. The goal is to make the distribution of features in the HOG histogram bins approximately uniform (see Fig. 4 for additional information and results).

only slight performance sensitivity to the precise size of the windows utilized. This process is shown visually in Fig. 2.

The process described above is used during training to determine temporal locations at which to extract data representative of target and nontarget responses. During testing, the HOG features are extracted from a fixed set of windows of size  $54 \times 24$ , which overlap 50% in the vertical along the A-scan corresponding to the prescreener alarm. The final decision at each spatial location is made by first processing each depth window along the A-scan of interest, and then taking the average of the top three confidences to create a final statistic. This process was developed based on empirical results, and agrees closely with other approaches used in the literature, e.g., [22]. As with many GPR-processing applications, both the training and testing procedures used here are beset by logistical problems necessitating *ad hoc* solutions, e.g., the selection of positive examples with smoothed energy, and aggregating over classifier confidences using order statistics. In actuality these problems can be seen as symptoms of the underlying multiple-instance nature of the target identification problem [39], [40]; see Section VI for further discussion.

### C. Image Resampling

Because of the nature of typical image acquisition in visual imagery it is common to assume *a priori* that pixel values are likely to be equally variable in any of the cardinal directions (e.g., in the classic four-element neighborhood). This assumption does not carry over to processing GPR data since GPRs often use radically different spatial and temporal sampling rates, so variation along the y-axis (time) and the x-axis (space) cannot be assumed to be commensurate. For example, in the top left image in Fig. 3, the temporal sampling rate (y-axis) is much higher than the spatial sampling rate (x-axis), so derivatives calculated along the y-axis will have much smaller magnitudes than those calculated along the

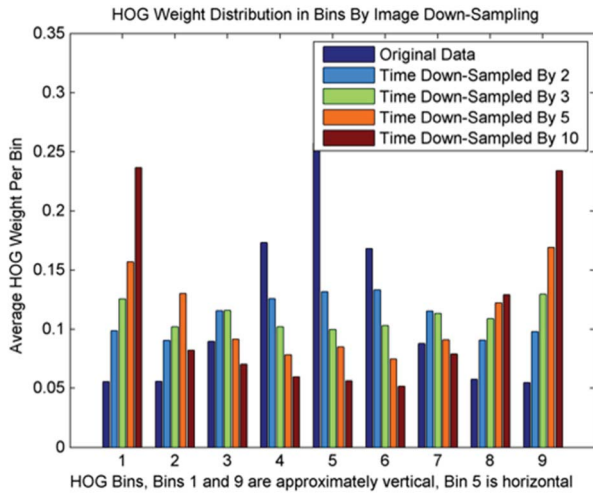


Fig. 4. Sample distribution of features in HOG histogram bins after different amounts of GPR temporal downsampling. For highly temporally sampled GPR data, the gradients are dominated by horizontal or spatial differences (Bin 5). For highly downsampled GPR data, the gradients are dominated by vertical gradients (Bins 1 and 9). For this particular radar, a downsampling rate of 3 achieved approximately uniform bin histograms.

$x$ -axis, where single pixel differences correspond to much larger changes in signal magnitude.

Since the sampling in GPR data is unequal, some spatial derivatives are artificially enhanced or reduced, certain angle estimates (1) will be overrepresented in the resulting HOG descriptors, resulting in reduced feature fidelity. It is possible to use the physics of electromagnetic propagation to analytically estimate the reduction in temporal sampling required to achieve equity between spatial and temporal sampling given known soil parameters. However, soil parameters are rarely known precisely, and can change rapidly over the course of a few meters in real data. Furthermore, simplistic analysis and downsampling ignores important factors for target identification such as the transmitted pulse width, which has significant impact since reflections of the transmitted pulse determine the data used to calculate gradients.

Instead of a physics-based approach, this paper utilizes a statistically motivated approach to determine the proper image downsampling. To achieve this we exploit the fact that, for optimal performance, the HOG feature descriptors should be approximately uniformly distributed across the histogram bins. For temporal sampling rates that are too high, the HOG feature vectors will tend toward descriptors with a great deal of weight in the horizontal bins, since derivatives along the time axis will be artificially smaller in magnitude. Similarly, for sampling rates that are too low, the HOG feature vectors will tend toward descriptors with a great deal of weight in the vertical bins, since the gradient estimates along the A-scan will be artificially large. Fig. 4 shows an example of this for HOG descriptors calculated from images like those shown in Fig. 3. Fig. 4 illustrates how changing the temporal sampling rate in a GPR image alters the distribution of gradient magnitude in the resulting HOG descriptors. Based on these results, the remainder of this paper uses a temporal downsampling of three time samples, since that amount of

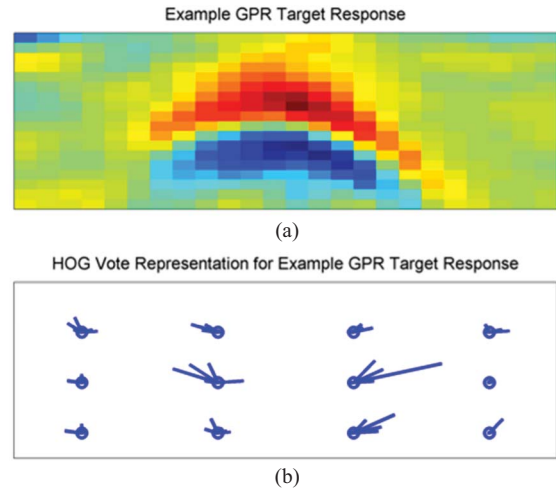


Fig. 5. (a) Sample GPR subimage and (b) resulting HOG descriptor representation extracted around a target response. The HOG representation is generated by taking the resulting gradient bins and gradient magnitudes within each cell and generating lines at the appropriate angle with the appropriate magnitude. This visualization was inspired by those in [28].

downsampling provides a roughly flat distribution of HOG features for the GPR under consideration.

#### D. Down-Track and Cross-Track Information

Although GPR data is easiest to visualize as a B-scan in two dimensions, vehicle mounted sensor arrays generate 3-D volumes of data as they move down-track. Although the HOG implementation described in this paper does not directly account for the 3-D nature of the data volume (but see Section IV), incorporating information from both the down-track and cross-track B-scans at an object of interest significantly improves performance compared to using either B-scan orientation alone. As a result, the process described in Sections IV-A to IV-C is repeated in both the down- and cross-track orientations, and the resulting feature descriptors are concatenated to generate a single feature vector for classification.

#### E. Sample GPR Data and HOG Feature Descriptors

Fig. 5 shows an example region of interest (after downsampling) from a subsurface buried target (a) and a visual representation of the HOG votes for each subsurface block (b). The top image is of size  $n_i \times n_j = 18 \times 24$  pixels after downsampling (Section IV-C). For each  $18 \times 24$  image flagged for feature extraction, this application uses a grid of cells of size  $3 \times 4$  (e.g., each cell consists of a region of  $6 \times 6$  GPR data pixels). During cell-based normalization, overlapping regions of  $3 \times 3$  cells are used to constitute each block. These parameter settings were chosen for numerical convenience, but results are relatively insensitive to exact parameter settings.

The example target response in Fig. 5 shows a clear hyperbolic response, which is exactly what the application of HOG to GPR data seeks to exploit and identify in subimages containing target responses. The bottom image in Fig. 5 represents the HOG histogram votes corresponding to each

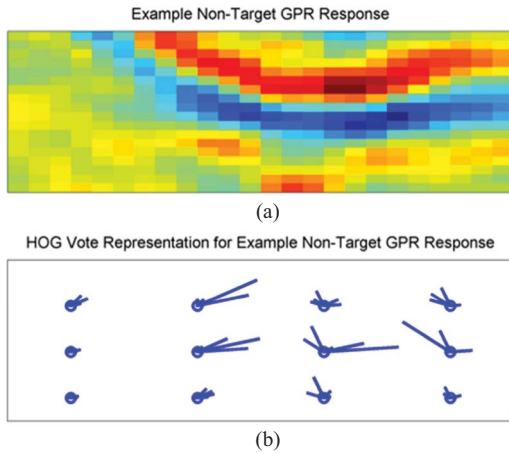


Fig. 6. (a) Sample GPR subimage and (b) resulting HOG descriptor representation extracted around a false alarm response. The HOG representation is generated by taking the resulting gradient bins and gradient magnitudes within each cell and generating lines at the appropriate angle with the appropriate magnitude. This visualization was inspired by those in [28].

of the 12 cells in the image. For each cell, each region in the bottom plot shows the resulting HOG feature vector as an image consisting of total vote magnitude-weighted lines at the angles corresponding to each bin center,  $(\phi_k + \phi_{k-1})/2$ . In this paper,  $n_\theta = 9$ , so there are nine unique possible lines in each block. The resulting histogram bins clearly indicate the dominant downward-sloping gradients as a function of proximity to the target center.

Conversely, Fig. 6 illustrates a high-energy response from a nontarget. This response has significantly less structure than the response in Fig. 5, and the resulting HOG descriptors illustrate a much less coherent distribution of HOG feature votes compared to the HOG descriptor shown in Fig. 5.

#### F. Prescreening Versus Classification

As described in the introduction, the proposed HOG features can be utilized in both a post-prescreener anomaly-classification role, or directly as a prescreening algorithm. When used as a classification algorithm, the feature extraction during training proceeds as described in Sections IV-A–IV-D, and during testing the trained classifier is used to classify data from alarms generated by the same or a similar prescreener to the one used during training.

When HOG is used as a prescreener, the training procedure follows the same approach as above. During prescreener testing, however, HOG features are extracted from every pixel in the GPR volume, and a separate faster classification algorithm is used to classify each pixel. The resulting confidence volume is then summed along the time or depth axis to form a 2-D confidence map, which is then smoothed and normalized prior to alarm declaration [38].

### V. RESULTS

This section presents results for both the HOG prescreening algorithm as well as postprocessing alarms using both HOG and EHD feature extraction and classification.

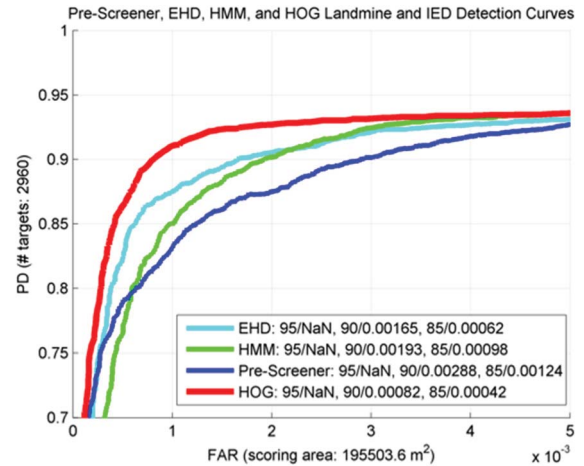


Fig. 7. Classification results for prescreening, HOG, EHD, and HMM algorithms. HOG performance on this dataset is very robust, dominating both the HMM and EHD algorithms. Note that, to show detail, the PD axes start at 70%, and the FAR axis only extends to 0.005 FA/m<sup>2</sup>. See text for additional information.

#### A. Dataset

The data used in this paper was collected with the vehicle-mounted GPR described in Section II. All data were collected at a western U.S. testing site with ground conditions varying between dirt, asphalt, and concrete. The data collected spans approximately 195 000 m<sup>2</sup>, and contains 2960 target encounters (landmine or other explosive) over 740 unique targets (the vehicle was driven over the same testing roads four times). The target populations include a mix of buried shells in varying orientations and combinations, buried explosive simulants, pressure plates, and both metallic and nonmetallic landmines. Overall, the target population is biased toward more difficult targets such as low-metal antitank landmines and low-metal pressure plates. All the targets under consideration are buried at operationally relevant depths. When the standard prescreener is run on this data, it finds about 93.5% of all targets, and generates approximately 1000 false alarms. This dataset provides a large and comprehensive body of data to enable statistical learning for both target and false alarm responses.

#### B. HOG Feature Classification and Comparison ROC Curves

For target classification using HOG features, a random forest [35] was iteratively trained and tested on varying subsets of the available data. Careful cross validation is crucial to obtaining robust performance estimates, and significant care must be taken when working with multiple passes over the same spatial regions, as with the GPR data considered here. Software developed at the University of Florida [23] provides the capability to automatically infer distinct training and testing folds during spatial cross validation by ensuring that spatial clusters of false alarms and unique target emplacements are properly divided during cross validation.

Fig. 7 shows the results obtained from evaluating a random forest (100 trees, 2 variable splits at nodes, central axis projection classifier; see [35]) on HOG features



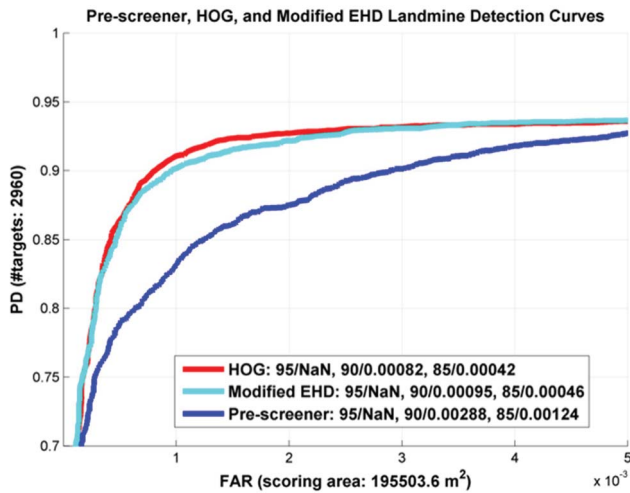


Fig. 8. Classification results comparing HOG with modified EHD features. HOG performance remains slightly better than EHD performance. The use of EHD features in the framework outlined in Section IV, along with retraining and cross validation, yields a significant improvement in algorithm performance over current fielded versions.

using the cross-validation procedure outlined in the preceding paragraph. The current implementations of EHD [22] and the HMM algorithm [21] are shown in cyan and green, respectively.

Clearly, on this very large dataset, HOG features provide performance benefits when compared to both EHD and HMM features, although two caveats are important to note. First, the HOG algorithm used in this paper was trained in cross validation on the current dataset, while the EHD and HMM implementations correspond to the current, fixed fielded implementations and therefore were not retrained. Furthermore, on other datasets, or under other weather or soil considerations, the relative performance of these algorithms is always subject to change (see [23] for a review, and [41], [42] for options to mitigate performance variability across contexts).

### C. Performance Utilizing Modified EHD Features

The results in Fig. 7 show performance for the HOG algorithm compared to a fielded version of the EHD algorithm. However, several modifications make direct comparison of these algorithms difficult. For example, as noted previously, the algorithm described in [28] makes use of several steps which were avoided during the HOG feature extraction process, including manual determination of target locations during training and the application of special-purpose classification algorithms. To more directly compare the HOG and EHD feature extraction approaches, a new version of the EHD algorithm was developed using the same EHD feature descriptor approach, but using locations and preprocessing as described in Sections IV-A, IV-B, and IV-D. Furthermore, in contrast to the special purpose algorithms developed in [22], EHD classification was performed using a random forest with the same parameters used for HOG classification. Results from this performance comparison can be seen in Fig. 7.

The results in Fig. 7 show that, when the same preprocessing and classification approaches are used, both EHD and HOG methods perform comparably, with HOG performing slightly better at several operating points of interest. Comparison of Figs. 7 and 8 illustrate how direct comparison of feature extraction methodologies in GPR data is significantly complicated by a large number of preprocessing approaches undertaken during algorithm development. These results indicate that the preprocessing techniques developed in this paper (Section IV) can provide robust performance and simplify algorithm training and implementation. These results also suggest that HOG features provide a robust technique for feature extraction and target classification in GPR data. Furthermore, as the next sections show, the HOG feature descriptors can also be used to develop prescreening algorithms, removing the artificial divide between alarm generation and final classification.

### D. HOG Prescreener Performance and Fusion

In contrast to postprocessing alarm responses, HOG feature extraction is computationally simple enough so that it is feasible to perform HOG feature extraction at every GPR pixel as a vehicle moves down-track, which enables the HOG features to be used in prescreening. Unfortunately, for our application, the random forest classification approach is too slow to be used at every pixel, though parallelization of the RF, as well as cascade-like classifiers, could be used to overcome these limitations. Instead, during prescreening, this paper makes use of a PLS-DA classifier [36]. PLS-DA is an effective technique for linear object classification, and has an extremely fast (linear) runtime, so it is more applicable to prescreener processing compared to other more complicated classifiers. From a computational perspective, the dominant sources of complexity in the HOG feature extraction process are the gradient calculation and binning of angle votes, both of which are linear in the number of pixels or number of bins under consideration.

Since prescreening algorithms process entire data collections at once, training and testing is complicated for prescreening algorithms; it is difficult to do leave-one-target-out cross validation, for example. Instead, in this paper, the prescreening version of HOG was trained using data from the large data collection outlined above, and then run on a separate data collection (taken several months later) over some of the same regions. The second validation dataset consists of 47700 m<sup>2</sup> and 654 target encounters.

Results from the on-board prescreening algorithm, HOG-based prescreening, and a fusion of the prescreening algorithms are shown in Fig. 8.

The current system prescreener is an energy-based prescreening algorithm that searches for anomalies based on a whitened energy metric and successive stages of smoothing and CFAR processing [38]. The algorithm has been successively optimized to detect target-like anomalies over several years, and is extremely well-tuned for detecting anomalies of approximate target size and shape whose energy varies even slightly from nearby background. In contrast, the HOG

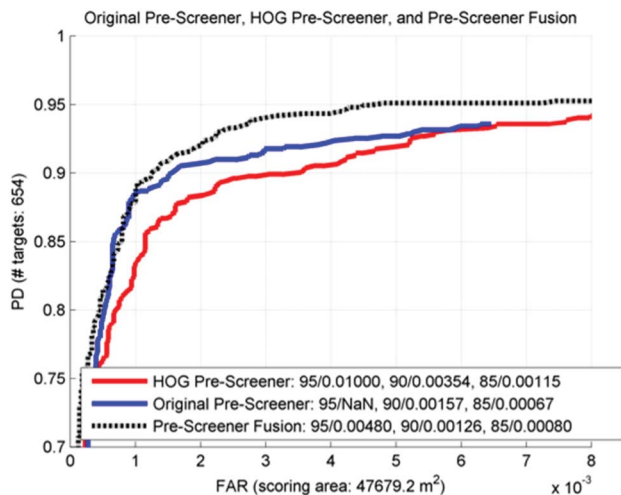


Fig. 9. Comparison of different pre-screening algorithms and pre-screener fusion. HOG pre-screening was trained on the data set explained previously, and tested on a separate data set for this experiment. Overall performance of HOG pre-screening does not match the performance of the on-board prescreener operating alone, but simple fusion of the two prescreeners significantly improves performance.

prescreening algorithm presented here does not make use of multiple CFAR steps and has not been fully optimized for overall performance due to the risks of overtraining; as a result, the HOG prescreener performs slightly worse than the on-board prescreener. However, since the HOG prescreener and on-board prescreener are sensitive to different aspects of the underlying data (HOG is sensitive to edge shapes, and relatively invariant to energy, while the on-board prescreener is primarily an energy detector), it is natural to wonder if the fusion of the on-board and HOG prescreening functions can provide additional performance improvements. The dashed black line in Fig. 8 illustrates a simple prescreener proximity fusion (if two alarms are within 0.25 m of one another, create a new alarm at the center of the alarm locations, with confidence equal to the sum of the individual alarm's confidences, otherwise retain the original alarm locations and confidences). The fusion of the two algorithms dominates either of the algorithms operating in isolation, and achieves 95% Pd at 0.0048 FA/m<sup>2</sup>, a 50% improvement in FAR compared to the HOG prescreener alone, and a significant improvement over the on-board processor, since, at the current operating threshold, 95% Pd is not achieved with the original prescreener.

Overall, both the HOG classification using a random forest and the HOG prescreener using PLS-DA enable significant improvements for subsurface threat identification in GPR data.

## VI. CONCLUSION

This paper explored the application of a particular technique from the computer vision literature, namely HOG features, to the problem of landmine and explosive detection in GPR data. Our results indicated that the performance of the HOG algorithms derived in this paper enable significant performance improvements over existing methods, and that HOG might provide a useful technique to help break down the artificial

divide between prescreening and target classification used in GPR systems due to computational bottlenecks.

This paper also developed a number of preprocessing and classification steps that enabled modifications to the EHD algorithm to overcome some previous limitations (e.g., manual target localization, special-purpose classification algorithms) and enabled simpler algorithm retraining. A comparison of EHD and HOG performance in Figs. 7 and 8 illustrated the importance that preprocessing steps have on algorithm performance and the difficulty in providing accurate feature-to-feature performance comparisons.

From a computational perspective, the HOG feature extraction procedure is quite tractable; for example, HOG features are widely used in image and video processing, where algorithms using these features are capable of achieving real-time performance. Furthermore, several open-source implementations of the HOG feature extraction code are available (e.g., [43]).

Based on the work presented in this paper, interesting avenues for further research and development are possible. First, throughout the training process, a rather *ad hoc* energy-based technique to localize actual responses of interest was utilized. Optimal extraction of locations of interest to identify and classify target versus nontarget responses is an unsolved problem in GPR processing. In the computer vision literature, corner points are often used to find regions of interest for additional processing—could a similar motivation be used to identify interesting locations in GPR data?

Also, during classifier development and execution, other *ad hoc* techniques to aggregate decision metrics from sets of images are required (e.g., average the top three confidences as in Section IV-B). This process is necessary because the underlying problem being solved is not a simple pattern recognition problem but can instead be well represented as a multiple-instance learning (MIL) problem (e.g., [39], [40], [44]–[46]). Treating the learning problem in GPR as an MIL problem is a young but active area of research that is currently being explored.

At a bigger picture level, HOG is certainly not the only technique available from the computer vision literature that might have implications for target classification in GPR data. It is possible that the work described in this paper will lead to significant leveraging of other alternative approaches for class identification from the computer vision literature being applied to solve problems in GPR data.

For example, numerous powerful and computationally simple feature extraction techniques such as BRIEF [47], random features [48], combinations of features [49], and others may enable faster and more robust target classification. Other approaches to prescreening are also possible—the Viola–Jones algorithm [27] is of particular interest due to simple computational requirements and the inclusion of cascade detectors to help maintain rapid rates of advance.

As mentioned previously, HOG is related to (but distinct from) previous approaches to target classification in GPR data, such as the EHD algorithm. As more and more image processing techniques are brought to bear on target classification, an interesting question is what properties of different

techniques make them successful for processing GPR data. Figs. 7 and 8 in this paper make it clear that direct head-to-head comparison of feature extraction approaches with different preprocessing steps is quite difficult, so care should be taken to ensure that the same preprocessing steps are utilized when attempting to make direct comparisons of feature performance.

In terms of classical machine vision properties and their relevance toward landmine detection in GPR, we hypothesize that invariance to exact target (or interest point) location and invariance to received energy will be shown to be useful properties, while invariances to rotations or affine transformations can be safely ignored or discarded. Also, novel techniques in image processing and understanding, including part-based models [37], may enable much more robust modeling of target response variations due to changing soil properties, target sizes, and target emplacements.

#### ACKNOWLEDGMENT

The authors would like to thank Dr. H. Frigui for several important suggestions and the anonymous reviewers for their insightful comments.

#### REFERENCES

- [1] *Counter-Improvised Explosive Devices: Multiple DOD Organizations are Developing Numerous Initiatives*, Government Accountability Office, Washington, DC, USA, 2012.
- [2] L. Carin, H. Yu, Y. Dalichaouch, A. R. Perry, P. V. Czipott, and C. E. Baum, "On the wideband EMI response of a rotationally symmetric permeable and conducting target," *IEEE Trans. Geosci. Remote Sens.*, vol. 39, no. 6, pp. 1206–1212, Jun. 2001.
- [3] F. Shubitidze, B. E. Barrowes, K. O'Neill, I. Shamatava, and J. P. Fernández, "A combined NSMC and pole series expansion approach for UXO discrimination," *Proc. SPIE*, vol. 6553, pp. 1–69, Apr. 2007.
- [4] D. O. Carlson, H. A. Duvoisin, III, K. L. Johnson, and M. Trishaun, "Autonomous mine detection system (AMDS) incorporating SFCW GPR and CWMD sensors for discrimination," *Proc. SPIE*, vol. 7664, p. 766414, Apr. 2010.
- [5] P. Torrione and L. Collins, "Performance Comparison of Automated Induction-Based Algorithms for Landmine Detection in a Blind Field Test," *Subsurf. Sens. Technol. Appl.*, vol. 5, no. 3, pp. 121–150, Jul. 2004.
- [6] J. D. David, *Ground Penetrating Radar*, 2nd ed. London, U.K.: Institution of Electrical Engineers, 2004.
- [7] R. Firoozabadi, E. L. Miller, C. M. Rappaport, and A. W. Morgenthaler, "Subsurface sensing of buried objects under a randomly rough surface using scattered electromagnetic field data," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 1, pp. 104–117, Jan. 2007.
- [8] C.-P. Kao, J. Li, Y. Wang, H. Xing, and C. R. Liu, "Measurement of layer thickness and permittivity using a new multilayer model from GPR data," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 8, pp. 2463–2470, Aug. 2007.
- [9] S. Lambot, E. C. Slob, I. van den Bosch, B. Stockbroeckx, and M. Vanclooster, "Modeling of ground-penetrating radar for accurate characterization of subsurface electric properties," *IEEE Trans. Geosci. Remote Sens.*, vol. 42, no. 11, pp. 2555–2568, Nov. 2004.
- [10] J. Minet, S. Lambot, E. C. Slob, and M. Vanclooster, "Soil surface water content estimation by full-waveform GPR signal inversion in the presence of thin layers," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 3, pp. 1138–1150, Mar. 2010.
- [11] O. Lopera, E. C. Slob, N. Milisavljevic, and S. Lambot, "Filtering soil surface and antenna effects from GPR data to enhance landmine detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 3, pp. 707–717, Mar. 2007.
- [12] F. Soldovieri, O. Lopera, and S. Lambot, "Combination of advanced inversion techniques for an accurate target localization via GPR for demining applications," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 1, pp. 451–461, Jan. 2011.
- [13] G. Borgioli, L. Capineri, P. Falorni, and S. Matucci, "The detection of buried pipes from time-of-flight radar data," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 8, pp. 2254–2266, Aug. 2008.
- [14] H. Chen and A. G. Cohn, "Probabilistic robust hyperbola mixture model for interpreting ground penetrating radar data," in *Proc. Int. Joint Conf. Neural Netw.*, Jul. 2010, pp. 1–8.
- [15] C. M. Rappaport, "Accurate determination of underground GPR wavefront and B-scan shape from above-ground point sources," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 8, pp. 2429–2434, Aug. 2007.
- [16] E. Pasolli, F. Melgani, and M. Donelli, "Automatic analysis of GPR images: A pattern-recognition approach," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 7, pp. 2206–2217, Jul. 2009.
- [17] J. S. Kobashigawa, H. S. Youn, M. F. Iskander, and Z. Yun, "Classification of buried targets using ground penetrating radar: Comparison between genetic programming and neural networks," *IEEE Antennas Wireless Propag. Lett.*, vol. 10, pp. 971–974, Oct. 2011.
- [18] T. Wang, J. M. Keller, P. D. Gader, and O. Sjahputera, "Frequency subband processing and feature analysis of forward-looking ground-penetrating radar signals for land-mine detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 3, pp. 718–729, Mar. 2007.
- [19] D. P. Williams, V. Myers, and M. S. Silvious, "Mine classification with imbalanced data," *IEEE Geosci. Remote Sens. Lett.*, vol. 6, no. 3, pp. 528–532, Jul. 2009.
- [20] A. B. Yoldemir and M. Sezgin, "A least squares approach to buried object detection using ground penetrating radar," *IEEE Sens. J.*, vol. 11, no. 6, pp. 1337–1341, Jun. 2011.
- [21] P. D. Gader, M. Mystkowski, and Y. Zhao, "Landmine detection with ground penetrating radar using hidden Markov models," *IEEE Trans. Geosci. Remote Sens.*, vol. 39, no. 6, pp. 1231–1244, Jun. 2001.
- [22] H. Frigui and P. Gader, "Detection and discrimination of land mines in ground-penetrating radar based on edge histogram descriptors and a possibilistic  $K$ -nearest neighbor classifier," *IEEE Trans. Fuzzy Syst.*, vol. 17, no. 1, pp. 185–199, Feb. 2009.
- [23] J. N. Wilson, P. Gader, W.-H. Lee, H. Frigui, and K. C. Ho, "A large-scale systematic evaluation of algorithms using ground-penetrating radar for landmine detection and discrimination," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 8, pp. 2560–2572, Aug. 2007.
- [24] R. Szeliski, *Computer Vision: Algorithms and Applications*. New York, NY, USA: Springer-Verlag, 2010.
- [25] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *Int. J. Comput. Vis.*, vol. 60, no. 6, pp. 91–110, Nov. 2004.
- [26] H. Bay, A. Ess, T. Tuytelaars, and L. V. Gool, "SURF: Speeded up robust features," *Comput. Vis. Image Understand.*, vol. 110, no. 3, pp. 346–359, Jun. 2008.
- [27] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, 2001, pp. 511–518.
- [28] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, vol. 1, Jun. 2005, pp. 886–893.
- [29] R. T. Sakaguchi, J. Morton, D. Kenneth, L. M. Collins, and P. A. Torrione, "Keypoint-based image processing for landmine detection in GPR data," *Proc. SPIE*, vol. 8357, p. 83571Z, May 2012.
- [30] C. Harris and M. Stephens, "A combined corner and edge detector," in *Proc. 4th Alvey Vis. Conf.*, 1988, pp. 147–151.
- [31] D. Gerónimo, A. M. López, A. D. Sappa, and T. Graf, "Survey of pedestrian detection for advanced driver assistance systems," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, no. 7, pp. 1239–1258, Jul. 2010.
- [32] P. Dollár, C. Wojek, B. Schiele, and P. Perona, "Pedestrian detection: An evaluation of the state of the art," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 4, pp. 743–761, Apr. 2012.
- [33] B. S. Manjunath, P. Salembier, and T. Sikora, *Introduction to MPEG 7: Multimedia Content Description Language*. New York, NY, USA: Wiley, 2002.
- [34] T. Kohonen, "The self-organizing map," *Proc. IEEE*, vol. 78, no. 9, pp. 1464–1480, Sep. 1990.
- [35] L. Breiman, "Random Forests," *Mach. Learn.*, vol. 45, no. 1, pp. 5–32, Jan. 2001.
- [36] S. Wold, M. Sjöström, and L. Eriksson, "PLS-regression: A basic tool of chemometrics," *Chemometrics Intell. Lab. Syst.*, vol. 58, no. 2, pp. 109–130, Oct. 2001.
- [37] P. F. Felzenszwalb, R. B. Girshick, D. McAllester, and D. Ramanan, "Object detection with discriminatively trained part-based models," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, no. 9, pp. 1627–1645, Sep. 2010.

- [38] P. A. Torrione, C. S. Throckmorton, and L. M. Collins, "Performance of an adaptive feature-based processor for a wideband ground penetrating radar system," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 42, no. 2, pp. 644–658, Apr. 2006.
- [39] T. G. Dieterich, R. H. Lathrop, and T. Lozano-Pérez, "Solving the multiple instance problem with axis-parallel rectangles," *Artif. Intell.*, vol. 89, nos. 1–2, pp. 31–71, Jan. 1997.
- [40] A. Manandhar, K. D. Morton, L. M. Collins, and P. A. Torrione, "Multiple instance learning for landmine detection using GPR," *Proc. SPIE*, vol. 8357, p. 1–6, May 2012.
- [41] C. R. Ratto, P. A. Torrione, and L. M. Collins, "Exploiting ground-penetrating radar phenomenology in a context-dependent framework for landmine detection and discrimination," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 5, pp. 1689–1700, May 2011.
- [42] H. Frigui, L. Zhang, and P. D. Gader, "Context-dependent multisensor fusion and its application to land mine detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 6, pp. 2528–2543, Jun. 2010.
- [43] G. Bradski, "The OpenCV Library," *Dr. Dobb's J. Softw. Tools*, vol. 25, no. 1, pp. 122–125, Nov. 2000.
- [44] S. Andrews, I. Tsochantaridis, and T. Hofmann, "Support vector machines for multiple-instance learning," in *Advances in Neural Information Processing Systems*, vol. 15. Cambridge, MA, USA: MIT Press, 2003, pp. 561–568.
- [45] H.-Y. Wang, Q. Yang, and H. Zha, "Adaptive p-posterior mixture-model kernels for multiple instance learning," in *Proc. 25th Int. Conf. Mach. Learn.*, New York, NY, USA, 2008, pp. 1136–1143.
- [46] J. Bolton, P. Gader, H. Frigui, and P. Torrione, "Random set framework for multiple instance learning," *Inf. Sci.*, vol. 181, no. 11, pp. 2061–2070, Jun. 2011.
- [47] M. Calonder, "Robust, high-speed interest point matching for real-time applications," Zurich, Switzerland: Swiss Federal Institute of Technology, Sep. 2010.
- [48] L. Liu and P. Fieguth, "Texture classification from random features," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 3, pp. 574–586, Mar. 2012.
- [49] X. Wang and T. X. Han, "An HOG-LBP human detector with partial occlusion handling," in *Proc. IEEE 12th Int. Conf. Comput. Vis.*, Oct. 2009, pp. 32–39.



**Peter A. Torrione** received the B.S.E.E. degree from Tufts University, Medford, MA, USA, in 1999, and the M.S.E.E. and Ph.D. degrees in electrical engineering from Duke University, Durham, NC, USA, in 2002 and 2008, respectively.

He has been a Research Associate with Duke University from 2002 to 2004, a Research Scientist from 2008 to 2009, and an Assistant Research Professor since 2009. His current research interests include physics-based statistical signal processing, pattern recognition, and machine learning.



**Kenneth D. Morton, Jr.** (M'01) was born in York, PA, USA, in 1982. He received the B.S. degree in electrical and computer engineering from The University of Pittsburgh, Pittsburgh, PA, USA, in 2004, and the M.S. and Ph.D. degrees in electrical and computer engineering from Duke University, Durham, NC, USA, in 2006 and 2010, respectively.

He is currently an Assistant Research Professor with Duke University with research focused on the use of Bayesian methods for a variety of applications including acoustic signal processing and

landmine detection.

Dr. Morton is a member of Tau Beta Pi and Eta Kappa Nu.

**Rayn Sakaguchi** (S'12) was born in Honolulu, Hawaii, in 1984. He received the B.S. degree in engineering from Trinity College, Hartford, CT, USA, in 2007. He is currently pursuing the Ph.D. degree in electrical and computer engineering with Duke University, Durham, NC, USA.

He was a Quality Engineer with the Avionics Division, Sikorsky Aircraft, Stratford, CT, from 2007 to 2010. His current research interests include computer vision, machine learning, and subsurface sensing.



**Leslie M. Collins** (M'96–SM'01) was born in Raleigh, NC, USA. She received the B.S.E.E. degree from the University of Kentucky, Lexington, KY, USA, in 1985, and the M.S.E.E. and Ph.D. degrees in electrical engineering from the University of Michigan, Ann Arbor, MI, USA, in 1986 and 1995, respectively.

She was a Senior Engineer with the Westinghouse Research and Development Center, Pittsburgh, PA, USA, from 1986 to 1990. In 1995, she became an Assistant Professor with the Department of Electrical and Computer Engineering (ECE), Duke University, Durham, NC, USA, and became an Associate Professor in ECE in 2002, and is currently a Professor with the ECE Department. Her current research interests include incorporating physics-based models into statistical signal processing algorithms, and she is pursuing applications in subsurface sensing as well as enhancing speech understanding by hearing-impaired individuals.

Dr. Collins is a member of Tau Beta Pi, Eta Kappa Nu, and Sigma Xi.