Detection and Segmentation of Concealed Objects in Terahertz Images

Xilin Shen, *Student Member, IEEE*, Charles R. Dietlein, *Member, IEEE*, Erich Grossman, Zoya Popović, *Fellow, IEEE*, and François G. Meyer, *Senior Member, IEEE*

***Abstract*—Terahertz imaging makes it possible to acquire images of objects concealed underneath clothing by measuring the radiometric temperatures of different objects on a human subject. The goal of this work is to automatically detect and segment concealed objects in broadband 0.1–1 THz images. Due to the inherent physical properties of passive terahertz imaging and associated hardware, images have poor contrast and low signal to noise ratio. Standard segmentation algorithms are unable to segment or detect concealed objects. Our approach relies on two stages. First, we remove the noise from the image using the anisotropic diffusion algorithm. We then detect the boundaries of the concealed objects. We use a mixture of Gaussian densities to model the distribution of the temperature inside the image. We then evolve curves along the isocontours of the image to identify the concealed objects. We have compared our approach with two state-of-the-art segmentation methods. Both methods fail to identify the concealed objects, while our method accurately detected the objects. In addition, our approach was more accurate than a state-of-the-art supervised image segmentation algorithm that required that the concealed objects be already identified. Our approach is completely unsupervised and could work in real-time on dedicated hardware.**

***Index Terms*—Object detection, segmentation, terahertz imaging.**

I. INTRODUCTION

AT room temperature, the peak of black body radiation is in the infra-red (IR) region of the spectrum. However, detectable power levels are emitted from objects in the terahertz and millimeter wave frequency range. The different emissivities of different materials in this frequency range enable application such as concealed weapon detection under clothing, guidance in adverse weather conditions, etc. [1]–[5]. However, passive radiometry implies very low signal levels and low SNR, as well as low contrast images. Detectors have limited sensitivity, and images are typically obtained by scanning, resulting in a small number of pixels. In fact, we show in our experiments that state of the art image segmentation algorithms fail with radiometric thermal images. In this work, we address the problem of detecting from terahertz images the contours of objects that are concealed under clothing.

In previous research, millimeter wave (MMW) imaging has been combined with infra-red imaging [6]–[8]. Although the IR images provide no useful information about the concealed objects, they help locate the human subject. Image restoration techniques have been applied simultaneously to MMW and IR images to improve their quality. Restoration techniques include Lorentzian and Wiener filtering [9], wavelet based methods [10], neural network based methods [11], [12], super-resolution algorithms [13]–[16], and others [17], [18]. After image restoration, the same regions of interest are extracted from both modalities (terahertz and IR) and the sub-images are aligned and fused. Several segmentation methods have been proposed for the fused images, e.g. in [19], the authors used Otsu's thresholding [20] method to separate concealed weapons from the rest of the images. The method is solely based on the histogram of the image, and the single value thresholding offers no guarantee to find the concealed objects. The Slamani Mapping Procedure (SMP) [21], [22] quantizes and decomposes the fused image into several homogeneous regions. Shape parameters such as circularity, Fourier descriptors and moments were used to analyze each region for recognition purpose [23]. The underlying assumption of SMP is that a concealed object can be represented by a single region. However, a low signal-to-noise ratio greatly affects the accuracy of the decomposition. It is likely that a concealed object is decomposed into two or more regions, which makes it difficult to recognize. More recently, video sequences obtained using millimeter wave imaging technology have been studied [24]. In order to segment the human body from each frame of the video sequence, an image-histogram-based segmentation approach and a model-based segmentation approach were implemented and compared, but segmentation of concealed objects was not addressed. In summary, existing methods for concealed object detection depend on the information from an IR/visible image which helps restrict the analysis to a small region on a human subject. In this paper, we show that the terahertz images alone provide sufficient information for concealed object detection and segmentation.

Because terahertz images usually have low SNR, it is advantageous to remove noise to improve the image quality. There are a number of sophisticated image denoising algorithms that can

Manuscript received December 13, 2007; revised August 04, 2008. Current version published November 12, 2008. This work was supported by the National Science Foundation, project ECS-0501578. NIST is an agency of the U.S. government; therefore, this work is not subject to copyright. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Ercan E. Kuruoglu.

X. Shen is with the Department of Radiology, Yale University, New Haven, CT 06519 USA.

C. R. Dietlein, Z. Popović, and F. G. Meyer are with the Department of Electrical Engineering, University of Colorado, Boulder, CO 80309-0425 USA (e-mail: fmeyer@colorado.edu).

E. Grossman is with the Optoelectronics Division, NIST, Boulder, CO 80305 USA.

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/TIP.2008.2006662

be used for terahertz images. The algorithms include Gaussian filtering [25], Wiener local filter [26], anisotropic diffusion [27], total variation model [28], translation invariant wavelet thresholding [29], Yaroslavsky neighborhood filters [26], and nonlocal means (NL-means) algorithm [30]. Terahertz images are piecewise smooth with no significant texture; thus, the anisotropic diffusion algorithm is appropriate for denoising the images. In our experiments, we compare the anisotropic diffusion algorithm with the NL-means algorithm. It has been proved in [30] that the NL-means algorithm is asymptotically optimal under a generic statistical image model, and, therefore, the results obtained by the NL-means algorithm should be close to the best that one can achieve.

We have explained above that existing methods to segment concealed objects in millimeter wave images suffer from significant limitations and require the use of IR images. Thus, these methods cannot be applied directly to terahertz images. The emergence of terahertz imaging systems for the detection of concealed nonmetallic weapons creates a need for fast algorithms that can remove the noise and segment images to extract the hidden objects. Such systems should be fully automated in order to be useful in environments such as airports. In this paper, we propose a fully automated method called Multilevel Thresholding, which combines a mixture model for the image histogram and a geometry analysis of the intensity isocontours. We compare our approach to three state-of-the-art segmentation algorithms: the level set approach of Chan and Vese [31], the Normalized Cut algorithm of Shi and Malik [32], and the random walk segmentation algorithm of Grady [33]. In [31], the authors propose the so-called "active contours without edges approach" to detect objects in a given image. By restricting the segmented image to be piecewise constant, they re-formulate the Mumford-Shah [34] minimization problem using level set functions. The outline of the objects is given by the zero-level sets. The piecewise constant model is appropriate for terahertz images, but the output of the segmentation algorithm is not accurate due to the low contrast across object boundaries. Over the last couple of years, graph-based image segmentation approaches [32], [33] have been developed and shown promising results. In [32], an image is represented by a graph, where each pixel becomes a node in the graph and the nodes are connected and edges are weighed. The segmentation problem then turns into a graph partition problem. In [33], the author solves the segmentation problem using an algorithm based on random walks. By prelabeling a small number of pixels, one finds the probabilities that a random walker first reaches the labeled pixels from each unlabeled pixel. We apply the two graph-based algorithms to terahertz images. We have found that the algorithm described in [32] does not work well due to the low contrast. Although the random walk algorithm produces reasonable segmentation of the concealed objects, its interactive mode of operation requires a significant amount of human interaction, and, thus, does not satisfy the requirement of automatic processing.

The paper is organized as follows. In Section II, the acquisition of passive terahertz images and the physical properties of the data are discussed. The anisotropic diffusion algorithm and the NL-means denoising method are presented in Section III, followed by a description of the proposed Multilevel Thresholding method in Section IV. We show the results of our approach in Section V and compare them with three state-of-the-art image segmentation algorithms.

II. Image Acquisition

A passive superconducting $(T = 4\ \mathrm{K})$, vacuum-bridge, antenna-coupled Nb microbolometer, described in [35], is utilized for terahertz detection. As explained in [35], the images are acquired by single-pixel, row-based raster scanning of the uncontrolled scene. Readout of the radiometric temperature is performed by electronics discussed in [36]. The images are broadband, where the bandwidth is nominally 0.1–1 THz. This bandwidth is governed by the logarithmic spiral antenna bandwidth and the low-pass filters installed in the optical path; the mean imaging frequency for purposes of spatial resolution and edge sharpness is approximately 450 GHz, as mentioned in [37]. The primary aperture of the imaging system is 30 cm in diameter, with a focal length of 25 cm. The subject in the images is approximately 1 m from the primary aperture. The human subject is fully clothed with concealed objects underneath clothing. In the terahertz frequency range, clothing has very different emissivity than the body and the concealed objects, and it appears transparent in the images. The concealed objects include a $\mathrm{ZrO_2}$ kitchen knife, a small metallic handgun, and a rectangular piece of standard RF anechoic material.

Three terahertz images, shown in Fig. 1, are referred to as "image 1," "image 2," and "image 3" in the rest of the text. The size of the three images are 108 $\times$ 118, 108 $\times$ 117, and 108 $\times$ 116, respectively. The radiometric temperature was mapped to grayscale. In addition to the concealed weapons and absorber, the images contain some details related to the clothing. For example, the shirt collar has a similar intensity as that of the concealed objects, but is of no interest in the detection process. The bright spot in the images near the subject is a temperature calibration target, which, when combined with knowledge of the background scene temperature, allows the images to be absolutely calibrated to a radiometric temperature scale in SI units. The noise-equivalent temperature difference, or temperature resolution sensitivity, achieved by these detectors referenced to an integration time of $\tau = 30$ ms is 125 mK, and their noise equivalent power is 25 fW/Hz$^{1/2}$ [36]. An NETD of this order is required for passive imaging indoors, where the radiometric temperature differences are no greater than 10 K.

III. Image Denoising

Images obtained by the terahertz imaging system usually have a low SNR. We computed the statistical distribution of the intensity distribution of an image with a constant background. We found (results not shown) that the noise can be accurately modeled as a Gaussian process. We further assume that the pixels are independent and, therefore, consider the noise to be white Gaussian. Denoising is applied to improve the quality of the terahertz images. Because the NL-means algorithm [30] is asymptotically optimal under a generic statistical image model, we use this algorithm to remove the noise. In addition, we found the anisotropic diffusion algorithm to perform very similarly for terahertz images. The results of the two denoising algorithms are provided in Section V.

Fig. 1. Image 1 (left), $T \in [285.84, 322]$. A ceramic knife and a handgun are concealed under the clothing. Image 2 (center), $T \in [286.67, 324]$. A handgun and a rectangular piece of standard RF anechoic material are concealed under the clothing. Image 3 (right), $T \in [285.88, 324]$. A ceramic knife is concealed under the clothing. The radiometric temperature T is mapped to grayscale intensity.

A. NL-Means Image Denoising Algorithm

The NL-means algorithm [30] takes advantage of the high degree of redundancy that exists in natural images. Given a small neighborhood in a natural image, one can find many other neighborhoods where the intensity function is locally the same. The intensity in all these neighborhoods can thus be averaged for removing the incoherent noise. Let I denote a noisy image. We define the similarity between pixel i and pixel j to be a weighted l_2 distance between the intensities around the two pixels

$$d^2(i,j) = \sum_{s \in Z^2} G_\alpha(s) \left(I(i+s) - I(j+s)\right)^2 \tag{1}$$

where G_α is a Gaussian kernel with standard deviation α. The Gaussian kernel controls the size of the windows around pixels i and j within which the intensity is compared. The NL-means algorithm [30] is then defined by

$$I_{NL}(i) = \frac{1}{Z(i)} \sum_j e^{-d^2(i,j)/h^2} I(j) \tag{2}$$

where $Z(i)$ is a normalization factor given by $Z(i) = \sum_j e^{-d^2(i,j)/h^2}$. The parameter h controls the amount of averaging. The algorithm is nonlocal since the estimated value $I_{NL}(i)$ at pixel i is a weighted average of all pixels j in the image. In practice, the computation is restricted within a search window to reduce the complexity of the algorithm.

B. Anisotropic Diffusion Algorithm

Convolving an image with a Gaussian kernel is equivalent to solving the heat equation with the image as the initial condition. Inspired by this equivalence, Perona and Malik [27] introduced an anisotropic diffusion equation to denoise images

$$\frac{\partial u(x,y,t)}{\partial t} = \operatorname{div}\left[g\left(|\nabla u|^2\right)\nabla u\right] \tag{3}$$

where $\nabla u(x,y) = [(\partial u/\partial x), (\partial u/\partial x)]^T$. The initial condition is given by the noisy image, $u(x,y,0) = I(x,y)$. The time parameter t controls the amount of smoothing. The function $g(s)$ determines the diffusion coefficient. Because $\lim_{s\to\infty} g(s) = 0$, (3) encourages diffusion in smooth regions where $|\nabla u|$ is small and stops diffusion at discontinuities where $|\nabla u|$ is large. We can write (3) as a weighted sum in the normal direction and in the tangential direction [38]

$$\begin{aligned}\operatorname{div}\left(g\left(|\nabla u|^2\right)\nabla u\right) &= g\left(|\nabla u|^2\right)(u_{xx}+u_{yy}) + 2g'\left(|\nabla u|^2\right)\\ &\quad \left(u_x^2 u_{xx} + u_y^2 u_{yy} + 2u_x u_y u_{xy}\right)\\ &= g\left(|\nabla u|^2\right)u_{TT} + b\left(|\nabla u|^2\right)u_{NN}\end{aligned}$$

where $b(s) = g(s) + 2sg'(s)$. The diffusion in the tangential direction is controlled by $g(s)$, and the diffusion in the normal direction is controlled by $b(s)$. Since we choose $g(s) \geq 0$ for all s, the equation results in smoothing in the tangential direction. However, in the normal direction, the sign of $b(s)$ determines whether an edge is blurred or sharpened. The evolution equation (3) can also be interpreted in terms of a gradient descent method to minimize an energy functional. Geman and Geman proposed in [39] nonconvex potentials that require more complex minimization methods.

There are several possible choices for the function $g(s)$. In our implementation of the anisotropic diffusion algorithm, we used the Tukey's biweight function, because it generates the best result. The Tukey's biweight function, given by

$$g(s) = \begin{cases} \frac{1}{2}\left(1 - \frac{s}{\sigma}\right)^2, & \text{if } |s| \leq \sigma \\ 0, & \text{otherwise} \end{cases} \tag{4}$$

was suggested by Black *et al.* [40] because of its connection to the robust error norm in robust statistics. The function b is given by

$$b(s) = \begin{cases} \frac{2}{\sigma^2}(s-\sigma)(s-\sigma/4), & \text{if } |s| \leq \sigma \\ 0, & \text{otherwise.} \end{cases} \tag{5}$$

Clearly, $b(s) > 0$ when $s < \sigma/4$, $b(s) < 0$ when $\sigma/4 < s < \sigma$, and $b(s) = 0$ when $s > \sigma$, thus the edge enhancing stops at σ.

IV. Multilevel Thresholding

A. Analysis of the Histogram

Terahertz imaging measures the radiometric temperature of the scene. In general, the background has a lower temperature, the human body has a higher temperature (clothing is transparent), and the concealed objects (knives or hand guns) have a

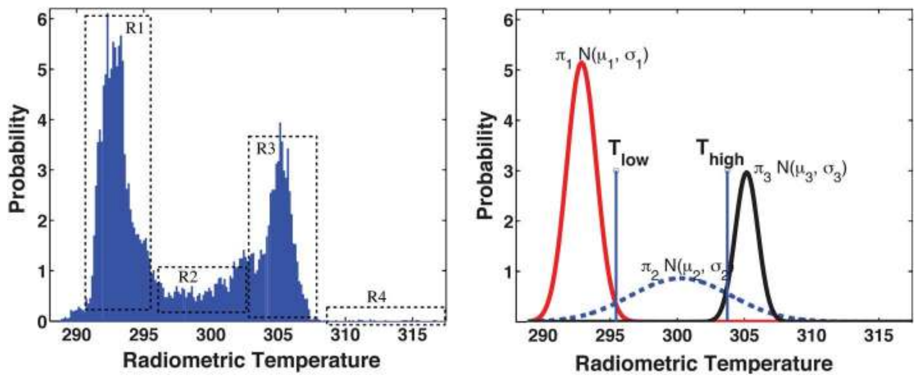


Fig. 2. Left: histogram of the intensity of denoised image 1. Right: the mixture model and the thresholds T_{low} and T_{high}.

radiometric temperature that is higher than the background but lower than the human body (Fig. 1).

Fig. 2 shows the histogram of the denoised "image 1." We interpret the histogram as the sum of four regions. Region one, denoted by "R_1," approximately ranging from 286 K to 296 K, has the lowest temperature and corresponds to the radiometric temperature of the background region. Region two, denoted by "R_2," ranging from 296 K to 303 K is the transition region between R_1 and R_3, and corresponds to the radiometric temperature of the concealed objects. Region three, denoted by "R_3," ranging from 303 to 308 K, has relatively high temperature and corresponds to the radiometric temperature of the human body. Region four, denoted by "R_4," has the highest temperature and corresponds to the radiometric temperature of the heat source for calibration. R_2 is the region which we are interested in. We model the probability density function of the image intensity (radiometric temperature of the scene) using a mixture of Gaussian densities

$$\Pr(T) = \sum_{k=1}^{3} \pi_k \mathcal{N}(T; \mu_k, \sigma_k) \tag{6}$$

where π_k is the weight of each component and $\mathcal{N}(T; \mu_k, \sigma_k)$ is the Gaussian density with mean μ_k and standard deviation σ_k. We specify the number of components of the mixture to be 3, representing the background, the human body and the concealed objects respectively. Similar mixture models have been used in [41] and [42] with Laplacian-Rayleigh densities and 2 components. Because there are only very few pixels corresponding to the heat source, we do not assign an individual component for it. The parameters (π_k, μ_k and σ_k) of the mixture model can be estimated using the expectation-maximization algorithm [43]. The estimate of the Gaussian densities for image 1 are shown in Fig. 2.

Given the estimation of the Gaussian mixture density, we can assign to each pixel a label $\beta \in \{1, 2, 3\}$ that corresponds to the most likely component of the mixture that gave rise to the temperature T at that pixel. The maximum a posteriori estimation of the label $\hat{\beta}(T)$ is given by

$$\hat{\beta}(T) = \arg\max_{\beta_k \in \{1,2,3\}} \Pr(\beta_k | T). \tag{7}$$

Fig. 3. Pixels i such that $I(i)$ is in given range. Left: $\mathrm{range} = [T_{\text{low}}, T_{\text{high}}]$. Center: $\mathrm{range} = [T_{\text{low}} + \delta, T_{\text{high}}]$. Right: $\mathrm{range} = [T_{\text{low}}, T_{\text{high}} - \delta]$. We have $T_{\text{low}} = 295.37$ K, $T_{\text{high}} = 303.35$ K, and $\delta = 3$ K.

$\Pr(\beta_k | T)$ is the posterior probability on the basis of the observation T, which can be re-written according to the Bayes' theorem as

$$\Pr(\beta_k | T) = \frac{\Pr(T | \mu_k, \sigma_k) \Pr(\beta_k)}{\Pr(T)} \simeq \pi_k \mathcal{N}(T | \mu_k, \sigma_k). \tag{8}$$

The distribution of the image intensity (see Fig. 2) is guaranteed to have the first and the third components of the Gaussian mixture well separated from each other. Therefore, for any given temperature T, we only need to make two comparisons to determine the component label according to (7). In fact, it is sufficient to compute two thresholds, T_{low} for separating the first and second components and T_{high} for separating the second and third components. T_{low} and T_{high} are defined as

$$\begin{aligned} T_{\text{low}} &= \min\{T : P(\mu_2, \sigma_2 | T) > P(\mu_1, \sigma_1 | T)\} \\ &= \min\{T : \pi_2 \mathcal{N}(T | \mu_2, \sigma_2) > \pi_1 \mathcal{N}(T | \mu_1, \sigma_1)\} \\ T_{\text{high}} &= \max\{T : P(\mu_2, \sigma_2 | T) > P(\mu_3, \sigma_3 | T)\} \\ &= \max\{T : \pi_2 \mathcal{N}(T | \mu_2, \sigma_2) > \pi_3 \mathcal{N}(T | \mu_3, \sigma_3)\}. \end{aligned} \tag{9}$$

Both T_{low} and T_{high} are marked in Fig. 2. The two thresholds are the lower and upper bounds of the transition region R_2.

Fig. 3 is the binary image showing the segmentation. Pixels with intensity values in the range $[T_{\text{low}}, T_{\text{high}}]$ are shown in white. White pixels not only come from the concealed objects but also from the boundary of the human subject. Because the imaging system is not able to produce images with sharp edges, the terahertz image shows a smooth transition from the human body to the background. Therefore, pixels located at the boundary of the human body also have temperature in R_2. It should be clear that pixels at the boundary are included in the segmentation not because of the inaccuracy of the two thresholds T_{low} and T_{high}. This is illustrated by Fig. 3-left. The binary image in Fig. 3-center is obtained using $T_{\text{low}} + \delta$ and T_{high}, and the binary image Fig. 3-right is obtained using T_{low} and $(T_{\text{high}} - \delta)$. Increasing the lower threshold by δ reduces the number of pixels from the boundary. However, some of the pixels from the concealed object (the handgun) are missing, as well. On the other hand, lowering the higher threshold causes the shape of the knife to be inexact. Therefore, both thresholds should remain as they are. Here, we denote the set of pixels that have intensity value between T_{low} and T_{high} by P

$$P = \{i : T_{\text{low}} \leq I(i) \leq T_{\text{high}}\}. \tag{10}$$

We need to partition P into two sets: P_c that contains pixels that belong to the concealed objects, and P_b that contains pixels located at the boundary of the human body

$$P = P_c \cup P_b. \tag{11}$$

B. Multilevel Thresholding

We describe here a new method, called Multilevel Thresholding, to partition P into P_c and P_b. The method takes advantage of the smooth transition of the temperature from the background to the human body. We note that by continuously increasing the lower threshold T_{low}, the boundary of the human body shrinks inward continuously. The Multilevel Thresholding method is designed to keep track of this evolution of this boundary. The general idea is as follows.

We partition the set P into a union of level sets

$$C_k = \{i : T_k \leq I(i) < T_k + \delta\}, \quad T_{k+1} = T_k + \delta \tag{12}$$

where, T_k goes from T_{low} to $(T_{\text{high}} - \delta)$ and i is the coordinate of each pixel in the level set. All C_k's are disjoint and $\cup_k C_k = P$. Now, each C_k can be further partitioned into two disjoint subsets B_k and H_k, where B_k is the subset of pixels that belong to the boundary of the human body and H_k is the subset of pixels that belong to the concealed objects

$$B_k = C_k \cap P_b, \quad H_k = C_k \cap P_c. \tag{13}$$

The Multilevel Thresholding first finds an initial boundary set B_0 and then recursively identifies the subset B_k in each C_k. By eliminating the B_ks from P, we are able to recover the set P_c that corresponds to the concealed objects. Fig. 4 is a cartoon example that shows how the level sets C_ks are spatially located and the relationship among the sets C_ks, B_ks and H_ks. Note that (13) does not provide a way to find B_k or H_k, since we do not have access to either P_b or P_c. The B_ks are actually identified based on the observation that the radiometric temperature changes smoothly from the background to the human body; therefore, in successive B_ks, the pixels are spatially connected. One can start from an initial set B_0 that defines the boundary of the human subject, and search in C_1 for all pixels that are spatially connected to B_0. These pixels must form the subsets B_1, which will serve as the reference set for finding B_2. We

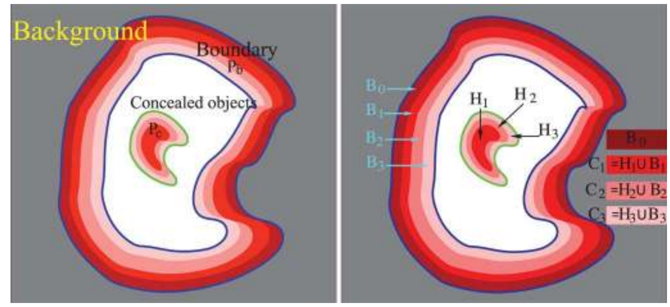


Fig. 4. Left: pixels in various shades of red have intensity between T_{low} and T_{high}. The boundary of the region P_b is shown in blue. The outline of the concealed object P_c is drawn in green. Pixels colored with the same shade of red belong to the same level set C_k. Right: each C_k can be partitioned into two disjoint sets B_k and H_k, where B_k contains pixels from the boundary region, and H_k contains pixels from the concealed object. Successive B_k's are spatially connected. B_0 is the starting set.

illustrate in Fig. 5 the Multilevel Thresholding using a 1-D example. First, a line segment is extracted from a denoised terahertz image (see Fig. 5). Starting from the left, points along the segment are in the background, in the human body, across the handgun and the collar, and finally are again in the background. The transition of the intensity between the different regions is shown in the top-right plot in Fig. 5. We mark the first few level sets C_1, C_2, C_3 and the initial boundary set B_0 on the bottom-left plot. In C_1, all pixels are spatially connected to B_0, so $B_1 = C_1$. Similarly in C_2, all pixels are spatially connected to B_1; therefore, $B_2 = C_2$. In C_3, part of the pixels are spatially connected to B_2. These pixels are classified to be in B_3, while the rest of the pixels form the set H_3. Recursively, we can recover all subsets B_k's, as described in the algorithm given in Fig. 9. The union of the B_ks gives P_b and the complement set of P_b is P_c. The sets P_b and P_c are shown in Fig. 7. The Multilevel Thresholding algorithm is given in Fig. 9. The step size δ between the level sets should be sufficiently small to guarantee that the consecutive B_k's remains connected. In other words, if we define that a pixel is spatially connected to its 8-neighbors, then pixels in B_{k+1} must be one of the 8-neighbors of pixels in B_k. If δ is too large, then some pixels in B_{k+1} may not be directly connected to pixels in B_k. In this case, these pixels will by mistake be identified to belong to the set H_{k+1}. Fig. 6 shows how the choice of δ affects the construction of B_k's. In our experiments, we choose δ to be $(T_{\text{high}} - T_{\text{low}})/50$.

V. Experiments

We generate noisy images by adding white Gaussian noise of increasing strength (σ = 0.5, 1.0, 1.5 , and 2.0 K) to the three images shown in Fig. 1. For each noise level, we create ten realizations of the noisy images. Note that the original terahertz images on which we add noise are already noisy.

A. Image Denoising

In Fig. 8, we show the images denoised using the anisotropic diffusion and the NL-means algorithm. Both algorithms work equally well in terms of mean squared error (MSE). The denoised images using the NL-means are visually more pleasant. However, the NL-means method is much slower than the anisotropic diffusion algorithm. On a personal computer, it

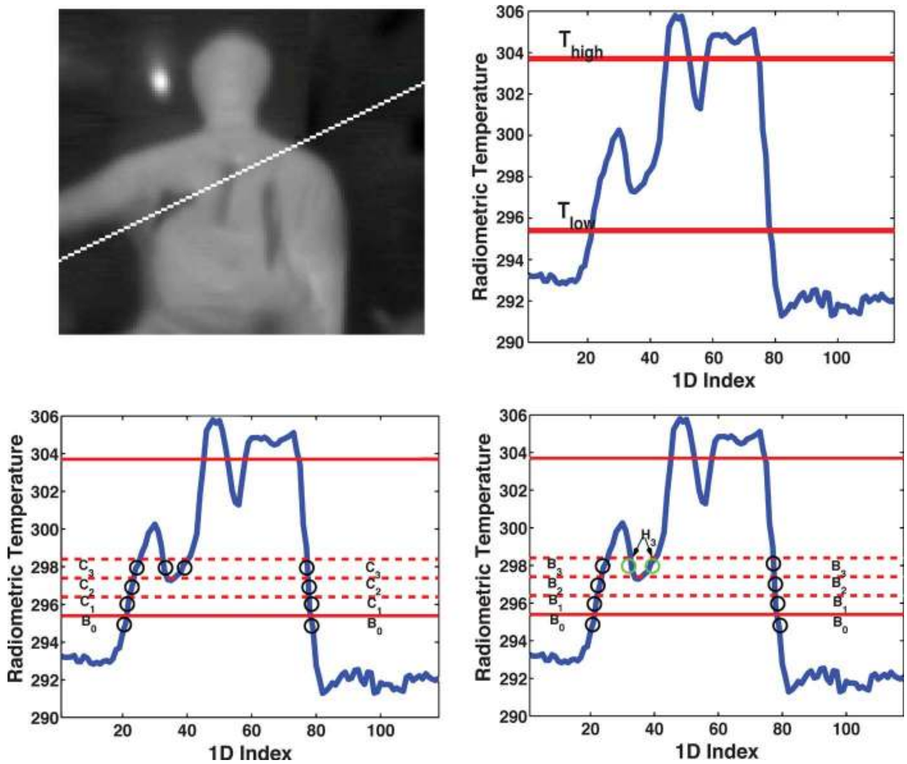


Fig. 5. Intensity along a segment from denoised image 1 (top left) is plotted against the arc length (top right). Bottom left: the level sets C_1, C_2, C_3 and initial boundary set B_0. Bottom right: the subsets B_1, B_2, B_3 that contain pixels located at the boundary of the human body, and the set H_3 that contains pixels corresponding to the concealed objects.

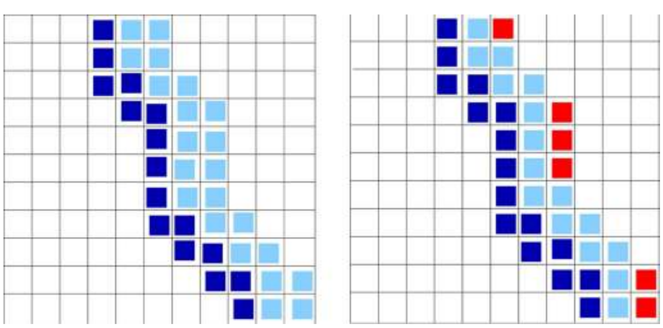


Fig. 6. Left: dark blue pixels are from B_k. Light blue pixels are from the level set C_{k+1}; if classified correctly, these pixels should all belong to the set of boundary pixels P_b. Right: because δ is too large, some of the light blue pixels are not in the 8-neighbors of the dark blue pixels. According to the spatial connectivity constraint, only part of the light blue pixels are classified as in B_{k+1}, while the rest of them, which are painted in red, are by mistake classified as in H_{k+1}.

Fig. 7. Set of pixels with intensity in $[T_{\mathrm{low}}, T_{\mathrm{high}}]$ is partitioned into P_b and P_c. Pixels in P_b (shown in red) are located at the boundary of the human body. Pixels in P_c (shown in white) belong to the concealed objects.

takes about 0.8 to 1.0 s to denoise an image using the anisotropic diffusion, while it takes about 90 to 94 s for the NL-means algorithm. Therefore, we only use the anisotropic diffusion algorithm in the remaining of the paper. We now confirm our visual impression with a quantitative comparison. Let $\tilde{I}_\sigma$ be the denoised image for the noise level σ, and let I_o be the original terahertz image. The mean squared error is given by $\mathrm{MSE}_1 = \sum_i^N (\tilde{I}_\sigma(i) - I_o(i))^2/N$.

Because the denoising will remove some of the noise of the original image, MSE_1 may not be a good indicator of the performance of the denoising. To address this issue, we compare the denoised image to a "cleaned" original image, $\tilde{I}_o$, which is obtained by a light denoising with the anisotropic diffusion. We define a second mean squared error

$$\mathrm{MSE}_2 = \frac{1}{N}\sum_i^N \left(\tilde{I}_\sigma(i) - \tilde{I}_o(i)\right)^2.$$

The statistics in Table I confirm that anisotropic diffusion and NL-means perform similarly for terahertz images. This result is specific to terahertz images and, therefore, does not contradict the results obtained in [30] for the class of natural images.

B. Concealed Object Detection and Segmentation

Fig. 11 (right column) shows the contours of the concealed objects detected by our algorithm. While we accurately segment

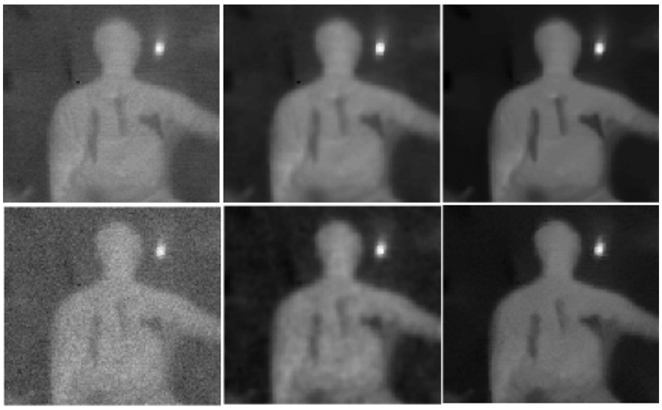


Fig. 8. Noisy image (left), denoised images: anisotropic diffusion (center) and nonlocal means (right). Top: noise level $\sigma = 0.5$, bottom $\sigma = 2.0$.

Input: image I
Algorithm:
1) Denoise the image I with anisotropic diffusion
2) Fit a three-class Gaussian mixture model to the image intensity distribution.
3) Compute T_{low} and T_{high} according to (9).
4) Compute the set P_c of pixels that belong to the concealed objects:
   a) Create an initial segmentation. Take T_{low} as the threshold and create a binary image F_0.
   $$F_0(i) = \begin{cases} 1, & I(i) \geq T_{\mathrm{low}} \\ 0, & \text{otherwise} \end{cases} \tag{14}$$
   b) Mark the boundary of the initial segmentation. The boundary set B_0 of the initial segmentation F_0 is defined as:
   $$B_0 = \{i : F_0(i) = 1 \text{ and } \exists\, j \in \eta_i\ F_0(j) = 0\} \tag{15}$$
   where η_i is the 8-neighborhood of pixel i.
   c) Find the level sets C_k's. C_k's are defined in (12).
   d) Keep track of the evolution of the boundary of the human body. Since pixels in consecutive B_k's are spatially connected according to the smoothness assumption, we find in C_k all pixels that are spatially connected to B_{k-1}. These pixels form the set B_k. The complement set is H_k.
   e) Stop. The set of pixels that corresponds to the concealed objects is $P_c = \cup_k H_k$.

Output: set P_b and set P_c.

Fig. 9. Multilevel thresholding algorithm.

the concealed objects, we also detect regions that correspond to the collar of the clothing (see Fig. 11). Fig. 12 illustrates the performance of the algorithm when noise is added ($\sigma = 2.0$) to the original images. It takes approximately 7 to 9 s to complete the detection and segmentation on a personal computer using MATLAB[1] code.

C. How Accurate is the Segmentation?

To quantify the accuracy of the segmentation we compare the results with hand-labeled segmentations (see Fig. 10). For each concealed object, we have access the set of pixels, A, enclosed inside the hand-drawn contour. We also have access to a region B detected automatically by our method.

[1]Certain commercial software is identified in this paper in order to specify the experimental procedure adequately. Such identification is not intended to imply recommendation or endorsement by NIST, nor is it intended to imply that the software identified is necessarily the best available for the purpose.

TABLE I
RESIDUAL MEAN SQUARED ERROR AFTER DENOISING

	noise level σ	Anisot. Diff. (MSE_1, MSE_2)		NL Means (MSE_1, MSE_2)	
Image 1	0.5	0.177	0.021	0.182	0.019
	1	0.242	0.082	0.259	0.062
	1.5	0.352	0.137	0.337	0.135
	2	0.448	0.233	0.455	0.263
Image 2	0.5	0.216	0.021	0.210	0.020
	1	0.284	0.083	0.287	0.076
	1.5	0.415	0.138	0.374	0.158
	2	0.510	0.230	0.501	0.292
Image 3	0.5	0.194	0.021	0.215	0.018
	1	0.261	0.082	0.281	0.065
	1.5	0.387	0.145	0.366	0.144
	2	0.492	0.250	0.490	0.276

Fig. 10. Left to right: hand labeled segmentation for images 1, 2, and 3.

A measure of mutual proximity of the two sets A and B is provided by Hausdorff distance

$$D(A, B) = \max\left(\max_{i \in A} d(i, B), \max_{j \in B} d(j, A)\right) \tag{16}$$

where $d(i, B) = \min_{j \in B} d(i, j)$ is the distance of a pixel i to the set B. The statistics are summarized in Table II: the smaller the Hausdorff distance, the more accurate is the segmentation. For noisy images, the Hausdorff distance is averaged over the ten realizations. The proposed method works well even at high noise level.

D. Comparison to Other Segmentation Methods

In this section, we compare the results obtained using the Multilevel Thresholding with three state-of-the-art image segmentation algorithms. Both the Ncut and the Active Contour without Edges algorithm are unsupervised techniques. The random random walk algorithm can operate in two modes: an interactive mode [33] and a fully automated mode [44]. The automatic mode requires that an estimate of the probability density function of each class (regions R_1 to R_4 in our case) be available. Because region R_4 (concealed objects) and the boundary of the human subject have very similar intensity distributions we only use the random walk algorithm in an interactive mode.

Fig. 11. Left: original images (1, 2, and 3 from top to bottom); right: segmentation of the concealed objects detected by our algorithm.

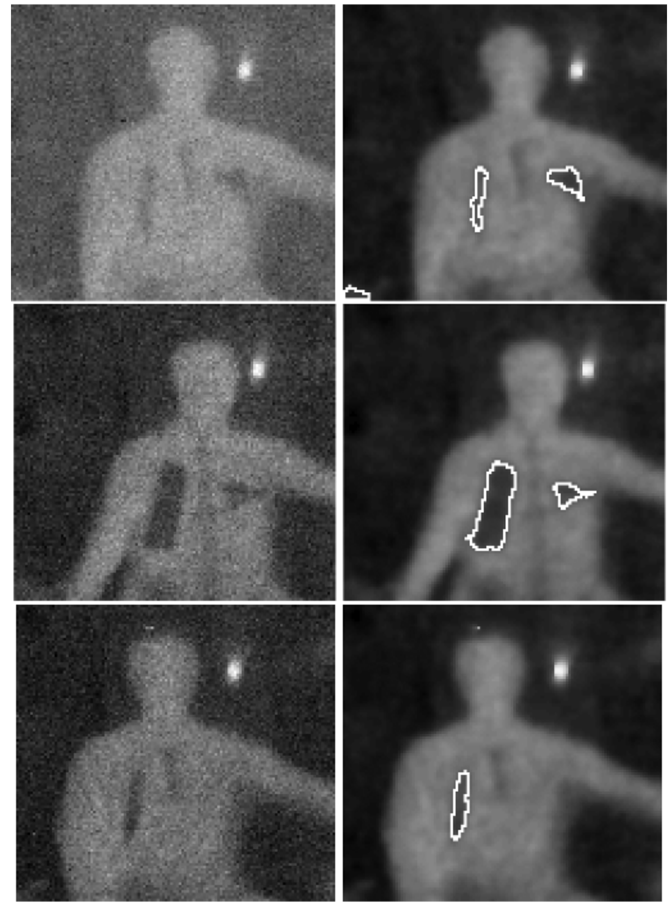


Fig. 12. Left: noisy image at $\sigma = 2.0$; right: segmentation of the concealed objects detected by our algorithm.

1) Normalized Cut (Ncut): The Ncut algorithm relies on spectral graph theory to partition the image into two regions. Each region is as homogeneous as possible, and the two regions are as different as possible. The image I is mapped onto a graph $G = (V, E)$. Each pixel becomes a vertex of V. A distance between any two vertices i and j is defined. It combines the difference between the intensity values at the corresponding pixels with their spatial distance. An edge exists between any two vertices i and j if their distance is smaller then a specified threshold r. A weight matrix W encodes the distance along the edges. Cour *et al.* [45] proposed a multiscale variation of the standard Ncut algorithm. The weight matrix W is decomposed into different scales and the image is segmented at the corresponding scale. We applied this algorithm [45] to the denoised "image 1." We generated a series of segmentations with an increasing number of classes (see Fig. 13). Unfortunately, none of the segmentation was able to detect the concealed objects. Obviously the concealed objects have a small size, making their detection more difficult. But the real difficulty is the absence of strong contrast between the concealed objects and the human body. We conclude that the multiscale Ncut method is not appropriate for detecting and segmenting the concealed objects.

2) Random Walk Algorithm: We use the random walk algorithm only in its interactive mode [33]. The automated mode [44] is not used here because of the similarity between the distribution of the intensity at the boundary of the human body and the distribution of the intensity in region R_4. The interactive mode of operation allows us to very carefully choose seeds in five classes: the heat source, the background, the human body,

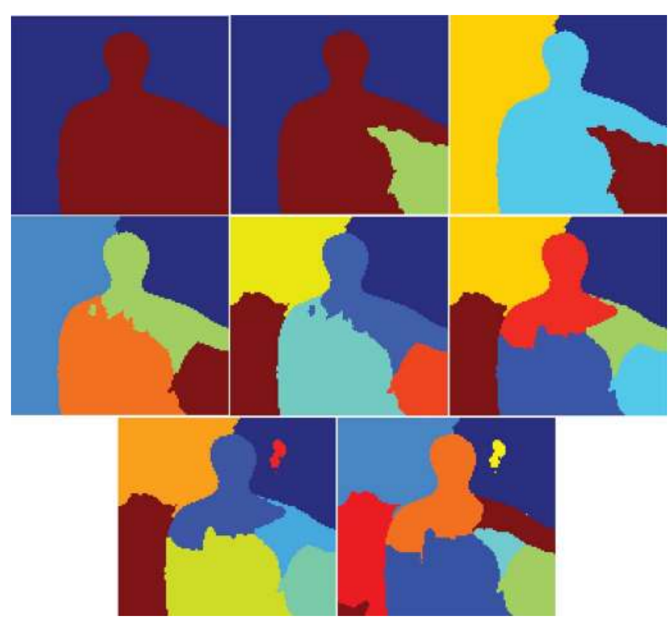


Fig. 13. Segmentation obtained by the multiscale Ncut algorithm. Top to bottom, left to right: the image is partitioned into $k = 2, 3, \ldots, 9$ classes.

TABLE II
QUANTITATIVE EVALUATION OF THE ALGORITHM

	noise level	Hausdorff distance			
	σ	Multilevel (object 1, object 2)		Random walk (object 1, object 2)	
Image 1	0	2	2	2	3
	0.5	1.5	1.5	2.3	3
	1	1.8	1.9	2.1	3
	1.5	1.9	2.0	2.5	2.8
	2	1.9	2.2	2.4	2.8
Image 2	0	1	4	2	2
	0.5	1.0	3.7	2	2
	1	1.0	3.6	2.1	1.9
	1.5	1.1	3.2	2.5	2.4
	2	1.6	3.1	2.5	2.1
Image 3	0	2		4	
	0.5	2		3.5	
	1	1.8		3.2	
	1.5	1.4		3.3	
	2	1.6		2.7	

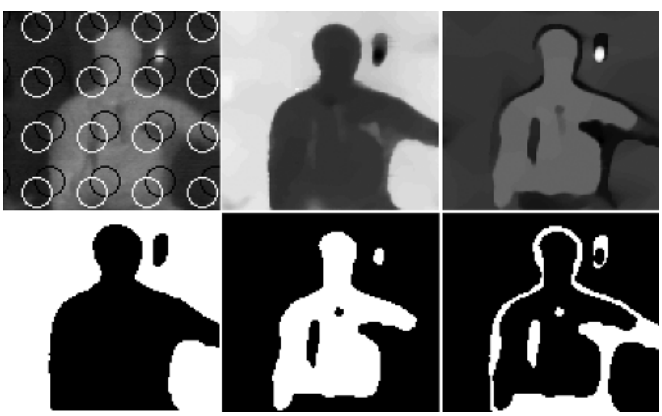


Fig. 14. Top (left to right): initialization of the level set functions for the "active contour without edges" algorithm; final level set function ϕ_1, and final level set function ϕ_2. Bottom (left to right): regions 2, 3, and 4 are shown in white.

concealed object 1 and concealed object 2 (for the first two images). The selection of the seeds is crucial to the success of the segmentation. Fig. 15 shows the segmentation of the three original images using this supervised method. The prelabeled seed pixels and the final segmentation of the five classes are displayed using different colors. The shape and size of the concealed objects in Fig. 15 are approximately the same as in the segmentation obtained using the Multilevel Thresholding (see Table II). However, one should be aware that the seed pixels have to be chosen very carefully: the optimal segmentation was obtained by choosing the seed pixels to be very close to the blurred boundaries of the concealed objects (see Fig. 15 top). Obviously, this manual intervention makes the detection of concealed objects unrealistic: if the user knows where the hidden objects are, then she doesn't need a method to detect them!

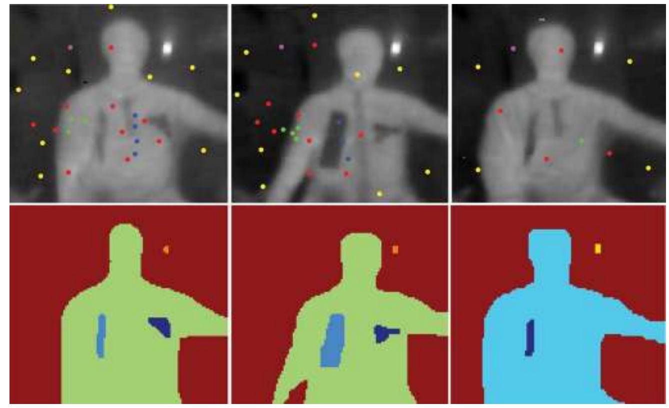


Fig. 15. Bottom: segmentation obtained by the random walker algorithm. Top: prelabeled pixels are marked on top of the denoised image.

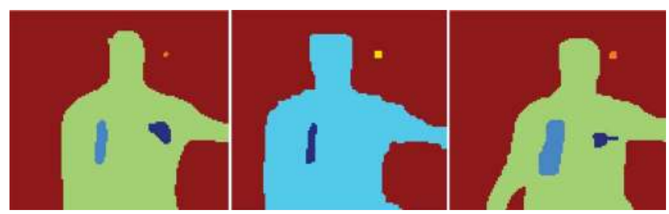


Fig. 16. Segmentation of the noisy terahertz images at ($\sigma = 2.0$) with the random walker algorithm.

The segmentation of the images that have been corrupted with Gaussian noise ($\sigma = 2.0$) is shown in Fig. 16.

VI. CONCLUSION

Despite the fact that terahertz images have low signal to noise ratio and low contrast, we have successfully achieved our goal to automatically detect and segment concealed objects in broadband 0.1–1 THz images. The proposed method combines the analysis of the image histogram and the geometry of the intensity isocontours. It is completely unsupervised and computationally efficient. The comparison to results obtained using both unsupervised and supervised methods have demonstrated that our approach outperforms the state-of-the-art supervised segmentation techniques.

REFERENCES

[1] H. Chen, S. Lee, R. Rao, M.-A. Slamani, and P. Varshney, "Imaging for concealed weapon detection," *IEEE Signal Process. Mag.*, pp. 52–61, March 2005.

[2] A. Agurto, Y. Li, G. Tian, N. Bowring, and S. Lockwood, "A review of concealed weapon detection and research in perspective," in *Proc. IEEE Int. Conf. MonE02 Networking, Sensing and Control*, 2007, pp. 443–448.

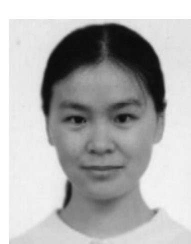
[3] R. W. McMillan, "Terahertz imaging, millimeter-wave radar," in *Advances in Sensing With Security Applications*, J. Bymes, Ed. New York: Springer Verlag, 2006, pp. 243–268.

[4] P. H. Siegel, "Terahertz technology," *IEEE Trans. Microw. Theory Tech.*, vol. 50, no. 3, pp. 910–928, Jun. 2002.

[5] R. Appleby, "Passive millimetre-wave imaging and how it differs from terahertz imaging," *Phil. Trans. R. Soc. Lond. A*, vol. 362, pp. 379–394, 2005.

[6] P. K. Varshney, H. Chen, and R. M. Rao, "On signal/image processing for concealed weapon detection from stand-off range," in *Proc. SPIE Optics and Photonics in Global Homeland Security*, T. T. Saito, Ed., 2005, vol. 5781, pp. 93–97.

[7] M.-A. Slamani, P. K. Varshney, and D. D. Ferris, "Survey of image processing techniques applied to the enhancement and detection of weapons in mmw data," in *Proc. SPIE Infrared and Passive Millimeter-Wave Imaging Systems: Design, Analysis, Modeling, and Testing*, D. A. W. R. Appleby and G. C. Holst, Eds., 2002, vol. 4719, pp. 296–305.
[8] H. M. Chen and P. K. Varshney, "Automatic two-stage IR and MMW image registration algorithm for concealed weapons detection," *Proc. IEE, Vis., Image, Signal Process.*, vol. 148, pp. 209–216, 2001.
[9] A. H. Lettington, M. R. Yallop, and S. Tzimopoulou, "Restoration techniques for millimeter-wave images," in *Proc. SPIE Passive Millimeter-Wave Imaging Technology V*, R. A. R. M. Smith, Ed., 2001, vol. 4373, pp. 94–104.
[10] H. Park, S.-H. Kim, and M. K. Singh *et al.*, "Performance analysis of wavelet based restoration for passive millimeter-wave images," in *Proc. SPIE Passive Millimeter-Wave Imaging Technology VIII*, R. Appleby, Ed., 2005, vol. 5789, pp. 157–166.
[11] K. Shimizu, K. Yamada, and K. Watabe *et al.*, "Image restoration for millimeter wave images by feedforward neural network," in *Proc. 4th Eur. Congr. Intelligent Techniques and Soft Computing*, 1996, vol. 1, pp. 334–338.
[12] K. Yuasa, H. Sawai, and K. Watabe *et al.*, "Image restoration for millimeter wave images by Hopfield neural network," in *Proc. 3rd Workshop on Image and Signal Processing on the Theme of Advances in Computational Intelligence*, 1996, pp. 321–324.
[13] H.-Y. Pang, M. K. Sundareshan, and S. Amphay, "Superresolution of millimeter-wave images by iterative blind maximum-likelihood restoration," in *Proc. Int. Soc. Optical Engineering*, 1997, vol. 3064, pp. 227–238.
[14] A. H. Lettington, D. Dunn, and M. P. Rollason *et al.*, "Use of constraints in the superresolution of passive millimeter-wave images," in *Proc. SPIE Int. Soc. Optical Engineering*, 2003, vol. 5077, pp. 100–109.
[15] M. K. Sundareshan and S. Bhattacharjee, "Superresolution of passive millimeter-wave images using a combined maximum-likelihood optimization and projection-onto-convex-sets approach," in *Proc. SPIE Int. Soc. Optical Engineering*, 2001, vol. 4373, pp. 105–116.
[16] A. H. Lettington, M. R. Yallop, and D. Dunn, "Review of super-resolution techniques for passive millimeter-wave imaging," in *Proc. SPIE Int. Soc. Optical Engineering*, 2002, vol. 4719, pp. 230–239.
[17] S. Lee, R. Rao, and M.-A. Slamani, "Noise reduction and object enhancement in passive millimeter wave concealed weapon detection," in *Proc. IEEE Int. Conf. Image Processing*, 2002, vol. 1, pp. 509–512.
[18] B. R. Hunt and D. Dekruger, "Bayesian restoration of millimeter wave imagery," in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing*, 1994, vol. 5, pp. 549–552.
[19] L. C. Ramac, M. K. Uner, and P. K. Varshney *et al.*, "Morphological filters and wavelet based image fusion for concealed weapons detection," in *Proc. SPIE Int. Soc. Optical Engineering*, 1998, vol. 3376, pp. 110–119.
[20] N. Otsu, "A threshold selection method from gray-level histogram," *IEEE Trans. Syst., Man, and Cybern.*, vol. 8, no. 1, pp. 62–66, Jan. 1978.
[21] M. A. Slamani, P. K. Varshney, and R. M. Rao *et al.*, "An integrated platform for the enhancement of concealed weapon detection sensors," in *Proc. SPIE Int. Symp. Enabling Technologies for Law Enforcement and Security*, 1998, vol. 3575, pp. 68–78.
[22] M. A. Slamani, V. Vannicola, and D. D. Weiner, "An automated approach to the partitioning and statistical characterization of a surveillance volume," presented at the Int. Conf. Signal Processing Application and Technology, 1995.
[23] M. A. Slamani and D. D. Ferris, "Shape descriptors based detection of concealed weapons in millimeter-wave data," in *Proc. SPIE Optical Pattern Recognition XII*, 2001, vol. 4387, pp. 176–185.
[24] C. D. Haworth, B. G. Gonzalez, and M. Tomsin *et al.*, "Image analysis for object detection in millimeter-wave images," in *Proc. SPIE Int. Soc. Optical Engineering*, 2004, vol. 5619, pp. 117–128.
[25] M. Lindenbaum, M. Fischer, and A. M. Bruchstein, "On Gabor contribution to image enhancement," *Pattern Recognit.*, vol. 27, pp. 1–8, 1994.
[26] L. P. Yaroslavsky*, Digital Picture Processing. An Introduction*. Berlin, Germany: Springer-Verlag, 1985.
[27] P. Perona and J. Malik, "Scale space and edge detection using anisotropic diffusion," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 12, no. 7, pp. 629–639, Jul. 1990.
[28] L. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Phys. D*, vol. 60, pp. 259–268, 1992.
[29] R. R. Coifman and D. Donoho*, Wavelets and Statistics*. New York: Springer-Verlag, 1995.
[30] A. Buades, B. Coll, and J. M. Morel, "A review of image denoising algorithms, with a new one," *Multiscale Model. Simul.*, vol. 4, pp. 490–530, 2005.
[31] T. T. Chan and L. A. Vese, "Active contours without edges," *IEEE Trans. Image Process.*, vol. 10, no. 2, pp. 266–277, Feb. 2001.
[32] J. Shi and J. Malik, "Normalized cuts and image segmentation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 22, pp. 888–905, 2000.
[33] L. Grady, "Random walks for image segmentation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 28, pp. 1–17, 2006.
[34] D. Mumford and J. Shah, "Optimal approximation by piecewise smooth functions and associated variational problems," *Commun. Pure Appl. Math.*, vol. 42, pp. 577–685, 1989.
[35] A. Luukanen and J. P. Pekola, "A superconducting antenna-coupled hot-spot microbolometer," *Appl. Phys. Lett.*, vol. 82, pp. 3970–3972, 2003.
[36] A. Luukanen, E. N. Grossman, A. J. Miller, P. Helisto, J. S. Penttila, H. Sipola, and H. Seppa, "An ultra-low noise superconducting antenna-coupled microbolometer with a room-temperature read-out," *IEEE Microw. Wireless Compon. Lett.*, vol. 16, no. 8, pp. 464–466, Aug. 2006.
[37] C. Dietlein, A. Luukanen, F. Meyer, Z. Popovic, and E. Grossman, "Phenomenology of passive broadband terahertz images," presented at the 4th ESA Workshop on Millimetre-Wave Technology and Applications, Feb. 2006.
[38] G. Aubert and P. Kornprobst*, Mathematical Problems in Image Processing: Partial Differential Equations and the Calculus of Variations*. New York: Springer-Verlag, 2002.
[39] S. Geman and D. Geman, "Stochastic relaxation, Gibbs distributions and the bayesian restoration of images," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 6, no. 6, pp. 721–741, Nov. 1984.
[40] M. J. Black, G. Sapiro, D. H. Marimont, and D. Heeger, "Robust anisotropic diffusion," *IEEE Trans. Image Process.*, vol. 7, no. 3, pp. 421–43, Mar. 1998.
[41] C. D. Haworth, Y. Petillot, and E. Trucco, "Image processing techniques for metallic object detection with millimetre-wave images," *Pattern Recognit. Lett.*, vol. 27, pp. 1843–1851, 2006.
[42] C. D. Haworth, Y. de Saint-Pern, D. Clark, E. Trucco, and Y. Petillot, "Detection and tracking of multiple metallic objects in millimetre-wave images," *Int. J. Comput. Vis.*, vol. 71, no. 2, pp. 183–196, 2007.
[43] G. Mclachlan and T. Krishnan*, The EM Algorithm and Extensions*. New York: Wiley, 1997.
[44] L. Grady, "Multilabel random walker image segmentation using prior models," in *Proc. IEEE Computer Society Conf. Computer Vision and Pattern Recognition*, 2005, vol. 1, pp. 763–770.
[45] T. Cour, F. Benezit, and J. Shi, "Spectral segmentation with multiscale graph decomposition," presented at the IEEE Int. Conf. Computer Vision and Pattern Recognition, 2005.
[46] L. A. Vese and T. F. Chan, "A multiphase level set framework for image segmentation using the mumford and shar model," *Int. J. Comput. Vis.*, vol. 50, pp. 271–293, 2002.



Xilin Shen (S'99) received the B.S. degree in electrical engineering, with honors from Zhejiang University, China, in 2002, and the Ph.D. degree in electrical engineering from the University of Colorado, Boulder, in 2007.

Since May 2008, she has been a Postdoctoral Associate at the Magnetic Resonance Research Center, Yale University, New Haven, CT. Her research interests are fMRI analysis, data mining, and image processing.

Charles R. Dietlein (M'99) received the B.S. degree in electrical engineering from Seattle Pacific University, Seattle, WA, in 2003, and the M.S. and Ph.D. degrees in electrical engineering from the University of Colorado, Boulder, in 2005 and 2008, respectively.

He is currently with the Optoelectronics Manufacturing Division, NIST, Boulder, CO. His research interests include active room-temperature and passive cryogenic millimeter-wave/terahertz imaging, measurement, and calibration techniques for both room-temperature and superconducting millimeter-wave/terahertz detectors, and broadband/multispectral image phenomenology.

Erich Grossman, photograph and biography not available at the time of publication.

Zoya Popović (F'02) received the Dipl.Ing. degree from the University of Belgrade, Serbia, Yugoslavia, in 1985, and the Ph.D. degree from the California Institute of Technology, Pasadena, in 1990.

Since 1990, she has been with the University of Colorado, Boulder, where she is currently the Hudson Moore, Jr. Chaired Professor of Electrical and Computer Engineering. She was a visiting professor at the Technical University of Munich, Germany, in 2001. Her research interests include high-efficiency, low-noise and broadband microwave and millimeter-wave circuits, quasi-optical millimeter-wave techniques for imaging, smart and multibeam antenna arrays, intelligent RF front ends, RF optics, and wireless powering for batteryless sensors.

Dr. Popović is the recipient of the 1993 and 2006 Microwave Prizes presented by IEEE MTT-S for best journal paper. She was the recipient of the 1996 URSI Issac Koga Gold Medal and the Humboldt Research Award for Senior U.S. Scientists from the German Alexander von Humboldt Stiftung. She was also awarded the 2001 HP/ASEE Terman Medal for combined teaching and research excellence.

François G. Meyer (SM'06) received the M.S. degree in applied mathematics (with honors) from the Ecole Nationale Superieure d'Informatique et de Mathematiques, Grenoble, France, in 1987, and the Ph.D. degree in electrical engineering from INRIA, France, in 1993.

He is currently an Associate Professor with the Department of Electrical Engineering, University of Colorado, Boulder. He was previously an Assistant Professor at Yale University, New Haven, CT; a Visiting Professor at the Institute Henri Poincaré; a Senior Fellow at the Institute of Pure and Applied Mathematics, University of California, Los Angeles; and a Visiting Research Scholar at Princeton University, Princeton, NJ.