

# MilliCam: Hand-held Millimeter-Wave Imaging

Moh Sabbir Saadat; Sanjib Sur; Srihari Nelakuditi  
University of South Carolina, Columbia, USA  
msaadat@email.sc.edu; sur@cse.sc.edu; srihari@sc.edu

Parmesh Ramanathan  
University of Wisconsin-Madison, USA  
parmesh.ramanathan@wisc.edu

**Abstract**—We present *MilliCam*, a system that captures the shape of small metallic objects, such as a gun, through obstructions, like clothing. *MilliCam* builds on the millimeter-wave (mmWave) imaging systems, which are widely used today in airport security checkpoints. Existing systems achieve high-resolution using a Synthetic Aperture Radar (SAR) principle, but require bulky motion controllers to position the mmWave device precisely. In contrast, *MilliCam* emulates the SAR principle by pure hand-swiping. However, alias-free, high-resolution imaging requires a linear, error-free hand-swiping motion. Furthermore, image focusing on an object of interest requires steering perfectly-shaped beam over the target-scene; but it is unavailable in off-the-shelf devices. We design a set of algorithms to enable high-quality handheld imaging: compensating for the errors in hand-swipe motion; and focusing the target-scene digitally without beam-steer. We have prototyped *MilliCam* on a 60 GHz testbed. Our experiments demonstrate that *MilliCam* can effectively combat motion errors and focus on the object in target-scene.

**Index Terms**—Millimeter-wave, SAR imaging, Handheld screening, Non-destructive testing, Focusing, Motion compensation.

## I. INTRODUCTION

Over the past few years, Radio-Frequency (RF) based imaging systems have made tremendous strides to see through a wall [1]–[3], track Non-Line-Of-Sight (NLOS) mobile targets [4]–[7], identify gestures [8], and differentiate humans through obstructions [3]. Millimeter-wave (mmWave) imaging systems represent the state-of-the-art in see-through-cloth screening and are widely used today in more than 250 airports worldwide [9], [10] to detect contrabands, such as weapons, explosives, and liquids [11]. These systems operate under the same principle as an airborne Synthetic Aperture Radar (SAR), albeit at much lower power, and within very close proximity. They illuminate the target using a mmWave beam; the signals then traverse through clothing, mostly absorbed by the human skin, but strongly reflect off the various concealed metallic or contraband objects. The systems use a precise mechanical movement to form a mesh surrounding the human body and capture the reflected signals [9]. A coherent combination of the signals allows the system to discriminate the target objects’ reflectivity against the human body. The non-ionizing radiation and small wavelength that can outline the shape of small concealed anomalies without impinging on privacy make mmWave an ideal medium for body scanning [10], [12].

In this paper, we imagine what if we can bring such imaging functionalities under ad-hoc settings in cheap, ubiquitous mobile mmWave devices? Driven by the recent commercialization effort of 5G cellular broadband [13], an unprecedented

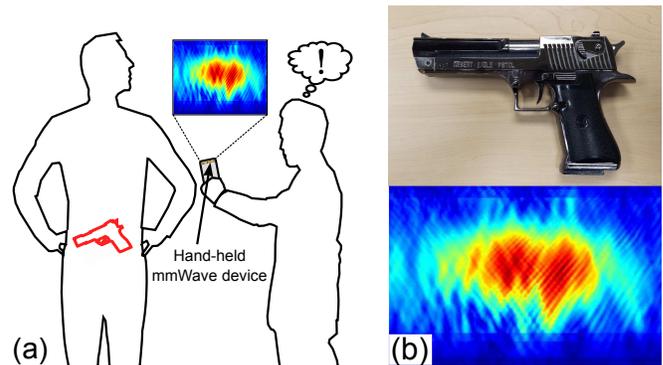


Fig. 1: (a) A potential application of *MilliCam*. User swipes a mmWave device in front of the target to obtain an image. (b) Optical image of a model gun (with a metal barrel and a plastic butt), and its mmWave image at 1 m. distance.

proliferation of the mmWave systems is imminent. Thus, the ad-hoc mmWave imaging may expand to a wide range of everyday settings, such as (1) In-situ security checking: Allowing law enforcement officers to conduct a stand-off screening of suspects, before proceeding to a pat-down search. (2) Baggage discrimination: Discriminating similar-looking baggage at the airport by detecting featured objects inside, e.g., a tablet, without opening the baggage, thus preserving user privacy. (3) Commercial loss prevention: Enabling non-intrusive counting of packaged inventory items in a factory or grocery store. (4) Non-destructive testing: Through-wall localization and diagnosis of powerlines or water pipes.

Realizing such a system in practice is challenging for three reasons. *First*, emulating the SAR principle requires swiping the hand-held mmWave device over the air to create a large SAR array — sufficient to discriminate small objects, like a gun, at a distance. Performance of a mmWave SAR imaging systems, however, depends highly on the stringent linear motion and millimeter-scale device positioning. *Second*, generating high-quality image depends on the focusing ability of the SAR system to remove background reflectivity. Airborne SAR and airport scanners achieve high-quality by steering a near-perfect shaped beam to collect focused reflection profile [14]–[16]. Standard-compliant, cheap, Commercial-Off-The-Shelf (COTS) mmWave devices cannot generate such beams [17]–[19]. *Third*, the SAR imaging algorithm relies on precise phase coherence among the reflected signals across the aperture points. It is well-known that the COTS mmWave devices do not provide coherent phase measurement [20], [21].

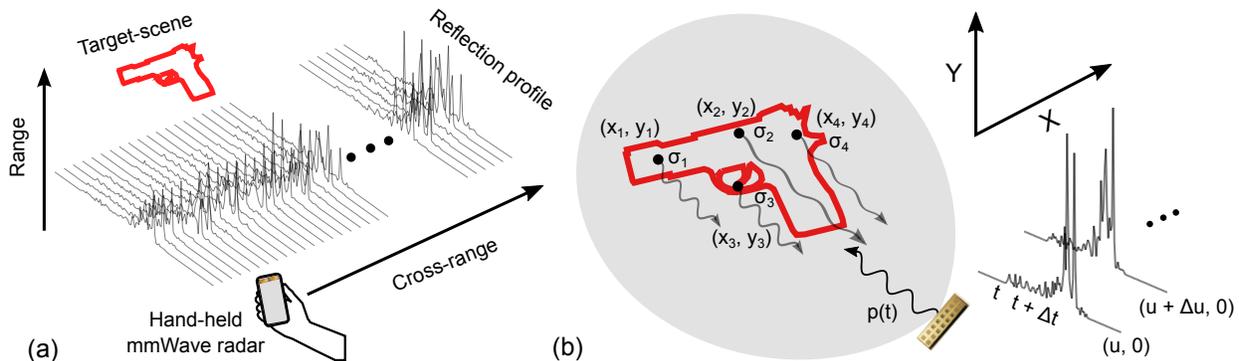


Fig. 2: (a) SAR range and cross-range imaging from a hand-held mmWave radar. (b) SAR reflection profile analysis.

We propose *MilliCam*, that aims to enable high-quality imaging under ad-hoc settings by overcoming the challenges in COTS devices. *MilliCam* relies purely on a hand-swiping motion to image a target-scene. *First*, it takes advantage of the co-located camera to obtain the coarse self-position of the device during hand-swipe. However, the device moves in a non-linear-trajectory, and the position measurement suffers from errors [22]. *MilliCam* overcomes this challenge by re-designing the existing SAR motion compensation algorithm [14] to operate at a meter-level distance. *Second*, *MilliCam* outputs a high-quality image by focusing the target-scene digitally onto a particular object. A key challenge is that the target location center is unknown *a priori*. To overcome this, *MilliCam* relies on the assumption that the target-scene is sparse; since most of the objects do not reflect back mmWave signals strongly, this assumption holds in practice. *MilliCam*, thus, iteratively focuses on the sparse clustered reflecting points to obtain a high-quality output image. *Finally*, *MilliCam* overcomes the phase incoherency by applying a transmitter-to-receiver Line-Of-Sight (LOS) path phase calibration.

We have prototyped *MilliCam* on a 60 GHz mobile mmWave testbed; the prototype is assembled from COTS components only. We have re-designed and implemented the chirp-based SAR imaging algorithm [14], [23]; our implementation works with the channel responses from the standard IEEE 802.11ad data communication packets. Our testbed experiments demonstrate that *MilliCam* can compensate for the motion error from hand-swipe, achieving  $16.5\times$  image quality improvement at 1 m. distance. At 1 m. from the target-scene, *MilliCam* improves the image quality  $7\times$  by focusing on the target object without prior knowledge of its center. Furthermore, *MilliCam* can discriminate between various shapes, such as square, circle, and rhombus — all hidden from LOS. It can measure the surface area accurately with less than 20% and 4% errors for  $5\times 5$  cm.<sup>2</sup> and  $20\times 20$  cm.<sup>2</sup> square-plates, respectively.

## II. RF IMAGING FUNDAMENTALS

### A. SAR Range and Cross-Range Imaging

Consider a stationary target-scene, such as in Figure 2, composed of a set of point reflectors with reflectivity  $\sigma_n$  at location

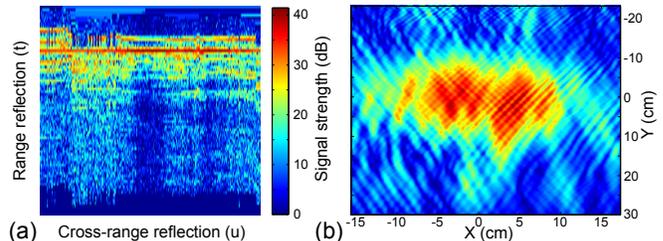


Fig. 3: (a) Measured SAR reflection in range and cross-range domain. (b) Output image.

$(x_n, y_n)$ . A mmWave radar located at coordinate  $(u, 0)$ , *i.e.*, the aperture point, illuminates the target-scene with a wide-beam and a wide-bandwidth signal  $p(t)$ . The signal then reflects off the target-scene, such as the barrel, butt, and trigger of the gun, to arrive at the radar. The measured echoed signal at time  $t$  and aperture  $u$  is [14]:  $s(t, u) = \sum_n \sigma_n \cdot p\left[t - \frac{2r_n}{c}\right]$ , where  $r_n (= \sqrt{(x_n - u)^2 + y_n^2})$  is the distance between the  $n^{\text{th}}$  reflecting point and the radar, and  $c (= 3 \times 10^8$  m/s) is the RF propagation speed. Each of the reflecting points reflects the same signal  $p(t)$  with a time-delay because of the different distance, and the wide-bandwidth signal can distinguish the reflections. A single aperture point, however, measures the target-scene reflectivity in the range of the radar and can image one-dimension only. For two-dimensional imaging, the radar is moved across the cross-range (Figure 2[a]), measuring the echoed signals at each of the aperture points. Figure 3(a) shows an example of the measured range and cross-range reflection signal  $s(t, u)$  from the gun (Figure 1) at 1 m. distance.

SAR imaging algorithm *maps* the reflection at the  $(t, u)$  domain to the spatial domain  $(x, y)$  to construct the image (Figure 3[b]). *First*, it applies two successive Fourier transforms over  $s(t, u)$  — one *w.r.t.* time  $t$ , and another *w.r.t.* space  $u$  — to recover the frequencies in the target-scene [14], [24],

$$\begin{aligned}
 s(\omega, u) &= \mathbf{FFT}_t(s(t, u)) = P(\omega) \sum_n \sigma_n \exp[-j2kr_n] \\
 s(\omega, k_u) &= \mathbf{FFT}_u(s(\omega, u)) \\
 &= P(\omega) \sum_n \sigma_n \exp[-jk_u x_n - j\sqrt{4k^2 - k_u^2} y_n]
 \end{aligned} \tag{1}$$

where  $\omega$  and  $k_u$  denote the temporal and spatial frequency

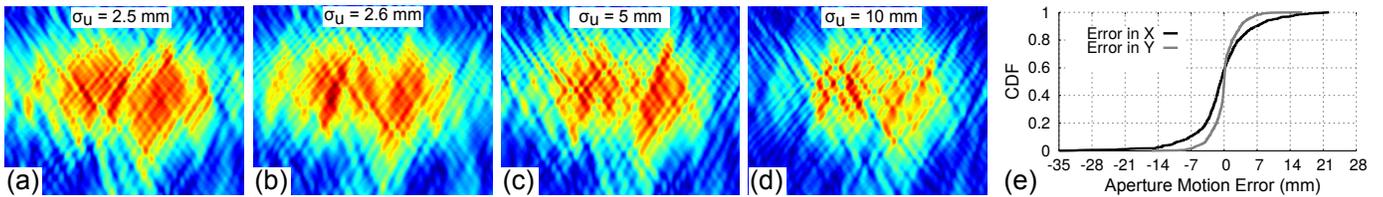


Fig. 4: (a–d) Effect of aperture localization error on SAR image. (e) Hand-swipe localization error distribution.

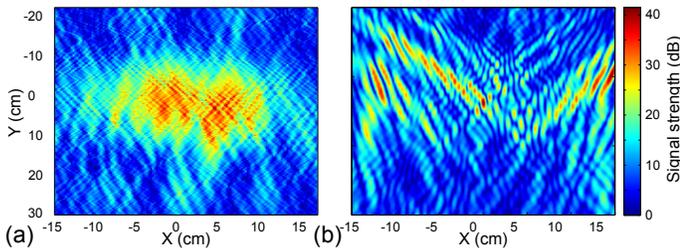


Fig. 5: (a) Defocused target-scene. (b) Effect of SAR phase measurement incoherency on the output image.

domain respectively [14], and  $k(= \omega/c = 2\pi/\lambda)$  is the wavenumber of RF signal. *Second*, it reconstructs the target reflectivity  $f(x, y)$  with a matched filter and a two-dimensional inverse Fourier transform [14],

$$\begin{aligned} s_m(\omega, k_u) &= P^*(\omega)s(\omega, k_u) \\ f(x, y) &= \text{IFFT}(s_m(\omega, k_u)) \end{aligned} \quad (2)$$

We also need spatial-frequency interpolation since the mapping from  $(\omega, k_u)$  to  $(x, y)$  is non-linear. We refer interested readers to [14] for more details.

### B. Key Factors Affecting Image Quality

**Aperture Length and Bandwidth:** The image resolution depends on the following two factors fundamentally. Across the cross-range, the resolution is limited by  $R\lambda/L_a$  [14], [25], where  $R$  is the mean distance between the aperture and the target-scene,  $\lambda$  is the wavelength of the RF signal, and  $L_a$  is the length of the aperture. At 60 GHz, a 50 cm. aperture length can achieve a 1 cm. cross-range resolution theoretically for a target-scene at 1 m. distance. The resolution across the range depends purely on the bandwidth  $B$  of the radar system and is limited by  $c/2B$  [14], [25]. Standard-compliant 60 GHz device, the operates on 1.76 GHz bandwidth [26], can achieve 8.5 cm. range resolution theoretically, sufficient to discriminate the butt from the barrel of a gun inside pocket (Figure 1[b]).

**Aperture Motion Error:** A correct Fourier transform in Equation (1), however, hinges upon the precise linear motion of aperture and the knowledge of its locations, *i.e.*,  $u$ . Intuitively, even a motion error above half-wavelength of the RF signal (*e.g.*, 2.5 mm. at 60 GHz) will distort the recovered frequencies in the target-scene and the final output image. Because, such error creates aliasing effect — adding signal intensities destructively where it should be constructive and vice-versa. To understand such effect, we simulate the aperture motion error with 2.5 mm., 2.6 mm., 5 mm., and 10 mm. standard deviation ( $\sigma_u$ ) under the scenario in Figure 3(a).

Figures 4(a–d) show how the shape of the gun gets distorted increasingly with large errors.

**Target-Scene Focusing:** Even with a perfect aperture motion, in practice, the quality of the output image depends on the focusing ability of the SAR system onto a target-scene. Said differently, without prior knowledge of the coordinates, the SAR system can image an arbitrary scene only and hope to recover the targets — leading to a poor image quality due to background reflections. To illustrate this point, we reconstruct the image in Figure 3(b) without prior knowledge of the target center. Figure 5(a) shows the defocused output result. Airborne radar solves this problem via steering a narrow-pencil beam electronically to collect the focused reflection profiles from the target-scene [14]. Cheap, COTS mmWave devices can not generate narrow beams that have multiple, strong spurious side-lobes [17], [27], [28].

**Coherent Measurement between Apertures:** SAR imaging algorithm also has an implicit assumption — the measurement between aperture points is phase-coherent. This is because the matched filtering and inverse Fourier transform in Equation (2) only work if the measured signals are phase-coherent across the aperture points. Unfortunately, COTS mmWave devices do not have such phase coherency [20], [21]. To illustrate phase-incoherency effect, we emulate a random phase-error with standard-deviation  $\pi$  on the measured SAR profile  $s(t, u)$  across  $u$ . Figure 5(b) shows the output image. The target object has vanished from the output image completely since the measured signal cannot be mapped to the spatial domain.

## III. MilliCam DESIGN

*MilliCam* aims to bring mmWave imaging functionality in cheap, ubiquitous mobile mmWave devices. It operates at short-range (1 ~ 1.5 m) and under ad-hoc settings, by transmitting mmWave signals from a standard-compliant, COTS hand-held mmWave device, measuring their reflections off the objects near-by, and processing these reflections to capture the shape of a small metallic object, such as a gun — hidden under obstructions like clothing. *MilliCam* emulates the existing SAR imaging principle in airborne radar and airport scanners, albeit at a much smaller-scale, and hand-moving the mmWave device purely in front of the target-scene.

Existing systems rely on the precise movement of the radar or high-end GPS to obtain aperture information for SAR imaging. *MilliCam*, in contrast, relies on the co-located optical camera in a mobile device to self-localize the mmWave device

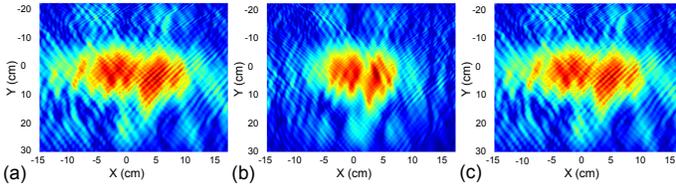


Fig. 6: Aperture motion error compensation: (a) Ground-truth image. (b) Distorted image with 5 mm. standard deviation of aperture motion error. (c) Motion compensated image.

and obtain the aperture information during hand-swipe. The challenge, however, lies in the non-linear motion trajectory and the unknown position estimation error from camera devices. In addition, to remove the background reflections and obtain a high-quality image, existing systems rely on steering a near-perfect beam over the target-scene. *MilliCam*, instead, re-designs a digital squint correction technique without relying on beam-steering or known location of the target's center. Furthermore, *MilliCam* system is designed to work with COTS mmWave devices that do not provide phase-coherence. Next, we describe these design components in detail.

#### A. Self-Localizing mmWave Device

To estimate the position of the mmWave device during hand-swipe, *MilliCam* relies on a co-located optical camera. A typical smartphone camera, using multi-million image sensors, can achieve sub-mm. optical pixel resolution. By measuring the translation and rotation of a reference pattern from an optical anchor image, *MilliCam* computes the position and trajectory of the mmWave device during hand-swipe. The reference pattern could be a known pattern on the user's body, such as the pattern on his shirt or t-shirt. Assume that the camera has captured a video sequence of a reference scene as the user swipes the device over the air. The camera projects a reference point in the scene  $\mathbf{p} = \{x, y, z\}$  onto the captured pixel of a video frame  $f$  as  $\mathbf{r}_f = \{u, v\}$ ; the projection follows a well-known geometrical model [29]:  $\mathbf{r}_f = \lambda[\mathbf{R}_f \ \mathbf{t}_f]\mathbf{p}$ , where  $\lambda$  is an *intrinsic matrix* containing the camera's factory-fixed parameters.  $\mathbf{R}_f$  and  $\mathbf{t}_f$  are the relative rotational and translational movement of the device at image frame  $f$  w.r.t. the first frame. Now, suppose the anchor image contains a set of  $N$  points —  $(\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_N)$ . The device's rotation and translation at frame  $f$  is then given by [29]:  $(\mathbf{R}_f^*, \mathbf{t}_f^*) = \underset{\mathbf{R}_f, \mathbf{t}_f}{\operatorname{argmin}} \sum_{i=1}^N \|\lambda[\mathbf{R}_f \ \mathbf{t}_f]\mathbf{p}_i - \mathbf{r}\|^2$ . The camera's relative location at frame  $f$  equals the joint effect of rotation and translation movement:  $-(\mathbf{R}_f^*)^T \mathbf{t}_f^*$ . However, our measurement shows that the localization error can reach 7 mm. within just two standard deviations (Figure 4[e]). Such error can severely distort the image (Figure 4[a–d]), and in the next section, we discuss an approach to overcome the error.

#### B. Short-distance Motion Error Compensation

The aperture motion error in *MilliCam* is reminiscent of the motion error in airborne SAR. Airborne radar uses high-end GPS system to obtain the aperture location data and

compensate for the non-linear motion error in the flight-path. Yet, the constraint under *MilliCam* is stringent — the target-scene is very close-by, and the error statistics (cm-level) is significant compared to the aperture length. In what follows, we first introduce how the airborne SAR compensates the non-linear motion error, and then re-design it for *MilliCam*.

Assume that the non-linear motion error at the aperture point  $u$  is  $(x_e(u), y_e(u))$ . The measured SAR profile of an ideal point reflector with motion error can be derived as [14],

$$\begin{aligned} s_n(\omega, u) &= \exp \left[ -j2k \sqrt{(x_n - u - x_e(u))^2 + (y_n - y_e(u))^2} \right] \\ &= \exp[j2kr_{en}(u)] \cdot \exp[-j2kr_n] \end{aligned} \quad (3)$$

where  $r_{en}(u) = -\sqrt{(x_n - u - x_e(u))^2 + (y_n - y_e(u))^2} + r_n$ , is the distance error for the  $n^{\text{th}}$  reflector. The key idea is to compensate for this distance error in the spatial frequency domain. Typically, the span in the range  $y_n$  is much larger than the cross-range  $x_n$ , aperture length, and the motion errors since the target-scene (ground) is far away from the flight trajectory (air). Thus, the distance error can be approximated as,  $r_{en}(u) \approx x_e(u)$ , and the compensation is given by [14],

$$\begin{aligned} s'(\omega, u) &= s(\omega, u) \cdot \exp[-j2kr_{en}(u)] \\ &= s(\omega, u) \cdot \exp[-j2kx_e(u)] \end{aligned} \quad (4)$$

This compensated data can then be directly used with Equation (2) to generate the output image.

Unfortunately, such approximation breaks down under *MilliCam* — target-scene is very close-by, range and cross-range have a similar span, and error is significant. *MilliCam*, however, leverages a key observation to re-design the above compensation — the target-scene in mmWave image is mostly sparse, since most objects do not reflect back the mmWave signals [30]–[32]. Said differently, the error in the spatial frequency is mostly dominated by the sparsely clustered strong reflecting points in the target-scene and nearby reflecting points tend to produce similar errors. As long as *MilliCam* can identify and compensate the spatial frequency for the clustered points, it can recover the motion compensated image.

To approach this ideal, *MilliCam* first constructs an image from the SAR profile without motion compensation and identify the clustered points. The image will appear distorted, nonetheless, clustered around specific coordinates. *MilliCam*, then finds out the centroid  $(x_c, y_c)$  of the clusters [33] and the corresponding SAR profile from Equation (3). The error in spatial frequency is then compensated via Equation (4) to generate a new image. *MilliCam*, then iterates over the above process to refine the centroid and the corresponding output image. Still, the residual, unknown motion estimation error can be significant and can distort the final image. *MilliCam*, thus, leverages a bicubic interpolation over the estimated trajectory and then uses the non-linear motion compensation to output an undistorted image. Figure 6 shows a visual example of the correction under 5 mm. standard deviation of aperture motion.

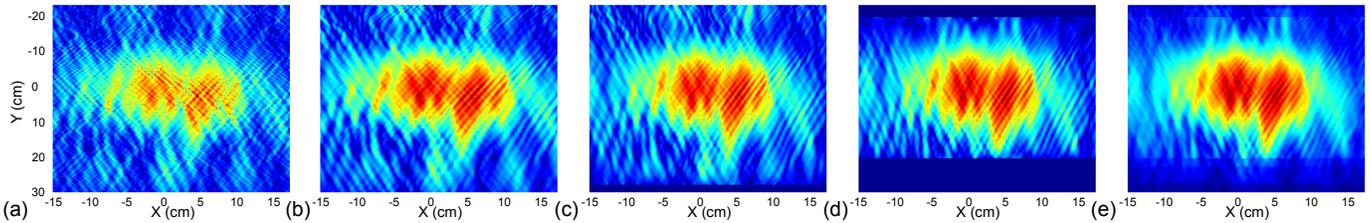


Fig. 7: (a) Defocused target-scene. (b) Squint correction with known target center. (c-d) Intermediate mmWave images from iterative squint correction. (e) Final output image.

### C. Target Focusing via Squint Correction

However, motion compensation alone does not ensure good quality image. Reflections from background and heterogenous target-reflectivity may defocus the image. Airborne SAR systems solve the problem either via electronically steering a narrow, pencil-beam to collect focused reflection profile from a specific area in the ground [14], [34] or by applying a digital squint correction [14], [35]. Since the COTS mmWave devices do not allow to generate a narrow, perfectly-shaped beam [17], [27], [28], we rely on the digital squint correction.

Squint correction, at a very high level, tries to focus the measured SAR profile  $s(t, u)$  at a specific coordinate of the target-scene before generating the output image. Consider the center of the target-scene is at  $(X_0, Y_0)$ . The temporal and spatial frequency response of an ideal point reflector at  $(X_0, Y_0)$  can be derived from Equation (1):  $s_0(\omega, k_u) = P(\omega) \cdot \exp[-jk_u X_0 - j\sqrt{4k^2 - k_u^2} Y_0]$ . Digital squint correction simply shifts the frequency response of the measured SAR profile, *i.e.*,  $s(\omega, k_u)$  towards the center frequency response  $s_0(\omega, k_u)$ . This is equivalent to  $s_c(\omega, k_u) = s(\omega, k_u) \cdot s_0(\omega, k_u)^*$ . The frequency shifted profile can then be used in the Equation (2) directly to generate a focused SAR image. Figure 7(a-b) show the focusing result when  $(X_0, Y_0)$  is known.

In practice, the center location information is unavailable to the SAR imaging system. Sequentially searching for the center of the target over a large scene is not only computationally expensive but also does not ensure appropriate focusing on a complex scene with multiple targets. *MilliCam* overcomes this challenge by reusing the sparse target-scene approximation in the mmWave image (Section III-B). When the target-scene is sparse, intuitively, its defocused version should also remain sparse. *MilliCam*, thus, can segment the defocused target-scene and then iteratively apply the correction on the segmented images, and finally, generate a composite focused image.

Algorithm 1 formalizes the intuition. It first synthesizes the output image  $f_0$  from the SAR profile  $s(t, u)$  with a random target-center location  $(X_0, Y_0)$ , yielding a defocused image. The defocused image is then segmented via global thresholding [36], and *MilliCam* finds a list of the possible targets' center location. *MilliCam*, then applies the squint correction with the center locations on  $s(t, u)$  for each of these targets, zooming in only on the X-Y dimensions around them. After each iteration, *MilliCam* outputs a smoother quality image

---

### Algorithm 1 Iterative Squint Correction

---

- 1: Input: SAR profile  $s(t, u)$ ; Output: Focused image  $F_I$ ;
  - 2: Parameters: Image PSNR change,  $\Delta_I = 3$  dB;
  - 3: Synthesize the SAR output image  $f_0$  from  $s(t, u)$ ;
  - 4:  $i \leftarrow 1$ ;  $f_i \leftarrow f_0$ ;  $f_{i+1} \leftarrow 0$ ;
  - 5: **while**  $|\text{PSNR}(f_{i-1}, f_i) - \text{PSNR}(f_i, f_{i+1})| \geq \Delta_I$
  - 6:   Segment image  $f_{i-1}$  using Otsu's method [36];
  - 7:   List of segmented images:  $f'[1, 2, \dots, N]$ ;
  - 8:   Find strongest pixel coordinate  $(X_j, Y_j)$  in  $f'[j]$ ;
  - 9:   Correct squint in  $f'[j]$  using  $(X_j, Y_j)$  to generate  $f'_c[j]$ ;
  - 10:   Generate  $f_i$  by combining all  $f'_c[j]$ ;
  - 11:    $f_{i+1} \leftarrow \text{alpha composite}(f_{i-1}, f_i)$ ;  $i \leftarrow i + 1$ ;
  - 12: **end while**
  - 13: Final focused composite image  $F_I \leftarrow f_i$ ;
- 

using alpha compositing [37] current image with the previous iteration and keeps track of the image PSNR improvement. The iteration stops when the PSNR improvement is small, less than 3 dB. Figures 7(c-d) show the two intermediate images from Algorithm 1 for the defocused target-scene in Figure 7(a), and Figure 7(e) shows the final image. Figure 8 further shows the focusing results with different shaped objects.

### D. Measurement Incoherency in COTS mmWave Devices

COTS mmWave devices do not provide any carrier phase-coherency across the aperture points [20], [21]. Since they use cheap high-frequency oscillators, tracking the carrier phase across the multiple aperture points is virtually infeasible. *MilliCam* leverages a simple observation to workaround this problem. Even if the phases of the reflected signals across the apertures are incoherent, the relative phases of the channel within a single aperture point are still coherent, as it purely depends on the length differences between the reflecting points, and can be calibrated *w.r.t.* to the line-of-sight signal from Tx to Rx. Mathematically, the phase-calibrated SAR profile is given by:  $s_p(t, u) = s(t, u) \cdot \exp(-j\angle s_T^m)$  where  $s_T^m = \max|s_T|$ . Said differently, *MilliCam* first leverages the strongest Tx to Rx self-interference signal to calibrate the incoherent phases, and then cancels the self-interference.

## IV. IMPLEMENTATION

We implement and evaluate *MilliCam* on a 60 GHz mobile mmWave testbed (Figure 9). We build our own hand-held 60 GHz mmWave prototype by assembling the COTS components only. The prototype board consists of a back-end processor

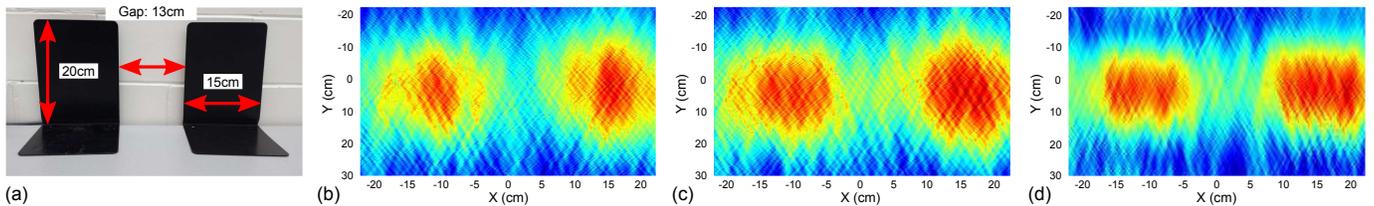


Fig. 8: (a) Two metallic square-shaped objects. (b-c) Intermediate images from Algorithm 1. (d) Final output image.

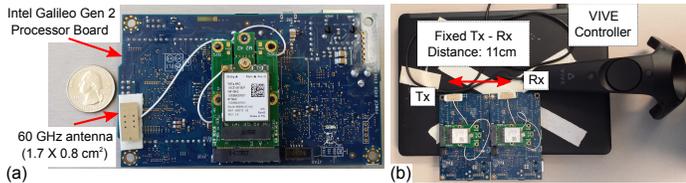


Fig. 9: (a) 60 GHz mmWave prototype board assembled from COTS components. (b) Baseline SAR imaging system that uses a laser motion tracker [42], [43].

board with Intel Galileo Gen 2 [38] and a front-end 60 GHz chipset from Qualcomm [39]. The processor board consists of a 400 MHz Intel Quark X1000 CPU, 256 MB DDR3 memory, and a mini PCIe slot. The 60 GHz chipset consists of a 32-element phased-array antenna and a MAC/baseband chip hooked to the processor board via the mini PCIe slot. The phased-array antenna can generate up to 64 transmit and receive beams, but throughout our experiments, we fix onto a single widest-beam direction. The chipset is IEEE 802.11ad-compatible and operates on a 1.76 GHz wide-bandwidth on the unlicensed 60 GHz. The device supports extracting the 1.76 GHz channel response from the standard “Channel Estimation” (CE) header field of the data packets [26]. The CE field spans only 650 ns. as per the IEEE 802.11ad standard, and thus, the speed of human hand-swipe does not affect the collected channel response in one SAR aperture position. Nonetheless, we still need to collect the channel responses in real-time across the entire hand-swipe duration. Thus, we bootstrap our Galileo board with the open-source Linux kernel [40], build support for the latest 60 GHz kernel driver and firmware [41], and implement suitable kernel daemons to trigger and collect the channel response at 5 ms. interval.

Since the COTS communication devices do not allow switching between Tx and Rx mode within nanoseconds, we mount two prototype boards, as Tx and Rx, on a laptop (Figure 9(b)). The boards are separated by 11 cm., and we calibrate the phases of the measured channel following Section III-D. To evaluate the performance of an imaging system with an accurate SAR antenna aperture information, we also use HTC VIVE controller and base stations [42], [43]. The HTC devices can accurately measure the X-Y movement with a millimeter-scale precision [44], and we implemented software support in the laptop to collect and synchronize the hand-swipe trajectory locations and measured channel responses in real-time.

Finally, based on the location and channel response traces, we evaluate *MilliCam*’s design components in MATLAB on a PC

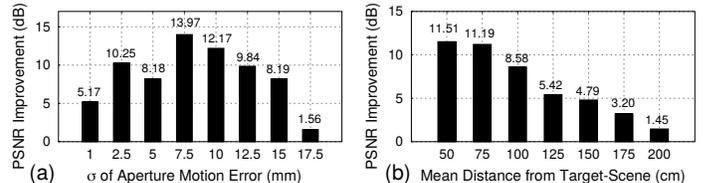


Fig. 10: *MilliCam* average PSNR improvement: (a) Under aperture motion errors. (b) At different distances.

host with Core-i5 CPU with 8 GB memory. We re-designed and implemented the chirp-based SAR imaging algorithm [23] to operate with the standard-compliant channel responses of COTS mmWave devices. Our implementation can generate one image frame within 1.25 seconds. Although the design components are evaluated on a customized platform, *MilliCam* itself does not require any specialized hardware beyond commodity 60 GHz mmWave radio’s capabilities.

## V. EXPERIMENTAL RESULTS

### A. Micro-benchmarks

**Error Correction Results:** We first measure *MilliCam*’s effectiveness in correcting aperture motion error quantitatively. We use one of the square-shaped metallic objects in Figure 8(a) and place it in the target-scene with a background concrete-wall at 1 m. distance. We use a mechanical slider [45] to linearly move our imaging setup in front of the target-scene and collect the SAR reflection profile. We then follow Section II-A to generate an output image, which serves as the ground-truth. To evaluate *MilliCam*’s imaging performance under different aperture motion error, we intentionally mark the SAR profile at aperture  $u$  as aperture  $u + \delta u$  instead. The error  $\delta u$  follows the same normal distribution as the camera calibration error we have measured in Section III-A, but with different standard deviations. We then perform 50 such experiments.

Figure 10(a) shows the average PSNR improvement under different standard deviations of motion error in X-direction. At 10 mm. error standard deviation, the average PSNR improvement is 12.17 dB, *i.e.*, 16.5 $\times$  over the image without error compensation. *MilliCam* achieves the peak improvement at 7.5 mm. error standard deviation with a 13.97 dB PSNR boost. Figure 6 also shows the visual quality improvement in *MilliCam* under 5 mm. error standard deviation. The improvement below 2.5 mm. of error standard deviation is not significant, only 3 $\times$ . Still, *MilliCam* achieves 8.19 dB PSNR improvement even with 15 mm. error standard deviation at 1 m. target distance.

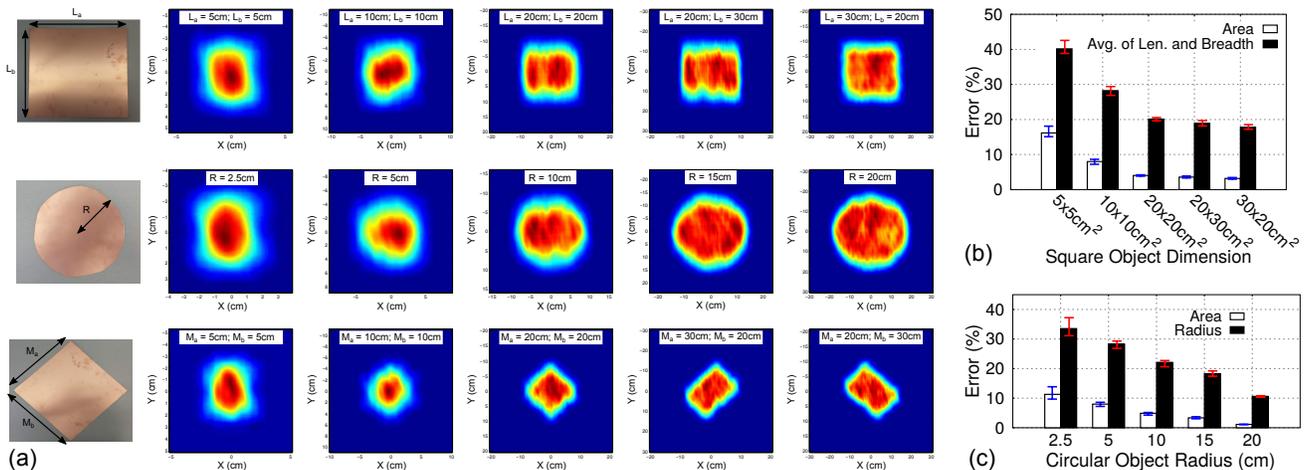


Fig. 11: (a) Square, circle, and rhombus shapes with various dimensions imaged by *MilliCam* at 1 m. (b) Area and average of length, breadth error of the square object. (c) Area and radius error of the circular object. Error bar shows max and min.

**Target Focusing Effectiveness:** Next, we evaluate *MilliCam*'s performance of focusing on the target-scene. We place the same square object onto the target-scene and vary the distance from 50 cm. to 2 m. between the target-scene and the imaging setup. The target-scene contains only the square object and the background concrete wall (Figure 8[a]). At each distance, we manually estimate the objects' center and create a ground-truth image. Then, we employ *MilliCam*'s iterative squint correction without the knowledge of the objects' center. Finally, we measure the image PSNR improvement with and without focusing. Figure 10(b) shows the result. At 50 cm. distance from the target-scene, *MilliCam*'s average PSNR improvement over defocused target-scene is 11.51 dB, *i.e.*, a  $14\times$  increase in objective scale. The improvement, however, drops significantly as we move further away from the target-scene — only  $\sim 1.4\times$  at 2 m. This is expected since the image resolution of the ground-truth itself increasingly becomes poor as the distance between target-scene and the imaging setup increases (Section II-B). Nonetheless, the image quality improvement is still significant even at 1 m. distance — more than  $7\times$ .

**Shape Detection and Errors:** We now evaluate *MilliCam*'s ability to differentiate object shapes. We use flat copper plates with 0.5 mm. thickness and cut it into three shapes — square, circle, and rhombus (Figures 11[a]). The dimension of the square and rhombus varies from  $5\times 5$  cm<sup>2</sup> to  $30\times 20$  cm<sup>2</sup>. The radius of the circular shapes varies between 2.5 cm. to 20 cm. Figures 11(a) show the resultant images for each of the shapes and dimensions. For larger dimensions (*e.g.*, the last column in Figures 11[a]), *MilliCam* can distinguish between various shapes clearly and identifies its dimensions correctly.

To quantitatively measure the dimension and area error, we generate 50 images for the square and circular objects, for each dimension. Figures 11(b–c) show the results. A smaller dimension object has a larger area and dimension error. Figure 11(b) shows the error in percentage for the square-shaped copper plates, where we calculate the error as the ratio of pixel in- and outside of the ground-truth dimensions. Figure

11(c) further shows the error for the circular object. Typically this error is less compared to the square object, since the two-dimensional point-spread function of the imaging system is circular, creating less smearing effect around the edges.

### B. Imaging A Hidden Gun

We now evaluate *MilliCam* under a toy setup, where the gun is hidden behind clothing. We mounted the model, plastic gun on the background concrete wall (Figure 12[a]), and then hide it with a piece of clothing (Figure 12[b]). Then, we run *MilliCam* at 1 m. distance and generate the output image.

Figure 12(c) shows the composite of the optical and the mmWave image. While the dimensions of the gun are preserved, the butt and the barrel can only be faintly detected. This is because mmWave signals have to penetrate the clothing twice: one during transmission and another during reflection. Thus, the measured SAR profile has a much larger noise floor, resulting in poorer image quality. We believe, a fine-grained transmit and receive power control, similar to the adaptive camera exposure control and correction [46], may improve the image quality. Such power control is currently unavailable on our platform, and we plan to investigate it in the future.

## VI. LIMITATIONS

**Unknown Depth of the Target:** *MilliCam* relies on the assumption that the target object's distance, *i.e.*, the depth, from the aperture plane is known. Throughout our experiments, we have used manual depth estimation to reconstruct images. In practice, however, the object may be hidden from the LOS, and the depth estimation will be unavailable, even when using out-of-band information, like camera image. A strawman approach could be signal time-of-arrival based depth estimation; yet, the time-of-arrival from the same reflection point varies with the aperture points. Moreover, multiple reflections from the same object may corrupt the depth estimation. We propose to investigate it in our future work.



Fig. 12: (a) Imaging a plastic gun. (b) Gun is hidden inside a piece of cloth. (c) Composite of mmWave and the optical image.

**Multiple Overlapping Targets at Different Depths:** The above problem is exacerbated when there are multiple overlapping objects at different depths. Even using a very wide-bandwidth signal, which can separate the signal time-of-arrival from nearby reflection points, it may not be feasible to distinguish whether the reflections are from the same object or different objects. We propose to leverage two-dimensional handheld movements, together with continuous object plane assumption, like [47], to jointly estimate multiple target depths and reconstruct images.

**Blockage on Target:** *MilliCam* relies on the reflection from an object to identify its shape. An obstructing metallic object can indeed hide the relevant object from *MilliCam*. In such cases, a multi-spectral analysis may help to identify the target in the hidden layer. A similar method is used today in airports to detect contraband items in checked-in baggage, where the target may be hidden below multiple layers of items. We leave such an extension for our future work.

**Imaging Mobile Targets:** *MilliCam* also relies on the assumption that the target-scene remains stationary during the hand-swipe movement. This is reasonable — typical hand-swipe speed can vary from 60-100 cm/s, and thus, minute sway motion from the human body may not affect *MilliCam*'s performance. *MilliCam*, however, cannot image a mobile target. We note that SAR imaging algorithms for moving target detection and imaging already exist and are used in the airborne systems [14]. Yet, re-using the algorithms under ad-hoc settings, for a close-by target with COTS mmWave device, is non-trivial, and we plan to investigate it in our future work.

## VII. RELATED WORK

**Conventional Radio Imaging:** Traditional radio imaging systems use multiple antennas (or virtual antennas via motion) to “scan” the target. By coherently combining reflected signals, it can recover pixel intensity in the target scene. This principle underpins the MIMO [48]–[50] or Synthetic Aperture Radar (SAR) imaging [14], [51]–[53]. Prior mobile imaging system attempted to emulate the airborne radar on ground-vehicles [54]; but, the form factor is too large for human portability. Recently proposed hand-held radar [55], [56] uses specialized hardware, such as a VNA, for imaging, yet they rely on bulky mechanical supports to facilitate precise movement. Furthermore, they require special waveforms [47], [56], which is only available on the dedicated high-end hardware. Zhu *et al.* [57], [58] recently proposed to use highly directional 60 GHz horn antennas and phased-array to detect object curvatures and boundaries. It is infeasible, however, to image

small heterogeneous-reflectivity objects using only received signal strength (RSS) alone. The radar in [58], in fact, has to travel 10s of meters to create a one-dimensional image of cm-scale objects. Tomographic imaging [59] purely relies on RSS, but it requires many radios around the target, and the resolution is inferior to phase-based approaches [3], [14]. *MilliCam*, in contrast, works with today's cheap, ubiquitous, and standard-compliant hand-held mmWave devices.

**Motion Compensation in Imaging:** The image distortion and defocusing caused by motion and squint error in *MilliCam* is reminiscent of the motion blur in optical cameras. Optical image blur due to motion, either from the target or the camera, can be compensated by many deblurring algorithms or by using out-of-band sensors like gyroscopes [60]–[62]. This is because multi-million image sensors on the camera do not change their relative position under any movement. In our ad-hoc image setting, however, the “sensors” are created through virtually hand-moving the antenna and thus, the existing deblurring algorithm does not work.

**Object Detection through non-RF Modality:** Infrared and thermal cameras can image the human body but are typically unsuitable for through-obstruction imaging. It is, in fact, impossible to differentiate concealed items using a thermal camera as the object's temperature approaches that of the body [63]. Jacor Inc. developed a hand-held *ultrasonic* device to detect concealed weapons [64]. The device can discriminate hard objects based on its stronger reflection intensity, but cannot estimate the shape or dimension.

## VIII. CONCLUSION

This paper proposes *MilliCam*, a system that leverages mmWave reflected signals to construct images of small metallic objects. Different from existing mmWave imaging systems, *MilliCam* leverages a single hand-swipe motion. *MilliCam* outputs a two-dimensional image by coherently combining the reflected wide-bandwidth mmWave signals across the hand-swipe. The non-linear motion due to hand-swiping and aperture localization errors, however, can distort the resultant image. *MilliCam* re-designs the motion compensation algorithm in airborne SAR to work at the meter-level range and produce undistorted images. Furthermore, *MilliCam* can digitally re-focus the target-scene onto an object of interest without its prior location information to output a high-quality image. We believe that *MilliCam* marks an important step towards enabling handheld mmWave imaging with cheap, ubiquitous mobile mmWave devices.

## REFERENCES

- [1] "See Through Wall Radar Imaging Technology," 2015. [Online]. Available: <http://fredecomposition.wordpress.com/technology/>
- [2] F. Adib and D. Katabi, "See Through Walls with WiFi!" in *Proc. of ACM SIGCOMM*, 2013.
- [3] F. Adib, C.-Y. Hsu, H. Mao, D. Katabi, and F. Durand, "Capturing the Human Figure Through a Wall," in *Proc. of ACM SIGGRAPH*, 2015.
- [4] F. Adib, Z. Kabelac, D. Katabi, and R. C. Miller, "3D Tracking via Body Radio Reflections," in *Proc. of USENIX NSDI*, 2014.
- [5] S. Nannuru, Y. Li, Y. Zeng, M. Coates, and B. Yang, "Radio-Frequency Tomography for Passive Indoor Multitarget Tracking," in *IEEE Trans. on Mobile Computing*, 2013.
- [6] J. Xiong and K. Jamieson, "ArrayTrack: A Fine-Grained Indoor Location System," in *USENIX NSDI*, 2013.
- [7] Thales Visionix Inc., "IS-900," 2014. [Online]. Available: <http://www.intersense.com/pages/20/14>
- [8] Q. Pu, S. Gupta, S. Gollakota, and S. Patel, "Whole-home Gesture Recognition using Wireless Signals," in *ACM MobiCom*, 2013.
- [9] "ProVision Automatic Target Detection," 2015. [Online]. Available: <http://www.sds.l-3com.com/advancedimaging/provision-at.htm>
- [10] Transportation Security Administration Press Office., "TSA takes next steps to further enhance passenger privacy," July, 2011. [Online]. Available: [shorturl.at/pCPS1](http://shorturl.at/pCPS1)
- [11] Transportation Security Administration, "What Can I Bring?" 2017. [Online]. Available: [shorturl.at/rzB15](http://shorturl.at/rzB15)
- [12] D. M. Sheen, D. L. McMakin, and T. E. Hall, "Three-Dimensional Millimeter-Wave Imaging for Concealed Weapon Detection," *IEEE Trans. on Microwave Theory and Techniques*, vol. 49, no. 9, 2001.
- [13] 3GPP: A Global Initiative, "The Mobile Broadband Standard: Release 15," 2017. [Online]. Available: <http://www.3gpp.org/release-15>
- [14] M. Soumekh, *Synthetic Aperture Radar Signal Processing*. John Wiley & Sons, Inc., 1999.
- [15] C. Y. Chang, M. Jin, and J. C. Curlander, "Squint Mode SAR Processing Algorithms," in *IEEE International Geoscience and Remote Sensing Symposium*, 1989.
- [16] S. Patole and M. Torlak, "Two Dimensional Array Imaging with Beam Steered Data," *IEEE Trans. on Image Processing*, vol. 22, no. 12, 2013.
- [17] S. Sur, I. Pefkianakis, X. Zhang, and K.-H. Kim, "WiFi-Assisted 60 GHz Wireless Networks," in *Proc. of ACM MobiCom*, 2017.
- [18] —, "Towards Scalable and Ubiquitous Millimeter-Wave Wireless Networks," in *Proc. of ACM MobiCom*, 2018.
- [19] D. Steinmetzer, D. Wegemer, M. Schulz, J. Widmer, and M. Hollick, "Compressive Millimeter-Wave Sector Selection in Off-the-Shelf IEEE 802.11ad Devices," in *Proc. of ACM CoNEXT*, 2017.
- [20] Facebook, Inc., "Introducing Facebook's new terrestrial connectivity systems — Terragraph and Project ARIES," [shorturl.at/bf1yU](http://shorturl.at/bf1yU), 2016.
- [21] M. E. Rasekh, Z. Marzi, Y. Zhu, U. Madhoo, and H. Zheng, "Noncoherent mmWave Path Tracking," in *ACM HotMobile*, 2017.
- [22] D. Ishizuka, A. Yamashita, R. Kawanishi, T. Kaneko, and H. Asama, "Self-localization of Mobile Robot Equipped with Omnidirectional Camera Using Image Matching and 3D-2D Edge Matching," in *IEEE International Conference on Computer Vision*, 2011.
- [23] Mehrdad Soumekh, "Synthetic Aperture Radar Signal Processing with MATLAB Algorithms," 2016. [Online]. Available: [goo.gl/p9UDUH](http://goo.gl/p9UDUH)
- [24] M. Soumekh, "A System Model and Inversion for Synthetic Aperture Radar Imaging," *IEEE Trans. on Image Processing*, vol. 1, no. 1, 1992.
- [25] M. Skolnik, *Introduction to Radar Systems*. McGraw-Hill, 1962.
- [26] IEEE Standards Association, "IEEE Standards 802.11ad-2012, Amendment 3: Enhancements for Very High Throughput in the 60 GHz Band," [goo.gl/r2JeYd](http://goo.gl/r2JeYd), 2012.
- [27] T. Nitsche, G. Bielsa, I. Tejado, A. Loch, and J. Widmer, "Boon and Bane of 60 GHz Networks: Practical Insights into Beamforming, Interference, and Frame Level Operation," in *ACM CoNEXT*, 2015.
- [28] D. Steinmetzer, D. Wegemer, M. Schulz, J. Widmer, and M. Hollick, "Compressive Millimeter-Wave Sector Selection in Off-the-Shelf IEEE 802.11ad Devices," in *Proc. of ACM CoNEXT*, 2017.
- [29] S. J. Prince, *Computer Vision: Models, Learning, and Inference*. Cambridge University Press, 2012.
- [30] S. Sur, V. Venkateswaran, X. Zhang, and P. Ramanathan, "60 GHz Indoor Networking through Flexible Beams: A Link-Level Profiling," in *Proc. of ACM SIGMETRICS*, 2015.
- [31] S. Sur, X. Zhang, P. Ramanathan, and R. Chandra, "BeamSpy: Enabling Robust 60 GHz Links Under Blockage," in *USENIX NSDI*, 2016.
- [32] C. R. Anderson, T. S. Rappaport, K. Bae, A. Verstak, N. Ramakrishnan, W. H. Tranter, C. A. Shaffer, and L. T. Watson, "In-Building Wideband Multipath Characteristics at 2.5 & 60 GHz," in *IEEE VTC*, 2002.
- [33] M. Tuceryan, "Moment Based Texture Segmentation," *International Conference on Pattern Recognition*, vol. 3, 1992.
- [34] A. J. Fenn, D. H. Temme, W. P. Delaney, and W. E. Courtney, "The Development of Phased-Array Radar Technology," *Lincoln Laboratory Journal*, vol. 12, no. 2, 2000.
- [35] G. W. Davidson and I. Cumming, "Signal Properties of Spaceborne Squint-Mode SAR," *IEEE Trans. on Geoscience and Remote Sensing*, vol. 35, no. 3, 1997.
- [36] N. Otsu, "A Threshold Selection Method from Gray-Level Histograms," *IEEE Trans. on Systems, Man, and Cybernetics*, vol. 9, no. 1, 1979.
- [37] R. Szeliski, *Computer Vision: Alg. and App.* Springer, 2010.
- [38] Intel Corporation, "Intel Galileo Gen 2 Board," 2014. [Online]. Available: [shorturl.at/nvwQZ](http://shorturl.at/nvwQZ)
- [39] "Qualcomm Atheros," [Online]. Available: [shorturl.at/auvCJ](http://shorturl.at/auvCJ)
- [40] "Getting Started with the Intel Galileo Board on Linux," 2015. [Online]. Available: <https://software.intel.com/en-us/get-started-galileo-linux>
- [41] "Atheros Linux wireless drivers — wil6210," 2015. [Online]. Available: <https://wireless.wiki.kernel.org/en/users/drivers/wil6210>
- [42] HTC, "VIVE Controller," 2017. [Online]. Available: [shorturl.at/jsY45](http://shorturl.at/jsY45)
- [43] —, "Base Station," 2017. [Online]. Available: [shorturl.at/amMR0](http://shorturl.at/amMR0)
- [44] Road To VR, "Analysis of Valve's 'Lighthouse' Tracking System Reveals Accuracy," 2016. [Online]. Available: [shorturl.at/pEj02](http://shorturl.at/pEj02)
- [45] "Lynx 3 Axis Slider," [Online]. Available: [shorturl.at/mX147](http://shorturl.at/mX147)
- [46] G. Messina, A. Castorina, S. Battiato, and A. Bosco, "Image Quality Improvement by Adaptive Exposure Correction Techniques," in *International Conference on Multimedia and Expo*, 2003.
- [47] B. Mamandipoor, G. Malysa, A. Arbabian, U. Madhoo, and K. Noujeim, "60 GHz Synthetic Aperture Radar for Short-Range Imaging: Theory and Experiments," in *IEEE Aslrm. Conf. on Sig., Sys. and Comp.*, 2014.
- [48] E. Fishler, A. Haimovich, R. Blum, D. Chizhik, L. Cimini, and R. Valenzuela, "MIMO Radar: an Idea Whose Time Has Come," in *Proc. of IEEE Radar Conference*, 2004.
- [49] D. Bliss and K. Forsythe, "Multiple-Input Multiple-Output (MIMO) Radar and Imaging: Degrees of Freedom and Resolution," in *IEEE Aslrm. Conf. on Sig., Sys. and Comp.*, 2003.
- [50] D.-W. Wang, X.-Y. Ma, and Y. Su, "Two-Dimensional Imaging via a Narrowband MIMO Radar System With Two Perpendicular Linear Arrays," *IEEE Trans. on Image Processing*, vol. 19, no. 5, 2010.
- [51] M. Cheney and B. Borden, *Fundamentals of Radar Imaging*. Society for Industrial and Applied Mathematics, 2009.
- [52] H. Griffiths and C. Baker, *Fundamentals of Tomography and Radar*. Springer Netherlands, 2006.
- [53] J. Nanzer, *Microwave and Millimeter-Wave Remote Sensing for Security Applications*. Artech House, 2012.
- [54] K. Browne, R. Burkholder, and J. Volakis, "A Novel Low-Profile Portable Radar System for High Resolution Through-Wall Radar Imaging," in *IEEE Radar Conference*, 2010.
- [55] A. Gromek, "High Resolution SAR Imaging Trials Using a Handheld Vector Network Analyzer," in *Int. Radar Symp. (IRS)*, 2014.
- [56] A. O. Boryszenko, D. L. Sostanovsky, and E. S. Boryszenko, "Portable Imaging UWB Radar System With Two-Element Receiving Array," in *Ultra-Wideband Short-Pulse Electromagnetics*. Springer, 2007.
- [57] Y. Zhu, Y. Zhu, B. Y. Zhao, and H. Zheng, "Reusing 60GHz Radios for Mobile Radar Imaging," in *Proc. of ACM MobiCom*, 2015.
- [58] Y. Zhu, Y. Yao, B. Y. Zhao, and H. Zheng, "Object Recognition and Navigation using a Single Networking Device," in *ACM MobiSys*, 2017.
- [59] B. Wei, A. Varshney, W. Hu, N. Patwari, and et al., "dRTI: Directional RadioTomographic Imaging," *CoRR*, vol. abs/1402.2744, 2014.
- [60] A. Karpenko, D. Jacobs, J. Baek, and M. Levoy, "Digital Video Stabilization and Rolling Shutter Correction using Gyroscopes," Dept. of Comp. Sc., Stanford University, Tech. Rep. CSTR 2011-03, 2011.
- [61] C. Jia and B. L. Evans, "Online Camera-Gyroscope Autocalibration for Cell Phones," *IEEE Trans. on Image Processing*, vol. 23, no. 12, 2014.
- [62] L. Oth, P. Furgale, L. Kneip, and R. Siegwart, "Rolling Shutter Camera Calibration," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2013.
- [63] P. W. Kruse, *Uncooled Thermal Imaging Arrays, Systems, and Applications*. SPIE Press, 2001.
- [64] N. Wilde, S. Niederhaus, H. Lam, and C. Lum, "Handheld Ultrasonic Concealed Weapon Detector," *Proc. of SPIE*, vol. 4708, 2002.