Sensor-level Privacy for Thermal Cameras

Francesco Pittaluga   Aleksandar Zivkovic   Sanjeev J. Koppal

Electrical and Computer Engineering, University of Florida
Gainesville, FL 32611

f.pittaluga@ufl.edu

Abstract

As cameras turn ubiquitous, balancing privacy and utility becomes crucial. To achieve both, we enforce privacy at the sensor level, as incident photons are converted into an electrical signal and then digitized into image measurements. We present sensor protocols and accompanying algorithms that degrade facial information for thermal sensors, where there is usually a clear distinction between humans and the scene. By manipulating the sensor processes of gain, digitization, exposure time, and bias voltage, we are able to provide privacy during the actual image formation process and the original face data is never directly captured or stored. We show privacy-preserving thermal imaging applications such as temperature segmentation, night vision, gesture recognition and HDR imaging.

1. Introduction

Sensing and understanding humans and their movements is an important goal of computer vision and computational imaging research. Long wave thermal vision sensors ($8-14\mu m$) have a unique advantage over conventional grayscale or color sensors since human detection is based on both physical parameters (body temperature) and algorithmic outcomes. While thermal sensors have had impact in defense, surveillance and other areas, the high-cost and large form factor has prevented wider adoption.

Recently, however, breakthroughs have been made in miniature uncooled bolometer systems that allow low cost and high quality thermal imagery. These have applications in mobile devices [8], intelligent sensors for homes [31] and other areas. We can now anticipate a future filled with millions, if not billions, of networked thermal cameras.

While the impact of such technology is exciting, there will be significant societal pushback against widespread adoption, since many objects that are opaque in the visible spectrum, such as clothes and walls, are transparent in long wave thermal wavelength (for e.g., a recent U.S. Supreme Court ruling [30] on imaging through walls). We focus on the privacy of faces (instead of heart-rate [40], gait [12] or other biometrics) since significant computer vision research exists for recognizing faces in thermal imagery [32, 2, 17, 33, 18].

We propose sensor policies and accompanying imaging algorithms that retain the thermal camera advantages for sensing humans while reducing the capabilities of face recognition vision algorithms. Our sensors can be used in situations where it is important to sense and track people without compromising their privacy. A few examples of these include monitoring employees in a factory, preventing violent assaults in bathrooms, counting children in a playground and detecting falls in assisted living environments.

To achieve these goals, we propose new hardware, firmware and algorithms for thermal cameras, without requiring any additional or new optics. Our algorithms achieve privacy during image formation (i.e. during readout, amplification and digitization). This means that privacy is enforced when the measurements of the signal are actually being made and there is never any direct image capture of sensitive face information. We present three contributions for enforcing sensor-level privacy:

1. Digitization: We present a circuit design which creates silhouettes during image digitization by masking measurements in the human temperature range. Face recognition becomes impossible (i.e. zero recognition rate) since the related voltages are never digitized. The design can be implemented on a sensor ASIC and we demonstrate results in simulation on real thermal data.

2. Sensor noise: We present policies for changing the sensor noise characteristics while the image is being created, by manipulating the microbolometer voltages and gain amplification. We show how this noise negatively affects face recognition. We demonstrate algorithms to track people with this noisy data. This technique requires an additional calibration step to specify the desired degree of facial deidentification.

3. Exposure bracketing: We present privacy preserving thermal high-dynamic range (HDR) imaging. Our algorithm removes a "no capture" region in the scene

radiance while maintaining HDR elsewhere. We show theory for selecting the exposure brackets and demonstrate both simulations and real implementations of HDR thermal imaging. This method, like the digitization approach, removes faces (by either overexposure or underexposure of the corresponding pixels), resulting zero recognition rate.

1.1. Thermal signature model

Although human core temperature variation is within a few degrees Celsius, *facial skin* temperature is more relevant to preserving privacy. Our methods assume that human facial skin temperature lies in a known, narrow band and that other objects of the scene are not in this band. Although skin temperature can depend on a variety of physiological factors [16], we use a model that depends on both the normal internal body temperature ($37°C$) and on known ambient temperature, which is straightforward to measure via, say, an on-board thermometer. We assume these two opposing factors result in thermal equilibrium and induce a constant facial skin temperature. Finally, the mapping between the known ambient temperature and this constant facial skin temperature is an input to our methods, and can be obtained from the relevant studies in physiology [9].

1.2. Related Work

Low resolution thermal imagers: Low resolution thermal imagers (such as those used in smart buildings [31]) have just enough light-field samples for their task, suggesting that hardware under sampling may achieve both privacy and utility. However, this approach indiscriminately subsamples all scene information, whereas we present techniques, such as privacy during digitization and privacy preserving thermal HDR, that remove facial information while providing high resolution and HDR for the rest of the scene.

Thermal imaging for computer vision: Thermal computer vision has demonstrated reliable people tracking [35], health monitoring [40] and remote sensing [19]. A large body of thermal face recognition exists [32, 2, 17, 33, 18], and our goal is to capture thermal imagery while confusing these types of algorithms, without losing the utility of thermal imagery for other vision techniques. With the advent of mobile thermal imagery such as the iPhone FLIR ONE [8] system, maintaining usefulness and privacy will continue to be an important goal.

ϵ**-Photography:** In computational photography, capturing images while changing camera parameters by small amounts [28] allows many light-field applications. We are inspired by efficient photography, where sensor constraints allow closed form solutions [13], and our goal is to deliver privacy-enabled thermal imagery (defined by either overexposed or under-exposed pixels that correspond to human

| Variable name	Meaning
Φ	Radiant power
G_{FID}, G_{SK}	Bias voltages
λ	Wavelength of incident light
λ_h, λ_l	Limits of wavelength sensitivity
$s(\lambda)$	Sensor spectral sensitivity
$r(\lambda)$	Scene power density
t	time
n	Sensor noise
g	Gain
I_{max}	Pixel value for sensor saturation
$\widetilde{I}_{max}(g)$	Gain dependent practical sensor saturation
I_{min}	Minimum pixel value
$\widetilde{I}_{min}(g)$	Gain dependent practical minimum pixel value
σ_{read}	Sensor noise influenced by bias voltages
σ_{ADC}	ADC induced sensor noise
e_i	i^{th} exposure value
f	Thermal camera response function
Φ_{min}, Φ_{max}	Max/min radiances for "no capture"
$(T_{min}, ..., T_{max})$	Integration times for an epsilon photograph
$(G_{min}, ..., G_{max})$	Gains for an epsilon photograph
β_{min}	Lowest radiance for "no capture"
β_{max}	Highest radiance for "no capture"
DR_{max}	Maximum achievable dynamic range

Table 1. Summary of symbols used

data), while maintaining the integrity of the remaining portions of the scene and a given time budget.

Privacy-preserving vision sensors: Most privacy preserving vision algorithms apply k-anonymity, pixelation, Gaussian blurring, face replacement or black-out [3, 34, 20, 1] after images of the scene are already captured. Our work is about maintaining privacy during the process of converting photons into pixels. Recently, vision sensors have been proposed that transform data at the sensor level itself by using embedded processing [23, 5, 38] or using custom sensors for watermarking [25], cartooning [37] or pixel averaging [7]. Other techniques use special optics attached to the camera [27, 41]. Our approach differs in two ways. First, in the thermal domain, we avoid depending on an embedded vision algorithm (such as face detection) whose failure in even a single frame may eliminate privacy. Instead we exploit reliable, temperature differences. Second, our use of exposure, gain and bias voltages for privacy works with existing thermal cameras and, in this sense, provides another security layer to the above complementary techniques.
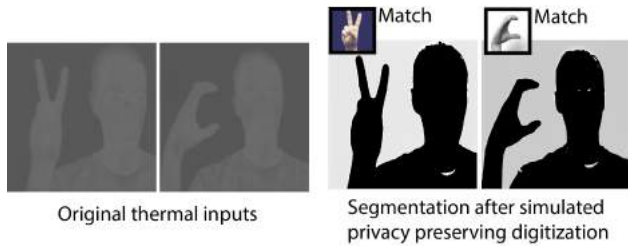
2. Privacy during digitization

Our first technique for pixel-level sensor privacy is through a component in an Application Specific Integrated Circuit (ASIC) that filters those voltages corresponding to human wavelengths. ASICs provide a low-power and scalable option for very specific image processing requirements. Further, sensor-level ASICs can directly participate
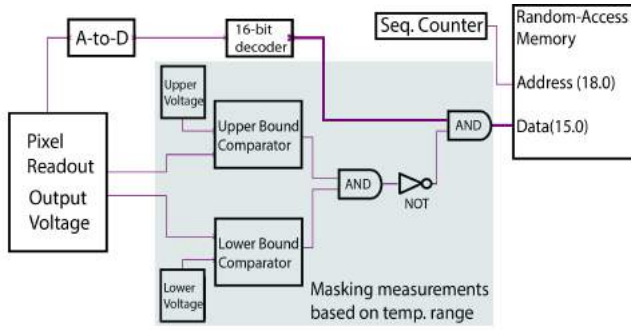
(I) Proposed digitization circuit block diagram

(II) Circuit simulations on real thermal imagery

(III) Application for hand gesture recognition

Figure 1. **Privacy during digitization:** During the digitization process, we can preserve privacy by masking out those sensor measurements that fall in the human body temperature range. In (I) we show a circuit diagram that does this, and in (II) we show simulations of this circuit on real thermal data. Pixels corresponding to the person's skin are removed whereas the hot coffee mug and the frozen water bottle are preserved. In (III) we show an application of this approach for gesture recognition.

in the digitization process, without any storage or communication with other parts of the thermal imaging hardware.

In Fig. 1(I) we show a high level block diagram of a circuit to remove sensor measurements in the human temperature range, while digitization occurs with an A-to-D converter. This design is simple and (for a factory calibrated thermal sensor) has no external parameters, other than the ambient temperature discussed earlier.

Given an accurate sensor calibration, all pixel values associated with a given human spectrum thermal band can be set to zero. Additional necessary circuit components, such as the serial pixel organization streaming section and syn-

Figure 2. **Digitization privacy in different scenes:** Here we show additional digitization results in scenes with people, computers and buildings. The left column are the input 16 bit images and the right column is the simulated output.

chronizing clock signals, have been left out. In Fig. 1(II) we show software simulations of the circuit using data taken with a real thermal sensor. The scene shows a person carrying a hot coffee mug and a frozen water bottle, and pixels corresponding to the human temperature range are removed. Some pixels corresponding to the person's clothing were also removed, since these were warmed by body proximity. Three additional results are shown in Fig. 2. The first shows a person with a computer, which also generates heat, and which is preserved by the technique. In the second and third, groups of people are moving in a building environment. In all of these results, most face pixels are removed.

Gesture recognition: We demonstrate a gesture recognition application of privacy during digitization in Fig. 1(III). We used a gesture database from [21] where data taken under a light/dark background can be easily binarized into silhouettes. This gives a training set of 10 hand gesture classes and 20 examples of each gesture. The classification was done using a multi-class bag-of-words SVM-based classifier [6] that operates on keypoints (we used SURF features). The test images were generated through simulation of the circuit in Fig. 1(I) and we captured 20 test images for the ten classes with a success rate of 97%.

3. Privacy with sensor noise

Noise added to grayscale and RGB images in software is known to provide high levels of face anonymity at the cost of reduction in image utility [34]. We will demonstrate that, for thermal cameras, noise added through sensor processes during image creation provides pre-capture anonymity while still allowing useful applications.

Let us consider raw images from a thermal sensor that give rise to independent and linear [32] pixel measurements. From [13] and [36], we can model the linear response function for each pixels as generating the appropriate value of radiant power, Φ as (from [36])

$$\Phi = \int_{\lambda_l}^{\lambda_h} s(\lambda) r(\lambda) d\lambda \qquad (1)$$

where $s(\lambda)$ is the sensor's spectral response function and $r(\lambda)$ is the incident power density per unit time at wavelength λ. The limits λ_l and λ_h are the wavelength limits beyond which the spectral response of the sensor is zero.

As in [13], since Φ is expressed in electrons per second, then Φt measured during an exposure of t seconds creates the pixel measurement I expressed as digital numbers (DN) [22]. This is written (from [13]) as,

$$I = min\{\Phi t/g + n, I_{max}\} \qquad (2)$$

where n is the sensor noise that is signal and gain dependent and that we wish to exploit to create anonymity and that is described in further detail below. I_{max} is the thermal sensor's level of total saturation.

The noise model above is a zero-mean random variable with three modes that are considered to independent. Unsaturated pixels are expressed as (from [13]),

$$\text{Var}(n) = \overbrace{\Phi t/g^2}^{\text{scene dependant}} + \overbrace{\sigma_{read}^2/g^2 + \sigma_{ADC}^2}^{\text{scene independent}} \qquad (3)$$

Sensor gain: An immediate implication of [13]'s noise model shown in Eq. 3 is that reducing the gain g for a fixed exposure time t increases noise. Although image degradation has been used in the context of privacy [34], we are the first to point out the pre-capture advantage of adding gain-related noise during amplification rather than later in software. A disadvantage of using gain is that it results in scene dependent noise. However, this can be compensated using an appropriate exposure. Consider t_{max}, the most exposure allowed for some particular task and time budget. In all further discussions we set the gain to be $g = \sqrt{t_{max}}$ which is the lowest value it can be while allowing the exposure to remove scene dependency. We now discuss the indirect control of scene independent noise parameter σ_{read} in Eq. 3 using thermal bolometer bias voltages.
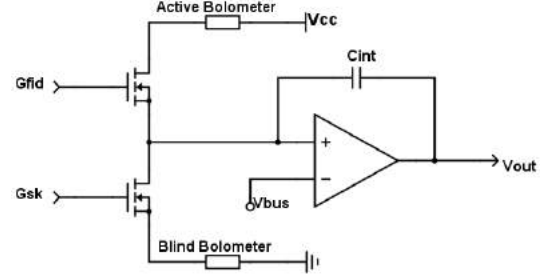

Figure 3. **Typical pixel readout circuit for thermal cameras (adapted from [24, 26, 15]) :** Most thermal cameras remove background thermal noise with a "blind" bolometer that captures ambient thermal signals. The weight given to this background factor depends on the ratio of two voltages, G_{FID} and G_{SK}. We exploit these to add scene independent read noise during image formation.

3.1. Bias voltages

Thermal camera pixel readout circuits depend on a designer's particular requirements. However, the fundamental circuit topology usually follows Fig. 3 (adapted from [24, 26, 15]). Here, G_{FID} and G_{SK} are gate voltages for both the nMOS and the pMOS transistors [39], respectively and are called *bias voltages*. Bias voltages are typically pre-set to some optimal value by the camera manufacturer, given the particular IC characteristics used in the circuit.

Bias voltages cause two effects; offsetting and dynamic range scaling. The blind bolometer offsets ambient temperature from the scene temperature read by the active bolometer. Additionally, if the blind bolometer bias voltage changes, the dynamic range will proportionally vary as well, because the compensation current to normalize the active bolometer readings changes. In Fig. 4(I), face images are taken for different exposures as the bias voltages G_{FID} and G_{SK} vary from zero to their maximum value. Useful thermal face images occur when the human signature falls within the range and offset provided by the bias voltages.

The transistors for the read-out circuit in Fig. 3 operate between their cut-off and saturation levels. By deliberately changing the bias voltages, the overall microbolometer reading will be skewed towards either of these extremes, given by the bit depth of the camera. This skew causes the voltage output of the integrator stage to either be much higher or lower than what the true value is from default bias calibration. If the bias is sufficiently close to the fringe of the thermal cameras resolution, parts of the image that are outside the expected resolution range will be set to zero. *In other words, pushing the bias voltages outside the recommended ranges will increase σ_{read} in Eq. 3.*

(I) Bias voltages and exposure varied for a given scene, showing the offset and scaling effects

(II) Bias voltages with exposure set to 5 for a flat lambertian plane

(III) Standard deviations for each image in (II) with maxima highlighted

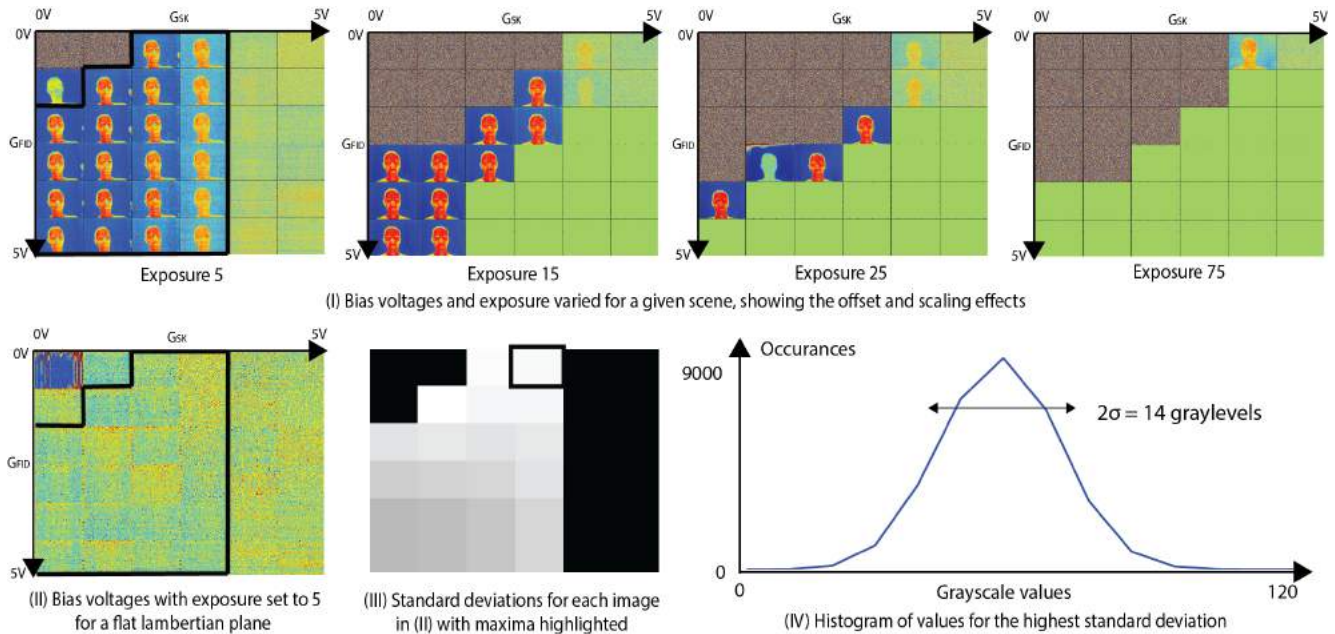(IV) Histogram of values for the highest standard deviation

Figure 4. **Noise calibration for privacy:** (I) shows the effect of bias voltages and (II) shows how to reverse engineer any thermal camera by collecting images of a plane at constant temperature over different exposures across the bias voltage range. The highest standard deviation of the image set (III) gives a histogram (IV) illustrating the ability of those bias voltages to remove data from images.

3.2. Calibration for privacy

The relationship between the bias voltages G_{SK} and G_{FID} to the readout noise σ_{read} depends on the specific sensor architecture. We reverse engineer the bias voltage values required for obtaining pre-capture privacy by simply varying the bias voltages over a range of voltages and exposures, while the camera views a untextured lambertian plane at a constant temperature.

Fig. 4(II) demonstrates this calibration using a 16 bit Xenics Gobi 640 thermal camera with a 640×480 spatial resolution and $0.005C$ thermal resolution. Such an image of a simple scene should be smooth, without noise. By recovering the standard deviations of the pixels, as in Fig. 4(III), we can obtain a measure of how much the bias voltages settings degrade the image, as visualized by the histogram of the highest standard deviation image shown in Fig. 4(IV). We only use the noise values that correspond to the bias voltages that are able to image our target scene (i.e. humans) depicted by the black bounding lines in Fig. 4.

Evaluating privacy of bias voltage induced noise In Fig. 5(I-IV) we show images created by setting the bias voltages based on the calibration in Fig. 4 (G_{SK} to zero and G_{FID} to $3.8V$). Comparing the Fourier spectrum of faces taken under normal bias voltages with the spectrum of the same faces taken with privacy preserving voltages shows significant loss in information. In Fig. 5(V) we show the results of using a commercially available infrared face recognition software [29] on a small database of five individuals.

(I) Images with normal bias voltages, with example Fourier spectrum

(II) Images with privacy preserving bias voltages, with example Fourier spectrum

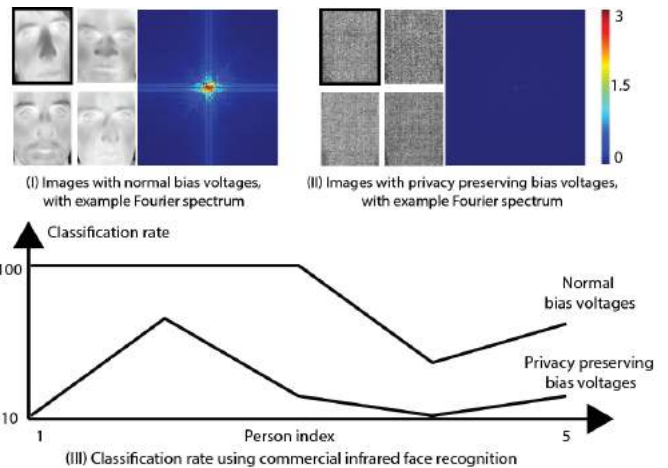(III) Classification rate using commercial infrared face recognition

Figure 5. **Bias voltage noise provides face privacy:** Normal bias voltage values create clear thermal images as in (I), and the Fourier spectrum of one of these images is shown. Setting G_{SK} to zero creates noisy images (II) whose Fourier spectrum shows significant degradation of information. In (III) we show the classification rates of a commercially available IR recognition software [29], which show the privacy preserving nature of sensor noise.

The individuals were imaged in the same room, at the same distance from the camera and with approximately the same pose. Both clear images and noisy images were obtained and ten training images and ten test images were used for
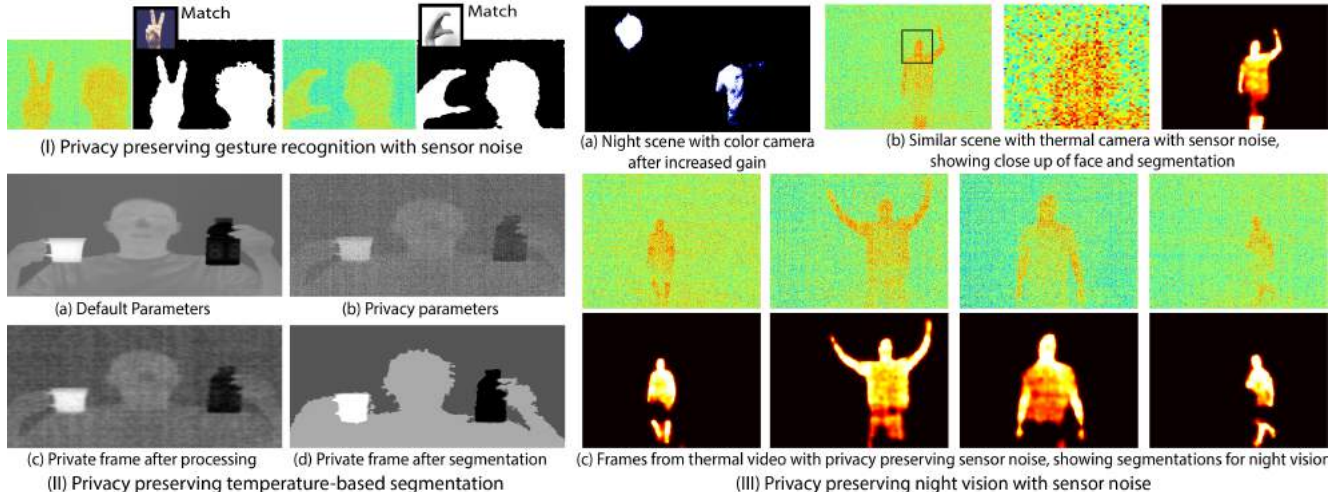
Figure 6. **Privacy-preserving thermal segmentation and people tracking with noise:** In (I) we show images captured with sensor noise and demonstrate segmentation and gesture recognition. In (II) we show a scene where a person carries a frozen water bottle and a boiling hot cup of water. A thermal camera views the same scene, but we have adjusted the noise level to be such that the faces are obscured, while still allowing segmentation of the background, person, boiling water and the frozen bottle. In (III)a we show an image of a night scene taken with a color camera at high gain, showing the lack of visible information. A closeup of a face taken with noisy, private thermal measurements is shown in (III)b. We demonstrate segmentation and tracking of a person at night in (III)c.

each person. In the figure, we see the average recognition rates for the test images fall when sensor noise is added.

3.3. Privacy-preserving applications

In Fig. 6 we show three applications, that are discussed below, using the previously calibrated sensor parameters.

Gesture recognition: We show gesture recognition in Fig. 6(I) using the same training dataset [21] and classifier [6] from the previous section. Here, however, segmentation was done on real, noisy thermal data by closing contours on a Sobel edge map. As a post-processing step, the segmentation was cleaned by removing connected components and applying diamond shaped erosion. We tested 20 images for the ten classes and achieved a 97% recognition rate.

Temperature-based segmentation: In Fig. 6(II), we show color images of a person carrying a frozen water bottle and a hot cup of water. We capture the noisy, private image and apply median filtering to obtain the images shown in the second row. Despite the heavy noise induced by the bias voltages, the large temperature differences allowed straightforward threshold-based image segmentation for the background, the person, the hot mug and the frozen bottle.

People tracking at night: In Fig. 6(III)a, we show a color image of a night scene. Even with high gain, the image has very little information. Thermal cameras can easily image such scenes, and we demonstrate this with privacy, as shown by the face closeup in Fig. 6(III)b. Four frames from a video are shown in Fig. 6(III)c as the person moves around the scene. We are able to segment a silhouette and

track the person at night by applying iterative guided image filtering [14] to smooth the image and combining thresholded masks from different iterations.

4. Privacy-preserving exposure bracketing

We now use exposure bracketing to either overexpose or underexpose some target radiance range, without requiring additional hardware and without degrading the remaining pixels. Like the previous methods, this algorithm removes sensitive information during image formation. Unlike the previous methods, it requires multiple images, each of which is also anonymized and which are combined in software, after the images have been captured.

4.1. "No capture" radiance range

Recall the definition of radiant power in Eq. 1. Let us consider the thermal case, with a broadband sensor with sensitivity $s(\lambda)$. The extreme thermal wavelengths are $\lambda_l = 7\mu m$ and $\lambda_h = 14\mu m$. The human thermal spectrum $r(\lambda)$ behaves like an ideal black body with an absolute temperature of 310.15 Kelvin. If we define $\Delta(\lambda)$ as a wavelength dependent variance on the human thermal spectrum $r(\lambda)$, then this results in two variations of Eq. 1

$$\Phi_{min} = \int_{\lambda_l}^{\lambda_h} s(\lambda) \left[r(\lambda) - \frac{\Delta(\lambda)}{2} \right] d\lambda \qquad (4)$$

and

$$\Phi_{max} = \int_{\lambda_l}^{\lambda_h} s(\lambda) \left[r(\lambda) + \frac{\Delta(\lambda)}{2} \right] d\lambda. \qquad (5)$$

We depict these equations in Fig. 7(a). Fig. 7(b) illustrates our key idea, which is to create an HDR image of the scene without any measurements in the range given by $[\Phi_{min}, \Phi_{max}]$, which we define as the "no capture" range.

A consequence of defining the "no-capture" range in terms of image irradiance is that surfaces with a spectral radiosity outside the spectral radiosity range $r(\lambda) \pm \frac{\Delta(\lambda)}{2}$, may fall within the no-capture range $[\Phi_{min}, \Phi_{max}]$, when expressed in terms of Eq. 4 and Eq. 5, which results in unintended black-out. In the visible spectrum, this makes privacy preserving exposure bracketing difficult since many surfaces will have a spectral radiosity that maps to the "no-capture range" of human skin. This is less of a problem in the thermal domain, as the radiosity to irradiance mapping effectively corresponds to a one-to-one map between radiosity and temperature, i.e. it is unlikely that objects lie within the human-temperature range.

4.2. Removing the "no capture" region

For single image capture, we can remove the "no capture" scene radiances by over or under exposure of the appropriate range. Consider a pixel range (I_{min}, I_{max}), where, from Eq. 2, sensor saturation occurs when gI_{max} electrons are collected.

Note that the pixel values just below I_{max} may also be untrustworthy and may actually be degraded by noise [13]. As in [13], in practice we use $\widetilde{I}_{max}(g) < I_{max}$, which must be measured for different values of sensor gain using a calibration step, and similarly for $\widetilde{I}_{min}(g) > I_{min}$. Complete overexposure of the "no-capture" range occurs when the integration time t is such that,

$$t \geq \frac{g\widetilde{I}_{max}(g)}{\Phi_{min}} \qquad (6)$$

and complete underexposure occurs when the integration time t is such that,

$$t \leq \frac{g\widetilde{I}_{min}(g)}{\Phi_{max}}, \qquad (7)$$

where g is gain, and Φ_{min} and Φ_{max}, expressed in terms of electrons collected per second, are the lower and upper boundaries of the no-capture range, respectively. We depict these equations with dashed lines in Fig. 7(b), although in reality the mapping will be piecewise constant.

4.3. Achievable dynamic range with variable gain

Consider an epsilon-photography experiment where a stack of images $\mathbf{S}$ is collected under a finite-discrete set of increasing exposure times and gain settings given by $\{T_{min}, ..., T_{max}\}$ and $\{G_{min}, ..., G_{max}\}$ respectively. Given $\mathbf{S}$, not all "no capture" regions can be feasibly over or under exposed. We define a minimum radiance, β_{min},

and maximum radiance, β_{max}, for which over or under exposure is possible, given this image stack.

We can pick the gain and the exposure for the radiance limits for the no capture region β_{min} and β_{max} from Eq. 2, Eq. 6 and Eq. 7, resulting in the following expressions,

$$\beta_{min} = \frac{G_{min}\,\widetilde{I}_{max}(G_{min})}{T_{max}} \qquad (8)$$

and

$$\beta_{max} = \frac{G_{max}\,\widetilde{I}_{min}(G_{max})}{T_{min}}. \qquad (9)$$

Finally, given the stack of images $\mathbf{S}$, the no-capture range $[\Phi_{min}, \Phi_{max}]$ and radiance limits for the no capture region $(\beta_{min}, \beta_{max})$, the maximum dynamic range, DR_{max} achievable with the "no capture" range removed is given by one of three cases:

1. **if** $\Phi_{min} \geq \beta_{min}$ **and** $\Phi_{max} \leq \beta_{max}$,
   then $DR_{max} = [0, (G_{max}\,\widetilde{I}_{min}(G_{max}))/T_{min}]$.

2. **if** $\Phi_{min} \geq \beta_{min}$ **and** $\Phi_{max} \geq \beta_{max}$,
   then $DR_{max} = [0, \Phi_{min}]$.

3. **if** $\Phi_{min} \leq \beta_{min}$ **and** $\Phi_{max} \leq \beta_{max}$,
   then $DR_{max} = [\Phi_{max}, (G_{max}\,\widetilde{I}_{min}(G_{max}))/T_{min}]$.

In the fourth case, no image in $\mathbf{S}$ can be anonymized, and privacy preserving exposure bracketing is impossible.

4.4. Exposure bracketing

Exposure bracketing based ϵ-photography methods create a single image from the stack $\mathbf{S}$. This renders a single effective sensor [11], whose camera response function can be defined in terms of the original image stack. If there are n exposures $\{T_1, ..., T_n\}$ from a camera with response f, then (from [11]),

$$h(\Phi) = \sum_{i=1}^{n} f(T_i\Phi) \qquad (10)$$

In our analysis, f is governed by the pixel measurement model in Eq. 2. Given a desired effective response function h_{des}, we can define an objective function based comparing the piecewise derivatives of the response function, as demonstrated by [11]. Other definitions of the objective function can also be used instead, such as maximizing the HDR image's SNR [13, 10]. Unlike [11], we remove any computation of the objective function in the "no capture" region, since we will later show how each exposure avoids capturing data in this region. The objective function is

$$\xi(n, \mathbf{T}) = \int_{\Gamma_{min}}^{\Phi_{min}} |h'_{des} - h'|^p \, \omega \, d\Phi + \int_{\Phi_{max}}^{\Gamma_{max}} |h'_{des} - h'|^p \, \omega \, d\Phi \qquad (11)$$
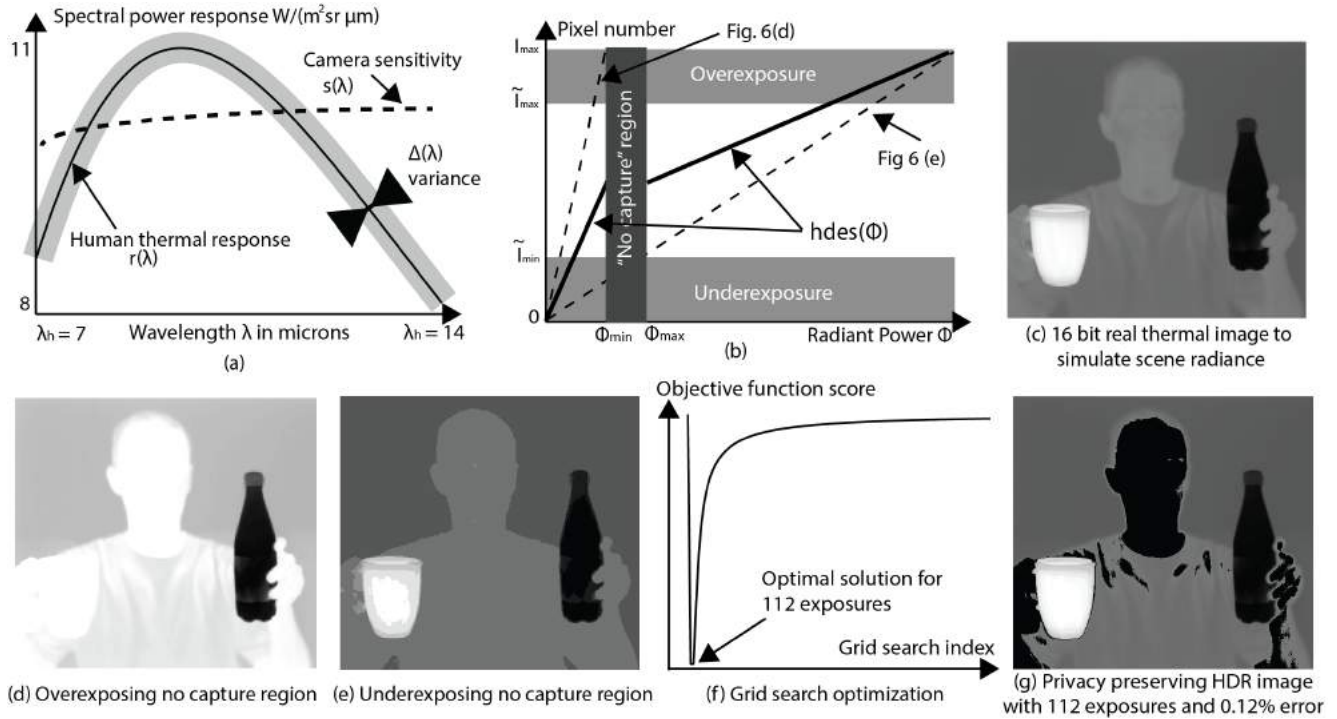
Figure 7. **Privacy-preserving exposure bracketing:** In (a) we show our model of the human thermal signature based on a black body with absolute temperature of 310.5. The radiances are mapped to pixel values in (b) using the camera exposure function (for clarity we avoid illustrating these as piecewise constant). We specify the "no capture" region as a set of radiances that we do not wish to capture. If the "no capture region" corresponds to human radiances, then we can use exposures so that the ground truth scene radiances in (c) are mapped to images such as (d) and (e), where the human pixels are either over or under exposed. The optimal values for these exposures can be obtained (f) to obtain a privacy preserving HDR thermal image, which has no information in the "no capture" region.

where Γ_{min} and Γ_{max} are the boundaries of the dynamic range of g, p is a positive number indicating the norm and Φ_{min} and Φ_{max} are the boundaries of the no-capture range. The weighting function is designed as in [11],

$$\omega = \begin{cases} 0 & h'_{des}(\Phi) < h'(\Phi) \\ 1 & \text{otherwise,} \end{cases} \quad (12)$$

and allows dense quantization but extracts a penalty for an effective sensor that samples less densely than desired. We constrain the minimization to positive exposures and such that the no-capture range is blacked out for each exposure:

$$\arg\min_{n,\mathbf{T}} \xi(n, \mathbf{T}) \text{ s.t.}$$

1. $T_i > 0$ (13)

2. $\forall\, T_i \; \left[T_i \geq \frac{\widetilde{I}_{max}}{\Phi_{min}} \oplus T_i \leq \frac{\widetilde{I}_{min}}{\Phi_{max}} \right],$

where $\oplus$ denotes logical OR, and $\widetilde{I}_{min}$ and $\widetilde{I}_{max}$ are the maximum and minimum reliable pixel value of the sensor, for exposure setting T_i. Constraint 1 enforces that the exposures be positive. Constraint 2 enforces that the "no capture range" be over or under exposed for each image.

We solve the above minimization through a simple exhaustive grid search over the parameters of exposures, which assumes fixed values for the maximum number of exposures and the step size for searching over the space of scene radiant intensities. This approach is tractable when there is a single "no capture" region, but becomes combinatorially expensive for multiple regions. We leave the creation of a general optimal algorithm for future work.

4.5. Exposure bracketing results

Simulated HDR results: In Fig. 7(c) we show a 16 bit image captured with a Gobi 640 thermal camera that is used to simulate scene radiance for the privacy-preserving HDR algorithm. The image contains over 2400 distinct grayscale values ranging from 0 to 9940, and the human temperature range falls between $\Phi_{min} = 4644$ and $\Phi_{max} = 5644$. The goal is to recreate this scene radiance with an image stack of 8 bit simulated exposures. We assume gain is not a factor (it is set to 1) and set the practical limits on pixel values to be $\widetilde{I}_{max} = 250$ and $\widetilde{I}_{min} = 5$.

In Fig. 7(d-e) we show two exposures designed so that the human temperature range is either over or under ex-
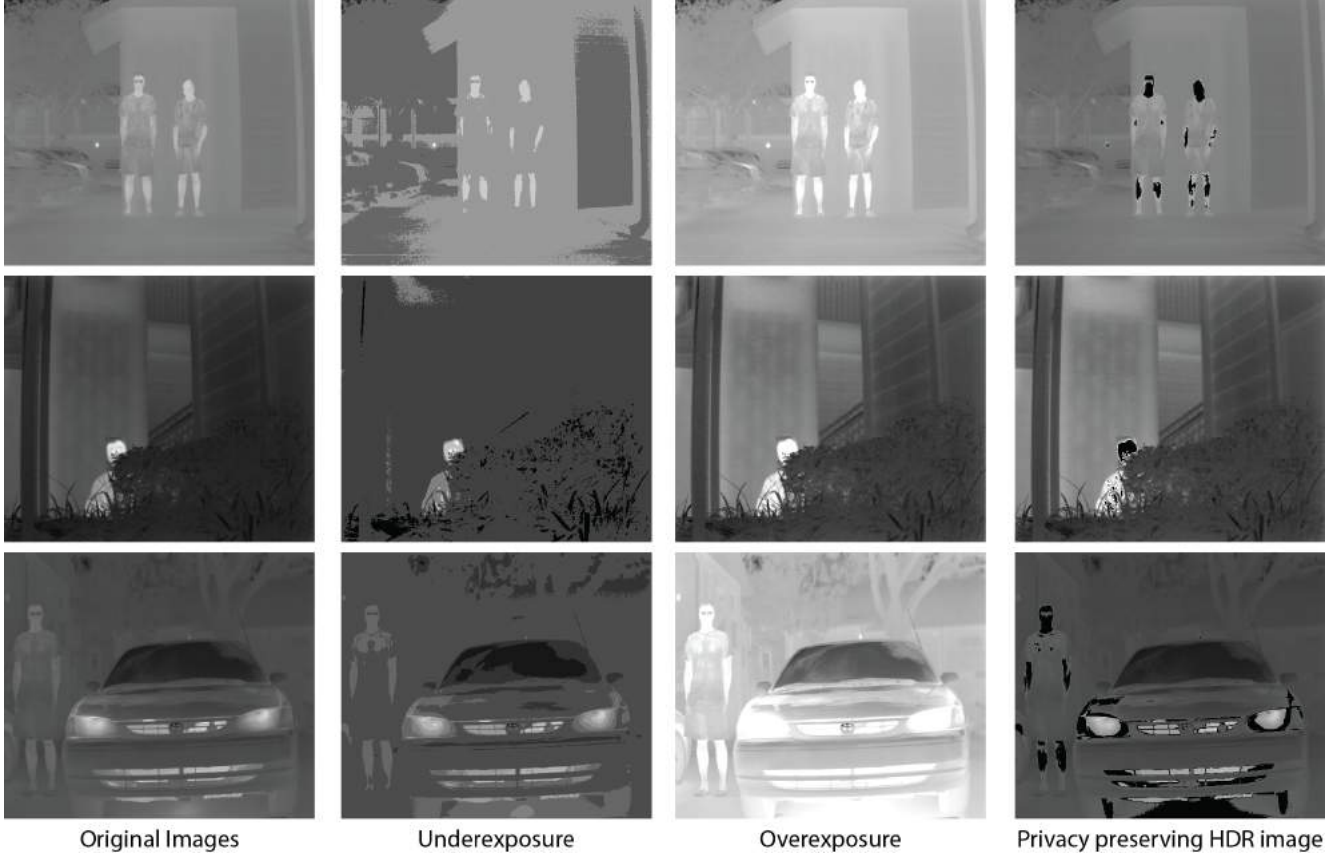
Figure 8. **Outdoor privacy preserving HDR:** Here we show additional simulated HDR results in the outdoors. The first column is the original 16 bit data used as scene radiance input, while the second and third column are 8 bit images rendered from these where the face information is either under or over exposed. The last columns corresponds to the HDR images created. The average pixel errors from top to bottom are 0.12%, 0.0319% and 0.0522%.

posed, as discussed by Eq. 6 and Eq. 7. For a set of 112 exposures and for a fixed step size of 0.1 along the radiant power axis, we perform a grid search optimization in Fig. 7(f), whose optimal parameters are used to generate the HDR image Fig. 7(g). The region of the original image that corresponds to human skin and body is removed, because this region is either over or under exposed in each of the 112 input exposures. Comparing the non-zeroed out pixels to the original image reveals a 0.12% average error.

In Fig. 8 we show results for testing the approach just described on three outdoor scenes containing buildings and vegetation. The first column shows the raw 16 bit images used as input to the simulation, while the next two show examples of over and under exposure of that data to simulate 8 bit image captures. The final column shows the HDR results, and the average pixel errors from top to bottom are 0.12%, 0.0319% and 0.0522%.. In the first two rows, the humans are still the only objects in the scene that generates heat. Note that the face pixels are completely removed in these results; in the second row, where the person is looking

through leaves, the individual leaves are preserved because this is a per-pixel approach. In the third example, a parked car with its engine running is imaged with a person. Since the car heats non-uniformly, some parts of it are within the human temperature band and are zeroed out. Although this is a case where we do not produce a high-quality image, privacy is maintained and most of the scene is captured.

Real HDR experiments trading off SNR/exposure: If we adjust integration time for each gain and bias voltage setting to keep image measurement $\frac{\Phi t}{g}$ constant, lower gain results in lower SNR [13]. Hence, once we have found the optimal set of exposures, we can trade-off capture-time for SNR or vice-versa.

Fig. 9 shows two sets of low SNR images taken at different exposures, 1 and 15. Privacy and fast capture are obtained by fixing the gain at $g = 1.5$ and using different bias voltages; $G_{sk} = 0V, G_{fid} = 3.005V$ for over exposure and $G_{sk} = 2.915V, G_{fid} = 1.5V$ for under exposure. This results in a fast capture of two images, without imaging any
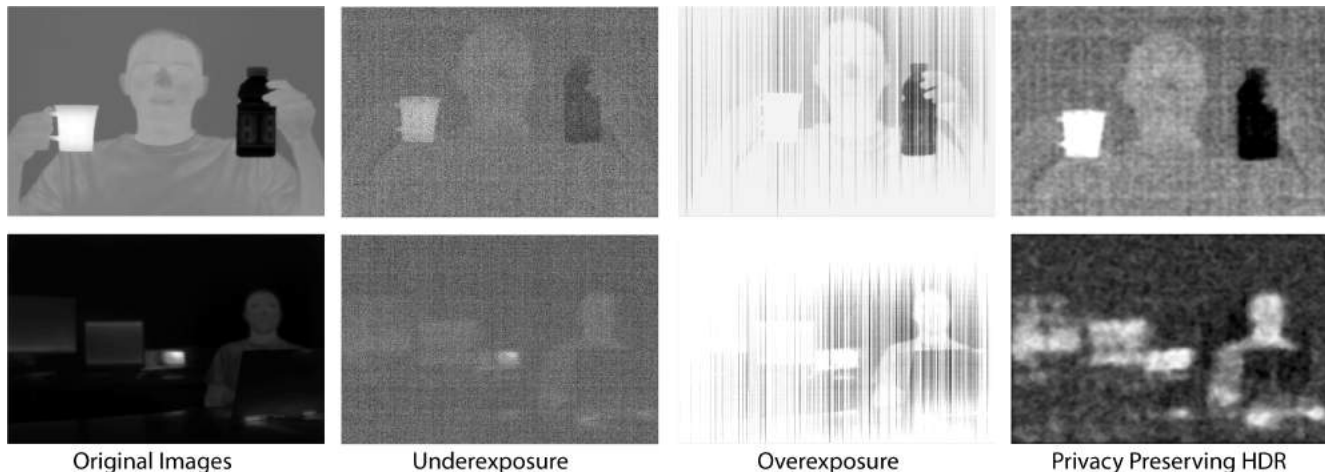
Figure 9. **Segmenting over a large dynamic range with two exposures:** Here we show two real privacy-preserving HDR results. By manipulating sensor parameters, we can quickly capture privacy-preserving over and underexposed images, which we fuse to generate a privacy-preserving HDR image. The images in the first row show a person holding a coffee mug and frozen water bottle. The images in the second row show a typical office scene: the background contains two monitors and a projector and the foreground contains a person seated behind a laptop.

facial features. We then throw away over or underexposed pixels in the respective over and underexposed images and fuse them to generate privacy-preserving HDR images.

5. Summary

We are the first to show that sensor level manipulation of thermal imagery offers an opportunity to create computational photographs that have privacy *and* utility. We have demonstrated simulations to show that a comparator circuit can create silhouettes of humans in thermal images during digitization of the image. We have shown real results for adding bias voltage noise to a thermal image, and demonstrated useful applications despite this noise. Finally, we have proposed a privacy preserving exposure bracketing system that allows for capture of HDR images.

Our three approaches are complementary. For example, the digitization and HDR techniques have low noise and produce good image quality outside the face region, when compared to the sensor noise approach that exploits bias voltages. On the other hand, the digitization and sensor noise approaches run in real-time. In contrast, the HDR technique requires multiple images, and, without additional processing, is only relevant for static scenes. Finally, the digitization technique requires additional hardware and firmware upgrades to the camera, whereas the sensor noise approach and the HDR technique can be used with currently available thermal cameras that allow bias voltage and exposure control, respectively.

5.1. Limitations

A structural limitation is that identification based on biometrics such as silhouette and gait may be extracted from images produced by our techniques. We have focused on removing face information, due the disproportionate availability of technology and databases for faces, when compared to other biometrics. Additional limitations include:

Background thermal signature: Our approaches assumes that most scene objects do not lie in the human thermal signature band. This assumption is generally reasonable, since most objects in a scene do not generate their own heat, unlike humans, and would therefore be close to the ambient temperature. Also, many objects that generate their own heat (fireplace, industrial machinery, etc.) are at a higher temperature than humans. Even if some objects' signatures do cross into that band, as in the last row of Fig. 8, scene understanding may still be possible.

Calibration: The digitization and HDR techniques require the same type of temperature-based calibration that all uncooled cameras need [4]. The sensor noise step requires an additional, application dependent step to specify the extent of face deidentification that is required.

Reflections: Smooth surfaces reflect thermal radiation, while scattering reduces the energy incident at the sensor. Therefore reflections of faces may not be removed by the digitization and HDR techniques.

6. Acknowledgements

References

[1] P. Agrawal and P. Narayanan. Person de-identification in videos. *IEEE Transactions on Circuits and Systems for Video Technology*, 2011.

[2] G. Bebis, A. Gyaourova, S. Singh, and I. Pavlidis. Face recognition by fusing thermal infrared and visible imagery. *Image and Vision Computing*, 24(7):727–742, 2006.

[3] M. Boyle, C. Edwards, and S. Greenberg. The effects of filtered video on awareness and privacy. 2000.

[4] H. Budzier and G. Gerlach. Calibration of uncooled thermal infrared cameras. 2015.

[5] A. Chattopadhyay and T. E. Boult. Privacycam: a privacy preserving camera using uclinux on the blackfin dsp. *CVPR*, 2007.

[6] G. Csurka, C. Dance, L. Fan, J. Willamowski, and C. Bray. Visual categorization with bags of keypoints. In *Workshop on statistical learning in computer vision, ECCV*, volume 1, pages 1–2. Prague, 2004.

[7] J. Fernández-Berni, R. Carmona-Galán, R. del Río, R. Kleihorst, W. Philips, and Á. Rodríguez-Vázquez. Focal-plane sensing-processing: A power-efficient approach for the implementation of privacy-aware networked visual sensors. *Sensors*, 14(8):15203–15226, 2014.

[8] FLIR. FLIR ONE at www.flir.com. 2015.

[9] J. Frim, S. Livingstone, L. Reed, R. Nolan, and R. Limmer. Body composition and skin temperature variation. *Journal of Applied Physiology*, 68(2):540–543, 1990.

[10] M. Granados, B. Ajdin, M. Wand, and C. Theobalt. Optimal hdr reconstruction with linear digital cameras. *CVPR*, 2010.

[11] M. D. Grossberg and S. K. Nayar. High dynamic range from multiple images: Which exposures to combine? *Proc. of ICCV Workshop on Color and Photometric Methods in Computer Vision*, 2003.

[12] J. Han and B. Bhanu. Individual recognition using gait energy image. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 28(2):316–322, 2006.

[13] S. W. Hasinhoff, F. Durand, and W. T. Freeman. Noise–optimal capture for high dynamic range photography. *CVPR*, 2010.

[14] K. He, J. Sun, and X. Tang. Guided image filtering. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 35(6):1397–1409, 2013.

[15] M. Henini and M. Razeghi. *Handbook of infra-red detection technologies*. Elsevier, 2002.

[16] Y. Houdas and E. Ring. *Human body temperature: its measurement and regulation*. Springer Science & Business Media, 2013.

[17] S. G. Kong, J. Heo, B. R. Abidi, J. Paik, and M. A. Abidi. Recent advances in visual and infrared face recognitiona review.

[18] S. G. Kong, J. Heo, F. Boughorbel, Y. Zheng, B. R. Abidi, A. Koschan, M. Yi, and M. A. Abidi. Multiscale fusion of visible and thermal ir images for illumination-invariant face recognition. *International Journal of Computer Vision*, 71(2):215–233, 2007.

[19] C. Kuenzer and S. Dech. *Thermal Infrared Remote Sensing*. Springer, 2013.

[20] G. Loukides and J. Shao. Data utility and privacy protection trade-of in k-anonymisation. 2008.

[21] S. Marcel and O. Bernier. Hand posture recognition in a body-face centered space. In *Gesture-Based Communication in Human-Computer Interaction*, pages 97–100. Springer, 1999.

[22] E. Martinec. Noise, dynamic range and bit depth in digital slrs. http://theory.uchicago.edu/, 2008.

[23] P. J. Narayanan and Mrityunjay. The de-identification camera. *CVPR*, 2011.

[24] A. Nawrat and Z. Kuś. *Vision Based Systemsfor UAV Applications*, volume 481. Springer, 2013.

[25] G. R. Nelson, G. A. Jullien, and O. Yadid-Pecht. Cmos image sensor with watermarking capabilities. In *Circuits and Systems, 2005. ISCAS 2005. IEEE International Symposium on*, pages 5326–5329. IEEE, 2005.

[26] W. J. Parrish and J. T. Woolaway II. Improvements in uncooled systems using bias equalization. In *AeroSense'99*, pages 748–755. International Society for Optics and Photonics, 1999.

[27] F. Pittaluga and S. J. Koppal. Privacy preserving optics for miniature vision sensors. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 314–324, 2015.

[28] R. Raskar. Computational photography: Epsilon to coded photography. In *Emerging Trends in Visual Computing*, pages 238–253. Springer, 2009.

[29] L. Rosa. Infrared face recognition system. 2015.

[30] R. H. Seamon. Kyllo v. united states and the partial ascendance of justice scalia's fourth amendment. *Washington University Law Quarterly, February*, 2002.

[31] A. Sixsmith and N. Johnson. A smart sensor to detect the falls of the elderly. *Pervasive Computing, IEEE*, 3(2):42–47, 2004.

[32] D. A. Socolinsky, A. Selinger, and J. D. Neuheisel. Face recognition with visible and thermal infrared imagery. *Computer Vision and Image Understanding*, 2003.

[33] D. A. Socolinsky, L. B. Wolff, J. D. Neuheisel, and C. K. Eveland. Illumination invariant face recognition using thermal infrared imagery. In *Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on*, volume 1, pages I–527. IEEE, 2001.

[34] L. Sweeney. k-anonymity: A model for protecting privacy. *International Journal on Uncertainty, Fuzziness and Knowledge-based Systems*, 2002.

[35] A. Treptow, G. Cielniak, and T. Duckett. Real-time people tracking for mobile robots using thermal vision. *Robotics and Autonomous Systems*, 54(9):729–739, 2006.

[36] P. L. Vora, J. E. Farrell, J. D. Tietz, and D. H. Brainard. Digital color cameras – 1 – response models. 1997.

[37] T. Winkler, A. Erdélyi, and B. Rinner. Trusteye. m4: Protecting the sensornot the camera. In *Advanced Video and Signal Based Surveillance (AVSS), 2014 11th IEEE International Conference on*, pages 159–164. IEEE, 2014.

[38] T. Winkler and B. Rinner. Trustcam: Security and privacy-protection for an embedded smart camera based on trusted computing. 2010.

[39] S. Xiu-Bao, C. Qian, G. Guo-Hua, and L. Ning. Research on the response model of microbolometer. *Chinese Physics B*, 19(10):108702, 2010.

[40] M. Yang, Q. Liu, T. Turner, and Y. Wu. Vital sign estimation from passive thermal video. In *Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on*, pages 1–8. IEEE, 2008.

[41] Y. Zhang, Y. Lu, H. Nagahara, and R.-I. Taniguchi. Anonymous camera for privacy protection. In *Pattern Recognition (ICPR), 2014 22nd International Conference on*, pages 4170–4175. IEEE, 2014.