

# Improving Infrared Images for Standoff Object Detection

---

Katherine Hanton<sup>1</sup>, Marcus Butavicius<sup>2</sup>, Ray Johnson<sup>3</sup> and Jadranka Sunde<sup>3</sup>

<sup>1</sup>School of Electrical & Information Engineering, University of South Australia, Adelaide, Australia

<sup>2</sup>Command, Control, Communications and Intelligence Division, Defence Science and Technology Organisation, Edinburgh, Australia

<sup>3</sup>Weapons System Division, Defence Science and Technology Organisation, Edinburgh, Australia

The ability to detect dangerous objects (such as improvised explosive devices) from a distance is important in security and military environments. Standoff imaging can produce images that have been degraded by atmospheric turbulence, movement, blurring and other factors. The number and size of pixels in the imaging sensor can also contribute to image degradation through under-sampling of the image. Establishing processes that enhance degraded or under-sampled infrared images so that objects of interest can be recognised with more certainty is important. Super-resolution image reconstruction and deconvolution methods are explored, as well as performance improvement measures.

*Keywords:* standoff detection, infrared imaging, super-resolution, performance improvement measure

## 1. Introduction

The detection and recognition of dangerous objects, in particular improvised explosive devices (IEDs) from a distance, is a formidable challenge as there is no fixed characteristic that can be exploited to distinguish such a device from the background.

The standoff detection of IEDs is further complicated by additional issues such as the need for identification at the maximum possible range and the possibility that devices may be concealed in otherwise innocent objects.

Detection technologies can be broadly divided into two groups: 1) explosives material detection and 2) detection of the device components. A key requirement is the ability to detect a threat at a safe standoff range for personnel and this generally requires an imaging system for detection of the device components.

We have investigated the capabilities of an infrared (IR) camera, which may be used in a variety of conditions to produce standoff images. Possible enhancements to acquired images, such as the use of super-resolution and deconvolution techniques, are discussed, as well as measures of image quality improvement that include both human and computer/algorithm assessment.

## 2. Standoff Imaging

Standoff detection involves detection and decision making within a nominal time at a distance. The following definition for standoff detection is provided in [11].

“Standoff explosive detection involves passive and active methods for sensing the presence of explosive devices when vital assets and those individuals monitoring, operating and responding to the means of detection, are physically separated from the explosive device. The physical separation should put the individuals and vital assets outside the zone of severe damage from a potential detonation of the device.”

A distinction can be made between “remote detection” and standoff detection [13]. In remote detection, the personnel performing explosives screening maintain a safe distance from the item being screened, but the screening equipment does not. In contrast, true standoff detection means that both personnel and equipment maintain a safe distance from any potential explosive device detonation.

The mass of explosive required to inflict equivalent damage increases approximately as the cube of the bomb-to-target distance [6]. This means that increasing the distance by a factor of two reduces the damage to the target by a factor of about eight. Thus, it is imperative to detect explosive devices at as great a distance as possible.

Device detection methods exploit the properties of the materials that the device is made of or by its shape. Imaging is a primary technique for cueing as devices are typically recognised by their shapes in passive thermal or active imagery.

### 3. Passive Infrared Imaging

All objects emit thermal IR radiation, depending on their temperature and emissive efficiency (emissivity). Infrared devices can operate in either passive or active modes. Passive modes use naturally occurring thermal radiation to form images, whereas in active mode, the device illuminates the area to create the image. Infrared radiation has a number of advantages over visible detection methods, since it may be possible to image through fog, smoke, rain and at night.

Images are formed where there is thermal contrast between the elements of a scene. Modern thermal cameras can comfortably resolve thermal contrasts of less than  $0.1^{\circ}\text{C}$ . Thermal contrast may also exist on the ground surface when objects are buried and IR imagery has been shown to produce some indication of the presence of buried landmines [7].

IR imaging has several drawbacks;

- effects such as climate, air currents and the presence of other radiation sources (e.g. the sun and other heat sources) need to be considered,
- it relies on the presence of thermal contrast to produce an image,
- the operator must identify objects by shape alone, and
- simple countermeasures may be effective as thermal or insulated materials can be used to conceal heat signatures.

A primary characteristic of staring array thermal imagers is that the detector arrays are generally very difficult and costly to manufacture

in large sizes. This means that most currently available thermal imagers produce significantly under-sampled images, with the pixel spacing being significantly less than the underlying image resolution. Studies have shown that this spatial under-sampling leads to aliasing in the imagery when it is shown on a display, and negatively impacts on observer performance [14].

The standoff imaging capabilities of such systems may possibly be significantly enhanced by reconstructing the underlying image using image processing techniques.

## 4. Image Reconstruction

Image reconstruction was performed in two steps; first by applying super-resolution, and then deconvolution.

### 4.1. Super-resolution

In [10], a range of sensor characteristics were identified for which there is a benefit from super-resolution reconstruction. Three regions (very beneficial, somewhat beneficial and no benefit) were identified for a range of sensors. For example, for a staring long wave (LW) IR detector system with 20 micron pixel spacing, the transition from a benefit to a no benefit occurs between F stops  $F/2$  and  $F/4$  (F stop is a ratio of the focal length of the lens to the diameter of the aperture). Many thermal imagers will benefit from super-resolution, as typical optical systems have low F numbers in order to maximise sensitivity.

Under-sampling produces aliased imagery, which results in a loss of subtle detail (high spatial frequency components). Super-resolution image reconstruction can increase spatial resolution without changing the design of the optics and the detectors. Sequences of low resolution images, with sub-pixel shifts between successive frames, are combined to form a higher resolution image. This is possible as relative motion between the scene and the IR focal plane array can be used to fill in the information between pixels, provided that the image motion is not fast enough to cause blurring in the images [8, 12].

A hybrid reconstruction scheme was proposed in [2] for super-resolution reconstruction, which

addresses the balance required between improving spatial resolution and keeping the computational time low. The scheme is referred to as iterative interpolation super-resolution and is based on interpolation and iterative improvements of the reconstructed image to generate a high resolution image from a sequence of low resolution frames. This method will form the basis of a proposed image processing technique.

#### 4.2. Deconvolution Methods

Deconvolution (or image deblurring), is the process of reconstructing or estimating the high spatial frequency components of an image when they have been degraded by the point spread function (PSF) of the imaging system.

The deconvolution problem is concerned with the separation of two signals; the image and the PSF of the imager, combined through convolution. Only estimates of the deconvolution result are possible as the problem is ill-posed, and therefore does not have a unique solution [3]. However, deconvolution generally results in sharper images that are perceived as being more realistic by humans, even if they contain signal processing artefacts.

If the camera characteristics are known exactly, it is possible to calculate the PSF, the two-dimensional spread of light that an instrument produces when its input is a point source. If this function can be determined, it is then a matter of computing its inverse function to achieve a representation of the original image. When the PSF is unknown, it may be possible to deduce it by trying different possible PSFs and assessing whether the image has improved. The accuracy of the approximation of the PSF will dictate the final result.

### 5. Experimental Work

In order to carry out some initial tests of the options for super-resolution and image enhancement, an experiment was conducted with targets laid out on a 4x5 m area of clear soil.



Figure 1. Test field.

A variety of targets were laid in a manner that would minimise shadow coverage caused by the equipment. The targets consisted of 2 metallic objects (1, 3), 2 resolution charts (2, 4) and a sheet of writing in fonts 30 to 72 (5) (see Figure 2).



Figure 2. Target layout.

#### 5.1. Image Acquisition

A Cedip Titanium 530 LW F/2 (mercury cadmium telluride – MCT) 320x256 IR camera of wavelength 7.7-9.3  $\mu\text{m}$  was used with a Jade LW 12 mm F/2 lens to acquire target images. The IR camera was mounted on a tripod which was placed on the tray of a utility vehicle. To achieve the image-to-image movement required in the image for super-resolution, the vehicle engine was left running, producing subtle vibrations to the tripod mounted camera. The vehicle was left idling for the duration of image capture

and a series of single frame images were taken successively. The image integration time was 0.35 ms so image blurring was minimal.

## 5.2. Image Processing

A manual image reconstruction was performed using Adobe Photoshop CS3, as per the method used in [9].

1. Each image from a sequence was enlarged from the original size of 320 x 256 to 1600 x 1280, using bicubic interpolation.
2. Each image was then sharpened using an unsharp mask filter with a 5 pixel radius, 0 threshold and 100% contrast.
3. Five images from the sequence were collated to form layers and stacked using the Photoshop script "align layers by content".

The resulting stacked image showed considerable resolution enhancement when compared to any of the original images. Since the noise in the stacked image was not excessive, a further enhancement was attempted using deconvolution to correct for the optics PSF. The image processing software package IRIS was used to implement the Van Cittert recursive algorithm [3] with an estimated Gaussian PSF. As a final step, the image was filtered to reduce fixed pattern noise and the highest spatial frequencies were then boosted using an unsharp mask filter.



Figure 3. Original image.

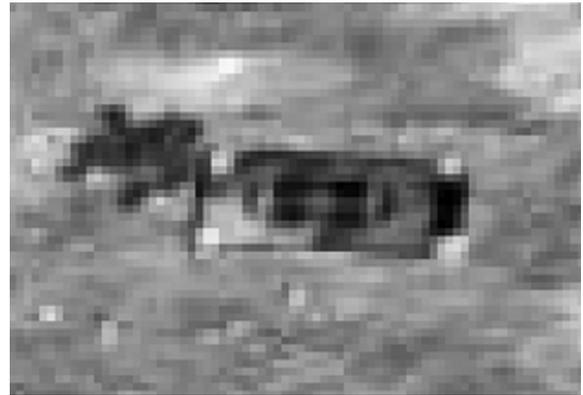


Figure 4. Original (zoomed-in) image.



Figure 5. Stacked image.



Figure 6. Van Cittert image.

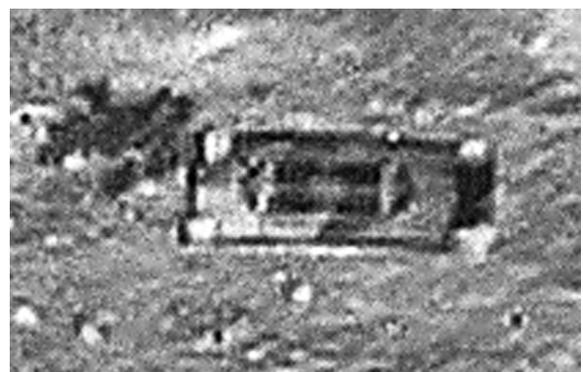


Figure 7. Final image.

The original image is shown in Figure 3, and a zoomed-in section of it is shown in Figure 4. Zoomed-in sections of the processed images are shown in Figure 5 to Figure 7. Figure 5 is the stacked image of five frames, and a Van Cittert deconvolution was applied to the image in Figure 6. The final image, after the application of a Gaussian blur and another unsharp mask, is shown in Figure 7.

The results displayed in the images in Figures 3–7 (zoomed-in section of images in Figures 4–7) show that there is a significant enhancement in the detail of the images using relatively straightforward processing and with image motion being based on accidental motion of the camera. Image processing to date has been performed manually, but an algorithm proposed in [2] is currently being modified to accept IR imagery for future automated image processing.

Image artefacts are enhanced by the processing, and noise is significant in the final image. The intention is to find a way to optimise the enhancement in order to help a human observer to recognise objects in the image with more certainty.

## 6. Measure of Image Improvement

Establishing performance improvement measures is the key to further development of image enhancement techniques. A widely used method in the image processing community is to simply present images for assessment by the reader. However, this method does not allow any improvements to be measured. We propose to measure image improvement by human observer assessment.

As a high resolution image is not available for comparison, the technique used was based solely on the presented image content. Other techniques have been investigated (see [1, 5]) including the Target Task Performance (TTP) metric (discussed in [3]), which is based on a measure of the information content of an image.

A pilot study was conducted with 16 participants interpreting IR images. Ten images were up-sampled from 320x256 to 1600x1280, and stacked as described earlier. Participants were presented with two images from the set (as detailed below) in succession.

- four targets were used in all images: soft drink can, mobile phone, plastic water bottle and glasses case,
- everyday clutter (rocks, grass, shrubs) was left in the field of view to replicate the real environment,
- targets were placed at different positions, to prevent observers from remembering their position from the previous scene.

The two images could be of any combination of the high resolution (HR) image and the raw IR image at the two distances. The image taken at 12m was always shown first, and the image taken at 4m was shown second. This constraint prevented the participants from using information from the more detailed image taken at a shorter viewing distance to interpret the longer distance image. All combinations (image resolution, scene and distance) occurred once every eight participants to control for interaction effects between the different variables.

At the start of the experiment, participants were presented with the four items from the scene. This stage was included to ensure that participants were familiar with the shapes of the items that they were searching for in the scene and that they had viewed the object from different angles. The images were presented on a laptop screen. For each image, the participants were asked to identify the four objects in the scene by pointing to them on the screen. The responses of the participants were recorded manually by the experimenter and the time from the presentation of the image to the final decision by the participants was recorded manually using a stopwatch.

Each of the four attempts by a participant to recognise one of the four items could be classified in two ways. The first was detection accuracy (i.e., whether the participant had detected an item from the target set) and the second was identification accuracy (i.e., whether the participant had correctly identified the image). For example, if a participant pointed to an item that was one of the four targets, but incorrectly identified this as another item in the set, this decision was classified as a ‘correct detection / incorrect identification’.

Probability of detection ( $P_d$ ) and probability of identification ( $P_i$ ) were assessed and presented in Table 1.

	$P_d$		$P_d \cap P_i$	
	12 m	4 m	12 m	4 m
Raw IR Image	0.22	0.50	0.16	0.25
HR Image	0.22	0.80	0.09	0.34

Table 1. Correct detection and identification ( $P_d$  and  $P_d \cdot P_i$ ).

## 7. Conclusions

Experimental work and preliminary results have shown considerable improvement in apparent image quality using a combination of super-resolution and resolution enhancement techniques and tools.

Results from the human observer performance indicate that an improvement was achieved with HR images for both  $P_d$  (0.5 compared to 0.8) and  $P_d \cap P_i$  (0.25 versus 0.34) at the 4m distance. However, at the 12m distance, the effect was reversed for  $P_d \cap P_i$  (raw: 0.16 versus HR: 0.09). This effect may be the result of clutter being enhanced in the 12m HR images.

Further work on the development of a super-resolution/image enhancement algorithm, and performance improvement measures are being explored, including other parameters affecting human observer performance.

## References

- [1] G. AVRIAM, S. ROTMAN, Evaluating the Effect of Infrared Image Enhancement on Human Target Detection Performance and Image Quality Judgement. *Optical Engineering*, 38(8), 1433–1440, 1999.
- [2] V. S. BANNORE, Computationally Efficient Iterative-interpolation Reconstruction Scheme. PhD Thesis. University of South Australia, 2008.
- [3] A. BENNIA, S. M. RIAD, Filtering Capabilities and Convergence of the Van Cittert Deconvolution Technique. *IEEE Transaction on Instrumentation and Measurement*, 41(2), p. 246–250, 1992.
- [4] J. H. CHA, E. JACOBS, Super-resolution Reconstruction and its Impact on Sensor Performance. *Proc. of SPIE 5784*, p. 107–113, 2005.
- [5] T. D. DIXON, E. F. CANGA, J. M. NOYES, T. TROSCIANKO, S. G. NIKOLOV, D. R. BULL, C. N. CANAGARAJAH, Methods for the Assessment of Fused Images. *ACM Transaction on Applied Perception*, 3(3), p. 309–332, 2006.

- [6] S. DUA, B. HALL, Standoff Detection of Improvised Explosive Devices: A Technological Review. DSTO-TR-2084, 2007.
- [7] K. HANTON, C. ABEYNAYAKE, Z. AYUBI, J. SUNDE, Phenomenology of Infrared Imaging. DSTO-TR-xxxx, in preparation.
- [8] R. C. HARDIE, K. J. BARNARD, J. G. BOGNAR, E. E. ARMSTRONG, E. A. WATSON, High Resolution Image Reconstruction from a Sequence of Rotated and Translated Frames and its Application to an Infrared Imaging System. *Opt. Engineering*, 37(1), p. 247–260, 1997.
- [9] B. KANEFSKY, T. J. PARKER, P. C. CHEESEMAN, Super-resolution Results from Pathfinder IMP. In: *29<sup>th</sup> Lunar and Planetary Science Conference*, Houston, Texas, 1998.
- [10] K. KRAPELS, S. YOUNG, G. HOLST, Characteristics of Infrared Imaging Systems which benefit from Super-resolution Reconstruction. Infrared Imaging Systems: Design, Modeling and Testing XVII. *Proceedings of SPIE*, 6207, 620703, 2006.
- [11] NATIONAL RESEARCH COUNCIL, Existing and Potential Standoff Explosives Detection Techniques. The National Academies Press, 2004.
- [12] S. C. PARK, M. K. PARK, M. G. KANG, Super-Resolution Image Reconstruction: A Technical Overview. *IEEE Signal Proc. Magazine*, 20(3), p. 21–36, 2003.
- [13] J. STORM, T. GEISLER, Tripwire Detection Using Polarimetric IR. *Proceedings of SPIE 4038*, p. 253–260, 2000.
- [14] R. VOLLMERHAUSEN, R. DRIGGERS, B. O’KANE, Influence of Sampling on Target Recognition and Identification. *Optical Engineering*, 38(5), p. 763–772, 1999.

Received: June, 2009  
Accepted: April, 2010

Contact address:

Katherine Hanton  
School of Electrical & Information Engineering  
University of South Australia  
Mawson Lakes, Adelaide  
Australia  
e-mail: Katherine.Hanton@dsto.defence.gov.au

Marcus Butavicius  
Command, Control, Communications  
and Intelligence Division  
Defence Science and Technology Organisation  
PO Box 1500, Edinburgh SA 5111  
Australia  
e-mail: Marcus.Butavicius@dsto.defence.gov.au

Ray Johnson  
Weapons System Division, Defence Science  
and Technology Organisation  
PO Box 1500, Edinburgh SA 5111  
Australia  
e-mail: Ray.Johnson@dsto.defence.gov.au

Jadranka Sunde  
Weapons System Division, Defence Science  
and Technology Organisation  
PO Box 1500, Edinburgh SA 5111  
Australia  
e-mail: Jadranka.Sunde@dsto.defence.gov.au

---

KATHERINE HANTON received her degree in B.App.Sci Scientific Imaging (Honours) from RMIT, in Melbourne, Australia in 2005. She started as a student in 2004 at the Defence Science and Technology Organisation (DSTO), Adelaide in 2005. She is currently conducting her PhD research in the area of Infrared image improvement for human operator performance gains.

---

---

MARCUS BUTAVICIUS received his Honours degree in Psychology from the University of Adelaide in 1996, and PhD in 2002. He joined the Defence Science and Technology Organisation (DSTO) in 2001 and his interests include human computer interaction approaches to the study of document visualization, information security, text analysis and image processing. Since 2004 he has been a Visiting Research Fellow at the School of Psychology at the University of Adelaide.

---

---

RAY JOHNSON retired from the Defence Science and Technology Organisation (DSTO) in 2001 and has carried out a variety of research tasks in DSTO as a contractor since 2003. He is a physicist with current interests in electro-optics, image and signal processing, biometrics and experiment design.

---

---

JADRANKA SUNDE received her degree in Mathematics from the University of Zagreb, Croatia in 1984, and her PhD from the University of Adelaide, Australia in 1997. She joined the Defence Science and Research Organisation (DSTO) Australia in 1997. Her main fields of interest are object detection and human system integration.

---

