

Super-resolution of Infrared Images: Does it Improve Operator Object Detection Performance?

Katherine Hanton¹, Jadranka Sunde², Marcus Butavicius³
and Nicholas R. Burns³

¹School of Electrical & Information Engineering, University of South Australia, Adelaide, Australia

²Weapons System Division, Defence Science and Technology Organisation, Edinburgh, Australia

³School of Psychology, University of Adelaide, Adelaide, 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. In this paper, super-resolution image reconstruction and deconvolution methods are explored, with an emphasis on quantifying and understanding human operator detection performance.

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.

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, are assessed to establish if human operator detection performance may be improved. In addition, psychometric tests are used to determine if any enhancements due to super-resolution can be related to the innate visual processing characteristics of the various test participants.

2. Standoff Detection and Passive Infrared Imaging

Standoff detection involves detection and decision making within a nominal time at a safe distance. A distinction can be made between “remote detection” and standoff detection [15]. In remote detection, the personnel performing 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.

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

All objects emit thermal IR radiation, depending on their temperature and emissivity. Infrared radiation has a number of advantages over visible detection methods, since it is possible to image through fog, smoke, rain and at night. Images

are formed where there is thermal contrast between the elements of a scene and modern thermal cameras can comfortably resolve thermal contrasts of less than 0.1°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 [6].

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) can produce a confusing picture,
- It relies on the presence of thermal contrast to produce an image and this may not always be available,
- The operator must identify objects by shape alone, and
- Simple countermeasures may be effective as reflective 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. This spatial under-sampling leads to reduced resolution and to aliasing in the imagery, which negatively impacts on observer performance [17].

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

3. Image Enhancement

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 $20\ \mu\text{m}$ pixel spacing, the transition from “somewhat beneficial” to “no benefit” occurs between F stops $F/2$ and $F/4$. Many thermal imagers will benefit from super-resolution, as typical optical systems have

low F numbers in order to maximise sensitivity and the resulting systems are generally under-sampled. The IR system used for this work has $30\ \mu\text{m}$ pixel spacing and an F stop of $F/2$, therefore falling within the “very beneficial” region for super-resolution processing benefit.

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. Fixed or random motion in both image axes can be used to provide the sub-pixel shifts, provided that the image motion is not too fast to cause blurring in the images [8, 13] and that the motion covers the full sub-pixel region.

3.1. Image Acquisition and Processing

Test images were taken with a 320×256 pixel array IR camera operating at $7.7\text{--}9.3\ \mu\text{m}$ and with a 12 mm $F/2$ lens. The IR camera was mounted on a tracking mount which was placed on the tray of a utility vehicle. Image sequences were taken with the camera mounted on a tracking head that was traversing slowly in both axes to provide variations in the raw sub-sampled images. The image integration time was 0.24 ms, so blurring within an image was minimal.

Image acquisition was carried out as follows;

- Single targets were selected from six objects of similar size and shape (shown in Figure 1),
- Each target was placed in the field of view of the imager at three distances 2.5 m, 5 m, 10 m (equates to approximately 10, 5 and 2.5 pixels respectively, across the shortest dimension of the objects),
- Each object was placed in a unique position, so that observers could not remember target locations when viewing different images,
- Outdoor scenes were used, with variation in natural features (e.g., trees, foliage and ground cover) and man-made objects (e.g., vehicles). Everyday clutter was left in the field of view to replicate distractor objects present in the real environment, with the following levels of clutter:

- Low – relatively clear ground with some leafy ground cover, but without any clutter objects.
- Medium – more cluttered area, as well as some clutter objects such as rocks, bricks and pine cones.
- High – cluttered area with man-made objects already present in the scene (two car wrecks and two rolls of barbed wire).



Figure 1. Target objects.

The image reconstruction was based on the method used in [9]. Image registration and combination was performed in Registax [3], and wavelet filters were also applied in Registax to enhance high spatial frequency components (see Figure 2 for a sample of images). Identical processing was applied to all image sequences.

1. Each raw image from a sequence was enlarged from the original size of 320x256 pixels to 1600x1280 pixels, 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. 100 single images from the sequence were registered and stacked in Registax, using a single reference point.
4. Wavelet filters at varying scales were then applied to the final stacked image and the

outputs combined to further enhance components with high spatial frequency (this is also a Registax feature).

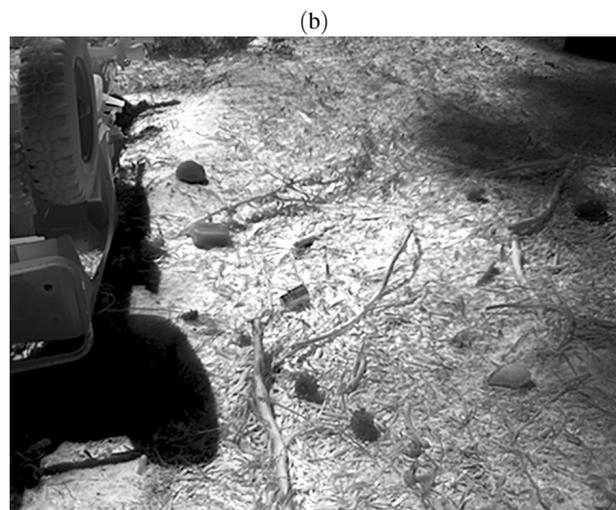
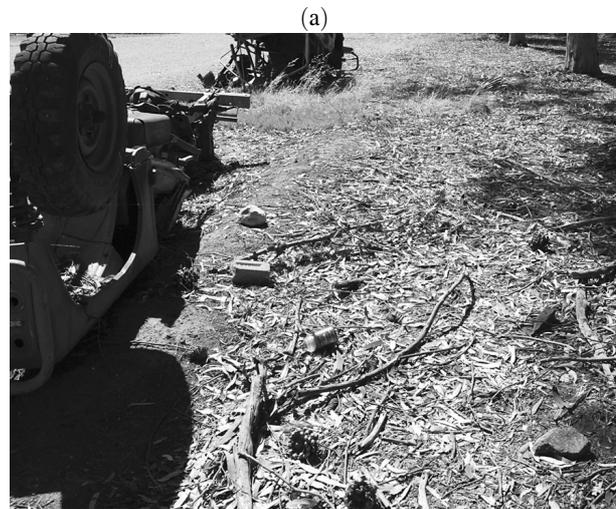


Figure 2. Example of experiment scene with target layout, shown in visible wavelength (a) and IR (b).

4. Human Observer Assessment

Given the substantial variation in performance results between individuals in the pilot study [7], which indicated that there could possibly be many factors which affect human operator detection performance, the main study included examination of individual factors that may affect detection performance. Consequently, an experiment was designed to examine correlations between visual tests and task-specific performance of the human operators, to determine if there were any performance indicators that

could show which observers were best able to process super-resolved images.

This assessment consisted of three stages:

1. Demographic data collection,
2. Detection performance test, and
3. Visual-based psychological tests.

4.1. Demographic Data Collection

Participants were asked questions regarding their age, eye or vision problems, experience with working with IR images, if they played computer games and their education level. There were also questions on their impressions of the study and whether they had any difficulties with the task.

4.2. Detection Performance Test

Eighteen images were utilised in the experiment. For half of the participants, they were shown in the original low resolution (LR) versions of these images while for the other half of participants, the super-resolved (SR) versions were presented. Speed, accuracy (for detection and identification) and confidence ratings were recorded. The study was designed with the following variables for the images and their presentation:

- Distance (2.5m, 5m and 10m),
- Clutter (low, medium and high),
- Single target present or absent (9 images of each),
- Resolution (LR/SR, 9 images of each),
- Clear sunny conditions and
- Maximum presentation time of 60 secs.

The visual performance task involved both target detection (i.e. location) and identification of the selected targets in the various backgrounds. Before the testing began, participants were shown images of the eight man-made objects used in the experiment.

Participants were asked to look at each presented stimulus image and, if one of the six objects was identified in the scene, they were asked to drag the number associated with the object from the icon list onto the position in

the image where the object was located. When the participant was satisfied with this decision, they clicked on the ‘Detection completed, move to next’ button. An example of an object (number 5 placed on the image) being located and identified is presented in Figure 3. If no object was identified in the image, participants were asked to indicate this by selecting the ‘No Targets’ button.



Figure 3. Observer attempt to locate target object.

After either the ‘Detection completed, move to next’ or ‘No targets’ buttons were pressed, or if the 60 second time limit had expired, the stimulus image was replaced with rating scales. Participants were prompted to provide a confidence rating on two 1 – 100 point sliding scales for their confidence in the object identified and for its location, as well as an additional rating of image quality. If the participants selected the ‘No targets’ response, they were presented with a scale to rate their confidence in their detection and image quality responses only.

5. Visual-based Psychological Tests

Participants were also required to complete seven perceptual and cognitive tasks to examine their general visual processing and problem solving ability, and to examine the relationship with performance on the detection performance test.

These tests, listed in the order they were presented to the participants, were:

1. Contrast Sensitivity Test [14],
2. Freiburg’s Visual Acuity Test (FrACT) [1],
3. Inspection Time (IT) [12],

4. Useful Field of View (UFOV) [2],
5. Global Precedence Test [11],
6. Raven’s Advanced Progressive Matrices (RAPM) [4], and
7. Mental Rotations Test [16].

The selection of cognitive and perceptual tests was based on the specific abilities that may be tested in the detection and identification of target objects in IR images. Each test deals with a part of perceptual/cognitive process that could be relevant to observer performance in the IR detection task.

5.1. Contrast Sensitivity Test

Contrast sensitivity is the visual ability to see objects that may not be outlined clearly or that do not stand out from their backgrounds. The Pelli-Robson Contrast Sensitivity test measures contrast sensitivity using a single large letter size, with contrast varying across groups of letters. The participant was asked to stand 1 m away from the chart (Figure 4) and read the letters aloud from left to right until they could no longer see any letters.



Figure 4. Pelli-Robson contrast sensitivity chart.

5.2. Freiburg Visual Acuity Test (FrACT)

This computer-based test measures visual acuity. Landolt-Cs (see Figure 5) were presented in one of eight orientations and in decreasing size. The task of the participant was to determine which of eight possible directions the gap in the Landolt-C was facing by pressing the corresponding key once the target figure was presented to them on the screen.



Figure 5. Landolt-C stimulus.

5.3. Inspection Time

Inspection Time (IT) is a measure of the speed of information processing. In this test, stimuli consist of two vertical lines (one shorter) joined at the top by a horizontal line (see Figure 6). A warning cue (a small + in the centre of the screen), precedes the target figure, which is then immediately replaced by a visual mask. The participant is instructed to indicate on which side of the target figure the shorter line appears and the decision time is measured. The stated emphasis is on accuracy, and not speed of responding.

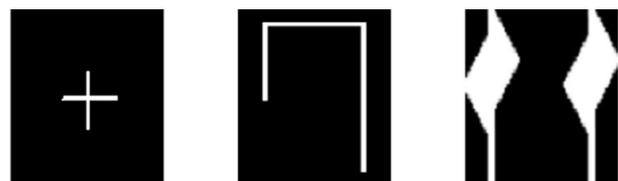


Figure 6. Pre-stimulus frame (left), stimulus (centre) and post-stimulus mask (right).

5.4. Useful Field of View (UFOV) Test

The useful field of view is the visual area from which information can be acquired within one eye fixation. Useful Field of View (UFOV), tests assess the speed of visual processing under increasingly complex task demands. The participant must detect, identify and localise briefly presented targets. For the purpose of this study, only Subtest 3, Selective Attention, was used. The participant has to identify a target presented in a centrally located fixation box that is presented for varying lengths of time, and also to localise a simultaneously presented target displayed in the periphery of the computer monitor (Figure 7). The target displayed in the periphery is embedded with distractors, making the task more difficult.



Figure 7. UFOV stimuli embedded with distractors.

This process of tracking the perceptual threshold is continued until a stable estimate of 75% correct is calculated. The length of time necessary to obtain the stable measure depends upon the consistency of the participant’s responses. The participants were informed that these tests were not a reaction time test, and it was the accuracy of their response that counted.

5.5. Global Precedence Test

This test examines a person’s ability to differentially respond to local and global features of an image. As can be seen in Figure 8, the global feature of the letter ‘H’ is composed of local features consisting of the letter ‘S’. On a different trial, the participant was required to identify either the global form (e.g. in this case, the ‘H’) or the local form (e.g. the ‘S’).

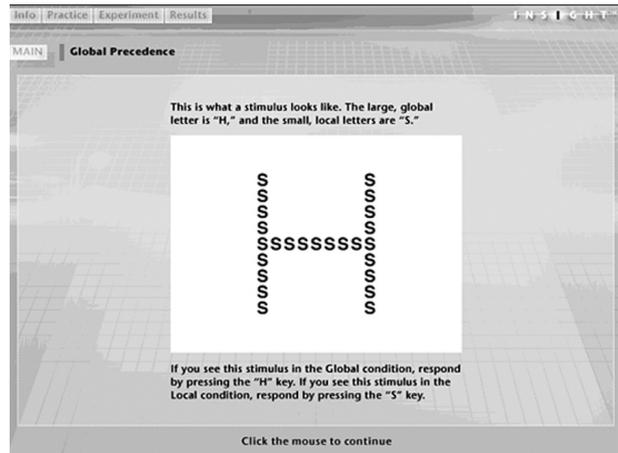


Figure 8. Global precedence stimulus.

5.6. Raven’s Advanced Progressive Matrices (RAPM)

This test looks at inductive or analytic reasoning and estimates the ability of the subject to solve problems without relying on an explicit base of knowledge derived from previous experience. For our purposes, RAPM was considered appropriate because it measures the ability to determine visual patterns; a task that may relate to

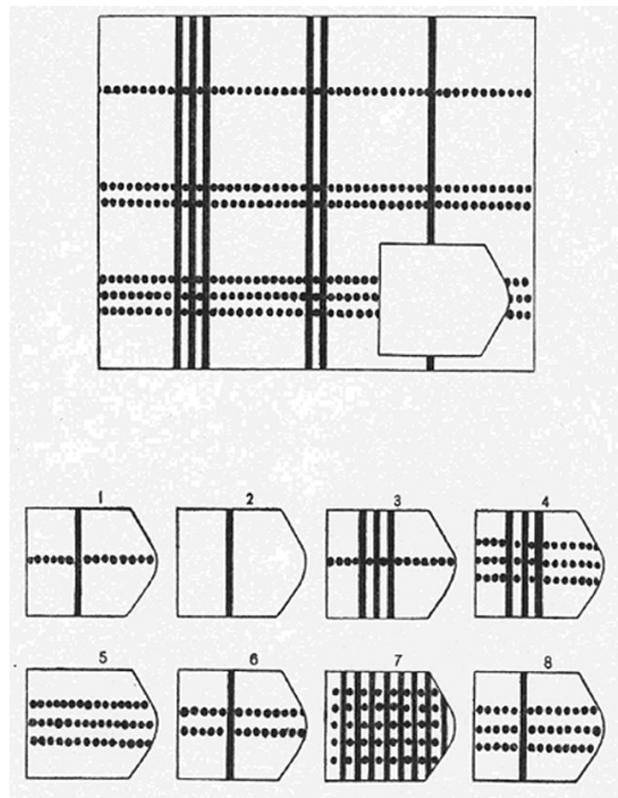


Figure 9. Example of a matrix.

an observer’s ability to determine a contrasting target object amongst a repeating pattern of visual clutter. The participant needs to determine which of the eight possible alternatives shown under the matrix fits into the missing space of the matrix (see Figure 9).

5.7. Mental Rotations Test

The Vandenberg-Kuse test of mental rotations, which utilises the rotation of cubes, is a common method of testing spatial comprehension. In this test, participants are presented with a number of figures (see Figure 10). The task is to decide which two figures depict exactly the same object as the ‘target’ figure, even though the object is viewed from different angles in these figures.

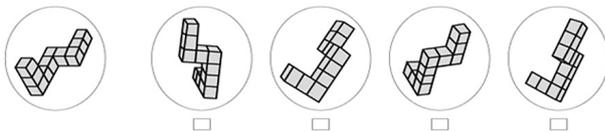


Figure 10. Mental rotation test example.

6. Analysis and Results Interpretation

The following data has been analysed:

- The overall performance of operators (i.e., the number and percentage of correct and incorrect detections and identifications) for SR and LR images;
- Effects of variables on correct and incorrect detections and identifications (e.g., different clutter environments, varying distances, etc.); and
- The relationship between performance on the cognitive and perceptual performance indices and on the detection performance test.

A total of 74 participants took part in the study, 11 of whom were female (37 viewing SR images and 37 viewing LR images). Participant ages were split into three ranges: <30 years (16 participants), 30 – 50 years (41 participants) and >50 years (17 participants).

Participants were also asked to indicate if they had prior experience with viewing/working with IR images/imaging. This was divided into four categories:

- “a lot of experience” (6 participants),
- “some experience” (12 participants),
- “little experience” (14 participants), and
- “no experience” (42 participants).

The highest level of education the participants had completed was also recorded:

- Secondary school (4 participants),
- Tertiary (22 participants),
- Post-graduate (47 participants).

6.1. Probability of Detection

The results were averaged across 74 participants for both LR and SR IR images and are shown in Figure 11 to Figure 13.

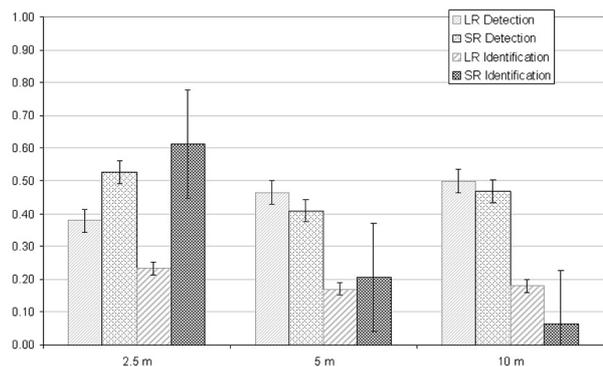


Figure 11. Correct detection and identification.

Results in Figure 12 indicate that using SR images produced limited detection performance improvement. A noticeable improvement was

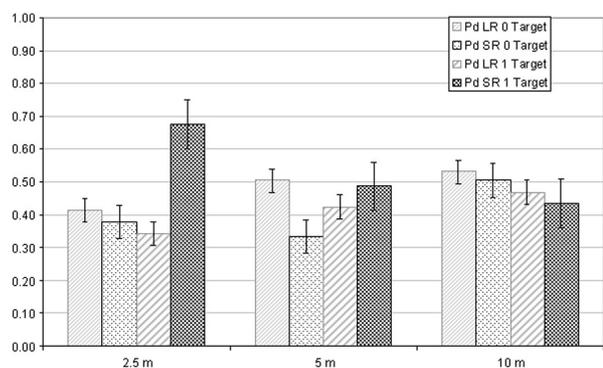


Figure 12. Correct detection (1 or 0 target) for SR and LR at 2.5 m, 5 m and 10 m.

achieved with SR images for detection (0.53 versus 0.38) and identification (0.61 versus 0.23) at the 2.5 m distance only.

Figure 12 and Figure 13 show the results for detection (with or without target present). An improvement in SR detection performance was generally achieved when there was one target present in the scene, with reasonable improvement at 2.5 m (0.68 versus 0.34), low clutter level (0.98 versus 0.75) and high clutter level (0.46 versus 0.28). However, when there was no target present in the scene, performance with LR images tended to be better than performance with SR images. This may have been a result of artefacts in the scene being enhanced by the SR processing, leading observers to believe these artefacts to be targets of interest.

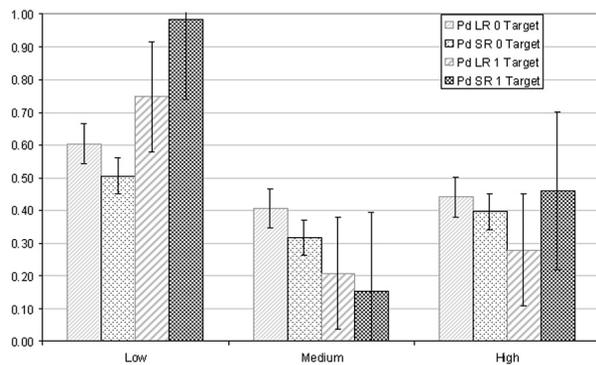


Figure 13. Correct detection (1 or 0 target) for SR and LR at low, medium and high clutter levels.

6.2. Detection Performance vs. Visual-based Psychological Tests

Pearson product-moment correlation coefficients between detection performance and the perceptual / cognitive tests are shown in Table 1 in the order of highest to lowest value.

Participant age was found to have the highest correlation with detection performance as seen in Table 1.

Individual observer detection performance was plotted against their performance in some of the perceptual and cognitive tests: mental rotations, RAPM, IT and global precedence (local condition only) and is shown in Figure 14 to Figure 17.

Age	-0.22
Mental rotation	0.18
RAPM	0.17
IT	-0.15
Global precedence (Local condition)	-0.13
FrACT Dec. VA	0.09
FrACT logMAR	0.06
UFOV	0.02
Global % correct	0.01

Table 1. Correlations between performance on the detection task and the psychological tests (N = 74).

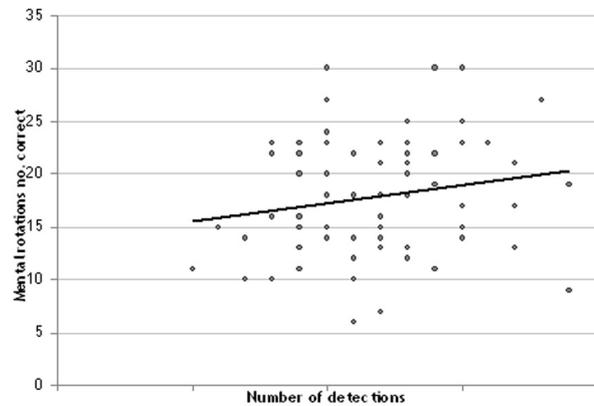


Figure 14. Mental rotations (number correct) vs. number of correct detections.

As depicted in Figure 14 and Figure 15, both mental rotations and RAPM have a positive correlation with detection performance with a small effect size [6]. Therefore, participants

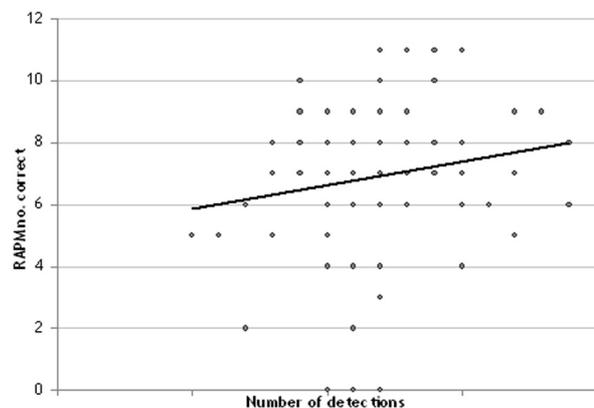


Figure 15. Ravens advanced progressive matrices (number correct) versus number of correct detections.

who scored higher in those tests were more likely to be successful in the detection task.

As seen in Figure 16, a longer inspection time tended to decrease detection performance. The result in Figure 17 was unexpected, as it suggests that increased detection performance was associated with poorer performance in the local condition of the global precedence task.

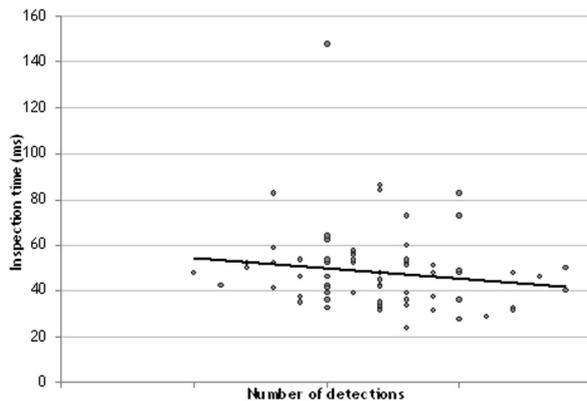


Figure 16. Inspection time (ms) versus number of correct detections.

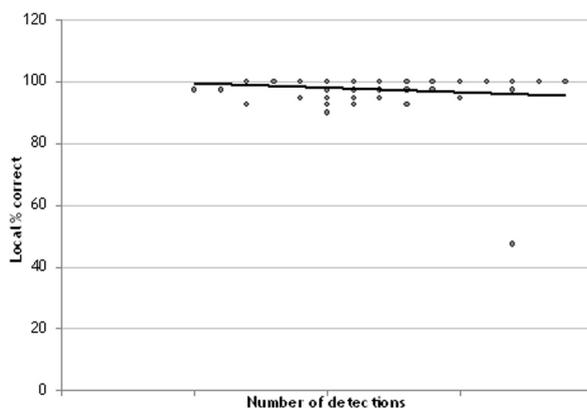


Figure 17. Global precedence (local condition only) versus number of correct detections.

7. Conclusions

Results from this study have shown that there is a general improvement in detection performance when using SR images. This improvement was evident when there was one target present in the scene at closer distances and with low and high (but not medium) clutter scenes. However, when no target was present, performance with LR images tended to better performance than with SR images. It is suggested

that the artefacts created by the SR process may have been mistaken by the participants as targets of interest. There was a small correlation between performance on the detection task and both the mental rotations task and the RAPM. This suggests that the relationship between an operator's ability to find targets in an IR image may relate to higher level visual-based problem solving abilities. Further analysis of results is required to fully understand which parameters affect detection performance, and to take into account all the other factors recorded such as the image quality data, confidence ratings, and experience levels.

References

- [1] M. BACH, The Freiburg Visual Acuity Test – Automatic Measurement of Visual Acuity. *Optometry & Vision Science*, 73, 49–53, 1996, www.michaelbach.de/fract/.
- [2] K. BALL, D. L. BEARD, D. L. ROENKER, R. L. MILLER, D. S. GRIGGS, Age and Visual Search: Expanding the Useful Field of View. *Journal of the Optical Society of America*, 5, p. 2210–2219, 1988.
- [3] C. BERREVOETS, Registax V5 software for Alignment/stacking/processing of Images, 2009, <http://www.astronomie.be/registax/download.html>.
- [4] D. A. BORS, T. STOKES, Raven's Advanced Progressive Matrices: Norms for First Year University Students and Development of a Short Form. *Educational & Psychological Measurement*, 58(3), 1998.
- [5] J. COHEN, Statistical Power Analysis for the Behavioral Sciences (2nd edition). Lawrence Erlbaum Associates, Hillsdale, N. J., 1988.
- [6] K. HANTON, C. ABEYNAYAKE, Z. AYUBI, J. SUNDE, Phenomenology of Infrared Imaging. DSTO-TR-2303, 2009.
- [7] K. HANTON, M. BUTAVICIUS, R. JOHNSON, J. SUNDE, Improving Infrared Images for Standoff Object Detection. *31st Int. Conf. on Inf. Tech. Interfaces*, (ITI June 2009), p. 641–646.
- [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: *29th Annual Lunar and Planetary Science Conf.*, Texas, 1998.

- [10] K. KRAPELS, S. YOUNG, G. HOLST, Characteristics of Infrared Imaging Systems which benefit from Super-resolution Reconstruction. *Infrared Imaging Systems: Design, Modelling and Testing XVII*.
- [11] J. MILLER, Global Precedence in Attention and Decision. *Journal of Experimental Psychology: Human Perception and Performance*, 7(6), 1161–1174, 1981.
- [12] T. NETTELBECK, M. LALLY Inspection Time and Measured Intelligence. *British Journal of Psychology*, 67, p. 17–22, 1976.
- [13] S. C. PARK, M. K. PARK, M. G. KANG, Super-resolution Image Reconstruction: A Technical Overview. *IEEE Signal Proc. Magazine*, 20(3), 21–36, 2003.
- [14] D. G. PELLI, J. G. ROBSON, A. J. WILKINS, Designing a New Letter Chart for Measuring Contrast Sensitivity. *Clinical Vision Sciences*, 2, p. 187–199, 1988.
- [15] J. STORM, T. GEISLER Tripwire Detection Using Polarimetric IR. *Proceedings of SPIE 4038*, 253–260, 2000.
- [16] S. G. VANDENBERG, A. R. KUSE, Mental rotations: a group test of three-dimensional spatial visualization. *Percept. Motor Skills*, 47, p. 599–604, 1978.
- [17] R. VOLLMERHAUSEN, R. DRIGGERS, B. O'KANE, Influence of Sampling on Target Recognition and Identification. *Optical Engineering*, 38(5), p. 763–772, 1999.

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.

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 Defence Science and Research Organisation (DSTO) Australia in 1997. Her main fields of interest are object detection and human system integration.

MARCUS BUTAVICIUS received his Honours degree in Psychology from the University of Adelaide in 1996, and PhD in 2002. He joined 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.

Associate Professor NICHOLAS BURNS'S (PhD, Adelaide 1998) main teaching area is in psychological research methodology, specifically statistical methods; individual differences, especially current research in human intelligence; and psychological assessment. His research interests are largely in the area of differential psychology. He has published his research in the leading international journals including *Intelligence, Personality and Individual Differences, Neuropsychology, and Behaviour Research Methods*.

Received: March, 2010

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

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

Marcus Butavicius
 School of Psychology
 University of Adelaide
 Adelaide
 South Australia, 5005
 e-mail: Marcus.Butavicius@adelaide.edu.au

Nicholas R. Burns
 School of Psychology
 University of Adelaide
 Adelaide
 South Australia, 5005
 e-mail: Nicholas.Burns@adelaide.edu.au