

# Fusion of Low Bit-Depth Images for Battle Damage Indication

**Jason Ralph**

Department of Electrical Engineering & Electronics  
University of Liverpool  
Liverpool, U.K.  
jfralph@liverpool.ac.uk

**Nigel Stocks**

School of Engineering  
University of Warwick  
Coventry, U.K.  
n.g.stocks@warwick.ac.uk

*Abstract – This paper addresses the problem of extracting information from an infrared weapon seeker in the terminal phase of a guided weapon engagement. The challenge is to obtain enough information (in the form of images) immediately prior to impact to allow the impact point to be verified. The introduction of a radio data link onto a weapon allows information to be transmitted back to a remote operator, but this data link will have a limited data rate due to finite bandwidth and other operational restrictions. The main problems are that the impact will destroy the seeker and this is very likely to disrupt the transmission of the data back to the operator. Any latency, introduced by pre-processing, compression or encryption of the images, will have a significant effect on the information that can be extracted. This paper shows that low bit-depth images can be fused to reconstruct an image of the weapon's impact point and that noise in the low-bit images increases the information content of the reconstructed images – an example of suprathreshold stochastic resonance.*

**Keywords:** Image fusion, image registration, image reconstruction.

## 1 Introduction

In the seconds immediately prior to impact, a seeker mounted on a guided weapon will produce a large amount of data that could be useful in determining the effectiveness of the weapon strike – battle damage indication (BDI). Images could be transferred via a data link to a remote operator (such as an airborne weapons officer or forward air controller). However, image transfer via a secure military data link can be somewhat slow. There are safety and security issues that require a large amount of two-way communication between transmitting and receiving data links. The ability to reduce the data required to represent the images is therefore crucial to the use of such a system for BDI. Unfortunately, infrared images tend to be much noisier than visible band images and standard encoding algorithms will tend to encode the noise together with the signal/image content, meaning that the transmission of useful information will be reduced still further.

Infrared cameras based on CMOS electronics have some very interesting and useful properties that can be used to assist in image encoding for BDI. Many cameras are limited by the ability to take the analogue signal (photo-current) from each pixel in turn and to digitise it – quantisation. In CMOS systems, the size of the imaging area can be reduced electronically to allow smaller images to be processed faster – thereby increasing the effective frame rate. This process, called ‘windowing in’, has some disadvantages since smaller areas reduce the image content and shorter pixel integration times will tend to produce noisier images [1-3]. Rather than changing the number of pixels being processed, this paper considers reducing the number of bits used for each pixel. In the seconds immediately prior to impact, a faster frame rate with fewer bits per pixel might be advantageous, since it is unlikely that all the bits could be transmitted via the data link before the system is destroyed.

The aim of this paper is to show that reduced bit images can be used in this time-critical situation and that the additional noise present in the imagery can be advantageous to image reconstruction rather than disruptive. This is a type of Stochastic Resonance (SR) – known as Suprathreshold Stochastic Resonance (SSR) – and is an effect wherein the information transmitted increases as more noise is added to the signal (over a certain range of the signal-to-noise ratio).

## 2 Suprathreshold Stochastic Resonance

Stochastic Resonance has been an active area for research for more than twenty-five years [4]. SR occurs where the response of a nonlinear system to a particular input signal is improved by the addition of noise. The simplest illustrative example of SR is a bi-stable dynamical system – i.e. a system with two possible stable states. When the noise level is low, the state of the system is perturbed around one of these stable states and only occasionally undergoes a transition to a region around the other stable state. This transition from one stable state to another will normally have some characteristic time scale – the escape time. The escape time is dependent upon the level of the noise, increasing the level of the noise reduces the escape time. If a small periodic input signal is applied, and the level of the signal on its own is too low to drive the system

from one stable state to another, tuning the noise level so that the escape time matches the period of the signal tends to cause the transition to synchronise to the input signal. The synchronisation and its dependence on the noise level is a characteristic of SR – reducing the noise level reduces the synchronisation and tends to lower the output signal-to-noise level. This is normally referred to as sub-threshold stochastic resonance because the system is only weakly perturbed around the stable states and the typical variations due to noise are much smaller than the threshold that separates the two stable states.

A more recent variant of the usual (sub-threshold) stochastic resonance occurs when the signal alone is far larger than the threshold and can initiate jumping between the stable states. This is referred to as Suprathreshold Stochastic Resonance (SSR). One of the simplest forms of nonlinear system to exhibit this behaviour is an array of simple one-bit quantisers or other thresholding devices [5]. The idea for SSR is that there are a number of quantisers/channels – each with the same threshold level – and a random (i.e. non-periodic) input signal, which is fed into all the available quantisers/channels. If the instantaneous level of the input signal is far below the threshold level, then the output of all of the quantisers will be in state ‘0’. On the other hand, if the instantaneous level of the input signal is far above the threshold level, then the output of all of the quantisers will be in state ‘1’, and if the signal is precisely at threshold then approximately 50% of the quantisers will be ‘0’ and the others in state ‘1’. Consequently, the number of quantisers that are in state ‘1’ is monotonically dependent on the instantaneous value of the signal and, hence, this acts as a (noisy) ‘code’ of the original signal. However, the level of the noise also plays a crucial role. No noise means all the quantisers will either be in state ‘0’ or all in state ‘1’, and hence the system can transmit a maximum of 1-bit of information per sample. Conversely, for very large noise (much larger than the signal) there will always be an approximate 50-50 split between devices in the ‘1’ and ‘0’ state and this will transmit no information. Hence, one can infer there will be an optimal level of noise that maximises the information.

The ‘useful’ information in this case is measured by the mutual information between the input signal (the scene used to generate the simulated images) and the output signal (the reconstructed/fused images). SSR has a peak in the information flow even when the signals are suprathreshold whilst conventional SR has a no such peak. This model for multiple thresholded channels has been used successfully for auditory signal processing in the presence of noise for cochlear implants [6,7]. In the present situation, the noisy signals are a set of one-bit images generated in the period immediately prior to impact. Each of the images in a sequence represents one of the quantiser channels. The images are aligned and then fused to generate an improved image of the impact point.

### 3 Battle Damage Indication (BDI)

In any military engagement, there is a clear requirement to limit the use of firepower – to prevent collateral damage, and to reduce the risk of friendly fire or excessive casualties. However, the use of firepower does carry risks and the ability to assess the effectiveness of any weapon strike is essential in minimising these risks. Battle Damage Indication (BDI) is related to the location of the impact on the target and the known vulnerability of the target. The assessment of the effectiveness of a weapon strike is the higher-level process of Battle Damage Assessment (BDA). This paper is concerned with using an imaging seeker mounted on the weapon to determine its impact point. The problem is that the weapon and seeker will not survive the impact and any information that is generated must be extracted as quickly as possible and then be transmitted to a remote operator. Conventional image compression techniques take time to process the images and are likely to introduce significant latency into the imaging/communication system. The other problem to be addressed is the finite frame rate of the imager – a seeker may move a significant distance between frames and the impact may not coincide with an image frame. One approach is to increase the frame rate of the seeker – at the expense of a higher capacity communication link between the weapon and the operator. However, such high-capacity communication systems are likely to be expensive and would require a very large bandwidth. This paper demonstrates that useful BDI information can be extracted very close to the impact point using one-bit images, which minimise the need for a high-capacity communication system and reduce the requirements for image compression before transmission.

#### 3.1 Infrared Seeker Model

The infrared seeker model uses basic physics for the imager and satellite imagery to provide a background scene over which a simple three-bar target is placed. The model is relatively simple to understand but it contains representative noise and structural properties [1]. The imagery includes non-uniform pixel gain-offset response and dead/saturated pixels – caused by inactive pixels or short-circuited read out circuitry. The imager model contains a number of basic functions:

- A model for the dynamics of the weapon/seeker in space.
- A non-uniform pixel response defined in terms of fixed gain and pixel bias currents.
- A physics-based model for the thermal emission for a nominal blackbody at different temperatures – calculating the photon flux at different temperatures.
- A line of sight calculation to project the pixel line of sight onto the ground plane (flat earth) and an interpolation between regions defined on the satellite image – this includes the ability to sub-sample each pixel to obtain more representative images for small (sub-pixel) features, and a simple sky model for points above the horizon.

- A gain-off set and non-uniformity correction and a variable bit-length output quantisation.

The physics of thermal emission are included by representing each object within the scene by a blackbody with a slightly different temperature and calculating the photon flux from each region of the image – the photon number being directly related to the photo-current in each pixel. This is an approximation, but for single waveband images at least, it is a reasonable simplification because any thermal ‘colour’ information (ie. non-trivial spectral properties) due to deviations from a true blackbody can be represented by a change in the effective temperature of the object in the waveband being used. The thermal emission for the scene is calculated using standard physical formulae and can be modified to allow for different background temperatures, temperature variations in the background and for different wavebands within the infrared bands. Emission is easier to model than reflection, but – again, in a single waveband – it is a reasonable approximation to represent any intensity variation by a change in temperature for a thermal emission process. Once the thermal emission is calculated, the mean number of photons for each pixel is calculated using standard formulae for the focal length of the camera, pixel pitch, stare time and the efficiency of the photo-detector/photo-diode [2,3]. The mean number of photons detected by the pixel is then adjusted to allow for the photon shot noise – which is the main type of noise present in modern imager and CMOS readout circuits. The photon shot noise has a standard deviation, which is just the square root of the mean photon number [2,3].



Figure 1 – An example of a simulated 8-bit infrared seeker image showing three-bar target at the centre.

In this paper, the images are based on a simulation of a relatively narrow-band mid-wave (3.5-4.5 $\mu\text{m}$ ) infrared imager with 3-5% pixel non-uniformity. The images are 256 $\times$ 256 pixels in size and the standard images would

have 8 bits per pixel. The field of view of the seeker is stepped between 3 degrees, 6 degrees and 12 degrees as the weapon approaches the target to allow for background/contextual information to be included in the images – this improves the robustness of the correlation tracker in the seeker model and improves the accuracy of the image registration process used in the image reconstruction/fusion processing. Dead and saturated pixels are present in the original images but they are corrected before further processing using a median filter. The standard 8-bit frame rate of the imager is taken to be 40 Hz, rising to 320 Hz for the 1-bit images to keep the same data rate – however, the value of the frame rate of the camera is not crucial to the approach as long as the image registration of the 1-bit images is robust.

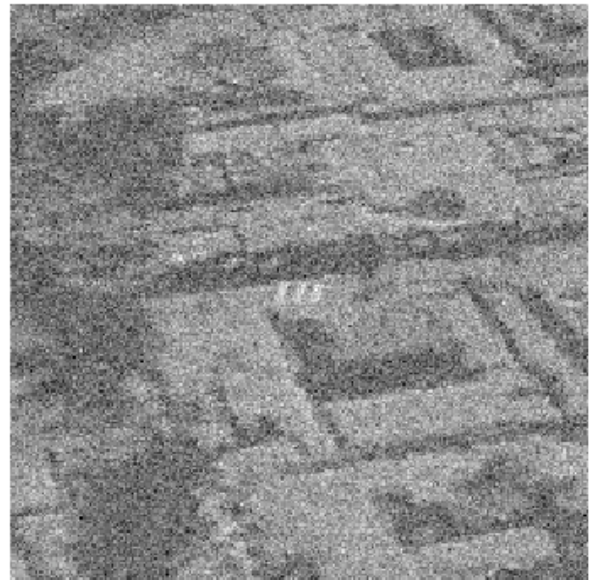


Figure 2 – Noisy, 8-bit simulated infrared image (frame rate = 320 Hz) and 1-bit thresholded image.

## 3.2 Communication Model

The modelling assumes that there is a data link to transmit the low bit-depth images to a remote operator and that this data link has a limited transmission capacity. In practice military data links have restrictions placed on them by the need for security and encryption [8], but these processes are not explicitly modelled here. The assumption is that the data rate is fixed to allow 40 frames of 8-bit imagery or 320 frames of 1-bit imagery to be transmitted per second. However, the results presented here are not particularly sensitive to the values used in the paper – as long as the 1-bit images contain enough background structure to allow the images to be registered relatively accurately, the image fusion will work well enough to allow the aim point/impact point to be identified.

## 3.3 Example Engagement

The engagement considered in this paper is based on a simple missile model where an air-launched, seeker-guided missile is launched against a ground target. The guidance system uses pursuit guidance and a correlation tracker to keep the seeker locked onto the target. The missile is launched at a speed of 450 knots, from an altitude of 3000 feet and a slant range of 3.5 km. It accelerates to approximately Mach 2 before impacting the target at approximately 650 m/s. The details of the weapon engagement are not important over and above the need to generate representative imagery and obtain a timeline for the terminal phase of the engagement – taken to be the final second before impact.

## 4 Image Registration

The first problem to be addressed when trying to combine the one-bit images into a single frame is to correct for the motion of the seeker – image registration. The geometrical transformation from image to image needs to be calculated – including translation, rotation and scaling. For good quality, visible-band imagery, the process of image registration has three main approaches. The simplest approach is to identify feature points in each image and then match them between images. A geometrical transformation can then be derived from the motion of these feature points [9]. The second method is to optimise the transformation using a similarity measure such as the mutual information [10] or the cross-correlation between two images [1]. The approach adopted here is based on the Fourier transform [12] because it is remarkably robust to noise in the image and it does not rely on the detection of consistent features from image to image because these features may not exist in the one-bit images. The Fourier registration method – also called phase correlation – uses the fact that the translation of a signal (or other function) is represented by a phase shift in the Fourier domain. If between two images ( $f$  and  $g$ ) are related by a translation,

$$f(x, y) = g(x - x_0, y - y_0)$$

taking the Fourier transform of  $g$  and  $f$  (represented by the complex-valued functions,  $F$  and  $G$  respectively),  $F$  and  $G$  are related to each other by a complex phase factor. This is related to the translation by

$$F(u, v) = G(u, v) \cdot \exp(2\pi j(ux_0 + vy_0))$$

The phase factor can be estimated by calculating the cross-power spectrum  $C$  of the two images by using,

$$C(u, v) = \frac{F(u, v) \cdot G(u, v)^*}{|F(u, v) \cdot G(u, v)^*|} = \exp(2\pi j(ux_0 - vy_0))$$

The inverse Fourier transform of the cross-power spectrum will produce an image that should have a maximum response at the point corresponding to the translation vector  $(x_0, y_0)$ .

This approach is remarkably robust, but it only corrects for translations of the images and not rotations and scaling of the images. Variants of the method exist for rotations and scaling transformations but the standard method is used here because the dominant image motion is translational (from movement of the line of sight). Scaling and rotation are important but they tend to be smaller effects. To correct for these effects, a simple optimisation search is used to find the maximum value for the peak in the phase correlation. The value of the peak should be one if the images are perfectly matched. In the presence of noise or other clutter, the maximum value will not be exactly one, but it should still be a maximum at the correct value of the translation vector  $(x_0, y_0)$ . The maximum value is not a perfect quantitative measure for the quality of the match between images, but it can be used as a good indication of relative quality of the image match for different images and it is used here to correct for rotations and scale transformations.

In practice, the image registration works well, and accurately estimates the scaling, rotation and translation of the images, as long as the signal-to-noise ratio for the underlying (un-quantised) images is greater than around  $\text{SNR} \approx 0.7$  and the pixel non-uniformities are less than about 5%. The requirement on the SNR is equivalent to a requirement on the structural qualities of the background scene. If the field of view of the seeker is too small to include enough background features, or the background is too bland, the image registration process will fail.

## 5 Image Fusion

Once the images have been aligned using the image registration processing, the 1-bit images are combined using a simple weighted sum. The weighting assigned to each 1-bit image is dependent upon the quality of the match with the first image in the sequence (the template). The goodness of fit is measured using the phase correlation (the maximum peak obtained in the Fourier registration processing). If the match is below a certain threshold, the image is discarded because any false match is likely to degrade the quality of the fused image. The threshold value is image/scene dependent because it is a

function of the perceived SNR of the images. For images where the match is above this threshold value, the weight assigned to the 1-bit image in the combined image is linearly proportional to the maximum phase correlation value.

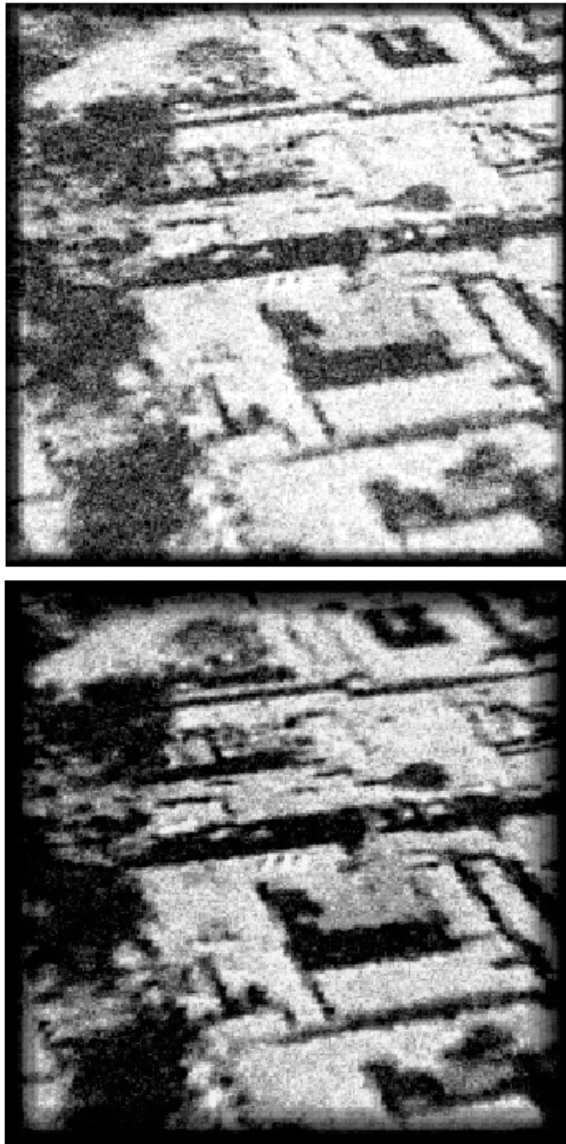


Figure 3 – Fused image from ten 1-bit frames (above) and twenty 1-bit frames (below) for SNR = 0.9 and approximately 0.5 seconds before impact.

## 6 Results

Examples of the registered and fused images are shown in Figures 3, 4 and 5. In figures 3 and 4, very low SNR 1-bit images have been aligned and fused to form grey scale images at approximately 0.5 seconds to impact (Figure 3) and 0.1 seconds to impact (Figure 4). Comparing these grey scale images to the 1-bit image shown in Figure 2, there is a clear improvement in the quality of the images and in the ability to discern the aim point for the weapon. Figure 4 also shows a fused image generated from the

final frames prior to impact – at 0.02 seconds to go – showing the impact point clearly.

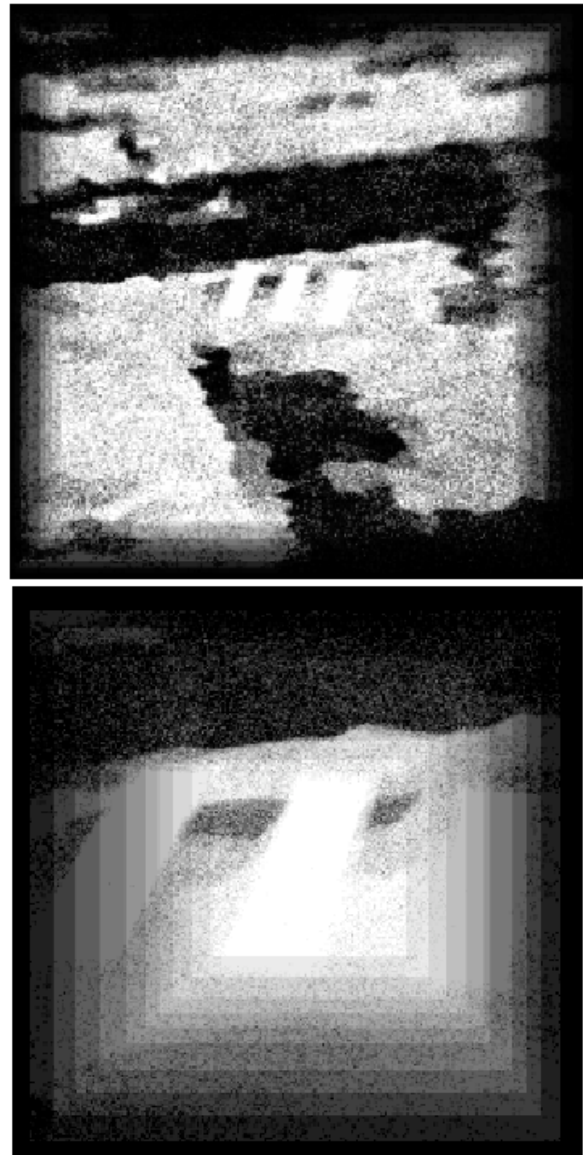


Figure 4 – Fused images from ten 1-bit frames for SNR = 1.1 and approximately 0.1 seconds before impact (above) and 0.02 seconds before impact (below).

The images in Figure 3 can be compared to the images shown in Figure 5, which are for different engagements with slightly different camera properties so that the images contain different levels of noise. Figure 5 shows fused images generated from 1-bit images with underlying signal-to-noise ratios of 1.7 and 2.5 compared to SNR  $\approx$  0.9 in Figure 3. For the lower noise images, little appears to be added by fusion of the 1-bit images because the noise is low enough that it generates very few threshold crossings and little information is added by adding more images. Some grey scale information is added at SNR  $\approx$  1.7, but very little at SNR  $\approx$  2.5.

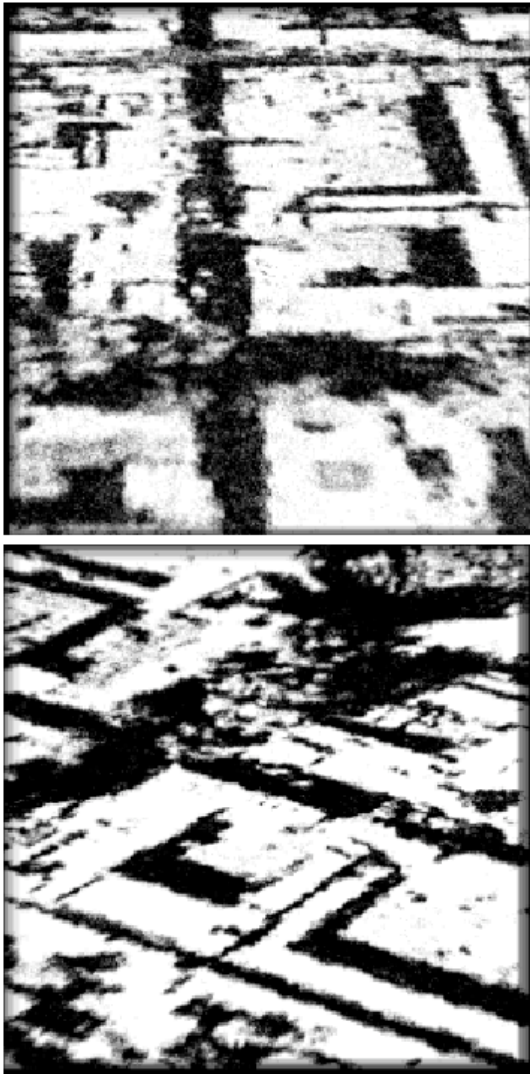


Figure 5 – Fused images from ten 1-bit frames for SNR = 1.7 (above) and SNR = 2.5 (below), approximately 0.5 seconds before impact.

## 7 Conclusions

This paper has considered the use of high frame-rate, 1-bit imaging in the terminal phase of a guided weapon engagement to assist in the assessment of the weapon impact point for battle damage indication. The ability to align and fuse the 1-bit images allows information to be extracted from weapon seeker with little need for pre-processing. This should assist in reducing latencies in transmission to a remote operator, thereby improving the ability to extract information as close as possible to the impact point.

The fused grey scale images show enough detail to allow the impact point to be identified even in the presence of significant levels of noise (due to short pixel integration times due to the high frame rate). The image registration and fusion post-processing is robust and accurate for scenes with signal-to-noise ratios down to 0.7 and pixel non-uniformities less than about 5%. Moreover,

it has been shown that image noise assists in the reconstruction/fusion of the grey scale images. This is a form of suprathreshold stochastic resonance – where the perceived information content of the reconstructed images increases as the noise level increases (over some finite range).

## References

- [1] J. F. Ralph, M. Bernhardt, ‘Smart Imaging in the Infrared’, *Contemporary Physics*, Vol. 43, pp. 259-272, 2002.
- [2] A. Hoffman, M. Loose and V. Suntharalingam, ‘CMOS Detector Technology’, *Experimental Astronomy*, Vol. 19, pp.111-134, 2005.
- [3] H. Schneider, M. Walther, C. Schonbein, R. Rehm, J. Fleissner, W. Pletschen, J. Braunstein, P. Koidl, G. Weimann, J. Ziegler and W. Cabanski, ‘QWIP FPAs for high-performance thermal imaging’, *Physica E*, Vol. 7, pp. 101-107, 2000.
- [4] L. Gammaitoni, P. Hänggi, P. Jung and F. Marchesoni, ‘Stochastic Resonance’, *Reviews of Modern Physics*, Vol. 70, pp. 223-287, 1998.
- [5] N. G. Stocks, ‘Suprathreshold stochastic resonance in multilevel threshold systems’, *Physical Review Letters*, Vol. 84, pp. 2310-2313, 2000.
- [6] N. G. Stocks, D. Allingham and R. P. Morse, ‘The application of suprathreshold stochastic resonance to cochlear implant coding’, *Fluctuation and Noise Letters*, Vol. 2, pp. L169-L181, 2002.
- [7] M.D.McDonnell, N.G.Stocks, C.E.M.Pearce, D.Abbott, ‘Optimal information transmission in nonlinear arrays through suprathreshold stochastic resonance’, *Physics Letters A*, Vol.352, pp.183-189, 2006.
- [8] K.L.Edwards, ‘Air-to-ground targeting – UAVs, data links and interoperability (project Extendor)’, *The Aeronautical Journal*, Vol. 108, pp. 493-504, 2004.
- [9] M. Holm, Towards automatic rectification of satellite images using feature based matching, *Proc. of the Int. Geoscience and Remote Sensing Symposium IGARSS’91*, Espoo, Finland, pp.2439–2442, 1991.
- [10] P. Thevenaz, M. Unser, ‘An efficient mutual information optimizer for multi-resolution image registration’, *Proc. of the IEEE International Conference on Image Processing ICIP’98*, Chicago, IL, pp.833–837, 1998.
- [11] J.P. Lewis, ‘Fast Template Matching’, *Vision Interface 95*, Canadian Image Processing and Pattern Recognition Society, pp.120–123, 1995.
- [12] E.D. Castro, C. Morandi, Registration of translated and rotated images using finite Fourier transform, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 9, pp.700–703, 1987.