

Point target detection using super-resolution reconstruction

Judith Dijk, Adam W.M. van Eekeren, Klamer Schutte and Dirk-Jan J. de Lange

TNO Defence, Security And Safety
P.O. Box 96864 2509 JG The Hague The Netherlands
Email: Judith.Dijk@tno.nl

ABSTRACT

Surveillance applications are primarily concerned with detection of targets. In electro-optical surveillance systems, missiles or other weapons coming towards you are observed as moving points. Typically, such moving targets need to be detected in a very short time. One of the problems is that the targets will have a low signal-to-noise ratio with respect to the background, and that the background can be severely cluttered like in an air-to-ground scenario.

The first step in detection of point targets is to suppress the background. The novelty of this work is that a super-resolution reconstruction algorithm is used in the background suppression step. It is well-known that super-resolution reconstruction reduces the aliasing in the image. This anti-aliasing is used to model the specific aliasing contribution in the camera image, which results in a better estimate of the clutter in the background. Using super-resolution reconstruction also reduces the temporal noise, thus providing a better signal-to-noise ratio than the camera images. After the background suppression step common detection algorithms such as thresholding or track-before-detect can be used.

Experimental results are given which show that the use of super-resolution reconstruction significantly increases the sensitivity of the point target detection. The detection of the point targets is increased by the noise reduction property of the super-resolution reconstruction algorithm. The background suppression is improved by the anti-aliasing.

Keywords: super-resolution reconstruction, target detection, point target, background

1. INTRODUCTION

In surveillance applications moving targets need to be detected in a very short time. In electro-optical surveillance systems missiles or other weapons coming towards you can be seen as moving points. These point targets may have a low signal-to-noise ratio with respect to the background. Another problem is that the background may also have many contrast or clutter, like in an air-to-ground scenario. Therefore, the first step in the detection of point targets is to suppress the background. To do so, normally two subsequent images are aligned and subtracted.

In previous work¹ we developed super-resolution reconstruction techniques to improve images for general applications. In this paper we apply these standard super-resolution reconstruction techniques to detect moving point targets. We propose to use a difference image made with the super-resolution reconstruction algorithm for the background suppression step. This has the advantage that the aliasing in the image can be better estimated, which results in a better estimate of the clutter in the background. Using super-resolution reconstruction also reduces the noise, so that the super-resolution image has a better signal-to-noise ratio than single camera images. After the background suppression step common detection algorithms such as thresholding or track-before-detect can be used. In this paper we show the results with a direct threshold.

This paper is organized as follows. In the next section we discuss the Hardie super-resolution reconstruction method and which difference or residue image can be used for point target detection. In section 3 a standard background suppression method is discussed, along with a standard detection method using thresholding. The methods are tested on a synthetically made data set. Point targets are inserted in a background image using a camera model. In section 4 the experimental setup is presented. The results on this data set are given in section 5. Finally conclusions are drawn in section 6.

This research is described in European patent application nr. 06077053.4: point target detection with super-resolution.

2. SUPER-RESOLUTION RECONSTRUCTION

Super-Resolution (SR) reconstruction is a well-known technique to increase the resolution of a sequence of aliased Low-Resolution (LR) images. The resolution enhancement factor RE is the ratio of the size of the HR image with respect to the size of the LR image. Generally super-resolution reconstruction can be split up in three parts: 1) registration, 2) fusion and 3) deblurring. The SR method described below incorporates the last two parts in one single step.

2.1. Registration

In simple scenarios, the scene movement between two frames can be well described with a single shift. The registration of the LR frames is done with an iterative gradient-based shift estimator.² A gradient-based shift estimator³ finds the displacement $t_{\vec{x}}$ between two shifted signals as the least squares solution of (1)

$$MSE = \frac{1}{N} \sum_S \left(s_2(\vec{x}) - s_1(\vec{x}) - t_{\vec{x}} \frac{\partial s_1}{\partial \vec{x}} \right)^2 \quad (1)$$

with s_2 a shifted version of s_1 , \vec{x} the sample positions and N the number of samples in supported region S .

The solution of (1) is biased, which is corrected in an iterative way. In the first iteration s_2 is shifted with the estimated sub-pixel displacement, which is accumulated in the next iteration with the estimated displacement between s'_2 (shifted s_2) and s_1 . This schema is iterated until convergence and results in a very precise ($\sigma_{disp} \approx 0.01$ pixel for noise free data) unbiased registration, which approaches the Cramér-Rao bound.⁴

2.2. Super-Resolution fusion and deblurring

The fusion and deblurring part of the SR reconstruction is performed with a method developed by Hardie et al.⁵ This method employs a discrete observation model that relates the ideally sampled High-Resolution (HR) image \mathbf{z} and the observed LR frames \mathbf{y} :

$$y_m = \sum_{r=1}^H w_{m,r} z_r + \eta_m \quad (2)$$

where $w_{m,r}$ represents the contribution of the r^{th} HR pixel in \mathbf{z} to the m^{th} LR pixel in \mathbf{y} . This contribution depends on the frame-to-frame motion and on the blurring of the Point Spread Function (PSF). η_m denotes additive noise.

The HR image estimate $\hat{\mathbf{z}}$ is defined as the \mathbf{z} that minimizes:

$$C_{\mathbf{z}} = \sum_{m=1}^L \left(y_m - \sum_{r=1}^H w_{m,r} z_r \right)^2 + \lambda \sum_{i=1}^H \left(\sum_{j=1}^H \alpha_{i,j} z_j \right)^2 \quad (3)$$

with L the number of LR samples and H the number of HR grid points.

The cost function in (3) balances two types of errors. The left term is minimized when a candidate \mathbf{z} , projected through the observation model (2), matches the observed data. The right term is a regularization term, which is necessary because directly minimizing the first term is an ill-posed problem. The parameters $\alpha_{i,j}$ (4) are selected to perform a Laplacian operation on \mathbf{z} and ensure that the regularization term is minimized when \mathbf{z} is smooth.

$$\alpha_{i,j} = \begin{cases} 1 & \text{for } i = j \\ -1/4 & \text{for } j:z_j \text{ is a cardinal neighbor of } z_i \end{cases} \quad (4)$$

The first term in (3), the so-called datafit term, can also be seen as a difference image. It is exactly this term that we use to detect the point targets. After minimization of (3) the so-called difference image is defined as:

$$diff_m = y_m - \sum_{r=1}^H w_{m,r} z_r \quad (5)$$

In the next section we discuss how this difference image can be used to detect point targets.

3. POINT TARGET DETECTION WITH SUPER-RESOLUTION RECONSTRUCTION

3.1. Background suppression

The first step of point target detection usually is to suppress the clutter of the stationary background in the image. The clutter suppression step should remove the information of the static background while preserving the target energy. Zhang⁶ states that an ideal clutter suppression algorithm would be capable of transforming an arbitrary inhomogeneous and no stationary background into an approximately stationary, homogeneous and Gaussian noise background without reducing the effective target signal power. However, the ideal background suppression algorithm does not exist yet.

A standard way of performing background suppression is to first determine the shift (dx, dy) between two subsequent frames y_m and y_{m-1} by registration. The next step is to correct one of the frames for this shift and then subtract the frames

$$diff_m(x, y) = y_m(x, y) - y_{m-1}(x + dx, y + dy) \quad (6)$$

We propose to do the background suppression by the use of super-resolution reconstruction. The formula for the difference image is given in equation 5. This difference image also depends on the shifts between frames, but instead of using only two frames, a number of frames is used to estimate the clutter in the background. By using super-resolution reconstruction, the difference image is better estimated because it models the specific aliasing contribution in the camera image. Another advantage of super-resolution reconstruction is that it reduces the temporal noise in the difference image, thus providing a better signal-to-noise ratio than the camera images.

3.2. Detection of point targets

After the super-resolution reconstruction or alignment the point targets need to be detected. Both the super-resolution reconstruction step or alignment step result in images in which the stationary objects are reduced, but the moving objects and noise are still visible. On this difference image the detection of the objects is done. An intuitive detection technique is to threshold this difference image. All pixels with a value above a certain threshold value are detected as targets.

$$T_m = \begin{cases} 0 & \text{for } |diff_m(x, y)| < \text{Threshold} \\ 1 & \text{for } |diff_m(x, y)| \geq \text{Threshold} \end{cases} \quad (7)$$

This Detect-Before-Track method works well for targets that have a sufficiently high Signal-to-Noise Ratio (SNR), so that the moving object can be detected in every frame. However, dim objects may not be detected with such a thresholding technique. To increase the sensitivity of the detection algorithm tracking can be used to associate the found detections.

If the SNR is so low that the target cannot be detected in one single image, the target might be detected by using successive frames, so-called Track-Before-Detect (TBD), see for instance Reed.⁷ These kind of algorithms can also be used after super-resolution reconstruction. In this paper we only show the results for direct thresholding.



Figure 1. The subsampled input image in which the point targets are inserted

4. EXPERIMENTAL SETUP

The performance of the different algorithms needs to be tested on images containing point targets. However, these images are not available. Therefore, we inserted point targets in a real image. A recorded satellite image from the Ikonos database is used as input. This image is subsampled by a factor 4 to obtain more details in the final output image. This image is shown in figure 1. A point target is placed in this super scale image. The images are downsampled a factor 15 using a camera model, in which the blur and fill-factor of the camera are modeled. The range of the LR images without point targets is 248-633 grey values. The maximum intensity of the point target in each low resolution image is scaled to the defined intensity of the point targets. The camera moves in a slightly different direction than the point target, with a constant speed of 0.75 low resolution pixels per frame.

In the generated test images the intensity and velocity of the point targets are varied. The intensity values depend on the noise in the image and vary between $0.5 * \text{noise}$ to $64 * \text{noise}$. The velocity of the point target varies between 0 low resolution pixel per frame (no movement with respect to the background) to 2 low resolution pixel per frame (the point moves with 2 pixels per frame with respect to the background). Finally, temporal noise is added to the low resolution images with a standard deviation of 1, 2, 4, 8, 16 and 32. Examples of the generated test images are shown in figure 2.

5. RESULTS

In figure 3 the difference images for different background suppression methods are shown. It can be seen that the difference images using super-resolution reconstruction contain less background information. This effect is best seen around the road in the left part of the image. The differences between the SR methods with different Resolution Enhancement factors are rather small, but the road is still somewhat more clear in the difference image with RE factor 1.

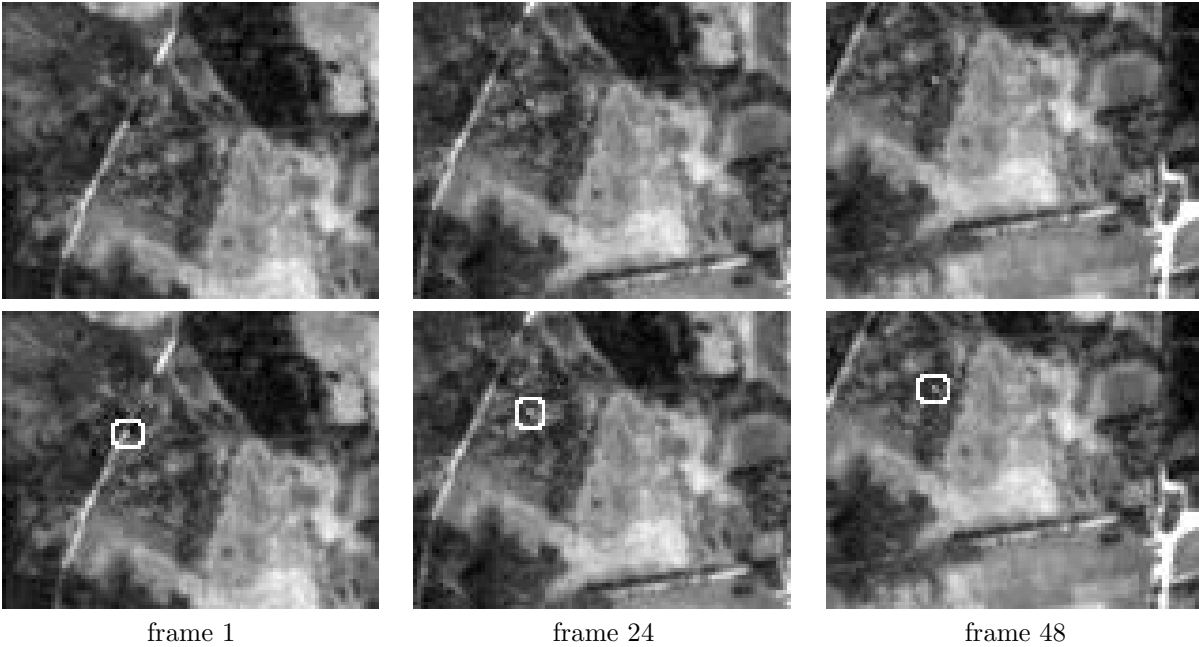


Figure 2. Three LR images of the same sequence, with (lower row) and without (upper row) indication of the position of the point target. The noise is 4, the amplitude of the point target is 64 (16×4). The velocity of the point is 0.5 LR pixel/frame. The point target is moving from the left to the right.

To evaluate the results we use sequences of 48 frames. For these sequences two different values are determined. The first value T_1 is a measure for the threshold value that is needed to detect the point target. In each frame the maximum difference at the position of the point target is taken. The value T_1 is the 90th percentile of the maxima of the individual frames. This is the maximum threshold value for which at least 90% of the point targets in the sequence are found. The higher T_1 is, the better the detection of the point targets.

The second value T_2 is a measure of the minimum threshold value needed to avoid too many false positives. In each frame the 95th percentile of all difference values which are not at the position of the point target is taken. Because the point targets can influence their neighbors, the difference values of these neighbors are also not taken into account. Also the difference values of pixels near the border are not used in calculating the 95th percentile.

The value T_2 is the maximum of these values over the frames. A lower value for T_2 means a lower false alarm rate, and thus a better method.

As a final measure of the sensitivity the maximum threshold T_1 is divided by the minimum threshold T_2 . If this value is larger than 1, the point target can be detected without much false alarms. The higher this value, the more sensitive the method is. This value is also shown normalized to the T_1/T_2 fraction of the shift-and-subtract method. In this way, the improvement of the super-resolution reconstruction methods can be seen with respect to this shift-and-subtract method.

In figure 4 the threshold values for point targets with velocity 1.5 and noise = 4 is shown. It can be seen that T_1 increases with the amplitude of the point target. The minimum threshold T_2 is more or less constant. This is expected, as the background does not depend on the point target. The measure T_1/T_2 is shown in figure 4 (c). It can be seen that the values for super-resolution reconstruction methods are much higher than the values for the shift and subtract method. This effect is mainly due to a better value of T_1 . It can also

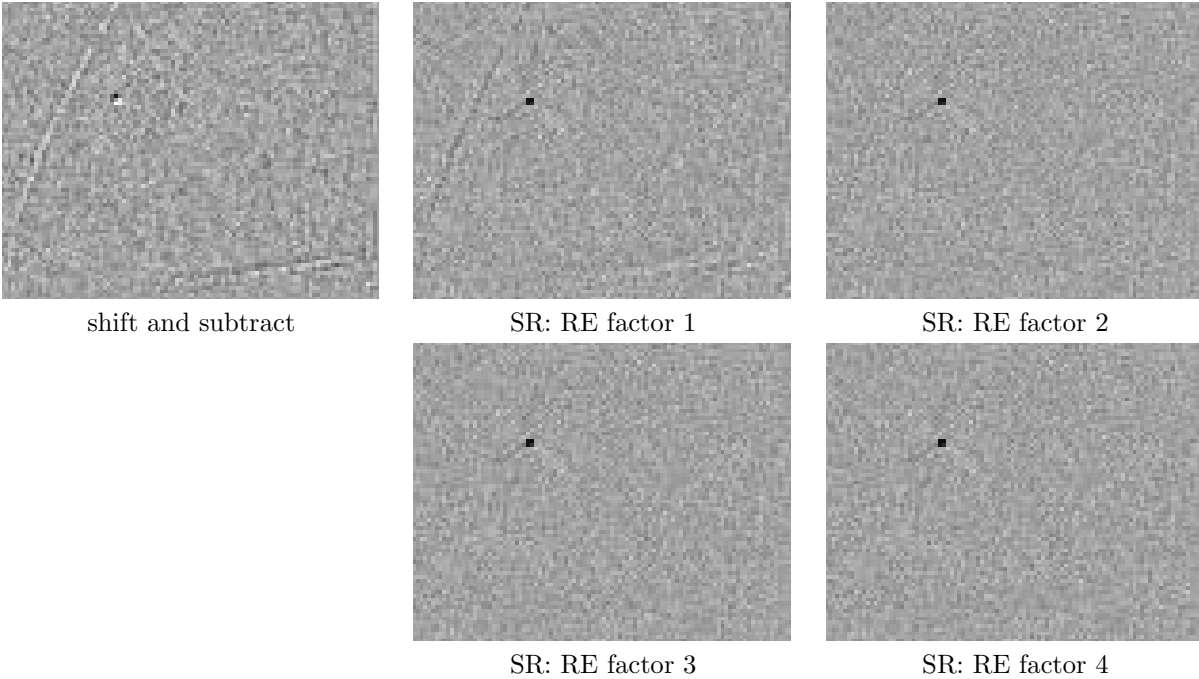


Figure 3. Difference images for the different background suppression methods. In the left image the difference image for the shift-and-subtract method is shown. In the four images on the right the results for the super-resolution reconstruction method are shown, for different Resolution Enhancement factors. The difference image is shown for the 24th frame in the sequence. The noise is 4, the amplitude of the point target is 64 (16×4). The velocity of the point is 0.5 LR pixel/frame. The difference images using super-resolution reconstruction contain less background information. This effect is best seen around the road in the left part of the image. The differences between the SR methods with different Resolution Enhancement factors are small, but the road is more clear in the difference image with RE factor 1

be seen that the sensitivity for a SR method with a RE factor > 1 is larger than for the super-resolution reconstruction method with RE factor 1. This effect is based on a better determination of the background, the value of T_2 is lower for RE factors is > 1 . An overall conclusion is that, as expected, the sensitivity increases linear with the amplitude for all background suppression methods. This indicates that the chosen thresholds are good measures of the sensitivity.

In figure 5 the threshold values for point targets for noise = 4 and amplitude of 64 (16×4) are shown. T_1 increases with the velocity of the point target. The minimum threshold T_2 is more or less constant. The measure T_1/T_2 is shown in figure 5 (c). In this figure it can also be seen that sensitivity of the super-resolution reconstruction methods is higher than the sensitivity of the shift and subtract method and that a RE factor > 1 leads to a higher sensitivity.

In figure 6 we show the results for different noise values. This will alter the amount of clutter in the background with respect to the noise. The ratio between the amplitude of the point target and the noise is kept constant. It can be seen that for small amounts of noise the sensitivity increases when the noise is increased. This was expected, as the point target amplitude is also increased with respect to the noise.

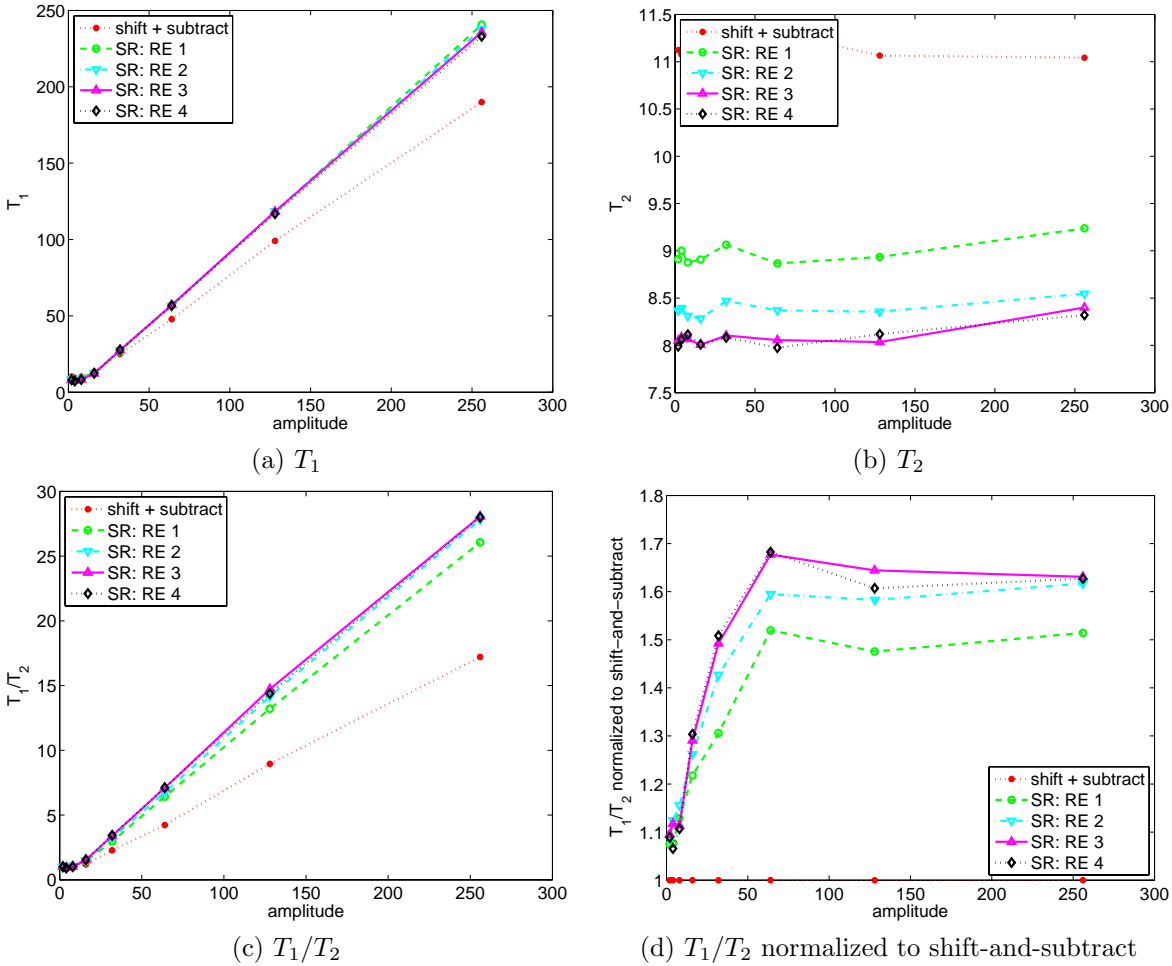


Figure 4. Threshold values as function of the signal-to-clutter ratio for the different background suppression methods. The noise is 4. The velocity of the point target is 1.5 LR pixel/frame.

6. CONCLUSIONS AND DISCUSSION

In this paper we show that the sensitivity of a point target detection is increased when a super-resolution reconstruction algorithm is used to suppress the background. After the background suppression step a thresholding was used, but also other common detection algorithms such as track-before-detect can be used.

The super-resolution reconstruction results are significantly better than for the shift-and-subtract method, especially in high clutter scenarios and for low apparent target velocities. The improvement in sensitivity is based on two properties of the super-resolution reconstruction algorithm: noise reduction and anti-aliasing. The effects of noise reduction are shown by comparing the super-resolution reconstruction algorithm with no resolution enhancement with a simple shift-and-subtract algorithm. For higher resolution enhancement factors the sensitivity was higher. This effect is mainly due to the fact that a better estimation of the background by using anti-aliasing will decrease the false alarm rate. The difference between resolution enhancement factors 3 and 4 are not significant.

REFERENCES

1. K. Schutte, D. de Lange, and S. van den Broek, "Signal conditioning algorithms for enhanced tactical sensor imagery," in *SPIE proceedings: Infrared Imaging Systems: Design, Analysis, Modeling, and Testing XIV*, **5076**, pp. 92–100, 2003.

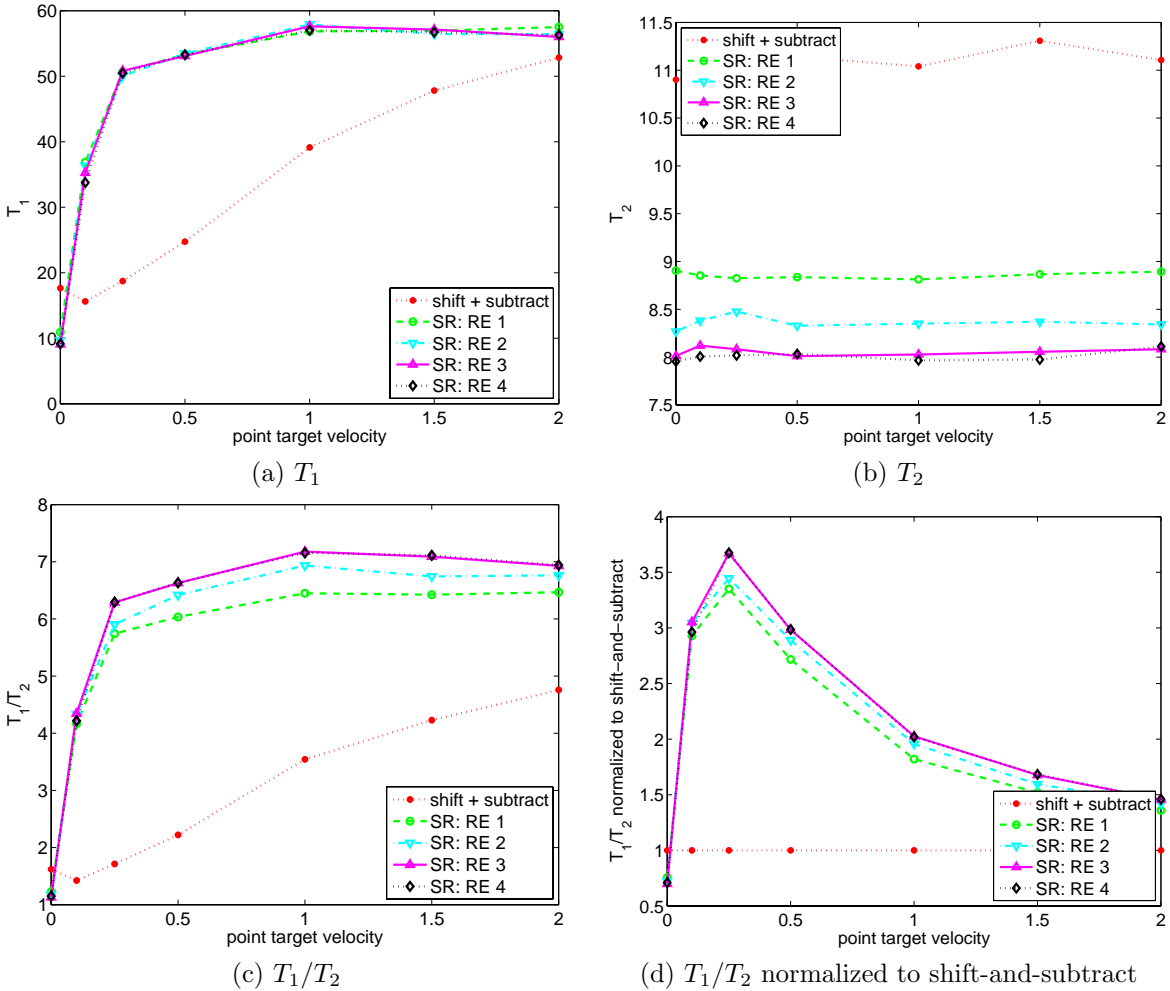


Figure 5. Threshold values as function of the target velocity for the different background suppression methods. The noise is 4. The amplitude of the point target is 64 (16×4).

2. T. Pham, M. Bezuijen, L. van Vliet, K. Schutte, and C. L. Hendriks, "Performance of optimal registration estimators," in *Visual Information Processing XIV*, **5817**, pp. 133–144, SPIE Defense and Security, 2005.
3. B. Lucas and T. Kanade, "An iterative image registration technique with an application to stereo vision," in *Imaging Understanding Workshop*, pp. 121–130, 1981.
4. S. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*, Prentice-Hall, 1993.
5. R. Hardie, K. Barnard, J. Bognar, E. Armstrong, and E. Watson, "High-resolution image reconstruction from a sequence of rotated and translated frames and its application to an infrared imaging system," *Optical Engineering* **37**(1), pp. 247–260, 1998.
6. W. Zhang, N. Cong, and L. Wang, "Algorithms for optical weak small targets. detection and tracking: review.," in *IEEE int. conf. Neural Networks & Signal Processing*, **1**, pp. 643–647, 2003.
7. I. Reed, R. Gagliardi, and L. Stotts, "A recursive moving-target-indication algorithm for optical image sequences," *IEEE transactions on aerospace and electronic systems* **26**(3), pp. 434–440, 1990.

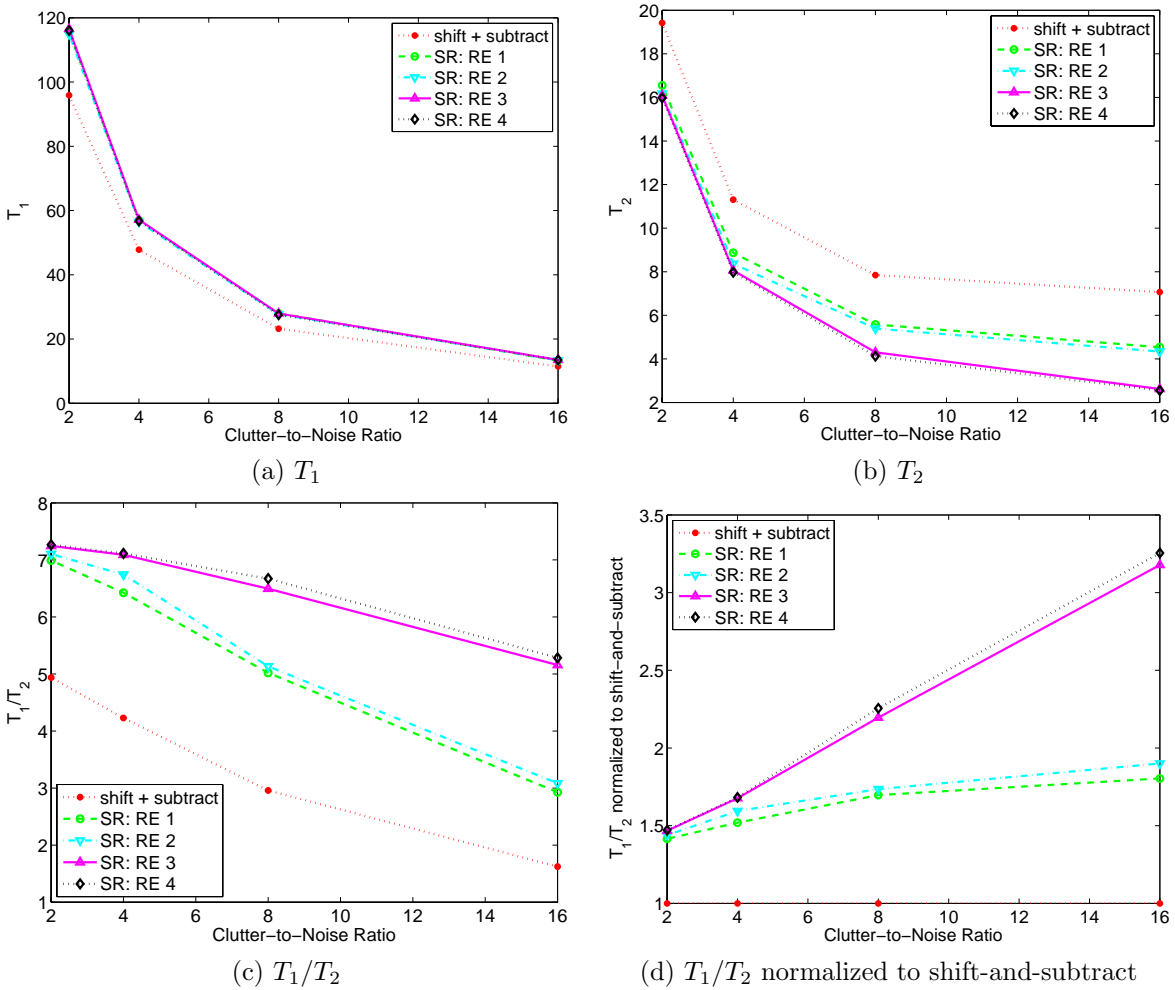


Figure 6. Threshold values for different Clutter-to-Noise ratio's. Note that these are arbitrary units, as the clutter is not measured. The velocity of the point target is 1.5, the amplitude is 16 * noise, so the Signal-to-Noise ratio of the point targets is constant. It can be seen that the super-resolution reconstruction methods give a better improvement for lower signal-to-clutter ratios.