Supplementary Material:

Appendix A: Stimulus Set Sources

The development of a suitable scene-target data set has been a significantly more difficult challenge than expected, and the final sets of scenes come from one of three sources:

1. **The TNO Search_2 data set (Toet, Bijl, & Valeton, 2001)**

   Developed by the TNO (Image dataset for testing search and detection models: Opt. Eng., Vol. 40, 1760 (2001); doi:10.1117/1.1388608), it consists of 44 scenes (6144 x 4096) of military vehicles in various natural scenes at different ranges. These are used in tests on the TAM (Zelinsky, 2008), and seem very appropriate for the type of comparisons in this investigation. Despite the set already having targets identified with coordinates and masking information, these are not used as the specifications for this experiment, due to the requirement that both a present-patches and a absent-patches must be selected from the scene.

2. **Flickr.com**

   A search of the online image database, filtering for appropriate size and scale, produced over 6,000 potential scenes. The terms used were “Hidden”, ”House Scene”, “Spot” (As in ‘spot the.’), “Camouflage” and “Woodland”. All the images were generally selected to have a flickr license level of 5-7 (flickr.photos.licenses.getInfo) which in general means that they can be used for non commercial purposes without credit.

3. **Personally Taken Photos**

   To create some controlled stimuli, I took a number of photos of toys (vehicles and animals) in natural scenes in a local park. The scenes have allowed for a number of search scene scenarios that were not available for the previous sources. Photos were taken on a Fuji Finepix F10 Digital Camera, with 6 megapixels (3,008 x 2,000)
Appendix B: Example Model Heatmaps – See main document for details

Original Scene

Feature Congestion: Heatmap

Edge Density: Edges

Segmentation: Min 500 Pixels

Segmentation: Min 1000 Pixels

Segmentation: Min 2000 Pixels
Appendix C: Stimulus Set Extraction

A few more examples of how the stimulus Scene-Target sets were selected.

Figure 1: Second Example of Target Present and Target Absent Patches for a scene

Figure 2: Third Example of Target Present and Target Absent Patches for a scene

The following images are additional examples of the stimuli given to the clutter model algorithms for the Target and Target-Black conditions to produce the Clutter Profile, as was shown in Figure 8 in the main paper.
Regional effects of clutter on human target detection performance

Figure 3: Second example of the sub-regions in the scene

Figure 4: Examples of the actual Sub-regions that were pushed into the clutter measures, the lower images show the condition where the target patch was removed from the sub-region. You can see that for the centre and right hand images, the black sections where the sub-region extends beyond the Scene boundary.
Appendix D: Full Scene Clutter Scatterplot Matrix

Figure 5: Scatterplot matrix based on Table 1 in the main paper. The axes are z scores, and correlation coefficients are shown on each plot, the bootstrapped confidence intervals shown underneath in square brackets. +: p < 0.001; *: p < 0.05. Note that the A’ vs. the four metrics are displayed along the top row.
Appendix E: Comparison of Clutter effect on Mean Response Time and Performance

Figure 6: Showing the comparative three way behaviour of Response Time, Performance, and each of the metrics. The correlation and regression values for each of the three elevations are shown, and shows that while the RT v.s. A' (red/X-Z plane) is a good correlation, and the Metric v.s. A' (green/X-Y plane) typically have recognisable correlation patterns, there is no apparent correlation discernable in any of the Metric v.s. RT (cyan/Y-Z plane). These are reflected in the values.
Appendix F: Full Correlation Profiles for all four clutter metrics

Figure 7: The region correlation profiles for all four clutter metrics, shown for both Target Present and target absent. Dotted line indicates the whole scene correlation level. These graphs show that, while the profiles for Edge Density and Segmentation are very similar to the Sub-band Entropy and Feature Congestion respectively, they persistently maintain lower values of coloration than the chosen metrics, an additional reason for excluding them from the further analysis.
Appendix G: Full Output for BIC Stepwise Regression

Reproduced below is the full output from the Whole Scene regression reported in page 17 of the main manuscript. The analysis was carried out using the LinearModel regression functions in MATLAB (2012b). Note that a colon indicates an interaction term. * in the model equation indicates interaction and lower terms.

Constant model
Change in BIC for adding Edge Density is -0.65189
Change in BIC for adding Segmentation is 2.4446
Change in BIC for adding Sub-band Entropy is -0.82257
Change in BIC for adding Feature Congestion is 2.6014

1. Adding Sub-band Entropy, BIC = -116.2705
   Change in BIC for adding Edge Density is 4.1407
   Change in BIC for adding Segmentation is 0.79809
   Change in BIC for adding Feature Congestion is -0.10731

2. Adding Feature Congestion, BIC = -116.3778
   Change in BIC for adding Edge Density is 4.8043
   Change in BIC for adding Segmentation is 4.8092
   Change in BIC for adding Sub-band Entropy:Feature Congestion is 3.019
   Change in BIC for removing Sub-band Entropy is 3.5312

Linear regression model: $A' = \text{Intercept} + \text{Sub-band Entropy} + \text{Feature Congestion}$

Estimated Coefficients:

<table>
<thead>
<tr>
<th></th>
<th>Estimate</th>
<th>SE</th>
<th>tStat</th>
<th>pValue</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>1.197</td>
<td>0.154</td>
<td>7.76</td>
<td>0</td>
</tr>
<tr>
<td>Sub-band Entropy</td>
<td>-0.106</td>
<td>0.036</td>
<td>-2.90</td>
<td>0.004</td>
</tr>
<tr>
<td>Feature Congestion</td>
<td>0.015</td>
<td>0.007</td>
<td>2.21</td>
<td>0.029</td>
</tr>
</tbody>
</table>

Model Results: $R^2 = 0.0846, F(3,116) = 5.36, p = 0.006$

Including Interactions for this regression model

Estimated Coefficients:

<table>
<thead>
<tr>
<th></th>
<th>Estimate</th>
<th>SE</th>
<th>tStat</th>
<th>pValue</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>0.653</td>
<td>0.439</td>
<td>1.49</td>
<td>0.139</td>
</tr>
<tr>
<td>Sub-band Entropy</td>
<td>0.014</td>
<td>0.098</td>
<td>0.14</td>
<td>0.886</td>
</tr>
<tr>
<td>Feature Congestion</td>
<td>0.138</td>
<td>0.094</td>
<td>1.47</td>
<td>0.143</td>
</tr>
<tr>
<td>Sub-band Entropy : Feature Congestion</td>
<td>-0.027</td>
<td>0.020</td>
<td>-1.32</td>
<td>0.189</td>
</tr>
</tbody>
</table>

Model Results: $R^2 = 0.098, F(4,115) = 4.18, p = 0.008,$
Appendix H: Output for BIC Stepwise Regression

Reproduced below is the full output from the sub-region regression reported in page 23 of the main manuscript. The analysis was carried out using the LinearModel regression functions in MATLAB (2012b). Note that a colon indicates an interaction term. * indicates interaction and lower order terms present.

Not all the stages for this stepwise regression are shown because of the number factors relating to patch size (400). The term x129 refers to Sub-band Entropy with a sub region size of 484 pixels. Term x152 refers to Feature Congestion with a sub region size of 1024 pixels.

Linear regression model: $y \sim 1 + x_{129} + x_{152}$

Estimated Coefficients:

<table>
<thead>
<tr>
<th></th>
<th>Estimate</th>
<th>SE</th>
<th>tStat</th>
<th>pValue</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>1.116</td>
<td>0.102</td>
<td>10.888</td>
<td>0</td>
</tr>
<tr>
<td>x129</td>
<td>-0.111</td>
<td>0.029</td>
<td>-3.849</td>
<td>0.0002</td>
</tr>
<tr>
<td>x152</td>
<td>0.023</td>
<td>0.009</td>
<td>2.411</td>
<td>0.0175</td>
</tr>
</tbody>
</table>

Model Results: $R^2 = 0.132$, $F(3,116) = 8.85$, $p = 0.000265$.

Including Interactions for this regression model

Linear regression model: $y \sim 1 + x_{129}*x_{152}$

Estimated Coefficients:

<table>
<thead>
<tr>
<th></th>
<th>Estimate</th>
<th>SE</th>
<th>tStat</th>
<th>pValue</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>1.137</td>
<td>0.305</td>
<td>3.721</td>
<td>0.0003</td>
</tr>
<tr>
<td>x129</td>
<td>-0.116</td>
<td>0.082</td>
<td>-1.422</td>
<td>0.1577</td>
</tr>
<tr>
<td>X152</td>
<td>0.016</td>
<td>0.091</td>
<td>0.181</td>
<td>0.8569</td>
</tr>
<tr>
<td>x129:x152</td>
<td>0.002</td>
<td>0.024</td>
<td>0.072</td>
<td>0.9425</td>
</tr>
</tbody>
</table>

Model Results: $R^2 = 0.132$, $F(4,115) = 5.85$, $p = 0.000936$.

Interaction regression model for SBE: 484 pixels & FC: 90 pixels

Linear regression model: $y \sim 1 + x_{129}*x_{200}$

Estimated Coefficients:

<table>
<thead>
<tr>
<th></th>
<th>Estimate</th>
<th>SE</th>
<th>tStat</th>
<th>pValue</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>1.311</td>
<td>0.257</td>
<td>4.049</td>
<td>0.0000</td>
</tr>
<tr>
<td>x129</td>
<td>-0.111</td>
<td>0.071</td>
<td>-1.566</td>
<td>0.1200</td>
</tr>
<tr>
<td>x200</td>
<td>-0.001</td>
<td>0.040</td>
<td>-0.018</td>
<td>0.9856</td>
</tr>
<tr>
<td>x129:x200</td>
<td>0.003</td>
<td>0.011</td>
<td>0.309</td>
<td>0.7579</td>
</tr>
</tbody>
</table>

Model Results: $R^2 = 0.131$, $F(4,115) = 5.78$, $p = 0.00102$. 

Regional effects of clutter on human target detection performance