

# Review of Camouflage Assessment Techniques

Alexander Toet, Maarten. A. Hogervorst

TNO Human Factors  
Kampweg 55, 3769DE Soesterberg  
THE NETHERLANDS

{lex.toet, maarten.hogervorst}@tno.nl

## ABSTRACT

*In military operations signature reduction techniques such as camouflage nets, low-emissive paints, and camouflage patterns are typically deployed to optimize the survivability of high value assets by minimizing their detectability. Various methods have been developed to assess the effectiveness of these camouflage measures. There are three main approaches to the evaluation of camouflage measures: (1) a subjective approach through observer experiments, (2) an objective computational approach through image analysis, and (3) an objective approach through physical measurements. Although subjective evaluation methods have a direct relation with the operational practice, they are often difficult to implement because of time and budget restrictions (these experiments typically require a lot of observers, repetitions and different targets and backgrounds), or simply because the associated conditions are not safe for the observers. The objective evaluation methods are typically based on the outcome of psychophysical laboratory experiments using simple artificial stimuli, presented under extremely restricted (impoverished) conditions, and in different experimental paradigms. Other objective methods are based on signal processing techniques and have no obvious counterpart in human vision. Only a few attempts have been made to validate any of these objective metrics against the performance of human observers in realistic military scenarios. As a result, there are currently no standard and internationally accepted methods and procedures to evaluate camouflage equipment and techniques, and to indicate their military effectiveness. In this review paper we present an overview of the various subjective (psychophysical) and objective (computational, image or video based) evaluation methods that are currently available and that have been used to validate camouflage effectiveness. In addition, we discuss the relative merits of field experiments versus laboratory experiments.*

## 1.0 INTRODUCTION

Both in large-scale battles and in peace keeping operations the survivability of high value military assets must be optimized by minimizing their detectability. This is typically achieved through camouflage measures, such as camouflage nets, low-emissive paints, camouflage patterns, etc. It is of great importance to know how effective such camouflage measures are in reducing a target's signature, and to know why and when they work. This requires a metric of detectability. Various metrics of target detectability have been developed over the years. There are three main approaches to the evaluation of camouflage measures:

- a subjective approach through observer experiments,
- an objective computational approach through image analysis, and
- an objective approach through physical measurements.

Although subjective evaluation methods have a direct relation with the operational practice, they are often difficult to implement because of time and budget restrictions (these experiments typically require a lot of observers, repetitions and different targets and backgrounds), or simply because the associated conditions are not safe for the observers. The objective evaluation methods are typically based on the outcome of

psychophysical laboratory experiments using simple artificial stimuli, presented under extremely restricted (impoverished) conditions, and in different experimental paradigms. Other objective methods are based on signal processing techniques and have no obvious counterpart in human vision. So far, only few attempts have been made to validate any of these objective metrics against the performance of human observers in realistic military scenarios. As a result, there are currently no standard and internationally accepted methods and procedures to evaluate camouflage equipment and techniques, and to indicate their military effectiveness. NATO SCI-12 [1] performed a large-scale comparative evaluation study on a common set of registered images (the Search\_2 dataset: [2]) to validate the value of a range of different computational and psychophysical metrics for the quantification of CCD effectiveness [1]. The study included models like CAMAELEON [3], CAMEVA [4,5], the Itti and Koch saliency model [6] and the ORACLE model [7]. The main outcome of that study was that

- visual conspicuity is the best subjective psychophysical predictor of human visual search performance,
- a simple RMS error metric predicts this performance just as well as, or even better than, most other and more complex computational target distinctness metrics.

Since this study, several new camouflage evaluation methods and metrics (e.g. of saliency) have been developed. However, a comparative evaluation study showed that most of the new metrics are outperformed by a simple multiscale contrast model [8]. This finding was later confirmed in a study on the effectiveness of urban camouflage [9]. Also, none of the new camouflage evaluation metrics have been subjected to a comparative evaluation study [10,11].

### **1.1 Observer experiments: field vs laboratory studies**

Given the fact that the goal of camouflage is to decrease the probability of detection, search performance in the field is often considered to provide the golden standard for camouflage assessment. In addition, camouflage also leads to longer search times, a smaller probability of identification, and a higher similarity to non-target objects or 'false targets'. Human target acquisition performance can in principle be measured in field trials using precisely controlled scenarios and closely monitored conditions (see Figure 1). The report of RTO-AG-SCI-095 [12] provides guidelines for the planning and design (e.g., selection of background type, time of day and year, weather conditions), execution (e.g., briefing of observers, observation procedures, required number of observers, number of observations) and data analysis (e.g., methods for testing the significance of data points and the significance of differences observed between different conditions) of field trials for camouflage assessment with human observers. However, a field test is not always the best option to measure search performance, because the information that can be obtained is often quite limited. Furthermore, a field trial is time consuming, labor intensive and logistically difficult, while the number of conditions in which the camouflage can be tested is limited and the variation in performance across conditions is likely to be very large. To obtain a good impression of camouflage performance field tests should be performed over many different conditions. To reduce the influence of environmental factors, camouflage assessments are preferably done in a relative way (e.g., one or several camouflaged targets versus an uncamouflaged reference target) rather than in an absolute manner (e.g., detection distance for this kind of target under these conditions). A recent NATO study concluded that the choice of backgrounds should be a representative mix of the areas in which it will be used [9]. The number of required scenes is also increased by the fact that performance may depend significantly on the precise positioning and lighting conditions (of target and background).



**Figure 1. Example of a field trial using precisely controlled scenarios. Potential target (camouflaged persons) positions indicated by white cards (from [13]).**

Field search trials can in principle be replaced by laboratory experiments, using photo- and video simulation techniques, and imagery that can for instance be recorded in field trials. Laboratory experiments allow complete control over all influential parameters.

According to a NATO study [12] the advantages of laboratory tests include:

- the ability to create a larger statistical data base;
- field imagery sequence collected under one set of environmental conditions;
- the same imagery is viewed by all observers;
- decreased cost compared to observer field trials;
- less time spent in the field collecting data; and
- the conduct of the experiment at multiple locations, using different observer pools, with the same input imagery under controlled laboratory conditions.

However, human performance in artificial laboratory conditions may differ considerably from performance in the real-world scenarios. This may be caused by differences between the simulated environment and the real world, such as limitations in resolution [14], luminance, color [15] reproduction, dynamic range, etc., while the relative importance of each of these factors is still largely unknown. Therefore, the results of laboratory experiments need to be validated by comparing them with the results from similar field experiments [16].

## **1.2 Computational approach**

A promising approach to the evaluation of CCD measures is the use of computational models of the human visual system to analyse recorded or computer-generated scenes, to predict human observer performance. The validation and further development of this approach requires a database of military relevant imagery, in combination with corresponding data on actual observer performance using this imagery. Also, adequate and sufficient laboratory or field measurements are required to perform and link these analyses.

## **1.3 Physical measures**

Camouflage in the visual and infrared aims to minimize the contrast of a target by spectrally matching its surface properties to those of its background. Evaluation criteria can correspond to physical properties such as reflectance, gloss, emissivity and the degree of polarization, as well as dynamic properties and broadband or multispectral properties [17]. Physical properties such as mean luminance, hue and saturation (preferably expressed in a perceptual color space, such as CIELAB color  $L^*a^*b^*$ ), and local temperature differences (in case of LWIR imagery) and their distributions can be measured in various conditions and backgrounds (or in the lab) and compared with the background parameters (e.g. [18]). Such measurements can be very valuable and are related to (but as such do not give a direct prediction of) camouflage performance. When spatial properties are not considered, numerical models can be used to evaluate camouflage performance based on the statistical physical properties of target and background (see section 3). Recent techniques to assess the effectiveness of camouflage include the use of multi- [19] and hyperspectral imagery [20,21]. Multispectral camouflage becomes increasingly important with the increasing availability of sensors operating across the entire electromagnetic spectrum. Physics-based broadband scene rendering tools like CAMEO-SIM [22,23] can be valuable tools for the development and assessment of multispectral camouflage.

## **1.4 Real vs synthetic imagery**

Methods to generate realistic synthetic imagery (e.g., CAMEO-SIM: [22], VBS3) enable the assessment of camouflage effectiveness at lower cost and in a range of widely different scenarios and environmental conditions not easily available in field trials. An accurate camouflage simulation program could effectively support, and possibly replace, field trials. Prior to its use, synthetic imagery must be validated to ensure its fidelity [23,24]. When the camouflage is targeted against sensor threats operating in other than visible spectra (e.g., LWIR) a good simulation requires accurate modelling of the scene as well as the sensor.

There are two ways to measure the fidelity of synthetic imagery [24]. One approach is to directly compare synthetic to real imagery. This is necessary when there is a requirement to accurately reproduce a specific site. This type of metric is very sensitive to geometric changes and the alignment of the two images. It is a very difficult task to exactly match a synthetic image to the real world not only geometrically, but also for the same meteorological conditions. This requires a great deal of ‘ground truth’ data such as the spectral reflectance properties of the different scene components. In many cases this is not necessary, for instance when a more generic assessment of whether camouflage is effective under Northern European conditions or desert conditions is required. There are natural variations in the real world due to variations in the meteorological conditions, time of day, or slightly different types of materials in the scene so an exact match to a particular scene is not necessary. Therefore, another approach is to use image statistics (e.g., the mean and standard deviations of the luminance and color channels in case of a visual color image) to determine whether the features present in the synthetic imagery are similar to those in the real-world imagery. Whether synthetic imagery is realistic enough will also depend on the military task that is being investigated. Lab measures (such as conspicuity and identifiability) have been proposed to quickly assess the fidelity of detection and identification [25,26].

**Table 1. Psychophysical camouflage effectiveness evaluation methods and their associated metrics and (dis-)advantages, with some example studies**

Psychophysical method	Metrics	Pros & cons	Example studies		
			Visual/NIR		Thermal
			Lab	Field	Lab
Range estimation	<ul style="list-style-type: none"> <li>- Detection range</li> <li>- Recognition range</li> </ul>	<ul style="list-style-type: none"> <li>+ Realistic</li> <li>+ Easy to perform in the lab (photosimulation)</li> <li>- Not always possible to perform in the field</li> </ul>	[27,28]	[27]	
Search and detection	<ul style="list-style-type: none"> <li>- Mean search time</li> <li>- (Cumulative) detection probability</li> <li>- Fraction correct</li> </ul>	<ul style="list-style-type: none"> <li>+ Realistic</li> <li>+ Easy to perform on imagery in the lab</li> <li>- Time consuming, needs many replications and observers</li> <li>- Difficult to perform in the field</li> </ul>	[9,16,29-41]	[27,42,43]	[38,39,44-46]
Conspicuity	<ul style="list-style-type: none"> <li>- Conspicuity angle</li> <li>- Conspicuity distance</li> </ul>	<ul style="list-style-type: none"> <li>+ Efficient procedure</li> <li>+ Can be performed in the field and in the lab</li> <li>+ Only few observers needed</li> <li>+ Target position is known by observers</li> <li>- Less suitable for well-camouflaged targets</li> </ul>	[47,48]	[48,49]	[49,50]
Ranking	Rank order	<ul style="list-style-type: none"> <li>+ Can be performed in the field and in the lab</li> <li>+ More efficient than paired comparisons, but</li> <li>- Relative (no absolute) measure</li> <li>- Lower precision than paired comparisons</li> <li>- No information on detection</li> </ul>	[51]		
Paired comparisons	<ul style="list-style-type: none"> <li>- Rank order</li> <li>- Relative performance</li> </ul>	<ul style="list-style-type: none"> <li>+ Higher precision than ranking</li> <li>+ Few observers needed</li> <li>- Relative (no absolute) measure</li> <li>- Less efficient than ranking</li> <li>- Difficult to perform in the field</li> <li>- No information on detection</li> </ul>	[16,52]		[16]
Eye tracking	<ul style="list-style-type: none"> <li>- Nr of fixations</li> <li>- Mean fixation time</li> <li>- Mean saccade length (over 120 different measures in total: [53])</li> </ul>	<ul style="list-style-type: none"> <li>+ Provides objective metrics that correlate with clutter and search performance</li> <li>+ Provides wealth of information per trial (compared to search time)</li> <li>- Requires specialized equipment</li> </ul>	[52,54,55]		[56]
Masked priming	<ul style="list-style-type: none"> <li>- Response time</li> <li>- Error rate</li> </ul>	<ul style="list-style-type: none"> <li>+ Eliminates cognitive and strategic biases</li> <li>+ sensitive to pattern variations</li> </ul>	[57]		
Subjective blending	Rating score	<ul style="list-style-type: none"> <li>+ Easy to perform in lab and field</li> <li>- Subjective measure</li> </ul>	[27,58-60].	[27]	



## **2.0 PSYCHOPHYSICAL EVALUATION METHODS**

This section presents an overview of the various psychophysical evaluation methods that have been used to assess camouflage effectiveness (see also Table 1).

### **2.1 Detection and recognition range**

The most commonly used measure of camouflage effectiveness is the detection range (see Figure 2) [61-63]. The detection range is commonly defined as the distance at which a certain percentage (e.g. 80%) of the observers first perceives the target, or alternatively, the distance at which a given observer perceives the target with a given (e.g. 80%) probability. The shorter the detection range, the more difficult the target is to distinguish.

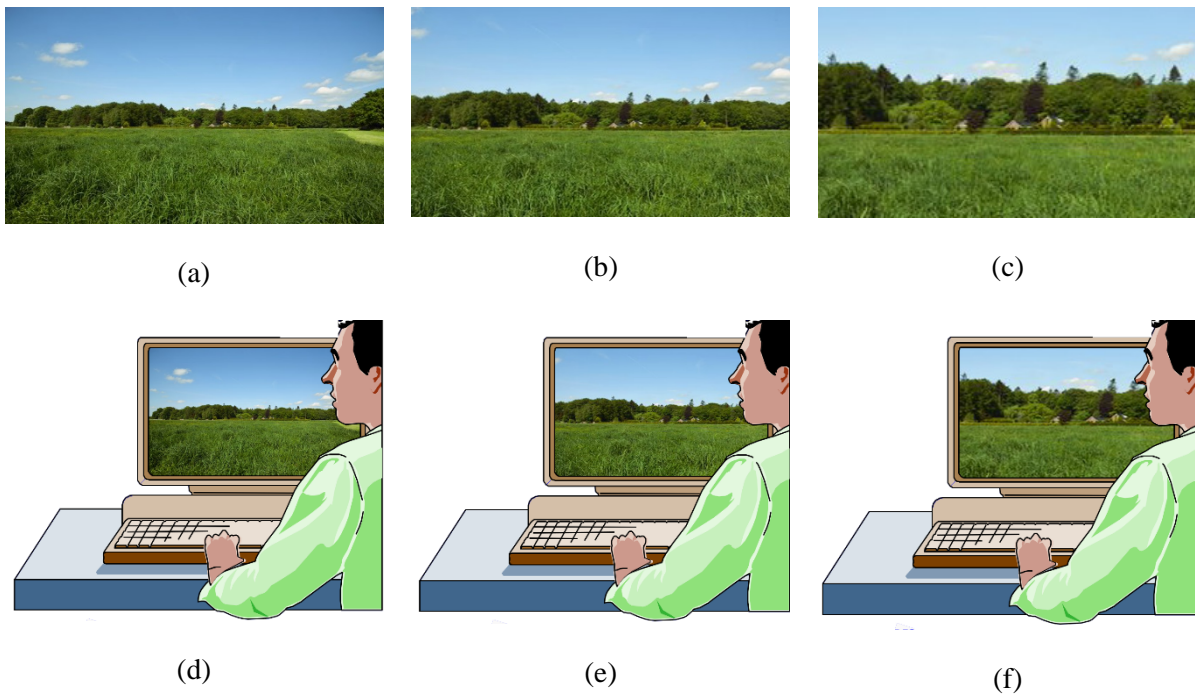
When measuring detection range in the field, the experiment is often done by dispersing targets in an open field so that line of sight is achieved for distances larger than the expected detection range. This way the observer (or imager, in case images are registered to enable observer experiments in the lab at a later time) can approach the target along a straight path from a distance at which the target cannot be resolved [64]. The target position should be known to the observer. In this type of experiments, the measured detection range depends on a number of parameters like the current weather condition, the sun position, the atmospheric visibility (aerosol content), the sensor system and platform used, and not the least the observer's level of experience and training.

In a jungle environment this procedure does not work, because of the lack of long lines of sight. A method to solve this problem is based on the use of panoramic images (Figure 3) [16,39]. The advantage of a panoramic image is its high density of information in one single image. Details of the large image can be inspected by zooming into a certain part of the whole picture. Panoramic images can also be used to study visual scanning behavior with field-of-view restrictions (Figure 4). Zooming allows the creation of a virtual approach sequence, which can be used in photo simulation and observer trial trials to determine a detectability range (instead of a detection range), even if the actual line of sight was very short [16]. Additionally, the virtual approach sequence does not contain changes in illumination and other disturbing natural factors which tend to occur during an original photo simulation imagery sequence. The zoom procedure is an efficient method to assess camouflage effectiveness. The advantage of this method is that the same observer can be used for several approach sequences. The zoom method requires that an observer zooms from a known target backwards until the target starts to blur into the background. This procedure is not new, having been previously used and named the "inverse spy glass" method. A modern version of this method is the software tool "Panoslider" (<http://wirestam.com>). Moving the Panoslider scrollbar to the right zooms in on the target, moving it to the left zooms out. The task of an observer is to move the scrollbar backwards and forwards until a position is found where the target starts to blur into the background.

Detection distances are significantly affected by the local background and shading of the targets. In practice it is often difficult or even impossible to distribute targets over a scene such that they all have the same local background. In that case one needs to successively place targets at the same position in the scene and repeat the detection experiments. However, this may present additional problems when lighting (shadows) changes over time or in partly cloudy conditions. No matter how carefully the experiment is conducted it will always be possible to argue that the results are not generally valid, but valid for this single experiment or class of experiments only. Even though the method has weaknesses with regard to producing statistically representative detection ranges, the method is well suited for comparative experiments, by ranking the camouflage effectiveness of different candidates.

NATO SCI-095 issued testing standards for measuring detection ranges in the lab [12]. This methodology involves showing an observer a sequence of images containing a camouflaged target, with each successive image depicting the target at a closer distance. The observer has a 14-s time window per scene to detect a

target, and the primary outcome is the distance at which observers achieve a 50% detection rate. Camouflage patterns that require presentation at closer distances to achieve 50% detection indicate greater concealment effectiveness. TNO developed an equivalent to this method that uses successive clips from a dynamic video approach sequence [65]. While representing a substantial step forward toward standardizing camouflage test and evaluation methodologies, the technique has several disadvantages; these include the lack of null (target absent) trials, near-ceiling detection effects due to the long exposure duration, subjectively determined target presence, and time-consuming field photography collection and test administration. Furthermore, research has demonstrated limited sensitivity of this technique in differentiating the conspicuity of camouflage patterns within a given pattern class (e.g., desert, woodland, transitional) [59].



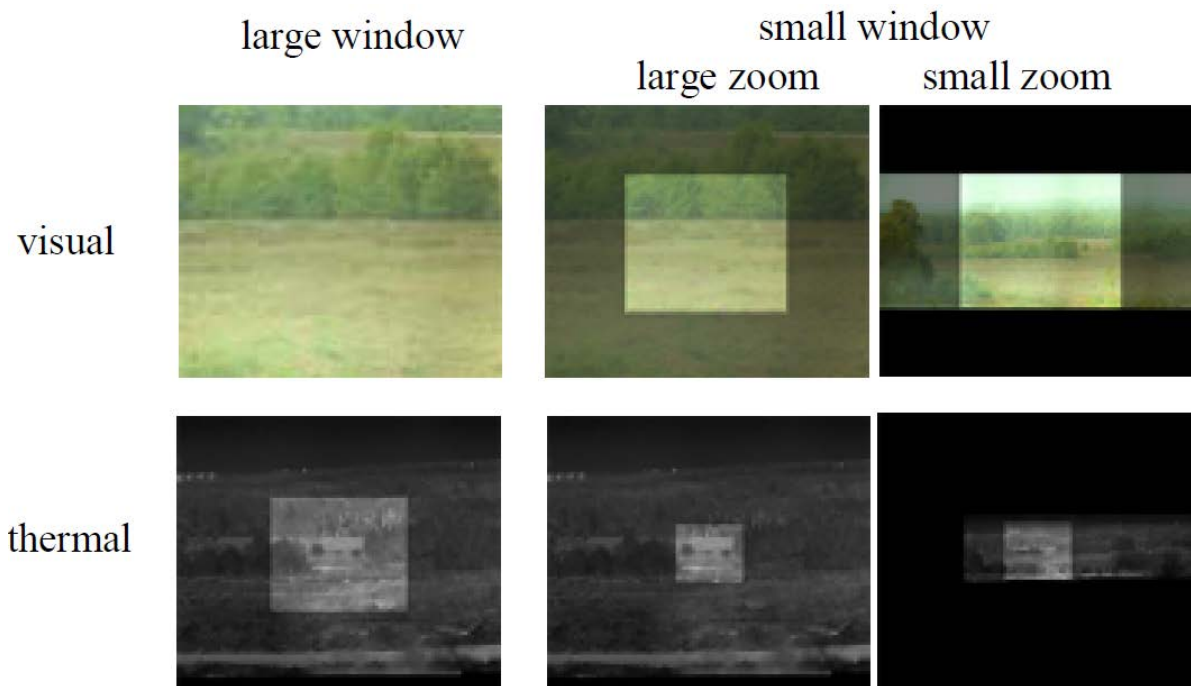
**Figure 2. (a-c) Measuring detection range in the field requires physical approach of the target location, while measurements in photosimulation studies (d-f) involves image enlargement**

## 2.2 Visual search

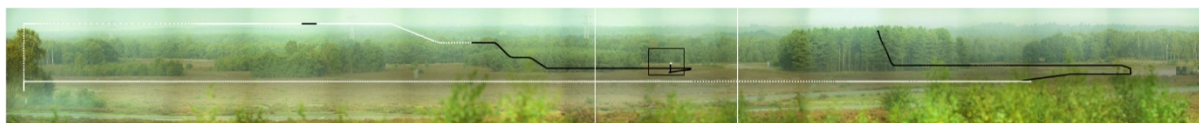
Visual search, with or without the aid of optical or electro-optical instruments, plays a significant role in various types of military operations. Camouflage is typically deployed to reduce the visual signature of targets, thereby maximizing the average time needed by an observer to find the target in the scene. Search time (e.g., [29-31,33,35,55,66]) and detection probability (e.g., [29,33,41,42,67]) are therefore often adopted as metrics to quantify the effectiveness of camouflage. Effective camouflage measures should result in an increase in search time and a decrease in detection probability.

Field search trials can in principle be replaced by experiments in the laboratory using either photographs and videos or computer simulated virtual environments [30,68]. The use of large panoramic scenes enables the investigation of the effects of FOV restrictions on human visual detection performance and search strategies ([67,69]; see Figure 3 and Figure 4). However, human performance in artificial laboratory conditions often differs considerably from performance in the real-world scenarios. The discrepancies in behavior may be caused by a range of factors that differ between the simulated environment and the real world, such as limitations of the representations (e.g. in resolution, luminance, color reproduction etc.), but also

environmental conditions (lab environment vs field conditions). The relative importance of each of these factors is still largely unknown. Other factors that are known to determine human visual search performance are for instance the location of the target in the field of view, scene familiarity and observer learning [70].



**Figure 3. The use of large panoramic scenes enables the investigation of the effects of FOV restrictions and zoom factors on human visual detection performance and search strategies (from [69]).**



**Figure 4. Example of a scan path over a large panoramic visual image with a small FOV. Black and white segments of the scan path correspond to zoom factors of 1 and 0.25, respectively (from [69]).**

### 2.3 Visual conspicuity

Visual search experiments inherently require numerous repetitions and/or a large number of observers to obtain statistically significant data [70]. Since visual search requires the target position to be unknown to the observer, this often is difficult to accommodate. Also, in field experiments it is usually impossible to perform the same scenario more than once, simply because one has no control over the environmental conditions. But even in the laboratory, where one has in principle complete control over all experimental parameters, search experiments are very time consuming because they involve a large number of observations.

It is therefore of great practical value to have a visual target characteristic that

- reliably predicts human visual search performance,
- can easily and quickly be determined in situ, with only a few observers, and



- does not depend on an unknown target location.

Visual target conspicuity has this potential [71]. Visual conspicuity refers to the discrepancy between the visual signatures of a target and its local background. This discrepancy can arise from differences in size, shape, luminance, color, texture, binocular disparity and motion. A target's local feature contrast should exceed the overall variation of the background to make it stand out and attract the observer's attention. Note that an object is not conspicuous per se, it is only conspicuous when it is sufficiently different from its surroundings. It is evident that a target will be easier to detect when it stands out from its background, and the more it stands out the quicker it will be found. It is therefore likely that visual conspicuity is the main determinant of visual search performance.

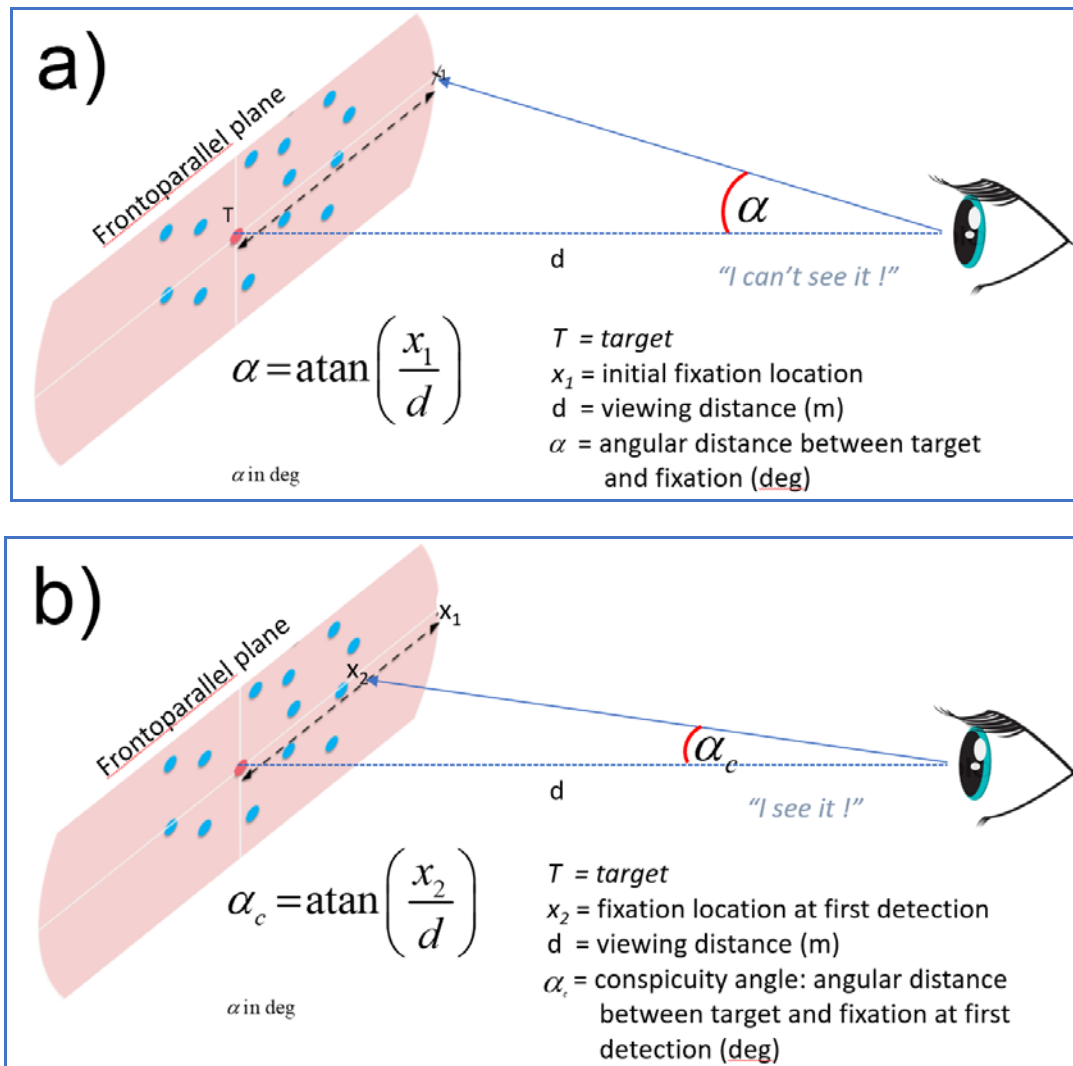
Human visual saliency has been operationalized through the concept of 'conspicuity area', which is defined as the spatial region around the centre of gaze where the target can be detected or identified in the background, within a single fixation [72-74]. Geisler and Chou [75] showed that much of the variation in search times are indeed accounted for by variations in the visual conspicuity area. Thus, visual conspicuity determines human visual behaviour and performance when freely inspecting and interpreting images. Visual conspicuity reflects the uncompromised physical, bottom-up distinctiveness of a target, and involves no high-level top-down processes such as memory, task demands, attention and strategies.

Human visual target conspicuity has operationally been defined as the peripheral area around the centre of gaze, from which specific target information can be extracted in a single glimpse [72-74]. Thus, for fixations falling within this area, the target is capable to attract visual attention. The size and shape of the conspicuity area have been measured for a range of static targets in static scenes [72-79]. The conspicuity area is small if the target is embedded in a background with high feature- or spatial variability. The conspicuity area is large if the target stands out clearly from its background. Importantly, in contrast to many saliency model predictions, the conspicuity area is not only affected by the immediate target background surrounding the target, but also by false/similar targets in the vicinity (the clutter). Psychophysical procedures for measuring the conspicuity area are typically rather intricate and time consuming [72-75]. Here we define human visual conspicuity as the maximum angular gaze deviation at which targets can still be distinguished from their immediate surroundings. We previously developed and validated an efficient psychophysical measurement procedure to quantify this concept [48,71,80-82]. Since saliency is reliably sustained in static stimuli [83], the procedure can be used with full prior knowledge of the target and its location in the scene (which is an advantage in test conditions in which it is difficult to move the target). The resulting saliency estimates determine human visual search and detection performance in realistic and complex scenarios [48,71].

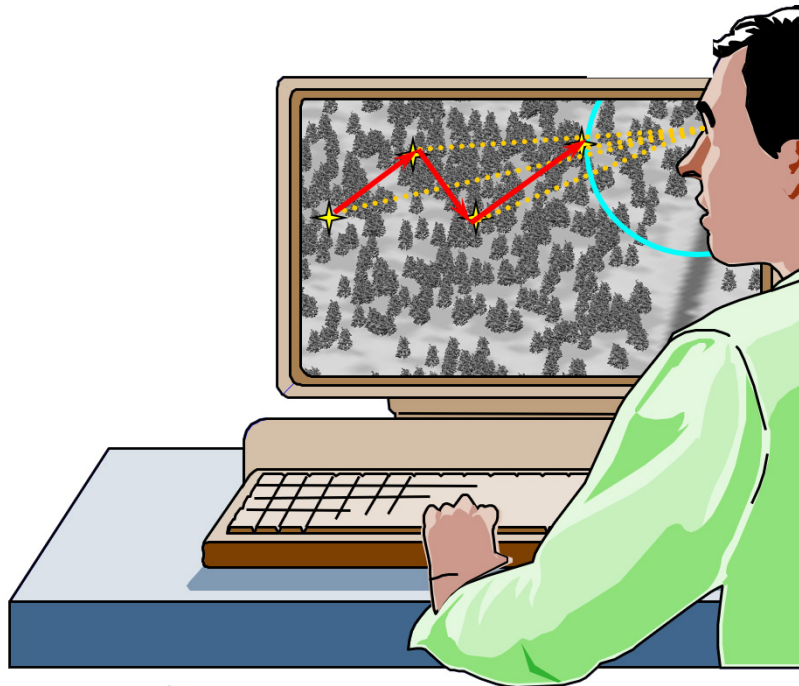
The conspicuity measurement procedure is as follows (Figure 6) [48,71,80-82]. First, the observer foveates the target, to ensure that he has full knowledge of its location and appearance. This is especially relevant for complex scenes where many details compete for the observer's attention. Next, the observer fixates a point in the scene that is both (a) at a large angular distance from the target location, and (b) in the frontoparallel plane through the target. This initial fixation point should be sufficiently remote from the target so that it cannot be distinguished at this stage. The observer then successively fixates locations in the scene that are progressively closer to the target position, until he can perceive (distinguish) the target in his peripheral field of view. The successive fixation points are along a line through the initial fixation point and the centre of the target. The outward-in fixation procedure simulates free search without a priori knowledge of target presence, and prevents hysteresis effects, which may occur for inward-out procedures. The angular distance between the fixation location at which the target is first noted and the center of the target is then recorded. The measurement is repeated at least three times. Subjects typically make a setting within one minute. The mean of the angular distances thus obtained is adopted as the characteristic spatial extent of the conspicuity area of the target, in the direction of the initial fixation point. The method requires some training since the observers should acquire a stable detection criterium. After training, the procedure yields reliable measurements (for 3 repetitions the standard measurement error is typically less than 10%) with only few observers (typically 3), and accounts for important visual effects like crowding [84] and lateral masking

## Review of Camouflage Assessment Techniques

[71,85,86]. Although different observers may use different detection criteria, there is typically a high correlation between their individual measurements [71].



**Figure 5. Schematic representation of the conspicuity distance concept and its associated measurement procedure.** Observers start a conspicuity measurement by first fixating at a location (in the frontoparallel plane through the target) that is far enough removed from the target so that it cannot be detected (a). Then they move their fixation towards the target until it can be discriminated from its background (b). Target conspicuity is then defined as the distance  $x_2$  (in meters in the frontoparallel plane) between the target center and the fixation location at which the target is first detected peripherally. The corresponding conspicuity angle  $\alpha_c$  can be used as a relative measure to compare target conspicuity when all measurements are performed at the same viewing distance  $d$ .



**Figure 6. Schematic representation of the conspicuity measurement procedure in a photosimulation study. The subject initially fixates on the left edge of the screen, far from the location at which the target appears (outside its visual lobe). The yellow crosses indicate the successive fixation locations. The scan path is indicated by the red arrows. The blue circle represents the target conspicuity area or visual lobe.**

Depending on the criterion used by the subjects to assess a target's conspicuity two types of conspicuity estimates can be distinguished: detection and identification conspicuity [71]. Detection conspicuity corresponds to the criterion whether the image structure in the target area differs noticeably from the local background. In this case, previous target fixation merely serves to familiarize the observer with the location of the target. The observer is explicitly instructed not to note the target features. Thus, detection conspicuity is likely to reflect bottom-up saliency. Identification conspicuity corresponds to the criterion whether the image details at the location of the target indeed represent the target or can actually be identified as the target. In this case, the observer is explicitly instructed to use the target features in the discrimination task. Thus, identification conspicuity also contains a top-down saliency component. This is probably the reason why identification conspicuity predicts mean search time more closely than detection conspicuity, since both tasks depend on a priori target information: an image detail is more likely to be fixated when it resembles a target [48]. Identification typically requires a larger feature contrast than detection [87]. Thus, for a given feature contrast, identification saliency is typically smaller than detection saliency. It was recently found that detection and identification indeed involve different cortical mechanisms [87].

Visual conspicuity has for instance successfully been deployed to measure saliency of camouflaged targets both in the lab (Figure 6) [9,47,48] and in the field [48,49], for the validation of simulated visual, near-infrared and thermal imagery [88,89] (see Figure 7), and to assess the detectability of targets in thermal imagery [49] (see Figure 8).

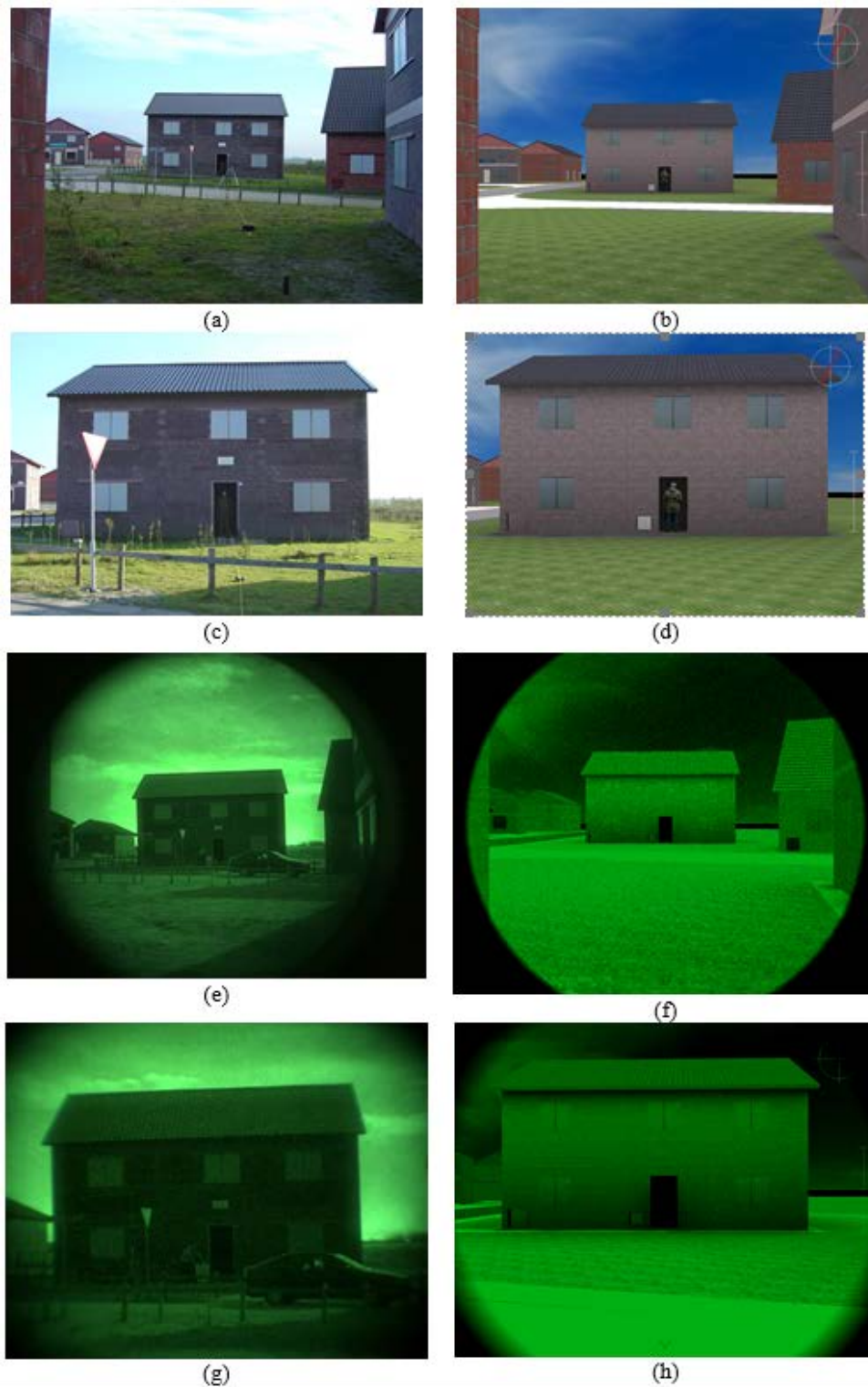
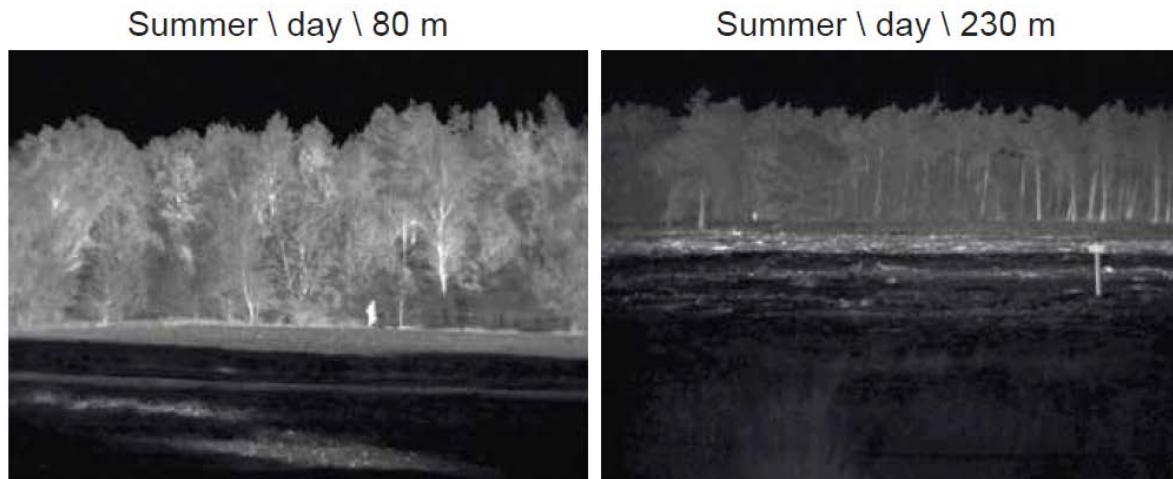


Figure 7. Real (left column) and simulated (right column) daytime (a-d) and nighttime intensified (e-h) imagery of the test location. Conspicuity measurements were used to equalize the detectability of the camouflaged person in the doorway between the real and simulated conditions in daytime and at night [89].





**Figure 8.** The conspicuity of camouflaged persons was measured in thermal imagery taken at different distances (80 and 230 m) and against different backgrounds (left: deciduous trees; right: pine trees).

### 2.4 Ranking experiments

In a ranking experiment, observers order images of competing camouflage designs in a particular scene, from best to worst (Figure 9). Various techniques can be used to analyse the data, such as the Analytical Hierarchy Process [90] or decision support tool such as MARCUS [91]. The ranking method is often used for assessing and comparing different camouflage methods. A disadvantage of a search experiment is that there is a certain amount of time that an observer spends searching in a non-interesting region and therefore brings little useful information, adding noise. This is not the case for a conspicuity or ranking experiment. In contrast with the conspicuity method (and search experiments), observers fixate the target in ranking experiments. Therefore, it may capture the recognition process well (governed by foveal processes) but it is unclear whether it realistically captures the detection process (governed by processes in the periphery, i.e. properties attracting attention).



**Figure 9.** Soldiers performing a ranking experiment (left) on printed images (right; from [16]).



### 2.5 Paired comparisons

When assessing camouflage effectiveness using the methods of paired comparisons (or pairwise comparisons), image stimuli are presented in pairs, and the observers are required to select which image of each pair contains the most distinct target [16,52,92-97]. For each trial in the method of paired comparisons, every stimulus is paired with every other stimulus exactly once. By combining the decisions from all observers, it is possible to estimate numerical scale values for the relative levels of perceived distinctness of the different targets. The established method for generating numerical scale values for the stimuli is the Law of Comparative Judgment (LCJ: [98]), introduced by Thurstone [99]. The LCJ is a set of equations relating the proportion of times any given stimulus  $k$  is judged greater on a given attribute than some other stimulus  $j$  to a scale ranking on the “*psychological continuum*.” The LCJ is based on the postulate that if a stimulus is presented to a human subject, it excites a discrimination process that produces some output value on the psychological continuum. It is also assumed that this value will not be exactly the same each time the same stimulus is presented, but rather these values will form a normal distribution on the psychological continuum. Efficient methods have been developed to design paired comparisons experiments [100,101] and analyze the results [102-105]. Additional information can be obtained by having the observer indicate the estimated difference in camouflage performance between the two stimuli. This information can be used to derive a scalar camouflage metric [16]. Also, the measurement process can be speeded up by combining the ranking method and limiting paired comparisons to those between nearby stimuli. While ranking and paired comparisons yield similar results, precision is greater with the paired comparisons method [106].

### 2.6 Eye tracking

Eye movements can reveal information about the search process that is not immediately obvious from reaction time analysis (Figure 10). Eye movement analysis has for instance shown that fixations on discrete distractors are more frequent than fixations on parts of the background that are similar to the target, even under conditions with high target-background similarity [107]. This finding suggests a biased search process in which salient patterns segmented from the background are inspected while parts of the background that are similar to the target are relatively neglected. Also, for high target-background similarities, foveal hit rates are significantly higher and foveal detection times significantly lower than peripheral ones [55]. In the literature over 120 different eye movement metrics have been presented [53]. It has been found that some eye-tracking metrics (e.g., first saccade amplitude and fixation duration) can objectively differentiate camouflage pattern conspicuity [54]. Also, several eye tracking metrics have been shown to correlate with visual clutter [55,108] and its effects on human visual search [109-113]. Eye tracking [114-116] data can be used to evaluate and tune the performance of target distinctness metrics [52,56] and visual saliency models [113,117-124].

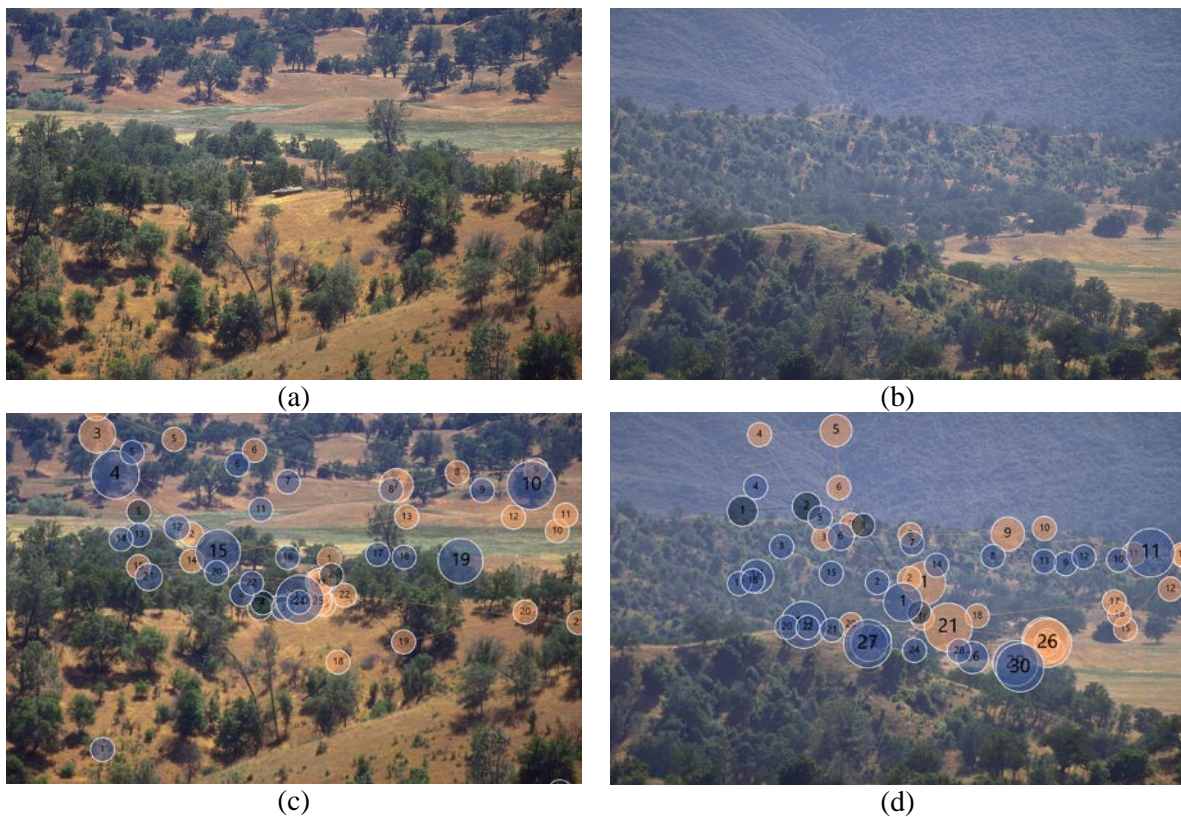


Figure 10. Two search scenes (a,b) with corresponding fixation locations and scan paths (c,d).

## 2.7 Masked priming

Masked priming involves the very brief presentation of a target scene that is flanked by forward (preceding the prime) and backward (following the prime) masks (typically in the form of a scrambled incoherent version of the prime), followed by an empty screen with a probe in a right or left position that is either congruent or incongruent with the target position (Figure 11). Between stimulus presentations the observer fixates a cross in the middle of an otherwise blank screen. Participants are instructed to press a right or left button in response to a right or left dot appearing on the screen, as quickly as possible without compromising accuracy. Critically, masked priming presents the prime for such a brief duration that participants are largely unaware of its presence or characteristics. The masked prime and brief exposure duration limit the extent of processing, thereby revealing the influence of very early perceptual processes. In the experiments of Brunye e.a. [57] participants were primed with camouflaged targets with variable conspicuity (as determined by saliency maps, appearing on the left or right of the screen, and then made a speeded response to a dot probe appearing on the same (congruent) or different (incongruent) side. The premise is that an exogenous shift of attention (i.e., without eye movement) toward a particular region of space will cause faster responses to a dot probe appearing in that same region of space, relative to it appearing in a different region of space. Moreover, the priming effects will typically increase with target signature strength (i.e., they decrease with increasing camouflage effectiveness). Masked priming has successfully been used to quantify the performance of different camouflage patterns both relative to each other and to the background [57].

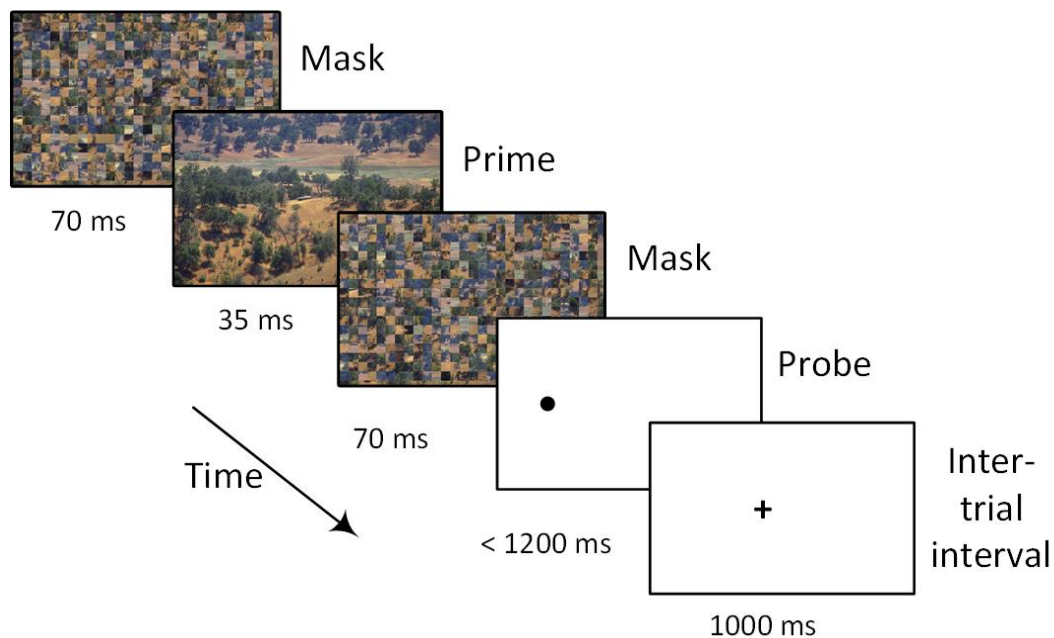


Figure 11. Masked priming procedure.

## 2.8 Subjective blending opinion score

A subjective measure for camouflage effectiveness can be obtained by letting observers rate how well a camouflaged target blends into (matches) its local background, using (Likert or Visual Analog) rating scales [27,58-60]. This method is simple and efficient and can be applied both in the field (using score boards) and in photosimulation studies (Figure 12). The advantage of photosimulation studies is that the same camouflage measure can be tested against (inserted in) a range of different backgrounds [58]. The subjective blending opinion score can also be used in a ranking paradigm [125].



Figure 12. Subjective blending rating procedure for a scene with an inserted camouflage pattern.



### 3.0 COMPUTATIONAL EVALUATION METHODS

This section presents an overview of the various computational evaluation methods that have been used to assess camouflage effectiveness (see also Table 1).

**Table 1. Computational evaluation methods that have been used to assess camouflage effectiveness, their associated metrics, (dis-)advantages and some example studies.**

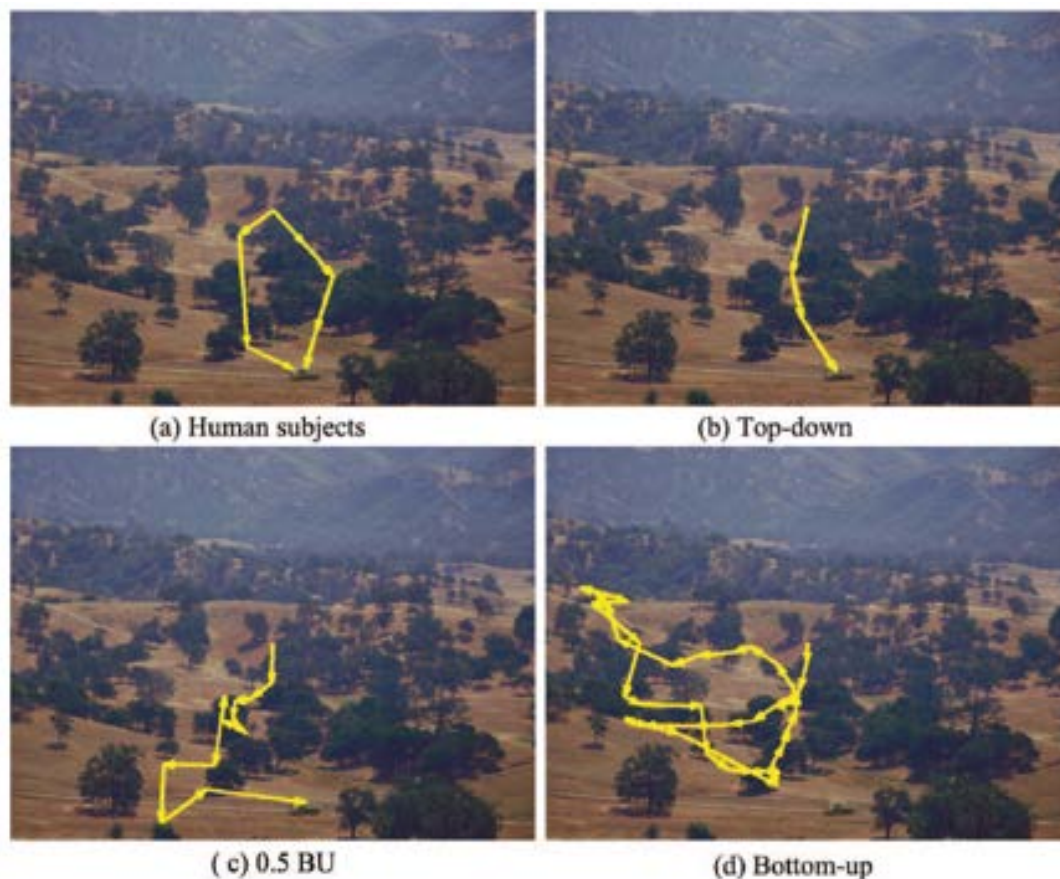
Computational method	Input	Output	Example studies	
			Visual	Thermal
Saliency	Image	Local image distinctness maps	[8,126-129]	[130]
Clutter metrics	Image	Global or local scene complexity	[38,51,131]	[38,92,131,132]
Target signature metrics	Image	Target-background similarity/distinctness	[5,11,32,34,93,133]	[38,44,92]
Search models	- Image - Scene parameters	- Mean search time - Cumulative detection probability	[134]	

#### 3.1 Saliency models

Human visual fixation behaviour is driven both by sensorial bottom-up mechanisms [135,136] and by higher order task-specific or goal-directed top-down (cognitive) mechanisms [137,138]. Visual saliency refers to the physical, bottom-up distinctness of image details [139]. It is a relative property that depends on the degree to which a detail is visually distinct from its background [86]. Visual saliency is believed to drive human fixation behaviour during free viewing, by attracting visual attention in a bottom-up way [140]. As such, it is an important factor in our everyday functioning. Human observer studies have indeed shown that saliency can be a strong predictor of attention and gaze allocation during free viewing, both for static scenes [121,123,141] and for dynamic scenes [142]. Moreover, saliency also appears to determine which details humans find interesting in visual scenes [143].

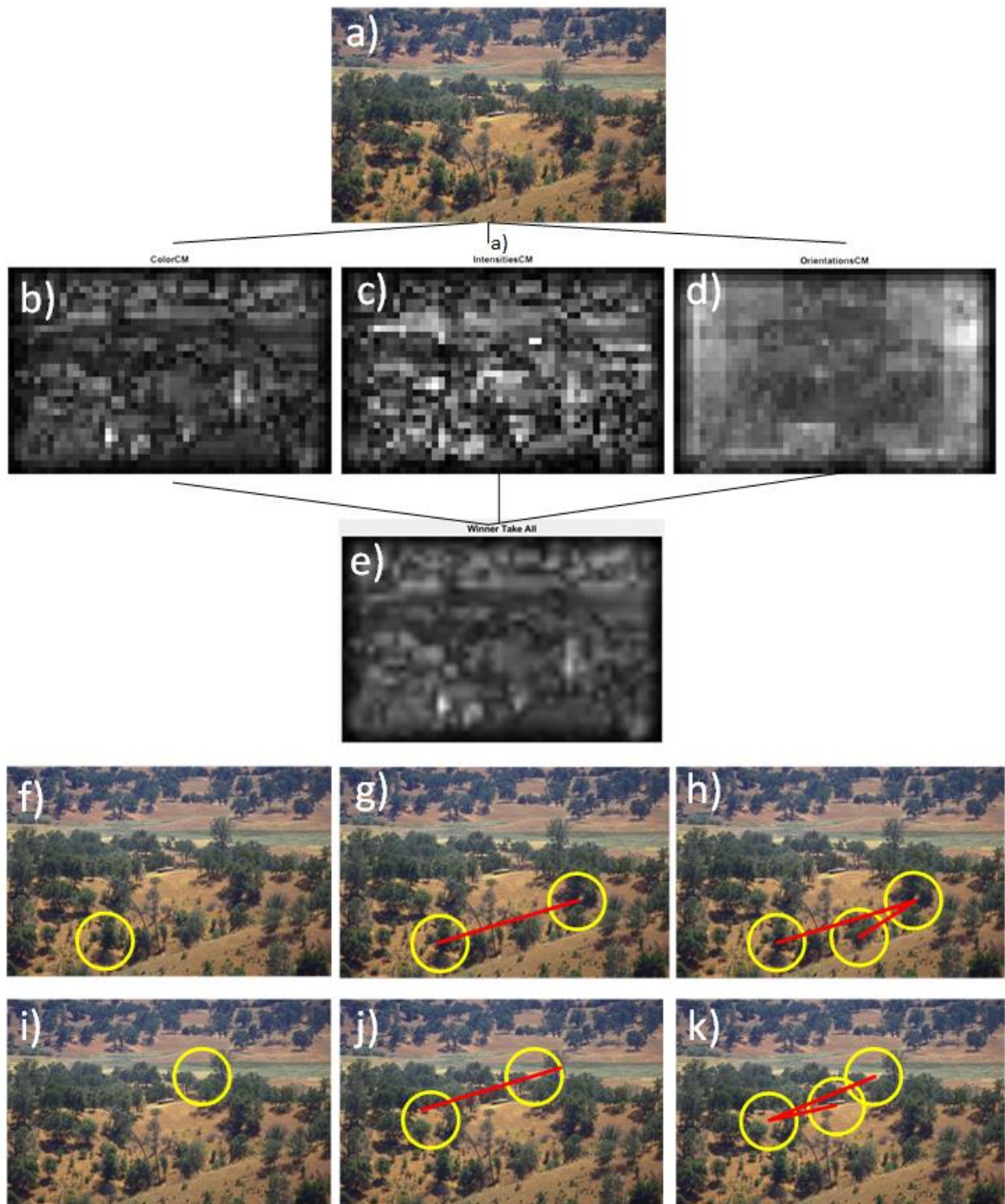
Based on the notion that being a local outlier makes a point salient, Koch and Ullman [136] introduced the concept of a saliency map. A saliency map is a two-dimensional topographic representation of saliency for each pixel in an image. Computational saliency models transform an input image into a saliency map in which signal intensity corresponds to local image saliency. Over the past decade, many different algorithms have been proposed to compute visual saliency maps from digital imagery (for reviews see [8,126,127]). These algorithms typically transform a given input image into a scalar-valued map that represents local image saliency [128,144,145] (see Figure 15). They can be broadly classified as biologically based [6,136,146,147], purely computational [144,148-151], or a combination of both [128,152,153]. Biologically based saliency algorithms typically use centre-surround absolute difference operators to compute local image distinctiveness. Computational saliency algorithms can also take global image distinctiveness into account [153-155]. Some algorithms detect saliency at multiple spatial scales [6,128,146,147,151], while others operate on a single scale [148,149]. They either separately create individual maps which are then combined to obtain the final saliency map [148,149,156,157], or they directly construct a feature combined saliency map (Figure 14) [18,148,151]. Recently saliency algorithms have been extended to produce spatio-temporal saliency maps for the analysis of dynamic imagery [142,155,158-165]. Other recent methods combine saliency maps with other images features like local regularity [166], entropy [167] or linear features [168]. Only a few saliency algorithms have been validated with observer data [129,130].

Where human observers look in images is often used as a ground truth estimate of image saliency [113,117-124]. Saliency models can be used to predict human visual fixation behavior by successively moving the focus of interest (Figure 13) between local maxima of the saliency maps (Figure 15), e.g. by using winner-take-all or inhibition of return approaches [169]. However, objectively determining which of the many available saliency models offers the ‘best’ approximation to human eye fixations remains a challenge [170]. Not only do the different saliency algorithms produce widely different saliency maps (Figure 15), but there are also many different evaluation metrics available that all produce different (ranking) scores [170]. Scenes containing objects in textured and cluttered backgrounds are still challenging for most models [127]. As a result, the saliency predictions produced by most models do not agree well with saliency as perceived by human observers. What current models are missing is a close interaction between different modules like object detection, segmentation, tracking, and categorization. Contextual information and prior knowledge from other modules may enable an attentive system to obtain better, more useful regions of interest [126]. For instance, a pure top-down saliency model yielded better predictions of human visual fixation behavior than a pure bottom-up or a hybrid model (see Figure 13).



**Figure 13. Comparison of human and model scan paths for different top-down and bottom-up saliency map weightings (from [113]).**





**Figure 14. Visual search scene showing a target vehicle in the center (a), saliency maps for color (b), intensity (c) and orientation (d) and the overall saliency map (e). (f-h) First three fixations when all individual saliency maps are weighted equally in the construction of the overall saliency map. (i-k) First three fixations when color is neglected and only intensity and orientation contribute to the overall saliency map.**

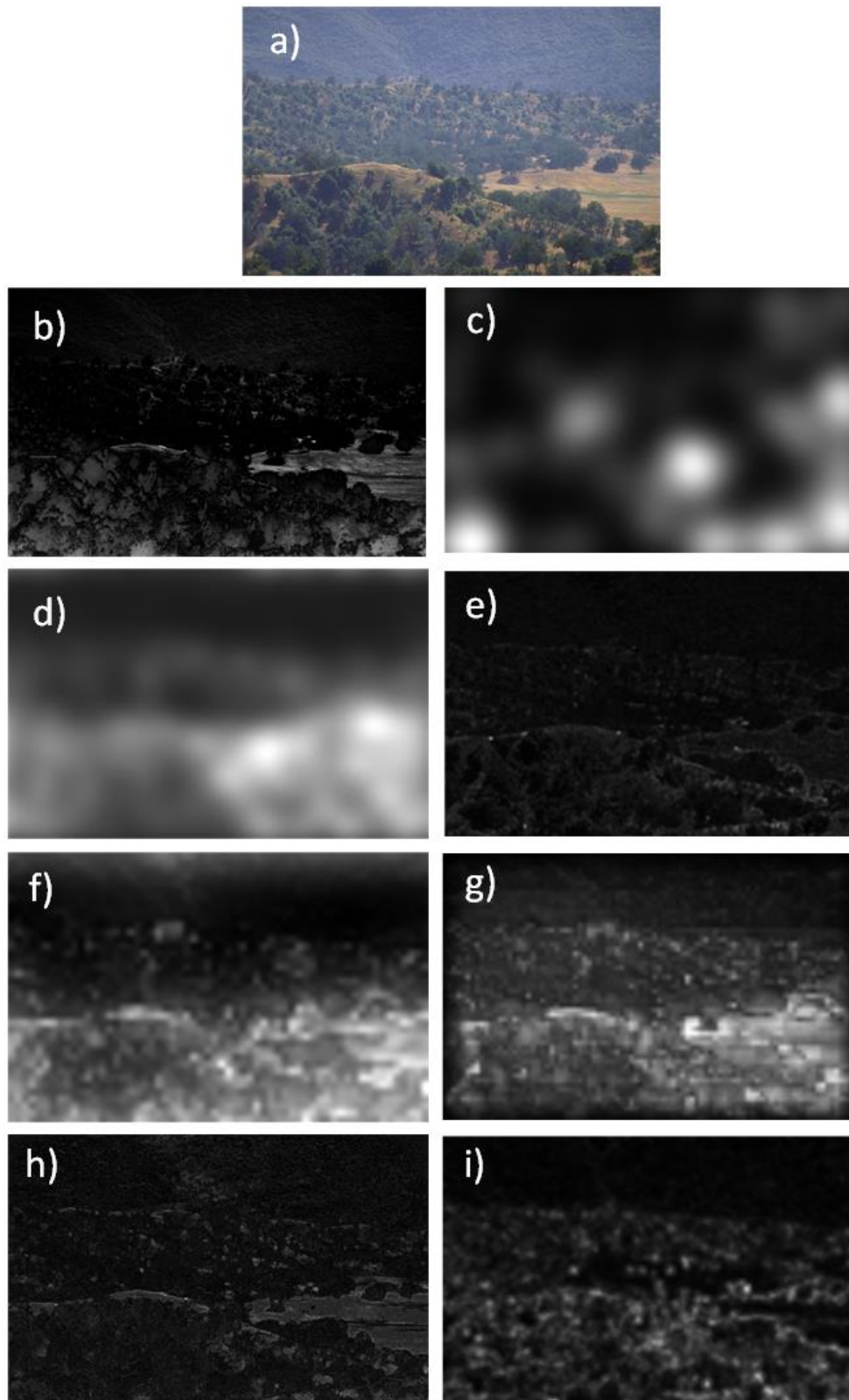


Figure 15. Visual search scene (a) with corresponding saliency maps produced by the algorithms of [144](b), [171] (c), [172,173] (d), [165] (e), [153] (f), [6] (g), [155,174] (h), [159,175] (i).

However, most of these metrics reflect target relevance instead of target saliency [139], since the human visual performance measures used in their construction (e.g. fixation order, subjective ranking, search and detection time) involve high-level top-down processes such as memory, task demands, attention and strategies. In most cases it is not bottom-up saliency, but the most ‘interesting’ or ‘meaningful’ object in an image that attracts attention [176]. In the presence of phase-dependent higher-order image features (e.g., lines, edges, symmetries), lower-order features only have a correlative effect on human fixation behaviour [177-179]. Moreover, task demands override sensory driven (bottom-up) saliency [141,180-182]. It is only in the absence of high-level features (e.g., in natural scenes where the global scene lay-out lacks diagnostic properties) and prescribed tasks that lower-order features guide attention to some extent [178]. As a result, bottom-up saliency models are limited in their ability to predict fixation behavior [141,181,183-187] (e.g., Figure 13). In realistic visual search conditions, top-down attention dominates, and people are rarely distracted by visually salient but irrelevant items [183]. But even in free viewing conditions, when top-down processing is minimal, bottom-up saliency models are poor predictors of human fixation behavior [188-191].

### 3.2 Global clutter metrics

Clutter metrics typically quantify either global or local image complexity. Image complexity (or the global amount of clutter) depends on the number of its structural components, their heterogeneity, (e.g., a single shape repeated vs. multiple distinct shapes), their regularity (e.g., simple polygons vs. more abstract shapes) and the regularity of the arrangement of elements (e.g., symmetry, distribution characteristics; see Figure 16).

An object can be distinguished or detected visually when its retinal image differs in some way from the image of its surroundings. The human visual system utilizes differences in size, shape, luminance, color, texture, binocular disparity and motion to achieve this figure-ground segregation. A sufficient difference between a target and its surround in any of these factors is sufficient to make a target stand out. Local feature contrast should exceed the overall variation in a pattern to allow a target to stand out from its background and be detected. For instance, Rosenholtz’s Feature Congestion clutter metric [192] is based on the assumption that an object’s visual distinctness or saliency depends on the degree to which its local feature contrast exceeds the local feature variation in a scene. Rosenholtz’s Subband Entropy clutter metric [193] recognizes that visual saliency depends on the degree of organization of the scene in which the target is situated. A highly organized (redundant) scene appears less cluttered. As a result, target conspicuity will increase with increasing scene regularity. Rosenholtz developed a third measure of clutter (Edge Density clutter) based on the finding that a target is more conspicuous when it is placed in an area with less detail [193]. Bravo and Farid’s Scale Invariant clutter metric [194] was designed to predict the performance (mean search time) of observers searching for non-salient targets with unknown features at unknown positions in unstructured scenes and may therefore especially applicable to camouflaged targets. Another clutter metric that is directly associated with the amount of detail in a scene is the *Number of Proto-Objects* (NPO) or the number of image segments or super-pixels with similar intensity, color and gradient orientation features (the proto-objects: [195]). Larger NPO values correspond to higher levels of visual clutter.

The *Compression Ratio* (CR) between the original (uncompressed) and (lossless JPEG or GIF format) compressed file sizes is another computational measure that positively correlates with subjective image complexity ratings ([196]). The file size of a digitized image is a measure of its structural information content ([197]). Compression algorithms use image redundancy or predictability to reduce the file size, such that more complex (or less predictable) images need more elements. Hence, larger compression ratios correspond to larger perceived image complexity (higher clutter).



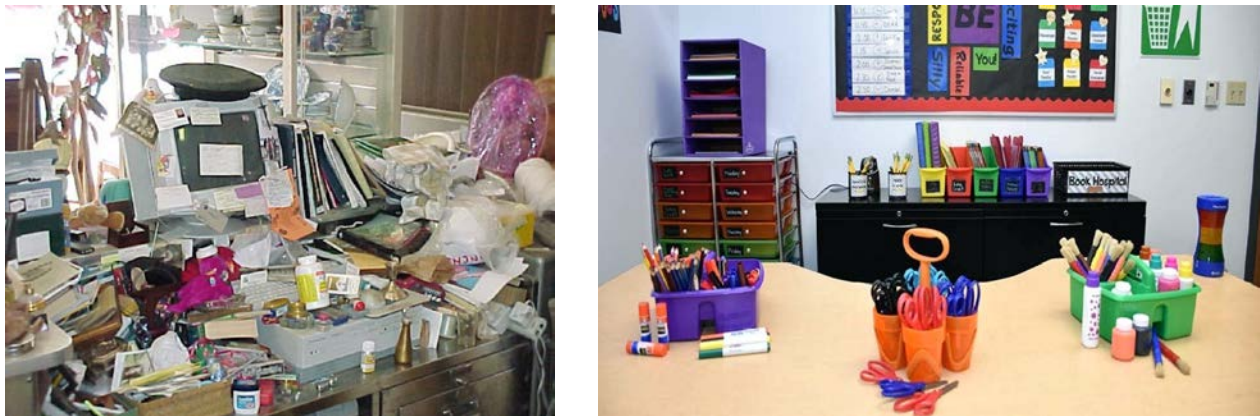
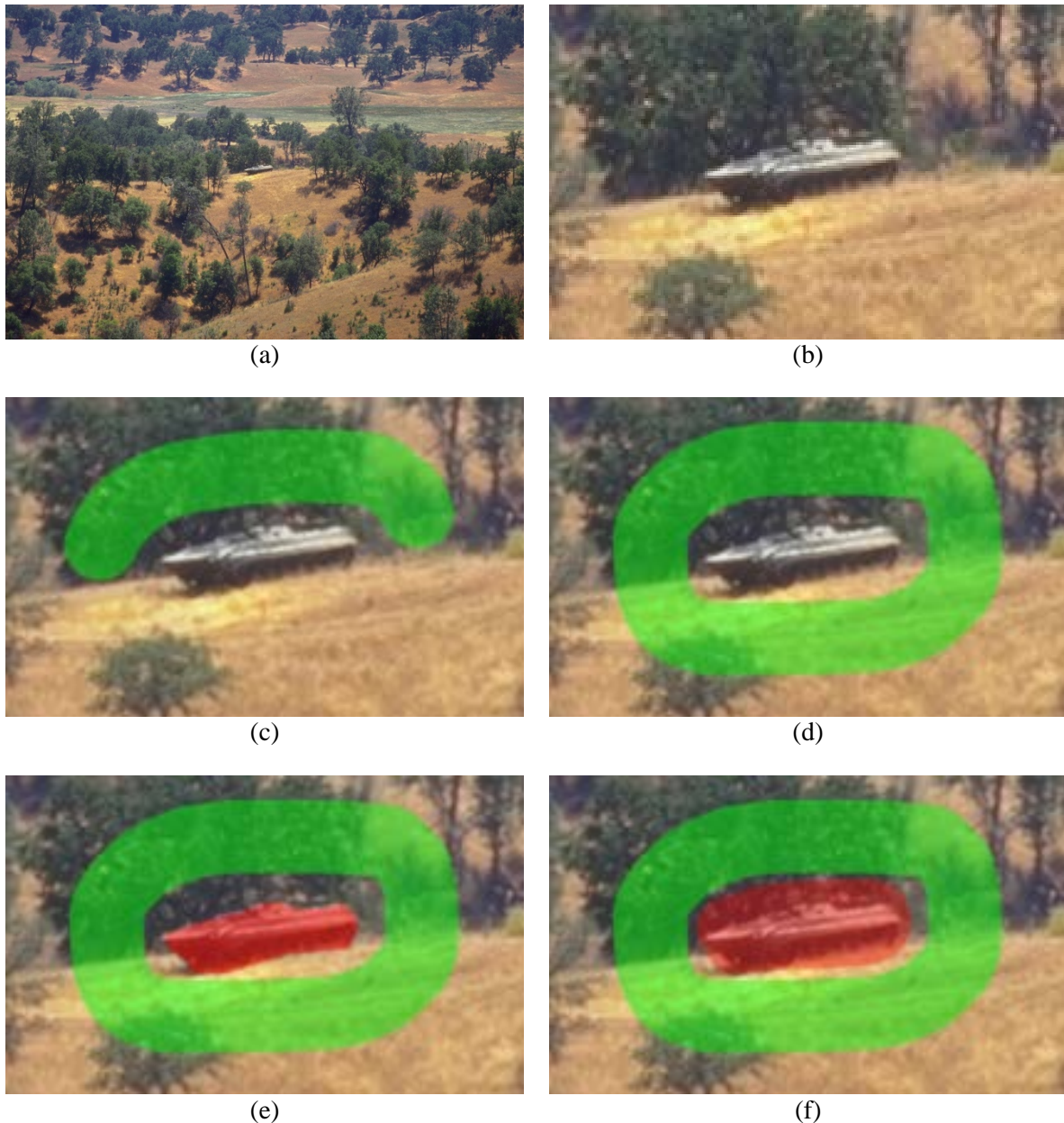


Figure 16. A cluttered scene (left) versus an organized one (right).

### 3.3 Local clutter metrics

Local clutter metrics typically quantify the distinctness of a target by computing the structural (dis-)similarity between the target and its local background [11,93]. A range of different local clutter (or target signature distinctness) metrics have been proposed, such as the Mean Square Error (MSE: [54]), Contrast [52], Average Co-occurrence Error [52], Doyle metric [95], Average Boundary Strength (ABS: [52,93]), the Gabor Edge Disruption Ratio (GabRat: [61]), the Improved Co-Occurrence Matrix (ICOM: [38,198-201]), the Peak Signal Noise Ratio (PSNR: [54]), the Probability of Edge (POE: [38,198,200-202]), the Universal Image Quality Index (UIQI: [55]), the Target structural similarity metric (TTSSIM: [133]) and various other texture measures [203,204]. While most of these target signature metrics are typically computed over the entire image support using a sliding window, some metrics (e.g., the CAMouflage Assessment by Evaluation of Local Energy, Spatial frequency and Orientation method or CAMAELEON [205] and the CAMouflage EVALuation method or CAMEVA [5]) require the user to manually indicate a target and local background region, making their results quite variable (highly dependent on the actual choice of the selected regions) and unreliable (see Figure 17).

The NATO-RTO SCI-12 Working Group identified standardized camouflage assessment methods and target signature metrics that are highly correlated with operational effectiveness [1]. They concluded that man-in-the-loop assessment is the only robust and effective method to evaluate camouflage. Recent evaluation studies agree with these previous findings, for instance that there is no clear relation between the output of the CAMAELEON signature metric and observer ratings for camouflaged targets [32,34]. Although several new methods have been introduced since the conclusion of this study (e.g., [10,11,61,166-168,204,206,207]) computational signature analysis methods still do not fully represent the range of significant visual and cognitive processes driving target acquisition performance [34]. Some promising methods that appear to reliably predict human visual detection of camouflaged targets are target-background similarity metrics like Structural Similarity (SSIM: [133]), the Universal Image Quality Index (UIQI: [11]), and the Gabor Edge Disruption Ratio (GabRat: [61]).



**Figure 17. (a) Search scene with (b) target in the middle. (c-f) Subjective manual definitions of local target background (green: c,d) and target support (red: e,f). The definition of the local background may be quite ambiguous (c vs d), while the accuracy of the target definitions may vary highly between observers (e vs f).**

## 3.4 Search models

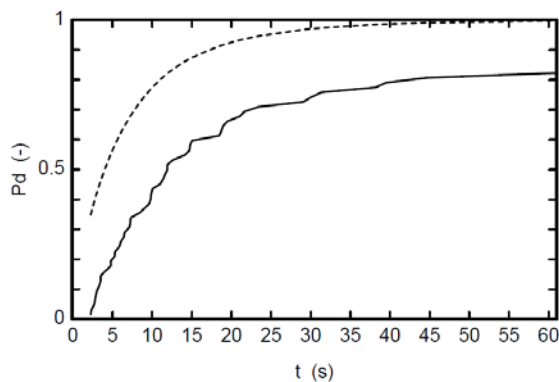
Another approach to avoid field experiments is to predict human observer performance using a validated search model. A range of computational models have been developed to predict the mean time a human observer needs to detect a given target while visually inspecting a scene (e.g., ORACLE: [7,208], Visdet: [209,210], ACQUIRE: [46,211-213]; for a review see [134]). These models use either imagery from the scene in question (or a scene with similar characteristics) or a list of parameters describing the visual characteristics of the scene (e.g., mean luminance of the target, of its local background, and of the overall scene, the dimensions of both the target sector and the entire search sector, the dimensions of the



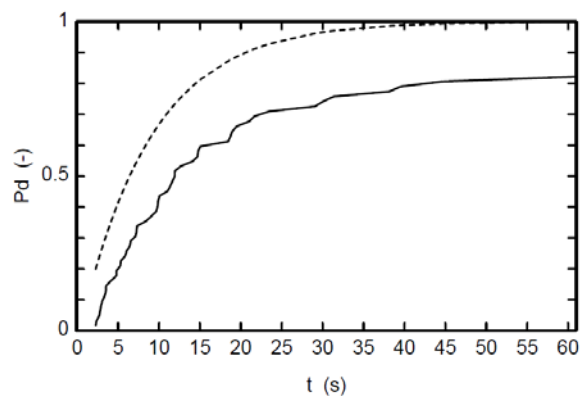
instantaneous field of view, etc.) to predict cumulative target detection probability (Figure 18). Visual search, however, is a very complex process that is not yet fully understood. This holds especially for visual search in complex realistic scenes. As a result, computational models of human visual search are not yet sophisticated enough to deal with all the factors that affect search performance, and the accuracy of their predictions is very limited [1,134].



(a)



(b)



(c)

**Figure 18. (a) Search scene with military vehicle (in the center of the white box), with the cumulative detection probability as measured in a human observer experiment (solid curves) and the predictions by the ORACLE (b) and Visdet (c) search models (dashed curves; from [214]).**

## 4.0 CONCLUSION AND DISCUSSION

In military operations signature reduction techniques such as camouflage nets, low-emissive paints, and camouflage patterns are typically deployed to optimize the survivability of high value assets by minimizing their detectability. Various methods have been developed to assess the effectiveness of these camouflage measures. There are three main approaches to the evaluation of camouflage measures: (1) a subjective approach through observer experiments, (2) an objective computational approach through image analysis, and (3) an objective approach through physical measurements. Although subjective evaluation methods have a direct relation with the operational practice, they are often difficult to implement because of time and budget restrictions (these experiments typically require a lot of observers, repetitions and different targets and backgrounds), or simply because the associated conditions are not safe for the observers. The objective evaluation methods are typically based on the outcome of psychophysical laboratory experiments using simple artificial stimuli, presented under extremely restricted (impoverished) conditions, and in different experimental paradigms. Other objective methods are based on signal processing techniques and have no obvious counterpart in human vision. Only a few attempts have been made to validate any of these objective metrics against the performance of human observers in realistic military scenarios. As a result, there are currently no standard and internationally accepted methods and procedures to evaluate camouflage equipment and techniques, and to indicate their military effectiveness. Benchmark datasets (e.g., like the Search\_2 dataset [2]) containing imagery and associated human camouflage performance measures would be useful to further develop, optimize and compare models that seek to predict camouflage effectiveness. Furthermore, such datasets can be used to assess the fidelity of simulation techniques used in modelling camouflage performance.

Camouflage assessment is typically performed in a limited number of controlled and static settings and dynamic environmental aspects are largely ignored. As a result, it is not clear how camouflage performance measured during evaluation studies translates to added benefits when used in military operations and different settings. Hence, it is crucial that new camouflage design and evaluation procedures take military strategies and contexts into account [215]. Thereto, a set of critical realistic operational scenarios and contexts (e.g., ambush, patrol and hasty defense) should be established (in close interaction with military experts) on which camouflage evaluation can be based. In addition, a set of benchmark data should be collected according to these scenarios and contexts to afford the design and (comparative) evaluation of camouflage assessment techniques and models. Such an approach should account for the operational and dynamic aspects, as well the variability in context for different scenarios, and could for instance involve the use of wargaming (e.g., VBS) and highly realistic photosimulation techniques [215]. Only few studies address the effects of motion on camouflage effectiveness [16,49]. Fraunhofer IOSB developed the CART (Camouflage Assessment in Real Time: [216]) system for camouflage assessment of static and moving objects in multispectral image data, which addresses both the problems of moving camera and moving objects.

In some conditions the added value of camouflage will be marginal (e.g., in dense forest) while in other situations camouflage can make an important contribution to survivability. The investigation of camouflage aspects that effectively enhance survivability should therefore be part of the evaluation procedure and should be incorporated into models that seek to predict performance in a military context [215]. Aspects that contribute little to camouflage effectiveness in certain scenarios or contexts should then be attributed less weight in the overall rating of the design than aspects that contribute more significantly.

Camouflage development tends to follow sensor development. As a camouflage for a particular sensor is developed and proved effective, new sensors are developed to counter the camouflage. Sensor and detection system development is progressing in two directions: (a) miniaturization and extended wavelength multispectral sensors and (b) active source/sensor systems [217]. Two areas of particular interest in camouflage research at present are active (dynamic) camouflage [218,219] and multispectral universal camouflage. Active camouflage is defined as a system that changes to match the environment in real time.

Multispectral universal camouflage is still the holy grail of camouflage research. A recent review [220] summarizes the state of the art.

## REFERENCES

- [1] NATO-RTO SCI-12 "Search and target acquisition", RTO Meeting Proceedings 45, NATO RTO, Neuilly-sur-Seine, France (1999).
- [2] A. Toet, P. Bijl, and J. M. Valetton, "Image dataset for testing search and detection models", *Optical Engineering*, 40(9), 1760-1767 (2001).
- [3] R. Hecker, "Camaeleon - Camouflage assessment by evaluation of local energy, spatial frequency and orientation", *Characterization, Propagation, and Simulation of Sources and Backgrounds II*, SPIE-1687, The International Society for Optical Engineering, Bellingham, WA, 342-349 (1992).
- [4] C. M. Birkemark, "Theoretical and practical aspects of CAMEVA", Danish Defence Research Establishment, Copenhagen, Denmark (1997).
- [5] C. M. Birkemark, "CAMEVA, a methodology for estimation of target detectability", *Optical Engineering*, 40(9), 1835-1843 (2001).
- [6] L. Itti, C. Koch, and E. Niebur, "A model of saliency-based visual attention for rapid scene analysis", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(11), 1254-1259 (1998).
- [7] K. J. Cooke, P. A. Stanley, and J. L. Hinton, *The ORACLE approach to target acquisition and search modelling*, World Scientific, Singapore (1995).
- [8] A. Toet, "Computational versus psychophysical image saliency: a comparative evaluation study", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 33(11), 2131-2146 (2011).
- [9] A. Toet, and M. A. Hogervorst, "Urban camouflage assessment through visual search and computational saliency", *Optical Engineering*, 52(041103), 1-8 (2013).
- [10] T. N. Volonakis, O. E. Matthews, E. Liggins *et al.*, "Camouflage assessment: Machine and human", *Computers in Industry*, 99, 173-182 (2018).
- [11] C. J. Lin, C.-C. Chang, and B.-S. Liu, "Developing and evaluating a target-background similarity metric for camouflage detection", *PLOS ONE*, 9(2), e87310 (2014).
- [12] NATO-RTO AC/323(SCI-095)TP/96 "Guidelines for camouflage assessment using observers", RTO-AG-SCI-095, (2006).
- [13] W. R. Watkins, G. D. Heath, M. D. Phillips *et al.*, "Search and target acquisition: single line of sight versus wide baseline stereo", *Optical Engineering*, 40(9), 1914-1927 (2001).
- [14] S. Ramsey, T. Mayo, C. Howells *et al.*, "Modeling apparent camouflage patterns for visual evaluation", *Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XXIV*, SPIE-10644, International Society for Optics and Photonics, 1064428 (2018).
- [15] S. Ramsey, T. Mayo, A. Shabaev *et al.*, "Modeling apparent color for visual evaluation of camouflage fabrics", *Optics and Photonics for Information Processing XI*, 10395, International Society for Optics and Photonics, 103950I (2017).
- [16] NATO Task Group SCI-219 "Camouflage in hot humid areas", STO-TR-SCI-219, NATO S&T Organization, Neuilly-sur-Seine, France (2014).
- [17] C. Åkerlind, J. Fagerström, T. Hallberg *et al.*, "Evaluation criteria for spectral design of camouflage", *Target and Background Signatures*, SPIE-965303, SPIE, Bellingham, WA, USA, 1-13 (2015).
- [18] G. K. Selj, and D. H. Heinrich, "A field-based method for evaluating thermal properties of static and mobile camouflage", *Target and Background Signatures IV*, SPIE-10794, 107940B-1.
- [19] X. Li, X. Lv, Y. Wang *et al.*, "Performance evaluation of pattern painting camouflage based on the texture characteristic of multi-spectral images", *MIPPR 2007: Automatic Target Recognition and Image Analysis; and Multispectral Image Acquisition*, SPIE-6787, The International Society



- for Optical Engineering, (2007).
- [20] A. Zavvartorbat, H. Dehghani, and A. J. Rashidi, "Evaluation of camouflage effectiveness using hyperspectral images", *Journal of Applied Remote Sensing*, 11(4), 045008 (2017).
  - [21] D. Pearce, and J. Feenan, "High dynamic range hyperspectral imaging for camouflage performance test and evaluation", *Target and Background Signatures II*, SPIE-9997, International Society for Optics and Photonics, 99970F (2016).
  - [22] I. R. Moorhead, M. A. Gilmore, A. W. Houlbrook *et al.*, "CAMEO-SIM: A physics-based broadband scene simulation tool for assessment of camouflage, concealment and deception methodologies", *Optical Engineering*, 40(9), 1896-1905 (2001).
  - [23] S. Newman, M. A. Gilmore, I. R. Moorhead *et al.*, "Validation of the use of synthetic imagery for camouflage effectiveness assessment", *Targets and Backgrounds VIII: Characterization and Representation*, SPIE-4718, The International Society for Optical Engineering, Bellingham, WA, 35-45 (2002).
  - [24] M. A. Gilmore, I. R. Moorhead, G. H. Watson *et al.*, "Assessment of synthetic image fidelity", SPIE-4718, The International Society for Optical Engineering, 23-34 (2002).
  - [25] A. Toet, M. A. Hogervorst, and P. Bijl, "Conspicuity and identifiability: efficient calibration tools for synthetic imagery", SPIE-5431, The International Society for Optical Engineering, 1-12 (2004).
  - [26] M. A. Hogervorst, A. Toet, P. Bijl *et al.*, "Identifiability: an efficient measure for identification performance", SPIE-5407, The International Society for Optical Engineering, 96-103 (2004).
  - [27] R. F. Mortlock, "Camouflage combat pattern case study", NPS-AM-18-219, Naval Postgraduate School, Monterey, CA, USA (2018).
  - [28] M. A. Hogervorst, and A. Toet, "Evaluation of the effectivity of camouflage for equipment in combination with fractal pattern uniform designs (in Dutch)", TNO 2013 R10647, TNO Behavioural and Societal Sciences, Soesterberg, The Netherlands (2013).
  - [29] G. K. Selj, and M. Söderblom, "Discriminating between camouflaged targets by their time of detection by a human-based observer assessment method", *Target and Background Signatures*, SPIE-9653, SPIE, 965305-965314 (2015).
  - [30] G. K. Selj, and D. H. Heinrich, "Search by photo methodology for signature properties assessment by human observers", *Signal Processing, Sensor/Information Fusion, and Target Recognition XXIV*, SPIE-9474, SPIE, 947411-947412 (2015).
  - [31] G. K. Selj, and D. H. Heinrich, "Disruptive coloration tricks the human eye - a study of detection times of two near similar targets in natural backgrounds", *Target and Background Signatures II*, SPIE-9997, SPIE, 99970A-99970A-15 (2016).
  - [32] D. H. Heinrich, and G. K. Selj, "The effect of contrast in camouflage patterns on detectability by human observers and CAMAELEON", *Automatic Target Recognition XXV*, SPIE-9476, International Society for Optics and Photonics, Bellingham, WA, 947604 (2015).
  - [33] G. K. Selj, and D. H. Heinrich, "Disruptive coloration in woodland camouflage: evaluation of camouflage effectiveness due to minor disruptive patches", *Target and Background Signatures II*, SPIE-9997, 99970A-99970A-15 (2016).
  - [34] D. H. Heinrich, and G. K. Selj, "Evaluation of camouflage pattern performance of textiles by human observers and CAMAELEON", *Target and Background Signatures III*, SPIE-10432, International Society for Optics and Photonics, 1043206 (2017).
  - [35] F. Racek, A. Jobánek, T. Baláž *et al.*, "Evaluation of validity of observer test for testing of camouflage patterns", *Target and Background Signatures IV*, SPIE-10794, International Society for Optics and Photonics, Bellingham, WA, 107940E (2018).
  - [36] V. C. Wheaton, and J. B. Culpepper, "Performance of target distinctness metrics evaluated against colour and monochromatic photosimulation results", *Target and Background Signatures*, SPIE-9653, SPIE, Bellingham, WA, USA, 96530U-9 (2015).
  - [37] J. B. Culpepper, A. J. Meehan, Q. T. Shao *et al.*, "Mirage: a visible signature evaluation tool", *Target and Background Signatures III*, SPIE-10432G, 1-20.
  - [38] G. Aviram, and S. R. Rotman, "Analyzing the effect of imagery wavelength on the agreement

- between various image metrics and human detection performance of targets embedded in natural images", *Optical Engineering*, 40(9), 1877-1884 (2001).
- [39] D. M. Deaver, S. Moyer, and C. Ra, "Modeling defined field of regard (FOR) search and detection in urban environments", *Infrared Imaging Systems: Design, Analysis, Modeling, and Testing XVII*, SPIE-6207, SPIE, Bellingham, WA, 8-19 (2006).
  - [40] NATO-RTO SCI-157 "Urban camouflage for the individual soldier system", NATO Science and Technology Organization, Neuilly-sur-Seine, France (2010).
  - [41] Q. Jia, X. L. Lv, W. D. Xu *et al.*, "Intelligent design of gradual disruptive pattern painting and comparison of camouflage effectiveness", *Cluster Computing*, Online, 1-9 (2018).
  - [42] J. B. Culpepper, and V. C. Wheaton, "A target detection model predicting field observer performance in maritime scenes", *Electro-Optical and Infrared Systems: Technology and Applications XI*, SPIE-9249, SPIE, Bellingham, WA, 92490C-924915 (2014).
  - [43] J. B. Culpepper, V. C. Wheaton, C. S. Madden *et al.*, "Methods for measuring time to detect in human observer trials", *Target and Background Signatures*, SPIE-10794, SPIE, Bellingham, WA, (2018).
  - [44] T. J. Meitzler, D. Ellis, E. Sohn *et al.*, "A perception test of infrared images of soldiers in camouflaged uniforms", *Targets and Backgrounds: Characterization and Representation IV*, SPIE-3375, 2-16.
  - [45] S. R. Rotman, M. L. Kowalczyk, J. F. Cartier *et al.*, "Understanding infrared clutter through fixation point analysis", *Infrared Imaging Systems: Design, Analysis, Modeling, and Testing V*, SPIE-2224, The International Society for Optical Engineering, Bellingham, WA, 142-154 (1994).
  - [46] J.-h. Wang, and W.-q. Jin, "Research and development on performance models of thermal imaging systems", *International Symposium on Photoelectronic Detection and Imaging 2009*, SPIE-7383, SPIE, (2009).
  - [47] A. Toet, S. C. de Vries, P. Bijl *et al.*, "Validation of military target representation in a simulator through conspicuity", *TNO Human Factors*, Soesterberg, The Netherlands (1996).
  - [48] A. Toet, F. L. Kooi, P. Bijl *et al.*, "Visual conspicuity determines human target acquisition performance", *Optical Engineering*, 37(7), 1969-1975 (1998).
  - [49] J. A. Beintema, and A. Toet, "Conspicuity of moving soldiers", *Infrared Imaging Systems: Design, Analysis, Modeling, and Testing XXII*, SPIE-8014-9, SPIE, Bellingham, WA, 1-12 (2011).
  - [50] A. Toet, "Visual conspicuity of targets in synthetic IR imagery", *TNO Human Factors*, Soesterberg, The Netherlands (1999).
  - [51] A. J. Meehan, and J. B. Culpepper, "Clutter estimation and perception", *Optical Engineering*, 55(11), 113106-113106 (2016).
  - [52] A. C. Copeland, and M. M. Trivedi, "Computational models for search and discrimination", *Optical Engineering*, 40(9), 1885-1895 (2001).
  - [53] K. Holmqvist, M. Nyström, R. Andersson *et al.*, *Eye tracking: A comprehensive guide to methods and measures*, Oxford University Press, Oxford, UK (2011).
  - [54] C. J. Lin, C.-C. Chang, and Y.-H. Lee, "Evaluating camouflage design using eye movement data", *Applied Ergonomics*, 45(3), 714-723 (2014).
  - [55] C.-C. Chang, Y.-H. Lee, C. J. Lin *et al.*, "Visual assessment of camouflaged targets with different background similarities", *Perceptual and Motor Skills*, 114(2), 527-541 (2012).
  - [56] S. R. Rotman, M. L. Kowalczyk, and V. George, "Modeling human visual search and target acquisition performance: fixation-point analysis", *Optical Engineering*, 33(11), 3803-3809 (1994).
  - [57] T. T. Brunyé, M. D. Eddy, M. S. Cain *et al.*, "Masked priming for the comparative evaluation of camouflage conspicuity", *Applied Ergonomics*, 62, 259-267 (2017).
  - [58] L. Hepfinger, C. Stewardson, K. Rock *et al.*, "Soldier camouflage for operation enduring freedom (OEF): pattern-in-picture (PIP) technique for expedient human-in-the-loop camouflage assessment", *27th Army Science Conference*, Army Natick Soldier Research Development and Engineering Center, (2010).



- [59] K. Rock, Leshner, L.L., Stewardson, C., Isherwood, K., Hepfinger, L., "Photosimulation camouflage detection test", NATICK/TR-09/021L, US Army Natick Research, Development and Engineering Center, Natick, MA, USA (2009).
- [60] M. A. Hogervorst, A. Toet, and J. Beintema, "Design and evaluation of camouflage patterns for application to uniforms", TNO-DV 2009 A422, TNO Human Factors, Soesterberg, The Netherlands (2009).
- [61] J. Troscianko, J. Skelhorn, and M. Stevens, "Quantifying camouflage: how to predict detectability from appearance", *BMC Evolutionary Biology*, 17(1), 7 (2017).
- [62] H. E. Guttman, and R. G. Webster, "Determining the detectability range of camouflaged targets", *Human Factors*, 14(3), 217-225 (1972).
- [63] T. Nilsson, "Evaluation of target acquisition difficulty using recognition distance to measure required retinal area", *Optical Engineering*, 40(9), 1827-1834 (2001).
- [64] F. B. Olsen, "Methods for evaluating thermal camouflage", Paper presented at the RTO SCI Symposium on "Sensors and Sensor Denial by Camouflage, Concealment and Deception", RTO-MP-SCI-145, NATO RTO, Neuilly-sur-Seine, France, 16: 1-14 (2005).
- [65] M. A. Hogervorst, A. Toet, and P. Bijl, "The TOD method for dynamic image quality assessment", TNO Defense, Security and Safety, Soesterberg, The Netherlands (2006).
- [66] G. K. Selj, "Disruptive camouflage tricks the human eye: a study of detection times of two near-similar targets in natural backgrounds", *Target and Background Signatures*, SPIE-9653, SPIE, 96530S-965311 (2015).
- [67] M. A. Hogervorst, A. Toet, and P. Bijl, "Human search performance with limited field of view: effects of scan speed, aperture size, and target conspicuity", *Optical Engineering*, 52(4), 041106-1-14 (2013).
- [68] A. Toet, and P. Bijl, "Search experiment example", *Encyclopedia of optical engineering*, Marcel Dekker Inc., New York, USA, 2566-2576 (2003).
- [69] M. A. Hogervorst, P. Bijl, and A. Toet, "On the relationship between human search strategies, conspicuity and search performance", *Infrared Imaging Systems: Design, Analysis, Modeling, and Testing XVI*, SPIE-5784, SPIE, Bellingham, WA., USA, 240-251 (2005).
- [70] A. Toet, and P. Bijl, "Visual search", *Encyclopedia of optical engineering*, Marcel Dekker Inc., New York, USA, 2949-2954 (2003).
- [71] A. Toet, and P. Bijl, "Visual conspicuity", *Encyclopedia of optical engineering*, Marcel Dekker Inc., New York, USA, 2929-2935 (2003).
- [72] F. L. Engel, "Visual conspicuity and selective background interference in eccentric vision", *Vision Research*, 14(7), 459-471 (1974).
- [73] F. L. Engel, "Visual conspicuity, visual search and fixation tendencies of the eye", *Vision Research*, 17(1), 95-108 (1977).
- [74] F. L. Engel, "Visual conspicuity. Directed attention and retinal locus", *Vision Research*, 11(6), 563-575 (1971).
- [75] W. S. Geisler, and K.-L. Chou, "Separation of low-level and high-level factors in complex tasks: visual search", *Psychological Review*, 102(2), 356-378 (1995).
- [76] J. R. Bloomfield, "Visual search in complex fields: size differences between target disc and surrounding discs", *Human Factors*, 14(2), 139-148 (1972).
- [77] Y. M. Bowler, *Towards a simplified model of visual search*, Taylor & Francis, London, UK (1990).
- [78] B. L. Cole, and S. E. Jenkins, "The effect of variability of background elements on the conspicuity of objects", *Vision Research*, 24(3), 261-270 (1984).
- [79] S. E. Jenkins, and B. L. Cole, "The effect of the density of background elements on the conspicuity of objects", *Vision Research*, 22(10), 1241-1252 (1982).
- [80] F. L. Kooi, and A. Toet, "Conspicuity: an efficient alternative for search time", TNO Human Factors, Soesterberg, The Netherlands (1998).
- [81] A. Toet, P. Bijl, F. L. Kooi *et al.*, "Quantifying target distinctness through visual conspicuity", SPIE-3375, 152-163.

- [82] F. L. Kooi, and A. Toet, "Conspicuity: an efficient alternative for search time", VII, Elsevier Science Ltd, 451-462 (1999).
- [83] F. J. A. M. Poirier, F. Gosselin, and M. Arguin, "Perceptive fields of saliency", *Journal of Vision*, 8(15), 14-1-14-19 (2008).
- [84] D. M. Levi, "Crowding - An essential bottleneck for object recognition: A mini-review", *Vision Research*, 48(5), 635-654 (2008).
- [85] A. H. Wertheim, I. T. C. Hooge, K. Krikke *et al.*, "How important is lateral masking in visual search?", *Experimental Brain Research*, 170(3), 387-402 (2006).
- [86] A. H. Wertheim, "Visual conspicuity: A new simple standard, its reliability, validity and applicability", *Ergonomics*, 53(3), 421-442 (2010).
- [87] S. Straube, and M. Fahle, "Visual detection and identification are not the same: Evidence from psychophysics and fMRI", *Brain and Cognition*, In Press, (2010).
- [88] A. Toet, "Visual conspicuity of targets in synthetic IR imagery", *Camouflage, Concealment and Deception Evaluation Techniques*, RTO-TM-017 AC/323(SCI-012)TP/32, 137-152.
- [89] A. Toet, F. L. Kooi, F. Kuijper *et al.*, "Objective assessment of simulated daytime and NVG image fidelity", *Enhanced and Synthetic Vision 2005*, SPIE-5802-13, 1-10.
- [90] J. Baumbach, "Using the Analytical Hierarchy Process (AHP) to evaluate target signatures", *Land Warfare Conference 2008*, (2009).
- [91] E. J. Emond, and D. W. Mason, "A new rank correlation coefficient with application to the consensus ranking problem", *Journal of Multi-Criteria Decision Analysis*, 11(1), 17-28 (2002).
- [92] G. Aviram, and S. R. Rotman, "Evaluation of human detection performance of targets embedded in natural and enhanced infrared images using image metrics", *Optical Engineering*, 39(4), 885-896 (2000).
- [93] A. C. Copeland, and M. M. Trivedi, "Signature strength metrics for camouflaged targets corresponding to human perceptual cues", *Optical Engineering*, 37(2), 582-591 (1998).
- [94] A. C. Copeland, and M. M. Trivedi, "Models and metrics for signature strength evaluation of camouflaged targets", *Algorithms for Synthetic Aperture Radar Imagery IV*, SPIE-3070, The International Society for Optical Engineering, Bellingham, WA, USA, 194-199 (1997).
- [95] A. C. Copeland, M. M. Trivedi, and J. R. McManamey, "Evaluation of image metrics for target discrimination using psychophysical experiments", *Optical Engineering*, 35(6), 1714-1722 (1996).
- [96] A. C. Copeland, M. M. Trivedi, and J. R. McManamey, "Psychophysical experiments for evaluating target distinctness in images", *SPIE-2469*, 452-461.
- [97] P. G. Engeldrum, *Psychometric scaling: a toolkit for imaging systems development*, Imcotek Press, Winchester, MA (2000).
- [98] W. S. Torgerson, "Theory and methods of scaling", Wiley, New York, USA (1958).
- [99] L. L. Thurstone, "A law of comparative judgement", *Psychological Review*, 34, 273-286 (1927).
- [100] U. Graßhoff, and R. Schwabe, "Optimal Designs for the Bradley-Terry Paired Comparison Model", *Otto-von-Guericke-Universität Magdeburg, Fakultät für Mathematik, Magdeburg, Germany* (2006).
- [101] D. A. Silverstein, and J. E. Farrell, "Efficient method for paired comparison", *Journal of Electronic Imaging*, 10(2), 394-398 (2001).
- [102] T. C. Brown, and G. L. Peterson, "An enquiry into the method of paired comparison: reliability, scaling, and Thurstone's Law of Comparative Judgment", *U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station, Fort Collins, CO* (2009).
- [103] R. Rajae-Joordens, and J. Engel, "Paired comparisons in visual perception studies using small sample sizes", *Displays*, 26(1), 1-7 (2005).
- [104] R.-C. Tsai, and T.-L. Wu, "Analysis of Paired Comparison Data Using Mx", *Structural Equation Modeling: A Multidisciplinary Journal*, 11(1), 73-91 (2004).
- [105] F. Wickelmaier, and C. Schmid, "A Matlab function to estimate choice model parameters from paired-comparison data", *Behavior Research Methods, Instruments, & Computers*, 36(1), 29-40 (2004).
- [106] J. C. Handley, J. S. Babcock, and J. B. Pelz, "Experimental congruence of interval scale

- production from paired comparisons and ranking for image evaluation", SPIE-5294, 211-221.
- [107] M. B. Neider, and G. J. Zelinsky, "Searching for camouflaged targets: Effects of target-background similarity on visual search", *Vision Research*, 46(4), 2217-2235 (2006).
  - [108] N. M. Moacdieh, and N. B. Sarter, "Eye tracking metrics: A toolbox for assessing the effects of clutter on attention allocation", *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 56(1), 1366-1370 (2012).
  - [109] N. M. Moacdieh, and N. Sarter, "Using eye tracking to detect the effects of clutter on visual search in real time", *IEEE Transactions on Human-Machine Systems*, 47(6), 896-902 (2017).
  - [110] N. D. B. Bruce, D. P. Loach, and J. K. Tsotsos, "Visual correlates of fixation selection: a look at the spatial frequency domain", 2007 IEEE International Conference on Image Processing, 3, IEEE Computer Society, Washington, USA, III-289-III-292 (2009).
  - [111] M. Cerf, J. Harel, W. Einhäuser *et al.*, "Predicting human gaze using low-level saliency combined with face detection", *Advances in neural information processing systems*, 20, MIT Press, Cambridge, MA, 241-248 (2008).
  - [112] O. Le Meur, P. Le Callet, and D. Barba, "Predicting visual fixations on video based on low-level visual features", *Vision Research*, 47(19), 2483-2498 (2007).
  - [113] G. Zelinsky, W. Zhang, B. Yu *et al.*, "The role of top-down and bottom-up processes in guiding eye movements during visual search.", 18, MIT Press, 1569-1576 (2006).
  - [114] C. M. Masciocchi, and J. D. Still, "Alternatives to eye tracking for predicting stimulus-driven attentional selection within interfaces", *Human-Computer Interaction*, 28(5), 417-441 (2013).
  - [115] A. Garcia-Diaz, X. Fdez-Vidal, X. M. Pardo *et al.*, "Saliency from hierarchical adaptation through decorrelation and variance normalization", *Image and Vision Computing*, 30(1), 51-64 (2012).
  - [116] Q. Zhao, and C. Koch, "Learning a saliency map using fixated locations in natural scenes", *Journal of Vision*, 11(3-9), 1-15 (2011).
  - [117] N. Riche, M. Duvinage, M. Mancas *et al.*, "A study of parameters affecting visual saliency assessment".
  - [118] R. J. Peters, A. Iyer, L. Itti *et al.*, "Components of bottom-up gaze allocation in natural image", *Vision Research*, 45(18), 2397-2416 (2005).
  - [119] J. Stoll, M. Thrun, A. Nuthmann *et al.*, "Overt attention in natural scenes: Objects dominate features", *Vision Research*, 107(0), 36-48 (2015).
  - [120] D. Sen, and M. Kankanhalli, "A bio-inspired center-surround model for salience computation in images", *Journal of Visual Communication and Image Representation*, in press, 277-288 (2015).
  - [121] B. W. Tatler, R. J. Baddeley, and I. D. Gilchrist, "Visual correlates of fixation selection: effects of scale and time", *Vision Research*, 45(5), 643-659 (2005).
  - [122] N. Ouerhani, R. von Wartburg, and H. Hugli, "Empirical validation of the saliency-based model of visual attention", *Electronic Letters on Computer Vision and Image Analysis*, 3(1), 13-24 (2004).
  - [123] D. Parkhurst, K. Law, and E. Niebur, "Modeling the role of salience in the allocation of overt visual attention", *Vision Research*, 42(1), 107-123 (2002).
  - [124] L. Itti, and C. Koch, "Target detection using saliency-based attention", *Search and Target Acquisition*, RTO-MP-45 AC/323(SCI)TP/19, North Atlantic Treaty Organization, paper 3, 1-10 (2000).
  - [125] M. A. Hogervorst, "Ontwerp NLD Fractal Pattern-Tan", TNO-DV 2011 A426, TNO Behavioural and Societal Sciences, Soesterberg, The Netherlands (2012).
  - [126] T. V. Nguyen, Q. Zhao, and S. Yan, "Attentive systems: A survey", *International Journal of Computer Vision*, 126(1), 86-110 (2018).
  - [127] A. Borji, M. Cheng, H. Jiang *et al.*, "Salient object detection: A benchmark", *IEEE Transactions on Image Processing*, 24(12), 5706-5722 (2015).
  - [128] A. Schwegmann, "Camouflage evaluation by bio-inspired local conspicuity quantification", *Target and Background Signatures IV*, SPIE-10794, 107940H.
  - [129] X. Feng, C. Guoying, H. Richang *et al.*, "Camouflage texture evaluation using a saliency map", *Multimedia Systems*, 21(2), 169-175 (2015).

- [130] X.-p. Cheng, D.-p. Zhao, Z.-j. Yu *et al.*, "Effectiveness evaluation of infrared camouflage using image saliency", *Infrared Physics & Technology*, 95, 213-221 (2018).
- [131] S. R. Rotman, and G. Aviram, "Clutter metrics for predicting human target acquisition performance", *Targets and Backgrounds VII: Characterization and Representation*, SPIE-4370, The International Society for Optical Engineering, Bellingham, WA, 114-119 (2001).
- [132] M. Bernhardt, M. I. Smith, P. Whitehead *et al.*, "A statistical sea surface clutter model in the long-wave infra-red", *SPIE-5431*, 270-278.
- [133] A. Toet, "Structural similarity determines search time and detection probability", *Infrared Physics & Technology*, 53(6), 464-468 (2010).
- [134] A. Toet, P. Bijl, and J. M. Valeton, "Test of three visual search and detection models", *Optical Engineering*, 39(5), 1344-1353 (2000).
- [135] S. J. Dickinson, H. I. Christensen, J. K. Tsotsos *et al.*, "Active object recognition integrating attention and viewpoint control", *Computer Vision and Image Understanding*, 67(3), 239-260 (1997).
- [136] C. Koch, and S. Ullman, "Shifts in selective visual attention: towards the underlying neural circuitry", *Human Neurobiology*, 4(4), 219-227 (1985).
- [137] M. Land, N. Mennie, and J. Rusted, "The roles of vision and eye movements in the control of activities of daily living", *Perception*, 28(11), 1311-1328 (1999).
- [138] A. Yarbus, *Eye movements and vision*, Plenum Press, New York, USA (1967).
- [139] J. H. Fecteau, and D. P. Munoz, "Saliency, relevance, and firing: A priority map for target selection", *Trends in Cognitive Sciences*, 10(8), 382-390 (2006).
- [140] E. Niebur, and C. Koch, "Computational architectures for attention", 163-186.
- [141] G. Underwood, and T. Foulsham, "Visual saliency and semantic incongruity influence eye movements when inspecting pictures", *The Quarterly Journal of Experimental Psychology*, 59(11), 1931-1949 (2006).
- [142] L. Itti, "Quantifying the contribution of low-level saliency to human eye movements in dynamic scenes", *Visual Cognition*, 12(6), 1093-1123 (2005).
- [143] L. Elazary, and L. Itti, "Interesting objects are visually salient", *Journal of Vision*, 8(3), 1-15 (2009).
- [144] R. Achanta, S. Hemami, F. Estrada *et al.*, "Frequency-tuned salient region detection", *IEEE International Conference on Computer Vision and Pattern Recognition (CVPR 2009)*, IEEE Press, 1597-1604 (2009).
- [145] L. Zhang, M. H. Tong, T. K. Marks *et al.*, "SUN: A Bayesian framework for saliency using natural statistics", *Journal of Vision*, 8(7), 1-20 (2008).
- [146] D. Walther, *Interactions of visual attention and object recognition: computational modeling, algorithms, and psychophysics*, California Institute of Technology, Pasadena, California (2006).
- [147] D. Walther, and C. Koch, "Modeling attention to salient proto-objects", *Neural Networks*, 19(9), 1395-1407 (2006).
- [148] Y.-F. Ma, and H.-J. Zhang, "Contrast-based image attention analysis by using fuzzy growing", 374-381.
- [149] Y. Hu, X. Xie, W.-Y. Ma *et al.*, "Salient region detection using weighted feature maps based on the human visual attention model", 993-1000.
- [150] X. Hou, and L. Zhang, "Saliency detection: a spectral residual approach", 1-8.
- [151] R. Achanta, F. Estrada, P. Wils *et al.*, "Salient region detection and segmentation", *Computer Vision Systems (ICVS 2008)*, Lecture Notes in Computer Science 5008, Springer, Berlin-Heidelberg, Germany, 66-75 (2008).
- [152] N. D. B. Bruce, and J. K. Tsotsos, "Saliency, attention, and visual search: An information theoretic approach", *Journal of Vision*, 9(3), 1-24 (2009).
- [153] J. Harel, C. Koch, and P. Perona, "Graph-based visual saliency", 19, 545-552.
- [154] M. Mancas, L. Couvreur, B. Gosselin *et al.*, "Computational attention for event detection", *5th International Conference on Computer Vision Systems (ICVS 2007)*, Applied Computer Science Group, (2007).



- [155] M. Mancas, B. Gosselin, and B. Macq, "Perceptual image representation", *EURASIP Journal on Image and Video Processing*, Article ID 98181, 1-9 (2007).
- [156] L. Itti, and C. Koch, "Comparison of feature combination strategies for saliency-based visual attention systems", *SPIE-3644*, 473-482.
- [157] S. Frintrop, M. Klodt, and E. Rome, "A real-time visual attention system using integral images".
- [158] K. Rapantzikos, N. Tsapatsoulis, Y. Avrithis *et al.*, "Spatiotemporal saliency for video classification", *Signal Processing: Image Communication*, 24(7), 557-571 (2009).
- [159] H. J. Seo, and P. Milanfar, "Static and space-time visual saliency detection by self-resemblance", *Journal of Vision*, 9(12-15), 1-27 (2009).
- [160] P. Bian, and L. Zhang, "Biological plausibility of spectral domain approach for spatiotemporal visual saliency", *Lecture Notes in Computer Science* 5506, 251-258.
- [161] C. Guo, and L. Zhang, "Spatio-temporal saliency detection using phase spectrum of quaternion Fourier transform", 1-8.
- [162] H. Shi, and Y. Yang, "A computational model of visual attention based on saliency maps", *Applied Mathematics and Computation*, 188(2), 1671-1677 (2007).
- [163] D. Gao, V. Mahadevan, and N. Vasconcelos, "On the plausibility of the discriminant center-surround hypothesis for visual saliency", *Journal of Vision*, 8(7), 13-1-13-18 (2008).
- [164] V. K. Singh, S. Maji, and A. Mukerjee, "Confidence based updation of motion conspicuity in dynamic scenes", 13-20.
- [165] C. Guo, and L. Zhang, "A novel multiresolution spatiotemporal saliency detection model and its applications in image and video compression", *IEEE Transactions on Image Processing*, 19(1), 185-198 (2010).
- [166] F. Xue, F. Wu, J. Wang *et al.*, "Camouflage texture design based on its camouflage performance evaluation", *Neurocomputing*, 274, 106-114 (2018).
- [167] Z. Dehui, W. Pengye, Z. Hengbo *et al.*, "Camouflage effectiveness assessment based on fusion with constant color background", *Journal of Physics: Conference Series*, 1060, 012009.
- [168] F. Xue, C. Yong, S. Xu *et al.*, "Camouflage performance analysis and evaluation framework based on features fusion", *Multimedia Tools and Applications*, 75(7), 4065-4082 (2015).
- [169] L. Itti, and C. Koch, "A saliency-based search mechanism for overt and covert shifts of visual attention", *Vision Research*, 40(10-12), 1489-1506 (2000).
- [170] Z. Bylinskii, T. Judd, A. Oliva *et al.*, "What do different evaluation metrics tell us about saliency models?", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 41(3), 740-757 (2019).
- [171] T. Avraham, and M. Lindenbaum, "Esaliency (extended saliency): meaningful attention using stochastic image modeling", *IEEE Transactions on Pattern Analysis and Machine Intelligence PAMI*, 32(4), 693-708 (2010).
- [172] A. Garcia-Diaz, X. R. Fdez-Vidal, X. M. Pardo *et al.*, "Saliency based on decorrelation and distinctiveness of local responses", *Lecture Notes in Computer Science* 5702, 261-268.
- [173] A. Garcia-Diaz, X. R. Fdez-Vidal, X. M. Pardo *et al.*, "Decorrelation and distinctiveness provide with human-like saliency", *LNCS* 5807, 343-354.
- [174] M. Mancas, C. Mancas-Thillou, B. Gosselin *et al.*, "A rarity-based visual attention map-application to texture description", *2006 IEEE International Conference on Image Processing*, IEEE Press, 445-448 (2006).
- [175] H. J. Seo, and P. Milanfar, "Nonparametric bottom-up saliency detection by self-resemblance", *IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 1st International Workshop on Visual Scene Understanding (ViSU'09)*, IEEE Press, 45-52 (2009).
- [176] W. Einhäuser, M. Spain, and P. Perona, "Objects predict fixations better than early saliency", *Journal of Vision*, 8(14), 18-1-18-26 (2008).
- [177] R. J. Baddeley, and B. W. Tatler, "High frequency edges (but not contrast) predict where we fixate: A Bayesian system identification analysis", *Vision Research*, 46(18), 2824-2833 (2006).
- [178] A. Aık, S. Onat, F. Schumann *et al.*, "Effects of luminance contrast and its modifications on fixation behavior during free viewing of images from different categories", *Vision Research*,

- 49(12), 1541-1553 (2009).
- [179] R. Carmi, and L. Itti, "Visual causes versus correlates of attentional selection in dynamic scenes", *Vision Research*, 46(26), 4333-4345 (2006).
- [180] W. Einhäuser, U. Rutishauser, and C. Koch, "Task-demands can immediately reverse the effects of sensory-driven saliency in complex visual stimuli", *Journal of Vision*, 8(2), 1-19 (2008).
- [181] T. Foulsham, and G. Underwood, "How does the purpose of inspection influence the potency of visual salience in scene perception?", *Perception*, 36(8), 1123-1138 (2007).
- [182] C. Roggeman, W. Fias, and T. Verguts, "Saliency maps in parietal cortex: Imaging and computational modeling", *Neuroimage*, 52(3), 1005-1014 (2010).
- [183] X. Chen, and G. J. Zelinsky, "Real-world visual search is dominated by top-down guidance", *Vision Research*, 46(24), 4118-4133 (2006).
- [184] M. Nyström, and K. Holmqvist, "Semantic override of low-level features in image viewing - both initially and overall", *Journal of Eye Movement Research*, 2(2), paper #2,1-11 (2008).
- [185] J. A. Stirr, and G. Underwood, "Low-level visual saliency does not predict change detection in natural scenes", *Journal of Vision*, 7(10), paper #3,1-10 (2007).
- [186] G. Underwood, E. Templeman, L. Lamming *et al.*, "Is attention necessary for object identification? Evidence from eye movements during the inspection of real-world scenes", *Consciousness and Cognition*, 17(1), 159-170 (2008).
- [187] L. Itti, C. Gold, and C. Koch, "Visual attention and target detection in cluttered natural scenes", *Optical Engineering*, 40(9), 1784-1793 (2001).
- [188] G. Harding, and M. Bloj, "Real and predicted influence of image manipulations on eye movements during scene recognition", *Journal of Vision*, 10(2, paper no. 8), 1-17 (2010).
- [189] J. M. Henderson, J. R. Brockmole, M. S. Castelano *et al.*, "Visual saliency does not account for eye movements during visual search in real-world scenes", 537-562.
- [190] T. Foulsham, and G. Underwood, "What can saliency models predict about eye movements? Spatial and sequential aspects of fixations during encoding and recognition", *Journal of Vision*, 8(2), 6-17 (2008).
- [191] G. Underwood, T. Foulsham, and K. Humphrey, "Saliency and scan patterns in the inspection of real-world scenes: eye movements during encoding and recognition", *Visual Cognition*, 17(6-7), 812-834 (2009).
- [192] R. Rosenholtz, Y. Li, J. Mansfield *et al.*, "Feature congestion: a measure of display clutter", 761-770.
- [193] R. Rosenholtz, Y. Li, and T. Nakano, "Measuring visual clutter", *Journal of Vision*, 7(2-71), 1-22 (2007).
- [194] M. J. Bravo, and H. Farid, "A scale invariant measure of clutter", *Journal of Vision*, 8(1), 23-1-23-9 (2008).
- [195] C. P. Yu, D. Samaras, and G. J. Zelinsky, "Modeling visual clutter perception using proto-object segmentation", *Journal of Vision*, 14(7), 1-16 (2014).
- [196] M. M. Marin, and H. Leder, "Examining complexity across domains: relating subjective and objective measures of affective environmental scenes, paintings and music", *PLOS ONE*, 8(8), e72412 (2013).
- [197] D. Donderi, "An information theory analysis of visual complexity and dissimilarity", *Perception*, 35(6), 823-835 (2006).
- [198] G. Aviram, and S. R. Rotman, "Evaluating TNO human target detection experimental results agreement with various image metrics", *RTO-MP-45 AC/323(SCI)TP/19 - 7*, 1-6.
- [199] G. Aviram, and S. Rotman, "Modeling human false target detection decision behavior in infrared images, using a statistical texture image metric", *IEEE Press*, 393-397 (2000).
- [200] S. R. Rotman, D. Hsu, A. Cohen *et al.*, "Textural metrics for clutter affecting human target acquisition", *Infrared Physics & Technology*, 37(6), 667-674 (1996).
- [201] G. Aviram, and S. R. Rotman, "Evaluating human detection performance of targets and false alarms, using a statistical texture image metric", *Optical Engineering*, 39(8), 2285-2295 (2000).
- [202] J. F. Cartier, and D. H. Hsu, "Human visual search: a two state process", *Infrared Imaging*

- Systems: Design, Analysis, Modeling, and Testing VI, SPIE-2470, 58-68.
- [203] S. Nyberg, and L. Bohman, "Assessing camouflage using textural features", Targets and Backgrounds VII: Characterization and Representation, SPIE-4370, The International Society for Optical Engineering, Bellingham, WA, 60-71 (2001).
  - [204] J. B. Culpepper, "Texture metric that predicts target detection performance", Optical Engineering, 54(12), 123101-123101 (2015).
  - [205] R. Hecker, "Camaeleon--camouflage assessment by evaluation of local energy, spatial frequency, and orientation", Characterization, Propagation, and Simulation of Sources and Backgrounds II, SPIE-1687, SPIE, (1992).
  - [206] D. Xu, and Z. Shi, "DSIM: A DisSIMilarity-Based Image Clutter Metric for Targeting Performance", IEEE Transactions on Image Processing, 22(10), 4108-4122 (2013).
  - [207] K. More, and S. Borse, "Method of evaluating camouflage of texture based on wssim and nature", IJISSET - International Journal of Innovative Science, Engineering & Technology, 3(8), 19-23 (2016).
  - [208] K. J. Cooke, "The ORACLE Handbook", British Aerospace - Sowerby Research Center, Sowerby, UK (1992).
  - [209] G. Waldman, and J. Wootton, Electro-optical systems performance modeling, Artech House, Boston, MA (1993).
  - [210] G. Waldman, J. Wootton, and G. Hobson, "Visual detection with search: an empirical model", IEEE Transactions on Systems, Man and Cybernetics SMC, 21(3), 596-606 (1991).
  - [211] J. A. D'Agostino, W. Lawson, and D. L. Wilson, "Concepts for search and detection model improvements", Infrared Imaging Systems: Design, Analysis, Modeling, and Testing VII, SPIE-3063, SPIE, Bellingham, WA, 14-22 (1997).
  - [212] G. R. Gerhart, E. L. Bednarz, T. J. Meitzler *et al.*, "Target acquisition methodology for visual and infrared imaging sensors", Optical Engineering, 35(10), 3026-3036 (1996).
  - [213] T. Maurer, D. L. Wilson, and R. G. Driggers, "Search and detection modeling of military imaging systems", Optical Engineering, 52(4), 041108-041108 (2013).
  - [214] A. Toet, P. Bijl, F. L. Kooi *et al.*, "A test of three visual search and detection models", TNO Human Factors, Soesterberg, The Netherlands (1998).
  - [215] NATO-RTO SCI-287 "Assessment methods for camouflage in operational context ", NATO S&T Organization, Neuilly-sur-Seine, France (2019).
  - [216] T. Müller, and M. Müller, "CART V: recent advancements in computer-aided camouflage assessment", Infrared Imaging Systems: Design, Analysis, Modeling, and Testing XXII, SPIE-8014, SPIE, (2011).
  - [217] V. C. Coffey, "Seeing in the dark: defense applications of IR imaging", Optics and Photonics News, April, 26-31 (2011).
  - [218] A. Schwarz, "Adaptive camouflage in the VIS and IR spectral range: main principles and mechanisms", Target and Background Signatures, SPIE-965306, SPIE, Bellingham, WA, 1-10 (2015).
  - [219] K. W. McKee, and D. W. Tack, "Active camouflage for infantry headwear applications", DRDC Toronto CR-2007-023, Humansystems Inc., Guelph, Ontario, Canada (2007).
  - [220] K. E. Andersson, and C. Åkerlind, "A review of materials for spectral design coatings in signature management applications", Optical Materials and Biomaterials in Security and Defence Systems Technology XI, SPIE-92530Y, SPIE, 1-20 (2014).

