PROCEEDINGS OF SPIE

SPIEDigitalLibrary.org/conference-proceedings-of-spie

Evaluation of camouflage pattern performance of textiles by human observers and CAMAELEON

Daniela H. Heinrich, Gorm K. Selj

Daniela H. Heinrich, Gorm K. Selj, "Evaluation of camouflage pattern performance of textiles by human observers and CAMAELEON," Proc. SPIE 10432, Target and Background Signatures III, 1043206 (5 October 2017); doi: 10.1117/12.2277908

SPIE.

Event: SPIE Security + Defence, 2017, Warsaw, Poland

Evaluation of camouflage pattern performance of textiles by human observers and CAMAELEON

Daniela H. Heinrich*[a], Gorm K. Selj[a]
[a]Norwegian Defence Research Establishment, P.O.Box 25, N-2027 Kjeller, Norway.

ABSTRACT

Military textiles with camouflage pattern are an important part of the protection measures for soldiers. Military operational environments differ a lot depending on climate and vegetation. This requires very different camouflage pattern to achieve good protection. To find the best performing pattern for given environments we have in earlier evaluations mainly applied observer trials as evaluation method. In these camouflage evaluation test human observers were asked to search for targets (in natural settings) presented on a high resolution PC screen, and the corresponding detection times were recorded. Another possibility is to base the evaluation on simulations. CAMAELEON is a licensed tool that ranks camouflaged targets by their similarity with local backgrounds. The similarity is estimated through the parameters local contrast, orientation of structures in the pattern and spatial frequency, by mimicking the response and signal processing in the visual cortex of the human eye. Simulations have a number of advantages over observer trials, for example, that they are more flexible, cheaper, and faster. Applying these two methods to the same images of camouflaged targets we found that CAMAELEON simulation results didn't match observer trial results for targets with disruptive patterns. This finding now calls for follow up studies in order to learn more about the advantages and pitfalls of CAMAELEON. During recent observer trials we studied new camouflage patterns and the effect of additional equipment, such as combat vests. In this paper we will present the results from a study comparing evaluation results of human based observer trials and CAMAELEON.

Keywords: Camouflage, CAMAELEON, Human observers, Detection time, Detectability range.

1. INTRODUCTION

Camouflage in military settings is used to conceal an object or soldier. The aim is to prevent the target from detection. To achieve this, the camouflage has to blend in with the surroundings. In the visible part of the electromagnetic spectrum this means the camouflage has to attract minimum attention by mimicking the characteristics of the surroundings. When developing camouflage pattern both colours and patterns have to be optimized [1]. The spatial fitting needed for pattern development is achieved through frequency and orientation matching [2-4] aiming also for the target shape disruption effect [5-10]. Colours are usually determined based on spectral reflectance measurements of natural backgrounds [11-12]. To determine the performance of camouflage pattern in operational settings, which will include a variety of different natural scenes, one can apply different evaluation methods.

The human visual system allows for detection and subsequently identification of objects at a distance, depending on a broad set of parameters. These involve the visual acuity, field of view, light levels and how spatial information is extracted from the visual environment and how it is processed and interpreted [13]. To evaluate camouflage pattern based on the eye's probability of detecting a camouflaged target in natural backgrounds one can use human observers or systems mimicking the human visual system [14-20].

Based on how the human eye's receptive filter banks respond to a target's contrast, spatial information and orientation a software tool – appropriately named CAMAELEON – was developed [2], verified [21,22] and made commercially available during the 1990's, claiming to distinguish between a target and its local background similarly to the human eye. Hence, CAMAELEON is a signature assessment tool valuable in a process of developing, testing and evaluating camouflage.

Software tools have the potential benefit of making a camouflage pattern evaluation much quicker as well as lowering costs significantly. Therefor it is interesting to compare such a tool, here CAMAELEON, with the results of real observer data from human observers. In this paper we present the evaluation of seven targets dressed in suits with different camouflage pattern placed in 15 varying natural scenes.

Target and Background Signatures III, edited by Karin U. Stein, Ric Schleijpen,
Proc. of SPIE Vol. 10432, 1043206 · © 2017 SPIE CCC
code: 0277-786X/17/$18 · doi: 10.1117/12.2277908

2. METHODS

The aim of this study is to evaluate the potential of CAMAELEON to rank the effectiveness of different camouflage patterns in a given position in a natural scene. For the study we use images taken recently and used during a human observer trial to determine the most effective camouflage pattern for Norwegian environments. For the discussion and evaluation of the CAMAELEON results we will address the ranking results from the human observer trial as a reference.

2.1 Search by photo trial with human observers

The human observer trial was conducted 2 years ago and is the basis for the CAMAELEON study. The preparation of that trial and the execution will be described in the following sections.

Image capture in natural backgrounds

The 15 scenes that were used to assess the camouflaged targets in this study were chosen to contain different types of local backgrounds around the target. We recorded the targets, one by one, in as identical conditions as we were able to, considering, target orientation, position, and area exposed, and as stable illumination conditions as possible. The purpose was to ensure that the target's camouflage patterns were evaluated only based on their relative camouflage properties. To achieve that, we carried out a near-continuous (within minutes) image recording of the targets in every scene by using a digital camera (Nikon D5200). Importantly, only one target was recorded per image to avoid confusion about what was actually to be assessed by the human observer during the trial.

Targets were located in different positions in the image frame (never centered), from one scene to the next, in order to avoid observers' expectations on where to start to search [15, 23]. The physical distance to the targets in the field was varied between 9 m and 93 m in the 15 different scenes. We recorded the scene images with the intention that it should be possible, for the observers, to distinguish between the target and the local background, whenever the observer's eye focus was in the proximity of the target position in the image frame. Hence, a detection of a target by a human observer was to reflect camouflage effectiveness and not to be based on observer's making guesses about too far objects. A collage of images of the 15 different scenes is shown in Figure1. The position of the target is marked with a red ring.


102


123


124


125


127


128

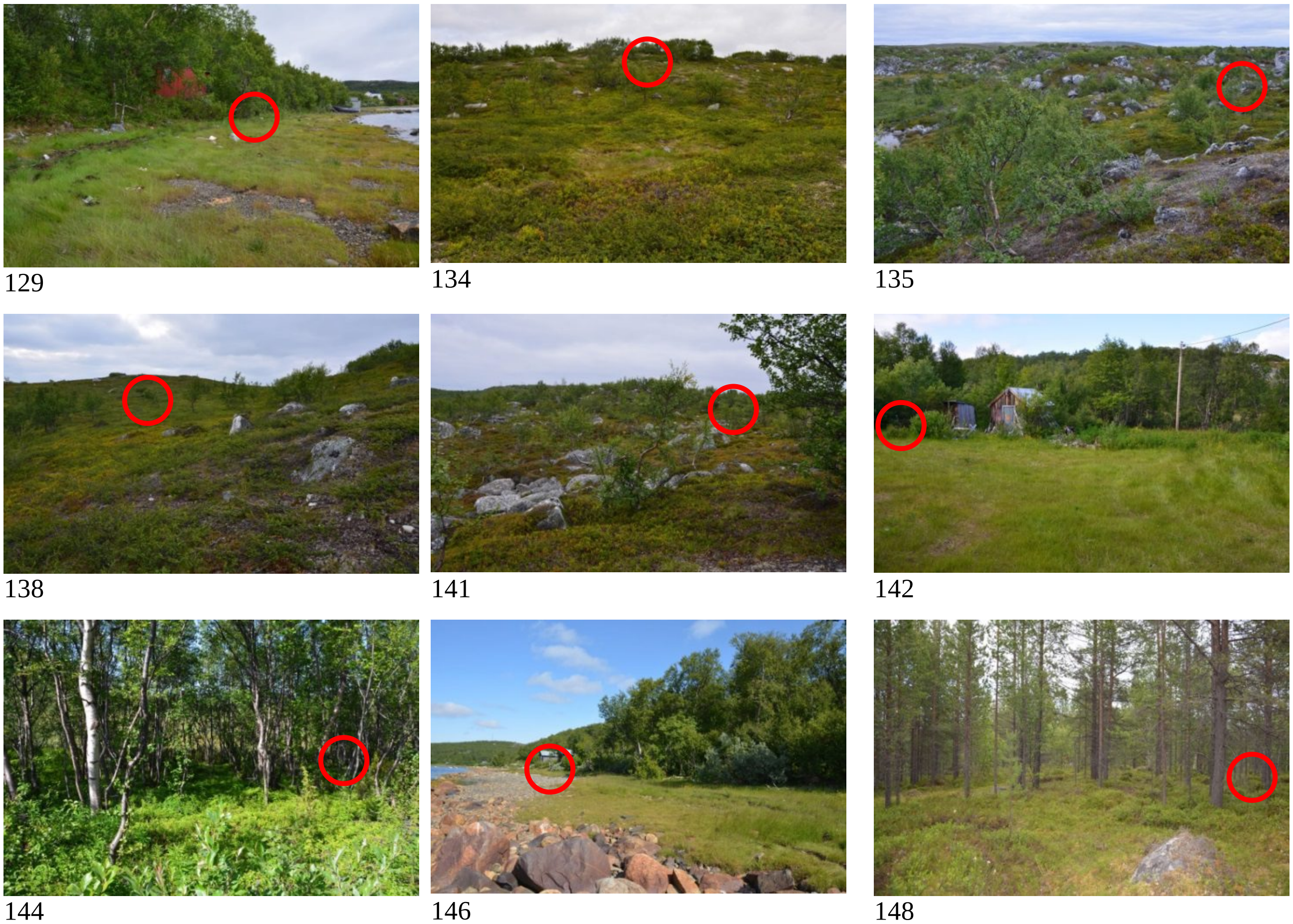


Figure 1 Images (2560 by 1600 pixels) of the 15 different scenes used in this study. The images of the scenes were chosen such that they would result in detection times within the second range for the search by photo camouflage assessment by human observers.

Generating proper test targets

The targets used in this study are shown in Figure 2 the 7 targets consisted of a mixture of targets with our own designed camouflage patterns, roller printed (HolTex, Germany) onto cotton textile (225 g/ m2) sewn into mannequin suits, as well as camouflage patterns that are commercially available. Three of the targets (5-7 in Figure 2) were also dressed with wests and pouches in the same pattern as the underlying suits. A styrofoam mannequin was dressed up with the suits, one at the time, and were then recorded in 15 various natural backgrounds (scenes) in central and northern parts of Norway.


Target 1


Target 2


Target 3


Target 4


Target 5


Target 6


Target 7

Figure 2 The camouflage patterns used in this study. We see that the study included targets dressed in suits with camouflage pattern and targets with additional wests with pouches above the suits. The wests and pouches matched the pattern of the underlying suits.

Evaluation of the camouflage effectiveness by observer trials

We used a purpose made, and human observer based, search by photo methodology ([15]) to evaluate the camouflage effectiveness of the targets. Each target was then evaluated – based on its distribution of time of detection – when subject to visual search by human observers. The test was set up so that each target in each scene was searched for by 15 unique observers (to an accuracy of one single observer).

Preparation of human observers

In order to reduce bias or learning effects during the trial each soldier was given a word by word *identical* introduction to the observer trial by an instructor prior to the trial itself. Observer position was adjusted to have an optimal and identical distance to the widescreen (ca. 40 cm). Also, the observer's eyes were approximately leveling the center of the pc-screen. Each of the observers conducted a test run consisting of two images similar to those in the main trial. During this test run, the observers were allowed to ask questions, reducing the risk of misunderstandings before the main trial started. During main trial, observers were not allowed to ask questions, but left to find targets solely by themselves.

The observer trial conduction

During the trial each observer was shown a randomized sequence of images of the 16 unique scenes. The scene images were shown one by one by means of a high definition (HD, 2560 x 1600 pixels) pc screen in a room that was dimly lit. Each of the images in the sequence was showing a scene with one or no camouflaged (human-shaped) targets in it. The observers were, one at the time, asked to search for a target with a known shape, but unknown camouflage pattern and image frame positioning, and then indicate detection by mouse-clicking on the target as soon as the observer was convinced it was a real target. To sort out detection data from anomalies or miss-detections, we established a minor tolerance surrounding each target for each scene. The dimension of the added tolerance around the target was typically lower than the dimension of the target itself, meaning that detection markings just outside the target outline were accepted as proper. For each search the corresponding time of detection was stored for further analysis. The observers were allowed to be exposed to the same targets in more than one scene throughout the 15 scenes in the trial. The observers were never asked to search for more than one target per scene.

Each scene was presented to the observer for maximum 50 seconds, giving the observer reasonable search time. Whenever the search time limit of 50 seconds was exceeded, the target (in that particular scene) was stored as a “non-detection”. All detection times as well as all non-detections were stored for further analysis. Finally, we used a purpose-made software tool to carry out the observer trial, showing the 15 scenes in a randomized order to each observer so that each target was assigned the same number of observer searches in each scene.

Finding the overall rank of the targets in all 15 scenes

To be able to rank the targets (over all 15 scenes), preserving their relative camouflage effectiveness, we followed the same approach as outlined in similar studies [14, 15]:

- A normalization of the median detection time for each target in each scene, representing the performance of the target relative to the other targets in a particular scene. A numeric value above 1.0 then reflected camouflage effectiveness above average, whereas a value less than 1.0 reflected the opposite. Such an approach also accounts for the relative difference (and not just their order) of the seven targets.
- An assigned weight (higher, equal to or less than one) for each scene. In this study each of the scenes were given equal weight as they were assumed equally relevant.

2.2 CAMAELEON simulations

Based on the already existing images we utilized CAMAELEON software (licensed by IABG, Germany) to learn more about the possibility of ranking the effectiveness of the 7 camouflage pattern in all of the 15 scenes in a more effective and more cost effective way [2]. CAMAELEON allows for estimates of detectability ranges (given in meters) of a target in a natural scene. Detectability is the ability to distinguish between object and background for some particular sensor, in this case the human eye. The detectability depends on size, luminance, contrast, texture, colour, shape primitives and motion.

CAMAELEON demands 2^n pixel square images for the calculations of target detectability and we chose 256 by 256 close-up images of all of the scenes (Figure 3). In each scene we marked the target by drawing a free-hand line along the outline of the target excluding areas which were covered by vegetation. Thereafter, a background had to be chosen, to which the target was compared regarding statistical overlap of local contrast, spatial frequency and local orientation [2]. Figure 3 shows the markings of the targets (in red) and the chosen backgrounds (in green). To compare the effectiveness of the camouflage patterns for all targets in one scene it was crucial to mark the target correctly and to ensure that the background elements were identical.

Selection of background elements

In our study we compared both the performance of all targets in one scene with each other and the overall performance of one target in all 15 scenes with the overall performance of the other 6 targets. We calculated detectability ranges for different background elements.

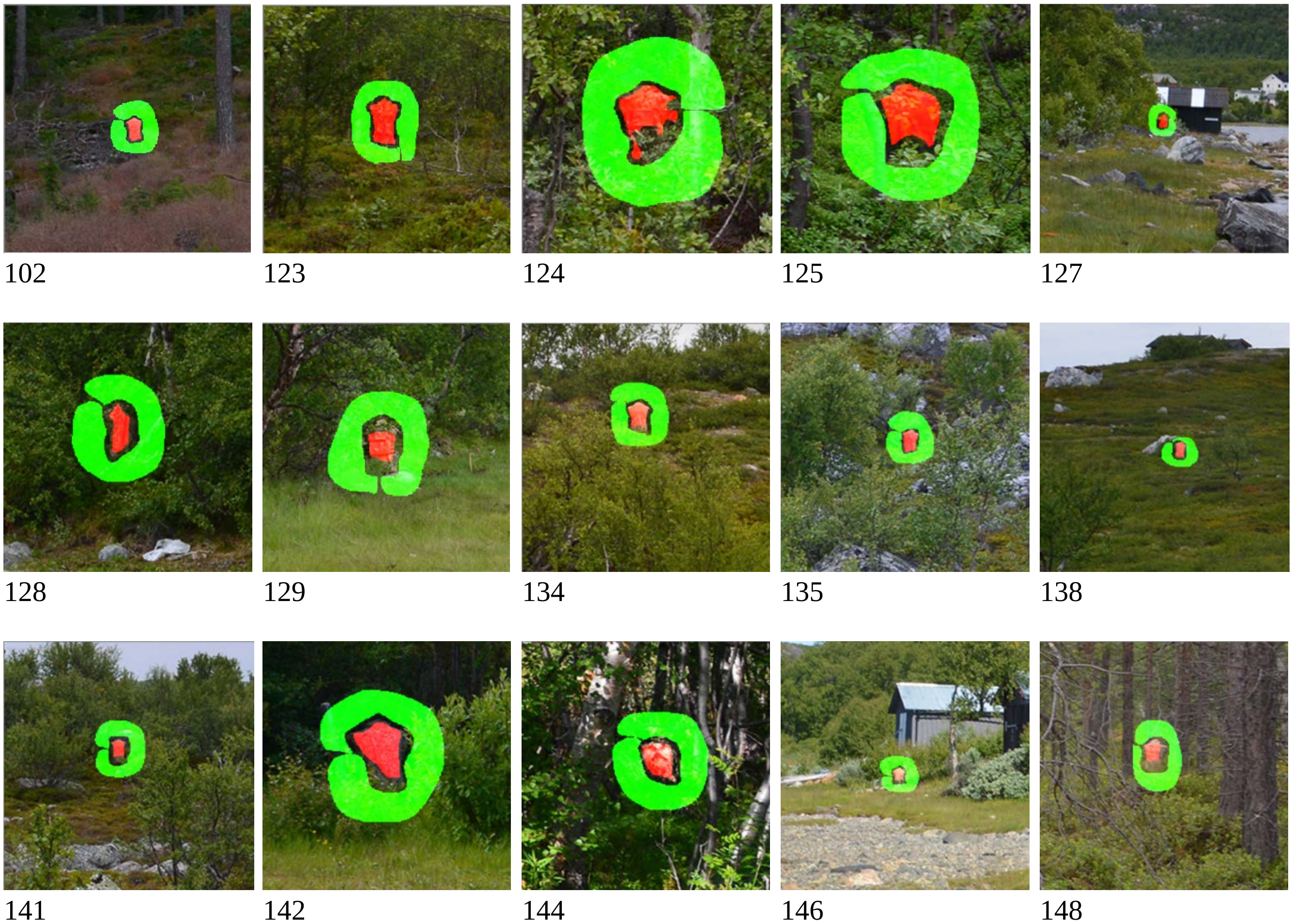


Figure 3. Close-up images (256 by 256 pixels) of the 15 different scenes used for the CAMAELEON study. The images show how the targets were marked (red area). The green halos around the targets mark the local backgrounds, to which the targets were to be compared, in order to find an estimate on the targets' individual detectability ranges. The corresponding scenes used in the search by photo camouflage assessment by human observers are shown in Figure 1.

Finding the overall rank of the targets in all 15 scenes

To compare the results from the search by photo camouflage assessment by human observers study to the CAMAELEON study results we need to rank the targets in a similar way [14, 15]:

- Normalization of detectability for each target with the average detectability range for all targets in each scene, representing the performance of the target relative to the other targets in a particular scene. A numeric value above 1.0 then reflected detectability ranges above average which corresponds to camouflage effectiveness below average. Values less than 1.0 reflected performance above average. Such an approach also accounts for the relative difference (and not just their order) of the seven targets.
- An assigned weight (higher, equal to or less than one) for each scene. In this study each of the scenes were given equal weight as they were assumed equally relevant.

3. RESULTS

The 7 camouflage pattern targets were evaluated in 15 natural backgrounds by utilizing both CAMAELEON detectability range calculations and based on human observers in a search by photo observer trial.

3.1 CAMAELEON results

For each scene we calculated the detectability ranges for all 7 targets. We found that the detectability ranges for the different scenes varied from about 100 to 400 meters amongst the targets. To make it possible to compare the performance of the different targets within one scene we normalized the detectability ranges. A shorter detectability range means the target is harder to distinguish and better adapted to the background. When normalizing the individual detectability ranges to the average detectability range for the scene targets that were performing above average will end up with a number less than 1.0. Targets which are easier to distinguish from the local background than average will be assigned a value larger than 1.0.

In our first calculations we used a background area surrounding the target (Figure 3). The background was based on the assumption that the adjacent background surrounding the target was most important for the detectability. The width of the background area was about the same as the width of the mannequin shoulders. This approach was applied to all 15 scenes. The corresponding normalized detectability ranges are given in Table 1.

To make the table easier to read a colour map is used. Green means that the target mimicked the surrounding background better than the average performance of all targets in the scene. Yellow is assigned to an approximate average performance, while a performance below average is coloured in nuances of red. To get the overall performance of each target over all scenes we calculated the sum of all normalized detectability ranges. This sum is given for each target in the bottom row of Table 1.

Table 1 Normalized average detectability ranges for all 7 targets in all 15 scenes with a specified background area surrounding the target.

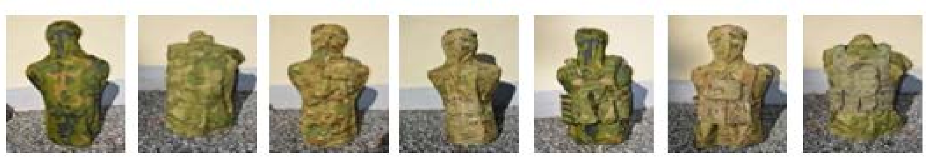
Scene/Target	T1	T2	T3	T4	T5	T6	T7
102	0,89	0,96	1,04	1,25	0,87	1,05	0,94
123	0,80	0,87	1,11	1,24	0,84	1,19	0,94
124	1,35	0,95	0,83	0,94	1,18	0,81	0,94
125	1,25	0,82	0,84	1,18	1,26	0,91	0,74
127	1,13	1,00	0,98	0,79	1,19	0,85	1,06
128	1,09	0,79	0,90	1,30	1,15	0,95	0,81
129	1,25	0,88	0,75	1,04	1,24	0,84	1,01
134	0,91	1,05	1,08	1,20	0,69	1,10	0,97
135	1,29	0,94	0,84	0,88	1,26	0,84	0,95
138	0,99	0,88	1,04	1,19	1,03	1,00	0,86
141	1,30	0,85	0,80	1,06	1,24	0,88	0,87
142	1,30	0,97	0,69	0,98	1,21	0,87	0,98
144	0,75	1,06	1,06	1,20	0,74	1,15	1,04
146	0,89	1,04	1,08	1,18	0,71	1,07	1,02
148	0,89	1,00	1,01	1,16	0,81	1,12	1,00
Sum	16,09	14,06	14,04	16,61	15,44	14,64	14,12

3.2 Observer trial results

The results obtained from the study with human observers was detection times for each target in each scene, as also used as a measure of camouflage effectiveness in similar studies [10, 14, 15]. The normalized median values for the detection times are given in Table 2,. In case an observer didn't detect the target at all the median detection time was set to 50 s (i.e. the search time limit). Targets that obtained relative values above 1.0 performed above average in that particular scene and those targets are given in green in Table 2. Targets that obtained values less than 1.0 performed below average (by more rapid observer detections) and are given in red in the table. The color map is used to facilitate a rapid overview of the target performances, whereas the vital detail on the differences in the camouflage effectiveness amongst the targets is given by the numeric values. In the lowermost row of Table 2 the overall performance of each target is given over all 15 scenes, given as the sum af all the normalized medians from the individual scenes.

Table 2 Normalized median detection times based on observer trial for each target in each scene. Values above 1.0 (green) indicate performance above average in that scene, whereas values less than 1.0 (red) indicate the opposite.

Scene/Target	T1	T2	T3	T4	T5	T6	T7
102	2,24	0,55	0,15	0,16	2,24	0,28	1,37
123	1,15	0,62	0,63	0,51	2,93	0,46	0,69
124	0,87	1,00	0,81	1,01	1,33	0,88	1,11
125	0,82	0,53	0,57	0,58	3,43	0,55	0,51
127	0,56	1,58	1,58	0,55	0,61	0,55	1,58
128	0,90	1,65	0,62	0,75	1,65	0,46	0,99
129	0,83	1,26	0,81	0,89	1,03	1,09	1,09
134	3,32	0,29	0,15	0,17	2,62	0,18	0,27
135	0,62	1,06	1,06	1,06	1,06	1,06	1,06
138	1,37	1,37	0,19	0,18	1,37	1,16	1,37
141	1,76	1,05	0,27	0,23	1,76	0,17	1,76
142	1,73	0,69	0,93	0,97	0,95	0,89	0,85
144	1,20	0,82	0,45	0,92	1,59	0,36	1,67
146	1,40	1,40	0,80	0,31	1,40	0,52	1,16
148	1,07	1,18	0,96	0,87	1,18	0,77	0,99
Sum	19,82	15,03	9,98	9,18	25,14	9,38	16,46



4. DISCUSSION

In this paper we have studied how signature assessments induced by a simulation tool (CAMAELEON), mimicking the early stages in the human visual sensory system [2], ranks detectability ranges for seven camouflaged targets in 15 different natural scenes. The study is based on images which have also been used during a trial assessing the performance of the targets by human observers. The human observer trial resulted in a ranking of the performance for the targets in each scene and an overall ranking for the targets in all 15 scenes. Using CAMAELEON we achieved a similar kind of rankings. In the following we are going to discuss the CAMAELEON results and how they correlate with human observers' assessment.

4.1 CAMAELEON results versus human observers

Background around the target

A comparison of the results for the two different methods to study the performance of the seven targets in the 15 different scenes requires that the ranking is based on the same conditions. In our first assumption we assumed the local background around the target to be decisive for the rating for both assessment techniques. The results from the CAMAELEON study are given in Table 1 and for the corresponding human based assessment results are shown in Table 2. CAMAELEON assessed the performance of the targets very different compared to the human observers. Only in four of the 15 scenes the ranking of the targets was similar. These scenes were 102, 123, 134 and 148. In the other scenes the ranking for some targets turned out be even the opposite for the two methods. For example in scene141 target 1 obtained the largest detectability range, e.g. it was distinguishable from the background at larger distances than the other targets. Target 1 was therefore the target that was ranked as the poorest by CAMAELEON in that scene. From the human observer trial, however, a detailed inspection (not given in this paper) of the results in scene 141 showed that target 1 was not detected at all by as many as 9 of the 15 observers, resulting in a high score in Table 2, and was one of the best performing targets, along with target 5 and 7.

Based on the very different rankings of the targets in most of the scenes it is not surprising that the corresponding two overall rankings turned out different. The ranking based on CAMAELEON resulted in the following order from better to worse: target 7, 2, 3, 6, 5, 1 and 4. For the human observer trial the corresponding ranking was: target 5, 1, 7, 2, 3, 6 and 4. Only the poorest performing target was the same in both methods.

What is the reason for this discrepancy? Looking at the different scenes (Figure 1) we see that the targets were placed in different types of surroundings, such as forest, coastline or mountain areas, different local backgrounds like trees, rocks and heather and at varying distances. Also, the scene images were recorded in different weather conditions, from one scene to the next, and by that in different light conditions during different seasons. We examined our ranking results connected to these parameters. We could not find any of these parameters likely to be the cause that in some scenes we found a high correlation between the rankings of the two assessment methods, whereas in other scenes the correlation was low. .

Since the simple approach of choosing a full circular element around the target as the background segment for the CAMAELEON analyzes did not lead to any obvious conclusion, we wanted to investigate further if our choice of background would influence the rankings and consequently also the correlation between the two methodologies. .

Selected background elements

The mentioned conflicting rankings of the two methods for the targets, among others in scene 141, led us to calculate the detectability ranges based on a different part of the background in the proximity of the target. We excluded the rock above the target. This part of the background was found to be very different from the camouflage pattern in colour, orientation and special frequency. None of the camouflage patterns were intended to mimic large rocks. We chose 4 more scenes to exploit different choices of the background. In scene 124 we excluded for the same reason the tree from the background selection (Figure 4). We also recalculated detectability ranges for scene 102 and 142. For scene 102 we chose the green segment of the background close to the target (Figure 4) and for scene 142 the vegetation behind the target in the shadow (Figure 5). We expect these backgrounds to include more homogenous background segments with the same type of vegetation and the same light conditions.

102 124 127 141

Figure 4 Images showing the markings of the targets (red) and the re-selected, and more exclusive, backgrounds (green) for scene 102, 124, 127 and 141.

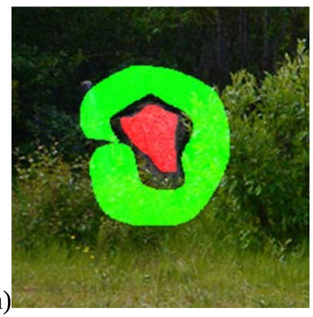
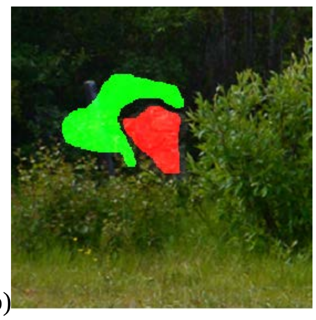
a)  b) 

Figure 5 Image a) shows the marking of the target (red) and the surrounding background (green) for scene 142. Image b) shows another choice of background while the target marking remained unchanged.

The resulting average detectability ranges, based on our re-selected background areas, are given in Table 3. The results show that in scene 102 and 127 the re-selected backgrounds only led to slight differences in the relative detectability ranges and did not change the ranking of the targets. However, in scene 142 the ranking changed dramatically which resulted in a higher correlation with the rankings from the human observer trial results. Anyhow, still none of the rankings based on CAMAELEON matched the rankings based on the human observers. In some scenes it seems obvious, and non-debatable, what parts of the background the target should blend in with, such as in scene 142. In other scenes the selection of background areas might be based on educated guesses or personal preferences. This is not a preferable approach. One of the scenes where the background selection process was not obvious is scene 141.

Table 3 Relative detectability ranges obtained by CAMAELEON for two different backgrounds, circular background around the target and a more selective background, compared to the human observer trial ranking results.

Scene/Target	T1	T2	T3	T4	T5	T6	T7
102	0,89	0,96	1,04	1,25	0,87	1,05	0,94
102 selected background	0,77	1,01	1,09	1,22	0,77	1,09	1,05
102 human observers	2,24	0,55	0,15	0,16	2,24	0,28	1,37
124	1,35	0,95	0,83	0,94	1,18	0,81	0,94
124 selected background	1,27	0,85	0,87	1,13	1,06	1,01	0,81
124 human observers	0,87	1,00	0,81	1,01	1,33	0,88	1,11
127	1,13	1,00	0,98	0,79	1,19	0,85	1,06
127 selected background	1,09	0,98	0,96	0,92	1,18	0,86	1,01
127 human observers	0,56	1,58	1,58	0,55	0,61	0,55	1,58
141	1,30	0,85	0,80	1,06	1,24	0,88	0,87
141 selected background	1,23	0,76	0,92	1,14	1,16	0,97	0,82
141 human observers	1,76	1,05	0,27	0,23	1,76	0,17	1,76
142	1,30	0,97	0,69	0,98	1,21	0,87	0,98
142 selected background	0,78	1,31	0,99	1,20	0,64	1,10	0,97
142 human observers	1,73	0,69	0,93	0,97	0,95	0,89	0,85

Inspecting scene 141 once more we decided to refine the background selection to only include the more homogenous background parts on the left side of the target (Figure 6, Background-2). The CAMAELEON results are given in Table 4 where they are compared to the human observer trial results.. With this background selection the rankings change dramatically and match the human observer trial results in a much larger extent.

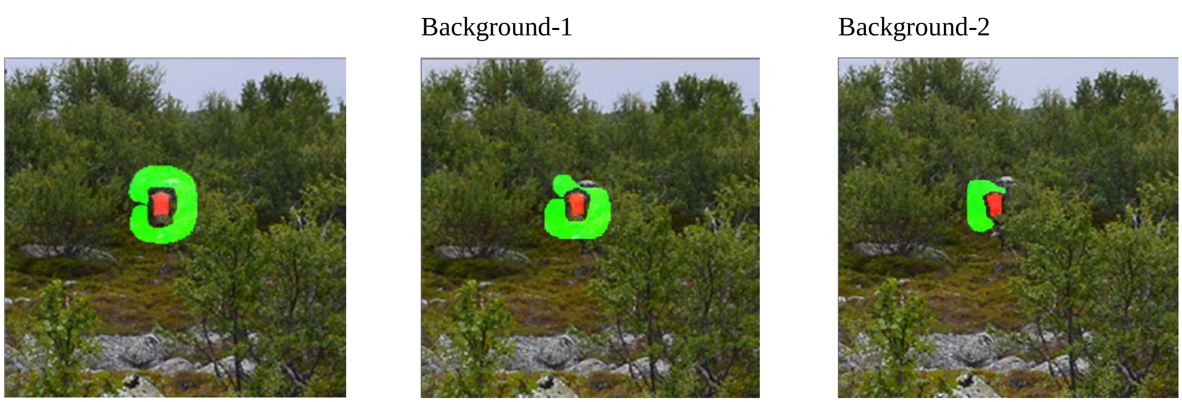


Figure 6 Illustration of the different background selections for scene 141.

Table 4 Relative detectability ranges for all targets in scene 141 for different background element according to Figure 6.

Scene/Target	T1	T2	T3	T4	T5	T6	T7
141	1,30	0,85	0,80	1,06	1,24	0,88	0,87
141 selected background -1	1,23	0,76	0,92	1,14	1,16	0,97	0,82
141 selected background-2	0,92	0,97	1,08	1,20	0,82	1,08	0,93
141 human observers	1,76	1,05	0,27	0,23	1,76	0,17	1,76

Summarizing, our first assumption was that a circular background element of the same dimension as the targets dimensions surrounding the target would be sufficient to receive detectability ranges in correspondence with the ranking based on human observer. Comparing the results of both methods showed that such an approach does not always work, meaning that first order results based on CAMAELEON can be in direct conflict with observer based results. Choosing more selective elements of the background in the proximity of the target resulted in results with a higher degree of correlation. This interesting finding leads to the question on what the selection of the background should be based on?

What influences the conspicuity of a target for a human observer?

Using the CAMAELEON software the first step in the analysis is the calculation of local orientation, local energy and local spatial frequency for the target as well as the selected background area. The results are presented in terms of colour coded images. An important question in this case is how those images together with the images of the scenes can generate a fundament for the background selection and a realistic ranking of the targets

The relative detectability ranges given in Table 3 for scene 102 did not vary much between the background chosen around the whole target and the segment with only green bushes, respectively. In this scene the target was surrounded by green small bushes, reddish heather, and gray-brown branches. To study the effect of different selections of background areas on the corresponding detectability range further, we calculated the relative detectability ranges for different background selections (Figure 8). To support our selection of backgrounds we used in addition the colour coded images for local orientation, local energy and local spatial frequency (

Figure 7). Inspecting these images carefully, we see that the shoulders of the target 6 differed from the adjacent local background in all three aspects. Comparing the colour coded images for target 6 and target 1, it seems that in this scene the target was distinguishable solely based on the differences between the shoulder section and the adjacent background. Therefore we enlarged the target marking to secure that the shoulder section was a part of the target (Figure 8 d)).

The result for the different background and target markings in scene 102 is given in Table 5. Even if the ranking did not change dramatically within the different alternatives of local background we see that the ranking based on the marking of the target larger than the target itself matched the ranking received by human observers to a higher degree. Only target 2 and target 7 obtained opposite ranking between the two methods. The CAMAELEON relative detectability ranges for target 2 and target 7 were 1.05 and 1.06, respectively. Thus, their performance in the given background and target marking was very similar. Comparing this ranking for target 2 and target 7 for the other three markings given in Table 5 we see that in two cases (i.e. selections of local background) target 2 performed slightly better than target 7 and in two other cases we found the opposite. We interpret the rankings as that these two targets performed similar and their individual ranking should therefore not be assigned too much weight.

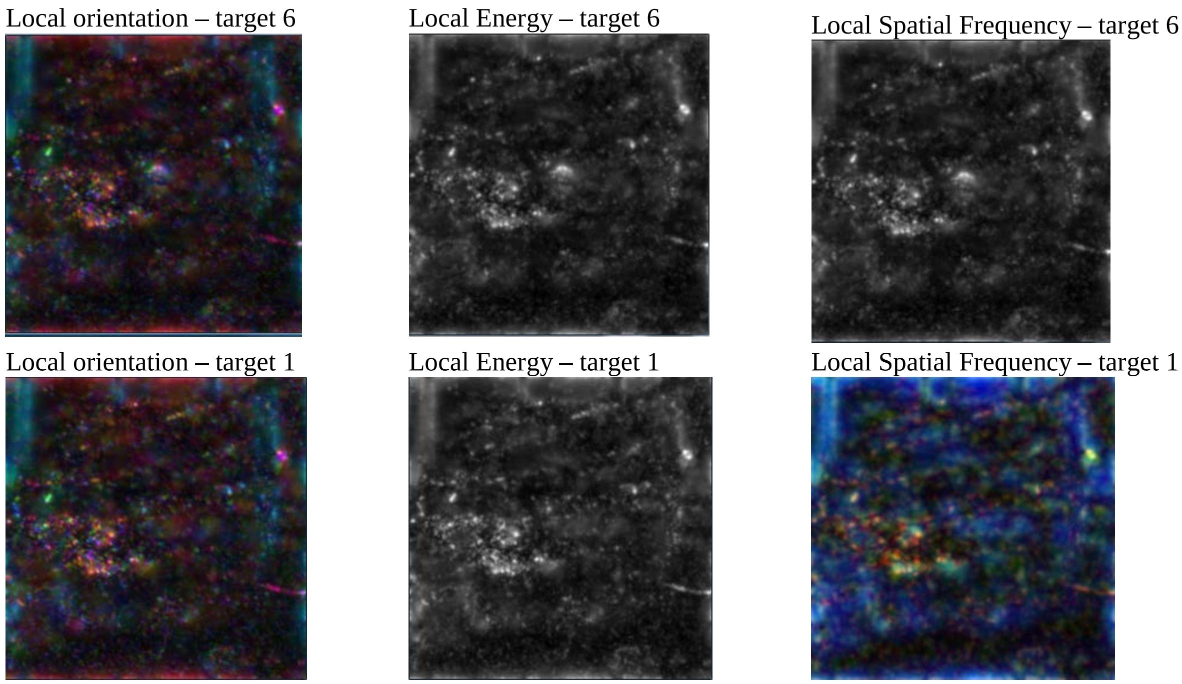


Figure 7 Illustration of the colour coded images with regard to local orientation, local energy and local spatial frequency of scene 102 for two of the targets (target 6 in the upper part, target 1 in the lower part).

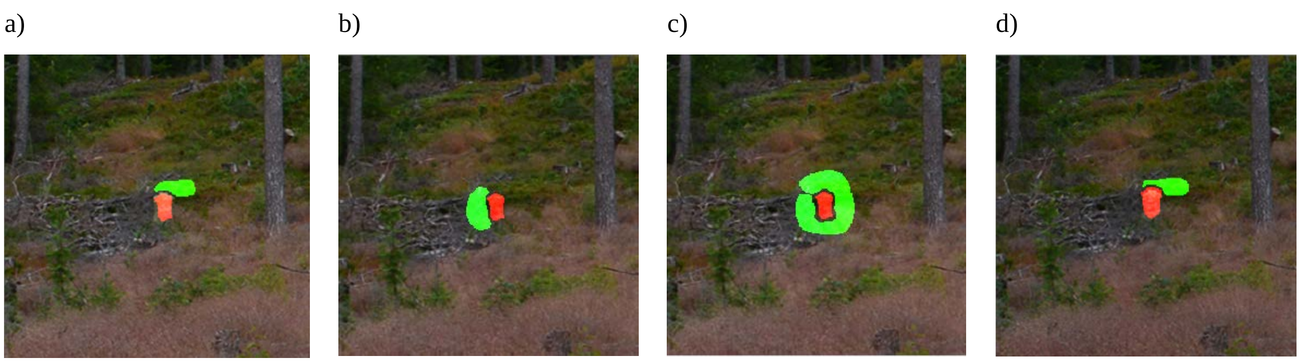


Figure 8 Images showing the markings of the target (red) and the selected more exclusive backgrounds (green) for scene 102. In image a) - c) the target is marked identically and the backgrounds are only green bushes, b) branches, and c) around the whole target, respectively. Image d) shows the same background marking as a) but a target marking larger the target itself.

Table 5 Relative detectability ranges for all targets in scene 102 for different background element according to Figure 8. The results are compared with human observer based results.

Scene/Target	T1	T2	T3	T4	T5	T6	T7
102 a)	0,77	1,01	1,09	1,22	0,77	1,09	1,05
102 b)	0,88	1,00	1,03	1,20	0,85	1,07	0,97
102 c)	0,89	0,96	1,04	1,25	0,87	1,05	0,94
102 d)	0,74	1,05	1,13	1,17	0,75	1,09	1,06
102 human observers	2,24	0,55	0,16	0,15	2,24	0,28	1,37

Target marking including edges/silhouette outline

The ranking results received for scene 102, where marking the target with a selection larger that the target itself, suggests that the choice of background in itself was not the main reason for the degree of correlation between the rankings received based on the two methods. In addition the marking of the target seems to be crucial on the result. By this we mean that the edges of the target have to be a part of the target. We applied this approach also to scene 124. Unfortunately, we did not obtain the expected correlation between observers and CAMAELEON. However, even if single scenes do not result in similar rankings, are we interested in how the approach applied to all scenes changes the results for the overall ranking. The evaluation of camouflage pattern cannot be done with single scenes. The camouflage should perform well in an operational environment spanning over a variety of natural backgrounds. The choice of camouflage pattern would never be based on a single evaluation in one position in one given natural scene. Therefore we recalculated all scenes based on that assumption. The results for all targets in the 15 single scenes and the resulting overall performing are given in Table **6**.

Table 6 Normalized average detectability ranges, given by CAMAELEON, for all 7 targets in all 15 scenes with re-selected background elements in the vicinity of the target and target markings that exceeded the target slightly to ensure that all effects introduced by the silhouette were assigned to the target.

Scene/Target	T1	T2	T3	T4	T5	T6	T7
102	0,74	1,05	1,13	1,17	0,75	1,09	1,06
123	0,80	0,87	1,11	1,24	0,84	1,19	0,94
124	0,94	0,94	1,05	1,22	0,79	1,11	0,95
125	0,88	0,88	1,03	1,19	1,03	1,08	0,91
127	1,09	0,98	0,96	0,92	1,18	0,86	1,01
128	1,12	0,86	0,92	1,28	1,07	0,90	0,85
129	1,12	0,69	0,96	1,22	1,05	1,05	0,91
134	0,70	1,11	1,10	1,21	0,68	1,12	1,08
135	1,26	0,89	0,87	0,93	1,15	0,97	0,92
138	1,10	0,88	0,98	1,18	0,96	1,01	0,89
141	0,92	0,97	1,08	1,20	0,82	1,08	0,93
142	0,78	1,31	0,99	1,20	0,64	1,10	0,97
144	0,75	1,06	1,06	1,20	0,74	1,15	1,04
146	0,84	1,07	1,08	1,15	0,73	1,07	1,06
148	0,89	1,00	1,01	1,16	0,81	1,12	1,00
Sum	13,92	14,57	15,33	17,48	13,25	15,92	14,53

Comparing the ranking for the overall performance given in Table 6 (CAMAELEON) and Table 2 (observer trials) we see that the rankings are the same. The overall detectability ranges for all 7 targets in all 15 scenes with re-selected background elements in the vicinity of the target and target markings that exceeded the target slightly Target 5 performed best, but only slightly better than target 1. The effectiveness of the camouflage pattern was furthermore found to decrease in the following order: 7, 2, 3, 6 and finally target 4. This is the same ranking as for the human observer trial (Section 3.2 and Table 2).

5. CONCLUSION

In this study we have, by a detailed comparison of ranking results for seven camouflage pattern in natural backgrounds using CAMAELEON and human observers, found that using CAMAELEON is not straight forward. A simple approach comparing the targets with the local background around the target with a width comparable to the width of the target did not result in a satisfactory correlation between the two methods. Inspecting the rankings and the images more carefully led to a more selective choice of background elements. Also, the target markings were done more thoughtfully based on the CAMAELEON images for the local orientation, local energy and local spatial frequency in the scenes. Applying this to all scenes we received an equivalent ranking for the targets with CAMAELEON as with the human observers. This finding gives us an additional tool for camouflage assessments. At the same time it is evident that CAMAELEON is not suitable to compare the target in the scene with arbitrary selections of background elements. Anyone interested in how a camouflage pattern on a given target performs in terms of detection by human observers has to do an accurate evaluation of the scene and a sensible selection for the target and background markings in order to obtain realistic results. Consequently, we recommend evaluations of the performance of camouflage patterns in natural backgrounds to be based on human observer trial rankings even if human observer trials require a lot of resources both because of the large number of observers which are necessary and the time it takes to carry out the trial.

REFERENCES

[1] Endler, J. A., "Disruptive and cryptic coloration," Proc. R. Soc. B. 273, 2425-2426 (2006).

[2] Hecker, R., "Camaeleon-Camouflage assessment by evaluation of local energy, spatial frequency and orientation," Proc. SPIE 1687, 342-349 (1992).

[3] Lillesaeter, O., "Complex contrast, a definition for structured targets and backgrounds," J. Opt. Soc. Am. A. 10, 2453-2457 (1993).

[4] Heinrich, D. H. and Selj, G. K., "The effect of contrast in camouflage patterns on detectability by human observers and CAMAELEON," Proc. SPIE 9476, 947604 (2015).

[5] Cuthill, I. C., Stevens, M., Sheppard, J., Maddocks, T., Parraga, C. A. and Troscianko, T. S., "Disruptive coloration and background pattern matching," Nature 434, 72-74 (2005).

[6] Stevens, M. and Merilaita, S., "Defining disruptive coloration and distinguishing its functions," Phil. Trans. R. Soc. B 364, 481-488 (2009).

[7] Selj, G. K., "Disruptive camouflage tricks the human eye: a study of detection times of two near-similar targets in natural backgrounds," Proc SPIE 9653, 96530S (2015).

[8] Webster R., Hassall C., Herdman C. M., Godin J. G. J. and Sherratt T. N., "Disruptive camouflage impairs object recognition," Biol Letters 23, 9(6) (2013).

[9] Troscianko J., Lown A. E., Hughes A. E., Stevens M., "Defeating crypsis: detection and learning of camouflage strategies," PLOS ONE, 8(9) (2013).

[10] Selj, G. K., Heinrich, D., "Disruptive coloration in woodland camouflage: evaluation of camouflage effectiveness due to minor disruptive patches," Proc SPIE 9997, 99970A (2016).

[11] Myrabo, H. K., Lillesaeter, O. and Hoimyr, T., "Portable field spectrometer for reflectance measurements 340-2500nm," Applied Optics. 21, 2844-2858 (1982).

[12] Winkelmann, M. E., "Analysis of exploitable spectral features of target and background materials," Proc. SPIE 9653, 96530L (2016).

[13] Troscianko, T., Benton, Ch. P., Lovell, P. G., Tolhurst, D. J. and Pizlo, Z., "Camouflage and visual perception," Phil. Trans. R. Soc. B 364, 449-461 (2009).

[14] Toet, A. and Hogervorst, M. A., "Urban camouflage assessment through visual search and computational saliency," Opt Eng 52 (2013).

[15] Selj, G. K. and Heinrich, D., "Search by photo methodology for signature properties assessment by human observers," Proc. SPIE 9474, 947411 (2015).

[16] Hecker, R., "Camaeleon-Camouflage assessment by evaluation of local energy, spatial frequency and orientation," Proc. SPIE 1687, 342-349 (1992).
[17] Kilian, J. C. and Hepfinger, L. B., "Computer-based evaluation of camouflage," Proc. SPIE 1687, 359-369 (1992).
[18] Birkemark, C. M., "CAMEVA, a methodology for computerised evaluation of camouflage effectiveness and estimation of target detectability," Proc. SPIE 3699, 229-238 (1999).
[19] Nyberg, S. and Bohman, L., "Assessing camouflage using textural features," Proc. SPIE 4370, 60-71 (2001).
[20] Houlbrook, A. W., Gilmore, M. A., Moorhead, I. R., Filbee, D., Stroud, C., Hutchings, G. and Kirk, A., "Scene simulation for camouflage assessment," Proc. SPIE 4029, 247-255 (2000).
[21] McManamey, J. R., "Validation plan for the German CAMAELEON model," Proc. SPIE 3062, 300-310 (1997).
[22] Hecker, R., "Efficent methods for validation target acquisition models," RTO-MP-45, 9-1 - 9-6 (2000).
[23] Mannan, S. K., Ruddock, K. H. and Wooding, D. S., "The relationship between the locations of spatial features and those of fixations made during visual examination of briefly presented images," Spatial Vision 10(3), 165-188,1996.