FFI-RAPPORT

16/00451

Evaluation methods of signature effectiveness

a first evaluation of camouflage assessments by CAMAELEON compared to human observers

—

Daniela Heinrich
Gorm Krogh Selj

Evaluation methods of signature effectiveness
a first evaluation of camouflage assessments by CAMAELEON compared to human observers

Daniela Heinrich
Gorm Krogh Selj

Norwegian Defence Research Establishment (FFI) 4 July 2016

Keywords

Kamuflasje
Signaturkontroll
Evaluering
Simuleringsverktøy CAMAELEON
Deteksjon

FFI-rapport:
FFI-RAPPORT 16/00451

Prosjektnummer
132302

ISBN
P: ISBN 978-82-464-2682-2
E: ISBN 978-82-464-2683-9

Approved by
Svein Rollvik, *Research Manager*
Jon Skjervold, *Director*

Summary

Signature control and a good understanding of the performance of signature control measures is a very important topic for the Norwegian Armed Forces. In order to achieve this we need methods for test and evaluation of signature measures such as camouflage patterns. In this report we present the first results from a camouflage study in the visible region of the electromagnetic spectrum obtained with a simulation tool. The simulation tool that has been used, CAMAELEON, is commercially available and was developed in Germany over a period of several years. CAMAELEON mimics the human visual system by modelling the eye's probability of detecting a camouflaged target in natural backgrounds.

There are important advantages using a simulation tool compared to other methods for evaluating camouflage. CAMAELEON, in particular, is a very flexible method to quickly achieve quantitative results for many camouflage patterns and their performance in a large number of backgrounds. However, simulation tools usually also have limitations. It is important to understand these limitations and under which conditions they appear. This is an essential requirement if we want to draw the right conclusions for a camouflage evaluation or for recommendations based on simulation results.

In this report, based on our first results, we discuss what kind of studies we believe will lead us to a good understanding of the advantages and disadvantages of CAMAELEON. This will give us a good foundation to decide in which occasions an extended use of this simulation tool will give adequate results in the future.

Sammendrag

Kontroll av egen signatur er svært viktig for Forsvarets operative evne, deriblant mulighetene for å operere skjult. For å sikre dette er det nødvendig med sikre, presise og aller helst raske test- og evalueringsmetoder av blant annet kamuflasjemønstre. I denne rapporten presenteres de første resultatene fra en simuleringsstudie av kamuflasje i den visuelle delen av det elektromagnetiske spekteret. Simuleringsverktøyet som ble benyttet, CAMAELEON, er kommersielt tilgjengelig og ble utviklet i Tyskland over flere år. CAMAELEON evaluerer kamuflerte mål, avbildet i naturlige bakgrunner ved at det modellerer hvordan det menneskelige øyet responderer på forskjeller i lokal kontrast, orientering og romlig frekvens mellom mål og bakgrunn.

Det er, potensielt sett, viktige fordeler med å kunne benytte et simuleringsverktøy for å evaluere kamuflasje sammenliknet med andre tilnærmingsmåter. Verktøyet som ble benyttet i denne studien, CAMAELEON, gjør det mulig å raskt gjennomføre en kvantitativ evaluering av kamuflasjeevnen til mål i et stort antall ulike typer naturlige bakgrunner. De fleste simuleringsverktøy har likevel begrensinger, og en av hensiktene med dette arbeidet er å forsøke å avdekke slike begrensninger samt under hvilke betingelser de gjør seg gjeldende. En slik kunnskap er viktig i de tilfeller simuleringer danner utgangspunktet for en evaluering eller anbefaling til Forsvaret.

I denne rapporten, basert på en første kamuflasjestudie ved bruk av CAMAELEON, vil vi diskutere hva slags framtidige studier som bør gjennomføres for å gi tilstrekkelig kjennskap til styrker og svakheter ved dette simuleringsverktøyet. Dette vil være et godt grunnlag for å ta i bruk verktøyet i framtiden i de tilfellene der dette passer og gir tilstrekkelig nøyaktighet i evalueringene.

Content

Preface

The primary purpose of this report is to present results from a first study of the CAMAELEON (**CAM**ouflage **A**ssessment **E**valuation of **L**ocal Energy, Spatial Frequency and **O**rientatio**N**) software as a potential, future concealment assessment tool. CAMAELEON was purchased by FFI August 2015. It is a software dedicated for camouflage evaluation of targets of some kind when placed out and recorded in natural backgrounds. It therefore has the potential of expanding and improving FFI's work on evaluation methodologies within signature management. The underlying principles of CAMAELEON are described in more detail in section 2.1 in this report.

The aim with this report is therefore two-fold: First of all, carry out a first comparison of camouflage evaluations based on CAMAELEON with the corresponding evaluations retrieved from human observers (i.e the human eye). The second aim is to present and discuss future, follow-up studies and possibilities with CAMAELEON. Those ideas involve both a more thorough study as well as a test of the suitability of CAMAELEON as an independent evaluation tool. By this we mean in what kind of operative settings that CAMAELEON seems to work well and reliably, and of equal importance, in what cases it does not.

The new ideas for signature evaluation that we suggest at the end of this report have the potential of assessing new signature solutions more rapidly and cheaper compared to what is possible today.

1 Introduction

In the recent years FFI has put some effort in developing methodologies for signature effectiveness evaluation, primarily in the visual band of the electromagnetic spectrum. This has led to the realization of the observer trial methodology [1] based on a similar approach at the DSTL, UK some years earlier [2]. Although the methodology seems to work very well, yielding trustworthy results of (high) operative relevance it is still very interesting to expand the evaluation methodology by including new and independent evaluation tools, such as simulation packages [3]. The primary goal is then to be more certain on our final recommendations (as it rests upon more than one test) as well as potentially performing evaluations more quickly whenever simulation tools – standalone – are trustworthy enough.

Understanding the nature and function of the human visual system in complex natural scenery is inextricably intertwined with progress within the fields of camouflage and concealment. The visual system allows for detection and subsequently identification of objects at a distance, depending on a broad set of parameters. This generally involves acuity of the system, the system's field of view, what types of light levels it operates under, the way spectral information is being extracted from the visual environment and how spatial information is being processed and then interpreted [4].

Several models have been reported, mimicking the human visual system in different ways, often with the aim of modelling the eye's probability of detecting a camouflaged target in natural backgrounds [3;5-7]. An example of such a model is CAMAELEON. It is a software based on how the human eye's receptive filter banks respond to a target's contrast, spatial information and orientation and claims to distinguish between a target and its local background similarly to the human eye. CAMAELEON was developed [3], verified [8;9] and made commercially available during the 1990's. Hence, CAMAELEON is a signature assessment tool valuable in a process of developing, testing and evaluating camouflage.

Although being verified and found to predict human performance in search processes for camouflaged targets in overall [8], it is still very interesting to test the computational modelling of human vision against real observer data from human observers. There is not very much reported in the literature on our visual function in natural environments [4]. Hence it will be of interest to try to find sub-areas of human search performance, founded in visual search tasks in natural environments that are thought to be difficult to model, and then compare it with one or more computational human vision models.

Based on our results from an extensive trial, where human observers were searching for camouflaged targets in natural scenery, we were able to rank targets in a relative manner based on their performance in the different natural sceneries. In this study we want to compare some of the results we obtained from our observer trial [10] with the corresponding (and new) results from a CAMAELEON-analysis carried out at FFI [11]. CAMAELEON has earlier been used in a limited manner in the recent camouflage evaluation study [10], and we now add on with more results compared to what was initiated back then.

We report on one particular anomaly discovered as an anomaly in the results from a recent observer trial. Two very similar camouflage targets performed surprisingly different. This seems to be induced by a contrast patch located at the head outline [12] of one of these targets. We have investigated this anomaly in more detail by evaluating the same two targets in the same natural scenery also by CAMAELEON, to study further if CAMAELEON picks up the same anomaly effect, and finally to compare results from human observers and CAMAELEON. We give possible explanations to the deviations that we observe, when comparing human vision and a computational model of human vision.

In this report we present a first evaluation of CAMAELEON, assessing its reliability as an independent (and simplified) test of signature effectiveness. We compare it with our primary evaluation methodology (the observer trial [1]) and suggest future possibilities on how CAMAELEON can be used to evaluate signature effectiveness in operational areas *without* having to travel to the specific site performing image capture of targets in a selected natural scenery as in FFI's observer trial methodology [1]. Hence CAMAELEON can, if successful, be a methodology reducing costs of camouflage evaluation significantly.

2 Methods used

In this section we describe how we carried out the analysis of the targets to be evaluated, both by the observer trial methodology and also by the CAMAELEON software tool. We picked out two targets from our recent camouflage study [10] and the performance of these two targets in 6 selected scenes is the basis of our comparison.

The effectiveness of camouflage depends on the similarity between the camouflaged object and the surrounding background. The human response in the process of target recognition is categorized in four levels: detectability, detection, recognition and identification. In the perception process detection means that a certain object has been distinguished from the background, it has been detected. In the case that the object class can be classified the response is specified as recognition and if the object type can be classified as identification. Detectability describes the lowest level of response. It means that it is possible to distinguish between an object and its nearby background. Thus, the level of recognition in the perception process influenced by camouflage is the detectability.

2.1 A short introduction to CAMAELEON

Before we go on with the comparison of CAMAELEON with how human observers perform, we present a short description of the CAMAELEON software tool package to the benefit of the reader.

2.1.1 How CAMAELEON evaluates camouflaged targets

The CAMAELEON software for camouflage assessment was developed by Dr. Richard Hecker, IABG, and uses digital image processing based on the human visual system.

Detectability as described above depends on size, luminance, contrast, texture, colour, shape primitives, and motion [9].

CAMAELEON aims at measuring the physiological detectability of an object against the nearby background by describing the similarity between object and background related to contrast (local energy) and textural features (local special frequency and local edge orientation), as these seem to be the most important features to determine detectability [3]. This means CAMAELEON assesses the characteristics for contrasts, and patterns (size, regularity and orientation) for an object and background.

Based on the fact that the human detection of objects in visual scenes occurs in the early stages of visual information processing, CAMAELEON uses a bandpass-filter bank which simulates the response of the neurons in the early stage of the visual cortex. An image containing a camouflaged object is filtered by these filters and the local features for the target and background are calculated from the output. The result is presented in two ways.

Colour coding

The first presentation form is by colour coding the image (Figure 2.1). These images provide a very quick impression of the similarity.

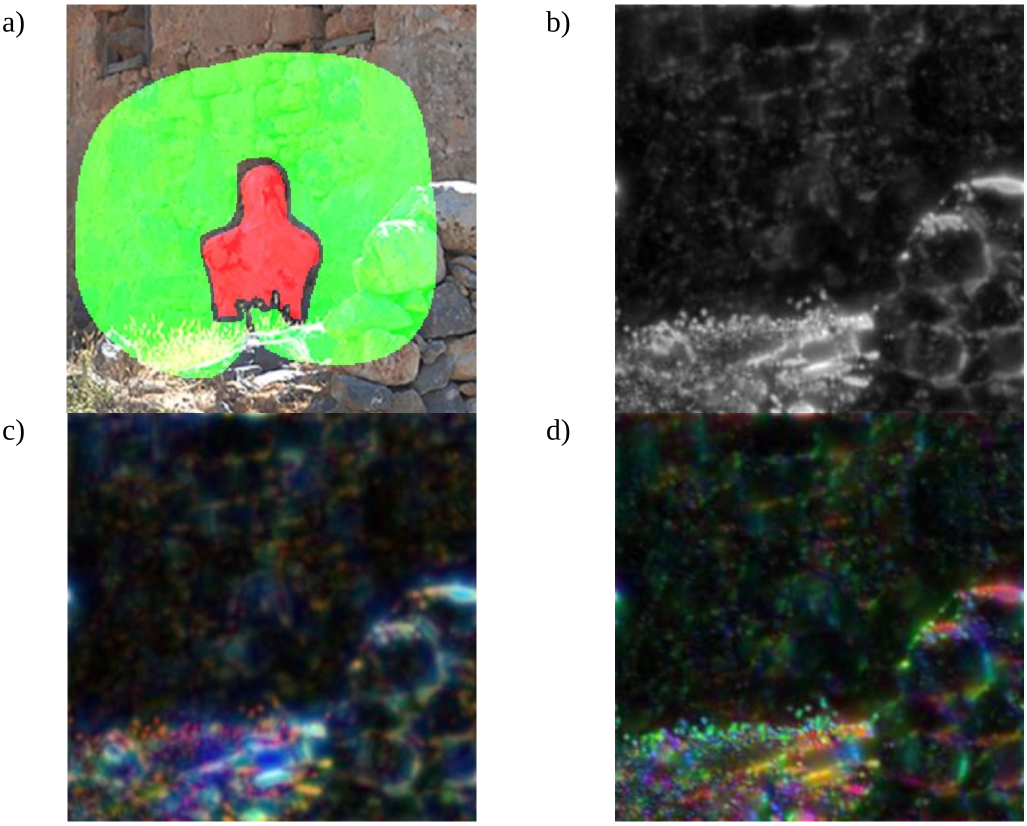


Figure 2.1 Illustration of the color coding of an image of our camouflage target (a) with regard to local energy (b), local special frequency (c) and local orientation (d).

Similarity measures

The second form of presenting the results is to obtain the measures for similarity by calculating the histograms of the local features for both the target and the background and their overlaps (Figure 2.2). In other words after the filtering the distribution for the local energy, spatial frequency and local edge orientations for target (red) and background (green) are plotted in the same graph and can be compared by investigating the overlaps (yellow).

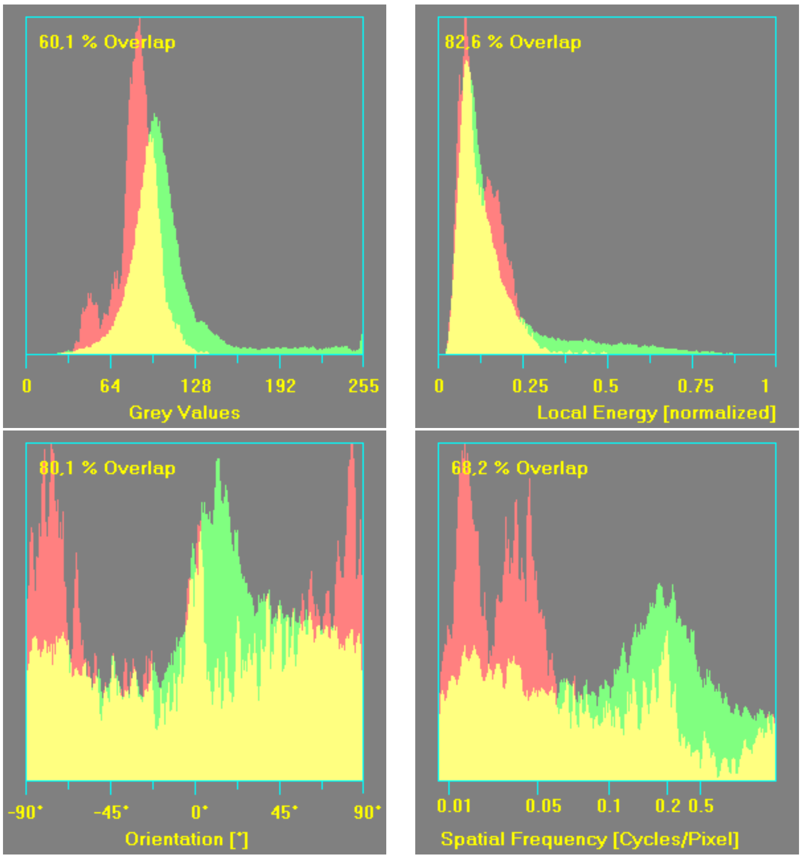


Figure 2.2 The calculated distributions of mean grey value, local energy, local orientation, and local spatial frequency for image a) in figure 2.1 are shown in the four panels above.

The overlap is given in percentage. The larger the overlap (yellow) the more difficult it is to distinguish the target from the background. In the case that the object and the background are identically the overlap would be 100 %.

The similarity measure (S_i) is the overlap calculated for the histograms within the target and object masks.

CAMAELEON requires images of good quality. Such images with camouflaged targets are available from the human observer trial [10].

The similarity measures are combined in a heuristically detectability model to calculate the detectability probability as a function of range. The single detectability probability DP_i for the local features is given by:

$$DP_i = 1 - exp[f_i * (1 - S_i) * r^2], \quad i = 1,2,3,4.$$

In this equation r^2 is the size of the target in mrad^2 and $(1 - S_i)$ are the single distance measures which are assumed to be independent of each other. To obtain detectability probabilities weighting factors (f_i) have to be used. These weighting factors f_i for the special features were found by correlation maximizing of measured and calculated detectability ranges [9].

The detectability probability DP is given by

$$DP = 1 - \prod_i (1 - DP_i) \ , \quad i = 1,2,3,4.$$

To calculate the detectability range for a special scenery (as shown in Figure 2.1a)) with a defined object against a certain background, different distances are simulated by varying the target size until the distance with 50 % detectability probability is found. The result is presented in a separate window (Figure 2.3). In this window the single statistical parameters, the detectability slope, as well as the detectability range are displayed.

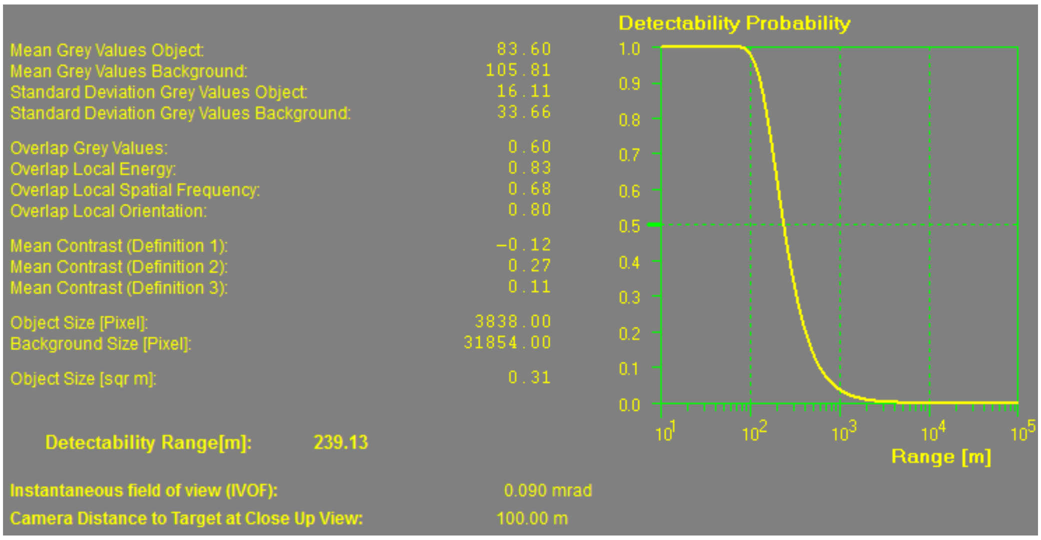


Figure 2.3 After calculating the detectability range CAMAELEON displays the following result window with the single statistical parameters, the detectability probability, and the detectability range.

In the study to be presented in this report, the discussions are based on CAMAELEON detectability range estimates of camouflage targets from FFI's observer trial [1].

2.1.2 How to use CAMAELEON

CAMAELEON requires an image of good quality and a 2^n pixels size. In these images the user has to mark both the target and the background element the target should be compared with. The marking is done by drawing a free-hand line, a polygon, an ellipse, or a rectangle and can be saved individually for target and background. This allows reloading both background and target masks to the same or to other images.

After the marking is done the following calculations can be executed: "Calculate Features", "Calculate Histograms", "Calculate Object Size" and "Calculate Detectability Range". The first two calculations are independent of other calculations, while the last two calculations require calculating histograms first. All calculations results for an image can be saved in a file, the Excel Result File. The parameters saved in this file are among others those displayed in the result window after calculating detectability ranges (Figure 2.3).

The procedure we used to perform CAMAELEON calculations is described in Section 2.4.

2.2 Targets to be evaluated by CAMAELEON and human observers

Two near identical camouflage patterns, from now on referred to as target 1 and target 2 and shown in Figure 2.4, were roller printed (HolTex, Germany) onto cotton textile (225 g/ m^2), similar to that used in military clothing, and then sewn into two mannequin suits consisting of torso and head. A styrofoam mannequin was dressed up with the two suits, one at the time, and the two camouflaged targets were then recorded in 6 various natural backgrounds (scenes) in Rhodes in August 2013. The scene images were all intended to be used in camouflage assessment with i) CAMAELEON and ii) human observers in a search by photo observer trial.

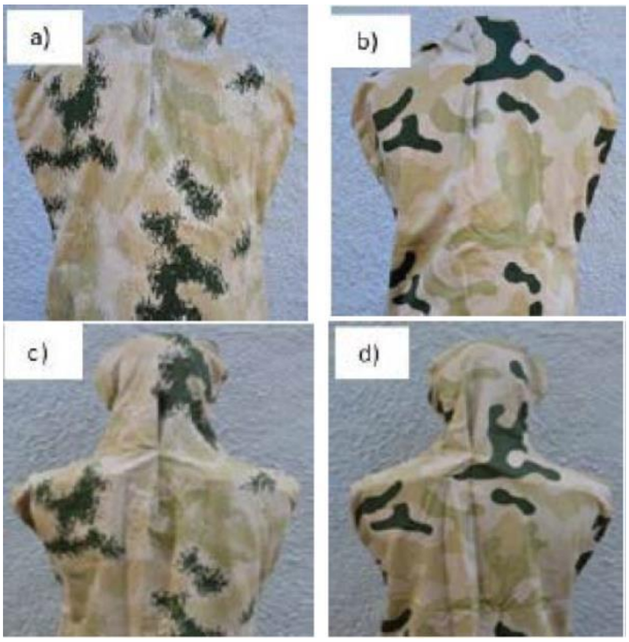


Figure 2.4 The camouflage patterns used in this study. In a) and b) we see the patterns of target 1 and target 2, respectively, with no head. In c) and d) we see the same two targets, target 1 and 2, respectively, with head. Note that in c) the head of target 1 is nearly half-filled with a highly contrasting dark green patch on the right side.

2.3 Image capture in natural backgrounds

The scenes and the corresponding target positioning were selected from the 12 available first entry scenes from our recent camouflage evaluation of camouflage patterns for first entry purposes [10]. Therefore, we now only re-state the basis of our image capture of targets in natural sceneries and the subsequent evaluation by observers.

2.3.1 Target recording by photography

The scenes were chosen to contain different types of local backgrounds around the target. We recorded the targets in as identical conditions as we were able to, considering, target orientation, position, area exposed, and as stable illumination conditions as possible. Our aim was to assure that the target's camouflage patterns were assessed solely based on their relative camouflage effectiveness. To achieve that, we carried out a near-continuous (within minutes) recording of the targets in each scene. This was done by a digital camera (Nikon D5200). Furthermore, only one target was recorded per image to avoid confusion about what is actually to be assessed by the human observer during the trial.

Targets were located randomly in the image frame (never centred) in order to avoid observers' expectations on where to start to search. The physical distance to the targets in the field was varied from17 m to 70 m in the 6 different scenes. We recorded the scene images with the intention that it was actually possible to detect the target, whenever the observer's eye focus was at the target's spot 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. In 3 of the in total 6 scenes the target was recorded with both torso and head (Figure 2.4), whereas in the 3 remaining scenes only the target's torso was recorded by photography. A collage of close up-images of the targets in the 6 different scenes is shown in Figure 2.5.

Figure 2.5 Close-up images of the 6 different scenes used in this study, both for CAMAELEON-analysis as well as in the search by photo camouflage assessment by human observers. The images show one of the two targets, either target 1 or target 2, and its local background. The 6 images are all 256 by 256 pixels prepared for further CAMAELEON-analysis of detectability. During the observer trial with humans, however, the scene images were much larger (2560 by 1600 pixels), suitable for detection times within the seconds range.

2.4 CAMAELEON simulations

We performed the CAMAELEON analysis on the two targets (target 1 and target 2) in all of the 6 scenes [3]. As shown in Figure 2.6, the target had to be marked (in red) along its outline as delicate as possible. This was done by drawing a free-hand line by an operator in the research team. Thereafter, a background had to be chosen, to which the target was compared regarding statistical overlap of local contrast, spatial frequency and local orientation [3]. Hence it was of great importance that both target was marked correctly and the backgrounds to be used to compare the target's camouflage effectiveness were identical.

In order to find the two target's detectability ranges and thereafter compare the calculated figures/numbers of target 1 *relatively* to target 2, we developed a procedure trying to minimize the sensitivity of the target marking on the final results. Each target was marked (as accurate as possible along the target outline) twice in every scene. Although the two targets were placed as identical as possible when recording the scenes by photography, their relative position in a particular scene deviated within a few centimetres. This was due to the fact that the targets appearance was never 100 % identical as the targets consisted of (soft) cotton fabrics, and hence deviated slightly due to small folds in the fabric etc. Therefore, in 5 of the 6 scenes it was not possible to use the markings of target 1 also for target 2 in the CAMAELEON calculations of detectability. Hence, the markings of target 1 and target 2 had to be drawn independently. Finally, we performed calculations of detectability of each of the target markings, using 5 different local backgrounds. An example is shown in Figure 2.6. This means that a total of 10 values of detectability were gained per target in each of the 6 scenes. We chose 256 by 256 close-up images of all of the scenes in the calculations of target detectability as CAMAELEON required 2^n pixel images. This means that only the relevant (and zoomed in) parts of the whole scene image were used in our CAMAELEON analysis.

In addition we performed calculations on target detectability in all scenes for at least one target, with the aim of studying the sensitivity of the different (drawn by human) target markings on the estimated detectability distance for a fixed background. Based on the results that were found, we were able to avoid markings of the targets in our study which would lead to extreme detectability range values.

Figure 2.6 Illustration on how one of the targets was marked (red) in scene 2 for further use by CAMAELEON. The green halo above and to either side of the target shows one of our 5 differently chosen local backgrounds, to which the targets were to be compared, in order to find an estimate on the targets detectability, given in metres.

2.5 Search by photo trial with human observers

The number of human observers that were used per scene varied between 24 (scene 3 and 5) and 36 (scene 1) in our observer trial. This was due to the fact that the observer results to be presented in this paper were a part of a larger camouflage study where a total number of 148 observers were used to assess 9 or 12 (some targets did not facilitate the use of a hood) unique targets per scene. However, the number of observers was the same for both target 1 and 2 (to an accuracy of one single observer) in each of the 6 scenes presented in this paper, and hence allows for a relative comparison of camouflage effectiveness per scene.

2.5.1 Preparation of human observers

Prior to the observer trial each soldier was given a word by word *identical* introduction to the observer trial by an instructor. Each observer was then adjusted to have an optimal and identical distance to the widescreen (ca. 40 cm), as the screen was intended to fill most of the observers' field of view Also, the observer's eyes were approximately levelling the centre of the pc-screen. Thereafter, 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 to the instructors, reducing the risk of misunderstandings before the main trial started. During trial itself, observers were not allowed to ask questions, but left to find targets solely by themselves. Finally, the observers were free to choose their own search strategy during the trial.

2.5.2 Conducting the observer trial

During the observer trial each observer was shown a randomized sequence of photographs of the 6 different scenes, one at the time in a high definition (HD) wide screen (2560 x 1600 pixels) in a dimly lit room. Each photograph represented a scene with either one or no target [1]. The observers searched for a target and indicated detection by mouse-clicking at the target as soon as he or she felt confident that it was a proper target and not an anomaly (e.g. a target-shaped bush). There was a small tolerance surrounding each target. Hence the observer had to click close to the target to indicate a "hit", but not necessarily spot on the centre of the target. Of importance was that each human observer was exposed no more than one single target per scene, as targets were identically positioned in each scene.

The total duration of each scene image presentation was limited to 60 s in order to give the observer reasonable search time, but on the same time avoid a reduction of the observer's concentration through tedious searches for a well hidden target. Whenever our fixed time limit was exceeded, the target (in that particular scene) was stored as a non-detection". Furthermore, we used a software tool, showing the scenes in a randomized order and keeping data in a sorted manner for further analysis.

2.6 Statistical analysis

2.6.1 CAMAELEON-data

For the CAMEALEON study we were not able to use exactly the same target masking for target 1 and target 2 in 5 of the 6 scenes. To estimate the uncertainty which was introduced by different selections of the target area, we calculated 10 detectability ranges for one target for each scene by choosing 10 different target markings while keeping the marked background area the same. From these 10 detectability ranges we calculated an average detectability range and the standard deviation (Table 3.1). The obtained standard deviations assured that comparing the two targets by applying the described procedure lead to meaningful results.

2.6.2 Analysing observer trial data

The detection time data from the observer trials, using humans, were inspected further in order to look for significant differences between the two target's camouflage effectiveness in each of the 6 scenes. We first carried out a Jarque-Bera test for each target per scene, testing whether the corresponding distribution of detection times was normally distributed or not. Whenever at least one of the target's detection times, in a particular scene, failed to fulfil normality we carried out a Wilcoxon on rank test (Mann-Whitney U-test) as such a non-parametric test has shown to be more trustworthy than the parametric counterpart (such as ANOVA) in such cases

[13;14]. Also, the Wilcoxon's rank test, being non-parametric, has the ability to account for non-detections (i.e. detection times larger than the search time limit set to 60 s) that appeared for some of the target's in some scenes during the observer trial. In total, we then gained p-values describing the degree of similarity in camouflage effectiveness between target 1 and target 2 for each of the 6 scenes. A further description on how to handle trial detection data, including the non-detections, can be found in a recent study on camouflage assessment methodologies [1;10].

3 Results

The performance results for the of the two near-similar camouflage pattern targets in 6 scenes obtained by CAMAELEON detectability range calculations and a search by photo observer trial will first be described for each of the methods, individually. At the end of the chapter we will compare the results and emphasise the effect of the disruptive pattern on the head of the targets.

3.1 CAMAELEON results

The results obtained by calculating similarity measures and based on these detectability ranges for target 1 and 2 in the 6 scenes are given in Figure 3.1. For each scene we selected five different, but relevant local background areas. All five backgrounds were selected in close vicinity to the target. Three of the backgrounds surrounded the target and differed only in size. One background mask was chosen without including any foreground (such as Figure 2.6), while the last one did not include any foreground and in addition was marking the most relevant background on only one side of the target. These 5 backgrounds were used for detectability range calculations for both targets. Additionally, each target was marked twice as similar as possible in each scene. This resulted in 10 different detectability ranges for each target in each scene, which are plotted in Figure 3.1. The 10 calculated detectability ranges for one target in one scene spanned different intervals. The spread of the calculated detectability ranges in-between the highest and the lowest value varied from 6 m for target 2 in scene 5 and up to 42 m for target 2 in scene 4. The average detectability ranges together with the standard deviations for each target in each scene are given above or underneath the data points.

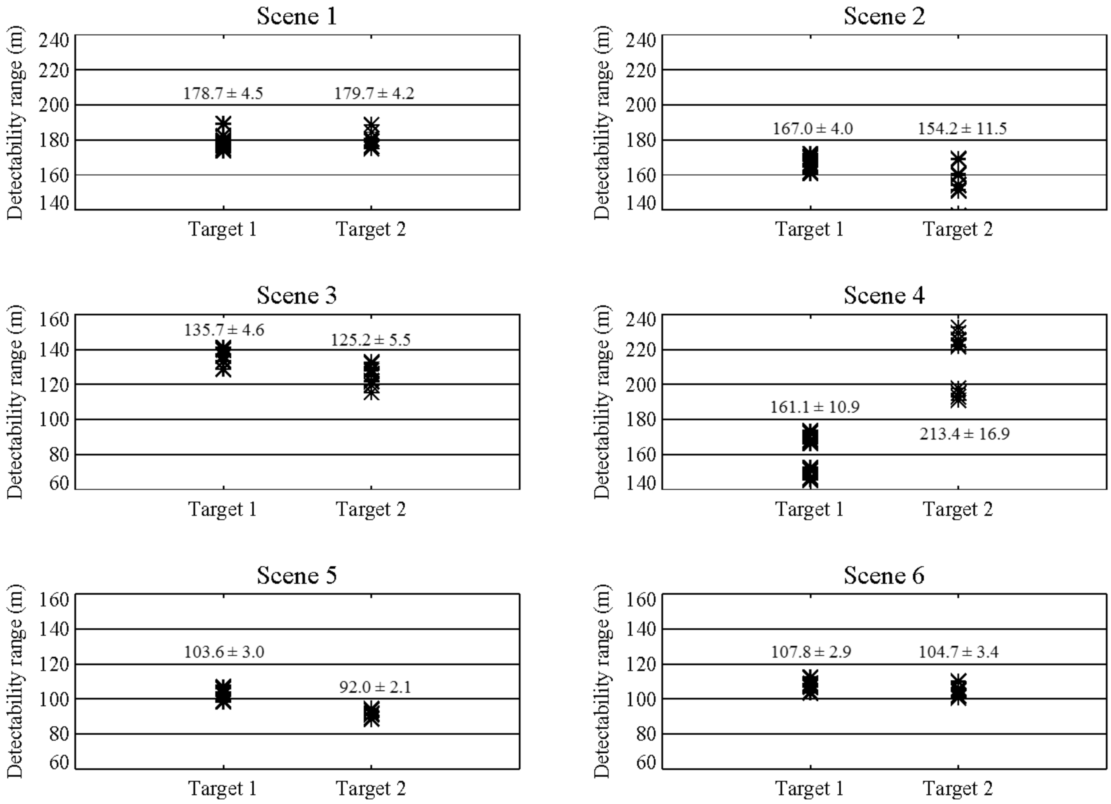


Figure 3.1 The 6 plots show the obtained 10 detectability ranges for the 10 combinations of target (two different target masks) and background (five different background masks) markings for each target and each scene, together with the average detectability range and standard deviation.

3.1.1 The role of target markings on the results

Different markings of the target and background resulted in different detectability range values. This is most likely due to the fact that each pixel of the image contributed to the calculation of measure of similarity. To estimate the influence of the target markings on the final result, we calculated detectability ranges for 10 target masks while keeping the background mask unchanged. This was done for one target and one background in each scene. Table 3.1 lists the obtained average detectability ranges with the corresponding standard deviations. These results were used in two ways. First to assure that the two markings of the targets for the calculations presented in Figure 2.6 would result in detectability ranges with low deviations. Furthermore, did the results indicate that the uncertainty introduced due to the different target masks did not lead to meaningless results for the study comparing the two targets.

Table 3.1 Average detectability ranges calculated for target 1 for each scene with 10 different markings of the target and one determined background mask.

Scene	Detectability range (m)
1	174,28 ±3,1
2	165,1 ±2,2
3	139,6 ±0,6
4	168,5 ±2,9
5	106,5 ±3,4
6	105,5 ±3,4

3.1.2 Average detectability ranges

Table 3.2 gives an overview over the average detectability ranges for target 1 and target 2 in the 6 scenes together with their distances to the camera, and if they were used with head/hood or not. The right-most column of the table presents the relation between the detectability ranges of the two targets. In 2 of the 6 scenes (scene 1 and 6) both targets performed almost equal (target 1/ target 2 values close to 1). In scene 2, 3, and 5 target 2 performed slightly better than target 1, while in scene 4 target 1 performed much better than target 2.

Table 3.2 Detectability ranges for target 1 and target 2 in the 6 scenes together with their distances to the camera and if they were used with head or without. The last column presents the relation between the detectability ranges of the two targets.

Scene	hood	Distance to target (m)	Detectability range (m) Target 1	Detectability range (m) Target 2	Target 1/ Taget 2
1	with	41	178,7 ±4,5	179,7 ±4,2	0,99 ±0,05
2	without	39	167,0 ±4,0	154,2 ±11,5	1,08 ±0,11
3	without	41	135,7 ±4,6	125,2 ±5,5	1,08 ±0,09
4	with	17	161,1 ±10,9	213,4 ±16,9	0,76 ±0,11
5	without	70	103,6 ±3,0	92,0 ±2,1	1,13 ±0,06
6	with	17	107,8 ±2,9	104,7 ±3,4	1,03 ±0,06

3.2 Observer trial results

The results obtained from the performance study based on human observers were detection times for each target in each scene [1]. The median values for the detection times are given in Table 3.3, together with the number of observers for each scene and the number of non-detections. This was the case if the observer didn't detect the target at all within the detection time limit of 60 s (i.e. the detection time limit). The two targets showed different performances for each of the scenes. To determine if the differences in the median detection times for the two targets were significant the Wilconxon rank test was applied [15]. For scene 2, 3, and 5 no significant differences in the performance of the two targets were found, in contrast to scene 1, 4 and 6, were the targets performed significantly different.

Table 3.3 Median observation times as result of the human observer trial for each target in each scene, together with the assessment of the significance of the obtained differences. Furthermore, the number of observes for each scene and the non-detections are listed.

Scene		Target 1		Target 2		
	No. of observers	Median detection time (s)	non-detection	Median detection time (s)	non-detection	Significant different
1	36	28,6	8	8,2	1	yes
2	27	23,6	2	27,7	3	no
3	24	5,3	1	9,8	1	no
4	35	38,0	7	15,3	2	yes
5	24	16,4	1	15,0	2	no
6	32	60	12	16,5	3	yes

3.3 CAMAELEON versus Observer trial results

The results for the two different methods that we used, described in Chapter 3.2, for the two targets in the 6 different scenes are presented in Table 3.4. CAMAELEON assessed the performance of the two near-identical targets as very similar in 5 of the 6 scenes. Only in scene 4 CAMAELEON predicted a better performance of target 1 in comparison to target 2. For this scene also the human observer trial resulted in different median detection times with target 1 having a near significant longer median detection time than target 2 as shown also earlier in the FFI rapport "Effekten av konturbrytende mønstre på deteksjonstid av kamuflerte mål" [12]. The human observer trial revealed in addition a significant difference in median detection times for the 2 targets in scene 1 and 6, in which the targets were used with head (hood). Target 1 performed significant better in both these scenes, too.

Table 3.4 Comparison of the relations of detectability ranges (given with standard deviation) from the CAMAELEON study and the observer times (with significance) from the human observer trial for target 1 and target 2.

Scene	Detectability Target1/Target 2	SD	Observation times Target1/Target 2	Significance
1	0,99	±0,05	3,49	yes (p=0,027)
2	1,08	±0,11	0,85	no (p=0,864)
3	1,08	±0,09	0,54	no (p=0,248)
4	0,76	±0,11	2,48	yes (p=0,064)
5	1,13	±0,06	1,09	no (p=0,707)
6	1,03	±0,06	3,64	yes (p=0,001)

4 Discussion

In this Chapter we will discuss our results, focusing on the similarities and differences between human observer trials and CAMAELEON simulations. We have studied how signature assessments induced by a simulation tool (CAMAELEON), mimicking the early stages in the human visual sensory system [3], correlate with human observers' assessment of two camouflaged targets in natural settings[12]. We found good correlation. In some cases, however, which seem to have been connected with the way the pattern was distributed within the target, the camouflage effectiveness of the targets were very much differently assessed by CAMAELEON and human observers.

4.1 CAMAELEON results versus human observers

In Figure 3.1, which represents the main finding in this study, we see that the two targets were assessed by the CAMAELEON software tool to be similar or near similar in camouflage effectiveness in 5 of 6 scenes. In scene 4 it seems that target 1 was assessed by CAMAELEON to have shorter detectability range (and hence better camouflage properties in that particular natural scenery) than target 2. The results derived from human observers, assessing the same two targets in the same 6 natural scenes, show that the targets were assessed significantly different, given by their median detection time, in 3 of the 6 scenes.

4.1.1 The previously reported anomaly

A careful inspection of the 3 scenes, in which the two (near identical) targets were assessed to pose different camouflage effectiveness by CAMAELEON and the human observers, shows that the targets were always with head (hood) in these scenes. In the remaining 3 scenes, both targets were exposed with no head to the human observers and CAMAELEON. The two targets, in this study were near similar, but not exact copies of one another. This holds also for the two hoods, seen in Figure 2.4 c) and d). A further visual inspection of the two target heads shows that target 1 had a large dark green patch in the pattern filling most of the right part of the head, whereas target 2 did not have such a large distinguished pattern. This large dark patch in the pattern may have had an important effect on how the human observers perceived the two targets [12]. In the following section we try to give an explanation to the observed difference on how the targets were assessed (to be similar) by CAMAELEON and (to be different) by human observers, regarding their camouflage effectiveness.

4.1.2 The effect of disruptive patterns on the results as a function of evaluation method

The CAMAELEON simulation tool assesses any target, relative to its background based on similarity measures based on local contrast, spatial frequency and local orientation [3]. Hence, CAMAELEON may not necessarily respond to minor variations, regarding how the camouflage pattern elements are distributed within the target, as long as the three above mentioned parameters are unchanged in overall. The human visual perception – from its first sensory system to the cognitive understanding on similarity of shapes - is likely to be more delicate than represented by the three parameters contrast, spatial frequency and orientation [16]. Studies on visual search processes indicate that the observer normally has some (pre-induced) representation on the visual properties of the target to be searched for and, furthermore, some description on its physical properties, allowing the observer to distinguish the target from a complex background [4]. Hence, the high contrast pattern patch on the head of target 1 may have altered its properties in a way that affected the detection time significantly. It has been reported that edges are a salient feature of target detection processes [17], and the high contrast marking may have disturbed this feature, inhibiting detection.

It has further been reported that detection through some organic visual perception mechanism, consists of four stages [4]. The steps are i) detecting an anomaly in the background, ii) identifying the region of a target, iii) describing its contours, and iv) verifying the target being real. Step iii) involves the observer being able to describe and perceive the contour of the target.

In our study, we found that target 1, with a contrasting marking at its head contour, had significantly longer detection times compared to its near-similar counterpart, target, 2, but with no such large marking. We believe this has to do with step iii) by some affection on the observer's ability to locate the targets outline in a search process.

Finally, it has been reported [18] that the human visual system is optimized for capturing the spatial information of natural visual images, and it would be interesting to know more on how such a high contrast patch, as was found in target 1 in our study, may have disturbed the observer's attention when the area where the target was eye scanned during the search process.

4.2 CAMAELEON results

What can we learn about the CAMAELEON results from this first study? Figure 3.1 presents the 10 detectability range calculations, belonging to 5 different background elements (masks) and 2 different target masks, for each target in each scene. Comparing the distribution of the detectability ranges in all 6 scenes it seems like there is an overall tendency that target 2 performed slightly better than target 1, with the exception of scene 4. In scene 4 target 2 performed significantly better than target 1.

4.2.1 The role of local background to the estimated detectability range

The detectability range distributions, caused by the use of 5 different background masks, revealed different characteristics in the different scenes. For scene 1 and 3 the values for the 10 detectability ranges for one target spread over intervals of 12-17 m with no significant difference in the performance of the two targets. However, we want to stress that in scene 1 the two targets were assessed by CAMAELEON to perform equal. For scene 5 and 6 the 10 detectability ranges could be found in intervals of 6-9 m. In these two scenes (Fig. 2) the backgrounds consisted mainly of stones and rocks. This represented a background with little variation in colour, size or structure. Also the contrasts (local energy) exhibited a homogenous distribution due to distinct illumination differences between sunlit and shadow areas. Hence, these backgrounds were quite uniform with regard to local contrast, spatial frequency and local orientation [3]. Scene 6 consisted in addition of larger regions in the shadow. For the background masking we only used small amounts of these shadow areas.

Accordingly, the 5 background masks in these two scenes resulted in similar detectability ranges for one target in one scene due to the uniformity of the background. The detectability ranges for scene 2 and 4 differed from the rest of the scenes because of their less compact distributions for both targets. This can be explained by the background mask choices. In both the scenes the foreground right in front of the targets was quite different to the background around the target. These foregrounds (scene 2 green bushes and scene 4 sunlit grass) were included in different amounts in the background masks, ranging from not accounted for to about a third of the mask. In addition did the distribution of the detectability ranges for target 1 varied less than for target 2 (Fig. 4) in both these scenes. In scene 2 the targets was used with head (hood) and in scene 4 without. This means that the calculations of measure of similarity suggested that the camouflage pattern of target 2, independent of the pattern distribution on the hood, was more sensitive to the background elements chosen for the different background masks.

Scene 4 differed from all the other scenes in a number of aspects. The average detectability ranges for both targets had the largest standard deviations, or in other words showed the widest distributions. The presumed reason for this has been discussed above. Furthermore, we find that the detectability ranges for one target accumulated in two groups. The grouping was due to the choice of background elements. 2 of the 5 background masks did not include the sunlit grass in the foreground. For the according 4 calculations we obtained detectability ranges grouped around 25 m lower values than the other 6 detectability ranges which includes the sunlit grass. It should be noted, that this was also the only scene where the targets were placed in the shadow. At last, in this scene, as the only scene with such a result, the difference in the performance of the two targets seemed to be significant.

The study showed that the choice of background had a large influence on the detectability range calculations. For comparison of the performance of different camouflage patterns in the future the same background element therefore has to be applied.

4.2.2 The role of target marking

The marking of different parts of the images, both for the target (object marking) and the background, influenced calculated detectability ranges. For our comparison study of the effectiveness of the camouflage pattern of target 1 and target 2 the target masking was different for the two targets, with one exception, namely scene 5. As the different masks did not introduce large deviations in the obtained detectability ranges we feel confident that the differences in the results of the CAMAELEON study and the human observer trial are not a consequence of the choice of target masks, but rather due to other, perceptional mechanisms.

5 Future follow-up studies

Using simulation software to evaluate the performance of camouflage pattern is obviously more convenient, more flexible and faster than performing observer trials. CAMAELEON provides advantages compared to observer trials which makes an extended use of this tool interesting for us. Thus it could be beneficial to perform selected studies only based on CAMAELEON. Before we do that it is very important to understand the drawbacks and limitations of the simulation software and identify sources which can lead to misinterpretation of the results.

5.1 Advantages using a simulation tool

CAMAELEON calculations are based on bitmap type images. All images taken in the past can be used for CAMAELEON calculations. During an observer trial the performance of the target depends on the direct surroundings and the results do not contain any information on how the target would perform in another position with different surrounding elements in the vicinity of the target in the same scene. This information can easily be obtained with CAMAELEON. After the target is marked, any arbitrary background element in the same image can be chosen to calculate similarity measures or detectability ranges.

Another advantage of CAMAELEON simulations is that it doesn't play a role where the image comes from. This means that we can pick an image of any landscape, place targets form conducted observer trials or artificial targets in the image and evaluate the performance of the targets. Composing artificial images most likely increases the uncertainty in the results based on, for example, different light conditions in the scene and for the target. But this still can be a possible way to go when the goal of the study is to investigate relative performance of differences for the targets in many different background types.

There are some other obvious advantages with simulations. Assumed that images are available CAMAELEON is an effective tool for low cost camouflage studies with low risk. In opposite to observer trials there are no observers, no traveling, and no extensive preparations necessary. The software can provide quick answers to new appearing problems and tasks.

5.2 Possible disadvantages with CAMAELEON

Comparing human observer trial results with CAMAELEON results stirred up our interest. It seems that CAMAELEON does not simulate all aspects of human camouflage detection skills. Thus, evaluating camouflage effectiveness only with CAMAELEON can lead to wrong results and conclusions, while human observer trial results seem to be more reliable.

5.3 Further studies

To use CAMAELEON properly and to understand which conclusions can be drawn from the CAMAELEON results we need to perform further studies. Comparing CAMAELEON results with results from other trusted methods, like our observer trail, will provide the necessary data basis for those studies.

More extensive comparison study of observer trial and CAMAELEON results

In the recent observer trial [10] we studied the relative performance of 6 "Norwegian" patterns in 14 Norwegian landscape scenes as well as 9 "First entry" patterns in12 different arid surroundings. As the images from this study are very well suited to be used for CAMAELEON calculations, this gives us a good basis for further comparison studies.

Already during the analysis of the results from that trial we used CAMAELEON for the first time to check our results against another evaluation method for camouflage performance. Only a few scenes were selected for the comparison. Some of the received results were in agreement with the observer trail results, while others clearly differed [1]. We didn't further analyse these findings.

However, it seems to be valuable to start a more extensive study now by choosing a representative selection of these scenes and evaluate the camouflage performance of handpicked targets also with the CAMAELEON software. Investigating the relative performances of the different targets in a given scene and comparing those to the relative performances received from the observer trial results will lead to a better understanding of the strengths and weaknesses of those two different evaluation methods. A founded understanding of how CAMAELEON results can be interpreted must form the basis for further studies. After we have achieved this understanding we plan to perform future CAMAELEON studies which are described the in the following paragraphs.

Further study of the influence of disruptive pattern

The comparison of the performance of the two targets utilizing the two evaluation methods (CAMAELEON and human observers) showed in some scenes significant differences in the evaluation results. The results we presented in this report are based on an interesting "performance anomaly" we found during a camouflage study when comparing two near-identical targets. Different findings on the influence of markings overlapping edges of a target are reported in studies on disruptive coloration and background pattern matching in reducing detectability [19-23]. To verify the presented preliminary results we suggest a follow-up study using one distinguished camouflage pattern applied in varying ways, in terms of pattern patch distribution, to the target. In case of our target mannequin (Figure 2.4) the high contrast colour patch can be applied to a part of the hood as well as a shoulder. A carefully selected number of varying scenes could allow studying different aspects of target placing, e.g. the influence of the target being placed in the shadow, on the comparison study results.

Camouflage pattern developed in the last years tend to favour small pattern structures, e.g. pixel structure of the US Universal Camouflage Pattern. This study shows that there may be potential pitfalls whenever pattern structures are too small, and that our reported effect of camouflage pattern contrast patches should be kept in mind.

Study the targets against other background elements
The aim of our observer trials is to recommend camouflage pattern for the Norwegian Army. To reinforce the foundation for our recommendations it would be useful to study the pattern, we found to perform best, against other backgrounds. For observer based studies we would dependent on new recordings of the targets in other natural backgrounds. However, a camouflage evaluation can be done by choosing other relevant background elements in each scene in the existing images and calculating CAMAELEON detectability ranges.

Studies of pattern effects and the influence of additional equipment
One interesting finding in this first CAMAELEON study was that the calculations suggested that the blurred pattern of target 1 was less sensitive to changes in the background. This could also be worth more extensive exploration in a later follow-up study.
Working with signature control it is not enough only to improve the camouflage pattern of the battle dress uniforms. A soldier will always wear some equipment, among others weapons, wests, and backpacks, in addition to the uniform. Investigating the effect of this additional equipment on the detectability could also be interesting to study with a simulation tool. In this case we have to make sure, that CAMAELEON evaluates differences in the same way as humans. This in itself would be worth a study and we intent to follow this up.

6 Conclusions

In this report we have presented results of a first study applying a software tool to evaluate the performance of camouflage patterns in natural backgrounds. In this study we chose to investigate an anomaly in the performance of two nearly similar camouflage patterns which we found in the results of an earlier conducted observer trial [1;12;23;23]. Comparing the results applying two different methods we have seen how the human visual perception may be sensitive to small – but important – deviations within a camouflage pattern, deviations that simulation tools mimicking the eye's sensory response will not necessarily pick up. Contrast patches in the camouflage pattern located differently particularly with regard to the outline of the target seemed to have an influence on detection times for human observers.

The use of simulation software to evaluate the performance of camouflage patterns implies a number of advantages. It is more flexible and results can be obtained much faster than by performing observer trials. Therefore we would like to use CAMAELEON regularly in our camouflage studies whenever this is possible. To be able to do this we have to have a good understanding of the value and reliability of the simulation results. To obtain a better knowledge of the CAMAELEON reliability und to learn more about different camouflage patterns we have concluded this report suggesting a number of possible future studies.

References

[1] G. K. Selj, D. Heinrich, and M. Søderblom, "A Method for test and evaluation of signature properties based on observers.," FFI rapport 2014.

[2] G. K. Selj and M. Søderblom, "Besøk hos DSTL, Portsdown, UK - utvikling av kamuflasjemønstre," FFI Reiserapport 2013/01718 (Begrenset), 2013.

[3] R. Hecker, "Camaeleon - Camouflage assessment by evaluation of local energy, spatial frequency and orientation,", 1687 ed. C. Watkins, Ed. 1992, pp. 342-349.

[4] T. S. Troscianko, C. P. Benton, P. G. Lovell, D. J. Tolhurst, and Z. Pizlo, "Camouflage and visual perception," *Phil. Trans. R. Soc. B*, vol. 364, pp. 449-461, 2009.

[5] C. M. Birkemark, "CAMEVA, a methodology for computerised evaluation of camouflage effectiveness and estimation of target detectability,", 3699 ed 1999, pp. 229-238.

[6] S. Nyberg and L. Bohman, "Assessing camouflage using textural features ,", 4370 ed 2001, pp. 60-71.

[7] A. W. Houlbrook, M. A. Gilmore, I. R. Moorhead, D. Stroud, G. Hutchings, and A. Kirk, "Scene simulation for camouflage assessment," SPIE 2000: 2000.

[8] J. R. McManamey, "Validation plan for the German CAMAELEON model," ,", 3062 ed 1997, pp. 300-310.

[9] R. Hecker, "Efficent methods for validation target acquisition models, "RTO-MP-45, 1999.

[10] G. K. Selj, M. Søderblom, D. Heinrich, and T. Høimyr, "Utvikling og evaluering av kamuflasjemønstre for stridsuniform," FFI rapport 14/00885 (Begrenset), 2014.

[11] D. Heinrich and G. K. Selj, "The effect of contrast in camouflage patterns on detectability by human observers and CAMAELEON," in *Signal and Image Processing, and Information Fusion Applications II*. SPIE 2015, Ed. 2015.

[12] G. K. Selj, "Effekten av konturbrytende mønstre på deteksjonstid av kamuflerte mål.," FFI rapport 14/01474 (Unntatt offentlighet), Oct.2014.

[13] S. S. Sawilowsky, "Misconceptions leading to to choosing the t test over the Wilcoxon Mann-Whitney test for shift in location parameter," *J Mod Appl Stat Method*, vol. 4, pp. 598-600, 2005.

[14] P. J. Bickel and K. A. Doksum, *Mathematical statistics - Basic ideas and selected topics*. Oakland, USA: Holden-Day, 1977.

[15] G. K. Selj and D. Heinrich, "Search by photo methodology for signature properties assessment by human observers," in *Signal processing, sensor/information fusion, and target recognition* SPIE 2015, 2015.

[16] T. J. Doll and R. Home, "Lessons learned in developing and validating models of visual search and target acquisition ,", RTO-MP-45, 1-1 - 1-8 ed 2000.

[17] D. Marr and E. Hildreth, "Theory of edge detection," *Proc R Soc Land B*, vol. 207, pp. 187-217, 1980.

[18] C. A. Parraga, T. S. Troscianko, and D. J. Tolhurst, "The human visual system is optimized for processing the spatial information in natural visual images," *Current Biol*, vol. 10, pp. 35-38, 2000.

[19] S. Fraser, A. Callahan, D. Klassen, and T. N. Sherratt, "Empirical tests of the role of disruptive coloration in reducing detectability," *Proc R Soc B Biol Sci*, vol. 274, no. 1615, pp. 1325-1331, July2007.

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

[21] S. Merilaita and J. Lind, "Background matching and disruptive coloration, and the evolution of cryptic coloration," *Proc R Soc B*, vol. 272, pp. 665-670, 2005.

[22] M. Stevens, I. C. Cuthill, A. M. M. Windsor, and H. J. Walker, "Disruptive contrast in animal camouflage," *Proc R Soc B*, vol. 273, pp. 2433-2438, 2006.

[23] G. K. Selj, "Disruptive coloration tricks the human eye - a study of detection times of two near-similar targets in natural backgrounds," in *Targets and Background Signatures,* 9653 ed. K. Stein and R. Schleijpen, Eds. SPIE, 2015.

About FFI

The Norwegian Defence Research Establishment (FFI) was founded 11th of April 1946. It is organised as an administrative agency subordinate to the Ministry of Defence.

FFI's MISSION

FFI is the prime institution responsible for defence related research in Norway. Its principal mission is to carry out research and development to meet the requirements of the Armed Forces. FFI has the role of chief adviser to the political and military leadership. In particular, the institute shall focus on aspects of the development in science and technology that can influence our security policy or defence planning.

FFI's VISION

FFI turns knowledge and ideas into an efficient defence.

FFI's CHARACTERISTICS

Creative, daring, broad-minded and responsible.

Om FFI

Forsvarets forskningsinstitutt ble etablert 11. april 1946. Instituttet er organisert som et forvaltningsorgan med særskilte fullmakter underlagt Forsvarsdepartementet.

FFIs FORMÅL

Forsvarets forskningsinstitutt er Forsvarets sentrale forskningsinstitusjon og har som formål å drive forskning og utvikling for Forsvarets behov. Videre er FFI rådgiver overfor Forsvarets strategiske ledelse. Spesielt skal instituttet følge opp trekk ved vitenskapelig og militærteknisk utvikling som kan påvirke forutsetningene for sikkerhetspolitikken eller forsvarsplanleggingen.

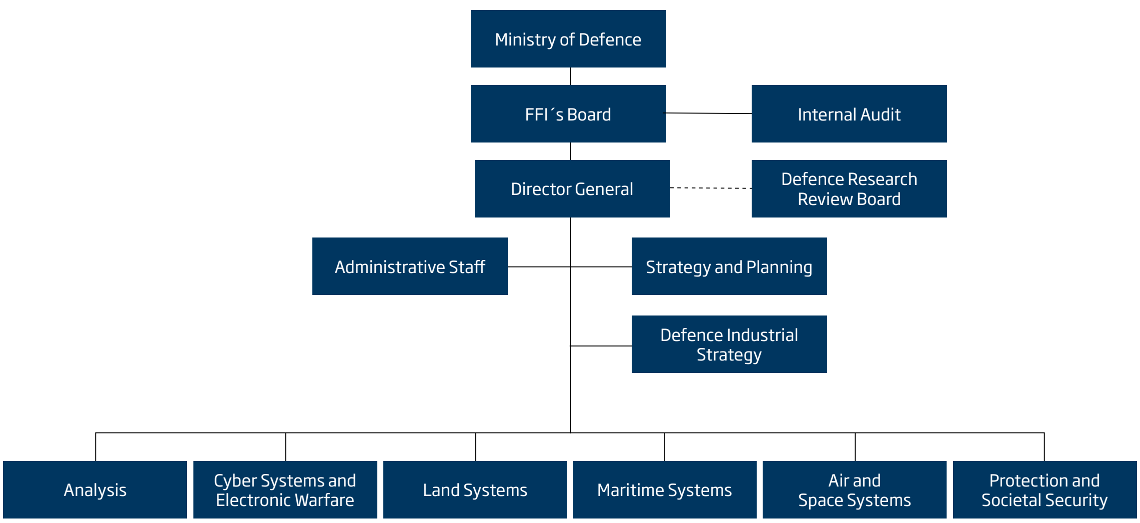
FFIs VISJON

FFI gjør kunnskap og ideer til et effektivt forsvar.

FFIs VERDIER

Skapende, drivende, vidsynt og ansvarlig.

FFI's organisation



Forsvarets forskningsinstitutt
Postboks 25
2027 Kjeller

Besøksadresse:
Instituttveien 20
2007 Kjeller

Telefon: 63 80 70 00
Telefaks: 63 80 71 15
Epost: ffi@ffi.no

Norwegian Defence Research Establishment (FFI)
P.O. Box 25
NO-2027 Kjeller

Office address:
Instituttveien 20
N-2007 Kjeller

Telephone: +47 63 80 70 00
Telefax: +47 63 80 71 15
Email: ffi@ffi.no

FFI Forsvarets forskningsinstitutt
Norwegian Defence Research Establishment