Optical Engineering

SPIEDigitalLibrary.org/oe

Urban camouflage assessment through visual search and computational saliency

Alexander Toet
Maarten A. Hogervorst

Optical Engineering 52(4), 041103 (April 2013)

Urban camouflage assessment through visual search and computational saliency

Alexander Toet
Maarten A. Hogervorst
TNO
Kampweg 5
3769 DE Soesterberg, The Netherlands
E-mail: lex.toet@tno.nl

Abstract. We present a new method to derive a multiscale urban camouflage pattern from a given set of background image samples. We applied this method to design a camouflage pattern for a given (semi-arid) urban environment. We performed a human visual search experiment and a computational evaluation study to assess the effectiveness of this multiscale camouflage pattern relative to the performance of 10 other (multiscale, disruptive and monotonous) patterns that were also designed for deployment in the same operating theater. The results show that the pattern combines the overall lowest detection probability with an average mean search time. We also show that a frequency-tuned saliency metric predicts human observer performance to an appreciable extent. This computational metric can therefore be incorporated in the design process to optimize the effectiveness of camouflage patterns derived from a set of background samples. [DOI: 10.1117/1.OE.52.4.041103]

Subject terms: camouflage; clutter; saliency; search time; detection probability.

Paper 121130SS received Aug. 2, 2012; revised manuscript received Sep. 9, 2012; accepted for publication Oct. 5, 2012; published online Nov. 30, 2012.

1 Introduction

There is currently a renewed interest in the design and evaluation of improved personal camouflage.[1–5] Traditionally camouflage design was (and still is) largely based on intuition and aesthetics, and was performed by industrial designers and (especially in the beginning) by artists[6] (incl. cubists such as Georges Braque).[7,8] The design principles were often inspired by nature and based on biological principles such as blending and disruption.[9–11]

More recently, scientific studies into the statistics of natural images, texture, visual perception and psychology[12] have entered the design process. As a result, it is now commonly acknowledged that effective camouflage patterns should contain details at multiple spatial scales (i.e., they should consist of both micro- and macropatterns), and should be similar in composition to natural images. Various camouflage patterns have been developed for a range of different theaters, such as woodland, jungle, desert, and arctic.[13] Current camouflage design follows two different trends. One trend is towards "universal" designs (such as the Army Combat Uniform of the U.S. Army) that perform well in a wide range of environments.[13] Another trend is towards more specialized designs, such as an individual design for urban environments (e.g., the Canadian Urban Environment Pattern which is currently under investigation). A universal pattern will most likely perform less well than a pattern that has been specifically designed for the environment in which it is deployed (like specialized patterns used for hunting, e.g., www.mossyoak.com). The challenge is therefore to design a pattern that is near-optimal for a variety of environments.

Urban environments differ largely from other theaters, both in their visual aspects and functional requirements. A specialized urban camouflage design may therefore outperform the more traditional patterns that were primarily designed for deployment in rural environments. Camouflage requirements for urban areas present a different challenge from those of natural environments like woodland or desert terrains. In urban combat, the tactical ranges are much closer than in woodland or desert warfare (over the last years the mean engagement distance has decreased to about 40 m). This suggests the need for camouflage designs with smaller macro patterns (i.e., patterns that resemble their background at closer viewing distances). Another common assumption is that urban backgrounds require camouflage schemes with vertical and horizontal straight edges, in combination with large macro patterns that mimic the pattern of the buildings and other man-made objects. Whereas natural environments display details at a wide range of scales (their spatial structure is approximately independent of the viewing distance),[14] man-made environments basically contain elements at two distinct scales (consisting of large structures and small details, and not much significant structure in between). As a result, multiscale and disruptive patterns may be less effective when viewed against contrasting areas of solid color like walls or pavement, and low contrast or uniform patterns may perform better in urban environments. These assumptions about optimal urban camouflage were evaluated in this study by comparing the relative performance of camouflage patterns that differ widely in their structure.

In the rest of this paper, we first present a new method for the construction of a multiscale camouflage pattern from a given set of background image samples. This method can be used both to create specialized camouflage designs (by using a set of background image samples from a single theater, e.g., an urban environment) and multitheater designs (by using a set of background image samples from a variety of theaters). Then we show how the method was applied to derive a camouflage pattern for a semi-arid urban area (designated by NATO SCI-157). Finally we present the

0091-3286/2012/$25.00 © 2012 SPIE

results of psychophysical (search) and computational (saliency metric) comparative evaluation studies that were performed to assess the effectiveness of the TNO multiscale camouflage pattern relative to that of 10 other (multiscale, disruptive, and monotonous) patterns that were also especially designed for the same area (by other members of the NATO SCI-157 group).[15]

2 Camouflage Patterns

In this section, we present the TNO method for deriving a multiscale camouflage pattern from a given set of (characteristic) background images. We also briefly describe the other camouflage patterns that were used in this study, that were designed by other members of NATO SCI-157.

2.1 *Test Site and Color Selection*

The test site used in this study was Fort Huachuca, a small town in a semiarid environment (southeast Arizona, United States; Fig. 1).

Fig. 1 Some typical views of the test site (Fort Huachuca, Arizona, United States).

The aim of this study was to assess the camouflaging properties of different spatial patterns, not to compare the effectiveness of different colors. Therefore, all patterns investigated in this study used the same set of 4 colors (Fig. 2). These colors were derived from the analysis of a collection of images representing various characteristic locations in the test environment, and from on-site color measurements.

Shape-disruptive patterns require bright and dark colors to create an appreciable luminance contrast. A dark color (rightmost color in Fig. 2) was therefore derived from a collection of shaded background images, while a bright color (leftmost color in Fig. 2) was derived from a set of images representing sunlit stone walls. The two remaining colors have a medium brightness. A neutral gray color (second from left in Fig. 2) of medium brightness was included since this occurs most frequently in an urban environment (concrete, stones, streets, grayish shadows, etc.). Finally, a mean representative background color of medium brightness (third from left in Fig. 2) was computed from the overall set of images of the test site.

2.2 *TNO Multiscale Pattern*

2.2.1 *Texture selection*

Natural images generally contain elements of all sizes, with larger elements occurring less frequently. The spatial structure of natural images shows the same statistical behavior as fractals (the power in spatial frequency decreases with one over the frequency).[15] Fractals are images in which the appearance essentially doesn't change when zooming in. We choose to use a similar texture, with a fractal nature. The advantage of such a texture is that it contains elements of various sizes. Thus, for each background there will be pattern elements that match the size of the background elements. The apparent size of relevant details in a scene with camouflage depends on the distance from the observer to both the camouflage pattern itself and to its local background pattern. For instance, when a camouflaged person stands right in front of a brick wall, the best camouflage pattern will be one that contains brick-like elements of the same size as the bricks in the wall. However, when the brick wall is farther away in the distance, the brick-like elements in the camouflage pattern should be smaller to match the angular size of the bricks in the background. By choosing a fractal structure, we ensure that the camouflage pattern will always contain elements that match the size of the background details.

Most existing camouflage patterns contain a dominant direction (i.e., contain elongated elements). For scenes with unrestricted viewing distance, this may have some value, since the perspective effect transforms circular patches on the ground into ellipses. However, in urban environments viewing distances are severely restricted by many objects that block the view (e.g., cars, trees, houses). Furthermore, older camouflage patterns have been designed for use in a fixed orientation (i.e., an upright position of the person wearing the camouflage pattern). Such directional patterns will no longer match the background when they are used in a different orientation (e.g., when a person lies down on the ground). Therefore, we chose to use a pattern without a dominant direction. To this end, we constructed a colored fractal texture pattern with a power spectrum that falls off with one over the spatial frequency ($1/f$) in both the hue, saturation

Fig. 2 The 11 different suits that were evaluated in this study. Suit number 9 is the TNO design.

and value (HSV) signals [i.e., in perceptually de-correlated color space; see Fig. 3(b)]. This was done by filling each channel with a Gaussian white noise image that had been filtered in the Fourier domain to produce an image with a $1/f$ natural slope amplitude spectrum.

2.2.2 *Color selection*

A collage was constructed of background samples that were characteristic for the test site [Fig. 3(a)], and its color distribution was transferred to a $1/f$ fractal color texture pattern as follows. First, both the background sample collage [Fig. 3(a)] and the fractal texture [Fig. 3(b)] were transformed to indexed images using 16 colors and standard indexing techniques.* The 16 entries in the color map of the collage and the colored fractal image (shown as insets on the right sides of these images in Fig. 3) were ordered in accordance of occurrence. Next, the entries in the color lookup table of the fractal image were replaced by those belonging to the collage (i.e., their respective color tables were swapped), to create a fractal texture with a color distribution that is similar to that of the collage. In the last step, this pattern was approximated by a pattern containing the four colors derived by NATO SCI-157 (Fig. 2). This was done by first representing the fractal texture image (now containing 16 colors) as an RGB image, and then transforming it to an indexed image representation (using standard dithering techniques) containing only the four prescribed colors.

2.3 *NATO SCI-157 Patterns*

Members of the NATO SCI-157 group designed different camouflage patterns for a comparative evaluation test in the Fort Huachuca urban area. These camouflage patterns were printed on cloth that was then used to sew suits. Figure 4 shows the 11 uniforms that were used in the test. Suit 1 is a plain beige suit and suit 11 is an in-service (semi-arid) suit designed for deployment in arid conditions. Both of these suits were included for reference purposes. The rationale for including a plain (monotone, sandy) suit in the test is that all patterns look monochrome when seen from a sufficient distance, and brightness is the only distinctive cue. Suits 2 through 10 were all designed by members of SCI-157 and consist of the same four dominant colors that were derived from imagery of the Fort Huachuca area and from on-site color measurements (Fig. 2). Suits 2 and 3 were designed to explore the prospect of shape disruption. Suits 4 through 10 are adorned with micropatterns that were synthesized with different techniques. Suite 9 is adorned with the TNO test pattern that was constructed according to the procedure described in Sec. 2.2 [Fig. 3(d)]. Suits 4 and 10 were designed to have, respectively a relatively high and low overall luminance value.

2.4 *Test Imagery*

Panoramic digital color images were registered of a mannequin wearing each of the 11 different camouflage suits (see Sec. 2.3) on each of 36 different selected locations in the Fort Huachuca background (target scenes: Fig. 5). The same locations were also registered without a person present (empty scenes).

3 Visual Search Experiment

We performed a visual search experiment using digital images of the target and empty scenes registered at all 36 different locations in the Fort Huachuca background. A total of 11 subjects participated in the experiment. A PC was used to control the presentation of the stimuli and to register the response of the participants. The setup was placed in a dimly lit room. Participants were comfortable seated at a distance of 40 cm behind a PC monitor (a 30-in. LCD screen, 75 Hz). The images were presented with resolution of 2560×700 pixels. In each trial, either a target or an empty scene was shown. The participant's

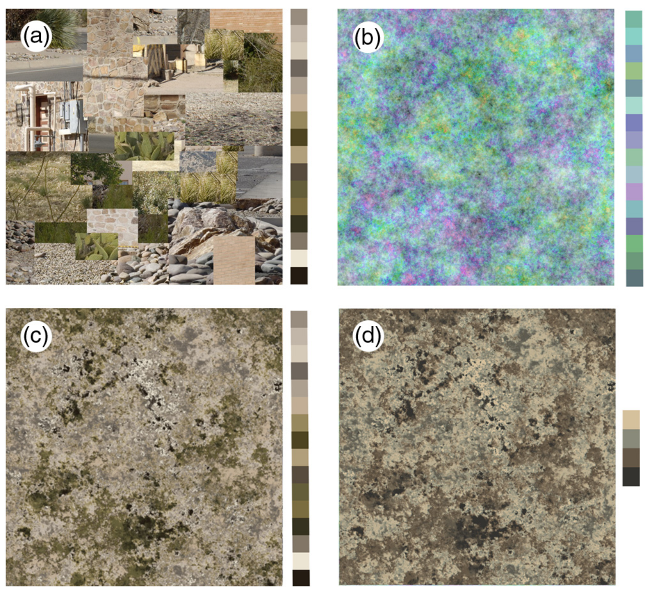

Fig. 3 (a) Representative set of background samples with the 16 primary colors shown on the right. (b) Fractal color texture. (c) Pattern obtained by swapping the color tables of (a) and (b). (d) Dithered representation of (c) using only 4 colors.

*The pixels in an indexed image do not contain RGB values but entries (values ranging from 0 to 15 in this case) to a color lookup table containing RGB values.

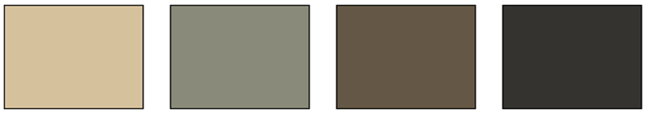

Fig. 4 The set of 4 colors used in the construction of all camouflage patterns investigated in this study.

task was first to decide whether the image contained a target and second, to indicate the target location as quickly as possible after the onset of the image presentation by clicking the mouse on the target as soon as the target was detected. When the participant did not find a target, he/she could indicate that no target was present by clicking on a square in the bottom-right corner of the screen. The computer registered the response time and the position of the mouse cursor on the screen (indicating either the location of the target or a "no target present" response) when the participant clicked the mouse button in response. Each subject participated in one session in which 36 target scenes and 36 empty scenes were shown. The subject was informed that there was a 50% chance that an image contained a target. Corresponding target and empty scenes (i.e., images of the same location either with or without a target) were shown mirror-reversed, to prevent possible recognition of the scene. The 396 target-location scenes (11 camouflage patterns × 36 locations) were randomly divided into 11 sets of 36 scenes, such that each set showed each location only once. Each of the 11 observers viewed one set in a single session.

4 Frequency-Tuned Saliency

Human visual fixation behavior is driven both by sensorial bottom-up mechanisms[16,17] and by higher order task-specific or goal-directed top-down mechanisms.[18,19] Visual saliency refers to the physical, bottom-up distinctness of image details.[20] It is a relative property that depends on the degree to which a detail is visually distinct from its background.[21] Visual saliency is believed to drive human fixation behavior during free viewing by attracting visual attention in a bottom-up way.[22] As such, it is an important factor in our everyday functioning. Human observer studies have indeed shown that saliency can be a strong predictor of attention and gaze allocation during free viewing, both for static scenes[23–25] and for dynamic scenes.[26] Moreover, saliency also appears to determine which details humans find interesting in visual scenes.[27]

Based on the notion that being a local outlier makes a point salient, Koch and Ullman[17] introduced the concept of a saliency map, which is a two-dimensional topographic representation of saliency for each pixel in an image. Over the past decade, many different algorithms have been proposed to compute visual saliency maps from digital imagery.[17,28–47] These algorithms typically transform a given input image into a scalar-valued map in which local signal

Fig. 5 The 36 different target locations, each showing a mannequin dressed in one of the 11 different camouflage suits as an example.

intensity corresponds to local image saliency.[33,45] In an extensive comparative evaluation study[48] we recently established that the maximum value over the target support computed by Achanta's Frequency Tuned Saliency model[45] is currently the best saliency-based predictor of human visual search and detection performance in complex realistic scenarios.

Achanta et al.[45] compute bottom-up image saliency as local multiscale color and luminance feature contrast. The underlying hypothesis of this approach is that human fixation is driven by local center-surround feature contrast.

First, the input sRGB image I is transformed to CIE Lab color space. Then, the scalar frequency-tuned saliency (FTS) map S for image I is computed as

$$S(x, y) = \|\boldsymbol{I}_{\mu} - \boldsymbol{I}_{\omega_{hc}}\|, \tag{1}$$

where $\boldsymbol{I}_{\mu}$ is the arithmetic mean image feature vector; $\boldsymbol{I}_{\omega_{hc}}$ equals a Gaussian blurred version of the original image, using a 5×5 separable binomial kernel; $\|\|$ equals the L_2 norm (Euclidian distance); and x, y represent the pixel coordinates.

Blurring with a 5×5 Gaussian kernel serves to eliminate noise and fine texture details from the original image, while still retaining a sufficient amount of high frequency details. We use $\frac{1}{16}[1, 4, 6, 4, 1]$ giving $\omega_{hc} = \pi/2.75$. The mean image feature vector corresponds to blurring with a Gaussian of infinite extent. The difference $\boldsymbol{I}_{\mu} - \boldsymbol{I}_{\omega_{hc}}$ effectively represents the output of a range of bandpass (DoG) filters at several image scales. Since the norm of the difference is used, only the magnitude of the local differences contributes to the saliency of the image detail.

We computed the maximum of frequency-tuned saliency value over the target support for each of the target scenes used in this study as follows. First we constructed target masks (binary images representing the target support area) for all target scenes (i.e., all 11 different camouflage suits on each of the 36 different locations in the Fort Huachuca background), by manually segmenting the images using Photoshop CS5. Then we applied Eq. (1) to compute a saliency map for each target scene, and we computed the maximal saliency value over the target support using the corresponding binary mask. Finally, we computed the mean of the maximal saliency value over all 36 tested locations for each of the 11 patterns.

5 Results

5.1 *Human Observer Performance*

The results of the visual search experiment enable the quantification of camouflage performance in terms of search time (mean time required to find a target) and detection probability (fraction of trials in which the target is actually found). Figure 6 presents the detection probability and the mean search time for the 11 camouflage patterns tested in this study. The overall false alarm rate was quite low (about 7%).

Effective camouflage performance should combine long search times with a low detection probability (lower right hand corner in Fig. 6). Figure 6 shows that patterns 1, 8, 9, and 11 perform well in the sense that they have a low detection probability. Patterns 1, 8 and 11 show good performance in the sense that they combine relatively long search times with relatively low detection probabilities. Pattern 11

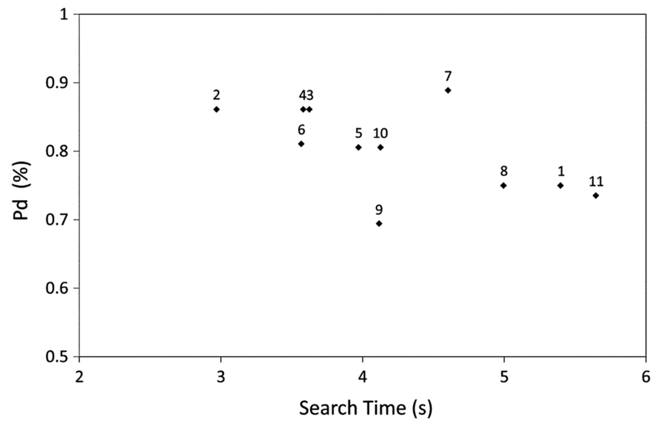

Fig. 6 Detection probability (%; standard error about 0.05) versus mean search time (s; standard error about 0.5) for each of the 11 camouflage patterns tested.

(the in-service semi-arid suit) yields the overall longest search time, while the fractal pattern 9 has the overall lowest detection probability. Pattern 1 (monotone plain beige suit) yields the second longest search time. Patterns 2, 3 and 4 show relatively low performance in the sense that they combine relatively short search times with relatively high detection probabilities.

The overall score of the camouflage patterns will depend on the weight attributed to each of the two performance measures (mean search time and detection probability). However, it is evident that the weight assigned to detection probability should be relatively high. We therefore conclude that TNO pattern 9 performs well since it combines the lowest detection probability with an intermediate search time.

5.2 *Computational Saliency*

Figures 7 and 8 show, respectively the detection probability and the mean search time as a function of the maximum of the FTS saliency value over the target support.

It appears that both detection probability and the mean search time correlate strongly with the maximal FTS saliency value over the target support, with a one-tailed Pearson

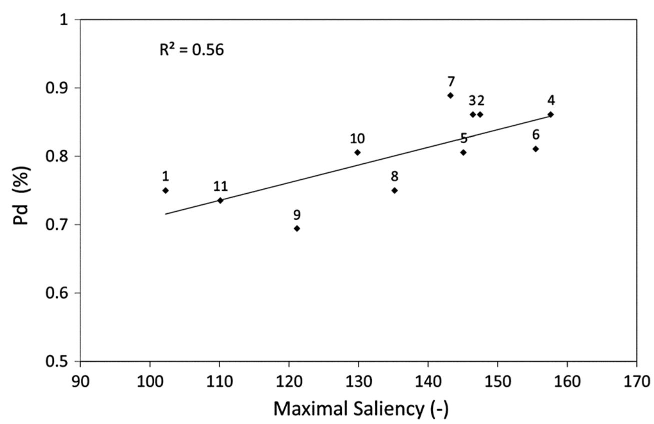

Fig. 7 Detection probability (%) as a function of the maximum of the FTS saliency value (-) over the target support. The line represents the result of a linear best fit to the data ($R^2 = 0.56$).

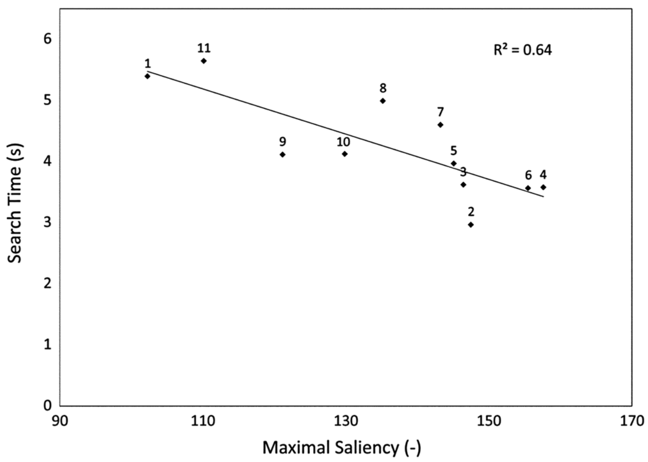

Fig. 8 Mean search time (s) as a function of the maximum of the FTS saliency value (-) over the target support. The line represents the result of a linear best fit to the data ($R^2 = 0.64$).

correlation of $r = 75$ and $r = 80$, respectively (both significant at the 0.01 level). Since $r^2 = 56$ and 64, respectively, both measures explain about 60% of the original variance. Hence, the FTS saliency metric appears to predict human observer performance in a visual search and detection task with camouflaged targets to an appreciable extent.

6 Discussion and Conclusions

We presented the efficient TNO method to design a dedicated multiscale camouflage pattern from a given set of characteristic background image samples, and used this method to construct a camouflage pattern for a semi-arid urban area. We evaluated the camouflage effectiveness of the resulting pattern relative to 10 other specially designed patterns, both through a human visual search experiment and through a computational saliency metric.

The camouflage performance measures resulting from the visual search experiments were the fraction of locations in which the target was detected and the average search time over trials in which the targets were detected. These metrics allowed us to distinguish between good (patterns 1, 8, 9, 11 in Fig. 4), medium (patterns 5, 6, 10) and poorly performing patterns (2, 3, 4, 7). The pattern based on the design method presented here (pattern number 9) combined the overall lowest detection probability with an intermediate search time, indicating that the TNO design method can produce effective camouflage patterns.

Interestingly, some common assumptions regarding characteristics that are essential for urban camouflage patterns were not confirmed in this study. It is often assumed that urban camouflage designs require large macropatterns with horizontal and vertical edges to blend in with an urban environment. However, our present results show that shape disruptive designs with large macropatterns (patterns 2 and 3 in Fig. 4) did not outperform designs with small micropatterns (Fig. 6). Also, the assumption that an effective urban pattern should contain vertical and horizontal elements is not confirmed. This also questions the need for a camouflage pattern customized for a particular urban environment. Instead, our study suggests that general purpose arid patterns may provide good camouflage also in urban arid environments.

It is well known that visual targets that are similar to their local background or to details in other parts of the scene are harder to find than targets that are highly distinct or conspicuous. Hence, visual target conspicuity also depends on the amount of detail and the structural composition of a scene.[49–51] A target is generally less conspicuous when it is placed in an area with more detail or when it is structurally similar to its background. It appears that detection performance depends to a large degree on the energy contrast between a target and its local background, whereas recognition depends mainly on the structural dissimilarity between a target and its surround.[52,53] This obscuring effect, which is generally known as clutter, determines human visual search and detection performance to a large extent. For complex scenes, the spatial relationships (shape and relative location) of features in an image can have a greater effect on detection than the relative luminance of the features.[54] Many attempts have been made to quantify the effects of clutter by means of digital clutter metrics. However, the concept of clutter is inherently elusive, and attempts to model it have only been partly successful.[54–68]

The present results indicate that the simple computational FTS saliency metric predicts the relative effectiveness of different visual camouflage designs to a large degree. However, this metric only incorporates the energy contrast between a target and its local background, and therefore only models human detection performance. An effective camouflage evaluation metric should also include a structural dissimilarity metric to account for human recognition performance (cognitive screening). We therefore attempted to extend the FTS metric with Wang and Bovik's structural image similarity index (SSIM), which measures the similarity between images in terms of luminance, contrast and structure.[53,69] We hypothesized that a larger similarity between the camouflage pattern and its local background would result in a longer search time and a lower detection probability. However, the predictions of the combined metric correlated less with human observer performance than the predictions of the FTS metric alone. We intend to investigate this issue further in a follow-up study. The availability of a validated computational camouflage evaluation metric will enable the automatic design of an optimal camouflage pattern for a given area from a set of characteristic background image samples.

References

1. J. Baumbach, "Colour and pattern composition to blend objects into a natural environment," in *Colour—Effects & Affects. Proc. Interim Meeting Int. Colour Assoc.*, I. Kortbawi, B. Bergström, and K. F. Anter, Eds., Swedish Colour Centre Foundation/Scandinavian Colour Institute AB, Stockholm, Sweden (2008).
2. V. A. Billock, D. W. Cunningham, and B. H. Tsou, "What visual discrimination of fractal textures can tell us about discrimination of camouflaged targets," in *Proc. Human Factors Issues in Combat Identification Workshop*, D. H. Andrews, R. P. Herz, and M. B. Wolf, Eds., pp. 98–111, Ashgate Publishing, Farnham, UK (2008).
3. C.-C. Chang et al., "Visual assessment of camouflaged targets with different background similarities," *Percept. Motor Skill.* **114**(2), 527–541 (2012).
4. J. Venezia and L. Hepfinger, "Using a constructive simulation to assess the survivability of various camouflage patterns," Report Technical Report NATICK/TR-12/004L, U.S. Army Natick Soldier Research, Development and Engineering Center, Natick, Massachusetts (2012).
5. D. Fakin and A. Ojstrsek, "The effect of camouflage shades on clothings' low visibility," in *Colour—Effects & Affects. Proc. Interim Meeting Int. Colour Assoc.*, I. Kortbawi, B. Bergström, and K. F. Anter, Eds., Swedish Colour Centre Foundation/Scandinavian Colour Institute AB, Stockholm, Sweden (2008).

6. R. P. Behrens, "The role of artists in ship camouflage during World War I," *Leonardo* **32**(1), 53–59 (2006).
7. D. Delouche, "Cubisme et camouflage," *Guerres Mondiales et Conflits Contemporains* **171**, 123–137 (1993).
8. R. R. Behrens, "The weave (and warp) of invention," *J. Creat. Behav.* **8**(2), 81–92 (1974).
9. H. B. Cott, *Adaptive Coloration in Animals*, Methuen & Co., London, UK (1940).
10. G. H. Thayer, *Concealing—Coloration in the Animal Kingdom: An Exposition of the Laws of Disguise through Color and Pattern: Being a Summary of Abbott H. Thayer's Discoveries*, The Macmillan Co., New York (1909).
11. A. H. Thayer, "The law which underlies protective coloration," *Auk* **13**(2), 477–482 (1896).
12. T. Troscianko et al., "Camouflage and visual perception," *Philos. Trans. Royal Soc. B* **364**(1516), 449–461 (2009).
13. A. Dugas et al., "Universal camouflage for the Future Warrior," *Proceedings of the 4th International Conference on Safety and Protective Fabrics: Solutions and Opportunities for the Safety and Protective Fabrics Industry*, Industrial Fabrics Association International, Pittsburgh (2004).
14. D. J. Tolhurst, Y. Tadmor, and T. Chao, "Amplitude spectra of natural images," *Ophthal. Physiol. Opt.* **12**(2), 229–232 (1992).
15. NATO SCI-157. "Urban camouflage for the individual soldier system," Report RTO-TR-SCI-157, NATO Science and Technology Organization (2010).
16. S. J. Dickinson et al., "Active object recognition integrating attention and viewpoint control," *Comput. Vis. Image Und.* **67**(3), 239–260 (1997).
17. C. Koch and S. Ullman, "Shifts in selective visual attention: towards the underlying neural circuitry," *Hum. Neurobiol.* **4**(4), 219–227 (1985).
18. M. Land, N. Mennie, and J. Rusted, "The roles of vision and eye movements in the control of activities of daily living," *Perception* **28**(11), 1311–1328 (1999).
19. A. Yarbus, *Eye Movements and Vision*, Plenum Press, New York (1967).
20. J. H. Fecteau and D. P. Munoz, "Salience, relevance, and firing: a priority map for target selection," *Trends Cogn. Sci.* **10**(8), 382–390 (2006).
21. A. H. Wertheim, "Visual conspicuity: a new simple standard, its reliability, validity and applicability," *Ergonomics* **53**(3), 421–442 (2010).
22. E. Niebur and C. Koch, "Computational architectures for attention," in *The Attentive Brain*, R. Parasuraman, Ed., pp. 163–186, MIT Press, Cambridge, Massachusetts (1998).
23. D. Parkhurst, K. Law, and E. Niebur, "Modeling the role of salience in the allocation of overt visual attention," *Vision Res.* **42**(1), 107–123 (2002).
24. B. W. Tatler, R. J. Baddeley, and I. D. Gilchrist, "Visual correlates of fixation selection: effects of scale and time," *Vision Res.* **45**(5), 643–659 (2005).
25. G. Underwood and T. Foulsham, "Visual saliency and semantic incongruency influence eye movements when inspecting pictures," *Q. J. Exp. Psychol.* **59**(11), 1931–1949 (2006).
26. L. Itti, "Quantifying the contribution of low-level saliency to human eye movements in dynamic scenes," *Vis. Cogn.* **12**(6), 1093–1123 (2005).
27. L. Elazary and L. Itti, "Interesting objects are visually salient," *J. Vision* **8**(3), 1–15 (2009).
28. N. D. B. Bruce and J. K. Tsotsos, "Saliency, attention, and visual search: an information theoretic approach," *J. Vision* **9**(3), 1–24 (2009).
29. N. D. B. Bruce and J. K. Tsotsos, "Saliency based on information maximization," in *Advances in Neural Information Processing Systems 18*, Y. Weiss, B. Schölkopf, and J. Platt, Eds., pp. 155–162, MIT Press, Cambridge, Massachusetts (2009).
30. N. D. B. Bruce and J. K. Tsotsos, "An Information Theoretic Model of Saliency and Visual Search," in *Attention in Cognitive Systems. Theories and Systems from an Interdisciplinary Viewpoint: 4th International Workshop on Attention in Cognitive Systems, WAPCV 2007*, L. Paletta and E. Rome Eds., pp. 171–183, Springer-Verlag, Berlin/Heidelberg, Germany (2008).
31. N. D. B. Bruce, "Features that draw visual attention: an information theoretic perspective," *Neurocomputing* **65–66**, 125–133 (2005).
32. C. Kanan et al., "SUN: top-down saliency using natural statistics," *Vis. Cogn.* **17**(6–7), 979–1003 (2009).
33. L. Zhang et al., "SUN: a Bayesian framework for saliency using natural statistics," *J. Vision* **8**(7), 1–20 (2008).
34. M. Mancas et al., "A rarity-based visual attention map-application to texture description," in *Proc. IEEE Int'l. Conf. Image Proc. 2006*, pp. 445–448, IEEE, Atlanta, Georgia (2006).
35. M. Mancas, B. Gosselin, and B. Macq, "Perceptual image representation," *EURASIP J. Image Video Process.* 1–9 (2007).
36. M. Mancas, *Computational Attention: Towards Attentive Computers*, Presses Universitaires de Louvain, Louvain-la-Neuve, Belgium (2007).
37. M. Mancas et al., "Computational attention for event detection," in *Proc. 5th Int'l Conf. Computer Vision Systems*, Applied Computer Science Group, Bielefeld University, Germany (2007).
38. M. Mancas, "Image perception: relative influence of bottom-up and top-down attention," in *Proc. WAPCV Workshop ICVS Conf.*, pp. 94–106 (2008).
39. M. Mancas et al., "Tracking-dependent and interactive video projection," in *Proceedings of the eNTERFACE'08 Workshop on Multimodal Interfaces*, pp. 73–83, The Numediart Institute for New Media Art Technology, University of Mons, Orsay-Paris, France (2009).
40. P. L. Rosin, "A simple method for detecting salient regions," *Pattern Recogn.* **42**(11), 2363–2371 (2009).
41. D. Walther, "Interactions of visual attention and object recognition: computational modeling, algorithms, and psychophysics," PhD Thesis, California Institute of Technology (2006).
42. D. Walther and C. Koch, "Modeling attention to salient proto-objects," *Neural Networks* **19**(9), 1395–1407 (2006).
43. L. Itti, C. Koch, and E. Niebur, "A model of saliency-based visual attention for rapid scene analysis," *IEEE Trans. Pattern Anal. Mach. Intell.* **20**(11), 1254–1259 (1998).
44. P. Bian and L. Zhang, "Biological plausibility of spectral domain approach for spatiotemporal visual saliency," in *Advances in Neuro-Information Processing*, M. Köppen, N. Kasabov, and G. Coghill, Eds., pp. 251–258, Springer, Berlin/Heidelberg, Germany (2009).
45. R. Achanta et al., "Frequency-tuned salient region detection," in *Proc. IEEE Int. Conf. Comput. Vis. Pattern Recognit. June 20–25, 2009*, pp. 1597–1604, IEEE Press, Washington (2009).
46. Y.-F. Ma and H.-J. Zhang, "Contrast-based image attention analysis by using fuzzy growing," in *Proc. Eleventh ACM Intl. Conf. Multimedia*, pp. 374–381, ACM Press, New York (2003).
47. J. Harel, C. Koch, and P. Perona, "Graph-based visual saliency," in *Advances in Neural Information Processing Systems*, B. Schölkopf, J. Platt, and T. Hoffman, Eds., pp. 545–552, MIT Press, Cambridge, Massachusetts (2007).
48. A. Toet, "Structural similarity determines search time and detection probability," *Infrared Phys. Technol.* **53**(6), 464–468 (2010).
49. B. L. Cole and S. E. Jenkins, "The effect of variability of background elements on the conspicuity of objects," *Vision Res.* **24**(3), 261–270 (1984).
50. S. E. Jenkins and B. L. Cole, "The effect of the density of background elements on the conspicuity of objects," *Vision Res.* **22**(10), 1241–1252 (1982).
51. D. G. Pelli and K. A. Tillman, "The uncrowded window of object recognition," *Nat. Neurosci.* **11**(10), 1129–1135 (2008).
52. W. L. Braje, B. S. Tjan, and G. E. Legge, "Human efficiency for recognizing and detecting low-pass filtered objects," *Vision Res.* **35**(21), 2955–2966 (1995).
53. T. Caelli and G. Moraglia, "On the detection of signals embedded in natural scenes," *Percept. Psychophys.* **39**(2), 87–95 (1986).
54. J. M. Cathcart, T. J. Doll, and D. E. Schmieder, "Target detection in urban clutter," *IEEE Trans. Syst. Man Cybern.* **19**(5), 1242–1250 (1989).
55. M. V. Shirvaikar and M. M. Trivedi, "Developing texture-based image clutter measures for object detection," *Opt. Eng.* **31**(12), 2628–2639 (1992).
56. D. E. Schmieder and M. R. Weathersby, "Detection performance in clutter with variable resolution," *IEEE Trans. Aerospace Electron. Syst.* **AES–19**(4), 622–630 (1983).
57. S. R. Rotman, G. Tidhar, and M. L. Kowalczyk, "Clutter metrics for target detection systems," *IEEE Trans. Aerosp. Electr. Syst.* **30**(1), 81–91 (1994).
58. G. Tidhar et al., "Modeling human search and target acquisition performance: IV. detection probability in the cluttered environment," *Opt. Eng.* **33**(3), 801–808 (1994).
59. G. Waldman et al., "A normalized clutter measure for images," *Comput. Vision Graph.* **42**(2), 137–156 (1988).
60. S. R. Rotman et al., "Textural metrics for clutter affecting human target acquisition," *Infrared Phys.Technol.* **37**(6), 667–674 (1996).
61. G. Aviram and S. R. Rotman, "Evaluating human detection performance of targets and false alarms, using a statistical texture image metric," *Opt. Eng.* **39**(8), 2285–2295 (2000).
62. G. Aviram and S. R. Rotman, "Evaluation of human detection performance of targets embedded in natural and enhanced infrared images using image metrics," *Opt. Eng.* **39**(4), 885–896 (2000).
63. G. Aviram and S. R. Rotman, "Evaluating TNO human target detection experimental results agreement with various image metrics," in *Search and Target Acquisition*, A. Toet, Ed., pp. 1–6, North Atlantic Treaty Organization, Neuilly-sur-Seine Cedex, France (2000).
64. R. Rosenholtz et al., "Feature congestion: a measure of display clutter," in *Proc. SIGCHI Conf. Human Factors in Computing Systems*, pp. 761–770, ACM Press, New York (2005).
65. R. Rosenholtz, Y. Li, and T. Nakano, "Measuring visual clutter," *J. Vision* **7**(2), 71–72 (2007).
66. M. J. Bravo and H. Farid, "A scale invariant measure of clutter," *J. Vision* **8**(1), 23-1–23-9 (2008).

67. S. Salem et al., "Rotational clutter metric," *Opt. Eng.* **48**(8), 086401 (2009).
68. R. V. van den Berg, F. W. Cornelissen, and J. B. T. M. Roerdink, "A crowding model of visual clutter," *J. Vision* **9**(4), 24-1–24-11 (2009).
69. A. Toet, "Computational versus psychophysical image saliency: a comparative evaluation study," *IEEE Trans. Pattern Anal. Mach. Intell.* **33**(11), 2131–2146 (2011).

Alexander Toet received his PhD in physics from the University of Utrecht in the Netherlands in 1987, where he worked on visual spatial localization (hyperacuity) and image processing. He is currently a senior research scientist at TNO (Soesterberg, The Netherlands) where he investigates multimodal image fusion, image quality, computational models of human visual search and detection, and the quantification of visual target distinctness. Recently, he started investigating cross-modal perceptual interactions between the visual, auditory, olfactory, and tactile senses with the aim to provide users of serious gaming programs (training and simulation) a more compelling experience. He has published over 60 papers in refereed journals, 70 papers in refereed conference proceedings, was co-editor of a book on the mathematical description of shape in images, was guest editor of a special issue of *Optical Engineering* on "Advances in Target Acquisition Modeling II," and is currently Area Editor of the *ISIF Journal of Advances in Information Fusion*. He organized and directed three international workshops on search and target acquisition, combinatorial algorithms for military applications, and on the mathematical description of shape in images. He is a Fellow of SPIE and a Senior Member of the Institute of Electrical & Electronics Engineers (IEEE).

Maarten A. Hogervorst received his PhD in physics from the University of Utrecht in the Netherlands in 1996 for his work on visual perception. From 1996 to 1999, he worked as a research assistant at Oxford University (UK), where he continued his work in this area. Currently, he is employed at TNO as a research scientist. He has worked on perception of depth, tracking, human combination of multiple sources of visual information, object detection and recognition, workload assessment, and low vision. His current research interests include visual information processing, electro-optical system performance, search and target acquisition modeling, image quality assessment, image enhancement, information fusion, color imaging, EEG and human factors of camera surveillance systems. He has published numerous papers on various topics. He is a Senior Member of SPIE.