In: J.G. Verly (Ed.), *Enhanced and Synthetic Vision 2005*, SPIE-5802-13. In press.. Bellingham, WA., USA: The International Society for Optical Engineering.

Objective Assessment of Simulated Daytime and NVG Image Fidelity

Alexander Toet[a*], Frank L. Kooi[a] , Frido Kuijper[b] , Roland J.M. Smeenk[b]

[a]TNO Human Factors, Kampweg 5, 3769 DE Soesterberg, The Netherlands
[b]TNO Defense, Security and Safety, Oude Waalsdorperweg 63, The Hague, The Netherlands

ABSTRACT

Synthetic imagery used for training and evaluating visual search and detection tasks should result in the same observer performance as obtained in the field. The generation of synthetic imagery generally involves a range of computational approximations and simplifications of the physical processes involved in the image formation, in order to meet the update rates in real-time systems or simply to achieve reasonable computation times. These approximations reduce the fidelity of the resulting imagery. This in turn affects observer performance. We have recently introduced visual conspicuity as an efficient task-related measure that can be deployed to calibrate synthetic imagery for use in human visual search and detection tasks. Target conspicuity determines mean visual search time. Targets in synthetic imagery with the same visual conspicuity as their real world counterparts will give rise to an observer performance in simulated search and detection tasks that is similar to the performance in equivalent real world scenarios. In the present study we compare the conspicuity and the detection ranges of real and simulated targets with different degrees of shading. When ambient occlusion is taken into account, and when the contrast ratios in a scene are calibrated, the detection ranges and conspicuity values of simulated targets are equivalent to those of their real-world counterparts, for different degrees of shading. When no or incorrect shading is applied in the simulation, this is not the case, and the resulting imagery can not be deployed for training visual search and detection tasks.

Keywords: Visual conspicuity, image fidelity, NVG simulation, simulation assessment, ambient occlusion

1. INTRODUCTION

Synthetic imagery is now widely used for a variety of military applications like training and evaluation of camouflage and sensor systems. The generation of synthetic imagery generally involves a range of computational approximations and simplifications of the physical processes involved in the image formation, in order to meet the update rates in realtime systems or simply to achieve reasonable computation times. These approximations may reduce the fidelity of the resulting imagery. Though the resulting imagery may at first sight appear realistic to the untrained eye, it may incorrectly represent reality. For instance, NVG imagery is sometimes simply generated by adding speckle noise to a greenish representation of daytime imagery. In a companion paper[1] we show that this results in a representation that does not contain the limitations and illusory effects that are so characteristic for real NVG imagery.

A simple non-realistic synthetic image may suffice some applications, since the degree of fidelity and the level of detail that is actually required in practice depends on the task for which the synthetic imagery has been generated. However, many military tasks involve human visual search, and detection. Synthetic imagery used for training these tasks should ideally result in the same observer performance that is obtained in the field. This implies that the synthetic imagery should be perceptually calibrated with respect to human visual search and detection. Synthetic image calibration is typically done by performing the same perceptual task both in the lab on the synthetic imagery and in the field, using similar scenes and scenarios in both cases. Typically this results in costly and time consuming experiments, requiring large numbers of observers and resources to obtain statistical significance. To circumvent these problems, we recently

* toet@tm.tno.nl; phone +31-346-356237; fax +31-346-353977; www.tm.tno.nl

introduced visual conspicuity as an efficient task-related measure that can be deployed to calibrate synthetic imagery for use in human visual search and detection tasks.

Synthetic image generation methods like radiosity or ray tracing can produce photorealistic images, but they are time-consuming, require large amounts of memory and are difficult to control for the non-expert. Real-time synthetic image generators therefore use simple models to produce lighting effects that only appear realistic on superficial inspection. More careful inspection reveals that the traditional lighting models used in computer graphics suffer from a lack of visual realism. For instance, Gouraud[2] or Phong[3] shading only take into account direct lighting and solve indirect light coming from other surfaces by applying a single ambient term. This leads to flat lighting of objects in the shadow of direct light. In general, we observed that the visibility of targets in simulated daytime and NVG imagery deviates considerably from their counterparts in corresponding real imagery. In many cases, image simulation software represents shaded targets with too high visibility. Shaded targets that are not visible in the real world may even be clearly visible in simulated imagery. A new method that improves on the traditional ambient term is called "ambient occlusion"[4] and is becoming rapidly more popular in current computer games[5,6]. Ambient occlusion is a lighting technique that gives objects a global illumination-like effect, as if they were lit from the entire hemisphere (rather than a point light). It simulates the amount of indirect light falling on a point on the surface of an object as being proportional to the fraction of the environment that can not be seen from that point (i.e. the fraction of the outward view that is occluded) and that can therefore not contribute to the indirect lighting of that point. Essentially it simulates a huge dome light surrounding the entire scene. As a result, a surface point located for instance under a table or inside a hallway will end up much darker than a surface point on top of the table or outside. Though not entirely physically correct it is a better approximation of the ideal global illumination situation when the entire light transport through a scene is simulated.

The research presented here was done to determine the validity of the "ambient occlusion" approach in the simulation of both daylight and night vision imagery for search and detection training purposes. We do this by investigating target detection ranges and target conspicuity, for targets with different degrees of shading (occlusion), both with the naked eye in daylight conditions and with intensified imagery in nighttime conditions, and in equivalent real and virtual situations. We find that ambient occlusion simulates real-life contrast ratios better than traditional lighting methods, both for daytime and for nightvision conditions. As a result, the corresponding target detection ranges and target conspicuity values resemble more closely the results obtained in reality.

2. VALIDATION EXPERIMENT

2.1 Experimental conditions in the real world

In the real world we determined both the visual conspicuity and the detection range of a target subject positioned in a dark hallway. The subject was a man wearing standard dark green camouflage clothing. The hallway was part of a house that was located in the Dutch tactical training village called "Marnehuizen". This mock-up village is the property of the Royal Netherlands Army (RNLA), and was constructed to train military operations in urban terrain (MOUT), and to test new concepts and materiel (see Figure 1). The houses, stores, basements and phone booths in this village are all practical, and some homes even contain furniture. The particular house that was chosen for this experiment allows an unobstructed view of the front of the house and straight into the hallway from far away, between the other buildings (Figure 2). This enabled us to achieve the large viewing distances that were required to measure the target detection ranges in broad daylight conditions. The subject was located either directly in the front entrance (i.e. aligned with the vertical plane of the front wall of the house, just inside the doorway), or either 1 or 2 meters behind the doorway down the hallway (Figure 3 and 4). These different positions were chosen because they result in different levels of target shading (due to the increasing amount of occlusion the shading on the subject increases when he is farther down the hallway). This scenario also illustrates an important aspect of military operations in urban terrain, namely the fact that soldiers are confronted with a large range of lighting conditions, ranging from broad daylight to dark enclosed spaces.

Figure 1 The Royal Netherlands Army in action in the MOUT training village Marnehuizen

2.2 Experimental conditions in the laboratory

The entire training village of Marnehuizen was simulated and the same location and viewing points were chosen to replicate the real world experiments with synthetic daytime and NVG imagery. The synthetic imagery was produced both with an ambient occlusion shading model and without any applied shading. Application of the ambient occlusion model resulted in target shading that varies gradually with the amount of target occlusion, whereas the absence of shading resulted in targets with constant luminosity, independent of their position in the scene. The observers were seated in a dimly lit room at a distance of 60 cm from a CRT display. At this viewing distance the scene resembled the corresponding real-world scene seen at normal eye height and with magnification 1. The measurements in the real world and in the simulation were performed by the same three subjects.

2.3 Photometry

In the real world the daytime measurements were performed on a partly clouded bright day. The nighttime measurements were done during a night with reduced visibility due to fog. Both in daytime and at night we measured the luminance values of respectively the target, its local background (the dark corridor surrounding the target), the part of the front wall next to the door opening, and a white reflecting paper target placed against the front wall next to the door. This white target served as a reference. Figure 5 shows that the target luminance falls off gradually with the distance down the hallway. This corresponds to the fact that the ambient lighting in the hallway decreases with the square of the distance from the entrance[1]. We used these luminance measurements to calculate the luminance contrast ratios of respectively the target, its immediate background (the dark corridor), and the wall relative to the white paper target.

Both the daytime and the NVG views of the real world scenario were simulated, to allow a complete replication of the real world measurements in the simulated environment. The luminance range that can be produced on a CRT screen is of course much smaller than that observed in reality. However, the human visual system is only sensitive to luminance contrast ratios in the scene. Therefore, we adjusted the luminance distribution in the scene so that the contrast ratios (measured from the CRT screen) of the simulated target, corridor and wall to the simulated white paper target resembled most closely the corresponding real-world values. This ensures that the conspicuity values and detection ranges measured on the synthetic imagery should reflect those observed in the real world.

The real-world luminance values that were used to adjust the contrast ratios in the simulation were measured from a distance of 15 m. Because of atmospheric attenuation the actual contrast ratios will be smaller for larger viewing distances. Because we did not take this effect into account the detection ranges obtained in the simulated conditions may be larger than the ones obtained in the real world. The conspicuity measurements were all performed at a (simulated) distance of 15 m, so here this effect plays no role.

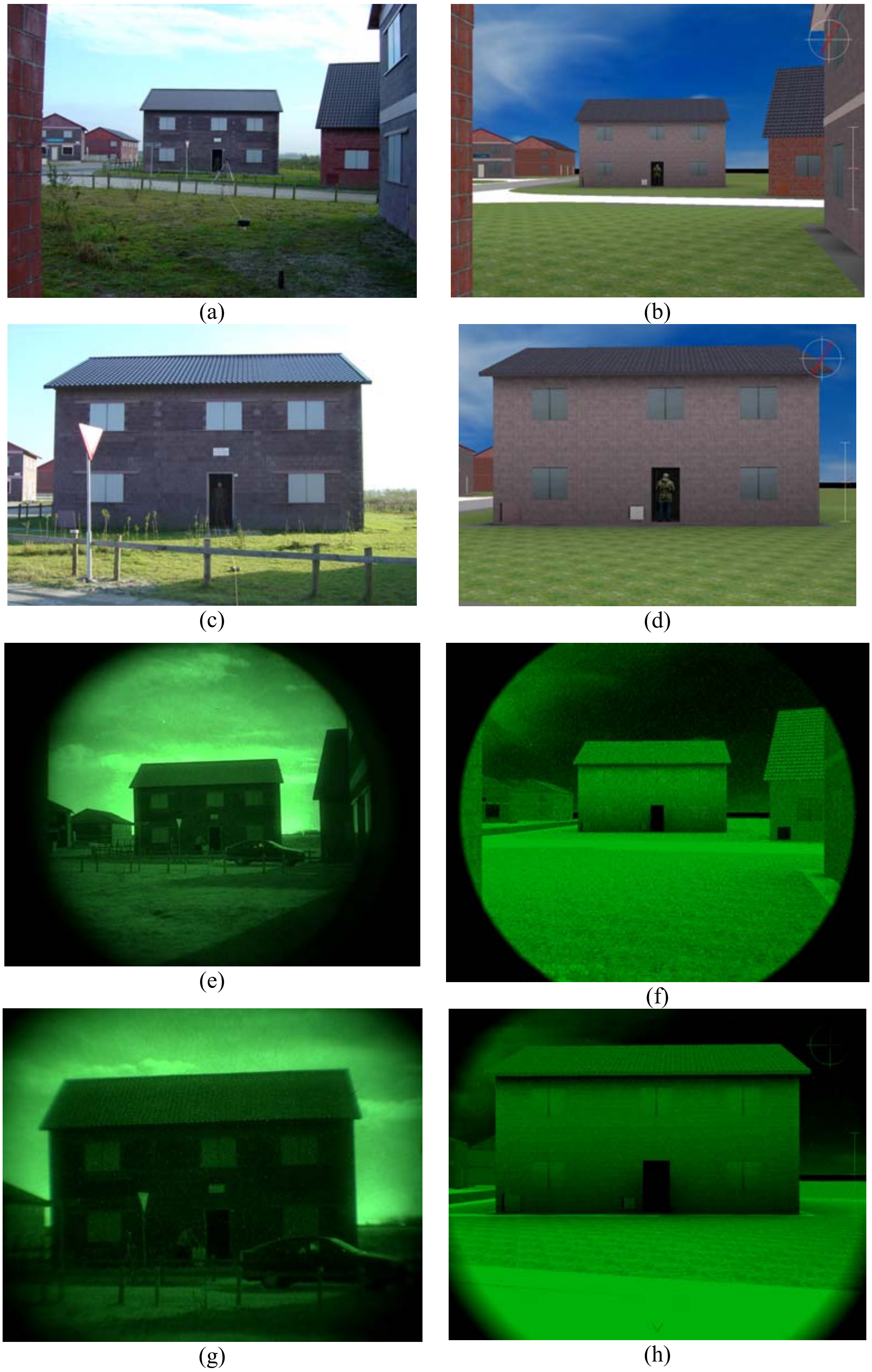

Figure 2 Real (left column) and simulated (right column) daytime (a-d) and nighttime intensified (e-h) imagery of the test location.

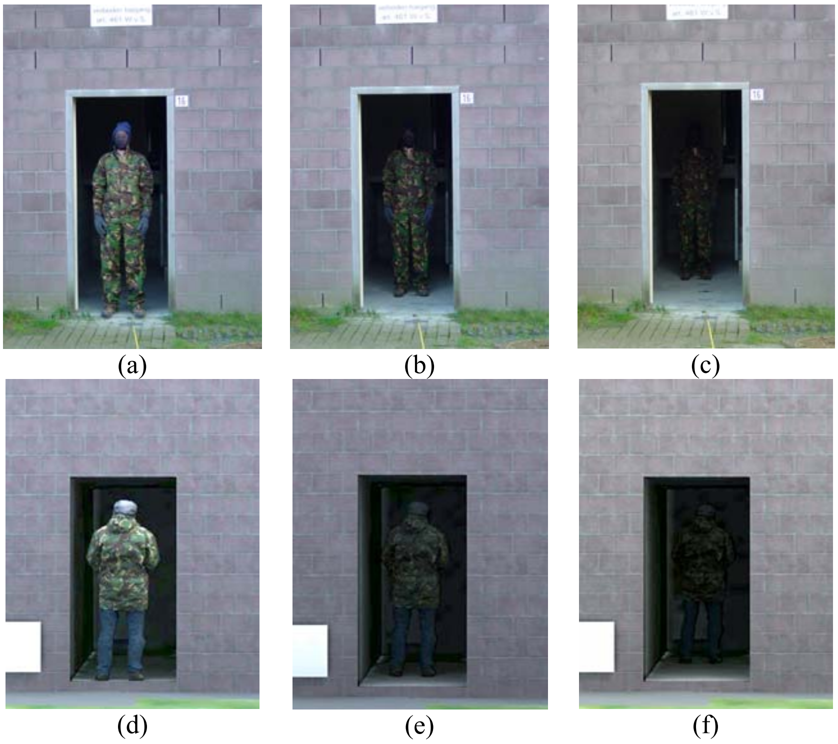

Figure 3 Real (a-c) and simulated (d-f) images of the target subject standing respectively in (a,d), 1 m behind (b,e) and 2 m behind (c,f) the door opening in a dark hallway.

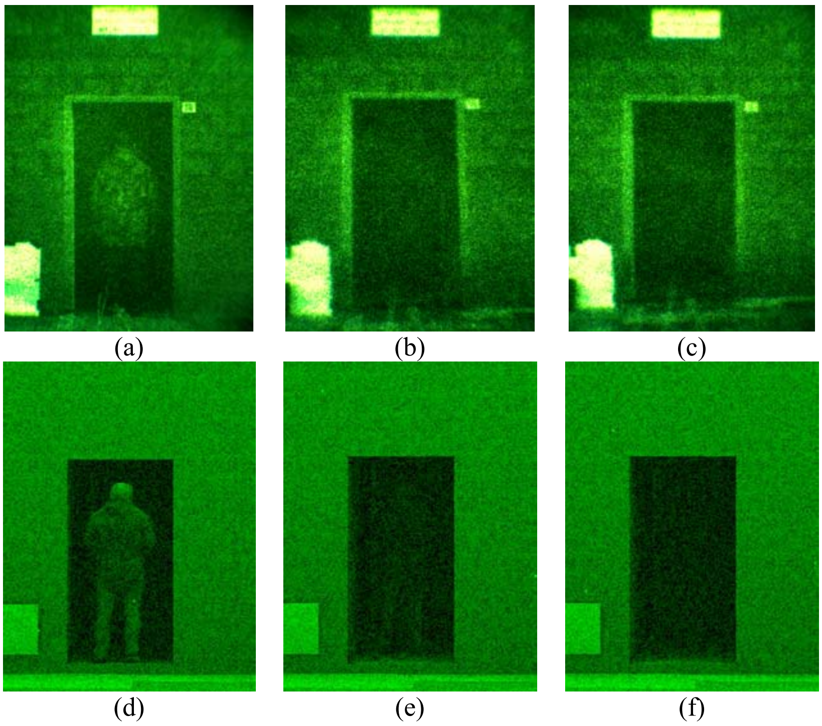

Figure 4 As Figure 3 for intensified night vision images.

2.4 Visual conspicuity

TNO Human Factors developed a psychophysical method to quantify the visual conspicuity of a target in a complex scene[7]. This measure can be obtained quickly and only a few observers are needed to achieve a reasonable accuracy. The measure is directly related to the "conspicuity area", which is defined as the area around the point of fixation within which the target it is perceived as significantly distinct from its local background, and wherein it is therefore capable of attracting visual attention[8-10]. In an extensive series of visual search experiment on realistic imagery it has been shown that the conspicuity measure developed by TNO correlates well with visual search performance[11,12] (as quantified through e.g. mean search time or detection probability). When measuring the conspicuity of objects in real-world scenes the observer first fixates the target and notes its characteristic details. This helps to distinguish the target from its surround during the actual conspicuity measurement. Then, the observer makes a rough estimate of the minimal threshold angle in his peripheral visual field at which the target can no longer be perceived. He then fixates a detail in the scene that is roughly at the same distance as the target. The visual angle between this fixation spot and the target should exceed the threshold visual angle. Then, the observer slowly moves his fixation in the direction of the target until he first notices the target. The angle between the target and the fixation point at which the target is first noticed determines the target conspicuity at the given viewing distance. The threshold viewing angle and the viewing distance determine the extent of the conspicuity area. This extent is defined as the distance between the target and the fixation point measured in the frontoparallel plane (i.e. the plane orthogonal to the viewing direction when the observer fixates the target) through the target. It has been shown that this extent is invariant for variations in viewing distance (for viewing distances that are not so small that the target covers a large area of the visual field or so large that the target can no longer be distinguished in the foveal visual field). Hence, the peripheral angle at which the target can be noticed will generally decrease with increasing viewing distance. The same procedure can be applied when measuring the conspicuity of objects in synthetic scenes, displayed on a CRT or a projection screen. In this case the frontoparallel plane is equal to the projection plane (i.e. the screen). From the conspicuity measured on the screen and the geometry of the projection one can then compute the conspicuity value corresponding to the real world situation.

2.5 Visual detection range

A well known metric for target visibility is its detection range, defined as the largest distance at which the target can still be seen. The detection ranges were measured as follows. In the real world, the observer walked in a straight line towards the entrance of the hallway, starting from far away at a distance at which the target could not be detected (was subthreshold). The subject stopped approaching the target as soon as he could detect it. He then measured the distance to the front of the house using a Leica Vector-1500 DAES pair of binoculars with a built-in laser range finder. In the simulation, the corresponding scene was presented on a CRT screen, and the observer could "approach" the target by using a joystick. Similar to the real world experiment, the subject stopped the virtual approach when he could detect the target, and read out the distance between the viewpoint and the target from the simulation parameters.

2.6 Results

Figure 5 shows the results of the luminance measurements of the human target standing at respectively 0, 1 and 2 m inside the dark hallway in real and simulated daytime conditions (see Figure 3). In the simulated condition the luminance values were registered from the screen of the CRT. This figure shows that although the simulated luminance values differ from the actual ones in magnitude, the ambient occlusion model correctly simulates the luminance variation with the degree of occlusion (depth in the hallway).

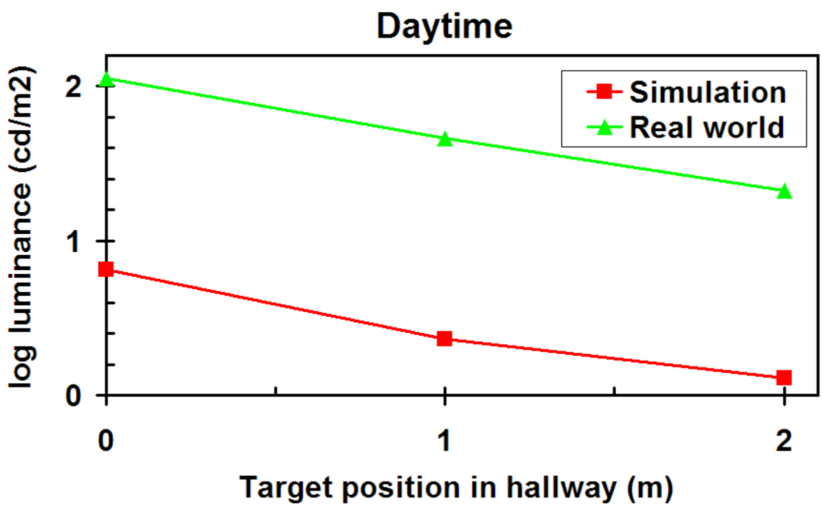

Figure 5 The luminance of the human target standing at respectively 0, 1 and 2 m inside a dark hallway (see Figure 3) in real and simulated daytime conditions.

Figure 6 shows the target detection ranges (m) in the real and simulated daytime conditions, for viewing with the naked eye. This result shows that, when ambient occlusion is taken into account, and when the contrast ratios in the scene are calibrated, the detection range in the simulation decreases with increasing target shading (i.e. with the target distance down the hallway), similar to the real-world situation. When no shading is applied in the simulation, the target luminance is independent of its distance down the hallway, and its detectability is therefore constant (represented by the upper broken line in Figure 6). The detection ranges obtained with the simulated imagery are somewhat larger than those obtained in reality because the simulated contrasts are based on the measurements made in the real world at 15m distance and therefore do not include atmospheric attenuation and ocular scattering.

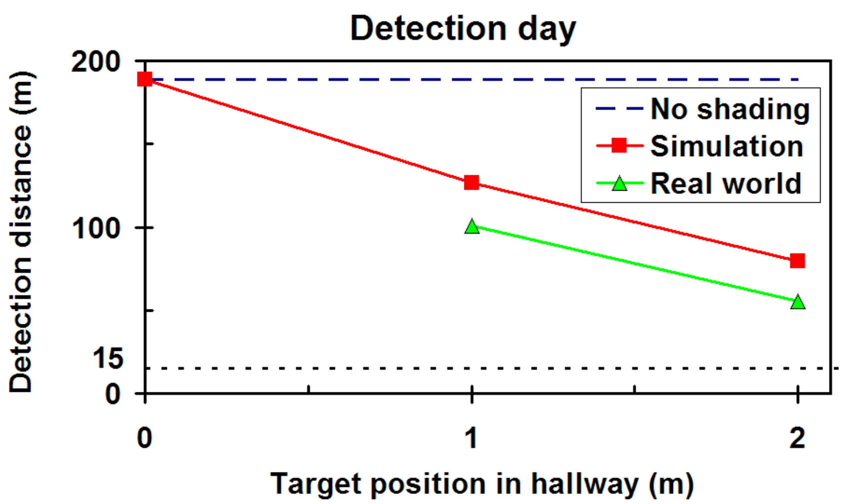

Figure 6 Target detection ranges (m) in real and simulated daytime conditions, for viewing with the naked eye. The target was a camouflaged person standing at respectively 0, 1 and 2 m inside a dark hallway (see Figure 3). The contrast ratios in the simulated scene were calibrated with photometric measurements that were performed at a distance of 15 m (represented by the dotted line) in the corresponding real world scene. The upper broken line represents the fact that a constant detection distance is obtained when no shading is applied.

Figure 7 shows the target detection ranges (m) using an NVG in the field and in simulated NVG viewing conditions. These results are similar to those obtained in the daytime viewing conditions. Again, we find that the gradual decrease of the target detection range with increasing target shading obtained for the simulated imagery is similar to that observed in the real world when ambient occlusion is applied and when the contrast ratios in the synthetic scene are calibrated with those of the corresponding real-world conditions. In the absence of shading the target visibility is independent of its position down the hallway, and a consequence its corresponding detection range is also constant.

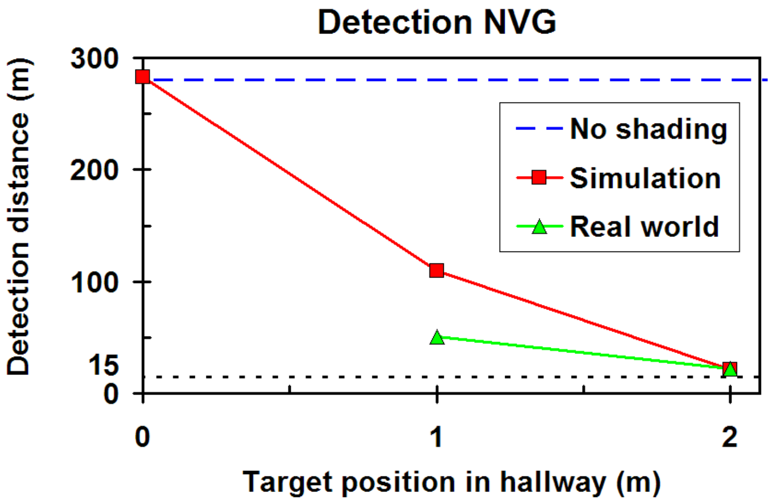

Figure 7 Target detection ranges (m) using an NVG in the field and in simulated NVG viewing conditions. The target was a camouflaged person standing at respectively 0, 1 and 2 m down a dark hallway (see Figure 4). The contrast ratios in the simulated scene were calibrated with photometric measurements that were performed at a distance of 15 m (represented by the dotted line) in the corresponding real world scene. The upper broken line represents the fact that a constant detection distance is obtained when no shading is applied in the simulation.

Figure 8 shows that visual target conspicuity (m) in simulated daytime conditions decreases with increasing target shading in a similar way to that observed in the real world when ambient occlusion is taken into account, and when the contrast ratios in the scene are calibrated. Conspicuity is independent of the position in the hallway when ambient occlusion is not taken into account.

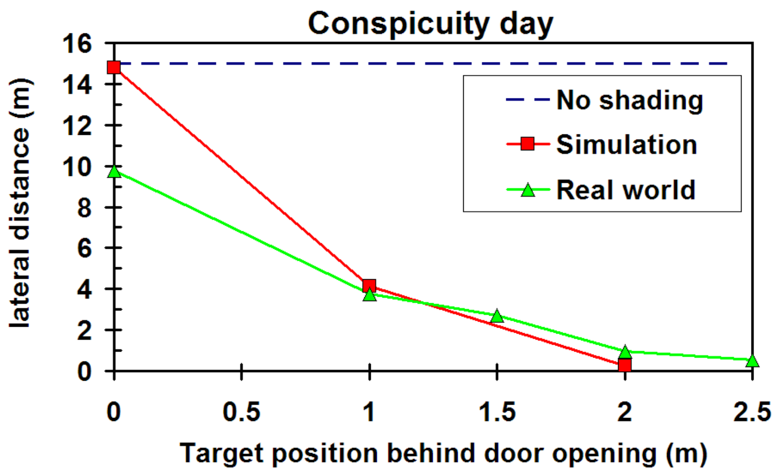

Figure 8 Visual target conspicuity (m) in real and simulated daytime conditions, for viewing with the naked eye. The target was a camouflaged person standing at respectively 0, 1 and 2 m down a dark hallway (see Figure 3). The upper broken line represents the fact that a constant conspicuity is obtained when no shading is applied in the simulation.

Figure 9 shows visual target conspicuity (m) using an NVG in the field and in simulated NVG viewing conditions. The results are similar to those obtained in daytime conditions and for viewing with the naked eye. Note that the fog that occurred near the end of the experiment produced an overall reduction of target visibility, but did not alter the way in which the target conspicuity decreases with increasing target shading.

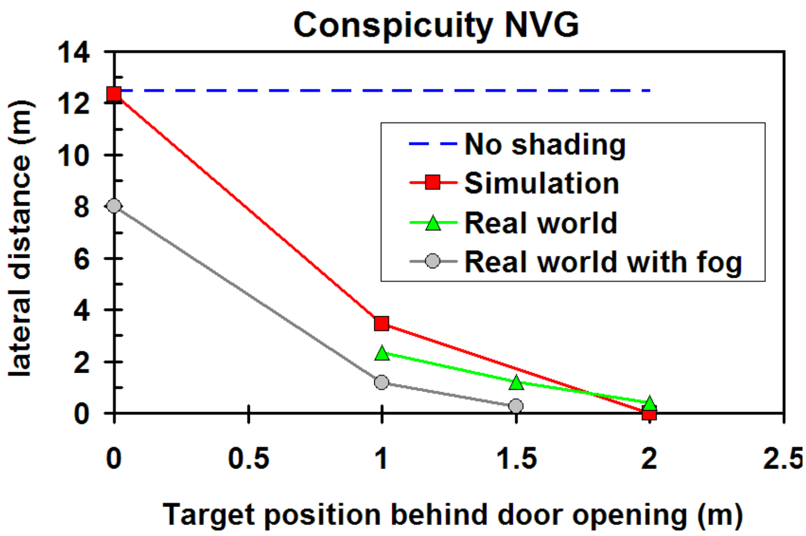

Figure 9 Visual target conspicuity (m) using an NVG in the field and in simulated NVG viewing conditions. The target was a camouflaged person standing at respectively 0, 1 and 2 m down a dark hallway (see Figure 4). The upper broken line represents the fact that a constant conspicuity is obtained when no shading is applied in the simulation. At the end of the experiment a fog occured, which resulted in an overall reduction of target conspicuity.

3. CONCLUSIONS

Our present results show that the detection ranges and conspicuity values of simulated targets are equivalent to those of their real-world counterparts, for different degrees of shading, when ambient occlusion is taken into account, and when the contrast ratios in a scene are calibrated. When shading is not or incorrectly applied in the simulation, the detection ranges and conspicuity values of targets do not correspond to the values of their real-world counterparts, and the imagery is not suitable for training visual search and detection tasks because of the large differences in the observability of shaded targets.

REFERENCES

1. Kooi, F.L. and Toet, A., What's crucial in night vision goggle simulation ?, In: J.G. Verly (Ed.), *Enhanced and Synthetic Vision 2005*, SPIE-5802-4 (this volume), The International Society for Optical Engineering, Bellingham, WA., USA, 2005.
2. Gouraud, H., Continuous shading of curved surfaces, *IEEE Transactions on Computers,* 20*(6),* pp. 623-629, 1971.
3. Phong, B.T., Illumination for computer generated pictures, *Communications of the ACM,* 18*(6),* pp. 311-317, 1975.
4. Landis, H., Production ready global illumination, In: *Course notes for SIGGRAPH 2002 Course 16, RenderMan in Production*, Chapter 5, pp. 87-101,2002.
5. Iones, A., Krupkin, A., Sbert, M. and Zhukov, S., Fast, realistic lighting for video games, *IEEE Computer Graphics and Applications,* 23*(3),* pp. 54-64, 2003.
6. Zhukov, S., Iones, A. and Kronin, G., An ambient light illumination model, In: *Rendering Techniques - Proceedings of the Eurographics Rendering Workshop*, pp. 45-55, Springer Verlag, Vienna, Austria, 1998.
7. Toet, A. and Kooi, F.L., Conspicuity: an efficient alternative for search time, In: A.G. Gale, I.D. Brown, C.M. Haslegrave & S.P. Taylor (Ed.), *Vision in Vehicles*, VII, pp. 451-462, Elsevier Science Ltd, Oxford, UK, 1999.
8. Engel, F.L., Visual conspicuity, visual search and fixation tendencies of the eye, *Vision Research,* 17*,* pp. 95-100, 1977.
9. Engel, F.L., Visual conspicuity and selective background interference in eccentric vision, *Vision Research,* 14*,* pp. 459-471, 1974.
10. Engel, F.L., Visual conspicuity. Directed attention and retinal locus, *Vision Research,* 11*,* pp. 563-575, 1971.
11. Toet, A. and Bijl, P., Visual conspicuity, In: R.G. Driggers (Ed.), *Encyclopedia of optical engineering*, pp. 2929-2935, Marcel Dekker Inc., New York, USA, 2003.
12. Toet, A., Kooi, F.L., Bijl, P. and Valeton, J.M., Visual conspicuity determines human target acquisition performance, *Optical Engineering,* 37*(7),* pp. 1969-1975, 1998.