Physical Modeling and Characterization of the Halo Phenomenon in Night Vision Goggles

Paul J. Thomas[2], Robert S. Allison[1], Peter Carr[1], Edward Shen[1], Sion Jennings[3], Todd Macuda[3], Greg Craig[3], Richard Hornsey[1]

[1]York University, Department of Computer Science, Toronto, Canada
[2]Topaz Technology Inc., Toronto, Canada
[3]National Research Council of Canada, Institute for Aerospace Research, Ottawa, Canada

ABSTRACT

When a bright light source is viewed through Night Vision Goggles (NVG), the image of the source can appear enveloped in a "halo" that is much larger than the "weak-signal" point spread function of the NVG. The halo phenomenon was investigated in order to produce an accurate model of NVG performance for use in psychophysical experiments. Halos were created and measured under controlled laboratory conditions using representative Generation III NVGs. To quantitatively measure halo characteristics, the NVG eyepiece was replaced by a CMOS imager. Halo size and intensity were determined from camera images as functions of point-source intensity and ambient scene illumination. Halo images were captured over a wide range of source radiances (7 orders of magnitude) and then processed with standard analysis tools to yield spot characteristics. The spot characteristics were analyzed to verify our proposed parametric model of NVG halo event formation. The model considered the potential effects of many subsystems of the NVG in the generation of halo: objective lens, photocathode, image intensifier, fluorescent screen and image guide. A description of the halo effects and the model parameters are contained in this work, along with a qualitative rationale for some of the parameter choices.

Keywords: Night Vision Goggles, halo, Gaussian model.

1. INTRODUCTION

Night Vision Goggles (NVG)[1,2,3,4] are a "system that improves night vision by adapting the scene to the conditions for optimum perception by the eye"[5]. The ability of the NVG to accurately re-create the scene, as seen by a human observer, becomes important especially for tasks that require judging distances to and sizes of an object. For example, pilots depend heavily on visual feedback when landing. An artifact, which can impair this ability, generated due to the structure of NVGs (in this work, the baseline NVG is a representative "Generation III" device[5]) that affects user perception is "haloing". Haloing occurs when a light source of small angular extent is viewed through the NVG, by a human observer, and the apparent angular size grows with source radiance. Since an NVG can accommodate a large signal dynamic range, spanning 7 orders of magnitude or more, and typical scenes often have a large dynamic range, haloing can become severe.

As part of ongoing experiments to create an empirical model of NVG performance (for use in research into human performance using NVGs), a "halo event" is measured by removing the eyepiece lens of the NVG and re-imaging the output of the optical image guide onto a detector array. A "white light" LED was used to create a point light source that generated halos. Irradiance at the NVG was varied by up to 7 orders of magnitude with a combination of neutral density filters and the duty-cycle modulation of the LED source. This large dynamic range allowed the image profile to be measured as it changed from the "weak-signal" to the "large signal" regime. The background irradiance was controlled in order to assess the role of the automatic gain control (AGC) of the NVG. The halo event data was processed using Matlab scripts to extract profiles and "thresholded" spot widths. The exact nature of the data will be discussed in a later section. From this analysis, our proposed NVG halo event formation model was confirmed.

The paper is organized as follows. Section 2 explains the halo phenomenon and the importance of the halfwidth to the determination of the "spot width". The experimental setup to measure the spot size of an event and the analysis of the

collected data is described in Section 3. Sections 4 and 5 give details about automatic gain control (AGC), and reflections and scattering effects – secondary effects in halo formation.

2. HALO CONCEPT

The radial distribution of radiance across the image of a small optical source, as seen at the output of the NVG, was modeled by a Gaussian spatial distribution[6,7] on the surface of the image guide of the NVG. Because the primary interest was in halo spots much larger than the point spread function (PSF) of the NVG, a detailed model of the PSF is not required here. Instead, the measured spot radial profile for low source radiance was used as the starting point for analysis. For a suitably small source, this spot profile is expected to approach the PSF, when camera effects are accounted for.
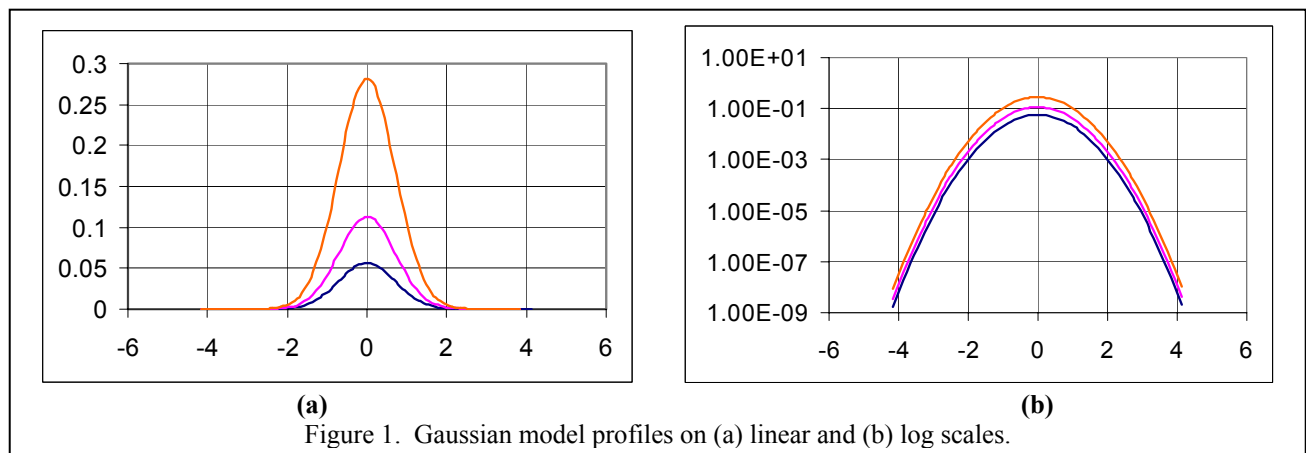
Equation (1) defines a Gaussian radial- intensity distribution of average count rate, $N(x)$, as a function of position, x, on the output of the image guide of the NVG. In azimuth (angle about centre of the spot), the halo is assumed to be symmetrical. Curvature of the output surface of the image guide was ignored. The centre, μ, is the geometrical location of the halo spot and can be mapped back to a corresponding position of the optical image on the photocathode of the NVG.

$$N(x) = \frac{N_{int}}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) = N_{peak} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

(1)

$$\sigma = \sigma_0 + \sigma_1$$

The halfwidth, σ, of the Gaussian controls the width of the point spread function. The equation is for the average count rate, integrated over a time period that is a multiple of the period of the pulsed light source. Asynchrony between the light source and the camera can introduce some amplitude uncertainty in the measured distribution. The expression for N is normalized such that the integral over all x is unity when $N_{int}=1$. Because variations in σ are of interest, a difference term, σ_1, is added to highlight changes in σ.

Figure 1(a) shows profiles of a Gaussian as the integrated signal, N_{int}, is varied, based on Equation 1. The x-axis is the radial position relative to the spot centre and the ordinate axis is the signal, $N(x)$. Note that the full width at half maximum, FWHM=2σ (approximately between -1 and 1 on the figure), is constant for each profile. The intersection points of a horizontal line, $N = 0.05$ for example, with the profile define the "thresholded" spot width, hereafter referred to as the "spot width", with the horizontal line representing the viewing threshold. In Figure 1(b), it can be seen that the spot width grows logarithmically with N_{int}.

(a)   (b)

Figure 1. Gaussian model profiles on (a) linear and (b) log scales.

Quantitatively, the spot width (TSpotWidth) for a Gaussian can be expressed as shown in Equation (2). The key features are the root(log) dependence of width on relative irradiance (i.e. N_{int}) at high irradiance levels, the scaling with threshold and the linear variation with small changes in halfwidth. That is, the spot width is insensitive to changes in source radiance, especially at high irradiance levels, and is proportional to the halfwidth – i.e. is sensitive to changes in halfwidth. At all values of irradiance of interest for halo, the ratio N_{peak}/N_{thresh} is much larger than unity.

$$TSpotWidth \cong 2\sqrt{2}\sigma\sqrt{\ln(\frac{N_{peak}}{N_{thresh}})} \qquad (2)$$

For sufficiently large irradiance, the spot width becomes less dependent on the threshold (N_{thresh}), and the halfwidth may show a dependence on irradiance. Equation (3) is a manipulation of Equations (1) and (2) to emphasize this regime. In particular, TSpotWidth is proportional to σ and deviations of experimental data from the root(log) function can be interpreted as changes in σ.

$$\ln(N_{thresh}) = \left[+\ln(N_{int}) - \ln(\sigma\sqrt{2\pi}) - (\frac{(TSpotWidth/2)^2}{2\sigma^2}) \right]$$

or :

$$TSpotWidth \cong 2\sqrt{2}\sigma_0\left(1+\frac{\sigma_1}{\sigma_0}\right)\sqrt{\ln(N_{int})}\sqrt{1-\frac{\ln\left(N_{thresh}\sigma\sqrt{2\pi}\right)}{\ln(N_{int})}}$$

$$(3)$$

Because TSpotWidth at large N_{int} is insensitive to changes in N_{int}, the TSpotWidth may not change noticeably with range and the size of the halo spot can confuse a visual observer.

3. EXPERIMENTAL SETUP

3.1. Halo profile and width determination

The baseline experimental setup for halo measurement is shown in Figure 2. A LED "point source" and a distributed "background source" illuminate one side of the (binocular) NVG. The second side of the NVG was generally blocked on the input end or used to monitor the scene. With the eyepiece optical element removed, the output surface of the fiber image guide of the NVG was imaged by a CMOS camera linked to a computer. A second camera and source could be used simultaneously on the second side of the NVG, although such measurements are not reported here.
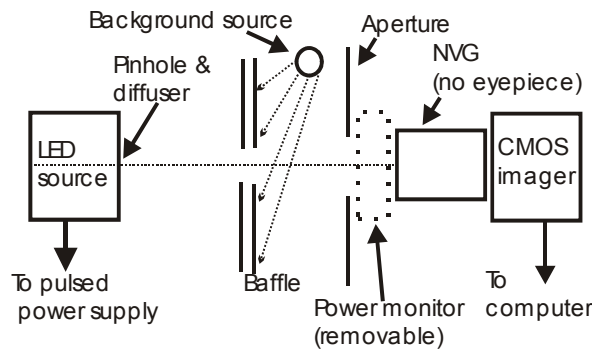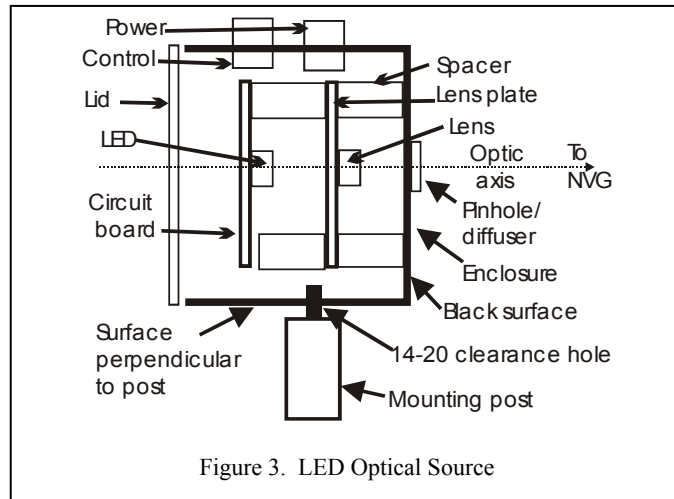

Figure 2. NVG Halo Measurement Setup

The imager used a short-focal-length C-mount lens (4 mm or 8 mm typically) to focus the central region of the output of the fiber image guide (of the NVG, without its eyepiece lens) onto the detector array. Usually, the detector array was a 1200x1080 pixel CMOS colour imager with 7.5 um pixel dimension and a firewire connection to a computer. Changing the integration time (typically about 100 msec) compensated for the limited signal dynamic range (8 bits) of the imager. The optic axis of the imager was approximately collinear with the optic axis of the output surface of the NVG fiber image guide, and the image of the source was within the central 10% of the area of the NVG output plane. This alignment was intended to reduce imager artifacts and reflections in the input optics of the NVG. The angular distribution of light emitted from the output plane of the NVG was broad enough to fill the lens of the imager of Figure 2.

Figure 3. LED Optical Source

The LED optical source (Figure 3) consisted of a pulsewidth-modulated LED whose output was directed onto a diffuser and pinhole. The LED was powered by a commercial pulse generator with 50 ohm output impedance and variable amplitude, period and duty cycle. Changes to the drive current and the modulation duty cycle allowed the average irradiance at the NVG to be conveniently varied over 7 orders of magnitude or more. As applicable, a silicon diode optical power monitor was used to measure the irradiance at the entrance aperture of the NVG. Although the modulation waveform was a square wave, the optical output of the LED had a tail[8] set by its capacitance, C_{led}, and the effective impedance of the drive circuit, R_{drv}. Below about 1 μsec, this tail could reduce the linearity of the calibration of irradiance versus pulse duration. The waveform and duty cycle of the modulation waveform were monitored with an oscilloscope. The uncertainty in modulation width was estimated to be about $\pm$ 5%, limited by the oscilloscope precision. A time-interval meter or equivalent was used when additional timing accuracy was warranted.

In principle, the NVG can tolerate very large signals. However, such signals can alter the performance characteristics in ways that are not well documented in the vendor specifications. The following procedures were used to minimize damage or device 'anomalous behaviour'.

1) each NVG was stored with protective caps on and with the battery pack disconnected;
2) during experiments, the NVG was turned on only when the room was dark;
3) during experiments, data was first acquired at low source intensity, so that any "debug" time was with a low-intensity source;
4) a manual shutter was used to block the source at high intensity levels, except when image data was being collected, so that any exposure was at most only a few seconds in duration;
5) "extremely intense" illumination was avoided, even though this constraint limited the completeness of the data collected.

The background scene generator was similar to the faint optical source with the diffuser removed, the pinhole aperture broadened and a lens added so that the full field of view of the NVG could be illuminated by reflection off a white target plane. Control of the background optical power was generally via a DC voltage, but a pulse generator could also be used as appropriate. The highest irradiance here was approximately 1 nW/cm^2. During use of the background scene generator, the source region was usually configured so that the diffuse background was blocked in a circular region centered on the point source. In this way, the halo properties could be measured in the image without any overlapping background illumination. The illumination arrangement of Figure 2 has this property.

Figure 4 shows a sequence of 4 halo images, with relative source irradiance of 1, 10, 100 and 1000 respectively. The background scene generator was turned off in this case. The residual signal in the image far from the halo centre is

similar in all images and is attributed to camera effects or leakage illumination from the scene. The symmetry about the spot polar angle – designated here as the "azimuth angle" – can be seen qualitatively from the images of Figure 4. Residual asymmetry is attributed to such factors as the scene background radiance, the geometry of the camera alignment (angle of incidence of the source light with respect to the optic axis of the NVG), the camera aspect ratio, the curvature of the NVG output surface, noise statistics (e.g. camera fixed pattern noise) and reflections, scattering and other phenomena within the NVG. An investigation into the consequences of these secondary effects is ongoing.
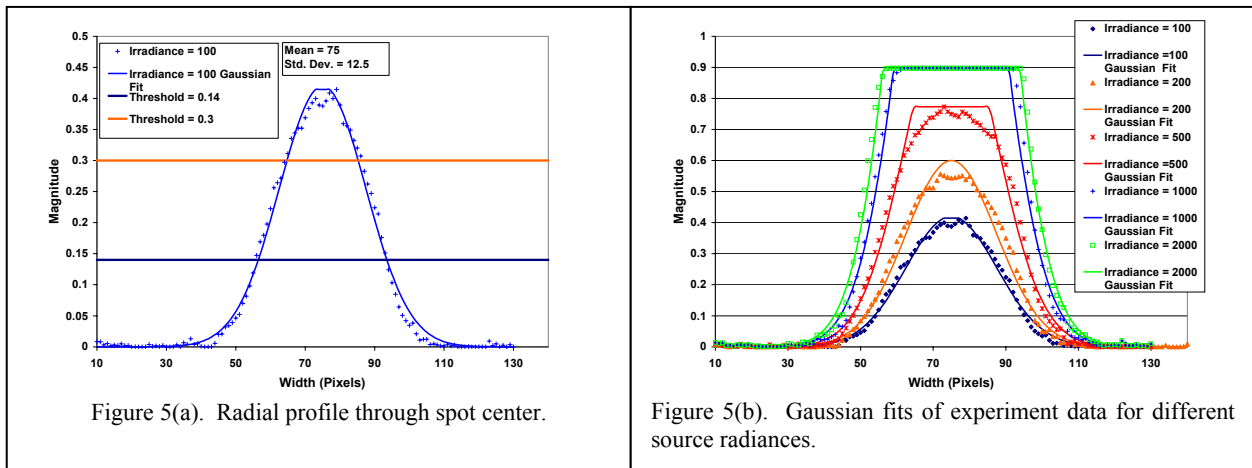

(a)

(b)

(c)

(d)

Figure 4. Images of point object with relative intensity a) 1, b) 10, c) 100, d) 1000

The colour balance in the imager was adjusted in the ratio green/red/blue = 4:2:1, in order to investigate saturation in the camera. When there is no saturation, the image is green, reflecting the true colour of the NVG output. After the green channel becomes saturated on a pixel, its colour balance shifts. When all colours are saturated on a pixel, its colour will be white. This variation in colour of the image can mislead the eye during a qualitative inspection of the imagery. The effect of saturation in the imager was verified by reduction of the integration time, to reduce the camera signal, and also by separate analysis of each of the colour channels. At even the largest irradiances used here, the "flat top" of the spot profile sue to saturation could be removed by appropriate selection of the camera gain and integration time.

The spot profile (the radial profile through the spot centre), or cross section, for each image was measured from the average of a sequence of 10 images acquired under identical conditions. A Matlab routine determined the profile of the spot in different azimuthal directions, and then averaged the spot profile over azimuth. A radon transform[9] was used to verify this method, and similar results were obtained. Figure 5(a) shows a representative measured cross section of a halo spot along with the corresponding Gaussian fit. The average deviation of measurements from the fit is approximately 3% of peak except in the wings where scene and camera effects become important. Some structure can also be seen near the peak.

Figure 5(a). Radial profile through spot center.

Figure 5(b). Gaussian fits of experiment data for different source radiances.

In Figure 5(b), the same halfwidth was used to fit halo spots whose radiances were up to 100x larger. Figure 6 demonstrates that a constant σ gives a good fit to the data, in that the fitted values of N_{peak} agree with the values expected from the relative irradiance. The peak value was selected from a-priori knowledge of the relative source irradiance, so that the only adjustable parameter was the halfwidth.

Figure 6. Best-fit N_{peak} and halfwidth for spot profiles

The measured spot width is shown in Figure 7 as a function of relative irradiance. The spot width in mm is found by multiplying the spot width in pixels by a scale factor. The experiment to determine the scale factor is described in the next section. The spot width data was obtained using the same Matlab routine used to determine the spot profile. Again, the radon transform confirmed the results of the "azimuthal averaging" method described above. Two thresholds were used, $N_{thresh} = 0.14 \times$(digitizer_range) and $N_{thresh} = 0.3 \times$(digitizer_range). As expected from the Gaussian model (Equation (2)), the width was smaller for the larger threshold and the shapes of the curves for the two thresholds were similar. Other threshold values were also used to confirm these results. In addition, the curves were not affected by the saturation region in the centre of the image, as long as N_{thresh} was lower than the saturation value.
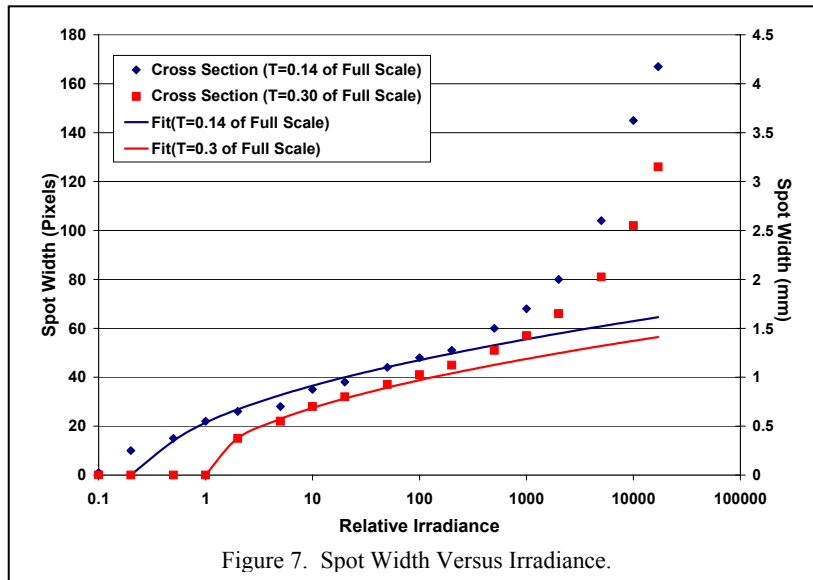
Figure 7. Spot Width Versus Irradiance.

Agreement between "fit" lines, calculated from Equation (2), and the measured data confirmed the Gaussian model for halo for the middle region of Figure 7, for constant σ. At low irradiance levels, the measured spot size is affected by the angular size of the light source, controlled by the geometry (pinhole size and source-to-NVG distance) of the experiment and by the camera resolution. At larger irradiances, the halfwidth becomes dependent on the irradiance, as indicated by Equation (3), and needed significant modification for the Gaussian profile to fit the measured halo spot width. Figure 8 shows this adjustment of the halfwidth. The "adjusted sigma" curve points were fitted with linear and quadratic terms of a piecewise Taylor expansion in log-irradiance.
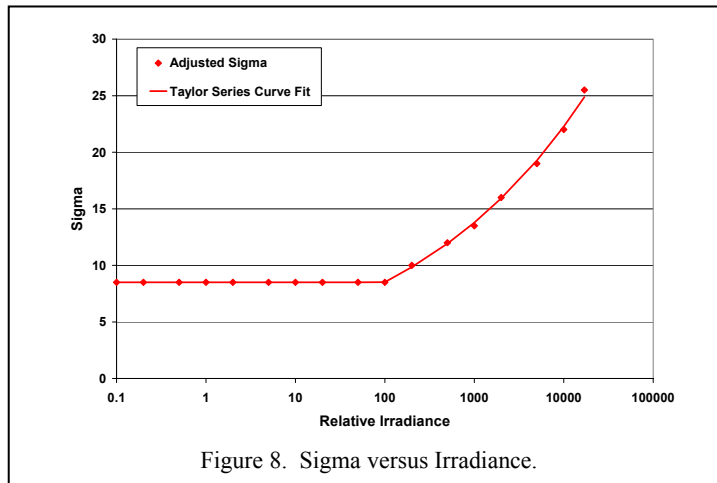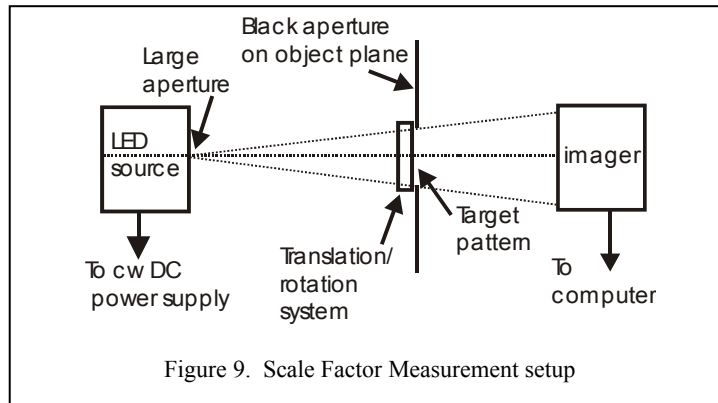

Figure 8. Sigma versus Irradiance.

The fit of the experimental data points in Figure 7 to the Gaussian model has not been corrected for AGC effects or incomplete azimuthal symmetry (described earlier).

3.2. Scale factor

To relate the size in pixels of a halo spot with the corresponding physical dimension of the halo on the output plane of the NVG, the experimental setup of Figure 9 was used for the measurement of the scale factor. This configuration employed the LED source and imager from the halo measurement scenario but replaced the NVG with a target pattern that was carefully aligned (rotation and translation) to coincide with the output surface of the fiber image guide. Because of the curvature of the image guide output surface, the alignment was made such that the normal to the target

plane was parallel to the optic axis of the imager. This optic axis was nominally aligned to coincide with the optic axis of the image guide. The systematic error in target alignment was an important source of uncertainty in the determination of the scale factor.

Figure 9. Scale Factor Measurement setup

The target object consisted of a pseudo-random Gaussian spatial distribution of grayscale (i.e. transmission). When the target was translated by a sequence of precisely known distances transverse to the optic axis of the imager, the corresponding changes in image position were determined by a correlation analysis of the images. This method allowed the scale factor to be determined to about 1% precision. The value of the scale factor was dependent on the imager lens and detector array, and the object distance, as can be seen from Figure 9. A typical value for the current experiments was scale_factor = 25.0 ± 0.5 μm/pixel.

4. Automated Gain Control (AGC) Effects

To assess the effects of AGC, the experimental setup of Figure 2 was used. The background irradiance was adjusted in increments to cause the NVG gain to change through the operation of its AGC. Figure 10 shows the change in a halo image when background illumination is added. The halo spot in the left hand image shrinks dramatically when background irradiance is present. In the right hand figure, part of the background illumination can be seen. As mentioned in Section 2, the background irradiance is blocked in the region around the halo spot, in order to facilitate the measurement of halo spot width.
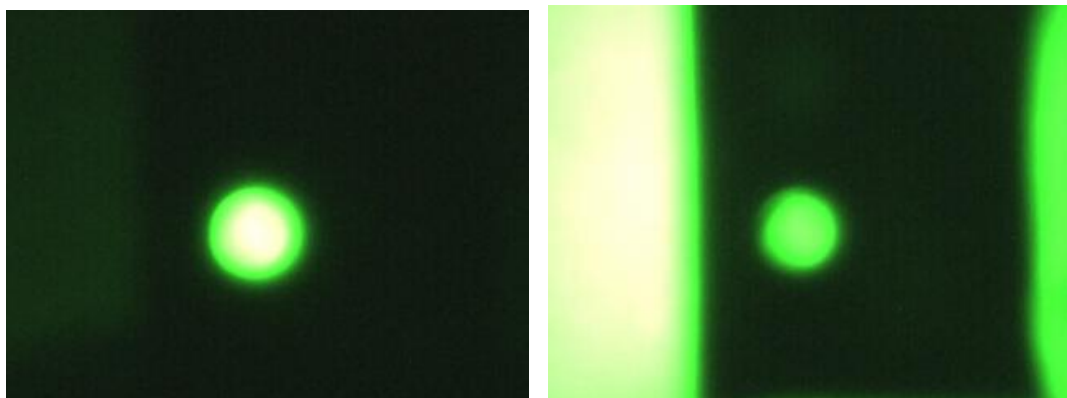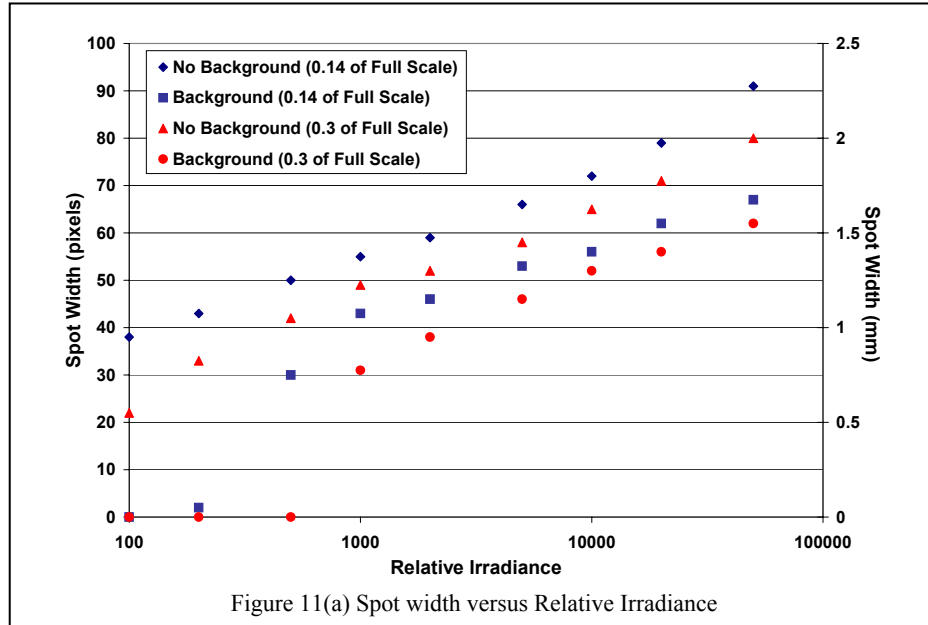

Figure 10. AGC Effects (a) no background and (b) with background

Figure 11(a) shows the spot width for the "no-background" and "background" cases under measurement conditions that were otherwise identical.

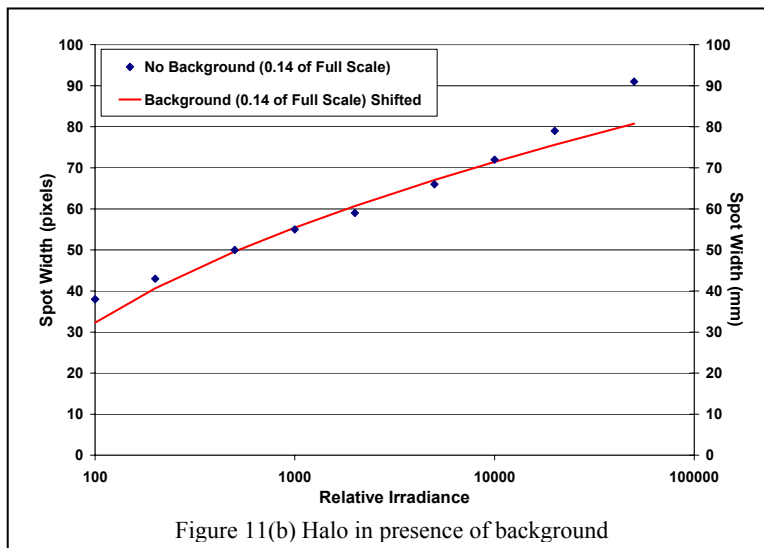Figure 11(a) Spot width versus Relative Irradiance

All the curves of Figure 11(a) are similar in shape to those depicted in Figure 7. As with Figure 7, two thresholds are shown. The curves for the "no background" and "background" cases have similar shapes as well. The irradiance for the "background" case can be renormalized to account for the AGC operation using the relationship shown in Equation (4).

$$I_{renorm} = I * k_{agc}$$
$$0 \le k_{agc} \le 1$$

(4)

In Figure 11(a) k_{agc}=1 for the "no background" case (assuming that the halo spot itself did not trigger AGC operation). The value of k_{agc} for the "background" case was determined using a least-squares fit. For the data of Figure 11(a) k_{agc}= 0.15 $\pm$ 3% brings the two curves together (Figure 11(b)). Similar measurements for a range of known background irradiances will map out k_{agc} as a function of irradiance.

Figure 11(b) Halo in presence of background

5. REFLECTIONS AND SCATTERING

At high irradiance levels of the point-like source, secondary spots can be seen in the image. Figure 12 shows two typical halos with reflection spot. These spots appear akin to reflections within the NVG and are modeled here, from an empirical perspective, as reflections. The reflections could be for electrons, photons or both. A possible mechanism for reflection is sketched in Figure 13.
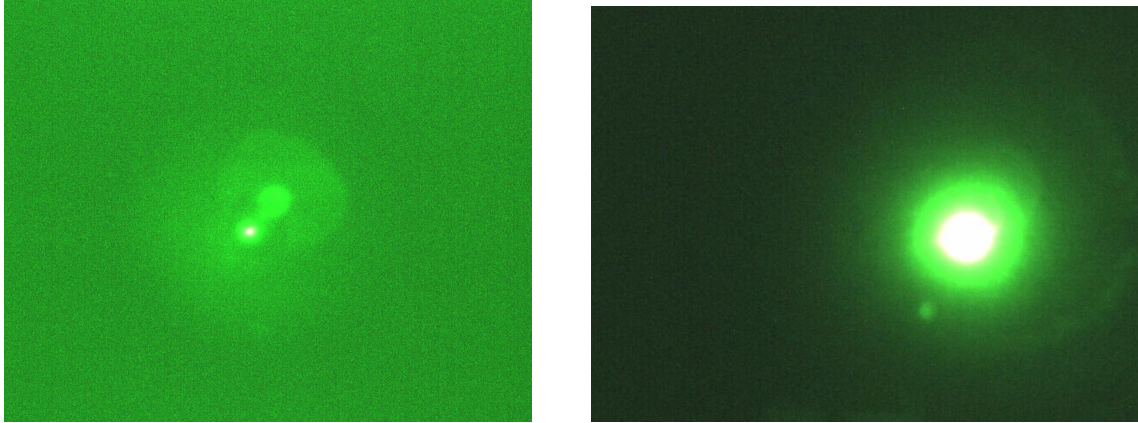

Figure 12. Typical halos with reflection spot, for two different NVGs

The position of the spot in the exit window (Δ_{refl}) is a function of the tilt angle (θ_{tilt}) of the pores of the MCP, the spacing (Δ_{fl}) between the MCP and the exit window (containing the fluorescent film), and the number of reflections (k_{spot}). This relationship is shown in Equation (5). The first reflection, $k_{spot}=1$, may be hidden by the main halo, while other reflections may be too faint to see. Reflections between plates of a multi-plate stack of MCPs are ignored.

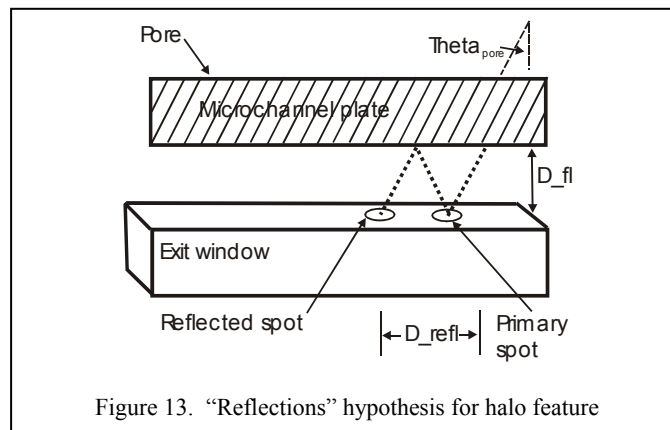$$\Delta_{refl} \cong 2k_{spot}\Delta_{fl} \tan(\theta_{tilt}) \tag{5}$$

Figure 13. "Reflections" hypothesis for halo feature

From Figure 12 the offset of the reflection spot from the original halo spot was measured to be 99.5 $\pm$ 2 pixels (2.5 mm $\pm$ 5 %) for the image on the right. This value is not consistent with either the first or the second bounce of the electron reflection from the fluorescent screen, when Δ_{fl} =500 um and θ_{tilt} = 12°. A fit can be made to optical reflection from the second surface of the exit window or the first surface of the image guide, since the equivalent Δ_{fl} is larger, suggesting that the observed reflection spot is "optical" rather than "electron" in cause.. Qualitative estimates of the relative intensities of the main halo and reflection can be fitted to reflectivities, to see if the multiple bounces should be above

the detection threshold. Additional experiments are needed to confirm this tentative reflection model and extract realistic parameter values.

In the left hand image of Figure 12, acquired from a different NVG, the extended dimension of the reflection phenomenon suggests that scattering of light might be important. A comprehensive model of reflection effects needs to include scattering and edge effects as well as the Fresnel reflections from surfaces.

6. SUMMARY

As part of continuing work to model NVG performance (to be used in human performance research), experiments for measuring the "thresholded" spot width and profile of a point light source of different irradiances were carried out. The data confirms that the halo profile can be modeled as a Gaussian distribution, which is unusual for a phenomenon that has dynamic range spanning 7 orders of magnitude. The spot width data indicated that there were two regimes (weak and large signal) controlled by the halfwidth of the corresponding halo profile. Correction for the effects of the automatic gain control was demonstrated. Future work will concentrate on isolating deviations from the halo model due to other secondary effects, such as reflection and scattering.

ACKNOWLEDGEMENTS

REFERENCES

1. E.H. Eberhart, "An Operational Model for Microchannel Plate Devices", *IEEE Transactions on Nuclear Science*, **Vol 28**, pp-712-717, 1980.

2. "Operating Characteristics of Sandwich Microchannel Plates", J.F. Pearson, J.E. Lees and G.W. Fraser, IEEE Trans. Nuc. Sci. Vo. 35, No. 1, February 1988.

3. H. Kume, K. Koyama, K. Nakatsugawa, S. Suzuki and David Fatiowitz, "Ultrafast microchannel plate photomultipliers", *Appl. Opt*, **Vol. 27(6)**, pp 1170-1178, 1988.

4. Paul J. Thomas, R.S. Allison, T. Macuda, S. Jennings, K. Yip, E. Savchenko, I. Tsang, R. Hornsey, "Experimental validation of synthetic imagery for night vision goggles", *SPIE*, **Vol. 5442**, paper 03, April 2004 .

5. P. Girard, Y. Beauvais and Pl de Groot, "Night Vision with Generation III Image Intensifiers, in R. Hradaynath ed., *Night Vision Technology*, Vol. MS 169, SPIE, 2001.

6. Edward H. Eberhard, "Image transfer properties of proximity focused image tubes", Appl. Opt. Vol. 16(8), pp 2127-2133 (August 1977).

7. G. Papp. IRE Trans. Nucl. Sci. NS-9 (4), 55 (1962).

8. A.S. Sedra and K.C. Smith, *Microelectronic Circuits*, Oxford University Press, New York, 1998.

9. S. Helgason, *The Radon Transform*, Birkhäuser, Boston, 1999.

10. Norman S. Kopeika, *A System Engineering Approach to Imaging*, SPIE Optical Engineering Press, Bellingham, WA, USA, 1998.