12th International Conference on Information Fusion
Seattle, WA, USA, July 6-9, 2009

Towards an Optimal Color Representation for Multiband Nightvision Systems

Alexander Toet
TNO Human Factors
Soesterberg, The Netherlands.
lex.toet@tno.nl

Maarten A. Hogervorst
TNO Human Factors
Soesterberg, The Netherlands.
maarten.hogervorst@tno.nl

Abstract - *We present a new Tri-band Color Low-light Observation (TRICLOBS) system The TRICLOBS is an all-day all-weather surveillance and navigation tool. Its sensor suite consists of two digital image intensifiers and an uncooled longwave infrared microbolometer. This sensor suite registers the visual, near-infrared and longwave infrared bands of the electromagnetic spectrum. The optical axes of the three cameras are aligned, using two dichroic beam splitters. A fast lookup-table based color transform (the Color-the-Night color mapping principle) is used to represent the TRICLOBS image in natural daylight colors (using information in the visual and NIR bands) and to maximize the detectability of thermal targets (using the LWIR signal). A bottom-up statistical visual saliency mode is deployed in the initial optimization of the color mapping for surveillance and navigation purposes. Extensive observer experiments will result in further optimization of the color representation for a range of different tasks.*

Keywords: Image fusion, false color, natural color mapping, real-time fusion, lookup tables, visual saliency.

1 Introduction

Night vision cameras are a vital source of information for a wide-range of critical military and law enforcement applications related to surveillance, reconnaissance, intelligence gathering, and security. The two most common night-time imaging systems cameras are low-light-level (e.g., image-intensified) cameras, which amplify the reflected visible to near infrared (VNIR) light, and thermal infrared (IR) cameras, which convert invisible thermal energy from the midwave (3 to 5 microns) or the long wave (8 to 12 microns) part of the spectrum into a visible image. Until recently a gray- or greenscale representation of nightvision imagery has been the standard. However, the increasing availability of fused and multi-band infrared and visual nightvision systems has led to a growing interest in the color display of night vision imagery [9, 11, 12, 16, 23]. In principle, color imagery has several benefits over monochrome imagery for surveillance, reconnaissance, and security applications. For instance, color may improve feature contrast, which allows for better scene recognition and object detection. When sensors operate outside the visible waveband, artificial color mappings generally produce false color images whose chromatic characteristics do not correspond in any intuitive or obvious way to those of a scene viewed under natural photopic illumination. This type of false color imagery may disrupt the recognition process, resulting in an observer performance that is even worse compared to that obtained with singleband imagery alone [13]. Several different techniques have been proposed to display night-time imagery in natural daylight colors [14-17, 20, 23], some of which have been implemented in realtime nightvision systems [2, 8, 18, 19, 21]. Most of these techniques are computationally expensive and/or do not achieve color constancy. We recently introduced a new color mapping that displays night-time imagery in natural daytime colors [7]. This technique is simple and fast, and can easily be deployed in realtime. Moreover, it provides stable colorization under variations in scene content [6, 7].

Here we describe the implementation of this new color mapping in the prototype TRICLOBS (TRI-band Color Low-light OBServation) all-day all-weather surveillance and navigation system. The system displays the co-aligned visual, near-infrared and thermal signals of respectively two image intensifiers and an uncooled microbolometer in full color. A fast lookup-table implementation of the Color-the-Night color mapping transform2 is deployed to represent the TRICLOBS image in natural daylight colors (using information in the visual and NIR bands) and to maximize the detectability of thermal targets (using the LWIR signal). A bottom-up statistical visual saliency model [22] is deployed to optimize the color mapping for surveillance and navigation purposes.

2 Color mapping

The principle of the new lookup-table based color mapping technique is explained in detail in [6]. For the sake of completeness we will briefly describe this procedure here. First, a false color image is constructed by mapping the different bands of a multisensor nightvision system to respectively the R, G, and B channels of an RGB image (set channel B to zero when only 2 bands are available, and use only the first three principal components when the system provides more than 3 bands). Second, transform the false color image thus obtained into an indexed image using a color lookup table containing a set RGB triples (this is a 3D lookup table, which reduces to 2D when only 2 bands are available). Finally, replace the false color lookup table of the input multiband nightvision image with a new color lookup table that maps the two image bands onto natural colors. The new color lookup-table can be obtained either by applying a statistical transform to the entries of the original lookup-table, or by a procedure that replaces entries of the original lookup-table by their corresponding natural color values. The statistical transform method transfers the first order statistics (mean and standard deviation) of the color distribution of a representative natural color daytime reference image to the false color multiband nighttime image [6, 15]. This mapping is usually performed in a perceptually de-correlated color space (e.g. lαβ [10]).The sample-based method deploys a set of corresponding samples from the combination of a multi-band sensor image of a given scene and a registered naturally colored (RGB) daytime reference image of the same scene to derive a color lookup table transform pair that transfers the color characteristics of the natural color reference image to the false color nighttime image [6, 7]. For an 8-bit multi-band system providing 3 or more bands the 3D color lookup table contains 256x256x256 entries (for a 2 band system the 2D table contains 256x256 entries). When the color lookup table contains fewer entries, the color mapping is achieved by determining the closest match of the table entries to the observed multi-band sensor values. Once the color transformation has been derived and the pair of color lookup tables that defines the mapping has been created, they can be used in a real-time application. The lookup table transform requires minimal computing power. An additional advantage of the color lookup transform method is that object colors only depend on the multi-band sensor values and are independent of the image content. As a result, objects keep the same color over time when registered with a moving camera.

In the next sections we first describe the overall system design and the components of a prototype portable tri-band real-time nightvision system that deploys the new lookup-table color mapping. Then we will explain how a simple bottom-up saliency model (SUN: [22]) can be deployed to optimize this color mapping for different applications. Finally, we will show the results of some preliminary field trials.

3 System design

3.1 Overview

The TRICLOBS system combines a three-band nightvision sensor suite, consisting of two digital image intensifiers and a thermal (LWIR) camera, in combination with a 3D digital position information system. The night vision sensor suite is sensitive in the visual (400-700 nm), the near-infrared (700-1000 nm) and the longwave infrared (8-14 μm) bands of the electromagnetic spectrum. The optical axes of all cameras are aligned. Figure 1 shows a schematic representation of the layout of the sensors suite and the beam splitters that are deployed to direct the appropriate band of the incoming radiation to each of the individual sensors. The incoming radiation is first split into a longwave (thermal) and a visual+NIR part by a heat reflecting (hot) mirror. The longwave part of the spectrum is reflected into the lens of the XenICs Gobi camera, while the visual+NIR light is transmitted to the combination of the two Photonis ICUs. The two ICUs are mounted under an angle of 90 degrees. A near-infrared reflecting mirror is used to separate the incoming light such that one ICU registers the visual part and the other ICU only detects the NIR part of the incoming radiation. The sensor suite and the mirrors are mounted on a common metal base. The whole configuration is placed in an enclosed housing. Figure 2 shows a test setup of this sensor configuration.

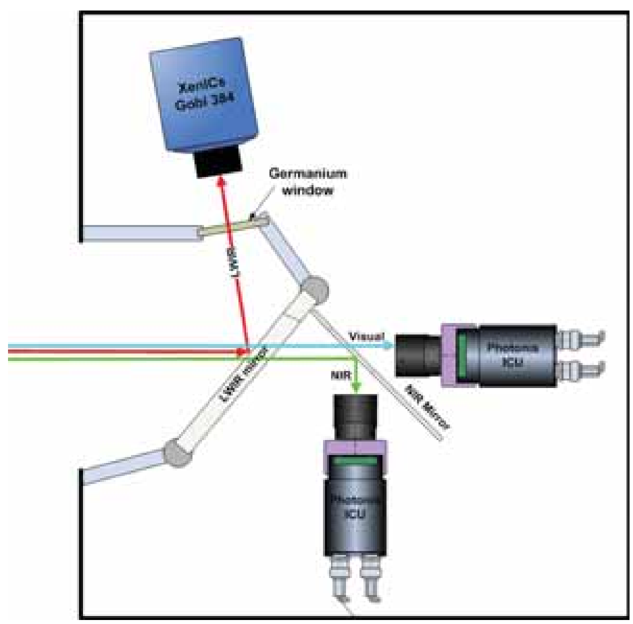

Figure 1. TRICLOBS Sensor suite layout.

3.2 Signal Processing

The TRICLOBS system delivers both analog video and digital signal output. The 16-bit TCP/IP Ethernet interface of the XenICs Gobi 384 is connected to a Netgear Gigabit Ethernet switch. The SDI channels of both Photonis ICU's are also connected to this Netgear hub through SDI/Etherrnet converters. This enables the user to transmit high quality TRICLOBS signals over an Ethernet connection.

The USB ports of both Photonis ICU's are connected to a high-speed 7-port USB 2.0 hub. This enables the user to interface with the ICU's and to adjust their settings or to download and install preferred settings.

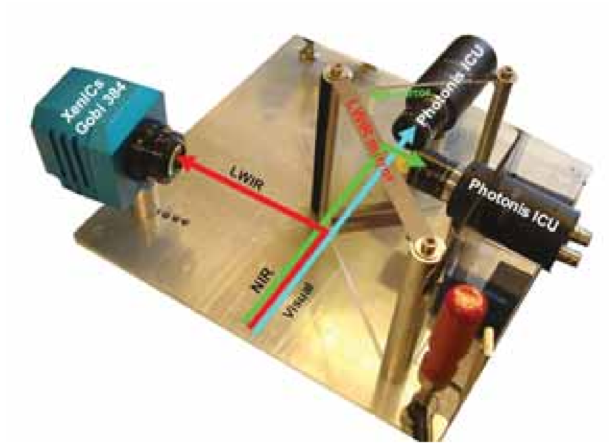

Figure 2. Test setup of the TRICLOBS Sensor suite.

All camera signals can either be stored on disk for offline processing, or can be processed online further as follows. Note that each of the following processing steps can selectively be activated or de-activated.

A Pleora frame grabber (www.pleora.com) can be used to digitize the analog video signals of both ICU's (through the SDI/Ethernet converters) and the digital output of the Gobi 384 camera, and output these signals to the Netgear Gigabit Ethernet switch.

Three Pinnacle Video Transfer Units (www.pinnaclesys.com/PVT) can be deployed to store (a) the analog video signals of all three cameras and (b) the audio signals of two (optional) external microphones, either on 3 internal 320 Gb harddisks, or on USB memory sticks. The microphones can for instance be positioned on the front and back of the camera suite. The microphone on front can then be used to register relevant audio information from the registered scene, and the second microphone can for instance be used to record spoken annotations.

An internal U-blox EVK-5P Positioning Engine (www.u-blox.com) and a Silicon Laboratories (www.silabs.com) F350-Compass-RD provide respectively position and orientation (i.e. sensor location and viewing direction) signals through the high-speed 7-port USB 2.0 hub.

Two external video displays are provided to simultaneously monitor two of the three video signals (either Visual/NIR, Visual/LWIR, or NIR/LWIR).

The Color-the-Night false color mapping is performed on an external processing unit, connected to the TRICLOBS via an Ethernet connection. This can either be a regular laptop (since the operation is efficiently implemented as a lookup table transform only a minimal amount of computation is required to achieve real-time performance), a regular PC (in case portability is not an issue), or a dedicated PC. The entire system can either run on an internal battery pack, or on 220V AC.

3.3 Digital Image Intensifiers

The TRICLOBS contains two Photonis Intensified Camera Units (ICUs): an ICU PP3000L and an ICU PP3000U (Fig. 6, see: www.photonis.com). The ICU is a new generation of low light level, intensified CMOS camera. It has a 2/3" CMOS sensor with a spectral response range of 400-900 nm, and delivers both a PAL or NTSC composite video signal output (ITU-R BT.656-4, 640x480 pixels), and an SDI – LVDS 270 Mbits/s signal. It is equipped with a C-mount lens adapter. Both ICU's are equipped with Pentax C2514M CCTV lenses, with a minimal focal length of 25mm and a lens aperture of F/1.4, resulting in a FOV of 30.7º x 24.8º.

Figure 3. Photonis Intensified Camera Unit (ICU PP3000L, see: www.photonis.com).

3.4 Thermal Camera

The XenICs Gobi 384 uncooled a-Si infrared microbolometer (Fig. 7, see: www.xenics.com) has a 384 x 288 pixel focal plane array, and a spectral sensitivity range of 8 – 14µm, which is the range of most interest for outdoor applications. It is equipped with an 18mm (f/1) lens providing a 29.9° x 22.6° wide angle view. The Gobi 384 has a 16-bit Ethernet and CameraLink interface.

Figure 4. The XenICs Gobi 384 infrared microbolometer (see: www.xenics.com).

3.5 LWIR Mirror

A custom made Melles Griot dichroic beam splitter consisting of Schott N-BK7 Borosilicate Crown glass with an Indium Tin Oxide (ITO) coating (www.ocioptics.com/ito.html) is used to split the LWIR part of the incoming radiation and reflect it into the lens of the XenICs Gobi 384 thermal camera. This filter transmits the visual/near-infrared band (400-900nm) and reflects the longwave (thermal) infrared part (7000-14000nm). According to the specifications provided by Melles Griot (Fig. 8) the reflection R > 84.75% for 7.5–9.5 µm, R>87.25% for 9.5–11.5 µm, R>90% for 420 – 700 µm, R>80 for 1064 µm, R > 50% for 1540–1570 µm, all measured at an angle of incidence of about 50°.

3.6 NIR Mirror

A hot mirror filter (45 deg angle of incidence, type Edmund Optics B43-958, 101x127x3.3 mm, see: www.edmundoptics.com) is deployed to split the visual/near-infrared band (400–900nm) by transmitting the visual (400–700nm) and reflecting the NIR part (700–900nm) of the spectrum (Fig. 9).

4 Image Optimization

The TRICLOBS system provides different color mappings so that the image representation can be adjusted to the task at hand and to the environmental conditions. For navigation and surveillance applications a natural color image appearance will usually be preferred. For tasks involving target detection and situational awareness a representation is needed that provides an enhanced display of the relevant image features. Initially we deploy a visual saliency model (SUN:see [22]) to derive optimal representation schemes for different tasks and conditions. Later we will use the results of observer tests in realistic scenarios to optimize the color mapping for individual tasks.

Human visual attention is largely driven bottom-up by the saliency of image details. An image detail appears salient when one or more of its low-level features (e.g. size, shape, luminance, color, texture, binocular disparity, or motion) exceeds the overall feature variation of the background. Recently several information theoretical approaches have been presented to compute visual saliency of local image features [3-5, 22].. These methods are based on the assumption that feature saliency is inversely related to feature occurrence (i.e. rare features are more informative and therefore more salient or surprising than features that occur more frequently). In this view interesting image details correspond locations of maximal self information (a measure closely related to local feature contrast: [1, 4]), and saliency driven free viewing corresponds to maximizing information sampling [3, 22]. These models have successfully been deployed to model human fixation behavior, pop-out, dynamic saliency, saliency asymmetries, and to solve to classic computer vision problems like dynamic background subtraction [3-5]. Here we apply a simple Bayesian model of natural image statistics (SUN) to compute bottom-up saliency from the self information of local image features [22]. We use the resulting bottom-up saliency map to derive a color mapping scheme for surveillance and navigation applications. The resulting image closely approximates a natural daytime image. For search and detection applications the saliency of relevant targets should be optimized. We hope to achieve this in a later stage by including a top-down component in the saliency model that calculates the mutual information between the target and the image content [22]. This will allow us to maximize the saliency of relevant targets by boosting the mutual information between their characteristic features and their representation in the TRICLOBS image. In addition, we are currently performing extensive field trials with human observers to evaluate the different image representations in realistic scenarios.

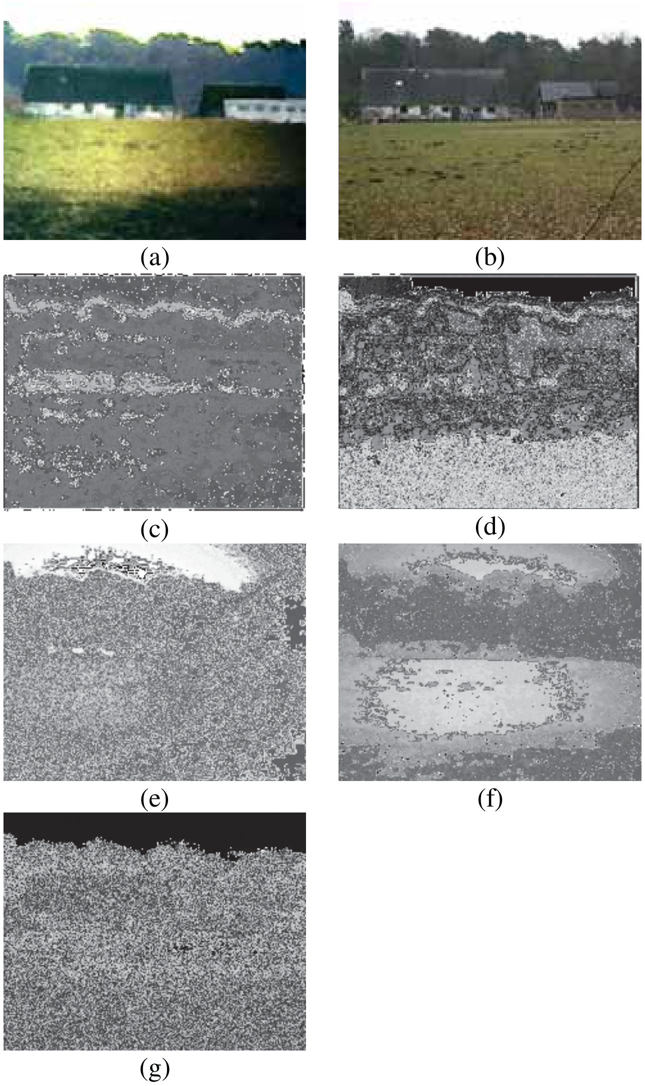

Figure 5. (a) False color nighttime image and (b) corresponding daytime image of same scene. (c,d) Saliency maps of (a,b). (e-g) respectively the Visual, NIR, and LWIR bands of (a).

5 Initial Results

We tested a prototype of the TRICLOBS realtime night vision system in some nocturnal data collection trials in the field.

Figure 5 shows a frame of the actual real-time false-color TRICLOBS image of a rural scene, registered in full darkness (luminance < 0.03 lux). Fig. 5b shows the daytime image of same scene for comparison. Figs. 5c,d show that the saliency map of the TRICLOBS nighttime image closely approximates the saliency map of the corresponding daytime image, indicating that most of the relevant daytime images features are represented with similar saliency in the nighttime image. Figs. 5e,f and g show respectively the three individual bands of the TRICLOBS image, i.e. the visual (wavelengths below 700 nm), NIR (wavelengths between 700 and 900 nm), and thermal (8–14μm) bands. The false-color TRICLOBS image in Fig. 5a was obtained through application of our new Color-the-Night remapping technique [7] to the raw false color image that was initially formed by assigning the images Figs. 5e,f, and g to each of the individual bands of a false color RGB image. Note that the resulting false color nightvision image (Fig. 5a) closely resembles the corresponding daytime photograph of the same scene (Fig. 5b). Also, note that it is much easier to distinguish different materials and objects in Fig.5a, than in each of the individual bands (Figs. 5e,f,g).

6 Conclusions

In this paper we presented the prototype TRICLOBS portable tri-band realtime night vision system that can be used to demonstrate the operational value of a newly developed real-time color mapping that applies natural daylight colors to multi-band night-time images. The TRICLOBS system provides real-time co-aligned visual, near-infrared and thermal images. These co-aligned images can either be stored on on-board harddisks, or they can be processed and displayed in real-time by a (notebook) computer. A real-time natural color mapping is implemented as a lookup table transform. The results of some preliminary field trials clearly demonstrate the potential benefits of this system for surveillance, navigation and target detection tasks. The resulting false color nightvision image closely resembles a daytime image, while thermal targets are clearly distinguishable. At this stage the color mapping is initially optimized through a bottom-up visual saliency model. At a later stage we will deploy a top-down version of this model to maximize the mutual information of relevant targets and their representation in the fused image. Finally, extensive observer testing in field trials will be performed to further optimize the color mapping scheme for different tasks.

References

[1] Bruce, N.D.B. & Tsotsos, J.K. (2009). Saliency based on information maximization. In Y. Weiss, B. Schölkopf & J. Platt (Ed.), *Advances in Neural Information Processing Systems 18* (pp. 155-162). MIT Press: Cambridge, MA.

[2] Fay, D.A., Waxman, A.M., Aguilar, M., Ireland, D.B., Racamato, J.P., Ross, W.D., Streilein, W. & Braun, M.I. (2000). Fusion of multi-sensor imagery for night vision: color visualization, target learning and search.*Proceedings of the 3rd International Conference on Information Fusion* (pp. TuD3-3-TuD3-10). Paris, France: ONERA.

[3] Gao, D., Mahadevan, V. & Vasconcelos, N. (2008). On the plausibility of the discriminant center-surround hypothesis for visual saliency. *Journal of Vision, 8(7),* 13-1-13-18.

[4] Gao, D. & Vasconcelos, N. (2005). Discriminant saliency for visual recognition from cluttered scenes. In L.K. Saul, Y. Weiss & L. Bottou (Ed.), *Advances in neural information processing systems 17* (pp. 481-488). Cambridge, MA: MIT Press.

[5] Gao, D. & Zhou, J. (2001). Adaptive background estimation for real-time traffic monitoring.*Proceedings of the IEEE Conference on Intelligent Transportation Systems 2001* (pp. 330-333). Washington, USA: IEEE Press.

[6] Hogervorst, M.A. & Toet, A. (2008). Nighttime imagery in natural daytime colors. In B.V. Dasarathy (Ed.), *Multisensor, Multisource Information Fusion: Architectures, Algorithms, and Applications 2008* Bellingham, WA, USA: The International Society for Optical Engineering.

[7] Hogervorst,M.A., Toet,A., & Kooi,F.L. (2006). TNO Defense Security and Safety. Method and system for converting at least one first-spectrum image into a second-spectrum image. Patent Number PCT/NL2007050392. Application Number 0780855.5-2202,

[8] Howard, J.G., Warren, P., Klien, R., Schuler, J., Satyshur, M., Scribner, D. & Kruer, M.R. (2000). Real-time color fusion of E/O sensors with PC-based COTS hardware. In W.R. Watkins, D. Clement & W.R. Reynolds (Ed.), *Targets and Backgrounds VI: Characterization, Visualization, and the Detection Process* (pp. 41-48). Bellingham, WA: The International Society for Optical Engineering.

[9] Li, G. & Wang, K. (2007). Applying daytime colors to nighttime imagery with an efficient color transfer method. In J.G. Verly & J.J. Guell (Ed.), *Enhanced and Synthetic Vision 2007* (pp. 65590L-1-65590L-12). Bellingham, MA: The International Society for Optical Engineering.

[10] Ruderman, D.L., Cronin, T.W. & Chiao, C.-C. (1998). Statistics of cone responses to natural images: implications for visual coding. *Journal of the Optical Society of America A, 15(8),* 2036-2045.

[11] Shi, J., Jin, W., Wang, L. & Chen, H. (2005). Objective evaluation of color fusion of visual and IR imagery by measuring image contrast. In H. Gong, Y. Cai & J.-P. Chatard (Ed.), *Infrared Components and Their Applications* (pp. 594-601). Bellingham, MA: The International Society for Optical Engineering.

[12] Shi, J.-S., Jin, W.-Q. & Wang, L.-X. (2005). Study on perceptual evaluation of fused image quality for color night vision. *Journal of Infrared and Millimeter Waves, 24(3),* 236-240.

[13] Sinai, M.J., McCarley, J.S. & Krebs, W.K. (1999). Scene recognition with infra-red, low-light, and sensor fused imagery.*Proceedings of the IRIS Specialty Groups on Passive Sensors* (pp. 1-9). Monterey, CA: IRIS.

[14] Sun, S., Jing, Z., Li, Z. & Liu, G. (2005). Color fusion of SAR and FLIR images using a natural color transfer technique. *Chinese Optics Letters, 3(4),* 202-204.

[15] Toet, A. (2003). Natural colour mapping for multiband nightvision imagery. *Information Fusion, 4(3),* 155-166.

[16] Tsagiris, V. & Anastassopoulos, V. (2005). Fusion of visible and infrared imagery for night color vision. *Displays, 26(4-5),* 191-196.

[17] Wang, L., Jin, W., Gao, Z. & Liu, G. (2002). Color fusion schemes for low-light CCD and infrared images of different properties. In L. Zhou, C.-S. Li & Y. Suzuki (Ed.), *Electronic Imaging and Multimedia Technology III* (pp. 459-466). Bellingham, WA: The International Society for Optical Engineering.

[18] Wang, L., Zhao, Y., Jin, W., Shi, S. & Wang, S. (2007). Real-time color transfer system for low-light level visible and infrared images in YUV color space. In I. Kadar (Ed.), *Signal Processing, Sensor Fusion, and Target Recognition XVI* (pp. 1-8). Bellingham, WA: The International Society for Optical Engineering.

[19] Waxman, A.M., et al. (1999). Solid-state color night vision: fusion of low-light visible and thermal

infrared imagery. *MIT Lincoln Laboratory Journal, 11,* 41-60.

[20] Waxman, A.M., Fay, D.A., Gove, A.N., Seibert, M.C., Racamato, J.P., Carrick, J.E. & Savoye, E.D. (1995). Color night vision: fusion of intensified visible and thermal IR imagery. In J.G. Verly (Ed.), *Synthetic Vision for Vehicle Guidance and Control* (pp. 58-68). Bellingham, WA: The International Society for Optical Engineering.

[21] Yue, Z. & Topiwala, P. (2007). Real-time EO/IR sensor fusion on a portable computer and head-mounted display. In B.V. Dasarathy (Ed.), *Multisensor, Multisource Information Fusion: Architectures, Algorithms, and Applications 2007* (pp. 657106- Bellingham, WA: The International Society for Optical Engineering.

22] Zhang, L., Tong, M.H., Marks, T.K., Shan, H. & Cottrell, G.W. (2008). SUN: A Bayesian framework for saliency using natural statistics. *Journal of Vision, 8(7),* 1-20.

[23] Zheng, Y., Hansen, B.C., Haun, A.M. & Essock, E.A. (2005). Coloring night-vision imagery with statistical properties of natural colors by using image segmentation and histogram matching. In R. Eschbach & G.G. Marcu (Ed.), *Color imaging X: processing, hardcopy and applications* (pp. 107-117). Bellingham, WA: The International Society for Optical Engineering.