

# Colorizing Grayscale Intensified Nightvision Images

Alexander Toet

TNO Human Factors

P.O. Box 23

3769 ZG Soesterberg

The Netherlands

[toet@tm.tno.nl](mailto:toet@tm.tno.nl)

**Abstract** - We present a method to give single band intensified nightvision imagery a natural day-time color appearance. For input, the method requires a true color RGB source image and a grayscale nightvision target image. The source and target image are both transformed into a perceptually decorrelated color space. In this color space a best matching source pixel is determined for each pixel of the target image. The matching criterion uses the first order statistics of the luminance distribution in a small window around the source and target pixels. Once a best matching source pixel is found, its chromaticity values are assigned (transferred) to the target pixel while the original luminance value of the target pixel is retained. The only requirement of the method is that the compositions of the source and target scenes resemble each other.

**Keywords:** Color transfer, false color, nightvision, intensified imagery.

## 1 Introduction

Most common night-time imaging systems use single-band sensors that are sensitive to either emitted infrared (IR) radiation or reflected light, and thus provide a grayscale representation of the inspected scene. A false color representation of night-time imagery that closely resembles a natural daylight color image will help the observer by making scene interpretation more intuitive.

False color representation of grayscale images may result in a better scene segmentation and understanding, since the human eye can discern several thousands of colors, whereas it can only distinguish about 100 shades of grey at any instance. False color rendering of night-time imagery that agrees with natural daytime colors (intuitively correct color mapping) significantly improves observer performance and reaction times in tasks that involve scene segmentation and classification [2, 10-13, 15]. However, color mappings that do not result in natural object colors (counterintuitive color mapping) seriously degrade situational awareness [5, 11, 13]. One of the main reasons is the counter intuitive appearance of scenes rendered in some artificial color schemes and the lack of color constancy [13]. Hence, an ergonomic color scheme should produce night vision imagery with a natural appearance and with colors that are to some extent invariant for changes in the environmental conditions (i.e. the image should always have more or less the same appearance).

Welsh *et al.* [14] recently introduced a method to transfer a source image's color characteristics to a target grayscale image. The method employs a transformation to a principal component space that has recently been derived from a large ensemble of hyperspectral images of natural scenes [9]. This color space provides three decorrelated, principal channels corresponding to an achromatic luminance channel and two chromatic opponent channels. In this color space a best matching characteristic pixel in the color source image is determined for each pixel in the grayscale target image. The chromaticity values of the matching source pixel are then transferred to the target pixel while the original luminance value of the target pixel is retained. The only requirement of the method is that the composition of the source and target scenes is similar to some extent. Hence, the depicted scenes need not be identical; they merely have to resemble each other. For surveillance systems, that usually register a fixed scene, a daylight color image of the same scene that is being monitored is easily available and can be used to derive an optimal color mapping.

Here we apply the method of Welsh *et al.* [14] to transfer the characteristics of natural daylight color images to single-band intensified nightvision imagery. The results show that the method can be used effectively to give single-band intensified nightvision imagery a realistic day-time color appearance.

## 2 Method

The aim of the present study is to give grayscale nightvision images the appearance of normal daylight color images. In this section we show how a recently introduced technique to color grayscale photographs can be applied for this purpose [14].

The general problem of adding color to a grayscale image has no exact and objective solution, since one single grayscale value may correspond to a range of different colors. However, Welsh *et al.* [14] recently presented a technique to transfer color between a source, color image and a target, grayscale image. The method works as follows.

First, both the color (source) and grayscale (target) RGB images are converted into the decorrelated  $l\alpha\beta$  space [9] for subsequent analysis. This color space was developed to minimise the correlation between the three

color coordinate axes. The color space provides three decorrelated, principal channels corresponding to an achromatic luminance channel ( $l$ ) and two chromatic channels  $\alpha$  and  $\beta$ , which roughly correspond to yellow-blue and red-green opponent channels. Thus, changes made in one color channel should minimally affect values in the other channels. The reason the  $l\alpha\beta$  color space was selected for the current procedure is because it allows us to selectively transfer the chromatic  $\alpha$  and  $\beta$  channels from the color image to the grayscale image without introducing cross-channel artefacts.

Second, a subset of pixels in the color (source) image are selected as samples. The samples should be (randomly) distributed over the entire area of the source image, to guarantee the availability of a characteristic sample for each of the details (materials) represented in the scene. For each of the sample pixels we compute the mean and standard deviation of the luminance ( $l$ ) value over a small neighborhood of the pixel. We found that a neighborhood size of 7x7 pixels works well for most images. This agrees with the results of Welsh *et al.* [14].

Third, for each pixel in the grayscale (target) image in scan-line order the best matching sample in the color (source) image is selected, based on the weighted average of luminance (50%) and standard deviation of luminance (50%). Once the best matching pixel is found, its  $\alpha$  and  $\beta$  chromaticity values are transferred (assigned) to the target pixel while the original luminance value of the target pixel is retained. We tried various ratios of the weighting coefficients but found no significant differences in the final results. This agrees with the results of Welsh *et al.* [14].

The algorithm works best when the luminance distributions of the target and source images are locally similar. Its performance will degrade when the pixel histograms of the target and source luminance images are substantially different. For example, a light source image will be of little use when processing a dark target image. As a preconditioning step we therefore compute a linear remapping of the luminance distribution of the source image. This remapping is such that the first and second order statistics of the luminance distribution of the source image become equal to those of the target image. This helps to create a better correspondence between the luminance ranges of the target and source images, but does not alter the luminance values of the target image. More concretely, if  $l(p)$  is the luminance of a pixel in the source image, then we remap it as follows:

$$l(p) = \frac{\sigma_t}{\sigma_s} (l(p) - \mu_s) + \mu_t \quad (1)$$

where  $\mu_t$  and  $\mu_s$  represent the mean luminances, and  $\sigma_t$  and  $\sigma_s$  are the standard deviations of the luminance distributions, both taken with respect to the target and source luminance distributions respectively. This

mapping yields better results than the standard approach of histogram matching [1], which yields non-smooth mapping with undesirable side effects [4].

Since most of the visually distinctive variation between pixel values arises from luminance differences, we can limit the number of samples we use as source color pixels and still obtain a significant range of color variation in the image. This decreases the number of comparisons made for each pixel in the grayscale target image and thereby decreases computation time. We found that approximately 100 samples taken on a regular grid is sufficient. This agrees with Welsh *et al.* [14], who suggested to use about 200 samples on an irregular grid.

### 3 Examples

Figure 1 demonstrates the potential of the color transfer method. Here the method is applied to colorize the target grayscale photograph in Figure 1a using the color photograph of Figure 1b as the source image. Notice that the source image contains the same kind of materials as the target image (both images show grass, trees, and sky) but the actual type of the displayed objects is different in both cases (e.g. the trees are of a different kind). The resulting colorized image, shown in Figure 1c, has a natural appearance, with blue sky, dark green trees, and grass that is light greenish and yellowish.

To test the color transfer method for nightvision imagery we registered a variety of nighttime outdoor scenes with a single band image intensifier. These scenes, which are shown in Figures 2-6a, display a wide range of different materials (grass, several types of trees, pavement, roads, buildings, sky, water). The image intensifier was a monocular Mini N/SEAS device (assembled by ITL, Israel), equipped with a 2<sup>nd</sup> generation XD4 tube (produced by DEP, The Netherlands). A Sony Cyber-shot DSC-P5 3.2 megapixels digital camera was mounted behind the image intensifier to record the nightvision image. The illumination level varied between 0.1mlux and 1 mlux. The atmospheric conditions were clear. Object ranges varied from about 10 up to several hundreds of meters. The greenish image from the image intensifier was converted to a grayscale image for further processing.

Figures 2-6b show the color photographs that were selected to serve as source images. The source and target images represent entirely different scenes, but they both contain similar details (e.g. the source and target images may both show trees and buildings, although the actual type of trees and buildings are not the same).

To demonstrate the effectiveness of the procedure for grayscale nightvision imagery, we applied the color transfer method to colorize the central square region of the target images shown in Figures 2-6a, using the color images in Figures 2-6b as the source images. The resulting colorized images are shown in Figures 2-6c.

Corresponding daytime color photographs of these scenes are shown in Figures 2-6d. Notice that in all these examples the resulting colorized nightvision images (Figures 2-6c) have a natural daylight color appearance with color characteristics that are similar to those of the corresponding source images.

Figure 6 shows an example in which an oil painting called “Old Oak Tree”, by the Dutch painter Barend Cornelius Koekkoek (1803–1862), was adopted as the source image. The corresponding colorized nightvision image has an appearance which is quite natural, and its color characteristics resemble those of the oil painting.

## 4 Concluding remarks

We showed that a recently introduced method to transfer a source image’s color characteristics to a grayscale target image [14] can be used effectively to give single-band intensified nightvision imagery a natural day-time color appearance.

The procedure transforms both the source (color) image and the target (grayscale) image into a perceptually decorrelated color space. In this color space a best matching source pixel is determined for each pixel of the target image. The matching criterion uses the first order statistics of the luminance distribution in a small window around the source and target pixels. Once a best matching source pixel is found, its chromaticity values are assigned (transferred) to the target pixel while the original luminance value of the target pixel is retained. The only requirement of the method is that the compositions of the source and target scenes resemble each other (both images should contain similar details). We applied the method to a set of single-band grayscale intensified nightvision images. The resulting false color nightvision images have the appearance of the corresponding color source images. Such a full color representation of night-time scenes may be of great ergonomic value by making the interpretation (segmentation) of the displayed scene easier (more intuitive) for the observer.

The color transfer method only uses 4 values for each sample pixel in the color source image: the two chromaticity values of the sample pixel and the first order statistics (mean value and standard deviation) of the luminance distribution in a small local neighborhood ( $7 \times 7$  window) centered on the sample pixel. This implies that we only need a representative list of 4-tuples to apply a natural day-time color appearance to a single-band nightvision image. Hence, there is no need to actually store the source images from which the color information is derived. A system that is equipped with a look-up table of characteristic numbers for different types of backgrounds is sufficient to enable the observer to adjust the color mapping to the scene being viewed. A GPS based system may even automatically choose the

appropriate list of color samples, depending on the actual position and viewing direction of the nightvision device.

The method may be improved by using a different matching procedure. One could for instance use (a range of) different texture metrics, local similarity metrics [7, 8], or mutual information [6] in the matching process.

Since there evidently exists no unique mapping between luminance statistics and color distributions, the goal of transforming a grayscale intensified nightvision image into a full color day-time image can never be fully achieved. However, the fairly simple method employed here allows one (1) to settle for a single mapping that works satisfactorily in a large number of conditions (e.g. by selecting the color statistics of a generic representative scene), or (2) to adapt (optimise) the color mapping to the situation at hand (e.g. by selecting the color statistics that perfectly match the scene at hand).

## Appendix I: RGB to $la\beta$ transform

In this Appendix we present the RGB to  $la\beta$  transform [9]. This transform has recently been derived from a principal component transform of a large ensemble of hyperspectral images that represents a good cross-section of natural scenes. The resulting data representation is compact and symmetrical, and provides automatic decorrelation to higher than second order.

The actual transform is as follows. First the RGB tristimulus values are converted to device independent XYZ tristimulus values. This conversion depends on the characteristics of the display on which the image was originally intended to be displayed. Because that information is rarely available, it is common practice to use a device-independent conversion that maps white in the chromaticity diagram to white in RGB space and vice versa [3].

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 0.5141 & 0.3239 & 0.1604 \\ 0.2651 & 0.6702 & 0.0641 \\ 0.0241 & 0.1228 & 0.8444 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (2)$$

The device independent XYZ values are then converted to LMS space by

$$\begin{bmatrix} L \\ M \\ S \end{bmatrix} = \begin{bmatrix} 0.3897 & 0.6890 & -0.0787 \\ -0.2298 & 1.1834 & 0.0464 \\ 0.0000 & 0.0000 & 1.0000 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \quad (3)$$

Combination of (2) and (3) results in

$$\begin{bmatrix} L \\ M \\ S \end{bmatrix} = \begin{bmatrix} 0.3811 & 0.5783 & 0.0402 \\ 0.1967 & 0.7244 & 0.0782 \\ 0.0241 & 0.1288 & 0.8444 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (4)$$

The data in this color space shows a great deal of skew, which is largely eliminated by taking a logarithmic transform:

$$\begin{aligned}\mathbf{L} &= \log L \\ \mathbf{M} &= \log M \\ \mathbf{S} &= \log S\end{aligned}\quad (5)$$

The inverse transform from **LMS** cone space back to RGB space is as follows. First, the **LMS** pixel values are raised to the power ten to go back to linear LMS space. Then, the data can be converted from LMS to RGB using the inverse transform of Equation (4):

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 4.4679 & -3.5873 & 0.1193 \\ -1.2186 & 2.3809 & -0.1624 \\ 0.0497 & -0.2439 & 1.2045 \end{bmatrix} \begin{bmatrix} L \\ M \\ S \end{bmatrix} \quad (6)$$

Ruderman *e.a.* [9] presented the following simple transform to decorrelate the axes in the **LMS** space:

$$\begin{bmatrix} l \\ \alpha \\ \beta \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{3}} & 0 & 0 \\ 0 & \frac{1}{\sqrt{6}} & 0 \\ 0 & 0 & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & -2 \\ 1 & -1 & 0 \end{bmatrix} \begin{bmatrix} \mathbf{L} \\ \mathbf{M} \\ \mathbf{S} \end{bmatrix} \quad (7)$$

If we think of the **L** channel as red, the **M** as green, and **S** as blue, we see that this is a variant of a color opponent model:

$$\begin{aligned}\text{Achromatic} &\propto r + g + b \\ \text{Yellow-blue} &\propto r + g - b \\ \text{Red-green} &\propto r - g\end{aligned}\quad (8)$$

After processing the color signals in the  $l\alpha\beta$  space the inverse transform of Equation (7) can be used to return to the **LMS** space:

$$\begin{bmatrix} \mathbf{L} \\ \mathbf{M} \\ \mathbf{S} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & -1 \\ 1 & -2 & 0 \end{bmatrix} \begin{bmatrix} \frac{\sqrt{3}}{3} & 0 & 0 \\ 0 & \frac{\sqrt{6}}{6} & 0 \\ 0 & 0 & \frac{\sqrt{2}}{2} \end{bmatrix} \begin{bmatrix} l \\ \alpha \\ \beta \end{bmatrix} \quad (9)$$



Figure 1. Illustration of the color transfer method. (a) The target grayscale photograph in (a) is colorized using the color photograph in (b) as the source image. The resulting colorized image is shown in (c).

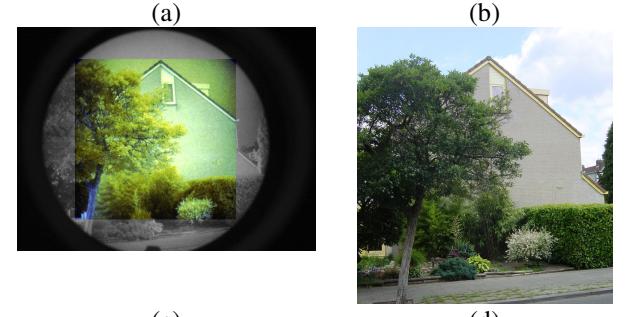


Figure 2. Application of the color transfer method to the intensified image in (a) using the color photograph in (b) as the source image yields the colorized image shown in (c). A daytime color photograph of the actual scene is shown in (d).



(a)



(b)



(c)



(d)

Figure 3. As Figure 2.



(a)



(b)



(c)



(d)

Figure 5. As Figure 2.



(a)



(b)



(c)



(d)

Figure 4. As Figure 2.



(a)



(b)



(c)



(d)

Figure 6. As Figure 2. The target image (b) is an oil painting called “Old Oak Tree”, by the Dutch master Barend Cornelius Koekkoek (1803–1862).

## References

- [1] K.R. Castleman. *Digital image processing*. Prentice-Hall, Englewood Cliffs, NJ, USA, 1996.
- [2] E.A. Essock, M.J. Sinai, J.S. McCarley, W.K. Krebs and J.K. DeFord. Perceptual ability with real-world nighttime scenes: image-intensified, infrared, and fused-color imagery. *Human Factors*, 41(3): 438-452.,1999.
- [3] M.D. Fairchild. *Color appearance models*. Addison Wesley Longman, Inc., Reading, MA, 1998.
- [4] A. Hertzmann, C.E. Jacobs, N. Oliver, B. Curless and D.H. Salesin. Image analogies. *Proceedings of the ACM SIGGRAPH 2001, Computer Graphics Proceedings, Annual Conference Series*, pages 341-346,ACM2001.
- [5] W.K. Krebs, D.A. Scribner, G.M. Miller, J.S. Ogawa and J. Schuler. Beyond third generation: a sensor-fusion targeting FLIR pod for the F/A-18. In B.V. Dasarathy, editors, *Sensor Fusion: Architectures, Algorithms, and Applications II*, pages 129-140,International Society for Optical Engineering,Bellingham, WA, USA,1998.
- [6] F. Maes, D. Vandermeulen and P. Suetens. Medical image registration using mutual information. *Proceedings of the IEEE*, 91(10): 1699-1722.,2003.
- [7] P. Rogelj and S. Kovacic. Local similarity measures for multimodal image matching. *Proceedings of the First International Workshop on Image and Signal Processing and Analysis*, pages 81-86,Pula, Croatia,2000.
- [8] P. Rogelj and S. Kovacic. Similarity measures for non-rigid registration. In M. Sonka and K. Hanson, editors, *Medical Imaging 2001: Image Processing*, pages 569-578,The International Society for Optical Engineering,Bellingham, WA,2001.
- [9] D.L. Ruderman, T.W. Cronin and C.-C. Chiao. Statistics of cone responses to natural images: implications for visual coding. *Journal of the Optical Society of America A*, 15(8): 2036-2045.,1998.
- [10] M.J. Sinai, J.S. McCarley, W.K. Krebs and E.A. Essock. Psychophysical comparisons of single- and dual-band fused imagery. In J.G. Verly, editors, *Enhanced and Synthetic Vision 1999*, pages 176-183,The International Society for Optical Engineering,Bellingham, WA,1999.
- [11] A. Toet and J.K. IJspeert. Perceptual evaluation of different image fusion schemes. In I. Kadar, editors, *Signal Processing, Sensor Fusion, and Target Recognition X*, pages 436-441,The International Society for Optical Engineering,Bellingham, WA,2001.
- [12] A. Toet, J.K. IJspeert, A.M. Waxman and M. Aguilar. Fusion of visible and thermal imagery improves situational awareness. In J.G. Verly, editors, *Enhanced and Synthetic Vision 1997*, pages 177-188,International Society for Optical Engineering,Bellingham, WA, USA,1997.
- [13] Varga, J.T. Evaluation of operator performance using true color and artificial color in natural scene perception. Report AD-A363036. Naval Postgraduate School, Monterey, CA, 1999.
- [14] T. Welsh, M. Ashikhmin and K. Mueller. Transferring color to greyscale images. *Proceedings of ACM SIGGRAPH2002*, ACM2002.
- [15] White, B.L. Evaluation of the impact of multispectral image fusion on human performance in global scene processing. Report AD-A343639. Naval Postgraduate School, Monterey, CA, 1998.