Available at
www.ComputerScienceWeb.com
POWERED BY SCIENCE DIRECT®

Information Fusion 4 (2003) 155–166

INFORMATION
FUSION

www.elsevier.com/locate/inffus

# Natural colour mapping for multiband nightvision imagery

Alexander Toet *

*TNO Human Factors, Kampweg 5, 3769 DE Soesterberg, The Netherlands*

Received 2 July 2002; received in revised form 20 September 2002; accepted 1 October 2002

**Abstract**

We present a method to give (fused) multiband night-time imagery a natural day-time colour appearance. For input, the method requires a false colour RGB image that is produced by mapping three individual bands (or the first three principal components) of a multiband nightvision system to the respective channels of an RGB image. The false colour RGB nightvision image is transformed into a perceptually decorrelated colour space. In this colour space the first order statistics of a natural colour image (target scene) are transferred to the multiband nightvision image (source scene). To obtain a natural colour representation of the multiband night-time imagery, the compositions of the source and target scenes should resemble each other to some degree. The inverse transformation to RGB space yields a nightvision image with a day-time colour appearance. The luminance contrast of the resulting colour image can be enhanced by replacing its luminance component by a grayscale fused representation of the three input bands.


*Keywords:* Image fusion; Infrared; Intensified imagery; False colour; Pyramid

## 1. Introduction

Modern night-time cameras are designed to expand the conditions under which human observers can operate. A functional piece of equipment must therefore provide an image that leads to good perceptual awareness in most environmental and operational conditions (to ''Own the weather'' or ''Own the night''). The two most common night-time imaging systems either display emitted infrared (IR) radiation or reflected light, and thus provide complementary information of the inspected scene. A suitably combined or *fused* representation of IR and (intensified) visual imagery may enable an observer to construct a more complete mental representation of the perceived scene, resulting in a larger degree of situational awareness [26]. A false colour representation of fused night-time imagery that closely resembles a natural daylight colour image will help the observer by making scene interpretation more intuitive.

The rapid development of multi-band infrared and visual nightvision systems has led to an increased interest in colour fused ergonomic representations of multiple sensor signals [1–3,7,9,17,18,28–33]. Simply mapping multiple spectral bands of imagery into a three dimensional colour space already generates an immediate benefit, since the human eye can discern several thousand colours, whereas it can only distinguish about 100 shades of grey at any instance. Combining bands in colour space therefore provides a method to increase the dynamic range of a sensor system [6]. Experiments have convincingly demonstrated that appropriately designed false colour rendering of night-time imagery can significantly improve observer performance and reaction times in tasks that involve scene segmentation and classification [7,19,24,25,28,34]. However, inappropriate colour mappings may hinder situational awareness [11,24,28]. One of the main reasons seems to be the counter intuitive appearance of scenes rendered in artificial colour schemes and the lack of colour constancy [28]. Hence, an ergonomic colour scheme should produce night vision imagery with a natural appearance and with colours 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).

Reinhard et al. [14] recently introduced a method to transfer one image's colour characteristics to another. The method was designed to give synthetic images a natural appearance. Here we show that this method can be applied to transfer the natural colour characteristics of daylight colour imagery to fused multiband

* Tel.: +31-3463-56237; fax: +31-3463-53977.
*E-mail address:* toet@tm.tno.nl (A. Toet).

1566-2535/$ - see front matter © 2003 Elsevier B.V. All rights reserved.
doi:10.1016/S1566-2535(03)00038-1

nightvision images. 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 [16]. In this decorrelated colour space the first order statistics of natural colour images (target scenes) are transferred to the multiband nightvision images (source scenes). 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 colour image of the same scene that is being monitored can be used to derive an optimal colour mapping.

Here we apply the method of Reinhard et al. [14] to transfer the characteristics of natural daylight colour images to false colour fused night-time imagery. We demonstrate the effectiveness of the method for the combined (fused) display of visual (400–700 nm) and near infrared (700–900 nm) intensified low-light CCD images and thermal middle wavelength band (3–5 μm) infrared images. The results show that the method can be used effectively to give night-time imagery a day-time appearance. Reinhard's [14] colour transfer method is in fact a simplification of a more general method that employs a principal component analysis, and that applies to any type of scene (not only to natural scenes). We show that the method can also be applied to images representing man made objects by using a full principal component analysis.

## 2. Imagery

A variety of outdoor scenes, displaying several kinds of vegetation (grass, heather, semi shrubs, trees), sky, water, sand, vehicles, roads, and persons, were registered at night with a recently developed dual-band visual intensified (DII) camera (see below), and with a state-of-the-art thermal middle wavelength band (3–5 μm) infrared (IR) camera (Radiance HS). Both cameras had a field of view (FOV) of about $6\times6$ degrees.

The DII camera was developed by Thales Optronics (Delft, The Netherlands) and facilitated a two-colour registration of the scene, applying two overlapping bands covering the part of the electromagnetic spectrum ranging from visual to near infrared (400–900 nm). The short (visual) wavelength part of the incoming spectrum was mapped to the R channel of an RGB false colour composite image. The long (near infrared) wavelength band corresponds primarily to the spectral reflection characteristics of vegetation, and was therefore mapped to the G channel of the RGB false colour composite image. This approach utilises the fact that the spectral reflection characteristics of plants are distinctly different from other (natural and artificial) materials in the visual and near infrared range [12]. The spectral response of the long-wavelength (G) channel roughly matches that of a Generation III image intensifier system.

Images were recorded at various times of the diurnal cycle under various atmospheric conditions (clear, rain, fog, ...) and for various illumination levels (1 lux–0.1 mlux). Object ranges up to several hundreds of meters were applied. The images were digitized on-site (using a Matrox Genesis frame grabber, using at least 1.8 times oversampling).

The recorded DII and IR images were first registered through an affine warping procedure, using fiducial registration points that were recorded at the beginning of each session. After warping, corresponding pixels in images taken with the different cameras represent the same location in the recorded scene. Then, a fused false colour RGB image was produced by assigning the IR image to the (empty) B channel of the false colour DII image. Finally, patches displaying different types of scenic elements were selected and cut out from the resulting false colour fused images. These patches were deployed as test images in the rest of this study. They display either buildings, vehicles, water, roads, trees, heather or humans. These details were selected because their signature varies strongly among the different image modalities. The false colour fusion scheme results in images in which grass, trees and persons are displayed as greenish, and roads, buildings, and vehicles are brownish. Examples of the individual image modalities with their greyscale fused representation are shown in Figs. 1 and 2. For brevity we can only present two different scenes here. A large number of examples with scenes of different composition are given elsewhere [23].

## 3. Colour transfer

The false colour images resulting from the aforementioned fusion scheme have an unnatural colour appearance (e.g. the top image in respectively Figs. 4–6, and Figs. 7c and 8c). The aim of the present study is to give these images the appearance of normal daylight colour images. In this section we introduce a simple technique to transfer the colour characteristics from natural daylight imagery to false colour nightvision imagery. A similar method was recently introduced to enhance the colour representation of synthetic imagery [14].

The method is as follows. Let the input multiband nightvision image be the source image, and let a normal daylight colour photograph be the target image. First, the source and target image are both transformed to the *LMS* cone response space. The different bands of multisensor signals and day-time colour images are usually correlated. Since we want to be able to transfer the characteristics of day-time colour images to false colour fused night-time images we first need to trans-


Fig. 1. Night-time images (made in complete darkness) of a scene representing a sandy path, trees, and fences, and taken with respectively (a) the visual (400–700 nm) and (b) the near infrared (700–900 nm) bands of a double-band image intensifier, and (c) a 3–5 μm midwave infrared camera. (d) The graylevel fused combination of the three images (a–c). Notice that the person, who is standing behind the trees and close to the fence, and who is clearly visible in the midwave infrared image (c), is also well represented in the fused image.

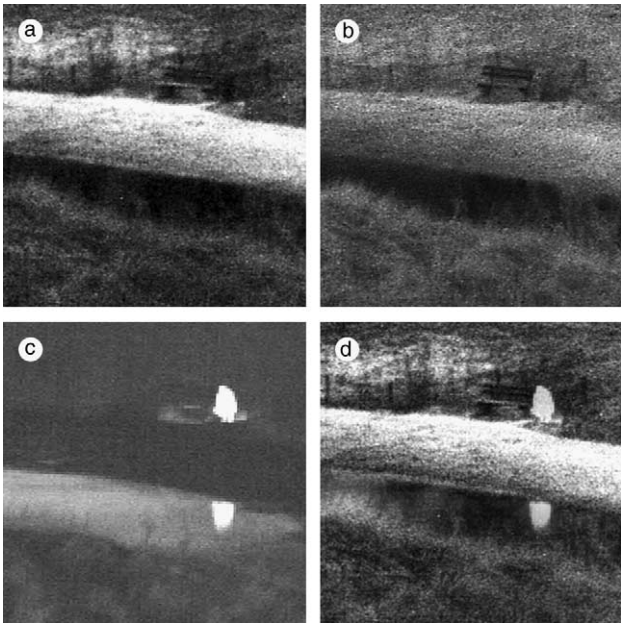


Fig. 2. As Fig. 1, for a scene representing a person crouching next to a bench near the lakeside. Notice the thermal reflection of the person in the water in the midwave infrared image (c).

form the input (multiband fused and day-time) colour imagery to a space which minimizes the correlation between channels. Therefore, through principal component analysis, we rotate the axes in the *LMS* cone space to achieve maximal decorrelation between the data points. Then, the mean and standard deviation of the source image is set equal to those of the target image. Finally, the source image is transformed back to RGB space for display. The result is a colour representation of the multiband nightvision image that resembles a normal daylight image.

In the following sections we first discuss the RGB to *LMS* transform. Then, we present a colour transfer method that employs a principal component transform in *LMS* cone space. Finally, we will show that for natural scenes the principal component transform can effectively be replaced by a fixed transform to $l\alpha\beta$ space [16]. This space 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.

### 3.1. RGB to LMS transform

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 [8].

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 0.5141 & 0.3239 & 0.1604 \\ 0.2651 & 0.6702 & 0.0641 \\ 0.0241 & 0.1228 & 0.8444 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \tag{1}$$

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

$$\begin{bmatrix} L \\ M \\ S \end{bmatrix} = \begin{bmatrix} 0.3897 & 0.6890 & -0.0787 \\ -0.2298 & 1.1834 & 0.0464 \\ 0.0000 & 0.0000 & 1.0000 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \tag{2}$$

Combination of (1) and (2) results in

$$\begin{bmatrix} L \\ M \\ S \end{bmatrix} = \begin{bmatrix} 0.3811 & 0.5783 & 0.0402 \\ 0.1967 & 0.7244 & 0.0782 \\ 0.0241 & 0.1288 & 0.8444 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \tag{3}$$

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

$$\begin{aligned} \boldsymbol{L} &= \log L \\ \boldsymbol{M} &= \log M \\ \boldsymbol{S} &= \log S \end{aligned} \tag{4}$$

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 Eq. (3):

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 4.4679 & -3.5873 & 0.1193 \\ -1.2186 & 2.3809 & -0.1624 \\ 0.0497 & -0.2439 & 1.2045 \end{bmatrix} \begin{bmatrix} L \\ M \\ S \end{bmatrix} \tag{5}$$

### 3.2. Transfer method I: principal component transform

The principal component transform [10,15,20] effectively rotates the *LMS* coordinate axes such that the pixel components are maximally decorrelated. The set of normalized eigenvectors of the covariance matrix of the set of pixel values, arranged in order of increasing eigenvalues, constitute the column vectors of the corresponding rotation matrix. Let $R_t$ be the rotation matrix that decorrelates the target pixels. The pixel values of the source and target images in this new coordinate system are then respectively given by

$$\begin{bmatrix} \boldsymbol{L}'_s \\ \boldsymbol{M}'_s \\ \boldsymbol{S}'_s \end{bmatrix} = R_t \begin{bmatrix} \boldsymbol{L}_s \\ \boldsymbol{M}_s \\ \boldsymbol{S}_s \end{bmatrix} \tag{6}$$

and

$$\begin{bmatrix} \boldsymbol{L}'_t \\ \boldsymbol{M}'_t \\ \boldsymbol{S}'_t \end{bmatrix} = R_t \begin{bmatrix} \boldsymbol{L}_t \\ \boldsymbol{M}_t \\ \boldsymbol{S}_t \end{bmatrix} \tag{7}$$

where the indices 's' and 't' refer to the source and target images respectively.

First, the mean is subtracted from the data points:

$$\begin{aligned} \boldsymbol{L}^* &= \boldsymbol{L}' - \langle \boldsymbol{L}' \rangle \\ \boldsymbol{M}^* &= \boldsymbol{M}' - \langle \boldsymbol{M}' \rangle \\ \boldsymbol{S}^* &= \boldsymbol{S}' - \langle \boldsymbol{S}' \rangle \end{aligned} \tag{8}$$

Then, the source data points are scaled with the ratio of the standard deviations of the source and target images respectively:

$$\begin{aligned} \boldsymbol{L}_s^+ &= \frac{\sigma_t^{\boldsymbol{L}}}{\sigma_s^{\boldsymbol{L}}} \boldsymbol{L}_s^* \\ \boldsymbol{M}_s^+ &= \frac{\sigma_t^{\boldsymbol{M}}}{\sigma_s^{\boldsymbol{M}}} \boldsymbol{M}_s^* \\ \boldsymbol{S}_s^+ &= \frac{\sigma_t^{\boldsymbol{S}}}{\sigma_s^{\boldsymbol{S}}} \boldsymbol{S}_s^* \end{aligned} \tag{9}$$

After this transformation, the resulting data points have standard deviations that correspond to those of the target image. Before reconstructing the RGB representation the averages computed over the target image are added to the source image:

$$\begin{aligned} \boldsymbol{L}_s^{\oplus} &= \boldsymbol{L}_s^+ + \langle \boldsymbol{L}'_t \rangle \\ \boldsymbol{M}_s^{\oplus} &= \boldsymbol{M}_s^+ + \langle \boldsymbol{M}'_t \rangle \\ \boldsymbol{S}_s^{\oplus} &= \boldsymbol{S}_s^+ + \langle \boldsymbol{S}'_t \rangle \end{aligned} \tag{10}$$

After this transformation, the mean and standard deviation of the source image conform to those of the target image. The result is transformed back to RGB space via the inverse rotation $R_t^{-1}$, $\log LMS$, *LMS*, and *XYZ* colour space using Eq. (5).

### 3.3. Transfer method II: $l\alpha\beta$ transform

Ruderman et al. [16] recently derived a colour space, called $l\alpha\beta$, which effectively minimises the correlation between the *LMS* axes. This result was derived from a principal component transform to the logarithmic *LMS* cone space representation of a large ensemble of hyperspectral images that represented a good cross-section of natural scenes. The principal axes encode fluctuations along an achromatic direction ($l$), a yellow-blue opponent direction ($\alpha$), and a red-green opponent direction ($\beta$). The resulting data representation is compact and symmetrical, and provides automatic decorrelation to higher than second order.

Ruderman et al. [16] presented the following simple transform to decorrelate the axes in the *LMS* space:

$$\begin{bmatrix} l \\ \alpha \\ \beta \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{3}} & 0 & 0 \\ 0 & \frac{1}{\sqrt{6}} & 0 \\ 0 & 0 & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & -2 \\ 1 & -1 & 0 \end{bmatrix} \begin{bmatrix} \boldsymbol{L} \\ \boldsymbol{M} \\ \boldsymbol{S} \end{bmatrix} \tag{11}$$

If we think of the $\boldsymbol{L}$ channel as red, the $\boldsymbol{M}$ as green, and $\boldsymbol{S}$ as blue, we see that this is a variant of a colour opponent model:

$$\begin{aligned}&\text{Achromatic} \propto r+g+b\\ &\text{Yellow-blue} \propto r+g-b\\ &\text{Red-green} \propto r-g\end{aligned} \tag{12}$$

After processing the colour signals in the $l\alpha\beta$ space the inverse transform of Eq. (11) can be used to return to the $LMS$ space:

$$\begin{bmatrix}\boldsymbol{L}\\ \boldsymbol{M}\\ \boldsymbol{S}\end{bmatrix}=\begin{bmatrix}1 & 1 & 1\\ 1 & 1 & -1\\ 1 & -2 & 0\end{bmatrix}\begin{bmatrix}\frac{\sqrt{3}}{3} & 0 & 0\\ 0 & \frac{\sqrt{6}}{6} & 0\\ 0 & 0 & \frac{\sqrt{2}}{2}\end{bmatrix}\begin{bmatrix}l\\ \alpha\\ \beta\end{bmatrix} \tag{13}$$

The processing in the $l\alpha\beta$ space is similar to the processing applied in the previous section, and given by Eqs. (8)–(10). First, mean is subtracted from the source and target data points:

$$\begin{aligned}l^{*}&=l-\langle l\rangle\\ \alpha^{*}&=\alpha-\langle\alpha\rangle\\ \beta^{*}&=\beta-\langle\beta\rangle\end{aligned} \tag{14}$$

Then, the source data points are scaled with the ratio of the standard deviations of the source and target images respectively:

$$\begin{aligned}l'_{\mathrm{s}}&=\frac{\sigma_{\mathrm{t}}^{l}}{\sigma_{\mathrm{s}}^{l}}l_{\mathrm{s}}^{*}\\ \alpha'_{\mathrm{s}}&=\frac{\sigma_{\mathrm{t}}^{\alpha}}{\sigma_{\mathrm{s}}^{\alpha}}\alpha_{\mathrm{s}}^{*}\\ \beta'_{\mathrm{s}}&=\frac{\sigma_{\mathrm{t}}^{\beta}}{\sigma_{\mathrm{s}}^{\beta}}\beta_{\mathrm{s}}^{*}\end{aligned} \tag{15}$$

After this transformation the pixels comprising the multiband source image have standard deviations that conform to the target daylight colour image. Finally, in reconstructing the $l\alpha\beta$ transform of the multiband source image, instead of adding the previously subtracted averages, the averages computed for the target daylight colour image are added. The result is transformed back to RGB space via log$LMS$, $LMS$, and $XYZ$ colour space using Eqs. (13) and (5).

## 4. Optimising luminance contrast

When combining the different bands of a multiband nightvision system into a single colour display it is essential that the relevant contrast details of the individual bands are preserved in the final colour image, and that no spurious pattern elements (that may interfere with subsequent analysis) are introduced by the merging process. The multiband images used in this study were obtained by mapping visual (400–700 nm) and near infrared (700–900 nm) intensified low-light CCD images and thermal middle wavelength band (3–5 μm) infrared images to respectively the Red, Green and Blue channels of an RGB false colour image. Note that the contrast of an image detail may vary strongly among the different bands. In some conditions a detail may even be represented with opposite contrast in different spectral bands. The combination of the individual image bands into a single colour image may therefore significantly reduce the luminance contrast of an image detail. As a result, a detail that is clearly visible in the individual image bands may be much less visible in the final colour representation, due to a lack of luminance contrast.

The next two sections describe the procedure we used to preserve the luminance contrast (and hence the visibility) of perceptually relevant image details in the final colour representation of the multiband night-time images. First we will explain how we fused the individual sensor bands into a single grayscale representation that preserves all relevant contrast details of the individual bands. Then we will show how we assign this optimally fused graylevel image to the luminance component of the RGB colour images resulting from the application of Method I to the multiband night-time images.

### 4.1. Pyramid graylevel fusion

We used a pyramidal image fusion scheme [5,21,22,27] to combine the images of the visual and near-infrared intensified low-light CCD sensor and the thermal middle wavelength infrared sensor into a single grayscale representation. An image pyramid is a collection of images at different spatial scales that together represent the original source image. Such a multi-resolution image representation can be obtained through a recursive *reduction* of the input image, i.e. a combination of low-pass or band-pass filtering and decimation.

The popular Laplacian pyramid, introduced by Burt and Adelson [5], is a sequence of images in which each image is Laplacian filtered and subsampled copy of its predecessor. The construction of this pyramid is as follows. First, a Gaussian or low-pass pyramid is constructed. The original image is adopted as the bottom or zero-level $G_0$ of the Gaussian pyramid. Each node of pyramid level $i$ ($1 \leqslant i \leqslant N$, where $N$ is the index of the top level of the pyramid, i.e. the lowest resolution level) is obtained as a (Gaussian) weighted average of the nodes at level $i-1$ that are positioned within a $5\times5$ window centered on that node. Because of the reduction in spatial frequency content each image in the sequence can be represented by an array that has only half the dimensions of its predecessor. The process which generates each image in the sequence from its predecessor is called a REDUCE operation, since both the sample density and the resolution are decreased. Thus for $1 \leqslant l \leqslant N$ we have $G_l = \mathrm{REDUCE}[G_{l-1}]$, meaning

$$G_l(i,j)=\sum_{m,n=-2}^{2} w(m,n)G_{l-1}(2i+m,2j+n) \tag{16}$$

where $w$ represents a standard binomial Gaussian filter of $5\times5$ pixels extent [4]. A set of band-pass filtered

images $L_0, L_1, \ldots, L_{N-1}$, that correspond to Laplacian of difference of low-pass filtered images, can be obtained by taking the difference of successive levels of the Gaussian pyramid. Since these levels differ in sample density it is necessary to interpolate new values between the given values of the lower frequency image before it can be subtracted from the higher frequency image. Interpolation is achieved simply by defining the EXPAND operation as the inverse of the REDUCE operation. Let $G_{l,k}$ be the image obtained by applying EXPAND to $G_l$ $k$-times. Then

$$G_{l,0} = G_l \tag{17}$$

and

$$G_{l,k} = \mathrm{EXPAND}[G_{l,k-1}] \tag{18}$$

meaning

$$G_{l,k}(i,j) = 4 \sum_{m,n=-2}^{2} w(m,n) G_{l,k-1}\left(\frac{i+m}{2}, \frac{j+n}{2}\right) \tag{19}$$

where only integer coordinates contribute to the sum. The sequence $L_i$ is then defined by

$$L_i = G_i - \mathrm{EXPAND}[G_{i+1}] \quad \text{for } 0 \leqslant i \leqslant N-1 \tag{20}$$

and

$$L_N = G_N \tag{21}$$

Thus, every level is a difference of two levels in the Gaussian pyramid, making it equivalent to a convolution with a Laplacian-like band-pass filter. The Laplacian pyramid is a complete representation of the input image. $G_0$ can be recovered exactly by reversing the steps used in the construction of the pyramid:

$$G_N = L_N \tag{22}$$

and

$$G_i = L_i + \mathrm{EXPAND}[G_{i+1}] \quad \text{for } 0 \leqslant i \leqslant N-1 \tag{23}$$

The Laplacian image fusion scheme is a three step procedure [5,21,22]. First, a Laplacian pyramid is constructed from each of the source images. Next, a Laplacian pyramid is constructed for the composite image by selecting or combining corresponding nodes in the component pyramids. Finally, the composite or fused image is recovered from its pyramid representation through the EXPAND and add reconstruction procedure.

In this study we used a 7-level Laplacian pyramid [5], in combination with a maximum absolute contrast node (i.e. pattern element) selection rule. This procedure ensures that perceptually relevant image details from all individual bands are represented in the final grayscale fused image.

### 4.2. Restoring luminance contrast

The RGB colour images resulting from the application of Method I to the multiband night-time images were transformed to a HSV (hue, saturation, value) representation [13]. This transformation effectively decouples the color information (H and S) from the luminance information (V). Hue represents the dominant color as seen by an observer, saturation refers to the amount of dilution of the color with white light, and value defines the average brightness. The luminance component may therefore be processed independently of the image's colour information. We used this property to replace the luminance component of the HSV transformed multiband night-time image with the grayscale image that was obtained by combining the individual sensor images through the pyramidal fusion scheme described in the previous section. Finally, we transformed the resulting images back to an RGB representation. The result is a false colour representation of the three band nightvision imagery, with optimal luminance contrast (resulting from the application of the pyramidal grayscale fusion method), and with colour characteristics similar to a daylight photograph of the depicted scene (resulting from the application of Method I).

## 5. Examples

Figs. 3 and 4 illustrates the application of colour transfer Method I to the false colour RGB source images obtained by mapping respectively the visual (400–700 nm) and the near infrared (700–900 nm) intensified

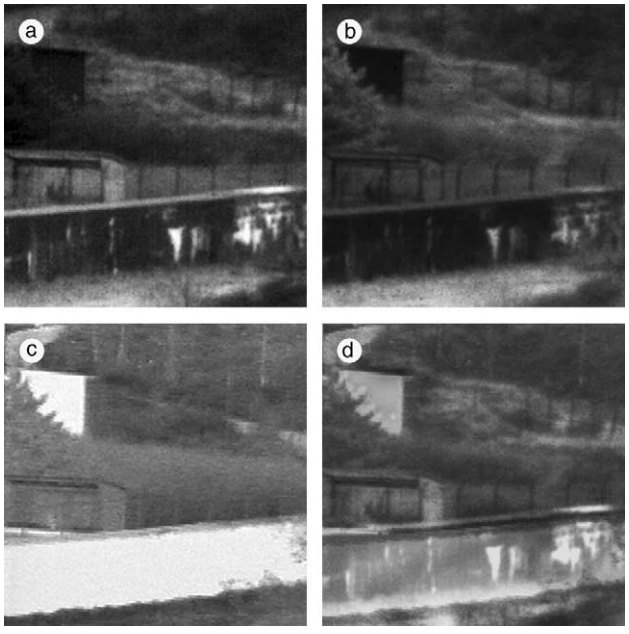


Fig. 3. As Fig. 1, for a scene representing a stone wall, grass, a shed, and a small building, with trees and fences in the background.

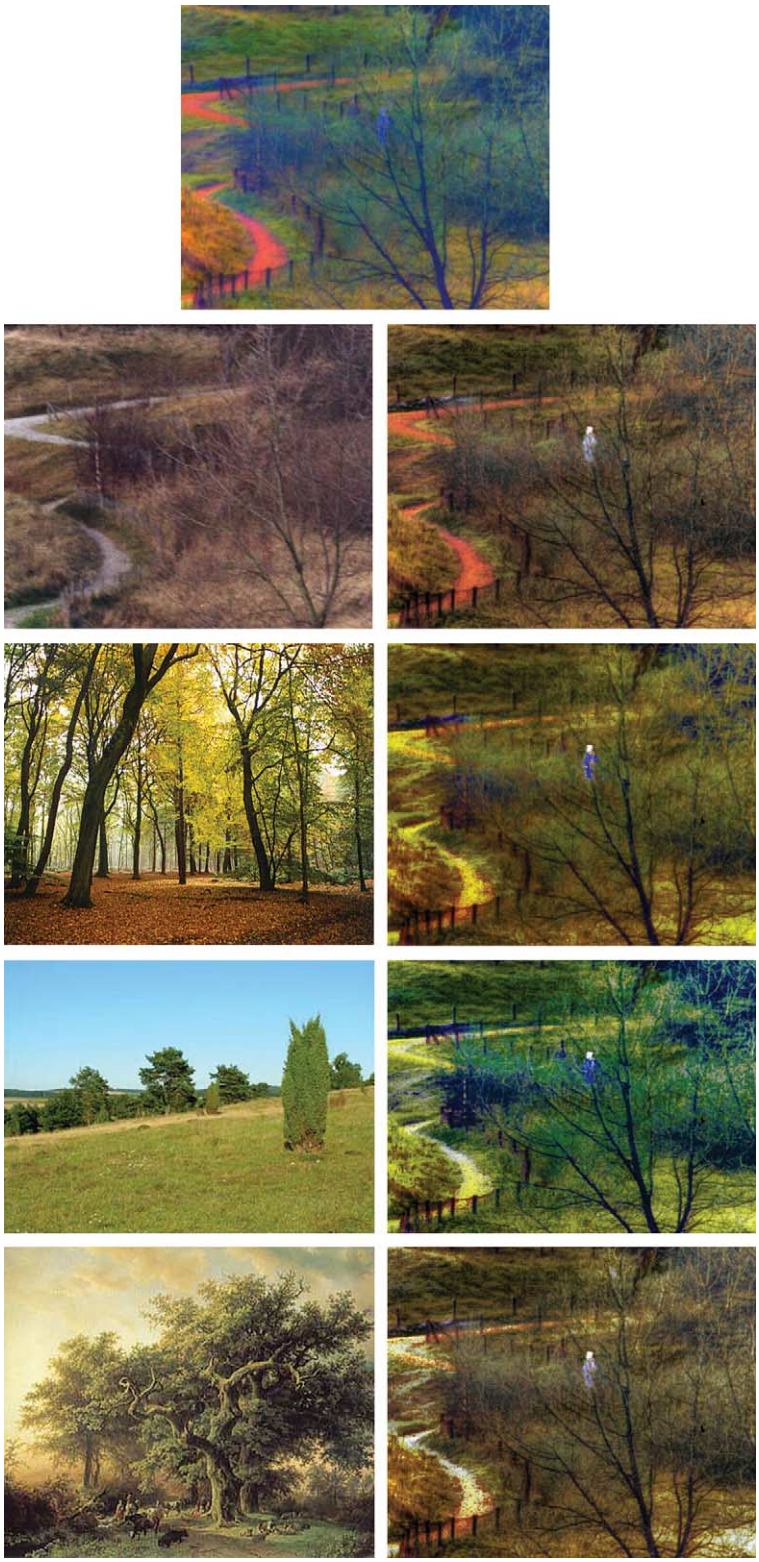

Fig. 4. Illustration of colour transfer Method I and the use of different target images. Top: false colour image obtained by mapping respectively the images a, b and c from Fig. 1 to the R, G and B channels of an RGB colour image. Notice the unnatural appearance of this image. Left column: the different target images. Right column: the corresponding results of the colour transfer Method I. The top image in the left column is a daylight photograph of the same scene (but without the person). The lower left target image is a painting called "Old Oak Tree", by the Dutch master Barend Cornelius Koekkoek (1803–1862). Notice the resemblance between the colour characteristics of the results of the colour mapping Method I (right column) and the target images (left column).

low-light CCD images and the thermal middle wavelength band (3–5 μm) infrared images from Figs. 1 and 2 to respectively the R, G, and B bands of a false colour representation (top). Results are shown for a number of different target images (left row). The first target image in these examples (top left row) is a daylight colour photograph of the same scene that is represented on top as a false colour RGB image, and that was recorded at night in full darkness with both the dual-band image intensifier and the thermal middle wavelength infrared camera. The colour in these target photographs is somewhat washed out because of atmospheric attenuation, since the viewing distance was about 800 m. Note that the colour of the target photograph returns in the processed false colour night-time imagery (top right column). Hence, the colour transfer method is effective in giving night-time imagery a day-time appearance. The results produced with Methods I and II are quite similar (for reasons of brevity not shown here, for an extensive comparison of both methods see [23]). This implies that for the type of imagery we investigated in this study the costly principal component analysis of Method I can be replaced by the simple matrix transform of Method II.

Figs. 3 and 4 also demonstrate the fact that the actual choice of the target image is not critical, as long as the colour distribution is to some extent similar to that of the source scene. The left column shows the different target images used in the colour transfer Method I. The right column represents the result of applying the colour transfer Method I to the image on top with the corresponding image in the left column as the target. The results show that the appearance of the final result resembles that of the target image in each case.

Fig. 3 also shows an example (5th row) in which an oil painting called "Old Oak Tree", by the Dutch painter Barend Cornelius Koekkoek (1803–1862), was adopted as the target image. The corresponding result has an appearance which is quite natural. However, the example shown in the 4th row of Fig. 3 shows that the method fails when the colour composition of the target and the source images are too dissimilar. In this case, the target image displays green grass and green trees with a bright blue sky in the background. The source image in contrast only shows brownish trees and vegetation. As a result, the transfer of statistics fails in this case.

We also applied the colour transfer Method I to some registered combinations of visual and infrared imagery that we had available from previous studies. The results are shown in Figs. 5 and 6. Since the corresponding daylight colour photographs were not available for these images we adopted some arbitrary colour images as targets.

Fig. 5 shows images of a night-time scene representing a building, a path, heather, and a person walking along the outside of a fence. Fig. 5a shows that the person is nearly invisible in the visual (400–700 nm)

Fig. 5. Illustration of colour transfer Method I and the use of different target images. Top: false colour image obtained by mapping respectively the images a, b and c from Fig. 2 to the R, G and B channels of an RGB colour image. Notice the unnatural appearance of this image. Left column: the different target images. Right column: the corresponding results of the colour transfer Method I. The top left column is a daylight photograph of the same scene (but without the person). The lower left target image is an arbitrary image of a craterlake with green trees in the foreground and greyish mountains in the background. Notice the resemblance between the colour characteristics of the results of the colour mapping Method I (right column) and the target images (left column).

intensified image, whereas the other details in the scene are clearly recognizable. The corresponding thermal midwave 3–5 μm image from Fig. 5b clearly shows the person, but it shows less detail of the remaining parts of the scene. Fig. 5c shows the false colour source image that was constructed by mapping the visual intensified image from Fig. 5a and the corresponding thermal midwave image from Fig. 5b to respectively the B and G channels of an RGB image representation (the R channel is set to zero or black). Fig. 5d shows an arbitrary colour image with similar content as the source image (grass, trees, building, path, etc.). Fig. 5e shows the result of applying colour transfer Method I to Fig. 5c with Fig. 5d as target. Fig. 5f illustrates the contrast

Fig. 6. Illustration of colour transfer Method I and the use of different target images. Top: false colour image obtained by mapping respectively the images a, b and c from Fig. 3 to the R, G and B channels of an RGB colour image. Notice the unnatural appearance of this image. Left column: the different target images. Right column: the corresponding results of the colour transfer Method I. The top left column is a daylight photograph of the same scene. The middle and lower left target images are arbitrary images of scenes containing buildings, grass, and trees. Notice the resemblance between the colour characteristics of the results of the colour mapping Method I (right column) and the target images (left column).

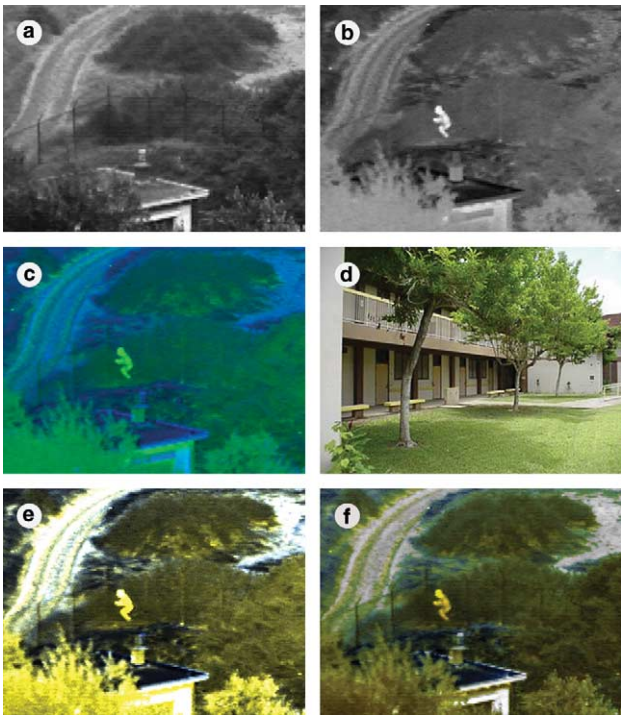


Fig. 7. Night-time images of a person walking along the outside of a fence. (a) The visual (400–700 nm) intensified image. (b) The corresponding thermal midwave 3–5 μm image of the same scene. (c) The false colour image obtained by mapping (a) and (b) to respectively the B and G channels of an RGB image representation (the R channel is set to zero or black). (d) Arbitrary colour image with similar colour distribution. (e) Result of colour transfer Method I applied to (c) with (d) as the target image. (f) Result of replacing the luminance component of (e) with the grayscale fused image of (a) and (b) (not shown here).

enhancement that is obtained by replacing the luminance component of Fig. 5e with the grayscale fused image of Fig. 5a and b (not shown here). Notice the improvement in the definition of the poles of the fence, the structure of the road, and the outlines of the person.

Fig. 6 shows images of a scene representing a battlefield with trucks, persons and a helicopter behind a smoke screen, and with a hill in the background. Fig. 6a shows that the persons and the helicopter disappear behind the smoke screen in the visual (400–700 nm) image, whereas the terrain details are clearly visible. The corresponding thermal longwave 8–12 μm image from Fig. 6b clearly shows the helicopter and the persons in the background, but it shows much less terrain detail. Fig. 6c shows the false colour source image that was constructed by mapping the visual grayscale image from Fig. 6a and the corresponding thermal image from Fig. 6b to respectively the B and G channels of an RGB image representation (again, the R channel is set to zero or black). Fig. 6d shows an arbitrary image with (probably) a similar colour distribution as the original scene (grass, sky, etc.). Fig. 6e shows the result of applying colour transfer Method I to Fig. 6c with Fig. 6d as target. Fig. 6f illustrates the contrast enhancement that is obtained by replacing the luminance component of Fig. 6e with the grayscale fused image of Fig. 6a and b (not shown here). Notice the improvement in the

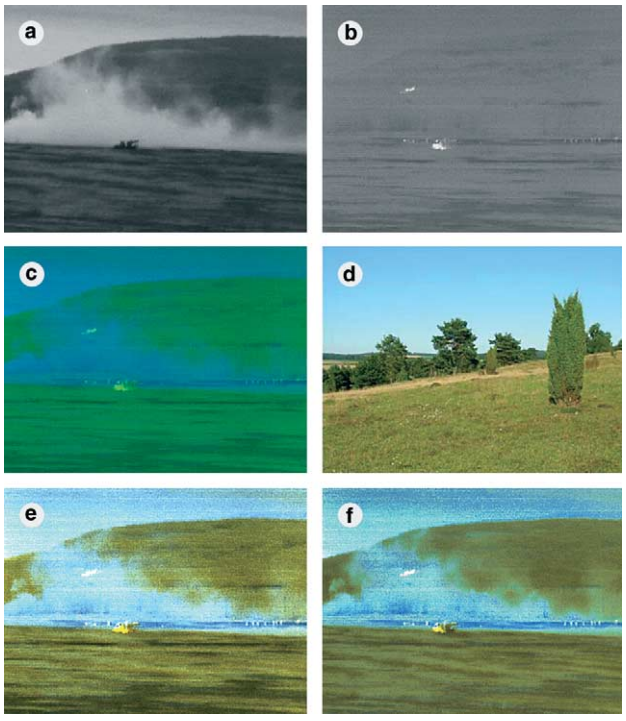


Fig. 8. Images of a battlefield with trucks, persons and a helicopter behind a smoke screen, with a hill in the background. (a) The visual CCD grayscale image. (b) The corresponding thermal 8–12 μm image of the same scene. (c) The false colour image obtained by mapping (a) and (b) to respectively the B and G channels of an RGB image representation (the R channel is set to zero or black). (d) Arbitrary colour image of a similar scene. (e) Result of colour transfer Method I applied to (c) with (d) as the target image. (f) Result of replacing the luminance component of (e) with the grayscale fused image of (a) and (b) (not shown here).

definition of the helicopter, the structure of the truck, and the outlines of the persons behind the smoke screen.

The examples given in Figs. 5 and 6 serve to demonstrate the general applicability of the colour transfer method to arbitrary combinations of (intensified) visual grayscale and thermal imagery. In both cases, the results show a colour image with natural characteristics, even though we had no corresponding target colour photograph available.

Figs. 5f and 6f illustrate the contrast enhancing effect of mapping the grayscale fused night-time imagery to the luminance component of the result of the colour transfer Method I.

The results show that this operation indeed enhances luminance contrast, resulting in images in which the depicted details are easier to perceive (i.e. they are easier to distinguish from their background).

## 6. Discussion

We showed that a recently introduced method to transfer one image's colour characteristics to another [14] can be used effectively to give multiband night-time imagery a natural day-time colour appearance. The contrast of the resulting colour imagery can be improved by mapping a grayscale fused representation of the individual image bands to the luminance component of the resulting colour images. The colour transfer method presented here can also be applied to remap the colour distribution of imagery resulting from existing colour fusion methods [1–3,7,9,17,18,28–33].

The colour transfer method employs a transformation to a principal component space. In this decorrelated colour space the first order statistics of natural colour images (target scenes) are transferred to the multiband nightvision images (source scenes). We applied the method to a set of RGB false colour night-time images recorded both with a dual band (visual and near infrared) image intensified low-light CCD camera (DII) and with a thermal middle wavelength band (3–5 μm) infrared (IR) camera. In each case, the resulting false colour night-time images adopted the appearance of the day-time colour images of the corresponding scene.

In this study we applied the method to source images with two or three spectral bands. When the input multiband image has more than three bands the method can also be applied. In that case a false colour RGB source image can be constructed by mapping the first three principal components of the multiband input image to the three channels of the RGB image. The colour mapping can then be applied to this false colour source image.

The colour transfer method only uses the first order statistics of natural colour images that are representative of the depicted scene. This implies that only six numbers (the three components of respectively the mean and standard deviation of the image components in *LMS* cone space) are required to apply a natural day-time colour appearance to multiband night-time imagery. Hence, there is no need to actually store the target images from which the colour information (the first order statistics) 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 colour mapping to the scene being viewed.

Night-time images recorded with an intensified low-light CCD camera and a thermal middle wavelength band (3–5 μm) infrared camera contain *complementary* information. This makes each of the individual image modalities only suited for specific observation tasks. In many operational conditions different nightvision systems are used side by side. By using a combined or fused display method the complementarity of the information in the image modalities can be fully exploited, thus enabling multiple observation tasks to be performed with a single night-time image representation. A full colour 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.

Since there evidently exists no one-to-one mapping between the temperature contrast and the spectral reflectance of a material, the goal of producing a night-time image, incorporating information from IR imagery, with an appearance identical to a colour day-time image can never be fully achieved. The 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 colour statistics of a generic representative scene), or (2) to adapt (optimise) the colour mapping to the situation at hand (e.g. by selecting the colour statistics that perfectly match the scene at hand).

## Acknowledgements

The authors thank Thales Optronics for providing the DII camera, and Hans Winkel (TNO-FEL), Jan Kees IJspeert and Nicole Schoumans (TNO-HF) for their help with the image registration.

This material is partly based upon work supported by the European Office of Aerospace Research and Development, Air Force Office of Scientific Research, Air Force Research Laboratory, under contract no. F61775-01-WE026, and by Senter, Agency of the Ministry of Economic Affairs of the Netherlands.

## References

[1] M. Aguilar, D.A. Fay, D.B. Ireland, J.P. Racamoto, W.D. Ross, A.M. Waxman, Field evaluations of dual-band fusion for color night vision, in: J.G. Verly (Ed.), Enhanced and Synthetic Vision 1999, The International Society for Optical Engineering, Bellingham, WA, 1999, pp. 168–175.

[2] M. Aguilar, D.A. Fay, W.D. Ross, A.M. Waxman, D.B. Ireland, J.P. Racamoto, Real-time fusion of low-light CCD and uncooled IR imagery for color night vision, in: J.G. Verly (Ed.), Enhanced and Synthetic Vision 1998, The International Society for Optical Engineering, Bellingham, WA, 1998, pp. 124–135.

[3] M. Aguilar, A.L. Garret, Biologically based sensor fusion for medical imaging, in: B.V. Dasarathy (Ed.), Sensor Fusion: Architectures, Algorithms, and Applications V, The International Society for Optical Engineering, Bellingham, WA, 2001, pp. 149–158.

[4] P.J. Burt, The pyramid as a structure for efficient computation, in: A. Rosenfeld (Ed.), Multiresolution Image Processing and Analysis, Springer, Berlin, GE, 1984, pp. 6–35.

[5] P.J. Burt, E.H. Adelson, Merging images through pattern decomposition, in: A.G. Tescher (Ed.), Applications of Digital Image Processing VIII, The International Society for Optical Engineering, Bellingham, WA, 1985, pp. 173–181.

[6] R.G. Driggers, K.A. Krapels, R.H. Vollmerhausen, P.R. Warren, D.A. Scribner, J.G. Howard, B.H. Tsou, W.K. Krebs, Target detection threshold in noisy color imagery, in: G.C. Holst (Ed.), Infrared Imaging Systems: Design, Analysis, Modeling, and Testing XII, The International Society for Optical Engineering, Bellingham, WA, 2001, pp. 162–169.

[7] E.A. Essock, M.J. Sinai, J.S. McCarley, W.K. Krebs, J.K. DeFord, Perceptual ability with real-world nighttime scenes: image-intensified, infrared, and fused-color imagery, Human Factors 41 (3) (1999) 438–452.

[8] M.D. Fairchild, Color Appearance Models, Addison Wesley Longman Inc., Reading, MA, 1998.

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

[10] E.L. Hall, Computer Image Processing, Academic Press, New York, USA, 1979.

[11] W.K. Krebs, D.A. Scribner, G.M. Miller, J.S. Ogawa, J. Schuler, Beyond third generation: a sensor-fusion targeting FLIR pod for the F/A-18, in: B.V. Dasarathy (Ed.), Sensor Fusion: Architectures, Algorithms, and Applications II, International Society for Optical Engineering, Bellingham, WA, USA, 1998, pp. 129–140.

[12] C.M. Onyango, J.A. Marchant, Physics-based colour image segmentation for scenes containing vegetation and soil, Image and Vision Computing 19 (8) (2001) 523–538.

[13] W.K. Pratt, Digital Image Processing, second ed., Wiley, New York, USA, 1991.

[14] E. Reinhard, M. Ashikhmin, B. Gooch, P. Shirley, Color transfer between images, IEEE Computer Graphics and Applications 21 (5) (2001) 34–41.

[15] J.A. Richards, Remote Sensing Digital Image Analysis, Springer Verlag, Berlin, 1986.

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

[17] J. Schuler, J.G. Howard, P. Warren, D.A. Scribner, R. Klien, M. Satyshur, M.R. Kruer, Multiband E/O color fusion with consideration of noise and registration, in: W.R. Watkins, D. Clement, W.R. Reynolds (Eds.), Targets and Backgrounds VI: Characterization, Visualization and the Detection Process, The International Society for Optical Engineering, Bellingham, WA, USA, 2000, pp. 32–40.

[18] D.A. Scribner, P. Warren, J. Schuler, Extending color vision methods to bands beyond the visible, in: Proceedings of the IEEE Workshop on Computer Vision Beyond the Visible Spectrum: Methods and Applications, Institute of Electrical and Electronics Engineers, 1999, pp. 33–40.

[19] M.J. Sinai, J.S. McCarley, W.K. Krebs, E.A. Essock, Psychophysical comparisons of single- and dual-band fused imagery, in: J.G. Verly (Ed.), Enhanced and Synthetic Vision 1999, The International Society for Optical Engineering, Bellingham, WA, 1999, pp. 176–183.

[20] P. Taylor, Statistical methods, in: M. Berthold, D.J. Hand (Eds.), Intelligent Data Analysis, Springer Verlag, Berlin, GE, 1999, pp. 67–127.

[21] A. Toet, Image fusion by a ratio of low-pass pyramid, Pattern Recognition Letters 9 (1989) 245–253.

[22] A. Toet, Hierarchical image fusion, Machine Vision and Applications 3 (1990) 1–11.

[23] A. Toet, Paint the night: applying daylight colours to nighttime imagery (Report TM-02-B006), TNO Human Factors, Soesterberg, The Netherlands, 2002.

[24] A. Toet, J.K. IJspeert, Perceptual evaluation of different image fusion schemes, in: I. Kadar (Ed.), Signal Processing, Sensor Fusion, and Target Recognition X, The International Society for Optical Engineering, Bellingham, WA, 2001, pp. 436–441.

[25] A. Toet, J.K. IJspeert, A.M. Waxman, M. Aguilar, Fusion of visible and thermal imagery improves situational awareness, in: J.G. Verly (Ed.), Enhanced and Synthetic Vision 1997, International Society

for Optical Engineering, Bellingham, WA, USA, 1997, pp. 177–188.

[26] A. Toet, J.K. IJspeert, A.M. Waxman, M. Aguilar, Fusion of visible and thermal imagery improves situational awareness, Displays 18 (1998) 85–95.

[27] A. Toet, J.J. Ruyven, J.M. Valeton, Merging thermal and visual images by a contrast pyramid, Optical Engineering 28 (1989) 789–792.

[28] J.T. Varga, Evaluation of operator performance using true color and artificial color in natural scene perception (Report AD-A363036), Naval Postgraduate School, Monterey, CA, 1999.

[29] A.M. Waxman, M. Aguilar, R.A. Baxter, D.A. Fay, D.B. Ireland, J.P. Racamoto, W.D. Ross, Opponent-color fusion of multi-sensor imagery: visible, IR and SAR, in: Proceedings of the 1998 Conference of the IRIS Speciality Group on Passive Sensors, 1998, pp. 43–61.

[30] A.M. Waxman et al., Solid-state color night vision: fusion of low-light visible and thermal infrared imagery, MIT Lincoln Laboratory Journal 11 (1999) 41–60.

[31] A.M. Waxman, J.E. Carrick, D.A. Fay, J.P. Racamato, M. Augilar, E.D. Savoye, Electronic imaging aids for night driving: low-light CCD, thermal IR, and color fused visible/IR, in: Proceedings of the SPIE Conference on Transportation Sensors and Controls, The International Society for Optical Engineering, Bellingham, WA, 1996.

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

[33] A.M. Waxman, A.N. Gove, D.A. Fay, J.P. Racamoto, J.E. Carrick, M.C. Seibert, E.D. Savoye, Color night vision: opponent processing in the fusion of visible and IR imagery, Neural Networks 10 (1) (1997) 1–6.

[34] B.L. White, Evaluation of the impact of multispectral image fusion on human performance in global scene processing (Report AD-A343639), Naval Postgraduate School, Monterey, CA, 1998.