Coloring Night-vision Imagery with Statistical Properties of Natural Colors by Using Image Segmentation and Histogram Matching

Yufeng Zheng, Bruce C. Hansen, Andrew M. Haun and Edward A. Essock
Dept. of Psychological & Brain Sciences, University of Louisville, Louisville, KY, USA

ABSTRACT

A natural color mapping method has been previously proposed that matches the statistical properties (mean and standard deviation) of night-vision (NV) imagery to those of a daylight color image (manually selected as the "target" color distribution). Thus the rendered NV image appears to resemble the target image in terms of colors. However, in this prior method the colored NV image may appear unnatural if the target image's "global" color statistics are too different from that of the night vision scene (e.g., it would appear to have too much green if much more foliage was contained in the target image). Consequently, a new "local coloring" method is presented in the current paper, and functions to render the NV image segment-by-segment by using a histogram matching technique. Specifically, a false-color image (source image) is formed by assigning multi-band NV images to three RGB (red, green and blue) channels. A nonlinear diffusion filter is then applied to the false-colored image to reduce the number of colors. The final grayscale image segments are obtained by using clustering and merging techniques. The statistical matching procedure is merged with the histogram matching procedure to assure that the source image more closely resembles the target image with respect to color. Instead of using a single target color image, the mean, standard deviation and histogram distribution of a set of natural scene images are used as the target color properties for each color scheme. Corresponding to the source region segments, the target color schemes are grouped by their scene contents (or colors) such as green plants, roads, ground/earth. In our experiments, five pairs of night-vision images were initially analyzed, and the images that were colored (segment-by-segment) by the proposed "local coloring" method are shown to be much more natural, realistic, and colorful when compared with those produced by the "global-coloring" method.

Keywords: color night vision, color visualization, histogram matching, image fusion, image segmentation, non-linear diffusion.

1. INTRODUCTION

Experiments have shown that appropriately designed false-color rendering of night-time imagery can significantly improve observer performance and reaction times in tasks that involve scene segmentation and classification.[1-6,25] Unfortunately, inappropriate color mappings have also been shown to be detrimental to human performance.[3,5,7,26] One of the reasons often cited for inappropriate or inconsistent color mapping is lack of physical color constancy.[5] Ideally, a color rendering method should render night vision imagery with a 'naturalistic', as well as stable, color appearance.

To address this problem, Reinhard et al.[8] recently introduced a method to transfer one image's color characteristics to another. The method was designed to give synthetic images a natural appearance. 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] In this decorrelated color space, the first order statistics of natural color images (target scenes) are transferred to the multiband night-vision images (source scenes). The only requirement of that method is that the source and target scenes possess similar chromatic properties (i.e., scenes were selected manually).

More recently, Toet[10] reported that Reinhard's method could be adapted to implement the natural color characteristics of daylight color imagery into multiband fused night-vision images. Essentially, Toet's natural color mapping method matches the statistical properties (i.e., mean and standard deviation) of night-vision (NV) imagery to that of a natural daylight color image (manually selected as the "target" color distribution). However, that particular color mapping method colors the image regardless of scene content, and thus the accuracy of the coloring is very much dependent on

how well the target and source images are matched. Specifically, the target image weights the local regions of the source image by the global color statistics of the target image, and thus will yield less naturalistic results if the target and source image are not matched with respect to structural content.

Based on Toet's framework (referred to here as "global coloring"), we presented a new "local coloring" method that addresses the above mentioned global color bias problem in order to produce colored NV images that appear more like realistic daylight imagery. The subsequent sections of this paper are organized as follows. First, a brief review of the global-coloring method is given. Next is a full description of the local-coloring method. Lastly, the experimental results and discussion are provided, followed by conclusions.

2. REVIEW OF GLOBAL-COLORING METHOD

The global-coloring method is summarized as follows. The aim of the global-coloring is to give night-vision images the appearance of normal daylight color images. A false-color image (source image) is first formed by assigning multi-band (two or three) NV images to three RGB channels. The false-color images usually have an unnatural color appearance (Figs. 1-3 (d)). Then, a true-color daylight image (target image) is manually selected with similar scenery (i.e., structural content) to the NV images. Both source and target images are transformed into a Luminance-Alpha-Beta ($l\alpha\beta$) color space, followed by calculating the global mean and standard deviation for each $l\alpha\beta$ plane. Next, a statistical-matching procedure is carried out between the source and target image. The matched source image is then transformed back to RGB space. Finally, the matched source image is transformed into HSV (Hue-Saturation-Value) space and the luminance component (i.e., the "value component" of the HSV decomposition) of the rendered source image is replaced with the fused NV image.

2.1 Color space transform

In this section the RGB to *LMS* (long-wave, medium-wave and short-wave) transform will be discussed first. Then an $l\alpha\beta$ space is introduced from which the resulting data representation is compact and symmetrical, and provides higher decorrelation than second order.

2.1.1 RGB to LMS transform

First, the RGB tristimulus values are converted to device independent XYZ tristimulus values[11], and then converted to *LMS*. The combination of conversions from *RGB* to *LMS* space is formulated by

$$\begin{bmatrix} L \\ M \\ S \end{bmatrix} = \begin{bmatrix} 0.3811 & 0.5783 & 0.0402 \\ 0.1967 & 0.7244 & 0.0782 \\ 0.0241 & 0.1288 & 0.8444 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}. \tag{1}$$

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

$$\begin{aligned} \boldsymbol{L} &= \log L, \\ \boldsymbol{M} &= \log M, \\ \boldsymbol{S} &= \log S. \end{aligned} \tag{2}$$

The inverse transform from *LMS* cone space back to RGB space is as follows. First, the ***LMS*** pixel values are raised to the tenth order in order to revert back into to linear *LMS* space. Then, the data can be converted from *LMS* to RGB using the inverse transform of Eq. (1):

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 4.4679 & -3.5873 & 0.1193 \\ -1.2186 & 2.3809 & -0.1624 \\ 0.0497 & -0.2439 & 1.2045 \end{bmatrix} \begin{bmatrix} L \\ M \\ S \end{bmatrix}. \tag{3}$$

2.1.2 LMS to lαβ transform

Ruderman et al.[9] recently utilized a color space, called $l\alpha\beta$ (Luminance-Alpha-Beta), which effectively minimizes 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 (α), and a red-green opponent direction (β). The resulting data representation is compact and symmetrical, and provides automatic decorrelation to higher than second order. Ruderman et al.[9] presented the following simple transform to decorrelate the axes in the ***LMS*** space:

$$\begin{bmatrix} l \\ \alpha \\ \beta \end{bmatrix} = \begin{bmatrix} 0.5774 & 0.5774 & 0.5774 \\ 0.4082 & 0.4082 & -0.8165 \\ 1.4142 & -1.4142 & 0 \end{bmatrix} \begin{bmatrix} \boldsymbol{L} \\ \boldsymbol{M} \\ \boldsymbol{S} \end{bmatrix}. \tag{4}$$

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:

$$\begin{aligned} &\text{Achromatic} \propto r + g + b, \\ &\text{Yellow-Blue} \propto r + g - b, \\ &\text{Red-Green} \propto r - g. \end{aligned} \tag{5}$$

After processing the color signals in the $l\alpha\beta$ space, the inverse transform of Eq. (4) can be used to return to the ***LMS*** space:

$$\begin{bmatrix} \boldsymbol{L} \\ \boldsymbol{M} \\ \boldsymbol{S} \end{bmatrix} = \begin{bmatrix} 0.5774 & 0.4082 & 0.3536 \\ 0.5774 & 0.4082 & -0.3536 \\ 0.5774 & -0.8165 & 0 \end{bmatrix} \begin{bmatrix} l \\ \alpha \\ \beta \end{bmatrix}. \tag{6}$$

2.2 Statistics matching

A simple technique to transfer the color characteristics from natural daylight imagery to false color night-vision imagery is introduced in this subsection. First, the mean is subtracted from the source data points. Then, the source data points are scaled with the ratio of the standard deviations of the source and target images respectively:

$$\boldsymbol{I}_C^k = (\boldsymbol{I}_S^k - \boldsymbol{\mu}_S^k) \cdot \frac{\boldsymbol{\sigma}_T^k}{\boldsymbol{\sigma}_S^k} + \boldsymbol{\mu}_T^k, \qquad \text{for } k = \{\, l, \alpha, \beta \,\}, \tag{7}$$

where $\mathbf{I}_C$ is the colored image, $\mathbf{I}_S$ is the source (false-color) image in $l\alpha\beta$ space; $\boldsymbol{\mu}$ denotes the mean and $\boldsymbol{\sigma}$ denotes the standard deviation; the subscripts 'S' and 'T' refer to the source and target images, respectively; and the superscript 'k' is one of the color components $\{\, l, \alpha, \beta \}$.

After this transformation the pixels comprising the multiband source image have standard deviations that conform to the target daylight color image. Finally, in reconstructing the $l\alpha\beta$ transform of the multiband source image, instead of adding the previously subtracted means, the means computed from the target daylight color image are added. The result is transformed back to RGB space via Eqs. (6) and (3).

3. THE LOCAL-COLORING METHOD

Based on the framework of the global-coloring method as described in Section 2, we present a new local-coloring method that makes the colored NV images appear more like daylight imagery. The major points for this new method to achieve these improvements are as following. (a) The source image is rendered segment-by-segment (i.e., local-coloring) – specifically, the false color image is region-segmented by its color properties (corresponding to its contents). A nonlinear diffusion filtering process[20] is applied to the false-colored image to reduce the number of colors. A set of preliminary clusters are then formed by anisotropically dividing the three components of the diffused image in $l\alpha\beta$ space based on their histogram analysis. These clusters are then merged to produce final segments if their centroid distances (corresponding to the diffused image) in 3D $l\alpha\beta$ color space are less than a given threshold. The final segments can be further smoothed with some appropriate morphological processing such as opening and/or closing. (b) The statistical-matching procedure is merged and/or replaced with the histogram-matching in order to assure that the source image more closely resembles the target image with respect to chromaticity. (c) Instead of a single color image, the averaged mean, standard deviation and histogram of a large sample of natural scene images are used as the target color properties for each color scheme. Corresponding to the source region segments, the target color schemes are grouped by their contents (or colors) such as green plants, mountain, roads, sky, ground/earth, water, buildings, people, etc. (as listed in

Table 1). (d) The mapping between the source region segments and target color schemes can be done automatically by a training and classifying process (i.e., pattern recognition). Currently this pattern recognition portion of this algorithm is under development and thereby was carried out manually in the following experiments reported in Section 4.

3.1 Image segmentation

3.1.1 Non-linear diffusion

Nonlinear diffusion methods have been proven as powerful methods in the de-noising and smoothing of image intensities while retaining and enhancing edges. Such an image smoothing process can be summarized as a successive coarsening of any given image while certain structures in that image are retained on a fine scale. Nonlinear diffusion is closely connected to a specific kind of multiscale analysis referred to as scale-space,[13,14] and was first used for image smoothing with simultaneous edge enhancement.[15] Close connections to regularization methods have been discovered,[16] and other related nonlinear methods such as interactive segmentation have also entered computer vision fields.[17,18]

Nonlinear diffusion filtering was first introduced by Perona and Malik.[15] Although their method in its original formulation is now regarded as ill-posed, it has triggered extensive research; see references[19,20] for overviews. In the following sections, the focus will be on one of the earliest regularizations of nonlinear diffusion introduced by Catté *et al*.[21]

Let $\Omega := (0, a_1) \times \cdots \times (0, a_m)$ be our image domain in $\mathbf{R}^m$ and consider a (scalar) image $f(x) \in L^\infty(\Omega)$. Then a filtered image $u(x,t)$ of $f(x)$ is calculated by solving a nonlinear diffusion equation with the original image as the initial state, and homogeneous Neumann boundary conditions:

$$\partial_t u = \mathrm{div}[g(|\nabla u_\sigma|^2)\ \nabla u] \ \text{ on } \ \Omega \times (0,\infty), \tag{8}$$

$$u(x, 0) = f(x) \ \text{ on } \ \Omega, \tag{9}$$

$$\partial_n u = 0 \ \text{ on } \ \partial\Omega \times (0,\infty), \tag{10}$$

where n denotes the normal to the image boundary $\partial\Omega$; 'div' means the divergence that is the dot product of a vector and its gradient, i.e., $\mathrm{div}(\boldsymbol{\alpha}) = \nabla \cdot \boldsymbol{\alpha}$; and '$\nabla$' is a gradient operator, e.g., for a 2-D image $\nabla = (\partial/\partial x)\mathbf{i} + (\partial/\partial y)\mathbf{j}$, where **i** and **j** are unit directional vectors along x and y axes.

The "time" t is a scale parameter: larger values lead to simpler image representations. In order to reduce smoothing at edges, the diffusivity g is chosen as a decreasing function of the edge detector $|\nabla u_\sigma|^2$, where ∇u_σ is the gradient of a Gaussian-smoothed version of u:

$$\nabla u_\sigma := \nabla (K_\sigma * u), \tag{11}$$

where '*' denotes the convolution, and

$$K_\sigma := \frac{1}{(2\pi\sigma^2)^{m/2}} \exp(-\frac{|x|^2}{2\sigma^2}) . \tag{12}$$

We use the diffusivity

$$g(s^2) := \begin{cases} 1, & for\ s^2 = 0 \\ 1 - \exp\left[\dfrac{-c}{(s/\lambda)^8}\right], & for\ s^2 > 0 \end{cases} . \tag{13}$$

For such rapidly decreasing diffusivities, smoothing on both sides of an edge is much stronger than smoothing across it. This selective smoothing process prefers intraregional smoothing to interregional blurring. One can ensure that the flux $\Phi(s) := sg(s^2)$ is increasing for $|s| \le \lambda$ and decreasing for $|s| > \lambda$ by choosing $c \approx 3.315$. Thus, λ is a contrast parameter that separates low-contrast regions with (smoothing) forward diffusion from high-contrast locations where backward diffusion may enhance edges.[15] Other choices of diffusivities are possible as well, but experiments indicate that Eq. (13) may lead to more segmentation-like results than the functions used in Ref.[15]. The parameter $\sigma > 0$ makes the filter insensitive to noise at scales smaller than σ. It is also a regularization parameter which guarantees the process is well-posed: Catté *et al.*[21] have shown that their filter has a unique solution which is infinitely times differentiable for $t > 0$. During the whole evolution, the average grey value remains unaltered.

Nonlinear diffusion filtering is usually performed with explicit schemes. They are only stable for very small time steps (t

≤ 0.25), which leads to poor efficiency and limits their practical use.[22] Based on a recent discrete nonlinear diffusion scale-space framework, Weickert *et al.* present semi-implicit schemes which are stable for all time steps.[22] These novel schemes use an additive operator splitting (AOS), which guarantees equal treatment of all coordinate axes. They can be implemented easily in arbitrary dimensions, have good rotational invariance and reveal a computational complexity and memory requirement which is linear in the number of pixels. Examples demonstrate that, under typical accuracy requirements, AOS schemes are at least ten times more efficient than the widely used explicit schemes. It has been shown that these AOS schemes are simple and efficient, do not require additional parameters, inherit important properties from the continuous equations, and are widely applicable.

One solution of color diffusions is to separately apply the nonlinear diffusion filtering (AOS) to RGB components of the false-colored NV image. The number of colors in the diffused image will be significantly reduced that will benefit the subsequent image segmentation procedures – clustering and merging.

3.1.2 Clustering and region merging

The diffused false-colored image is transformed into the *lαβ* color space. Each component of the diffused image is clustered in the *lαβ* space by individually analyzing their histogram. Specifically, for each component l, α or β, their mean (μ), standard deviation (σ), minimum (m_n) and maximum (m_x) are first calculated with their intensity values. Then clusters can be 'anisotropically' divided by using three different intervals based on which segment the current processing pixel belongs to. For each component $k = \{l, \alpha \text{ or } \beta\}$, the clustering interval can be formulated as following

$$\delta I^k(i,j) = \begin{cases} [(\mu^k - 3\sigma^k) - m_n^k]/n_m, & \text{if } I^k(i,j) \in [m_n^k, \mu^k - 3\sigma^k) \\ [(\mu^k + 3\sigma^k) - (\mu^k - 3\sigma^k)]/n_c, & \text{if } I^k(i,j) \in [\mu^k - 3\sigma^k, \mu^k + 3\sigma^k], \\ [m_x^k - (\mu^k + 3\sigma^k)]/n_m, & \text{if } I^k(i,j) \in (\mu^k + 3\sigma^k, m_x^k] \end{cases} \quad (14)$$

where two parameters are introduced, n_c and n_m, which denote the number of clusters in the central segment and two marginal segments (considering a histogram distribution), respectively. An ideal clustering will be resulted with Eq. (14) if the analyzed component image has a normal distribution. Otherwise, some boundary conditions need to be considered or the Eq. (14) should be revised, for instance, if $[(\mu^k - 3\sigma^k) - m_n^k] \le 0$. Anyway, the clustered image is still a false-colored image in the *lαβ* space.

Because the clustering is done by separately analyzing three components of the false-colored image, the clusters may not represent the color consistency. Thus, a relatively large number are assigned to n_m and n_c. Region merging is necessary to incorporate the fragmental clusters into meaningful regions in colors. Hereby a squared Euclidean distance between two clusters m and n is defined in the *lαβ* space,

$$d_{mn} = (\mu_m^l - \mu_n^l)^2 + (\mu_m^\alpha - \mu_n^\alpha)^2 + (\mu_m^\beta - \mu_n^\beta)^2, \quad (15)$$

where μ is the mean of a cluster in a particular component. Two clusters will be merged if $d_{mn} < T_d$, a specified threshold.

3.2 Histogram matching

Histogram matching (specification) is usually used to enhance an image when histogram equalization fails[12]. Given the shape of the histogram that we want the enhanced image to have, histogram matching can generate a processed image that has the specified histogram. In particular, by specifying the histogram of a target image (with true-color), a source image (with pseudo-color) resembles the target image in terms of histogram distribution after histogram matching. Histogram matching can be implemented with the following calculations as shown in Eqs. (16-18).[12] For the source image, we can calculate the transform

$$s_k = T(r_k) = (L-1) \cdot \sum_{j=0}^{k} \frac{n_j}{n}, \; k = 0,1,2,...,L-1, \quad (16)$$

where n is the total number of pixels in the image, n_j is the number of pixels that have gray level r_j, and L is the total number of possible gray levels in the image. Thus, an enhanced image is obtained by mapping each pixel with level r_k in the input image into a corresponding pixel with level s_k in the output image via Eq. (16). Actually, this is the process of histogram equalization. Obviously, $T(r_k)$is an increasing monotonic function. Similarly, for the target image or the given group of categorized images, we have

$$v_k = G(z_k) = (L-1)\cdot\sum_{j=0}^{k}\frac{n_j}{n},\ k = 0,1,2,...,L-1\ , \tag{17}$$

the same notations in Eq. (16) are applied here. Considering $v_k = s_k$ for histogram matching, the processed image is accordingly,

$$z_k = G^{-1}(v_k) = G^{-1}(s_k) = G^{-1}[T(r_k)],\ \ k = 0,1,2,...,L-1\ , \tag{18}$$

It is not difficult to find a discrete solution of inverse transform, $G^{-1}(v_k)$, because both $G^{-1}(v_k)$ and $T(r_k)$, can be simply implemented by look up tables (LUT).

4. EXPERIMENTAL RESULTS AND DISCUSSION

In our local-coloring method, the averaged mean, standard deviation and histogram of a large sample of natural scene images (in the *lαβ* space) are used as the target color properties for each color scheme other than a single color image. Corresponding to the source region segments, the target color schemes are grouped by their contents (or colors) such as green plants, mountains, roads, sky, ground/earth, water, buildings, etc. (as shown in Table 1). Our primary goal is to distinguish between the background objects in a NV image (e.g., for a pilot navigation). Of course, the foreground objects will be isolated or stand out even if they cannot be recognized. The mapping between the source region segments and target color schemes can be done automatically by a training/classifying process (i.e., pattern recognition). The priority values, the first digit of code, (the less the more important) give an order for the recognition task. Currently this pattern recognition portion of this algorithm is still under development. Thus, the mapping between segments and color schemes has to be done manually in the current set of experiments.

Table 1: The list of categorized objects – the mean and standard deviation of each group (in the *lαβ* space) is used for local-coloring NV images.

Objects	Code	# Images
Trees/Plants	101	31
Mountain	102	13
Roads	103	29
Sky	104	19
Sea/Water	105	19
Ground/Earth	106	13
Buildings	201	27
Clouds/Smoke	202	19
People	301	6
Others	500	39

In the following experiments, an iterative advanced-DWT method (*a*DWTi),[23,24] which is optimized with the established metric – image quality index (IQI), is used for multi-band NV image fusion (refer to Figs. 1-3 (c)). The advanced-DWT (*a*DWT) method incorporates principal component analysis (PCA) and morphological processing into a regular DWT fusion algorithm. Specifically, at each scale of the wavelet transformed images, a principle vector is derived from two input images and then applied to two of the images' approximation coefficients. For the detail coefficients, the larger absolute values are chosen and subjected to a neighborhood morphological processing procedure which serves to verify the selected pixels by using a "filling" and "cleaning" operation. Furthermore, the *a*DWT has two adjustable parameters – the number of DWT decomposition levels and the length of the selected wavelet that determinately affect the fusion result. An iterative *a*DWT fusion process (*a*DWTi) has been implemented and optimized with IQI. Previous experimental results showed that the iterative *a*DWT (*a*DWTi) achieved the best fusion compared to the pyramid or the regular DWT methods judged on both the IQI metric and visual inspection.[24]

The false colored images (Figs. 1-3 (d)) are obtained by assigning image intensified (II) or near infrared (NIR) images to blue channels, infrared (IR) images red channels, and providing averaged II/NIR and IR images to green channels. The diffused images (Figs. 1-3 (e)) are resulted with the non-linear diffusion algorithm (with AOS implementation) by setting the following parameter values: $\sigma = 1.0$, $\lambda = 2.0$, $t = 40$ and iterations = 8 (i.e., the number of repeated runs). The parameter values used in segmentation (refer to Figs. 1-3 (f)) are $n_c = 5$, $n_m = 2$ and $T_d = 0.9$.

To have a comprehensive view and a clear comparison, all experimental images are displayed on two adjacent pages: one grayscale page and a color page. Note that the same figure numbers on two pages represent different parts (subplots) of the same figure.

4.1 Imagery and results

Three pairs of image intensified (II) and infrared (IR) night-vision images taken outdoors (provided by TNO Human

Factors) were completely analyzed and are shown in Figs. 1-3 (a-b) (named NV1812, NV4911 and NV0140). The fused images (Figs. 1-3 (c)) were computed by the iterative advanced DWT method (*a*DWTi). The false colored images are shown in Figs. 1-3 (d). The false colors were reduced with the nonlinear diffusion algorithm (Figs. 1-3 (e)) that facilitated the subsequent segmentation. The segmentation was done in $l\alpha\beta$ space through clustering and merging operations (see Figs. 1-3 (f)). With this segment map, the histogram and/or statistical matching could be performed segment by segment (i.e., locally) in $l\alpha\beta$ space. The target color schemes corresponding to source region segments were recognized and mapped manually. The final locally colored images (segment-by-segment) are shown in Figs. 1-3 (g-h). From the visual examination of the resultant images, the merged matching (statistical- followed by histogram-matching) yielded a better colored image (e.g., more saturated colors) than the statistical-matching alone. To run the "global-coloring" algorithm, a natural scene image (see Fig. 4 (a)) was manually selected as the target image. From Figs. 4 (b-d) that were globally colored, it should be obvious that the mapped images are not as natural and colorful as those images produced by our "local-coloring" method, and they are also biased by the colors contained in the target image.

Two additional NV image pairs (named Lab05 and NVESD11, see Fig. 5 (a-b) and (e-f)) were analyzed here. "Lab05 consists of near infrared (NIR) and infrared (IR) images while "NVESD11" consists of intensified (II) and infrared (IR) night-vision images. The same procedures as described before were applied here but only the diffused images and final colored images (with the combination of statistical- and histogram-matching) are shown in Fig. 5 (c-d) and (g-h) because of limited space. Looking at Fig. 5 (d) or (h), a very realistic and impressive color image is observed, especially in the green plants, gray road, blue sky and blue river.

4.2 Discussion

Image segmentation is quite challenging because image contents vary dramatically from image to image. However, the combination of diffusion, clustering and region merging techniques make the segmentation very flexible. For example, a reasonable segmentation can be obtained by giving a set of proper parameter values (in Eqs. (8-15)), or even by redefining the distance between clusters (Eq. (15)).

Another challenging task is to recognize difference segments contained in NV images although, currently, this task is performed manually. However, it is feasible to automatically distinguish some limited or listed background objects as shown in Table 1. We will report our further research result on this recognition component in a subsequent paper.

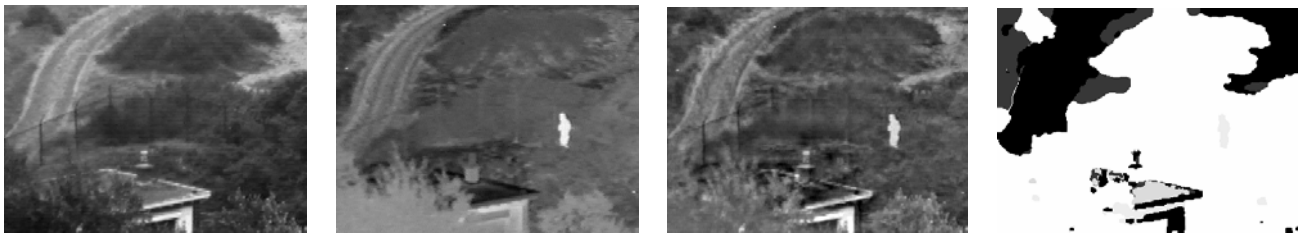
The histogram-matching method usually gives a more saturated color image than statistical-matching does. However, the histogram-matching procedure can cause various 'over-shoot' problems (such as over saturated colors) especially for a small image segment. Therefore, while matching small segments (e.g., identified by its area) in a false-colored NV image, we can either increase the matching region by merging similar segments together (i.e., by enlarging T_d), or use the statistical-matching only.

4.3 Future work

We plan to develop a classifier that can recognize the segments in NV images so that the mapping between NV segments and natural color schemes can be done automatically and also for a greater variety of types of scene. We also plan to analyze and colorize more NV images, followed by a series of performance-based psychophysical tasks in order to assess the functional utility of the proposed color mapping procedure.

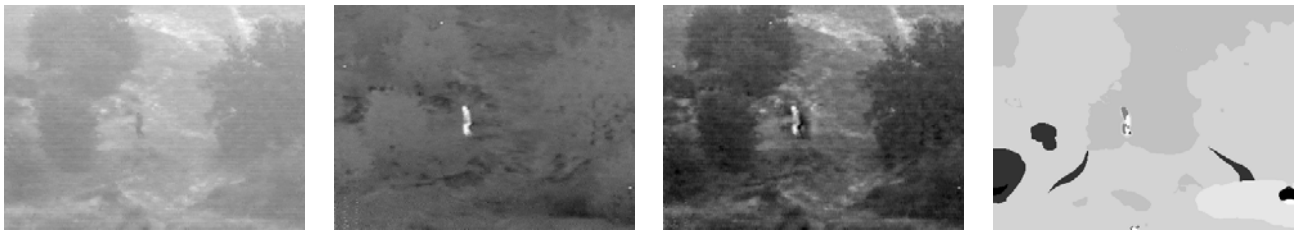
5. CONCLUSION

A novel "local-coloring" method is presented in this paper, which enables the rendering of NV imagery segment-by-segment by using statistical and/or histogram matching techniques. Furthermore, histogram matching enhances the color mapping process that makes the colored images more similar to the target color schemes. In addition, using the averaged color properties as target color schemes also reduces, on average, the possible bias as in the case when just one target color image were used. With this local-coloring method, experimental results indicated that the colored NV images are much more colorful and realistic compared with those produced by the global-coloring method.



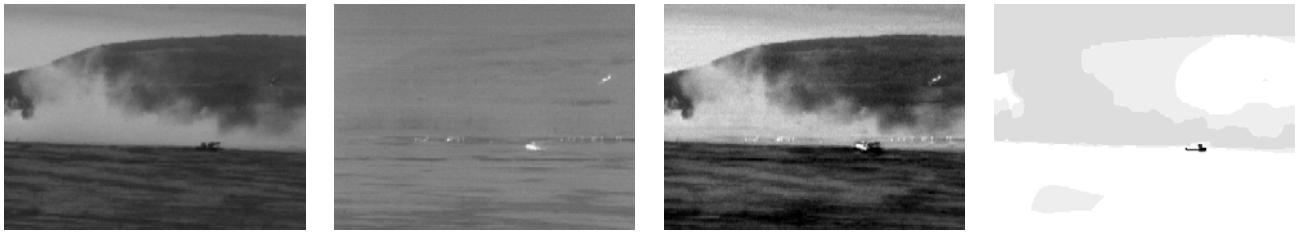
(a) Image A (II) (b) Image B (IR) (c) Fused by *a*DWTi (f) Segmented

Fig 1: Local-coloring (*segment-by-segment*) night-vision image pair NV1812 (360×270 pixels): (a) and (b) are two-band NV images, (c) is the fused image with (a) and (b), (d) is the false-colored image by using (a) and (b), (e) is the diffused one from image (d), five segments in (f) were derived from (e), and four color schemes (plants, roads, ground and people) were matched in final colored image (g) and (h). (Refer to color images (d), (e), (g) and (h) on next page.)



(a) Image A (II) (b) Image B (IR) (c) Fused by *a*DWTi (f) Segmented

Fig 2: Local-coloring (*segment-by-segment*) night-vision image pair NV4911 (360×270 pixels): (a) and (b) are two-band NV images, (c) is the fused image with (a) and (b), (d) is the false-colored image by using (a) and (b), (e) is the diffused one from image (d), nine segments in (f) were derived from (e), and three color schemes (plants, ground and people) were matched in final colored image (g) and (h). (Refer to color images (d), (e), (g) and (h) on next page.)



(a) Image A (II) (b) Image B (IR) (c) Fused by *a*DWTi (f) Segmented

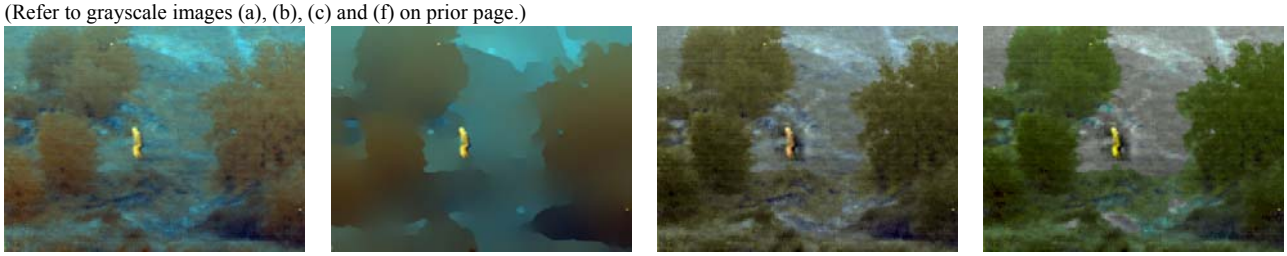
Fig 3: Local-coloring (*segment-by-segment*) night-vision image pair NV0140 (360×270 pixels): (a) and (b) are two-band NV images, (c) is the fused image with (a) and (b), (d) is the false-colored image by using (a) and (b), (e) is the diffused one from image (d), six segments in (f) were derived from (e), and three color schemes (plants, smoke and 'others') were matched in final colored image (g) and (h). (Refer to color images (d), (e), (g) and (h) on next page.)

(a) Lab05 – Image A (NIR) (b) Lab05 – Image B (IR) (e) NVESD11 – Image A (II) (f) NVESD11 Image B (IR)

Fig 5: Local-coloring (*segment-by-segment*) night-vision images: (a) and (b) are two-band NV images – Lab05 (312×245 pixels), and ten segments were formed from (c); (e) and (f) are two-band NV images – NVESD11 (531×401 pixels), and six segments were formed from (g). (Refer to color images (c), (d), (g) and (h) on next page.)

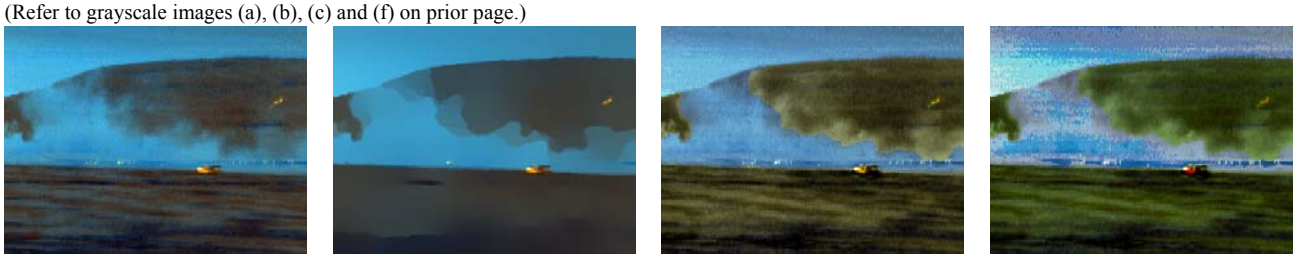
(d) False colored (e) Diffused (g) Colored by stat-match (h) Colored by hist- & stat-match

Fig 1: Local-coloring (*segment-by-segment*) night-vision image pair NV1812 (360×270 pixels): (d) is the false-colored image by using (a) and (b), (e) is the diffused one from image (d), and four color schemes (plants, roads, ground and people) were mapped/matched in final colored image (g) and (h). (Refer to grayscale images (a), (b), (c) and (f) on prior page.)



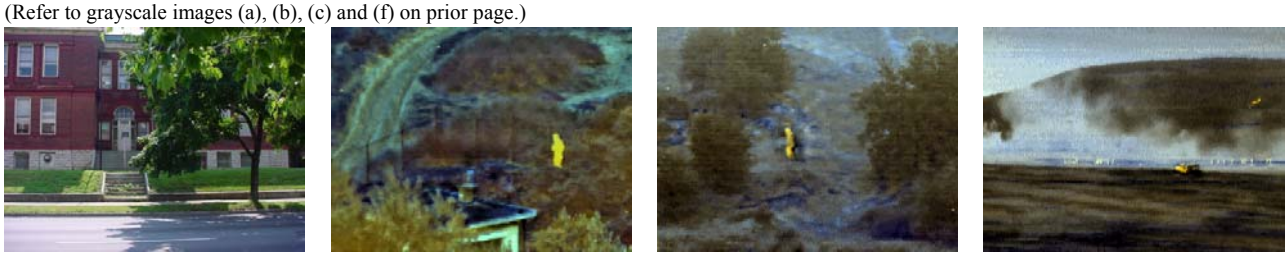
(d) False colored (e) Diffused (g) Colored by stat-match (h) Colored by hist- & stat-match

Fig 2: Local-coloring (*segment-by-segment*) night-vision image pair NV4911 (360×270 pixels): (d) is the false-colored image by using (a) and (b), (e) is the diffused one from image (d), and three color schemes (plants, ground and people) were mapped/matched in final colored image (g) and (h). (Refer to grayscale images (a), (b), (c) and (f) on prior page.)



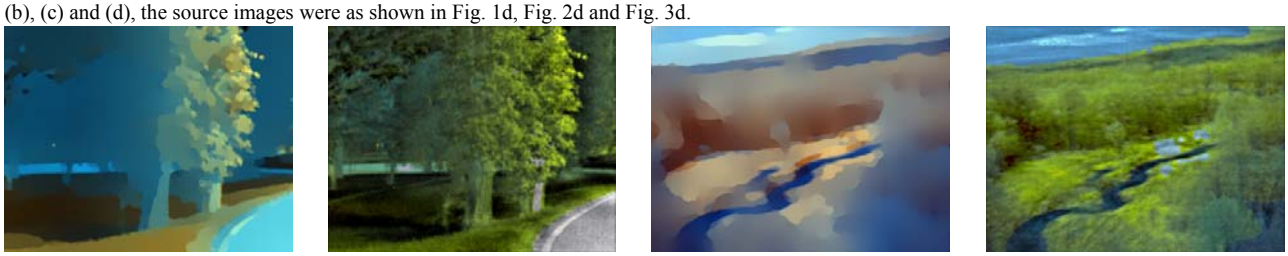
(d) False colored (e) Diffused (g) Colored by stat-match (h) Colored by hist- & stat-match

Fig 3: Local-coloring (*segment-by-segment*) night-vision image pair NV0140 (360×270 pixels): (d) is the false-colored image by using (a) and (b), (e) is the diffused one from image (d), and three color schemes (plants, smoke and 'others') were mapped/matched in final colored image (g) and (h). (Refer to grayscale images (a), (b), (c) and (f) on prior page.)



(a) Natural color image (b) Colored NV1812 (c) Colored NV4911 (d) Colored NV0140

Fig 4: Global-coloring night-vision images (stat-match only) by designating Fig. 4a as the target color image. Corresponding to the colored images in (b), (c) and (d), the source images were as shown in Fig. 1d, Fig. 2d and Fig. 3d.



(c) Lab05 — diffused (d) Lab05 — colored (g) NVESD11 — diffused (h) NVESD11 — colored

Fig 5: Local-coloring (*segment-by-segment*) night-vision images: (c) and (g) are the diffused false-colored images, (d) and (h) are colored images (by hist- & stat-match) by mapping to three (plants, road and 'others') and four (sky, plants, water and 'others') color schemes respectively. (Refer to grayscale images (a), (b), (e) and (f) on prior page.)

ACKNOWLEDGEMENTS

This work was supported by grant #N00014-03-1-0224 from the Office of Naval Research. We wish to thank Lex Toet and the TNO Human Factors Research Institute for providing night vision imagery and also Don Reago, Army Night Vision and Electronic Sensors Directorate, for providing an image pair.

REFERENCES

1. 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.
2. 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.
3. 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.
4. 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.
5. J.T. Varga, "Evaluation of operator performance using true color and artificial color in natural scene perception (Report ADA363036)", Naval Postgraduate School, Monterey, CA, 1999.
6. 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.
7. 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.
8. E. Reinhard, M. Ashikhmin, B. Gooch, P. Shirley, "Color transfer between images", *IEEE Computer Graphics and Applications* 21 (5) (2001) 34–41.
9. 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.
10. A. Toet, "Natural colour mapping for multiband nightvision imagery", *Information Fusion* 4 (2003) 155-166.
11. M.D. Fairchild, *Color Appearance Models*, Addison Wesley Longman Inc., Reading, MA, 1998.
12. Rafael C. Gonzalez, Richard E. Woods, *Digital Image Processing (Second Edition)*, Prentice Hall, Upper Saddle River, NJ, 2002.
13. Nielsen M., Johansen P., Olsen O. F. and Weickert J., eds., "Scale-Space Theories in Computer Vision", Vol. 1682 of *Lecture Notes in Computer Science*, Springer, Berlin, 1999.
14. ter Haar Romeny B., Florack L. and Koenderink J., eds., "Scale-Space Theory in Computer Vision", Vol. 1252 of *Lecture Notes in Computer Science*, Springer, Berlin, 1997.
15. Perona P. and Malik J., "Scale space and edge detection using anisotropic diffusion", *IEEE Transactions on Pattern Analysis and Machine Intelligence* 12 (1990), 629–639.
16. Scherzer O. and Weickert J., "Relations between regularization and diffusion filtering", *Journal of Mathematical Imaging and Vision* 12 (2000), 43–63.
17. Caselles V., Kimmel R. and Sapiro G., "Geodesic active contours", *International Journal of Computer Vision* 22 (1997), 61–79.
18. Kichenassamy S., Kumar A., Olver P., Tannenbaum A. and Yezzi A., "Conformal curvature flows: from phase transitions to active vision", *Archive for Rational Mechanics and Analysis* 134 (1996), 275–301.
19. ter Haar Romeny B. M., ed., "Geometry-Driven Diffusion in Computer Vision", Vol. 1 of *Computational Imaging and Vision*, Kluwer, Dordrecht, 1994.
20. Weickert J., *Anisotropic Diffusion in Image Processing*, Teubner, Stuttgart, 1998.
21. Catté F., Lions P.-L., Morel J.-M. and Coll T., "Image selective smoothing and edge detection by nonlinear diffusion", *SIAM Journal on Numerical Analysis* **32** (1992), 1895–1909.
22. Weickert J., ter Haar Romeny B. M. and Viergever M. A., "Efficient and reliable schemes for nonlinear diffusion filtering", *IEEE Transactions on Image Processing* 7 (1998), 398–410.

23. Zheng Y, Essock EA, and Hansen BC, "An Advanced Image Fusion Algorithm Based on Wavelet Transform – Incorporation with PCA and Morphological Processing", Proc. SPIE Vol. 5298, pp. 177-187, *Image Processing: Algorithms and Systems III*, Edward R. Dougherty, Jaakko T. Astola, Karen O. Egiazarian; Ed., January 2004.
24. Zheng Y, Essock EA, and Hansen BC, "An advanced DWT fusion algorithm and its optimization by using the metric of image quality index", *Optical Engineering* (2004) (In Press).
25. Waxman, Allen M., Gove, Alan N., Siebert, Michael C., Fay, David A., Carrick, James E., Racamato, Joseph P., Savoye, Eugene D., Burke, Barry E., Reich, Robert K., McGonagle, William H., Craig, David M., "Progress on color night vision: visible/IR fusion, perception and search, and low-light CCD imaging", Proc. SPIE Vol. 2736, pp. 96-107, *Enhanced and Synthetic Vision 1996*, Jacques G. Verly; Ed., May 1996.
26. Essock EA, McCarley JS, Sinai MJ, and DeFord JK, "Human Perception of Sensor-Fused Imagery", *Interpreting Remote Sensing Imagery: Human Factors (edited by Robert R. Hoffman, Arthur B. Markman)*, Lewis Publishers, Boca Raton, Florida, February 2001.