The Objective Evaluation Index (OEI) for Evaluation of Night Vision Colorization Techniques

Yufeng Zheng, Wenjie Dong, Genshe Chen and
Erik P. Blasch

Additional information is available at the end of the chapter

1. Introduction

A *night vision colorization* technique can produce colorized imagery with a naturalistic and stable color appearance by processing multispectral night vision (NV) imagery. The multispectral images typically include visual-band (e.g., red, green, and blue (RGB), or intensified) imagery and infrared imagery (e.g., near infrared (NIR) and long wave infrared (LWIR)). Although appropriately false-colored imagery is often helpful for human observers in improving their performance on scene classification and reaction time tasks (Waxman *et al.*, 1996; Essock *et al.*, 1999), inappropriate color mappings can also be detrimental to human performance (Toet et al., 2001; Varga, 1999). A possible reason is lack of physical color constancy. Another drawback with false coloring is that observers need specific training with each of the false color schemes so that they can correctly and quickly recognize objects; whereas with colorized nighttime imagery rendered with natural colors, users should be able to readily recognize and identify objects without any training.

There are several night vision (NV) colorization techniques developed in past decades. Toet (2003) proposed a NV colorization method that transfers the color characteristics of daylight imagery into multispectral NV images. Essentially, this color-mapping method matches the statistical properties (i.e., mean and standard deviation) of the NV imagery to that of a natural daylight color image (manually selected as the "target" color distribution). Zheng and Essock (2008) presented a "local coloring" method that can colorize the NV images more like daylight imagery by using histogram matching. The local-coloring method renders the multispectral images with natural colors segment by segment (i.e., "segmentation-based"), and also provides automatic association between the source and target images. Zheng (2011) recently introduced a *channel-based color fusion* method, which is fast enough

for real-time applications. Note that the term "color fusion" in this chapter refers to combing multispectral images into a color-version image with the purpose of resembling natural scenes. Hogervorst and Toet (2008 & 2012) recently proposed a new color mapping method using a lookup table (LUT). The LUT is created between a false-colored image (formed with multispectral NV images) and its color reference image (aiming at the same scene but taken at daytime). The colors in the resulting colored NV image resemble the colors in the daytime color image. This LUT-mapping method runs fast for real-time implementations. The LUT-mapping method and the statistic-matching method are also summarized in their recent paper (Toet & Hogervorst, 2012). Most recently Zheng (2012) developed a joint-histogram matching method for NV colorization.

The quality of colorized images can be assessed by subjective and/or objective measures. However, subjective evaluation normally costs time and resources. Moreover, the subjective evaluation methods cannot be readily and routinely used for real-time and automated systems. On the other hand, objective evaluation metrics can automatically and quantitatively measure the image qualities (Liu *et al.*, 2012 & Blasch *et al.*, 2008). Over the past decade, many objective metrics for grayscale image evaluations have been proposed (Alparone *et al.*, 2004; Wald *et al.*, 1997; Tsagaris & Anastassopoulos, 2006). However, the metrics for grayscale images cannot be directly extended to the evaluations of colorized images. Recently, some objective evaluations of color images have been reported in the literature. To objectively assess a color fusion method, Tsagaris (2009) proposed a color image fusion measure (CIFM) by using the amount of common information between the source images and the colorized image, and also the distribution of color information. Yuan *et al*. (2011) presented an objective evaluation method for visible and infrared color fusion utilizing four metrics: image sharpness metric, image contrast metric, color colorfulness metric, and color naturalness metric. In this chapter, we introduce an *objective evaluation index* (OEI) to quantitatively evaluate the colorized images. Given a reference (daylight color) image and several versions of the colorized NV images from different coloring techniques, all color images are first converted into International Commission on Illumination (CIE) LAB space, with dimension L for lightness and a and b for the color-opponent dimensions (Malacara, 2002). Then the OEI metric is computed with the four established metrics, phase congruency metric (PCM), gradient magnitude metric (GMM), image contrast metric (ICM), and color natural metric (CNM).

Certainly, a color presentation of multispectral night vision images can provide a better visual result for human users. We would prefer the color images resembling natural daylight pictures that we are used to; meanwhile the coloring process shall be efficient enough ideally for real time applications. In this chapter, we will discuss and explore how to objectively evaluate the image qualities of colorized images. The remainder of this chapter is organized as follows. Six NV colorization techniques are briefly reviewed in Section 2. Next, four image quality metrics are described in Section 3. A new colorization metric, *objective evaluation index* (OEI), is introduced in Section 4. The experiments and discussions are presented in Section 5. Conclusions are finally drawn in Section 6.

2. Overview of night vision colorization techniques

All color mapping methods described in Subsections 2.2-2.6 are performed in $l\alpha\beta$ color space. Thus the color space conversion from RGB to $l\alpha\beta$ must be done prior to color mapping, and then the inverse transformation to RGB space is necessary after the mapping. The details of $l\alpha\beta$ color space transformation are given elsewhere (Toet, 2003; Zheng & Essock, 2008). Certainly, two images, a *source* image and a *target* image, are involved in a color mapping process. The source image is usually a color fusion image (in Subsections 2.2-2.5) or a false-colored image (in Subsection 2.6); while the target image is normally a daylight picture containing the similar scene. The target image may have a different resolution as depicted in Subsections 2.2-2.5; however, the LUT described in Subsection 2.6 is established using the registered target (reference) image.

2.1. Channel-based color fusion (CBCF)

A fast color fusion method, termed as *channel-based color fusion* (CBCF), was introduced to facilitate realtime applications (Zheng, 2011). Notice that the term of "color fusion" means combing multispectral images into a color-version image with the purpose of resembling natural scenes. Relative to the "segmentation-based colorization" (Zheng & Essock, 2008), color fusion trades the realism of colors with speed.

The general framework of channel-based color fusion is as follows, (i) prepare for color fusion, preprocessing (denoising, normalization and enhancement) and image registration; (ii) form a color fusion image by properly assigning multispectral images to red, green, and blue channels; (iii) then fuse multispectral images (gray fusion) using aDWT algorithm (Zheng *et al.*, 2005); and, (iv) replace the *value* component of color fusion in HSV color space with the gray-fusion image, and finally transform back to RGB space.

In NV imaging, there may be several bands of images available, for example, visible (RGB), image intensified (II), near infrared (NIR), medium wave infrared (MWIR), long wave infrared (LWIR). Upon the available images and the context, we only discuss two of two-band color fusions of (II $\oplus$ LWIR), (NIR $\oplus$ LWIR). The symbol '$\oplus$' denotes the fusion of multiband images.

A color fusion of NIR and LWIR is formulated by,

$$
\begin{aligned}
F_R &= S_{[0,1.0]}^{[0.2,0.9]}\left(I_{LWIR}\right), & \text{(a)} \\
F_G &= S_{[0.1,I\_Gmax]}^{[0.2,1]}\left(I_{NIR}\right), & \text{(b)} \\
F_B &= S_{[0,1.0]}^{[0.1,0.7]}\left(\left[1.0\text{-}I_{LWIR}\right]\bullet I_{NIR}\right); & \text{(c)} \\
V_F &= Fus\left(I_{NIR},I_{LWIR}\right); & \text{(d)}
\end{aligned}
\tag{1}
$$

where $S_{[0.1,I\_Gmax]}^{[0.2,1]}$ denotes *piecewise contrast stretching* defined in Eq. (2) and $I_{\_Gmax}$ = $min([\mu_{NIR}+2\sigma_{NIR}],0.8)$, $min()$ is an operation to get the minimal number; $[1.0\text{-}I_{LWIR}]$ is to invert

LWIR image; symbol '•' means element-by-element multiplication; V_{F} is the *value* component of $\mathrm{F_C}$ in HSV space, *Fus*() means image fusion operation using *a*DWT algorithm (Zheng *et al.*, 2005). Although the limits given in contrast stretching are obtained empirically according to the night vision images that we had, it is viable to formulate the expressions and automate the fusion based upon a set of conditions (imaging devices, imaging time, and application location). Notice the transform parameters in Eqs. (1) were applied to all color fusions in our experiments (see Fig. 3d).

$$\mathbf{I}_S = S_{[I_{Min}, I_{Max}]}^{[L_{Min}, L_{Max}]} = (\mathbf{I}_0 - I_{\mathrm{Min}})\frac{L_{\mathrm{Max}} - L_{\mathrm{Min}}}{I_{\mathrm{Max}} - I_{\mathrm{Min}}} + L_{\mathrm{Min}}, \tag{2}$$

where $\mathbf{I}_S$ is the scaled image, $\mathbf{I}_0$ is the original image; I_{Min} and I_{Max} are the maximum and minimum pixel values in $\mathbf{I}_0$, respectively; L_{Min} and L_{Max} are the expected minimum and maximum pixel values in $\mathbf{I}_S$, respectively. After the image contrast stretching, $\mathbf{I}_S \in [L_{\mathrm{Min}}, L_{\mathrm{Max}}]$.

2.2. Statistic matching

A *statistic matching* (stat-match) is used to transfer the color characteristics from natural daylight imagery to false color night-vision imagery, which is formulated as:

$$I_C^k = (I_S^k - \mu_S^k) \cdot \frac{\sigma_T^k}{\sigma_S^k} + \mu_T^k, \ \ \text{for} \ \ k = \left\{ l, \ \alpha, \ \beta \right\}, \tag{3}$$

where I_C is the colored image, I_S is the source (false-color) image in *lαβ* space; μ denotes the mean and σ 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, α, β*}.

After this transformation, the pixels comprising the multispectral source image have means and standard deviations that conform to the target daylight color picture in *lαβ* space. The colored image is transformed back to the RGB space through the inverse transforms (Zheng & Essock, 2008; see Fig. 3e).

2.3. Histogram matching (HM)

Histogram matching (i.e., histogram specification) is usually used to enhance an image when histogram equalization fails (Gonzalez & Woods, 2002). Given the shape of the histogram that we want the enhanced image to have, histogram matching can generate a processed (i.e., matched) image that has the specified histogram. In particular, by specifying the histogram of a target image (with daylight natural colors), a source image (with false colors) resembles the target image in terms of histogram distribution after histogram matching.

Histogram matching (hist-match) can be implemented as follows. First, the *normalized cumulative histograms* of source image and target image (h_{S} and h_{T}) are calculated, respectively.

$$h_S = S(u_k) = (L-1) \cdot \sum_0^{L-1} \frac{n_k}{N},$$ (4)

where N is the total number of pixels in the image, n_k is the number of pixels that have gray level u_k, and L is the number of gray (bin) levels in the image. Typically, $L = 256$ for a digital image. But we can round the image down to m ($m < L$, e.g., $m = 64$) levels, and thus its histogram is called m-bin histogram. Clearly, $S(u_k)$ is a non-decreasing function. Similarly, $h_T = T(v_k)$ can be computed (see the "Target" curve in Fig. 1c).

Second, considering $h_S = h_T$ (i.e., $S(u_k) = T(v_k)$) for histogram matching, the matched image is accordingly computed as

$$v_k = T^{-1}[S(u_k)], \, k = 0,1,2,...,L-1.$$ (5)

It is straightforward to find a discrete solution of the inverse transform, $T^{-1}[S()]$ as both $T()$ and $S()$ can be implemented with look up tables.

Similar to the statistic matching (described in Subsection 2.2), histogram matching also serves for color mapping (see Fig. 3f) and is performed component-by-component in $l\alpha\beta$ space. Specifically, with each color component (say the α component, treated as a grayscale image) of a false-colored image, we can compute $S(u_k)$. With a selected target image, $T(v_k)$ can be calculated with regard to the same color component (say α). Using Eq. (5) the histogram matching can be completed regarding the color component (α). Histogram matching and statistic matching can be applied separately or jointly. When applied together, for instance, it is referred to as "statistic matching then histogram matching" (Zheng & Essock, 2008).

2.4. Joint histogram matching (JHM)

As described in Subsection 2.3, histogram matching is applied to each color component (plane) separately. It is highly possible to distort the color distributions of the mapped image (see Fig. 3f). To avoid color distortion, we introduce a new color mapping method, joint histogram matching (joint-HM).

In $l\alpha\beta$ space, α and β represent the color distributions; while l is the intensity component. A *joint histogram* (also called 2D histogram) of two color planes (α versus β) is calculated and then matched from source to target. The intensity component (l) is matched individually. The joint histogram is actually the joint (2D) intensity distribution of the two images, which is often used to compute the joint entropy (Hill & Batchelor, 2001) for image registration.

How to calculate the normalized cumulative histogram (denoted as h) from a 2D joint histogram (denoted as H_J) needs further discussion. To do histogram matching, h is expected to be a non-decreasing function. We propose to form a one-dimensional (1D) histogram by stacking H_J column-by-column and then perform histogram matching as defined in Eq. (10). Of course, to correctly index a 1D transform ($T^{-1}()$), the proper calculation of u_m (with m bins)

using two gray (bin) levels is expected. If H_J is computed as (β vs. α), its matching process is denoted as joint-HM($\beta\alpha$). Eventually, the histogram of the mapped image is sort of tradeoff between two histograms, "Source" and "Target". This is expected since we want no color distortion (i.e., preserving its own colors to some extent) during color mapping (see Fig. 3g).

2.5. Statistic matching then joint-histogram matching (SM-JHM)

The joint-HM can be applied together with statistic matching such as "stat-match then joint-HM", which usually result a better NV colorization. The statistic matching globally "paints" the image, while the joint-HM colors is more like the daylight picture in details (see Fig. 3h).

2.6. Lookup table (LUT)

Hogervorst and Toet (2008) proposed a color mapping method using a lookup table (LUT). The LUT is created using a false-colored image (formed with two-band NV images) and the reference (i.e., target) daylight image. This method yields a colored NV image similar to the daytime image in colors. The implementation of this LUT method is described as follows.

1. Create a false-colored image (of 3 color planes) by assigning LWIR image to R, NIR image to G plane, and zeros to B, respectively;

2. Build RG colormap (i.e., a 256×256 LUT) and convert the false-colored image to an indexed image (0 to 65535) associated with the RG colormap;

3. For all pixels in the indexed false-colored image whose index value equals 0:

    a. Locate all corresponding pixels in the reference (i.e., target) color image (that must be strictly aligned with the false-colored image);

    b. Calculate the averaged $l\alpha\beta$ values of those corresponding pixels and then convert them back to RGB values;

    c. Assign the RGB values to *index 0* in the lookup table;

4. Vary the index value from 2 to 65535 and repeat the processes described in Step 3. At the end, the LUT will be established.

Once the LUT is created, the LUT-based mapping procedure is simple and fast (see Fig. 3i), and thus can be deployed in realtime. However, the LUT creation thoroughly relies on the aligned reference image aiming at the same scene. Any misalignment, using a different reference color image, or coloring a different NV imagery (i.e., aiming at different direction), will usually result a poor colorization (see Fig. 5i).

3. Four image quality metrics

Three image quality metrics for grayscale images and one metric for color images are reviewed in this section. The color-related metrics are defined in the CIELAB space (Malacara, 2002)

specified by the International Commission on Illumination. The perceptually uniform CIELAB space consists of an achromatic luminosity component L^* (black-white) and two chromatic values a^* (green-magenta) and b^* (blue-yellow). The coordinates $L^*a^*b^*$ (CIE 1976) can be calculated using the CIE XYZ tristimulus values (Malacara, 2002).

3.1. Phase Congruency Metric (PCM)

The *phase congruency* (PC) model is also called the "local energy model" developed by Morrone *et al.* (1986). This model postulates that the features in an image are perceived at the points where the Fourier components are maximal in phase. Based on the physiological and psychophysical evidences, the PC theory provides a simple but biologically plausible model of how mammalian visual systems detect and identify the features in an image. PC can be considered as a significance measure of local structures in an image.

According to the definition of PC (Morrone *et al.*, 1986), there are many different implementations of PC map developed so far. A widely-used method developed by Kovesi (1999) is adopted in this chapter. Given a 1D image $f(x)$, M_n^e and M_n^o represent the even-symmetric and odd-symmetric filters at scale n, respectively. M_n^e and M_n^o form a quadrature pair: $e_n(x)$ and $o_n(x)$. Responses of the quadrature pair form a response vector:

$$\begin{bmatrix} e_n(x) \\ o_n(x) \end{bmatrix} = \begin{bmatrix} f(x) * M_n^e \\ f(x) * M_n^o \end{bmatrix}, \tag{6}$$

and the local amplitude at scale n is

$$A_n(x) = \sqrt{e_n^2(x) + o_n^2(x)}. \tag{7}$$

Let

$$F(x) = \sum_n e_n(x), \ \ H(x) = \sum_n o_n(x). \tag{8}$$

The one-dimensional (1D) *phase congruency metric* (PCM) can be computed as

$$PC(x) = \frac{\sqrt{F^2(x) + H^2(x)}}{\sum_n A_n(x) + \varepsilon}, \tag{9}$$

where ε is a small positive constant.

In order to calculate the quadrature pair of filters M_n^e and M_n^o, Gabor filters (Gabor, 1946) or log-Gabor filters (Mancas-Thillou & Gosselin, 2006) can be applied. In this chapter, we use log-Gabor filters (e.g., wavelets at scale $n = 4$) due to its following two features: (i) log-Gabor filters, by definition, have no direct current (DC) component; and (ii) the transfer function of the log-Gabor filter has an extended tail at the high frequency end, which makes it more capable to encode natural images than ordinary Gabor filters (Zhang *et al.*, 2011). The transfer function of a log-Gabor filter in the frequency domain is

$$G(\omega) = e^{-\frac{[\log(\omega/\omega_0)]^2}{2\sigma_r^2}}, \tag{10}$$

where ω_0 is the filter's center frequency and σ_r controls the filter's bandwidth.

To compute the PCM of two-dimensional (2D) grayscale images, we can apply the 1D analysis over several orientations and then combine the results according to some rules. The 1D log-Gabor filters described above can be extended to 2D ones by applying Gaussian function across the filter perpendicular to its orientation (Kovesi, 1999; Fischer *et al.*, 2007; Wang *et al.*, 2008). The 2D log-Gabor function has the following transfer function

$$G_2(\omega, \theta_j) = e^{-\frac{[\log(\omega/\omega_0)]^2}{2\sigma_r^2}} \cdot e^{-\frac{(\theta-\theta_j)^2}{2\sigma_\theta^2}}, \tag{11}$$

where $\theta_j = (j\pi)/(2J)$ and $j = 0, 1, 2,..., J-1$. J is the number of orientations and σ_θ determines the filter's angular bandwidth. By modulating ω_0 and θ_j and convolving G_2 with the 2D image, we get a set of responses at each point (x, y) as $e_{n,\theta_j}(x,y), o_{n,\theta_j}(x,y)]$. The local amplitude at scale n and orientation θ_j is

$$A_{n,\theta_j} = \sqrt{e_{n,\theta_j}^2(x,y) + o_{n,\theta_j}^2(x,y)}. \tag{12}$$

and the local energy along orientation θ_j is

$$E_{\theta_j} = \sqrt{F_{\theta_j}^2(x,y) + H_{\theta_j}^2(x,y)}. \tag{13}$$

where

$$F_{\theta_j}(x,y) = \sum_n e_{n,\theta_j}(x,y), \quad H_{\theta_j}(x,y) = \sum_n o_{n,\theta_j}(x,y). \tag{14}$$

The two-dimensional *PCM* at (x, y) is defined as

$$PC_{2D}(x,y) = \frac{\sum_j E_{\theta_j}(x,y)}{\sum_n \sum_j A_{n,\theta_j}(x,y) + \varepsilon}, \tag{15}$$

where ε is a small positive constant. It should be noted that $PC_{2D}(x,y)$ is a real number within [0,1]. The phase congruency metric (PCM) of an image is defined as

$$PCM = \frac{1}{MN} \sum_{(x,y)} PC_{2D}(x,y) = \frac{1}{MN} \sum_{(x,y)} \frac{\sum_j E_{\theta_j}(x,y)}{\sum_n \sum_j A_{n,\theta_j}(x,y) + \varepsilon}, \tag{16}$$

where $M{\times}N$ is the size of the image. The range of PCM is [0,1].

3.2. Gradient Magnitude Metric (GMM)

The image *gradient magnitude* (GM) is computed to encode contrast information. PC and GM are complementary and they reflect different aspects of the HVS (human visual system) in assessing the local image quality. The GM measures the sharpness of an image. The perception of sharpness is related to the clarity of detail of an image. Image gradient computation is a traditional topic in image processing. Gradient operators can be expressed by convolution masks. One of commonly used gradient operators is the Sobel operator. The partial derivatives of image $f(x, y)$, G_x and G_y, along horizontal and vertical directions using the Sobel operators are

$$G_x = \frac{1}{4}\begin{bmatrix} 1 & 0 & -1 \\ 2 & 0 & -2 \\ 1 & 0 & -1 \end{bmatrix} * f(x,y), \quad G_y = \frac{1}{4}\begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix} * f(x,y) \tag{17}$$

The GM of $f(x, y)$ at pixel (x, y) is defined as

$$G(x,y) = \sqrt{G_x^2 + G_y^2}. \tag{18}$$

The averaged GM over all pixels is called image *gradient magnitude metric* (GMM),

$$GMM = \frac{1}{MN} \sum_{x,y} G(x,y) = \frac{1}{MN} \sum_{x,y} \sqrt{G_x^2 + G_y^2}, \tag{19}$$

where $M \times N$ is the size of the image.

3.3. Image Contrast Metric (ICM)

An image with excellent contrast has a wide dynamic range of intensity level and appropriate intensity. Both the dynamic range of intensity level or the overall intensity distribution of the image can be provided by a histogram. A global contrast metric is proposed using the histogram character. The histogram of image with levels in the range [0, N-1] is a frequency-distribution function defined as the overall intensity distribution of an image

$$h(X_k) = n_k, \tag{20}$$

where X_k is the k-th level of input and n_k is the number of the pixels in the image having level x_k. The probability density function (PDF) is computed by

$$P(X_k) = n_k / n, \tag{21}$$

where n is the total number of the pixels of the image. The dynamic range value β is defined as

$$\beta = \sum_{k=0}^{L-1} S(X_k), \tag{22}$$

where

$$S(X_k) = \begin{cases} 1, & if \ P(X_k) > 0 \\ 0, & otherwise \end{cases}. \tag{23}$$

The dynamic range matrix α of histogram is defined as

$$\alpha = \frac{\beta}{2N - \beta}, \tag{24}$$

where $\alpha \in [0,1]$ and a larger value of α means a wider dynamic range in the histogram, which leads to better contrast. The image contrast metric is defined as

$$C = \alpha \sum_{k=0}^{N-1} \frac{X_k}{N} P(X_k). \tag{25}$$

For color images, the image contrast metric is determined by both gray contrast and color contrast. Because human perception is more sensitive to the luminance on contrast evaluation, we employ L^* channel in the CIELAB space to evaluate the color contrast. Thus, image contrast

is determined by the histogram of gray intensity and the histogram of color luminance L^* (see Fig. 1). For the gray intensity I, the gray contrast metric is defined as

$$C_g = \alpha_I \sum_{k=0}^{N_I - 1} \frac{I_k}{N} P(I_k),\qquad(26)$$

where α_I and $P(I_k)$ can be calculated as above for gray intensity. For L^* channel, the color contrast metric is

$$C_c = \alpha_c \sum_{k=0}^{N_{L^*} - 1} \frac{L_k^*}{N_{L^*}} P(L_k^*),\qquad(27)$$

where α_c and $P(L_k)$ can be calculated as above for L^* channel. The global *image contrast metric* (ICM) is defined as

$$ICM = \sqrt{\omega_1 C_g^2 + \omega_2 C_c^2},\qquad(28)$$

where ω_1 and ω_2 are the weights of C_g and C_c. For simplicity, we choose $\omega_1 = \omega_2 = 0.5$. *ICM* varies within [0,1]. The evaluation of image contrast metric of color fusion image is shown in Fig. 1.
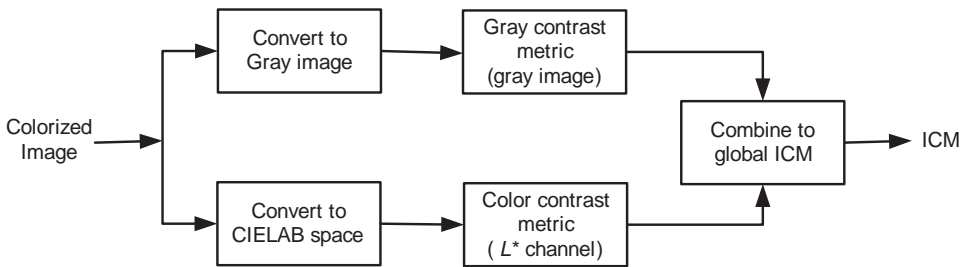

Figure 1. Diagram of calculation of the contrast metric.

3.4. Color Natural Metric (CNM)

Given a daylight image $f_1(x, y)$ and a colorized image $f_2(x, y)$, if a colorized image is similar to the daylight image then the colorized image is considered as of a good quality. Since a human is sensitive to hue in addition to luminance, we compare the a^* and b^* channels of the reference image with that of the colorized image using the gray relational analysis (GRA) theory (Ma *et al.*, 2005).

We first convert two images, f_1 and f_2, to $L^*a^*b^*$ space. $L_i(x,y)$, $a_i(x,y)$, and $b_i(x,y)$ are the $L^*a^*b^*$ values of f_i at pixel (x, y). The gray relation coefficient between a_1 and a_2 at pixel (x, y) is defined as

$$\xi_a(x,y) = \frac{\min_i\min_j |a_1^*(i,j)-a_2^*(i,j)|+0.5\max_i\max_j|a_1^*(i,j)-a_2^*(i,j)|}{|a_1^*(x,y)-a_2^*(x,y)|+0.5\max_i\max_j|a_1^*(i,j)-a_2^*(i,j)|+\varepsilon}, \tag{29}$$

where ε is a small positive constant.

The gray relation coefficient between b_1 and b_2 at pixel (x, y) is defined as

$$\xi_b(x,y) = \frac{\min_i\min_j |b_1^*(i,j)-b_2^*(i,j)|+0.5\max_i\max_j|b_1^*(i,j)-b_2^*(i,j)|}{|b_1^*(x,y)-b_2^*(x,y)|+0.5\max_i\max_j|b_1^*(i,j)-b_2^*(i,j)|+\varepsilon}. \tag{30}$$

In the definitions of $\xi_a(x,y)$ and $\xi_b(x,y)$, min() and max() are operated over whole image. However, it is possible that min() and max() are operated over a small neighborhood of (x, y).

The gray rational degrees of a^* and b^* information for two images are defined as

$$R_a = \sum_{(x,y)} \omega(x,y)\xi_a(x,y), \tag{31}$$

$$R_b = \sum_{(x,y)} \omega(x,y)\xi_b(x,y), \tag{32}$$

where $\omega(x,y)$ is the weight of the gray rational coefficient, which satisfies

$$\sum_{(x,y)} \omega(x,y) = 1. \tag{33}$$

For simplicity, we choose $\omega(x,y)=\frac{1}{M \times N}$ where M and N are the length of vectors x and y respectively.

The *color natural metric* (CNM) is defined as

$$CNM = \sqrt{R_a R_b}. \tag{34}$$

CNM varies within [0,1]; the larger the CNM, the more similar the two images.

4. Objective Evaluation Index (OEI)

With the four metrics defined in Section 3, a new *objective evaluation index* (OEI) is proposed to quantitatively evaluate the qualities of colorized images. Given the reference image f_1 and the colorized image f_2, the OEI is calcualted in two steps. First the local similarity maps of the two images are computed, and then the similarity maps are integrated into a single similarity score.

The two images are first converted into $L^*a^*b^*$ space. For L^* information, the PC maps are calculated and denoted as PC_1 and PC_2 for f_1 and f_2 images, respectively. The similarity measure between PC_1 and PC_2 at pixel (x, y) is defined as

$$S_{PC}(x,y) = \frac{2PC_1(x,y)PC_2(x,y) + K_1}{PC_1^2(x,y) + PC_2^2(x,y) + K_1},\qquad(35)$$

where K_1 is a positive constant. In practice, the determination of K_1 depends on the dynamic range of PC values. S_{PC} varies within [0,1]. Similarly, the similarity measure based on the two GM values is defined as

$$S_{G}(x,y) = \frac{2G_1(x,y)G_2(x,y) + K_2}{G_1^2(x,y) + G_2^2(x,y) + K_2},\qquad(36)$$

where K_2 is a positive constant. S_G varies within [0,1]. Then, $S_{PC}(x,y)$ and $S_G(x,y)$ are combined into one similarity measure, $S_L(x)$, as follows

$$S_L(x,y) = [S_{PC}(x,y)]^{\lambda_1}[S_G(x,y)]^{\lambda_2},\qquad(37)$$

where λ_1 and λ_2 are parameters to adjust the relative importance of PC and GM features.

With the aid of the similarity $S_L(x,y)$ at each pixel (x, y), the overall similarity between f_1 and f_2 can be calculated with the averaged $S_L(x,y)$ over all pixels. However, the image saliency (i.e., local significance) usually varies with the pixel location. For example, edges convey more crucial information than smooth areas. Specifically, a human is sensitive to phase congruent structures (Henriksson *et al.*, 2009), and thus a larger $PC(x, y)$ value between f_1 and f_2 implies a higher impact on evaluating the similarity between f_1 and f_2 at location (x, y). Therefore, we use $PC_{\max}(x,y) = max[PC_1(x,y), PC_2(x,y)]$ to weigh the importance of $S_L(x,y)$ in formulating the overall similarity. Accordingly, the objective evaluation index (OEI) between f_1 and f_2 is defined as follows

$$OEI = \left(\frac{\sum\limits_{(x,y)} PC_{\max}(x,y)S_L(x,y)}{\sum\limits_{(x,y)} PC_{\max}(x,y)}\right)^{\gamma_1} \times (S_{ICM})^{\gamma_2} \times (CNM)^{\gamma_3},\qquad(38)$$

where

$$PC_{max}(x,y) = \max[PC_1(x,y), PC_2(x,y)], \tag{39}$$

$$S_{ICM} = \frac{2ICM(f_1) \times ICM(f_2) + K_3}{ICM(f_1)^2 + ICM(f_2)^2 + K_3}, \tag{40}$$

where *CNM* is previously defined and K_3 and γ_i ($i = 1,2,3$) are positive constants. The diagram of calculating OEI is shown in Fig. 2. The range of *OEI* is [0,1]. The larger the *OEI* value of a colorized image is, the more similar (i.e., the better) the colorized image is to the reference image. *Error pooling* is the integration of methods with tradeoffs between γ_1, γ_2, and γ_3.
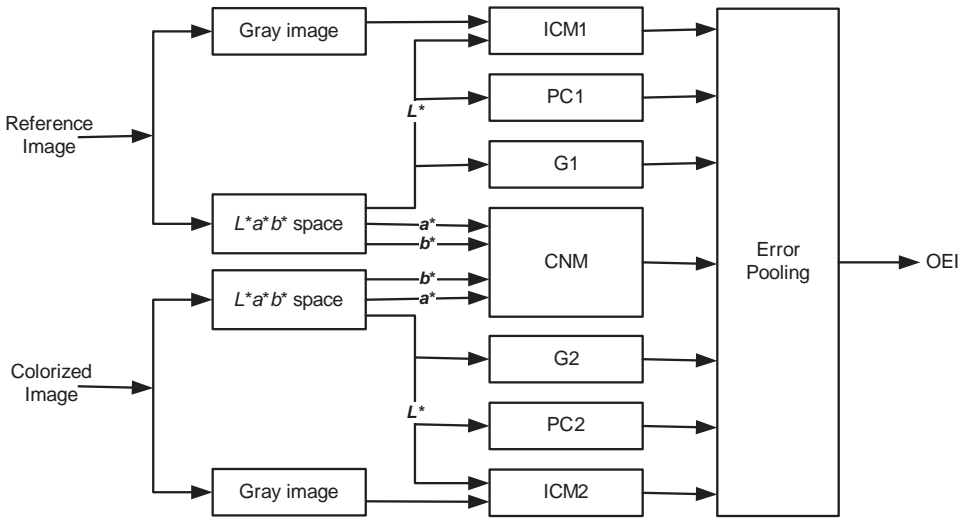
Figure 2. Diagram of calculating OEI in $L^*a^*b^*$ space.

γ_1, γ_2, and γ_3 are the weights of three components in the OEI metric. Selection of γ_i is critical for the OEI calculation. The values of γ_i are empirically decided, and the typical values of γ_1 and γ_2 are between 0.8~1.1 and γ_3 is between 0.05~0.2. K_i ($i = 1,2,3$) are constants to increase the metric stability. In our experiments presented in Section 6, we chose $\gamma_1=\gamma_2=1$, $\gamma_3=0.2$; $K_1 = 0.85$, $K_2 = 160$, $K_3 = 0.001$; and $\lambda_1=\lambda_2=1$.

5. Experimental results and discussions

In our experiments, five triplets of multispectral NV images (as shown in Figs. 3-7; collected at Alcorn State University), color RGB, near infrared (NIR) and long wave infrared (LWIR), were colorized by using six different coloring methods as described in Section 2. The three-band input images are shown in Figs. 3-7a, b and c, respectively. The image resolutions and its taken time are given in figure captions. The RGB images and LWIR images were taken by a FLIR SC620 two-in-one camera, which has LWIR camera (of 640×480 pixel original resolution and 7.5~13 μm spectral range) and an integrated visible-band digital camera (2048×1536 pixel original resolution). The NIR images were taken by a FLIR SC6000 camera (640×512 pixel original resolution and 0.9~1.7 μm spectral range). Two cameras (SC620 and SC6000) were placed on the same fixture and turned to aim at the same location. The images were typically captured during sunset time and dusk time during a fall season. One exception is shown in Fig. 7, which was taken at noon time.
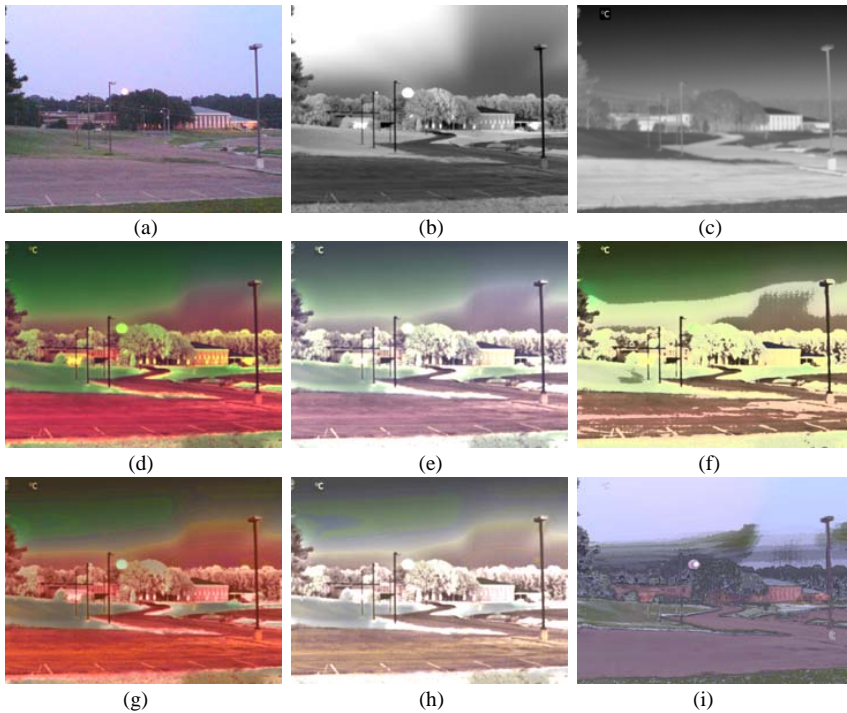

(a)　　　　　　　　(b)　　　　　　　　(c)

(d)　　　　　　　　(e)　　　　　　　　(f)

(g)　　　　　　　　(h)　　　　　　　　(i)

Figure 3. Night-vision coloring comparison (Case# AT008 – taken at sunset time; 640×480 pixels): (a-c) Color RGB, NIR, and LWIR images, respectively; (d-f) The colorized images using channel-based color fusion of (NIR⊕ LWIR), statistic-matching, and histogram-matching, respectively; (g-i) The colorized images using joint-HM, stat-match then joint-HM, and LUT-mapping, respectively. The settings in the color-mappings of (e-i) are source = (d) and target = (a). Notice that the contrasts of all color images were increased by 10%, and the brightness of (a) and (i) were increased by 10%.
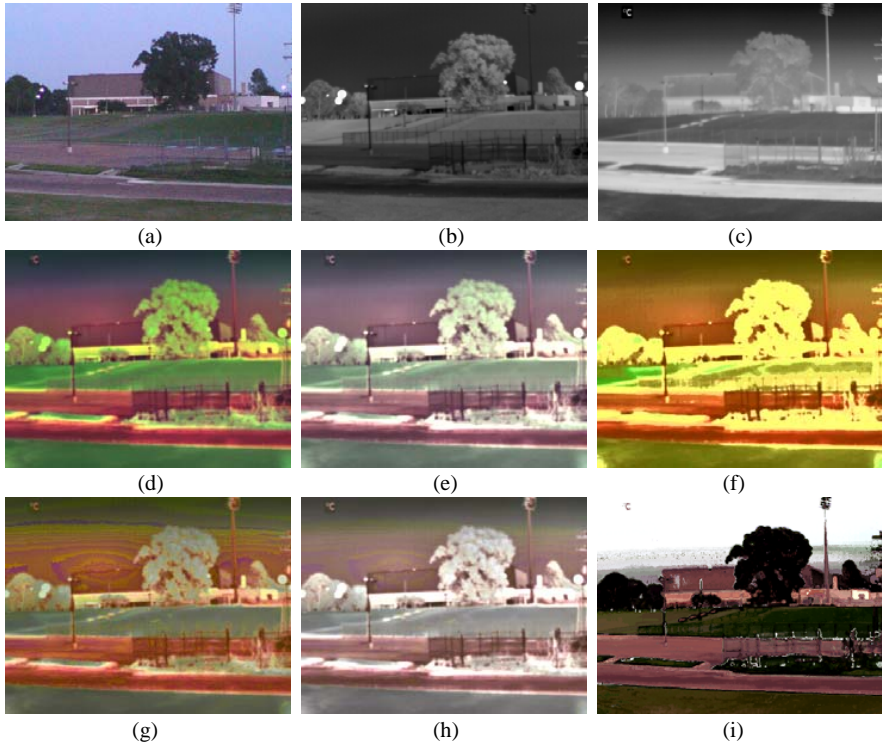
Figure 4. Night-vision coloring comparison (Case# AT009 – taken after sunset time; 640×480 pixels): (a-c) Color RGB, NIR, and LWIR images, respectively; (d-f) The colorized images using channel-based color fusion of (NIR⊕ LWIR), statistic-matching, and histogram-matching, respectively; (g-i) The colorized images using joint-HM, stat-match then joint-HM, and LUT-mapping, respectively. The settings in the color-mappings of (e-i) are source = (d) and target = (a). Notice that the contrasts of all color images were increased by 10%, and the brightness of (a) was increased by 10%.

Of course, image registration and fusion (Hil & Batchelor, 2001) were applied to the three band images shown in Figs. 3-7, where manual alignment was employed to the RGB image shown in Figs. 5-6a since they are so dark and noisy. To better present the color images (including the daylight RGB images and the colorized NV images), contrast and brightness adjustments (as described in figure captions) were applied. Notice that piecewise contrast stretching (Eq. (2)) was used for NIR enhancement. As referred in Eq. (1d), the fused images (shown elsewhere (Zheng & Essock, 2008)) were obtained using the aDWT algorithm (Zheng et $al.$, 2005). The channel-based color fusion (CBCF, defined in Eqs. (1)) was applied to the NIR and LWIR images (shown in Figs. 3-7b & c), and the results are illustrated in Figs. 3-7d. The resulted images from two-band color fusion (Figs. 3-7d) resemble natural colors, which makes scene classification easier. The paved ground appears reddish since they have strong heat radiations (at dusk time) and thus causes strong responses in LWIR images. In the color-fusion images, the trees, buildings and grasses can be easily distinguished from ground (parking lots) and
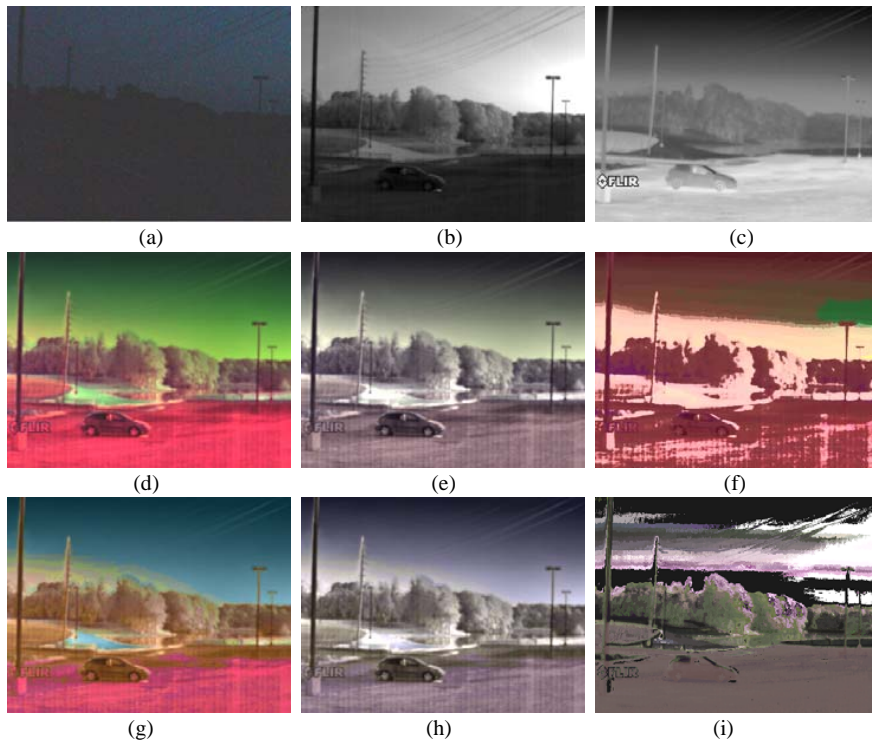
Figure 5. Night-vision coloring comparison (Case# AT012 – taken at dusk time; 640×480 pixels): (a-c) Color RGB, NIR, and LWIR images, respectively; (d-f) The colorized images using channel-based color fusion of (NIR⊕ LWIR), statistic-matching, and histogram-matching, respectively; (g-i) The colorized images using joint-HM, stat-match then joint-HM, and LUT-mapping, respectively. The settings in the color-mappings of (e-i) are source = (d) and target = Fig. 8(a) due to the dark RGB image in (a). Notice that the contrasts of all color images were increased by 10%, and the brightness of (a) and (i) were increased by 20% and 10%, respectively.

sky. For example, the car is clearly identified in Fig. 5d, where the water area (between ground and trees and shown in cyan color) is certainly noticeable. However, it is hard to realize any water area in the original images (Figs. 5a-c).

All color mapping methods were applied to the five triplets and their results are presented in Figs. 3-7. The source images are the color-fusion images (Figs. 3-7d), while the target images are the color RGB images (Figs. 3-4a & Fig. 8a-b). Figs. 5-6a cannot be used as the target images since they are too dark and noisy. Figs. 3-7e show the colored images with the statistic matching (SM) method, which are more similar to the daylight pictures in contrast with the color-fusion images. The five results (Figs. 3-7e) are equivalently good, which means that the statistic matching is reliable. The histogram matching (HM) results shown in Figs. 3-7f are oversaturated, which may be more suitable for segmentation-based colorization (Zheng & Essock, 2008). The joint histogram matching (JHM) are illustrated in Figs. 3-7g, where the mapped
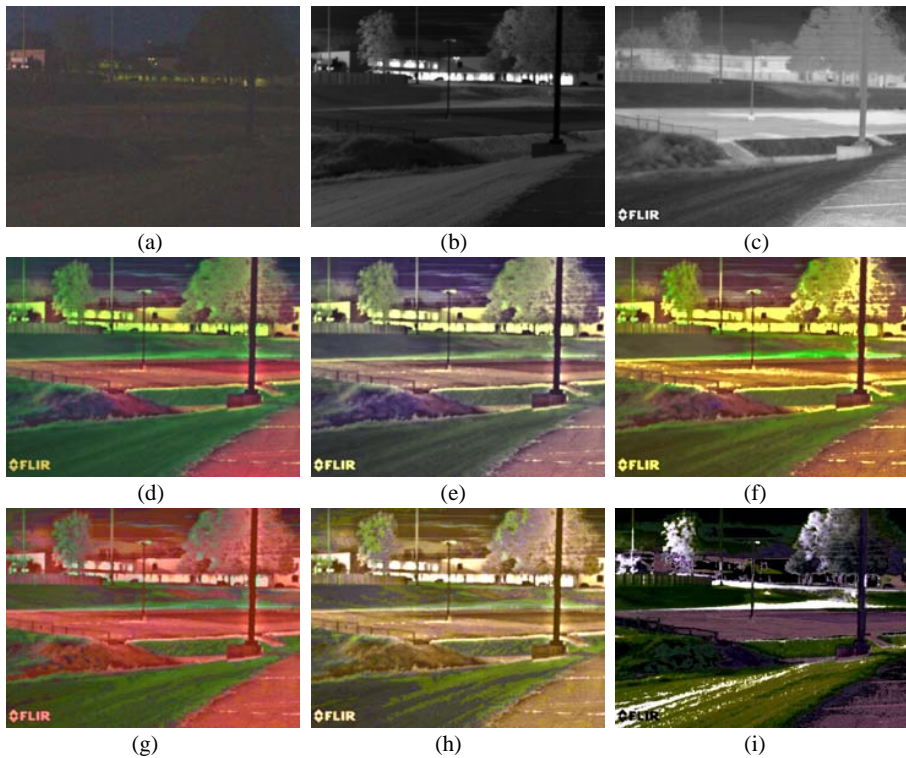
Figure 6. Night-vision coloring comparison (Case# ST029 – taken at dusk time; 640×480 pixels): (a-c) Color RGB, NIR, and LWIR images, respectively; (d-f) The colorized images using channel-based color fusion of (NIR⊕ LWIR), statistic-matching, and histogram-matching, respectively; (g-i) The colorized images using joint-HM, stat-match then joint-HM, and LUT-mapping, respectively. The settings in the color-mappings of (e-i) are source = (d) and target = Fig. 8(b) due to the dark RGB image in (a). Notice that the contrasts of (d-i) were increased by 10%, and (a) was increased by 20%. The brightness of (a) and (i) were increased by 20% and 10%, respectively.

images are better than the color fusions but preserve much reddish colors (existed in source images). The "stat-match then joint-HM" (SM-JHM) means that a joint-HM is performed with inputs of (source = the SM-colored image in Fig. 3e; target = the RGB image in Fig. 3a). The SM-JHM results are presented in Figs. 3-7h, which sometimes are better than the results from either stat-match or joint-HM (e.g., Fig. 3h). The examples of LUT-mapping colorization are given in Figs. 3-7i. Figs. 3-4i and Fig. 7i (an ideal case of LUT mapping) shows impressive colors; whereas Figs. 5-6i exhibit noisy and distorted since the reference images (shown in Figs. 8a-b) are misaligned with the NV images (shown in Figs. 5-6). When using the LUT established in a different case at daytime (aiming at different direction at nighttime), the more misalignment the worse the LUT-colored results appear. The LUT-based colorization described in Subsection 2.6 is perhaps suitable for a surveillance application where a camera is aiming at a fixed direction.
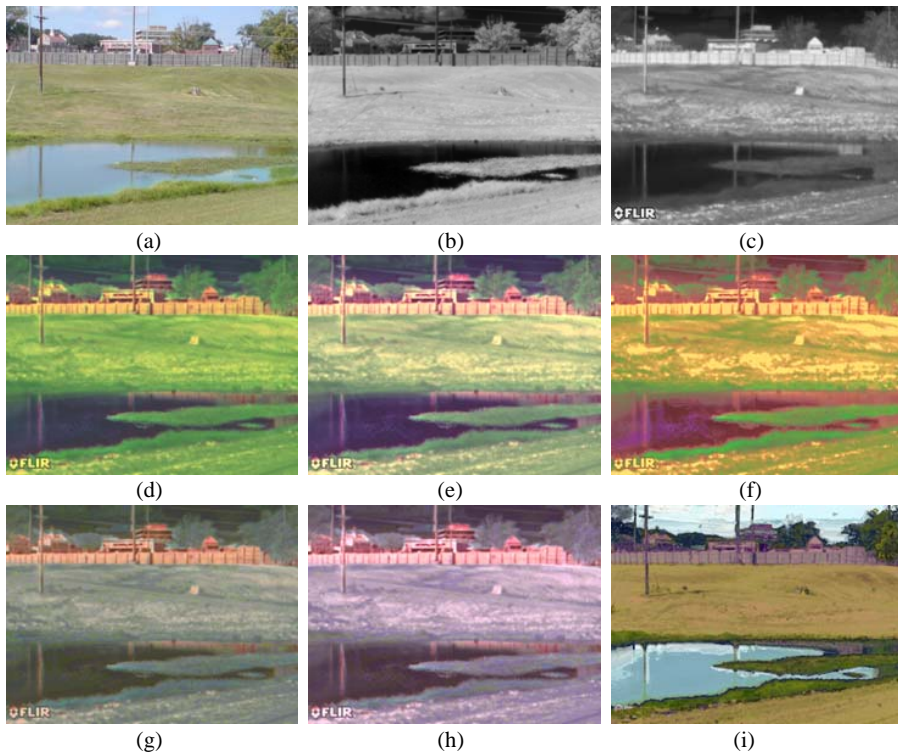
Figure 7. Night-vision coloring comparison (Case# ST102 – taken at noon time; 640×480 pixels): (a-c) Color RGB, NIR, and LWIR images, respectively; (d-f) The colorized images using channel-based color fusion of (NIR⊕ LWIR), statistic-matching, and histogram-matching, respectively; (g-i) The colorized images using joint-HM, stat-match then joint-HM, and LUT-mapping, respectively. The settings in the color-mappings of (e-i) are source = (d) and target = (a).

Figure 8. Color RGB images for night-vision colorization (taken before sunset time; 640×480 pixels): (a) from Case# AT002 (target of Fig. 5, AT012); (b) from Case# ST014 (target of Fig. 6, ST029). Notice that their contrasts were increased by 10%.

Visual inspections of colorized images can generally tell which one is better or the best when there are big enough differences between several versions of colorized images. For example, casual inspections may easily confirm that, top 3 methods are SM, SM-JHM, and LUT; HM and JHM are poor; and CBCF is medium. However, the subjective evalutions become more and more difficult with a larger number of color images and also hard with small or diverse differences. In other words, it is hard for subjective evalutions to give an exact order of six colroziation methods. Let us examine the objective evaluations.

The objective evaluations using the OEI metric defined in Eq. (14) (refer to Section 4) are presented in Table 1 (corresponding to Figs. 3-7 respectively), where the orders of metric values (1 for the smallest OEI) are given within round parentheses. Keep in mind that, the larger the OEI value of a colorized image is, the better quality (i.e., the higher order number) the colorized image has. According to the OEI values in Table 1, the quality order of colorized images varies with figures (cases). To have an overall impression, the sums of the order numbers in five cases (i.e., Figs. 3-7) are calculated and shown at the rightmost column in Table 1. The quality order of each colorization method (6 for the best) is given within the curly brackets. The order of colorization methods from the best to the worst: SM (stat-match), SM-JHM (stat-match then joint-HM), LUT, CBCF (channel-based color fusion), HM (histogram matching), JHM (joint-HM). This order sorted by OEI values is quite consistent with the order of subjective evaluations.

Method (Plot)	Fig. 3 (AT008)	Fig. 4 (AT009)	Fig. 5 (AT012)	Fig. 6 (ST029)	Fig. 7 (ST102)	Sum {Order}
CBCF (d)	0.4753 (3)	0.5497 (3)	0.5178 (2)	0.5132 (4)	0.5872 (3)	15 {3}
SM (e)	0.5470 (6)	0.6022 (5)	0.6058 (6)	0.5529 (5)	0.6337 (6)	**28 {6}**
HM (f)	0.4519 (2)	0.4890 (1)	0.3587 (1)	0.5099 (3)	0.5736 (2)	9 {2}
JHM (g)	0.4372 (1)	0.5250 (2)	0.5189 (3)	0.4674 (1)	0.5503 (1)	8 {1}
SM-JHM (h)	0.5428 (5)	0.5954 (4)	0.5978 (5)	0.5678 (6)	0.6154 (4)	*24 {5}*
LUT (i)	0.5148 (4)	0.6025 (6)	0.5238 (4)	0.4882 (2)	0.6322 (5)	21 {4}

Table 1. The OEI (Order) values of six color-mapping methods over five cases shown in Figs. 3-7 (The Sum & {Order} at last colmun is calculated with the orders of five cases).

The subjective evaluations of night vision coloration are based on casual visual inspections. More qualitative measurements, subjective evaluations (by a group of subjects), and statistical analysis will be introduced in the future. The quantitative (objective) evaluations using the objective quality index (OEI) require a reference (daylight) image. Thus we will continuously improve the OEI metric by relaxing the requirement of a reference image. We will further conduct more comprehensive comparisons.

6. Conclusions

In this chapter, we review six night-vision colorization techniques, a channel-based color fusion (CBCF) procedure; statistic matching (SM), histogram matching (HM), joint histogram matching (JHM), and stat-match then joint-HM (SM-JHM) method, and LUT-based approaches. An objective evaluation metric for NV colorization, objective evaluation index (OEI), is introduced. The experimental results with five case analyses showed the order of colorization methods from the best to the worst: SM, SM-JHM, LUT, CBCF, HM, JHM. The order of objective evaluations comply with the order of subjective evaluations.

The accurate objective metric such as OEI will help develop, select, and/or tune up a better NV colorization technique. The ideally colorized NV imagery can significantly enhance the night vision targeting by human users and will eventually lead to improved performance of remote sensing, nighttime perception, and situational awareness.

Acknowledgements

Author details

Yufeng Zheng[1*], Wenjie Dong[2], Genshe Chen[3] and Erik P. Blasch[4]

*Address all correspondence to: Yufeng.Zheng@R2Image.com

1 Alcorn State University, USA

2 The University of Texas-Pan American, USA

3 I-Fusion Technologies, Inc, USA

4 US Air Force Research Laboratory, USA

References

[1] Alparone, L.; Baronti, S.; Garzelli, A.; & Nencini, F. (2004). A global quality measurement of Pan-sharpened multispectral imagery, *IEEE Geosci. Remote Sens. Lett.*, 1(4), 313-317.

[2] Blasch, E.; Li X.; Chen, G. & Li, W. (2008). Image Quality Assessment for Performance Evaluation of Image Fusion, *Proc. of 11th international conference on Information fusion*, Germany.

[3] Essock, E. A.; Sinai, M. J. & et al. (1999). Perceptual ability with real-world nighttime scenes: imageintensified, infrared, and fused-color imagery, *Hum. Factors* 41(3), 438–452.

[4] Fischer, S.; Sroubek, F.; Perrinet, L.; Redondo, R. & Cristoal, G. (2007). Self-invertible 2D log-Gabor wavelets, *Int. J. Computer Vision*, 75(2), 231-246.

[5] Gabor, D. (1946). Theory of communication, *J. Inst. Elec. Eng.*, 93(3), 429-457.

[6] Gonzalez, R. C. & Woods, R. E. (2002). *Digital Image Processing* (Second Edition), Prentice Hall, ISBN: 0201180758, Upper Saddle River, NJ.

[7] Henriksson, L.; Hyvarinen, A. & Vanni, S. (2009). Representation of cross-frequency spatial phase relationships in human visual cortex, *J. Neuroscience*, 29(45), 14342-14351.

[8] Hill, D. L. G. & Batchelor P. (2001). Registration methodology: concepts and algorithms, in Medical Image Registration, Hajnal, J. V.; Hill, D. L. G.; & Hawkes, D. J. Eds, Boca Raton, FL.

[9] Hogervorst, M.A. & Toet, A. (2008). Method for applying daytime colors to nighttime imagery in realtime, *Proc. SPIE 6974*, 697403.

[10] Kovesi, P. (1999) . Image features from phase congruency, Videre*: J. Comp. Vis. Res.*, 1(3), 1-26.

[11] Liu, Z.; Blasch, E.; Xue, Z.; Langaniere, R.; & Wu, R (2012). Objective Assessment of Multiresolution Image Fusion Algorithms for Context Enhancement in Night Vision: A Comparative Survey, *IEEE Trans. Pattern Analysis and Machine Intelligence*, 34(1): 94-109.

[12] Ma, M.; Tian, H.P.; & Hao, C.Y. (2005). New method to quality evaluation for image fusion using gray relational analysis, *Opt. Eng.* 44, 087010.

[13] Malacara, D. (2002) Color Vision and Colorimetry: Theory and Applications, SPIE Press, Bellingham, WA.

[14] Mancas-Thillou, C. & Gosselin, B. (2006). Character segmentation-by-recognition using log-Gabor filters, *Proc. Int. Conf. Pattern Recognition*, 901-904.

[15] Morrone, M.C.; Ross, J.; Burr, D.C.; & Owens, R. (1986) Mach bands are phase dependent, *Nature*, 324 (6049), 250-253.

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

[17] Toet, A. & Hogervorst , M.A. (2012). Progress in color night vision," *Opt. Eng*. 51 (1), 010901.

[18] Toet, A. & IJspeert, J. K. (2001). 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, pp.436–441.

[19] Tsagaris, V. (2009). Objective evaluation of color image fusion methods, *Opt. Eng.* 48, 066201.

[20] Tsagaris, V. & Anastassopoulos, V. (2006) "Global measure for assessing image fusion methods," *Opt. Eng.* 45, 026201.

[21] Varga, J. T. (1999). Evaluation of operator performance using true color and artificial color in natural scene perception (Report ADA363036), Naval Postgraduate School, Monterey, CA.

[22] Wald, L. ; Ranchin, T. ; & Mangolini, M. (1997). Fusion of satellite images of different spatial resolutions: assessing the quality of resulting images, *Photogramm. Eng. Remote Sens.* 63(6), 691-699.

[23] Wang, W.; Li, J.; Huang, F.; & Feng, H. (2008). Design and implementation of log-Gabor filter in fingerprint image enhancement, *Pattern Recognit. Letters*, 29(3), 301-308.

[24] Waxman, A.M.; Gove, A. N. & *et al.* (1996). 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.

[25] Yuan, Y.; Zhang, J.; Chang, B.; & Han Y. (2011). Objective quality evaluation of visible and infrared color fusion image, *Opt. Eng.*, 50(3), 033202.

[26] Zhang, L.; Zhang, L.; Mou, X. & Zhang, D. (2011) FSIM: A Feature Similarity Index for Image Quality Assessment, *IEEE Trans. on Image Processing*, 20 (8), 2378 - 2386.

[27] Zheng, Y. (2011). A channel-based color fusion technique using multispectral images for night vision enhancement, *Proc. SPIE 8135*, 813511.

[28] Zheng, Y. (2012). An Overview of Night Vision Colorization Techniques using Multispectral Images: from Color Fusion to Color Mapping", 2012 International Conference on Audio, Language and Image Processing (ICALIP 2012), Shanghai, China.

[29] Zheng, Y. & Essock, E. A. (2008). A local-coloring method for night-vision colorization utilizing image analysis and image fusion, *Information Fusion* 9, 186-199.

[30] Zheng, Y.; Essock, E. A. & Hansen, B. C. (2005). An advanced DWT fusion algorithm and its optimization by using the metric of image quality index, *Optical Engineering* 44 (3), 037003-1-12.

[31] Zheng, Y; Dong, W.; & Blasch, E. (2012). Qualitative and quantitative comparisons of multispectral night vision colorization techniques," Optical Engineering, 51(8), 087004

[32] Zheng, Y; Reese, K; Blasch, E; & McManamon, P. (2013). Qualitative evaluations and comparisons of six night-vision colorization methods, Proc. SPIE 8745.