A false color image fusion method based on multi-resolution color transfer in normalization YC_BC_R space

Xuelian Yu a,*, Jianle Ren a, Qian Chen a, b, Xiubao Sui a

a Jiangsu Key Laboratory of Spectral Imaging & Intelligent Sense, Nanjing University of Science and Technology, Nanjing, China
b Key Laboratory of Photoelectronic Imaging Technology and System, Beijing Institute of Technology, Ministry of Education of China, China

A R T I C L E   I N F O

Keywords:
Image fusion
Color transfer
Multi-resolution transform
False color
YC_BC_R transformation

A B S T R A C T

In this paper, a false color image fusion method for merging visible and infrared images is proposed. Firstly, the source images and reference image are decomposed respectively by Laplacian pyramid transform. Then the grayscale fused image and the difference images between the normalized source images are assigned to construct YC_BC_R components. In the color transfer step, all the three channels of the color space in each decomposition level are processed with the statistic color mapping according to the color characteristics of the corresponding sub-images of the reference image. Color transfer is designed based on the multi-resolution scheme in order to significantly improve the detailed information of the final image, and to reduce excessive saturation phenomenon to have a comparatively natural color appearance compared with the classical global-coloring algorithm. Moreover, the differencing operation between the normalized source images not only provides inter-sensor contrast to make popping the potential targets but also weakens the influence of the ambient illumination variety to a certain degree. Finally, the fused results and several metrics for evaluation of fused images subjectively and objectively illustrate that the proposed color image fusion algorithm can yield a more complete mental representation of the perceived scene, resulting in better situational awareness.

© 2014 Elsevier GmbH. All rights reserved.

1. Introduction

Multi-sensor image fusion can be defined as the process that is to synthesize complementary and redundant information from different registered images which are obtained by different imaging sensors (sensitive in different spectral wavebands). Thus, the final image can provide a more reliable and accurate representation of the scene [1,2]. Infrared (IR) image records thermal radiations emitted from objects, and it can discover the target which is warmer or cooler than its surrounding environment, while visible image has much more high-frequency information of the background to favor accurate target localization [3]. In this paper, we concentrate on the false color fusion of visible and IR images.

It is anticipated that the successful multi-spectral fusion will lead to improved performance of computer vision, remote sensing, surveillance, medical imaging analysis, target detection, and military field [4,5]. The traditional fused imagery has been represented in gray tones. However, human eyes are more sensitive to color according to the principle of human visual system (HVS). In this regard, the human eyes can discern several thousands colors, whereas it can only distinguish about 100 shades of gray at any instance [2,4,6]. Therefore, a suitable color representation of fused night-vision (NV) imagery may help an observer construct a better description of the scene. On the contrary, an inappropriate color mapping will significantly decrease the human performance. With the increasing availability of multi-band vision systems, color image fusion has emerged as a novel and important research area, and many of fusion schemes have been developed in recent years to combine IR and visible images [7,8].

MIT Lincoln Laboratory developed the image fusion method of visible and IR radiation which was derived from the biological modes of color vision [9,10]. In Ref. [11], Waxman et al. used the ON and OFF center-surround response channels to simulate the rattlesnake bimodal cell operating mechanism.

Toet et al. proposed a false color image approach in that the common and particular parts of the visual and IR images were calculated firstly. A fused color image was then produced resulting from the red and green channels of a color display [12]. Although the resultant image enhanced the visibility of certain details, the
color was not natural. In 2003, Toet originally introduced the color transfer technology to produce NV imagery with a natural appearance in \( \text{IR} \) color space where three channels were perceptually irrelevant. Satisfactory results can be obtained so long as the false color fused image was similar to a selected natural day-time color image to some degree\cite{6,13}. As logarithm and exponent operations were involved in the transformation between RGB and \( \text{IR} \) spaces, the system’s storage requirements and computational complexity were increased. Subsequently, Hoegh and Toet presented a fast natural color mapping to render multi-band night-time imagery \cite{7,14}. The lookup-table based mapping procedure was extremely simple and it provided object color constancy. However, the derivation of the color scheme may require some time and frequently vary according to the sensor settings and a matching environment.

The rendered fused image may appear unnatural adopted by the global-statistic color mapping \cite{12} because the source and reference images contain significantly different regions in the colored content. Subsequently, Zheng and Essock suggested a local-coloring algorithm combining statistic and histogram distribution, and the target color schemes were grouped by their scene contents \cite{5}. Due to involving time-consuming procedures such as nonlinear diffusion, histogram analysis and wavelet-based fusion, it is computationally expensive and easy to make false judgment affected by image segmentation. Reference \cite{8} introduced a ratio of local to global divergence of the IR image to pop out both hot and cold targets in color. However, when the scene contains large similar temperature regions, this method does not stand out the potential targets and involve undesirably excessive saturation phenomenon.

In order to improve global-coloring performance and enhance the situation perception, a fusion approach for visual and IR images based on multi-resolution color transfer in normalization \( \text{YCgCb} \) space is proposed. Concerning the human visual characteristics, two source images and reference image are decomposed into a resolution pyramid using Laplacian pyramid (LP) transform. The differences of the normalized source images were then assigned to \( \text{Cb} \) and \( \text{Cr} \) channels respectively in \( \text{YCgCb} \) color space. The \( \text{Y} \) component which corresponds to the luminance channel of the fused image was obtained upon pattern-selective combination rule and normalization. Different from the color transfer method based on pyramid scheme \cite{15} which uses color mapping technology after multi-resolution reconstruction, each sub-image of three channels was mapped according to the mean and standard deviation of corresponding sub-image of the reference image in each decomposition level. This is what we called MRCT (Multi-resolution color transfer) in order to reduce excessive saturation phenomenon and produce the colored images that have a more natural appearance. Finally, the fused pyramid was reconstructed to generate the colorize fusion imagery. In addition, the luminance contrast of the final color fusion image can effectively be enhanced by using high contrast gray-level fusion image as the \( \text{Y} \) component. Three objective metrics in the course of the discussion were used to evaluate the performance of the proposed algorithm with other false color fusion methods. Our experiments demonstrated that the color fusion images ultimately obtained more rich color, contained more details and facilitated object recognition and a large degree of situational awareness.

The rest of this paper is organized as follows: Section 2 briefly describes the classical color mapping method developed by Toet. The false color fusion algorithm is introduced in detail in Section 3. Section 4 presents the experimental results to demonstrate the feasibility of the proposed fusion method. A discussion on the assessment of color fusion images quality is given in Section 5. Finally, we conclude the paper with a short summary in the last section.

2. Typical false color image fusion method

Inspired by the opponent-color fusion approach, an artificial color scheme was proposed by Toet et al., which fused visible and thermal images into composite images \cite{12}.

The fusion scheme involves only simple operations. First, the common component of the two original input images is determined. Second, the common component is subtracted from the original images to obtain the unique component of each image. Third, the unique component of each image modality is subtracted from the image of the other modality to enhance the representation of sensor-specific details in the final fused result. Finally, the resulting images can then be combined into a composite false color image by mapping the processed thermal image to the red band, the processed visual image to the green band and the difference between the characteristic components of both image modalities to the blue band of a RGB display:

\[
\begin{bmatrix}
R(x, y) \\
G(x, y) \\
B(x, y)
\end{bmatrix} = \begin{bmatrix}
IR(x, y) - Vis^*(x, y) \\
Vis(x, y) - IR^*(x, y) \\
IR^*(x, y) - Vis^*(x, y)
\end{bmatrix},
\]

where \( x, y \) represent the pixel coordinates and \( Vis^*(x, y) \) and \( IR^*(x, y) \) denote the unique components of the visual and IR images, which can be defined as:

\[
Vis^*(x, y) = Vis(x, y) - \text{Comn}(x, y),
\]

\[
IR^*(x, y) = IR(x, y) - \text{Comn}(x, y).
\]

In Eqs. (2) and (3), \( \text{Comn}(x, y) \) denotes the common component of the original images which can be computed as the morphological intersection:

\[
\text{Comn}(x, y) = \min \{ Vis(x, y), IR(x, y) \}.
\]

3. The proposed false color image fusion algorithm

In this section, we present the proposed false color image fusion algorithm. Fig. 1 shows the block diagram of the proposed method.

The method is as follows. Let the input multi-band NV images be the source images, and let a normal daylight photograph be the reference image. First, the image pyramids are obtained respectively using the LP decomposition algorithm for all the source images. The reference image is converted into \( \text{YCgCb} \) space, and three color components of reference image are decomposed into sub-images using the LP transformation. Second, the \( \text{Y} \) component of the sub-image in \( \text{YCgCb} \) color space is obtained in pattern-selective combination rule at each pyramid sample position. The \( \text{Cg} \) component is provided by subtracting IR pyramid from visual pyramid after the normalization process. In the same way, the \( \text{Cr} \) component is implemented by subtracting visual pyramid from IR pyramid. Third, color transfer technology is adopted in each decomposition layer based on the mean and standard deviation of the corresponding sub-images of the reference image with three channels. Finally, the composite image is recovered through the inverse MR pyramid transform.

![Fig. 1. Schematic diagram of the proposed false color image fusion method.](image-url)
final color fused image that is displayed by use of the inverse \( YC_bC_r \) transformation resembles a normal daylight image.

In the following subsections, we have provided more detailed explications of the image fusion process.

3.1. MR pyramid transform

There is verification that the HVS performs a similar decomposition in its early processing with the MR pyramid method [16]. This part is based on Burt and Kolczynski’s work on LP used to combine two different-focused images [17]. The fusion process begins with the construction of image pyramids \( D_{Vis} \) and \( D_{IR} \) for the two source images. Let \( D_l(m, n, k, l) \) be the sub-images (or coefficients) of image \( f(x, y) \) by LP transform. The indices \( mnlk \) indicate the sample position, level, and orientation in the pyramid. Image features, such as edges, are segregated by scale at the different levels. In the transform domain, the sub-images are combined by the pattern-selective fusion rule including two measures: salience measure and match measure.

**Salience measure:** At sample locations where the source images are distinctly different, the most salient component pattern from the source coefficients is adopted. The salience of a pattern is defined as a local energy within neighborhood \( p \):

\[
S_l(m, n, k, l) = \sum_{(s,r) \in p} D_l(m + s, n + t, k, l)^2,
\]

where \( p \) is a small window which size can be \( 3 \times 3 \) or \( 5 \times 5 \).

**Match measure:** The match measure is used to determine which of the two combination modes at each sample position, selection or averaging. The match of two sub-images is computed as a local normalized correlation within neighborhood \( p \):

\[
M_{AB}(m, n, k, l) = \frac{2 \sum_{(s,t) \in p} D_{Vis}(m + s, n + t, k, l) D_{IR}(m + s, n + t, k, l)}{S_{Vis}(m, n, k, l) + S_{IR}(m, n, k, l)}
\]

If the match measure between images is low at a given position, the coefficient from the source pyramid with the higher salience is copied to the composite pyramid. If the match measure is high, coefficients for two source pyramids are averaged to obtain the value inserted in the composite pyramid, and vice versa.

3.2. Combination rule

Different fusion strategies are chosen on different LP levels. Adding priority average method is adopted for the top level of LP which represents the low frequency information of original image. The pattern-selective combination rule can be stated as a weighted average, in which weights depend on the match and saliency measures, for the high frequency information of other level LP. The combined result of every sub-band is satisfied with the formula, given in Eq. (7):

\[
D_C(m, n, k, l) = W_{Vis}(m, n, k, l)D_{Vis}(m, n, k, l) + W_{IR}(m, n, k, l)D_{IR}(m, n, k, l)
\]

where \( W_{Vis} \) and \( W_{IR} \) at each position \((m, n, k, l)\), if \( M_{AB} \leq \alpha \), then:

\[
\begin{align*}
W_{Vis}(m, n, k, l) &= 1, W_{IR}(m, n, k, l) = 0 \quad S_{Vis} \geq S_{IR} \\
W_{Vis}(m, n, k, l) &= 0, W_{IR}(m, n, k, l) = 1 \quad S_{vis} < S_{IR}
\end{align*}
\]

3.3. YCbCr color space construction

Due to the simplicity and independency of the linear \( YC_bC_r \) space, the false color fusion algorithm based on it has an inherent advantage in the computational aspects over the approach based on the nonlinear \( le\beta \) space. The contrast of features uniquely presented in either of the images is reduced, while the conventional false color method is implemented based on global-coloring framework. An appropriate fused image obtained by combining the individual sensor images can preserve all relevant details of the individual bands. Consequently, color transfer is combined with MR process. In other words, color mapping is accomplished at each pyramid sample position. Based on this strategy, the MRCT scheme is put forward. \( Y_D, C_{BD} \) and \( C_{RD} \) components in each level are product as follows:

\[
\begin{align*}
Y_D(m, n, k, l) &= \frac{D_C(m, n, k, l)}{\max(D_C(m, n, k, l))} \\
C_{BD}(m, n, k, l) &= \frac{D_{Vis}(m, n, k, l)}{\mu D_{Vis}(l)} - \frac{D_{IR}(m, n, k, l)}{\mu D_{IR}(l)} \\
C_{RD}(m, n, k, l) &= \frac{D_{IR}(m, n, k, l)}{\mu D_{IR}(l)} - \frac{D_{Vis}(m, n, k, l)}{\mu D_{Vis}(l)}
\end{align*}
\]

where \( Y_D(m, n, k, l) \) represents the luminance of the colorized sub-image, which can be acquired by the Eq. (7). In some conditions, a detail that is noticeable in the individual image bands may be much less visible in the final color representation, due to the lack of luminance contrast [18]. Therefore, the luminance component of \( YC_bC_r \) color space is replaced with the grayscale fused sub-images. \( \mu \) represents the mean of decomposition coefficients of the source images. In Eq. (11), \( C_{BD}(m, n, k, l) \) and \( C_{RD}(m, n, k, l) \) denote color information consistent with Color Difference Mode [19], stand for difference between blue and luminance and that between red and luminance respectively.

\[
\begin{align*}
C_{BD} &= 0.5 \frac{B - Y}{0.886} \\
C_{RD} &= 0.5 \frac{R - Y}{0.701}
\end{align*}
\]

The differential coefficients of subtracting IR sub-images from visible ones are fed into \( C_b \) channel. It means that the coefficients whose values in visible sub-images are greatly larger than those in IR sub-images that will appear to be blue color, which are commonly background for instance trees and sky. In addition, the coefficients whose values in IR sub-images are greatly larger than those in visible sub-images will present red color, which are possibly the potential targets with high radiance in IR image. As a result, the chromatic channels of the fused image show inter-sensor contrast, popping out targets in possible color.

3.4. MRCT

Color transfer technology can modify the fused image (called source image) luminance and chrominance distribution according to a color day-time visible image (called reference image). In this
step, color transfer is performed to make the fused sub-image and the corresponding reference sub-image have the same mean and standard deviation for each channel and each level in YCbCr color space. First, the reference image is converted to YCbCr space as shown in Eq. (13):

$$
\begin{bmatrix}
Y_{\text{Ref}} \\
C_{b,\text{Ref}} \\
C_{r,\text{Ref}}
\end{bmatrix} =
\begin{bmatrix}
0.2990 & 0.5870 & 0.1140 \\
-0.1687 & -0.3313 & -0.5000 \\
0.5000 & -0.4187 & -0.0813
\end{bmatrix}
\begin{bmatrix}
R_{\text{Ref}} \\
G_{\text{Ref}} \\
B_{\text{Ref}}
\end{bmatrix}
$$

(13)

Second, the reference components \(Y_{\text{Ref}}, C_{b,\text{Ref}}, C_{r,\text{Ref}}\) are decomposed into a set of LP decomposition levels. Let \(D_{P,\text{Ref}}(m, n, k, l)\) be the sub-images of the reference image by LP transform. \(P\) denotes as an individual color channel of the image in YCbCr space. Third, at each position \((m, n, k, l)\) the color transfer step can be described as:

$$
p_{D,F}(m, n, k, l) = \frac{\sigma_{D,C}^2(l)}{\sigma_{D,C}^2(l)} (p_{D,C}(m, n, k, l) - \mu_{D,C}^2(l)) + \mu_{D,C}^2(l) \quad P = Y, C_b, C_r
$$

(14)

where the indices \(C\) and \(F\) refer to the sub-image and the final fused sub-image of each color band in each layer, respectively. \(\mu\) and \(\sigma\) represent the mean and standard deviation of channel \(P\) in each level. \(p_{D,C}(m, n, k, l)\) can be carried out according to the Eq. (11). By applying a linear mapping to each channel of the sub-image as in Eq. (14), the final fused sub-image will take the style and appearance of the reference sub-image.

3.5. Inverse pyramid transform

An inverse pyramid transform is accomplished in which the combined image \(F\) of channel \(P\) is recovered from its pyramid representation \(P_{D,F}(m, n, k, l)\):

$$
P_P = \sum_{l=0}^{N} P_{D,F}(m, n, k, l) \quad P = Y, C_b, C_r
$$

(15)

3.6. RGB color display

The last step in fusion is to transform the result back into RGB representation and get the final color fused image.

$$
\begin{bmatrix}
R_F \\
G_F \\
B_F
\end{bmatrix} =
\begin{bmatrix}
1.0000 & 0.0000 & 1.4020 \\
1.0000 & -0.3441 & -0.7141 \\
1.0000 & 1.7720 & 0.0000
\end{bmatrix}
\begin{bmatrix}
Y_F \\
C_{b,F} \\
C_{r,F}
\end{bmatrix}
$$

(16)

As the YCbCr transformation is linear, without the logarithm and exponential operations, its computational complexity is far lower than that of the \(l\alpha\beta\) transformation.

4. Experiments and results

To validate the effectiveness of our method, four sets of visual and IR source images are fused. Most of the original registered visible and IR images are provided by TNO Human Factors, except that the first dataset (shown in Fig. 2) is supplied by Yuan (the author of Ref. [22]). To be convenient, the source images have been processed of registration before image fusion. Since the corresponding daylight color images are not available for these image pairs, some arbitrary color images as reference are adopted. For comparison, we also obtain the color fused images by Toet’s [12] and Yin’s [8] methods. Four fusion results are completely analyzed and shown in Figs. 2–5.
Fig. 4. Visual and IR image fusion of dataset3. (a) Visible image; (b) IR image; (c) reference image; (d) result by Toet’s method; (e) result by Yin’s method; (f) result by our method.

Each decomposition level are fused by the pattern-selective fusion rule to achieve a fused image with less artifacts compared to the widely used maximum selection fusion rule.

In dataset 1, the contours of the trees are correctly represented in the CCD image (Fig. 2(a)), on the contrary, indistinct in IR image (Fig. 2(b)). The different fusion effects are illustrated in Fig. 2(d)–(f) based on three different schemes. Fig. 2(d) emerging excessive saturation phenomenon by Toet’s method may hide some salient information represented in the input image, such as the configuration of windows and the texture of trees. Fig. 2(e) shows that the fusion result obtained by Yin’s method provides more details than Fig. 2(d), but the hot target (the person) in this case appears to be blurred. In Fig. 2(f) some details are more easily seen and can be better recognized in the fused image than in the individual input images. Moreover, the outlines of the person are again clearly visible.

The fence is clearly visible in CCD image (Fig. 3(a)), but not presented by IR image (Fig. 3(b)) because of low contrast between objects and backgrounds. Fig. 3(d) appears in a glaring white color due to over-saturation. Fig. 3(e) has a relatively appropriate appearance, and the potential target (the person) is enhanced to be easily detected. Nevertheless, the overall image looks inadequately

Fig. 6. Results of applying our method to fuse dataset2 images with the different reference images.
distinct lack of details, some areas that unnecessarily outlined, such as the roof of the building, are recolored in a reddish color. This example indicates that the MRCT method is superior to other methods, containing more prominent information for the input images and providing a natural appearance with rich details.

In dataset 3 the person and burning object are represented at high contrast in IR image (Fig. 4(b)) because they have high temperature difference as the surrounding. However, they are not represented in CCD image because they are shielded with smog. By observing Fig. 4(d)–(f), it can be seen clearly that the person and burning object can be most easily explored in our method’s result (Fig. 4(f)). In this case, smoke shrouds the buildings. In others words, most of the area has similarly high temperature. Hence, it is difficult to investigate the potential objects by Yin’s method due to its algorithm principle.

Fig. 5 shows images of a nighttime scene representing a sailing on the sea covered with heavy cloud. It is found that the paths and the ship are represented in unnatural red color in Fig. 5(e). From Fig. 5(f), the waves around the ship are recolored in blue color and clearly visible with plentiful details. Furthermore, two persons and the path can be easily distinguished because they have different color saturation. Consequently, our scheme can achieve a visually satisfactory fused image with suitable contrast and much more abundant colors than Toet’s method.

Fig. 6 illustrates some examples of our method adopting the three different reference images (see the left-most column). Columns 2–4 of Fig. 6 show the final color fusion images. The second row describes the MRCT results using a natural picture representing a grassland area with some cattle and sheep as the reference image. Although the content of the reference image is largely dissimilar to that of the source scene, the corresponding fused results still have relatively natural appearance. The reference image in the bottom row represents the broad grass and the bright sky overhead. The reference image and the source image are extremely dissimilar, so that the color of the composite image seems unnatural, such as the paths represented in improper blue. As a result, the variation of reference image may affect the final colorizations in a small degree, and our scheme can yield a colorized image with a natural daytime color appearance as long as the color characteristics of the reference image are to some extent similar to those of the source scene.

5. Discussion

To objectively assess the performance of the proposed color image fusion method, the objective evaluation metrics are adopted. To the best of our knowledge, only a few studies evaluation metrics are introduced for color fused image [20–22]. As mentioned in Ref. [22], sharpness, contrast and colorfulness, are experimentally proved to play a dominant role on color image-quality assessment for observers. All the color attributes are computed in the CIELAB color space.

Image sharpness metric (ISM) represents the clarity of detail and edge information of an image and can be defined as:

\[
ISM = \frac{1}{|W|} \sum_i \nabla I(x,y)
\]

\[
\nabla I(x,y) = [C^2_x + C^2_y]^{1/2}
\]

where \(|W|\) is the total number of all windows and \(\nabla I(x,y)\) is the local gradient of the gray-scale image \(I\) at coordinate \((x, y)\) by the Sobel operator.

An image with excellent contrast performance has a wide dynamic range of gray or color intensity level and appropriate intensity. Image contrast metric (ICM) can be expressed as:

\[
ICM = \frac{1}{2} (CCM_1 + CCM_2)
\]

\[
CCM_1 = \left[ \frac{1}{MN} \sum_{i=0}^{M} \sum_{j=0}^{N} C^2_{ij} \right]^{1/2}
\]

\[
CCM_2 = \left[ \frac{1}{|W|} \sum_i \nabla f(x,y) \right]^{1/2}
\]

\[
C^* = (a^2 + b^2)^{1/2}
\]

where \(C_x\) and \(C_y\) are the gray contrast metric and color contrast metric, respectively.

Colorfulness metric (CCM) represents the color vividness degree of image and can be computed as:

\[
CCM = \frac{1}{2} (CCM_1 + CCM_2)
\]

The detailed description of the three objective metrics can be found in Ref. [22]. Quality assessments based on three metrics are evaluated for the previously described four datasets in Table 1. We can see that the proposed method performs best in both ISM and CCM results. It is consistent with HVS. The details and edges of the images using our method are clearest, such as leaves in the tree in Fig. 2(f) and both the fence and the poles in Fig. 3(f). The subjective evaluation that these images using our method look more vivid agrees with the CCM metrics. In image-contrast evaluation, the ICM results of our method are slightly lower than those of Toet’s method in the case of dataset 1 and 2. The colors of Figs. 2(d) and 3(d) look

<table>
<thead>
<tr>
<th>Assessment metrics</th>
<th>ISM</th>
<th>ICM</th>
<th>(\frac{C_x}{C_y})</th>
<th>CCM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fig. 2 (d)</td>
<td>45.7065</td>
<td>0.5927</td>
<td>0.4785</td>
<td>15.0795</td>
</tr>
<tr>
<td>(e)</td>
<td>56.8124</td>
<td>0.4071</td>
<td>0.4189</td>
<td>14.9968</td>
</tr>
<tr>
<td>(f)</td>
<td>99.8361</td>
<td>0.5883</td>
<td>0.4726</td>
<td>15.9349</td>
</tr>
<tr>
<td>Fig. 3 (d)</td>
<td>93.2341</td>
<td>0.5658</td>
<td>0.4760</td>
<td>17.2661</td>
</tr>
<tr>
<td>(e)</td>
<td>59.2055</td>
<td>0.3871</td>
<td>0.4096</td>
<td>15.8050</td>
</tr>
<tr>
<td>(f)</td>
<td>95.8213</td>
<td>0.5247</td>
<td>0.4709</td>
<td>16.4842</td>
</tr>
<tr>
<td>Fig. 4 (d)</td>
<td>29.9822</td>
<td>0.5501</td>
<td>0.3888</td>
<td>8.4785</td>
</tr>
<tr>
<td>(e)</td>
<td>23.7460</td>
<td>0.3379</td>
<td>0.2785</td>
<td>8.6088</td>
</tr>
<tr>
<td>(f)</td>
<td>32.1325</td>
<td>0.6455</td>
<td>0.5194</td>
<td>9.4067</td>
</tr>
<tr>
<td>Fig. 5 (d)</td>
<td>43.8189</td>
<td>0.5186</td>
<td>0.4247</td>
<td>13.8604</td>
</tr>
<tr>
<td>(e)</td>
<td>13.5854</td>
<td>0.2259</td>
<td>0.1702</td>
<td>12.6812</td>
</tr>
<tr>
<td>(f)</td>
<td>57.4756</td>
<td>0.5444</td>
<td>0.3991</td>
<td>15.5732</td>
</tr>
</tbody>
</table>

The bold denotes the respectively calculated maximum by three kinds of objective evaluation metrics for each image pair.

Table 1: Color fusion image-quality evaluation results of four pairs.
bright in certain regions, thus their color contrast \( C_r \) are highest. That is why Toet’s approach has the best performance in the ICM. However, the colors are not natural according to human perception and lacks of more details. The images provided by Yin’s algorithm have the poorest performance in the ISM, because they look blurred in some details. Although the red colors of the target such as the person in Figs. 3(e) and 5(e) can aid target detection, both the ICM and the CCM are not best due to the fact that the colors of the major images are too light and lack of gradation. Furthermore, according to the conclusion of Ref. [22], figures employing the color transfer technology have excellent performance in CNN (color naturalness metric) evaluation as long as the source images are similar to the reference images.

6. Conclusion

A false color image-fusion method based on MRCT in the normalization YC\(_B\)C\(_R\) space is proposed in order to provide the colorized image which represents more plentiful details, enhanced color contrast and more natural. The three uncorrelated channels of YC\(_B\)C\(_R\) color space are constructed and modified according to the statistical properties of a color day-time reference image, while the visible, thermal and reference images are decomposed through LP transform. The required information provided from the source images, (especially the meaningful transformed coefficients in detecting salient features), are selected by the pattern-selective fusion rule. As a result, abundant spatial details are contained in final image, which can improve the observer’s perception. MRCT implemented in different resolution levels decreases excessive saturation phenomenon in the resulting image used in typical global-coloring algorithm, and it has a comparatively natural color appearance. In addition, color contrast is effectively enhanced by emerging the luminance component by a grayscale fused representation of the two input bands. The colorfulness of the resulting color imagery is improved by the strategy of assigning the difference between the source images (which are normalized in order to remove the influence of the ambient illumination variety to a certain degree) to the two chromatic channels. The experimental results indicate that our method can significantly reserve details of the background and have excellent detection probability, and render the fused image similar color appearance with a color day-time image at the same time. Thus it indeed provides a better representation of the depicted scene compared with the other two schemes. The fusion scheme involves some simple operations and can probably be applied in real time and realized in hardware. MRCT can be intended for not only the pyramid decomposition but also the other MR transforms such as discrete wavelet (DWT) and dual-tree complex wavelet (DTCWT) to fulfill different user needs.

Our method renders the fused image based on the color transfer technology. As a result, the resulting colorized night-vision image may look unnatural when the color distribution will be quite biased to that of the reference image. The next research step is to try to expand the method to accept imagery from three or four spectral bands, e.g. visible, short-wave infrared, middle-wave infrared and long-wave infrared bands.

Acknowledgements

The authors would like to thank Kan Ren for polishing the grammar. This work was supported by National Natural Science Foundation of China (Grant No. 61101119) and the Natural Science Foundation of Jiangsu province of China (Grant No. BK2011698). Many thanks to Alexander Toet and the TNO Human Factors Research Institute for providing the source visible and IR images.

References