14

Region-Based Fusion for Infrared and LLL Images

Junju Zhang, Yiyong Han, Benkang Chang and Yihui Yuan
Institute of Electronic Engineering and Optic-electronic Technology, Nanjing University of Science and Technology, Nanjing China

1. Introduction

Thermal cameras and image intensifiers are common night vision (NV) cameras, which enable operations during night and in adverse weather conditions. NV cameras deliver monochrome images that are usually hard to interpret and may give rise to visual illusions and loss of situational awareness. The two most common NV imaging systems display either emitted infrared radiation or reflected low level light (LLL). In this way the different imaging modalities give complementary information about the objects or area under inspection. Thus, techniques for fusing infrared and LLL images should be employed in order to provide a compact representation of the scene with increased interpretation capabilities.

Image fusion can be classified into two types based on pixel-level: pixel-based and region-based. The pixel-based image fusion is characterized by simplicity and highest popularity. Because pixel-based methods fail to take into account the relationship between points and points, the fused image with either of them might lose some gray and feature information. However, for most image fusion applications, it seems more meaningful to combine objects rather than pixels. The region-based fusion, on the contrary, can obtain the best fusion results by considering the nature of points in each region altogether. Therefore, region-based fusion has advantages over the other two counterparts. At present, region-based methods use some segmentation algorithm to separate an original image into different regions, and then design different rules for different regions.

During the last decade, a number of gray fusion algorithms have been proposed, and the fusion methods based on the multiscale transform (MST) are the most typical. The commonly used MST tools include the Laplacian pyramid and the wavelet transform (DWT). In general, due to the perfect properties of the DWT such as multi-resolution, spatial and frequency localization, and direction, the DWT-based methods are superior to the pyramid-based methods. However, the DWT also has some limitations such as limited directions and non-optimal-sparse representation of images. Thus, some artifacts are easily introduced into the fused images using the DWT-based methods, which will reduce the quality of the resultant image consequently. The Dual-Tree Complex Wavelet Transform (DT-CWT) has been introduced by Nick Kingsbury, which has the following properties: Approximate shift invariance; Good directional selectivity in 2-D with Gabor-like filters also

true for higher dimensionality; Perfect reconstruction using short linear-phase filters; Limited redundancy: independent of the number of scales. Therefore, the Dual-Tree Complex Wavelet Transform is more suitable for image fusion.

In the context of NV imaging, a number of color fused-based representations have been proposed. A simple mapping of infrared and visual bands into the three components of an RGB image can provide an immediate benefit, since the human eye can discriminate several thousands of colors but only a few dozens of gray levels. On the other hand, inappropriate color mappings may hinder situational awareness due to lack of color constancy. Hence, an image fusion method for night vision imagery must result in color images with natural appearance and a high degree of similarity with the corresponding natural scenes. To make the coloration of false-color images appear more natural, Reinhard recently introduced a method that enabled the transfer of colors from one image to another. Subsequently, Toet demonstrated that Reinhard's method could be adapted to transfer the natural color characteristics of daylight imagery into multi-band infrared and LLL images. Essentially, Toet's natural color mapping method matches the statistical properties of the NV imagery to that of a natural daylight color image. However, this particular color mapping method colors the image regardless of scene content, weights all regions of the source image by the ''global'' color statistics, and thus the accuracy of the coloring is very much dependent on how well the target and source images are matched.

In this chapter, we present a region-based gray fusion method using the DT-CWT and a region-based color fusion method for infrared and LLL images. Segmentation is very important because segmentation precision has a great influence on the following fusion process. Here, we adopt two segmentation methods: the morphologic mehod and the nonlinear diffusion method. In the gray fusion method, the infrared and LLL images are decomposed by DT-CWT, the segmentation regions are mapped into each level, and fusion is carried out region by region in terms of some fusion rules. The region-based color fusion method is based on Toet's global-coloring framework.

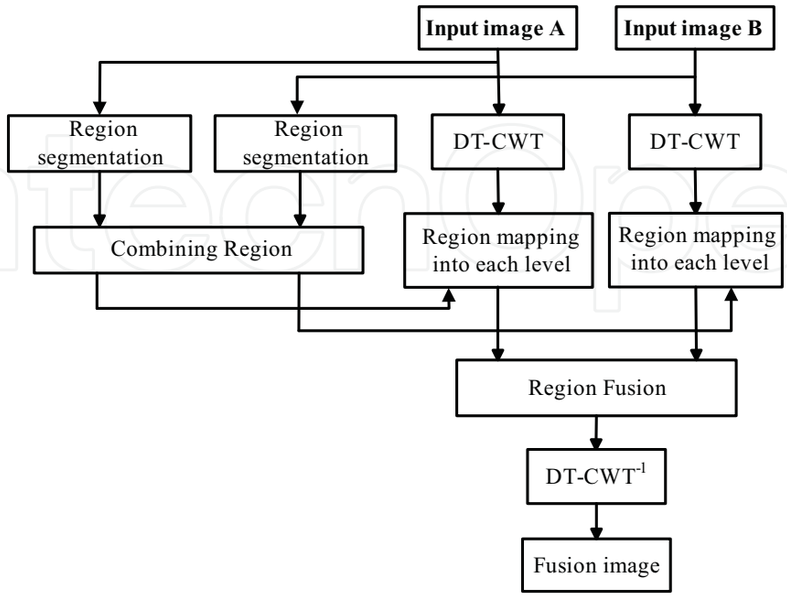


Fig. 1. Diagram of the proposed region-based method with DT-CWT

2. Image fusion based on region segmentation and complex wavelets

The region-based fusion method for infrared and LLL images adopts the DT-CWT because of its approximate shift invariance and limited redundancy. Diagram of the proposed region-based method with DT-CWT is shown in Fig. 1. Segmentation is firstly performed on the infrared image and LLL images respectively with top-bottom-hat filtering and the threshold method, consequently, the DT-CWT coefficients from the different regions are merged separately. Finally the fused image is obtained by performing inverse DT-CWT.

2.1 Image segmentation using morphology

The morphologic filters have been proven as powerful methods in the denoising and smoothing of image intensities while retaining and enhancing edges. The combination of different morphologic filters makes the segmentation flexible. The top-hat transform and the bottom-hat transform are all the combination of open operation, close operation and the original image.

The top-hat transform means subtracting a morphologically opened image from the original image and it can be used to enhance contrast in an image. The bottom-hat transform means subtracting the original image from a morphologically closed version of the image and it can be used to find intensity troughs in an image. The formula of the top-hat transform and bottom-hat transform are given by respectively

$$H_{top} = f - (f \circ p) \tag{1}$$

$$H_{bottom} = (f \bullet p) - f \tag{2}$$

Here f is the original image, "$\circ$" and "$\bullet$" are open operation and close operation, H_{top} and H_{bottom} are results of the top-hat transform and bottom-hat transform. Add the original image f to the top-hat filtered image H_{top}, and then subtract the bottom-hat filtered image H_{bottom}, we can obtain the enhanced image. At the same time, noises of the original image f are eliminated. The enhanced image H_E is given by

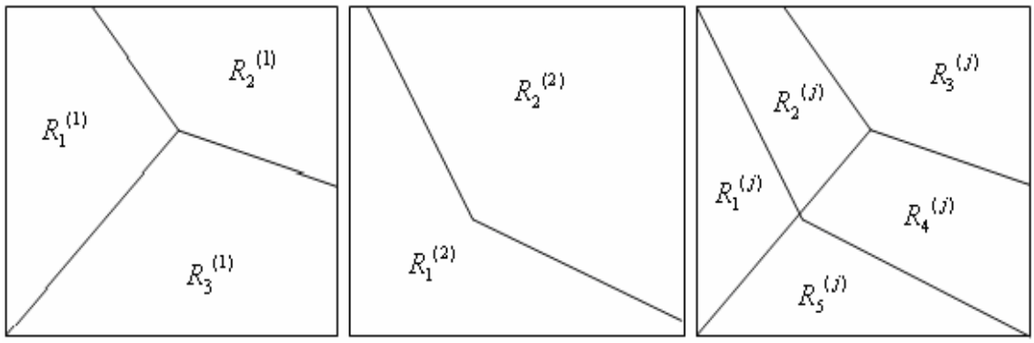
$$H_E = H_{top} - H_{bottom} + f \tag{3}$$

Then the threshold method is used to segment the enhanced image H_E. We can get the binary segmentation image based on this method. Because the physical significance of the pixel at the same location of the heterogeneous source images is different, the shapes of segmentation regions obtained by the former method are also different. So we must deal the segmentation region with the associate methods. The information of segmentation region should be added to the associated-segmentation image and is used to guide the fusion rules.

The following steps are used to generate the associated-segmentation image:

1. If there is no overlapping area between the region $R^{(1)}$ and the region $R^{(2)}$, then the associated- segmentation image is mapped into two regions, $R_1{}^{(j)} = R^{(1)}$, $R_2{}^{(j)} = R^{(2)}$;
2. If there is some overlapping area between the region $R^{(1)}$ and the region $R^{(2)}$, then the associated- segmentation image is mapped into three regions, $R_0{}^{(j)} = R^{(1)} \cap R^{(2)}$; $R_1{}^{(j)} = R^{(1)} \cap R_0{}^{(j)}$; $R_2{}^{(j)} = R^{(2)} - R_0{}^{(j)}$;

3. If region $R^{(1)}$ overlaps the region $R^{(2)}$ completely, the associated-segmentation image is mapped into the same region, $R^{(j)} = R^{(1)} = R^{(2)}$;
4. If one region completely contains the other region, for example, $R^{(1)} \subset R^{(2)}$, then the associated- segmentation image is mapped into two regions, $R_1^{(j)} = R^{(1)}$, $R_2^{(j)} = R^{(2)} - R^{(1)}$. Here, $R^{(1)}$ is presented a part of one source image, $R^{(2)}$ is presented a part of the other source image. $R^{(j)}$ is presented a part of associated-segmentation image. Fig. 2 is some typical examples of the associated region maps. There are some small regions in the associated-segmentation image, and they don't contain enough region information, which will cause false image in the fused image. We may merge these small regions using morphological operators.



a) b) c)

Fig. 2. Associated region maps for the fused image. (a) region map1, (b) region map2, (c) associated region map

2.2 Pixel fusion with complex wavelets

The dual-tree complex wavelet transform (DT-CWT) is a relatively recent enhancement to the discrete wavelet transform (DWT), with important additional properties: It is nearly shift invariant and directionally selective in two and higher dimensions. It achieves this with a redundancy factor of only 2d for d-dimensional signals, which is substantially lower than the undecimated DWT. The multidimensional (M-D) DT-CWT is non-separable but is based on a computationally efficient, separable filter bank (FB).

For 2-D signals, we can filter separately along columns and then rows by the way like 1-D. Kingsbury figured out in that, to represent fully a real 2-D signal, we must filter with complex conjugates of the column and row filters. So it gives 4:1 redundancy in the transform. Furthermore, it remains computationally efficient, since actually it is close to a classical real 2-D wavelet transform at each scale in one tree, and the discrete transform can be implemented by a ladder filter structure. The quad-tree transform is designed to be, as much as possible, translation invariant. It means that if we decide to keep only the details or the approximation of a given scale, removing all other scales, shifting the input image only produces a shift of the reconstructed filtered image, without aliasing. The most important property of DT-CWT is that it can separate more directions than the real wavelet transform. The 2-D DWT produces three band-pass subimages at each level, which are corresponding to LH, HH, HL, and oriented at angles of 0º, ± 45 º, 90 º. The 2-D DT-CWT can provide six

subimages in two adjacent spectral quadrants at each level, which are oriented at angles of ±15 °, ±45 °, ±75 °. The strong orientation occurs because the complex filters are asymmetry responses. They can separate positive frequencies from negative ones vertically and horizontally. Therefore, positive and negative frequencies will not be aliasing.

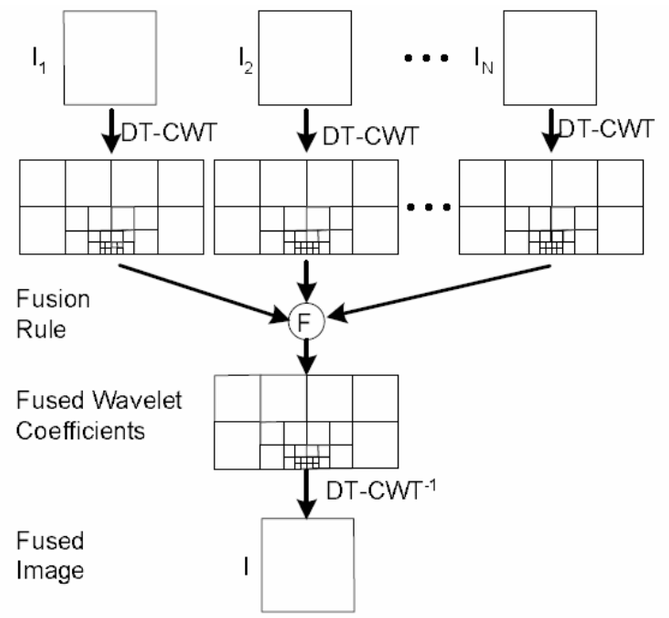


Fig. 3. The pixel-based image fusion scheme using the DT-CWT

The pixel-level fusion scheme used here employs the DT-CWT to obtain a MR decomposition of the input images. The wavelet coefficients are then combined, using a maximum-selection fusion rule to produce a single set of coefficients corresponding to the fused image. This process is shown in Fig. 3. The maximum-selection scheme selects the largest absolute wavelet coefficient at each location from the input images as the coefficient at that location in the fused image. As wavelets tend to pick out the salient features of an image, this scheme works well producing good results.

2.3 Image fusion based on region segmentation and complex wavelet

Decomposed by the multi-resolution DT-CWT, low-frequency part of the images denotes their approximate components, which contains spectral information of the source image. High-frequency part of the images denotes their detail components, which contains edge detail information of the source images. So, fusion algorithms after the source images decomposed are very important for the quality of fusion. At present, the fusion rules are commonly that average operator or weighted average operator is used in low-frequency domain, max absolute operator is used in high-frequency domain. For the two fused heterogeneous source images of the same scene, spectral information of one image is usually much richer than the other. For example, spectral information of visible light image is much richer than the infrared image. If the fusion rules of weighted average is adopted, part of spectrum information of visible light images will be lost, which results in that the spectrum information of fused image is less than visible light image.

To overcome these problems, we adopt spatial frequency to guide region-based fusion. The spatial frequency, which originated from the human visual system, indicates the overall active level in an image. The human visual system is too complex to be fully understood with present physiological means, but the use of spatial frequency has led to an effective objective quality index for image fusion. The spatial frequency of an image block is defined as follows: Consider an image, the row (R_F) and column (C_F) frequencies of the image block are given by

$$R_F = \sqrt{\frac{1}{M}\sum_{(i,j)\in\Omega}(F(i,j)-F(i,j-1))^2} \tag{4}$$

$$C_F = \sqrt{\frac{1}{M}\sum_{(i,j)\in\Omega}(F(i,j)-F(i-1,j))^2} \tag{5}$$

Here Ω is a certain segmentation region. The total spatial frequency S_F of the image is

$$S_F = \sqrt{R_F^2 + C_F^2} \tag{6}$$

We use the fusion of two registration source images A and B as an example, the image fusion process based on region segmentation and DT-CWT is accomplished by the following steps:

Step 1: Partition the source images A and B, then we get the region segmentations named R_A and R_B, using associated processing, then we can get the associated-segmentation image R_J. Calculate S_F of each region in the associated-segmentation image.

Step 2: Compare the spatial frequency of the corresponding regions of the two source images to decide fusion coefficients:

$$R_i^F = \begin{cases} \dfrac{S_{Fi}^{\ A}}{S_{Fi}^{\ A}+S_{Fi}^{\ B}}\bullet R_i^A + \dfrac{S_{Fi}^{\ B}}{S_{Fi}^{\ A}+S_{Fi}^{\ B}}\bullet R_i^B & \dfrac{1}{k} < S_{Fi}^{\ A}/S_{Fi}^{\ B} < k \\ R_i^A & S_{Fi}^{\ A}/S_{Fi}^{\ B} > k \\ R_i^B & S_{Fi}^{\ A}/S_{Fi}^{\ B} < \dfrac{1}{k} \end{cases} \tag{7}$$

Here R_i^F is the i th region of the fused image, $S_{Fi}^{\ A}$ and $S_{Fi}^{\ B}$ are the spatial frequencies of the i th region of image A and B, respectively, k is a threshold.

Step 3: Multi-level DT-CWT transform on the source images A and B, then we can get DT-CWT coefficients at different scale Layers, which contain low-frequency coefficient and high-frequency coefficient at different scale layers.

Step 4: Deal low-frequency and high-frequency part with fusion rules and fusion operators, then we get low-frequency coefficient and high-frequency coefficient at different scale Layers after fusion.

Step 5: Deal low-frequency coefficient and high-frequency coefficient at different scale Layers with DT-CWT inverse transform, then the reconstruction image is to be fused image.

2.4 Experiment results

To evaluate the presented fusion algorithm, we fuse the infrared and visible images of the same scene with this algorithm, and compare the fusion image with the fusion images with the DWT method (method 1) and SIDWT method (method 2). Fig. 4(a) is an infrared image, which presents the clear shapes such as a human being, trees and some high-temperature objects; Fig. 4(b) is a visible light (low light level) image, which provides more details than the infrared image. Besides this, it also shows some light sources. Fig. 4(c) is the segmentation region of the infrared image and Fig. 4(d) is the segmentation region of the visual image. Fig. 4(e) is associated region map of infrared/visible images, Fig. 4(f) is fused image with method 1 and Fig. 4(g) is fused image with method 2; Fig. 4(h) is fused image with the presented method.

According to the fusion images, the presented fusion algorithm has better effectiveness, which preserves not only the spectral information of the visible light image, but also the thermal target information of the infrared image. The details of the fusion image with the presented algorithm is clear, which shows that region segmentation has a function of extracting targets, also shows that the DT-CWT has the capability of capturing edge information. Though the fusion images with method 1 and 2 also reserve main scene information of the two images, they lose some details slightly. Edge of objects looks blurry slightly.

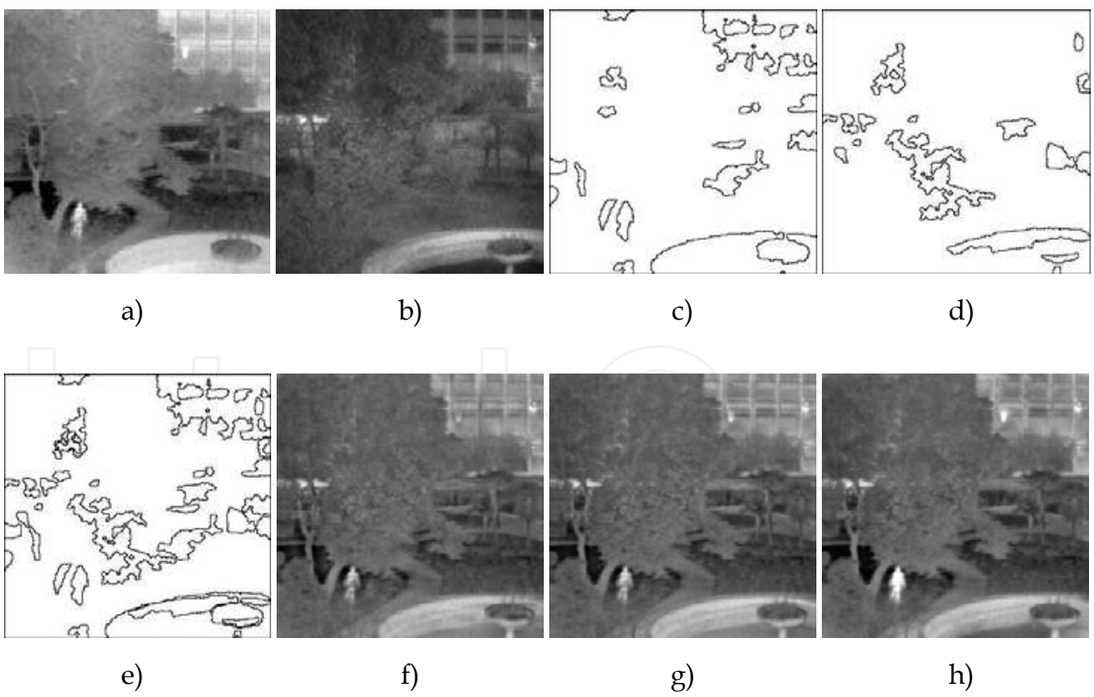


Fig. 4. Source images and fused results with different methods. (a) infrared image, (b) visible light image, (c) region map of the infrared image, (d) region map of the visible light image, (e) associated region map of infrared/visible light images, (f) fused image with method 1, (g) fused image with method 2, (h) fused image with the presented method

We use entropy, standard deviation, average gradient, structural similarity (SSIM) and $Q^{AB/F}$ to objectively evaluate the fusion images. Entropy reflects the average information of the image; standard deviation reflects the gray contrast of the fusion images and average gradient reflects the capability of expressing details of images; $SSIM(x,y,f)$ is an efficient metric of image fusion performance assessments. Given two images x and y of size $M \times N$, let μ_x denote the mean of x, let σ_x^2 and σ_{xy} be the variance of x and covariance of x and y. The SSIM index between signals x and y is :

$$\mathrm{SSIM}(x,y) = l(x,y)c(x,y)s(x,y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \tag{8}$$

In this paper, we use an 11×11 circular-symmetric Gaussian weighting function to modify μ_x, μ_y, σ_{xy}, σ_x and σ_y. With such a windowing approach, the quality maps exhibit a locally isotropic property. In practice, one usually requires a single overall quality measure of the entire image. We use a mean SSIM index to evaluate the overall image quality.

$$\mathrm{SSIM}(\mathrm{X},\mathrm{Y}) = \frac{1}{M}\sum_{j=1}^{M}\mathrm{SSIM}(x_j, y_j) \tag{9}$$

where X and Y are the reference and the distorted images, respectively; x_j and y_j are the image contents at the jth local window; and M is the number of local windows of the image.

We use the Wang-Bovik SSIM index in Eq. (9) to define a quality measure $SSIM(x,y,f)$ for image fusion. Here x, y are two input images and f is the composite image resulting from the fusion of x and y. The measure $SSIM(x,y,f)$ should express the "quality" of the composite image given the inputs x, y.

We denote by $s(x|w)$ some saliency of image x in window w. It should reflect the local relevance of image x within the window w, and it may depend on, e.g. contrast, variance, or entropy. Given the local saliencies $s(x|w)$ and $s(y|w)$ of the two input images x and y, we compute a local weight $\lambda_x(w)$ between 0 and 1 indicating the relative importance of image x compared to image y: the larger $\lambda_x(w)$, the more weight is given to image x. A typical choice for $\lambda_x(w)$ is

$$\lambda_x(w) = \frac{s(x|w)}{s(x|w) + s(y|w)} \tag{10}$$

In a similar fashion we compute $\lambda_y(w)$. Note that in this case $\lambda_y(w) = 1 - \lambda_x(w)$. Now we define the fusion quality measure $SSIM(x,y,f)$ as

$$SSIM(x,y,f) = \frac{1}{|W|}\sum_{w \in W}(\lambda_x(w)SSIM(x,f|w) + \lambda_y(w)SSIM(y,f|w)) \tag{11}$$

Thus, in regions where image x has a large saliency compared to y, the quality measure $SSIM(x,y,f)$ is mainly determined by the "similarity" of f and input image x. On the

other hand, in regions where the saliency of y is much larger than that of x, the measure $SSIM(x,y,f)$ is mostly determined by the "similarity" of f and input image y.

$Q^{AB/F}$ is based on the idea that a fusion algorithm that transfers input gradient information into the fused image more accurately performs better. For the fusion of input images A and B resulting in a fused image F, gradient strength g and orientation $\alpha(\in[0,\pi])$ are extracted at each location (n,m) from each image using the Sobel operator and used to define relative strength and orientation "change" factors G and A, between each input and the fused image:

$$(G^{AF}_{n,m}, A^{AF}_{n,m}) = \left(\frac{g^{F}_{n,m}}{g^{A}_{n,m}}\right)^{M}, 2\pi^{-1}\left|\alpha^{A}_{n,m} - \alpha^{F}_{n,m}\right| - \pi/2\right) \tag{12}$$

where M is 1 for $g^F > g^A$ and -1 otherwise. An edge information preservation measure Q^{AF} models information loss between A and F with respect to the 'change' parameters with sigmoid functions defined by constants Γ, κ_g, σ_g, κ_α, and σ_α:

$$Q^{AF}_{n,m} = \frac{\Gamma}{\sqrt{(1+e^{\kappa_g(G^{AF}_{n,m}-\sigma_g)})(1+e^{\kappa_\alpha(A^{AF}_{n,m}-\sigma_\alpha)})}} \tag{13}$$

Total fusion performance $Q^{AB/F}$ is evaluated as a sum of local information preservation estimates between each of the inputs and fused, Q^{AF} and Q^{BF}, weighted by local perceptual importance factors w^A and w^B usually defined as local gradient strength:

$$Q^{AB/F} = \frac{\sum_{\forall n,m} Q^{AF}_{n,m} w^{A}_{n,m} + Q^{AF}_{n,m} w^{B}_{n,m}}{\sum_{\forall n,m} w^{A}_{n,m} + w^{B}_{n,m}} \tag{14}$$

Table 1 gives the evaluation results of the three former algorithms. The evaluation results show the validity of the presented algorithm.

	Entropy	Average gradient	Standard deviation	$SSIM(x,y,f)$	$Q^{AB/F}$
Method 1	6.8313	0.0218	32.2700	0.6133	0.4282
Method 2	6.8774	0.0221	33.4626	0.6012	0.5010
The presented algorithm	6.9329	0.0226	34.1770	0.6304	0.5090

Table 1. Evaluation results of entropy, average gradient, standard deviation, $SSIM(x,y,f)$ and $Q^{AB/F}$

3. Region-based color fusion for infrared and LLL images

Toet demonstrated that transfer of colors could be adapted to transfer the natural color characteristics of daylight imagery into multi-band infrared and LLL images. However, this particular color mapping method colors the image regardless of scene content, weights all

regions of the source image by the ''global'' color statistics, and thus the accuracy of the coloring is very much dependent on how well the target and source images are matched. Based on Toet's global-coloring framework, we present a new region-based method that the image segmentation is firstly carried out and then region coloring is realized.

3.1 Review of global-coloring method

The aim of the global-coloring is to give NV images the appearance of normal daylight color images. A false-color image (source image) is first formed by assigning multi-band NV images to three RGB channels. The false-color images usually have an unnatural color appearance. Then, a true-color daylight image (reference image) is manually selected with similar scenery (e.g., syntactic content and color appearance) to the NV images. Both source and reference 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 ''statistic- matching'' procedure is carried out between the source and reference image. The mapped source image is then transformed back to RGB space. Finally, the mapped source image is transformed into YCbCr space and the ''value'' component (similar to the luminance component in $l\alpha\beta$ decomposition) of the mapped source image is replaced with the ''fused NV image'', which is a grayscale image made with multi-band NV images (e.g., image intensified and infrared image). This fused image replacement is necessary to make the colorized image have a proper and consistent contrast. Notice that the ''luminance'' component in $l\alpha\beta$ space cannot be used directly for this replacement because its dynamic range is very different from that of the fused image, whereas the "value" component in YCbCr space has the same gray-level range as the fused image. The lab space is utilized for color mapping because of its decorrelation property of three channels, whereas the YCbCr space is suitable for human interface.

The fusion process can be summarized in the following steps:

1. Set the R channel with the infrared image data, G and B channel with low-light-level image data and generate the rough color fusion image. Choose a reference image with good contrast

$$\begin{bmatrix} R_S \\ G_S \\ B_S \end{bmatrix} = \begin{bmatrix} IR \\ LL \\ LL \end{bmatrix} \tag{15}$$

2. The RGB values can be converted to LMS space by using the following equation

$$\begin{bmatrix} L \\ M \\ S \end{bmatrix} = \begin{bmatrix} 0.3811 & 0.5783 & 0.0402 \\ 0.1967 & 0.7244 & 0.0782 \\ 0.0241 & 0.1288 & 0.8444 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \tag{16}$$

3. A logarithmic transform is employed here to reduce the data skew that existed in the above color space

$$\begin{aligned} \boldsymbol{L} &= \log L \\ \boldsymbol{M} &= \log M \\ \boldsymbol{S} &= \log S \end{aligned} \tag{17}$$

4. $l\alpha\beta$ space can decorrelate the three axes in the LMS space

$$\begin{bmatrix} l \\ \alpha \\ \beta \end{bmatrix} = \begin{bmatrix} 0.5774 & 0.5774 & 0.5774 \\ 0.4082 & 0.4082 & -0.8165 \\ 1.4142 & -1.4142 & 0 \end{bmatrix} \begin{bmatrix} \boldsymbol{L} \\ \boldsymbol{M} \\ \boldsymbol{S} \end{bmatrix} \tag{18}$$

5. A simple technique, termed "statistic matching", used to transfer the color characteristics from natural daylight imagery to false-color night-vision imagery is formulated as

$$\boldsymbol{I}_C^k = (\boldsymbol{I}_S^k - \mu_S^k) \bullet \frac{\sigma_T^k}{\sigma_S^k} + \mu_T^k, for \quad k = \{l, \alpha, \beta\} \tag{19}$$

where I_C is the colored image, I_S is the source (false-color) image in $l\alpha\beta$ space; μ denotes the mean and σ denotes the standard deviation; the subscripts 'S' and 'T' refer to the source and reference images, respectively; and the superscript 'k' is one of the color components $\{l, \alpha, \beta\}$.After this transformation the pixels comprising source image have means and standard deviations that conform to the reference daylight color image in $l\alpha\beta$ space.

6. The inverse transform from the lab space to the LMS space can be expressed by

$$\begin{bmatrix} \boldsymbol{L} \\ \boldsymbol{M} \\ \boldsymbol{S} \end{bmatrix} = \begin{bmatrix} 0.5774 & 0.4082 & 0.3536 \\ 0.5774 & 0.4082 & -0.3536 \\ 0.5774 & -0.8165 & 0 \end{bmatrix} \begin{bmatrix} l \\ \alpha \\ \beta \end{bmatrix} \tag{20}$$

7. The transform depicted above can be inverted by raising the LMS pixel values to the tenth order back to linear LMS space, and then using the inverse transform of Eq. (10) to RGB space

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 4.4679 & -3.5873 & 0.1193 \\ -1.2186 & 2.3809 & -0.1624 \\ 0.0497 & -0.2439 & 1.2045 \end{bmatrix} \begin{bmatrix} L \\ M \\ S \end{bmatrix} \tag{21}$$

8. Transfer the fusion image data from RGB space to YC_BC_R space

$$\begin{bmatrix} Y \\ C_B \\ C_R \end{bmatrix} = \begin{bmatrix} 0.2990 & 0.5870 & 0.1140 \\ -0.1687 & -0.3313 & 0.5000 \\ 0.5000 & -0.4187 & -0.0813 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \tag{22}$$

9. The brightness Y_f acquired by transforming the fusion image and reference image from RGB space to $l\alpha\beta$ space is not usually appropriate, because Y_f is the weight sum of the infrared and low-light-level images. Thereby, we adopt the fusion image by laplacian pyramid fusion method replacing Y_f.
10. Transfer the adjusted rough fusion image data from YC_BC_R space back to RGB space, and we can get the ultimate re-staining rough fusion image.

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 1.0000 & 0.0000 & 1.4020 \\ 1.0000 & -0.3441 & -0.7141 \\ 1.0000 & 1.7720 & 0.0000 \end{bmatrix} \begin{bmatrix} Y \\ C_B \\ C_R \end{bmatrix} \tag{23}$$

3.2 The region-based coloring fusion method

Based on the framework of the global-coloring method as described in Section 3.1, we present a region-based coloring fusion method (see Fig. 5.) that makes the fusion images appear more like daylight imagery. The major points for this new method are as follows. (a) The infrared and LLL images are rendered segment-by-segment. (b) The segmented regions of the two images are combined and form a new segmented map. (c) These regions are classified according to the target types and the spatial frequencies, and some valuable targets are extracted according to the luminance of the infrared images or the motion trend of the infrared and LLL video. At present, our classification is still carried out manually. (d) The infrared and LLL images are mapped into the RGB space. A lot of mapping color methods has been provided, but the simplest mapping method is still suitable. For example, the infrared image is mapped into the R channel and the LLL image is mapped into the G channel and the average of the infrared and LLL images is mapped into the B channel. (e) Some typical scene images must be chosen as the reference images. These images should include some features similar to a certain segmented region. (f) Transfer of color is run region by region in the $l\alpha\beta$ space according to Reinhard's method. (g) The gray fusion image is used to replace the Y of the color fusion image in the YCbCr space. Here, the gray fusion method may be the fusion method in Section 2 or some other classical ones.

Image segmentation is quite challenging because image contents vary greatly from image to image. We adopt two segmentation methods. One is the morphologic method in Section 2, the other is the nonlinear diffusion method. The two methods have been proven as powerful methods in the denoising 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.

Basically, diffusion is a PDE (partial differential equation) method that involves two operators, smoothing and gradient, in 2D image space. The nonlinear diffusion equation is

$$\frac{\partial I(x)}{\partial t} = \vec{\nabla} \cdot \left(\omega(x) \vec{\nabla} I(x) \right) \tag{24}$$

Where $\vec{\nabla}$ is a vector containing gradients taken at different neighboring configurations (i.e., nearest-neighbors, second-neighbors, etc.) and $\omega(x)$ are the nonlinear diffusion coefficients.

The diffusion process smoothes the regions with lower gradients whereas stops smoothing at region boundaries with higher gradients. Nonlinear diffusion means the smoothing operation depends on the region gradient distribution. In other words, the diffused result is a nonlinear function of local gradients. Diffusion must be used with clustering and region merging techniques together, which make the segmentation flexible.

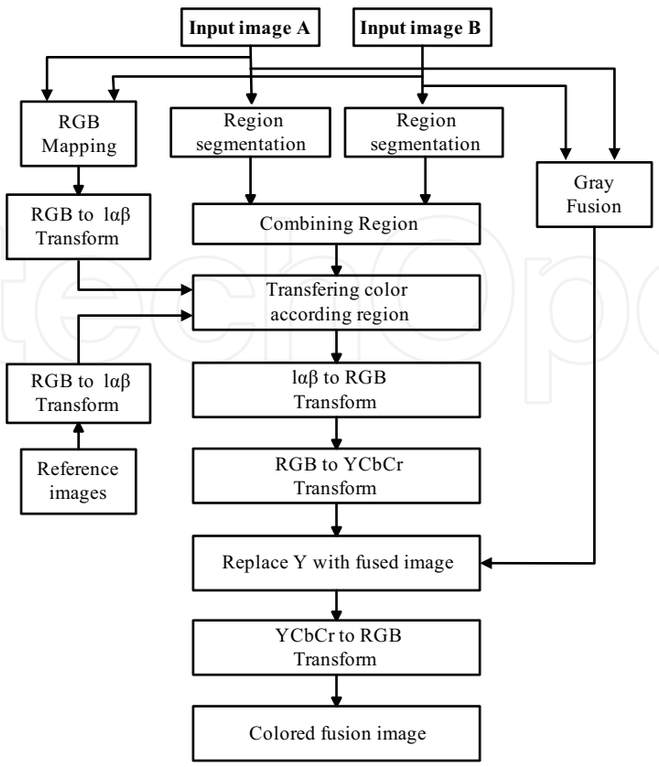


Fig. 5. Diagram of the proposed local-coloring method

3.3 Experiment result

Here two experiments have been carried out with the region-based coloring fusion method. The only difference is that the morphologic method is adopted in experiment 1 and the nonlinear diffusion method is adopted in experiment 2.

3.3.1 Experiment 1

To evaluate the region-based coloring fusion method, we fuse the infrared and LLL images of the same scene, and compare the fusion images with the presented method and the global-coloring method. Fig. 6(a) is an infrared image, which presents the clear shapes such as a human being, trees, building, pool and some high-temperature objects; Fig. 6(b) is a LLL image, which provides more details than the infrared image. Besides this, it also shows some light sources. Fig. 6(c) is the fusion image acquiring by Section 2. Fig. 6(e), (g), (i) and (k) are the fusion images with global-coloring method separately using Fig. 6(d), (f), (h) and (j) as reference image. Fig. 6(l) is the segmentation region of the infrared image with morphology method and Fig. 6(m) is the segmentation region of the visual image with morphology method. Fig. 6(n) is associated region map of infrared/ visible images. In the region image, that "person" was perfectly partitioned. The backgrounds such as road, building and so on are also well segmented. Fig. 6(o) is fused image with local-coloring method using Fig. 6(d), (f), (h) and (j) as reference images. Compared to Fig. 6(e), (g), (i) and (k), Fig. 6(o) has a clear color distinction between tree, person, building, pool and lawn. From Fig. 6(o) we can see that region-based coloring fusion method result can significantly improve observers' performance and reaction time.

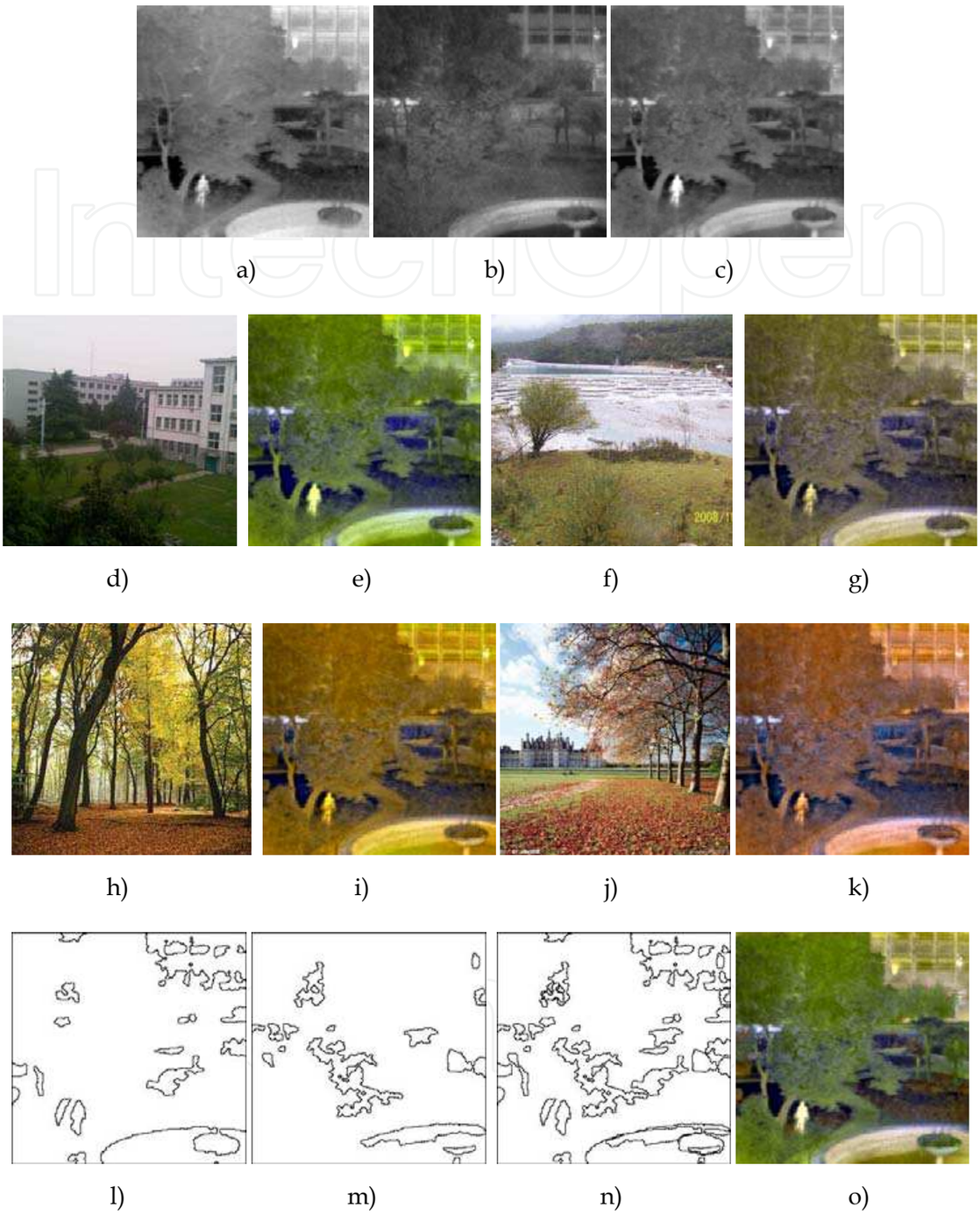


Fig. 6. Source images, reference images and fused results with different methods. (a) infrared image, (b) LLL image, (c) fusion image using method in Section 2, (e), (g), (i) and (k) fusion images with global-coloring method separately using (d), (f), (h) and (j) as reference image, (l) region map of the infrared image, (m) region map of the LLL image, (n) associated region map of (l) and (m), (o) fused image with the region-based coloring fusion method using (d), (f), (h) and (j) as reference color images

3.3.2 Experiment 2

To evaluate the presented region-based coloring fusion method, we fuse the infrared and visible light images of the same scene with this algorithm, and compare the fusion image with the fusion images with global-coloring method. Fig. 7(a) is an infrared image, which presents the clear shapes such as trees, building, sky and some high-temperature objects; Fig. 7(b) is a visible light image, which provides more details than the infrared image.

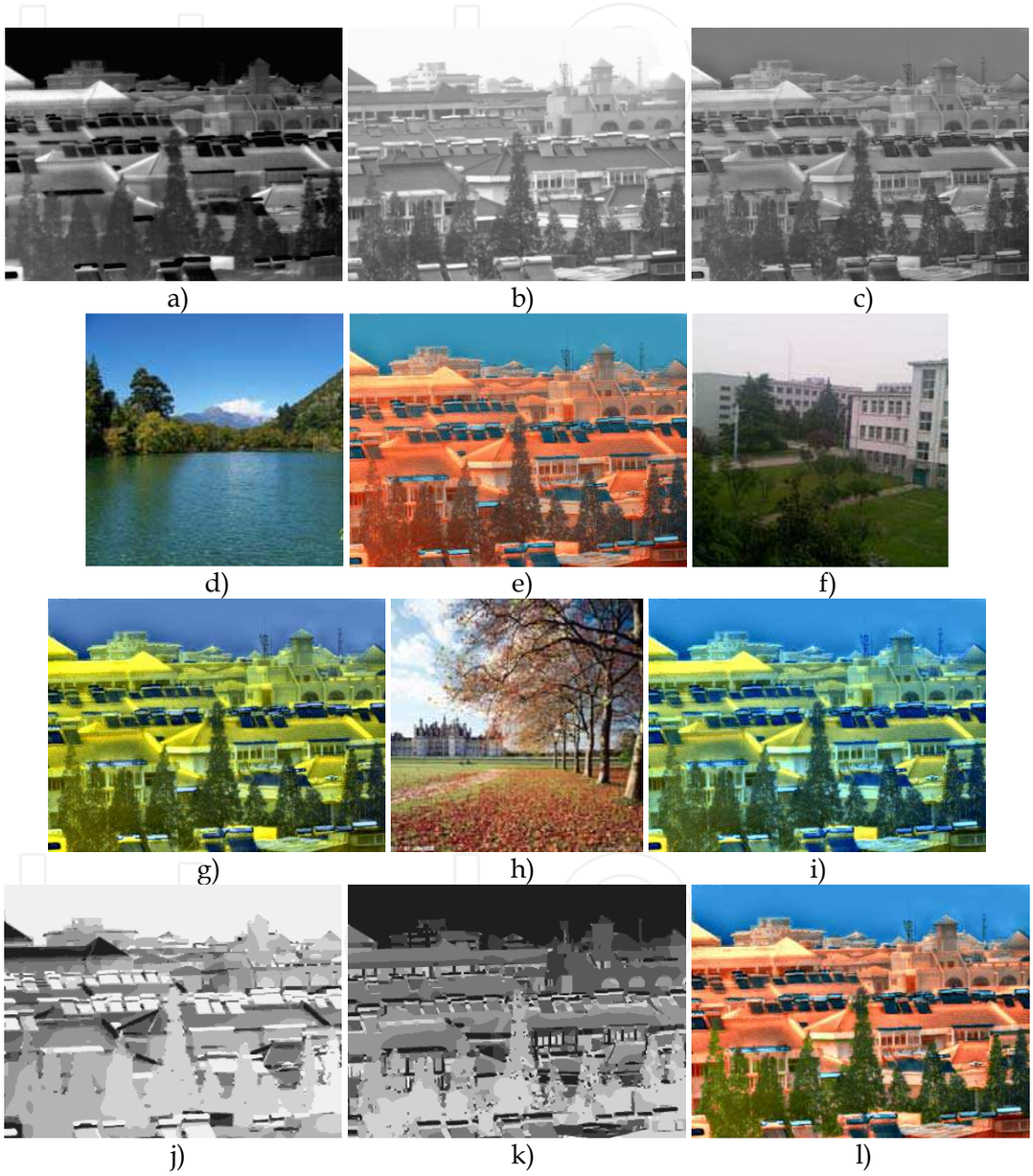


Fig. 7. Source images, reference images and fused results with different methods. (a) infrared image, (b) visible light image, (c) fusion image using gradient pyramid method, (e), (g) and (i) fusion images with global-coloring method separately using (d), (f) and (h) as reference image, (j) region map of the infrared image, (k) region map of the visible light image, (l) fused image with the region-based coloring fusion method using (d), (f) and (h) as reference images

Besides this, it also shows some light sources. Fig. 7(c) is the fusion image using gradient pyramid method. Fig. 7(e), (g) and (i) are the fusion images with global-coloring method separately using Fig. 7(d), (f) and (h) as reference image. Fig. 7(j) is the segmentation region of the infrared image with clustering method and Fig. 7(k) is the segmentation region of the visual image with clustering method. Fig. 7(l) is fused image with local-coloring method using (d), (f) and (h) as reference color images. Compared to Fig. 7(e), (g) and (i), Fig. 7(l) has a clear color distinction between tree, building and sky. From Fig. 7(i) we can see that local-coloring method result can significantly improve observers' performance and the colors are more nature than global-coloring method result.

4. Conclusion

This chapter presents a gray image fusion method and a color image fusion method based on the region segmentation. The region-based fusion methods use the different feature regions of original image and compound the pixel level and feature level of image fusions. The effective way to separate target and background proves crucial for the quality of image fusion. The proposed method preserves the details of the LLL image and the legible target of the infrared image, therefore, the fused image enables the exact location of the target to be easily observed and provides all-around information for further processing tasks.

In the gray method, segmentation is firstly performed on the IR image and LLL images with top-bottom-hat filtering and the threshold method, consequently, the DT-CWT coefficients from the different regions are merged separately. Finally the fused image is obtained by performing inverse DT-CWT. This method keeps the approximate shift invariance and the limited redundancy. Region segmentation performs us to using different rules for each region of each level. Experimental results evidence this method which could provide better fusion than classical fusion methods in terms of objective fusion metric values such as entropy, average gradient, standard deviation, $SSIM(x,y,f)$ and $Q^{AB/F}$

In the color fusion method, segmentation is firstly performed on the IR image and LLL image with the morphologic method or the diffusion method. At the same time, the IR image and LLL image are mapped into the RGB space, and the gray fusion of the two images is conducted. Here, the color map rule and the gray fusion method are not very important. The false-color images usually have an unnatural color appearance, but a chance of region by region color transferring is given to ensure the fusion image similar to natural images. The fusion images are transformed into YC_BC_R space and the brightness is replaced by the gray fusion images. Experimental results evidence this method which could provide better sense of hierarchy than the global color fusion method in terms of subjective evaluation.

5. Reference

A. Toet, New false color mapping for image fusion, *Opt. Eng*. 1996, 35(3).

A. Toet, et al. 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.

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.

A. Toet, and E.M. Franken. Perceptual evaluation of different image fusion schemes. *Displays*, 2003, 24.

Alexander Toet. Natural color mapping for multiband night vision imagery. *Information Fusion*, 2003, 4:155–166.

A.M. Waxman, et al. Color night vision: opponent processing in the fusion of visible and IR imagery. *Neural Networks*, 1997, 10(1):1–6.

A.M. Waxman. Solid-state color night vision: fusion of lowlight visible and thermal infrared imagery, *MIT Lincoln Laboratory Journal* 11 (1999) 41–60.

C. Wang, and Z.F. Ye. Perceptual contrast-based image fusion: A variational approach, *Acta Automatica Sinica*. 2007, 33(2).

Cvejic N, Lewis J, Bull D, et al. Region-based multimodal image fusion using ICA bases. *IEEE Sensors Journal* 2007, 7(5): 743-751.

D. Barash, D. Comaniciu. A common framework for nonlinear diffusion, adaptive smoothing, bilateral filtering and mean shift. *Image Vision Computing*, 2004, 22(1): 73-81.

D. Fay, et al. Color night vision: Opponent processing in the fusion of visible and IR imagery. *Neural Networks*, 1997, 10(1): 1-6.

D. Malacara. Color Vision and Colorimetry: Theory and Applications. *SPIE Press*, Bellingham, 2002.

D.A. Fay, et al. Fusion of multi-sensor imagery for night vision: color visualization, target learning and search. *Proceedings of the Third International Conference on Information Fusion*, Paris, France, 2000.

E. Reinhard, et al. Color transfer between images. *IEEE Computer Graphics and Applications*, 2001, 21(5):34–41.

Hill P, Canagarajah N, Bull D. Image fusion using complex wavelets, in *Proc. of British Machine Vision Conference*, 2002: 487-496

J. Bezdek. Pattern Recognition with Fuzzy Objective Functions. Plenum, New York, 1981.

J. Schuler, et al.. Infrared color vision: An approach to sensor fusion. *Optics and Photonics News*, August 1998.

J. Weickert, B.M. ter Haar Romeny, M.A. Viergever. Efficient and reliable schemes for nonlinear diffusion filtering. *IEEE Transactions on Image Processing*, 1998, 7:398-410.

J.S. Shi, et al. Objective evaluation of color fusion of visible and IR imagery by measuring image contrast, in *Proceedings of SPIE*, 2005, p.594.

Kingsbury N.G. Complex wavelets for shift invariant analysis and filtering of signals. *Applied and Computational Harmonic Analysis*, 2001, 10(3):234-253.

Kingsbury N.G. Image processing with complex wavelets. *Philosophical Transactions of the Royal Society of London*, 1999, 357: 2543-2560.

Krebs, et al. An evaluation of a sensor fusion system to improve drivers' nighttime detection of road hazards. *Proceedings of the 43 rd Annual Meeting Human Factors and Ergonomics Society*, 1999, p.1333.

Krista Amolins, Yun Zhang, Peter Dare. Wavelet based image fusion techniques—an introduction, review and comparison. *Journal of Photogrammetry & Remote Sensing*, 2007, 62:249–263

L.X. Wang, et al. Color fusion algorithm for visible and infrared images based on color transfer in YUV color space. *Proceedings of SPIE*, 2007, p. 67870S.

M. Aguilar, et al. Field evaluations of dual-band fusion for color night vision, in: J.G. Verly (Ed.), Enhanced and Synthetic Vision 1999, *The International Society for Optical Engineering*, Bellingham, WA, 1999, 168-175.

M. Nielsen, et al. Scale-Space Theories in Computer Vision, *Lecture Notes in Computer Science*, 1999, vol. 1682.

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

N. Otsu. A threshold selection method from gray-level histograms, *IEEE Trans. System, Man and Cybernetics*. 1979, 9(1).

O. Scherzer, J. Weickert. Relations between regularization and diffusion filtering. *Journal of Mathematical Imaging and Vision,* 2000, 12:43-63.

P. Hill, N. Canagarajah, D. Bull. Image fusion using complex wavelets. *Proceedings of the 13th British Machine Vision Conference*, Cardiff, UK, 2002.

Piella G. A general framework for multiresolution image fusion: from pixels to regions. *Information Fusion*, 2003, 4(4): 259-280.

Piella G. A region-based multiresolution image fusion algorithm. *Proceedings of the Fifth International Conference on Information Fusion*. 2002, 1557-1564.

R. O_Callaghan, D.R. Bull. Combined morphological–spectral unsupervised image segmentation. *IEEE Transactions on Image Processing*. 2005, 14(1):49–62.

Ranchin, T., Wald, L. Fusion of high spatial and spectral resolution images: the ARSIS concept and its implementation. *Photogrammetric Engineering & Remote Sensing*, 2000, 66(1):49–61.

S G Mallat, A theory for multiresolution signal decomposition: The wavelet representation, *IEEE Trans. PAMI*, 1989, 11(7):674-693

S. Li, J.T. Kwok, Y. Wang. Combination of images with diverse focuses using the spatial frequency, *Information Fusion*, 2001, 2(3):169–176.

Toet A, van Ruyven J J, Valeton J M. Merging thermal and visual images by a contrast pyramid. *Optical Engineering*, 1989, 28(7): 789-792.

V. Tsagaris, V. Anastassopoulos. Fusion of visible and infrared imagery for night color vision. *Displays*. 2005, 26:191–196.

V. Tsagaris and V. Anastassopoulos. A global measure for assessing image fusion methods. *Opt.Eng*. 2006, 45(2).

V. Tsagaris. Objective evaluation of color image fusion methods. *Opt.Eng*. 2009, 48(6).

W. Pedrycz. Conditional fuzzy c-means, *Pattern Recognition Letters*. 1996, 17:625– 631.

Wu, J. et al. Remote sensing image fusion based on average gradient of wavelet transform. *IEEE International Conference on Mechatronics and Automation*, 2005:1817–1821.

Yufeng Zheng, Edward A. Essock. A local-coloring method for night-vision colorization utilizing image analysis and fusion. *Information Fusion*, 2008, 9:186–199.

Z. Wang, et al. Image quality assessment: From error visibility to structural similarity. *IEEE Transactions on Image Processing*, 2004, 13(4):600-612.

Z.H. Xie, and T. G. Jr. Stockham.Toward the unification of three visual laws and two visual models in brightness perception, *IEEE Trans. System, Man and Cybernetics*. 1989, 19(2).

Zhang Junju, et al. An Image Fusion Method based on Region Segmentation and Complex Wavelets. *Proc. of SPIE*. Vol. 7384, 2009, 738421-1—738421-8.

Zhang Junju, et al. Real-time Color Image Fusion for Infrared and Low-light-level Cameras. *Proc. of SPIE*. Vol. 7383, 2009, 73833B-1—73833B-7.

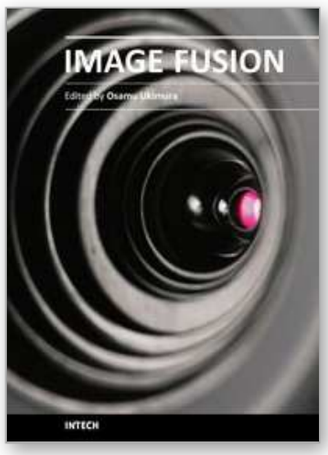


Image Fusion
Edited by Osamu Ukimura

ISBN 978-953-307-679-9
Hard cover, 428 pages
Publisher InTech
Published online 12, January, 2011
Published in print edition January, 2011

Image fusion technology has successfully contributed to various fields such as medical diagnosis and navigation, surveillance systems, remote sensing, digital cameras, military applications, computer vision, etc. Image fusion aims to generate a fused single image which contains more precise reliable visualization of the objects than any source image of them. This book presents various recent advances in research and development in the field of image fusion. It has been created through the diligence and creativity of some of the most accomplished experts in various fields.

INTECH
open science | open minds

InTech Europe
University Campus STeP Ri
Slavka Krautzeka 83/A
51000 Rijeka, Croatia
Phone: +385 (51) 770 447
Fax: +385 (51) 686 166
www.intechopen.com

InTech China
Unit 405, Office Block, Hotel Equatorial Shanghai
No.65, Yan An Road (West), Shanghai, 200040, China
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元
Phone: +86-21-62489820
Fax: +86-21-62489821

IntechOpen