Comparative Analysis of Image Fusion Techniques

Petr Skonnikov [1]

[1] *JSC «S&P Complex «Alpha-M», Sviazy, 25, Ryazan, 390047, Russian Federation*

Abstract

The relevance of multispectral image fusion problem during search and rescue operations is shown. Well-known algorithms for multispectral image fusion are considered and implemented. The comparison involved algorithms based on averaging, maximum method, analysis of low and high frequency components, assessment of information content, addition of differences, extraction of local contrasts, Laplace pyramid, wavelet transform, principal component analysis, 3D low pass filter, power transformation, tv channel priority, Pytyev morphology, diffuse morphology and local weighting summation. Based on publicly available multispectral image datasets, a combined database to compare the algorithms considered including 496 pairs of images has been compiled. The results of image fusion using the considered algorithms are obtained. The aim of the work is to compare well-known image fusion algorithms in terms of objective quality metric. The comparison of fusion results was carried out according to combined quality metric. Based on comparison results, the authors concluded that the best values of combined quality metric for multispectral image fusion are provided by the algorithms based on local weight summation, principal component analysis and Laplace pyramid.

Keywords

Multispectral image fusion, local weighted summation, infrared range, enhanced vision.

1. Introduction

Search and rescue operations are often complicated by adverse weather conditions. In this case, the view of ambient terrain can be difficult due to such interfering factors as, e.g., rain, snow, fog and smoke, and at night – underlighting.

To provide all-weather and round-the-clock ambient terrain view, multispectral enhanced vision systems are used. Such systems form an image based on the information received from the sensors of various spectral ranges. In such systems, two or three channels from the following spectral ranges are usually used: visible light (wavelengths 380-780 nm), long-wave infrared range (LWIR, wavelength from 8 to 15 nm), medium-wave infrared range (MWIR, from 3 to 8 nm), short-wave infrared (SWIR, from 1.4 to 3 nm), as well as near infrared (NIR, from 0.75 to 1.4 nm) [1]. The image of visible light range is familiar to human perception, but is most susceptible to the influence of mentioned interfering factors. NIR and SWIR sensors are usually used for night vision. In addition, in these ranges, the interfering effect caused by smoke and precipitation is reduced. In MWIR range, objects heated to several hundred degrees or more are clearly visible. LWIR range sensors (thermal imagers) receive objects thermal radiation in a wide temperature range. The images obtained from thermal imagers practically do not depend on external illumination of the scene observed and can differ significantly from the images of other spectral ranges.

To combine the benefits of different spectral ranges, a fused image resulting from the combination of images of different channels is displayed to an operator. At present, various algorithms to solve the problem of image fusion [1-10] have been developed. Such algorithms are used to solve various problems, while the problem of how applicable each of these algorithms is in the field of search and

GraphiCon 2022: 32nd International Conference on Computer Graphics and Vision, September 19-22, 2022, Ryazan State Radio Engineering University named after V.F. Utkin, Ryazan, Russia
EMAIL: skonnikovpn@yandex.ru (P. Skonnikov)
ORCID: 0000-0002-4989-3757 (P. Skonnikov)

rescue operations has not been studied. Thus, the problem of comparing the known multispectral image fusion algorithms concerning improved vision system for search and rescue operations is relevant.

2. Image fusion algorithms

It is important to note that not all of the known image fusion algorithms allow using three or more source channels. In this regard, in order to compare a larger number of algorithms, only two source channels were combined. For the experimental studies described below, the following image fusion algorithms have been implemented.

1. Averaging

According to this simplest algorithm, the pixels of resulting image I_{C} are represented by the mean of source images I_1 and I_2 [2]:

$$I_{\mathrm{C}}(x,y)=\frac{I_1(x,y)+I_2(x,y)}{2} \tag{1}$$

where x and y are pixel coordinates.

2. Maximum Method

When fusing images by the maximum method, the brightest pixel from the images of source channels is selected:

$$I_{\mathrm{C}}(x,y)=\max\left\{I_1(x,y),I_2(x,y)\right\} \tag{2}$$

3. Analysis of Low and High Frequency Components

The images coming from the sensors are divided into high and low frequencies by convolution of the source images with spatial low-pass filter and subtracting the filtering result from source images. Fusion of channels is performed separately for high and low frequencies. In both cases, it is possible to use different image fusion algorithms. In further studies, the implementation of the algorithm described in [2] was used.

4. Assessment of Information Content

The image obtained by the algorithm based on the assessment of information content is determined by weight summation of source images with coefficients σ_1 and σ_2:

$$I_{\mathrm{C}}(x,y)=\frac{\sigma_1 I_1(x,y)+\sigma_2 I_2(x,y)}{\sigma_1+\sigma_2} \tag{3}$$

Coefficients σ_1 and σ_2 are the same for all pixels, but they are recalculated in each frame. This calculation is based on calculating standard deviation of pixel brightness in windows sliding over the image. In each channel, weight σ is taken equal to maximum standard deviation for different positions of a sliding window [2].

5. Addition of Differences

At the first stage of image fusion algorithm based on the addition of differences, difference map D is calculated [2]:

$$D(x,y)=\left|I_1(x,y)-I_2(x,y)\right| \tag{4}$$

Further operations can be represented as a weighted sum of pixels:

$$I_{\mathrm{C}}(x,y)=f(x,y)I_1(x,y)+\left(1-f(x,y)\right)I_2(x,y) \tag{5}$$

where $f(x,y)$ is a piecewise linear function of current element of a difference map $D(x,y)$.

Depending on the value of $D(x,y)$, the resulting pixel $I_{\mathrm{C}}(x,y)$ takes on the values of either $I_1(x,y)$ or $I_2(x,y)$, or the result of their weight summation.

6. Extraction of Local Contrasts

In a coordinate plane of two original channels I_1OI_2, the points of brightness values in each pixel $\left\{I_1(x,y),I_2(x,y)\right\}$ and average brightness in the vicinity of corresponding pixels $\left\{\overline{I}_1(x,y),\overline{I}_2(x,y)\right\}$ are marked. Then, at a certain distance in a given direction φ from the point of averaged brightness, a reference point is plotted. Angle φ is the same for all pixels and must be determined in each frame or periodically after several frames. The brightness of the pixel in a complexed image $I_C(x,y)$ is calculated using a piecewise linear function of the distance between a reference point and an original point $\{I_1(x,y), I_2(x,y)\}$ [2,3].

The best angle φ is determined by enumerating different values of φ and comparing resulting images I_C according to some objective quality metric [3].

7. Laplace Pyramid and Wavelet Transform

The approaches based on the Laplace pyramid and wavelet transform can be represented as the rules for decomposing the original images into levels and components. For the individual components of two channels, it is possible to use different fusion algorithms. In particular, in [2], for the fusion using Laplace pyramid, maximum method and averaging were used, and for the fusion using wavelet transform, averaging and selection of a predetermined channel were used. For the experiment described below, we used the implementation of algorithms based on the Laplace pyramid and wavelet transforms available in image fusion toolbox [4].

8. Principal Component Analysis

Image fusion based on principal component analysis is based on decreasing the number of input data dimensions by discarding the least significant components. Multispectral images are reflected from three-dimensional space (width, height, channels) to two-dimensional (width, height) one. In this work, we used open implementation of this algorithm [4].

9. 3D Low Pass Filter

Paper [5] describes an operating principle of image fusion algorithm based on a three-dimensional low-pass filter and an example of its implementation in a frequency domain for the case of visible and thermal ranges. In further experiments, the implementation of this filter by convolution method in spatial domain was used.

10. Power Transformation

The algorithm based on power transformation [6] assumes obtaining a resulting image according to the following formula:

$$I_C(x,y) = I_1(x,y)\exp\left(1 - I_2(x,y)/256\right) \tag{6}$$

followed by reducing image $I_C(x,y)$ to the required brightness range.

11. TV Channel Priority

Image fusion with the priority of a TV channel is performed according to the following formula [7]:

$$I_C(x,y) = I_1(x,y) + \left|I_2(x,y) - \overline{I}_2\right| - \overline{\Delta} \tag{7}$$

where I_1 is a visible image, I_2 is a thermal image, $\overline{I}_2$ is average brightness of image I_2, and $\overline{\Delta}$ is the average value of absolute deviation in image brightness I_2.

12. Pytyev Morphology and Diffuse Morphology

In this work, algorithms based on Pytyev morphology and diffuse morphology are implemented not in exact accordance with [8, 9], but are based on general ideas of these algorithms. In both cases, fusion is carried out according to the formula:

$$I_C(x,y) = I_1(x,y) + \left|P_{12}(x,y) - I_2(x,y)\right| \tag{8}$$

where $P_{12}(x,y)$ is the result of linear filtering of image I_2 by image shape I_1:

$$P_{12}(x,y) = \sum_{i,j\in\Lambda_{x,y}} I_2(x,y) K_D(x,y,i,j) \tag{9}$$

Here $\Lambda_{x,y}$ is the set of pixels adjacent for the pixel with coordinates x and y, and $K_D(x,y,i,j)$ is the result of normalization of heat kernel $K_H(x,y,i,j)$.

To use a particular case of Pytyev morphology in this work, heat kernel K_H is determined as follows:

$$K_H(x,y,i,j) = \begin{cases} 1, \text{ if } I_1(x,y) = I_1(i,j); \\ 0, \text{ otherwise.} \end{cases} \tag{10}$$

The choice of the method for determining heat kernel of type

$$K_H(x,y,i,j) = \exp\left(-\left(I_1(x,y) - I_1(i,j)\right)^2\right) \tag{11}$$

leads to a special case of diffuse morphology.

13. Local Weighting Summation

According to this algorithm, the resulting image is the result of weight summation of original images, and weight coefficients are not the same for all image pixels, but are determined from normalized and smoothed weight maps Ω_1 and Ω_2:

$$I_{\mathrm{C}}(x,y) = \Omega_1(x,y)I_1(x,y) + \Omega_2(x,y)I_2(x,y) \tag{12}$$

Maps Ω_1 and Ω_2 are calculated based on analyzing the information content of individual fragments in original images of all channels. The operating principle of this algorithm is described in more detail in [10, 11].

3. Experimental Studies

A comparison of the algorithms for multispectral image fusion is realized by assessing the quality of fused images obtained using these algorithms. Since the desired result of fusion is unknown a priori, we will consider non-reference quality metrics only. At the same time, in order to assess the effectiveness of each algorithm, it is necessary to take into account not only fusion result I_{C}, but also source images I_1 and I_2, i.e., fusion quality metrics can be represented as the dependence on three images $Q(I_1, I_2, I_{\mathrm{C}})$. Such quality metrics can be objective or subjective.

Subjective quality metrics are the result of processing expert assessments. Subjective assessment shows the correspondence of given fusion algorithm to the verbal formulation of fusion goal. The main disadvantage of subjective assessment is the difficulty of processing large image datasets by experts in a short time. Objective image fusion quality metrics have a mathematical formalization and can be implemented in software. In this regard, the values of such quality metrics can be automatically calculated on a computer.

At present, various non-reference objective image fusion quality metrics are known. A review and comparison of such algorithms is given, for example, in [12, 13]. For further research, combined objective quality metric Q_{C} proposed in [14] was used. This metric is consistent with subjective quality assessment. Such consistence is achieved due to weighted summation of the components in various known objective metrics of image fusion quality, in particular, the Piella metric [14].

To compare the algorithms discussed above in terms of Q_{C} metric, a dataset of multispectral images has been created. This dataset is based on publicly available datasets: TNO [15], OSU [16] and VAIS [17]. According to the selection results, 496 pairs of visible and infrared images are included in combined dataset.

For each i-th pair of images I_{1i} and I_{2i}, the fusion was carried out using each of the algorithms considered above, as a result of which the resulting images $I_{\mathrm{C}i}$ were obtained. Next, the values of combined fusion quality objective metric $Q_{\mathrm{C}i}(I_{1i}, I_{2i}, I_{\mathrm{C}i})$ were calculated.

For each fusion algorithm, minimum and average values of combined objective quality metric $Q_{\mathrm{C\,min}}$ and $\overline{Q}_{\mathrm{C}}$ were calculated, respectively:

$$Q_{\mathrm{C\,min}} = \min_i \{Q_{\mathrm{C}i}\} \tag{13}$$

$$\overline{Q}_{\mathrm{C}} = \frac{1}{496}\sum_{i=1}^{496} Q_{\mathrm{C}i} \tag{14}$$

Consideration of not only averaged, but also minimum values of quality metric is necessary in order to determine the susceptibility of a particular image fusion algorithm to special situations for which it is not possible to ensure the visibility of all objects.

In addition, this experiment was carried out not only for images I_{C}, obtained directly as a result of fusion algorithm, but also for corresponding images with increased contrast to bring them to a full range of brightness from 0 to 255.

The results of calculating the specified quality metrics are shown in Table 1.

Table 1
The results of calculating

Image Fusion Algorithm	With Contrast Correction		Without Contrast correction	
	$\bar{Q}_{C}$	$Q_{C\,min}$	$\bar{Q}_{C}$	$Q_{C\,min}$
Averaging	0,812	0,585	0,826	0,531
LF and HF analysis	0,796	0,571	0,814	0,517
Assessment of information content	0,834	0,549	0,832	0,506
Adding differences	0,720	0,555	0,712	0,509
Highlighting local contrasts	0,752	0,277	0,738	0,275
Laplace pyramid	0,838	0,653	0,816	0,545
Wavelet transform	0,806	0,595	0,798	0,503
Principal component analysis	0,840	0,394	0,834	0,400
3D low-pass filter	0,780	0,543	0,801	0,492
Maximum method	0,659	0,243	0,673	0,317
Power transformation	0,506	0,067	0,704	0,370
TV channel priority	0,820	0,504	0,808	0,502
Pytyev morphology	0,778	0,304	0,800	0,324
Diffuse morphology	0,800	0,411	0,787	0,410
Local weight summation	0,845	0,585	0,836	0,529

From the analysis of table 1 it follows that in both experiments carried out on combined quality metric, the best averaged fusion quality is provided by an algorithm based on local weight summation of original images [10,11]. Further, in decreasing order of the values of this metric there are algorithms based on the analysis of the main components, based on the Laplace pyramid, based on the assessment of information content, an algorithm with television channel priority and averaging.

4. Conclusion

The choice of the algorithm for multispectral image fusion has a significant impact on object detection during search and rescue operations. According to the results presented in table 1, some image fusion algorithms, although they have high average value of image fusion quality index, lose their efficiency in some special scenes. Among the algorithms in which such an effect manifests itself in the least way, one can list the algorithms based on the Laplace pyramid, wavelet transform, local weight summation and averaging. At the same time, on average for the set of images used, the best results were shown by the algorithm based on local weighting summation.

5. References

[1] A.G. Zotin, A.I. Pahirka, M.V. Damov, E.I. Savchina. Improving the visual quality of images obtained in difficult lighting conditions based on infrared data. Software products and systems, 2016, Vol. 29, No. 3, pp. 109-120 (In Russian).

[2] V.V. Insarov, S.V. Tikhonova, I.I. Mikhailov. Problems of constructing technical vision systems using a fusion of information channels of various spectral ranges. Information Technology, 2014 (Att.), No. 3, pp. 1-32 (In Russian).

[3] V.V. Insarov, K.V. Obrosov, V.Ja. Kim, V.M. Lisitsyn. Formation of integrated television-thermal imaging images in front-view systems of aircraft. Vestnik of computer and information technologies, 2013, No. 4, pp. 3-10 (In Russian).

[4] Functionality of the image fusion toolbox. Metapix, 1999, URL: http://metapix.de/func.htm.

[5] V.N. Drynkin, Je.Ja., Falkov, T.I. Careva. Formation of a combined image in a two-zone airborne aerospace system // Mechanics, Control and Information. 2012. No. 9. pp. 33-39 (In Russian).

[6] A.S. Vasiliev, A.V. Trushkina. Methods of image fusion in multispectral optoelectronic systems. Proc. of “GraphiCon– 2016”, pp. 314-318 (In Russian).
[7] L.N. Kostjashkin, S.I. Babaev, A.A. Loginov, O.V. Pavlov. Technologies of systems of improved/synthesized vision for control of aircraft. Technical vision in control systems of mobile objects – 2010, Moscow, 2011, Vol. 4, pp. 45-56 (In Russian).
[8] A.Ju. Rubis, O.V. Vygolov, Ju.V. Vizilter. Morphological fusion of images of different spectral ranges. Mechanics, Control and Informatics, 2012, No. 8, pp. 143-148 (In Russian).
[9] Ju.V. Vizilter, O.V. Vygolov, S.Ju. Zheltov, A.Ju. Rubis. Fusion of multispectral images for systems of improved vision based on methods of diffuse morphology. Bulletin of the Russian Academy of Sciences, 2016, No. 4, pp. 103-114 (In Russian).
[10] S.N. Kirillov, P.S. Pokrovsky, A.A. Baukov, P.N. Skonnikov. Algorithms for the formation of television images of improved quality for remote control of robotics complexes. Radio Engineering, 2019, Vol. 83, No. 5 (II), pp. 205-214 (In Russian).
[11] S.N. Kirillov, P.N. Skonnikov. The procedure for determining the size of windows for the algorithm of local weight summation of multispectral images. Vestnik of RSREU, 2020, No. 73, pp. 3-9 (In Russian).
[12] S.N. Kirillov, P.S. Pokrovsky, P.N. Skonnikov, A.A. Baukov. Comparison of assessments of the quality of multi-spectral images fusion and algorithms for improving the quality of low-contrast images for video surveillance systems. DSPA-2019 proc., Moscow, 2019. pp. 403-48 (In Russian).
[13] S.N. Kirillov, P.S. Pokrovsky, P.N. Skonnikov, A.A. Baukov. Comparison of algorithms for improving low-contrast images and assessing the quality of multi-spectral images fusion for digital video surveillance systems. Digital Signal Processing, 2019, No. 1, pp. 52-58 (In Russian).
[14] S.N. Kirillov, P.N. Skonnikov. Optimization of the combined objective metric of the quality of complexing images of different spectral ranges. DSPA-2020 proc., Moscow, 2020, pp. 592-596 (In Russian).
[15] A. Toet, J.K. Ijspeert, A.M. Waxman, M. Aguilar. Fusion of visible and thermal imagery improves situational awareness. Displays, 1997, Vol. 18, No. 2, pp. 85-95.
[16] J. Davis, V. Sharma. Background-Subtraction using Contour-based Fusion of Thermal and Visible Imagery. Computer Vision and Image Understanding, 2007, Vol. 106, No. 2–3, pp. 162-182.
[17] M.M. Zhang, J., Choi, K. Daniilidis, M.T. Wolf, C. Kanan. VAIS: A Dataset for Recognizing Maritime Imagery in the Visible and Infrared Spectrums. PBVS-2015 proc., 2015, pp. 10-16.