Optica Applicata, Vol. XXXIV, No. 3, 2004

Multiscale contrast image fusion scheme with performance measures

XINMAN ZHANG, JIUQIANG HAN

School of Electronics and Information Engineering, Xi'an Jiaotong University,
28 Xian'ning West Road, Xi'an 710049, People's Republic of China

A new method of image fusion from various sensing modalities is proposed. This method adopts a perceptual fusion operator by using a sequence of multiscale contrast pyramid images. The method is tested by merging parallel registered visible and infrared images. Several performance measures clearly indicate that this method outperforms the other three approaches producing better visual effects.

Keywords: image fusion, multiscale image analysis, ratio of low-pass pyramid, visual perception, image processing.

1. Introduction

For images of the same scene obtained at different bandwidths, such as visible and infrared images, it is nearly impossible to capture all the details and salient features. Image fusion can be used to integrate different input sources into a single image, which is able to assimilate all these individual features potentially to be useful for human observation [1].

A plethora of algorithms for multiresolution image fusion [2]–[9], such as Laplacian pyramid and wavelet pyramid, have been developed. It is a well-known fact that the human visual system is sensitive to local luminance contrast. Thus the RoLP (ratio of low-pass) pyramid is a better image fusion method compared with other pyramid algorithms in general. But their fusion operators are relatively simple by performing logic or a weighted combination. Therefore, a more elaborate scheme is necessary to improve previous methods.

This paper describes the extension of RoLP pyramidal image decomposition. The proposed algorithm performs a novel fusion operator by judging uniform parameter on the analysis of segmentation techniques. Finally, fusion quality is calculated objectively through three measures that demonstrate the improvements offered by the present scheme over other three fusion approaches. The results all show that this method is perceptually meaningful and explicit.

2. Towards fusion method representation

This section describes an image fusion method based upon a previous successful multiresolution method, with an added ability to tailor the selection criteria to the contrast sensitivity. The basic scheme employs a multiresolution algorithm that uses

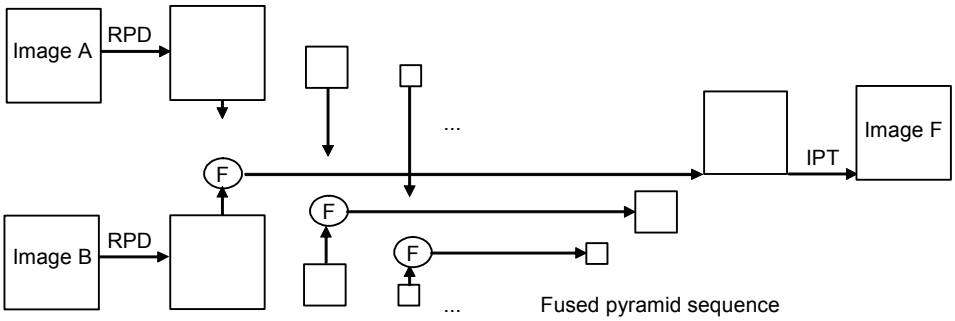

Fig. 1. Schematic diagram of multiscale contrast pyramid image fusion (RPD – RoLP decomposition, IPT – inverse pyramid transform).

decomposition, fusion, and reconstruction. Figure 1 shows a schematic diagram of the basic structure of the proposed image fusion scheme.

2.1. Generating kernel

One important component is the generating kernel ω, which is chosen subject to certain constraints such as separable, normalized, symmetric and equal contribution. Let $\omega(0) = 0.4$, $\omega(1) = \omega(-1) = 0.25$, $\omega(2) = \omega(-2) = 0.05$. Then,

$$\omega = \frac{1}{400}\begin{bmatrix} 1 & 5 & 8 & 5 & 1 \\ 5 & 25 & 40 & 25 & 5 \\ 8 & 40 & 64 & 40 & 8 \\ 5 & 25 & 40 & 25 & 5 \\ 1 & 5 & 8 & 5 & 1 \end{bmatrix}.$$

2.2. Image decomposition

Suppose the image is represented initially by the array G_0 which is the bottom level of the Gaussian pyramid. In a similar way each value within level l, representing G_l, is then obtained from values within level $l-1$ by applying the generating kernel. This is performed as follows, for all nodes i, j:

$$G_l(i,j) = \sum_{m,n=-2}^{2} \omega(m,n)\, G_{l-1}(2i+m, 2j+n). \tag{1}$$

Because of the reduction in the spatial frequency content, each image in the sequence can be represented by an array that has half dimensions of its predecessor. Thus we can expand array G_l into array G^*_{l-1} by interpolating new node values between the given values, which is the same size as G_{l-1}

$$G^*_{l-1}(i,j) = 4 \sum_{m,n=-2}^{2} \omega(m,n)\, G_l\left(\frac{i+m}{2}, \frac{j+n}{2}\right). \tag{2}$$

Only terms for which $(i+m)/2$ and $(j+n)/2$ are integers are included in this sum.

Then the RoLP pyramid C_l is the ratio of two successive levels in the Gaussian pyramid. It is then defined as

$$\begin{cases} C_l = G_l / G^*_l & \text{for} \quad 0 \le l \le N-1, \\ C_N = G_N & \text{for} \quad l = N. \end{cases} \tag{3}$$

2.3. Detail pyramid fusion

Consider a contrast vision system to be the key factor in determining the salient and dominant features of an image [10]. Note that this perceptual information association is supported by human visual system studies and is extensively used in image fusion scheme. Thus, we define $d(B_k)$, namely uniform parameter, the saliency formula for a given level image $G_l(i,j)$ as follows:

$$d(B_k) = \frac{1}{m' \times n'} \sum_{(i,j) \in B_k} \frac{|G_l(i,j) - \mu_k|}{\mu_k} \tag{4}$$

where μ_k is the mean of block B_k, and $m' \times n'$ is the block size.

In the following, we shall assume that there are two inputs, a visible image (VIS) and an infrared (IR) image. Each level of two RoLP pyramids to be fused is decomposed into sets of smaller blocks and compared using uniform parameter response. The block of fused image is formed here as

$$BC_{Fk} = \begin{cases} BC_{\mathrm{VIS}k} & \text{if} \quad d(BC_{\mathrm{VIS}k}) \ge d(BC_{\mathrm{IR}k}) \\ BC_{\mathrm{IR}k} & \text{otherwise} \end{cases} \tag{5}$$

where $d(BC_{\mathrm{VIS}k})$ and $d(BC_{\mathrm{IR}k})$ denote uniform parameters of blocks $BC_{\mathrm{VIS}k}$ and $BC_{\mathrm{IR}k}$ of corresponding image pair $C_{\mathrm{VIS}l}$ and $C_{\mathrm{IR}l}$, respectively. Therefore, image fusion is performed by selecting the feasible blocks according to Eqs. (4) and (5), so this process yields a sequence of merged images $C_{F0}, C_{F1}, \ldots, C_{FN}$.

2.4. Fused image reconstruction

As the final step the reconstructed composite image F_0 can be recovered exactly from its RoLP pyramid representation:

$$\begin{cases} F_N = C_{FN} & \text{for } l = N, \\ F_l = C_{Fl} F_l^* & \text{for } 0 \le l \le N-1. \end{cases} \tag{6}$$

3. Measures and results

In order to objectively evaluate the capabilities of different fusion algorithms, three measures of image fusion performance are provided for some quantitative comparison.

3.1. Entropy

Entropy is known to be a measure of the amount of uncertainty about the image. It is then given by

$$H = -\sum_{i=0}^{L-1} p_i \log_2 p_i \tag{7}$$

where L is the number of gray levels; and note that

$$p_i = \frac{\text{number of pixels } D_i \text{ of each gray level } i}{\text{number of pixels } D \text{ in the image}}.$$

3.2. Cross entropy

Notice that cross entropy measures the difference of two images, so a lesser cross entropy is preferred. It is computed as

$$H_{ce} = -\sum_{i=0}^{L-1} p_{Ri} \log_2 \left(\frac{p_{Ri}}{p_{Fi}} \right)$$

where p_{Ri} and p_{Fi} are the corresponding probability of input and fused image occurring. Here, we choose the root mean square cross entropy of two inputs to the output image, respectively

$$\overline{H_{ce}} = \sqrt{\frac{H_{ce1}^2 + H_{ce2}^2}{2}}. \tag{8}$$

3.3. Mutual information

Mutual information is a measure that determines how much information is obtained from the fusion of input images. We use this measure as the third evaluation method to assess the performance of different image fusion algorithms;

$$MI(R, F) = \sum_{i_1 = 0}^{L-1} \sum_{i_2 = 0}^{L-1} p_{R,F}(i_1, i_2) \log_2 \left(\frac{p_{R,F}(i_1, i_2)}{p_R(i_1) p_F(i_2)} \right)$$

where $p_{R,F}$ indicates the normalized joint gray level histogram of images R and F, p_R and p_F are the normalized marginal histograms of the two images. We choose the average cross entropy of two inputs to the output image, respectively,

$$\overline{MI} = \frac{MI_1 + MI_2}{2}. \tag{9}$$

3.4. Results of different schemes

To verify the proposed approach, four fusion algorithms are tested for objective performance evaluation applied to a pair of visible and infrared registered imagery. Obviously, the two images, as presented in Fig. 2**a** and **b**, have different spectral characteristics and details in the same depicted scene. Both are optical images and are registered with each other. The visible image has a lower contrast; thus the easily discernable finer features and background are almost indistinguishable in the infrared image. On the other hand, the infrared image also has some unique features, such as the textural patterns with a higher contrast, which are present in the visible image. In scheme 1 (Fig. 2**c**), Laplacian algorithm is performed with equal weights (*i.e.*, $w_{\mathrm{VIS}} = w_{\mathrm{IR}} = 0.5$) assigned to a pair of inputs respectively for reconstructing the composite image [3]. Scheme 2 (Fig. 2**d**) employs Toet algorithm with a maximum absolute contrast node selection technique [5]. In scheme 3 (Fig. 2**e**), the fusion rule is defined by calculating the wavelet transform modulus maxima using "Daubechies 8" filter [8]. Scheme 4 (Fig. 2**f**) implements the new algorithm using a decomposition block size of 4×4. All of the four schemes choose the third layer as the fusion level to compare

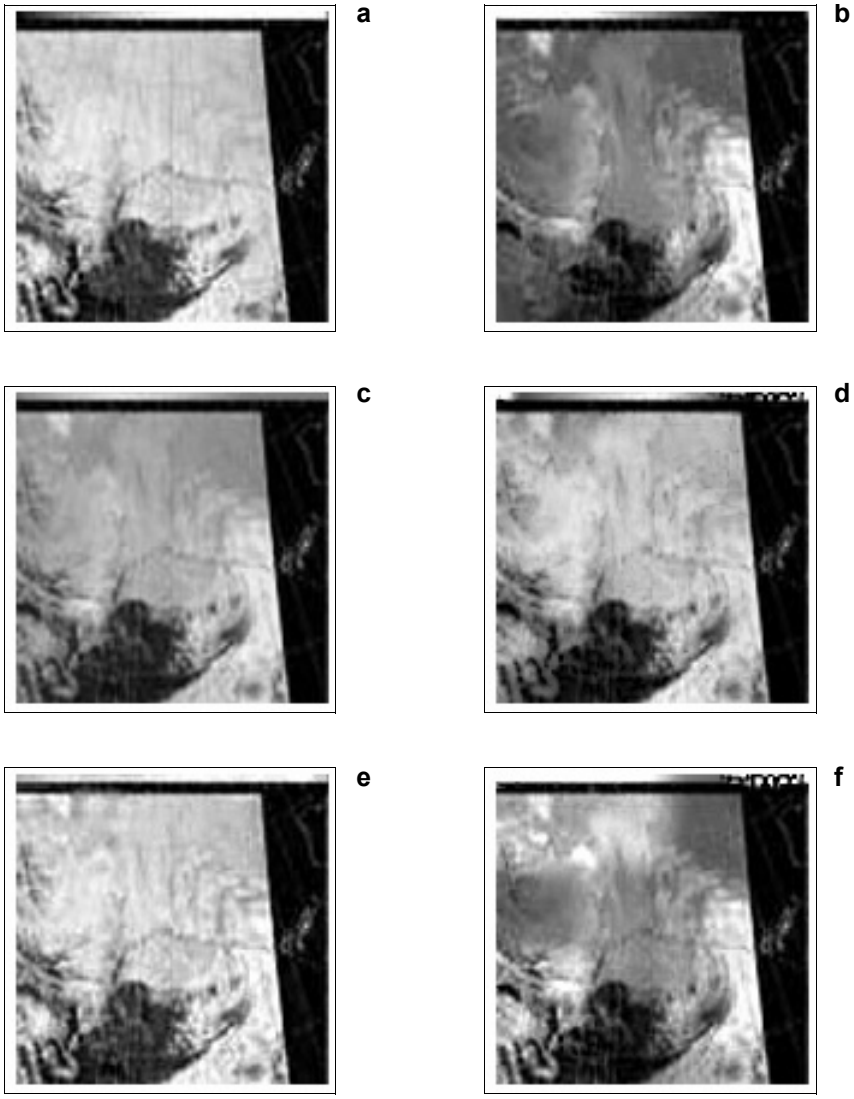

Fig. 2. Input visible and infrared images and fused output images of different schemes: visible image (**a**), infrared image (**b**), fused output image of scheme 1 (**c**), fused output image of scheme 2 (**d**), fused output image of scheme 3 (**e**), fused output image of the proposed algorithm – scheme 4 (**f**).

these methods for consistency verification. In general, the third layer is a better fusion level proved by more tested cases. From the observer's viewpoint, by comparing the fused images all four schemes do a good job of preserving visual information from each input image. However, Fig. 2**c** and 2**d** do not have the same amount of details compared to other figures. The reduction in intensity then causes the other areas of the image to become less contrasted in Fig. 2**e**. Obviously, Fig. 2**f** contains more details

T a b l e 1. Image fusion quality measures.

	Entropy	Cross entropy	Mutual information
Scheme 1	7.4965	0.7317	2.7969
Scheme 2	7.3801	0.7284	2.7532
Scheme 3	7.5523	0.8758	2.6221
Scheme 4	7.6384	0.4909	2.8530

of progressively finer resolution and greater contrast enhancement than other figures, which especially appears to preserve features in the input images that are dominant.

At the same time, this illustrative example is provided here for some quantitative comparison of four fusion schemes in Tab. 1. According to three evaluation criteria, the proposed method achieves the best entropy, cross entropy and mutual information. It is superior to the traditional Laplacian pyramid, ratio of low-pass pyramid and wavelet transform fusion method, with 1.1–3.4% improvement of entropy, 48–78% reduction in cross entropy and 3.5–8.1% enhancement in mutual information through further computations. The evaluation results coincide with the visual effect very well. Very clearly, the result convincingly demonstrates that the image in Fig. 2**f** contains more perceptual details and features than other figures. Overall the represented

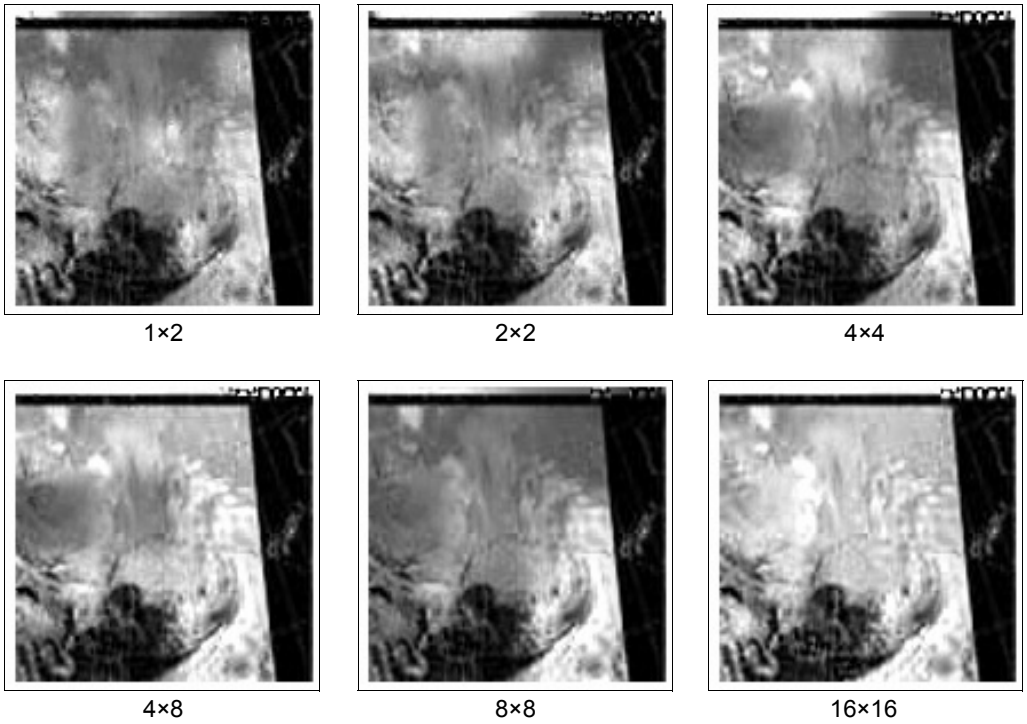

Fig. 3. Fused images formed by different size of blocks.

approach explicitly shows better fusion performance, both esthetically and numerically. Thus, we can see that image fusion is not different from singular image enhancement at all, which makes full use of complementary or redundant information of input images to acquire a synergistic combination. To prove the correctness and universality of this algorithm, more cases are tested dealing with visible and infrared images, judged by the above measures, with satisfying results derived.

Here, we discuss the influence of the size of blocks in detail. We choose mainly different decomposition block sizes to test it. Figure 3 shows the fused effects from different decomposition block sizes. At the same time, Tab. 2 shows performance measures for quantitative comparison. Block size 4×4 is the most optimal as a whole. In general, if the block size is too large, a particular block may lead to aberration of image gray level gradients. On the other hand, using a very small block size may lead to the saw-tooth effect. Moreover, a large amount of image experiments suggest that a better fusion result is quite sensitive and suitable to this block size 4×4, and this conclusion is stable and consistent.

T a b l e 2. Influence of the size of blocks on fusion performance.

Block sizes	Entropy	Cross entropy	Mutual information
1×2	7.5787	0.5678	2.6558
2×2	7.5921	0.5222	2.6078
4×4	7.6384	0.4909	2.8530
4×8	7.6336	0.4740	2.6693
8×8	7.5305	0.6960	3.0559
16×16	7.4901	0.7666	2.6744

There is an important issue that should be investigated in the future. If one of the images has much smaller local contrast than the other one, the algorithm may have some restrictions because of the uniform parameter d, which is not considered here. Nevertheless, preliminary experiments suggest that the scheme is suitable to the visible and infrared images in general, provided that the local contrast difference between the two images is not too extreme.

4. Conclusions

A visual perception-based multiscale contrast image fusion scheme is presented for multispectral image data. No assumption is made regarding the nature of the relation between the intensities in both input modalities. The desired visual improvements over image fusion agree remarkably well with that obtained from objective results in comparison with other methods. Detection, recognition, and search tasks can therefore benefit considerably from this new image representation.

References

[1] VARSHNEY P.K., Electron. Commun. Eng. J. **9** (1997), 245.
[2] BURT P.J., ADELSON E.H., Proc. SPIE **575** (1985), 173.
[3] BURT P.J., KOLCZYNSKI R.J., *Enhanced image capture through fusion*, [In] *IEEE 4th International Conference on Computer Vision*, Vol. 4 (1993), pp. 173–182.
[4] TOET A., RUYVEN L.J., VALETON J.M., Opt. Eng. **28** (1989), 789.
[5] TOET A., Opt. Eng. **31** (1992), 1026.
[6] LIU G.X., YANG W.H., Acta Opt. Sin. **21** (2001), 1336.
[7] YOCKY D.A., J. Opt. Soc. Am. A **12** (1995), 1834.
[8] LI H., MANJUNATH B.S., MITRA S.K., Graphical Models and Image Processing **57** (1995), 235.
[9] WILSON T., ROGERS S., MYERS L., Opt. Eng. **34** (1995), 3154.
[10] LI S.T., KWOK J.T., WANG Y.N., Pattern Recognition Letters **23** (2002), 985.

Received November 7, 2003
in revised form March 8, 2004