D^{2}MDFN: Data-Driven based Mask Deep Fusion Network for Visible and Infrared Image Fusion Without Reference Ground-Truth

Paper ID: 675

Abstract

In this complementary file, we display some further experimental results obtained by different fusion methods and D^{2}MDFN. Specifically, we show four pairs of images as well as their corresponding quantitative metrics to demonstrate the superiority of our method. Furthermore, additional experiments are displayed to emphasize the significance of the mask, SSIM and gradient strategies.

1 Experiments

1.1 Comparison with state-of-the-art

The images fused by various methods are shown in Fig.1. It is easy to observe that the proposed approach D^{2}MDFN always achieves the outstanding performances. Compared with CVT, GTF, ASR, CSR and FPDE, D^{2}MDFN is not only capable of generating more remarkable interested pixels existing in the infrared image, but also obtaining much clearer texture details. In contrast to GFF, although it gains a satisfactory performance in the human visual perception, it almost fails to get the thermal radiation information, being far from our requirement. Besides, the quality of the fused images computed by FusionGAN is also inferior to that obtained by D^{2}MDFN in both thermal radiation and texture details.

The quantitative comparisons associated with the fused images in Fig.1 are listed in Tab.1. Obviously, our presented network achieves the noticeable performances in most of cases. In particular, for the first and fourth pairs, except AG, D^{2}MDFN are superior to other comparison methods in all remaining metrics. For the second and third pair, there are also competitive improvements compared with these comparison approaches.

1.2 Algorithm Analysis

Fig.2 shows another example if the mask or SSIM term is removed from the proposed method. We can see that, without the mask strategy, the texture part in the generated image is much blurry than that in the original one. Compared with Fig.2(a2) and (b2), the fused image would corrupted by some artifacts if the SSIM loss is removed from the objective function.

We also conduct an experiment to show the importance of the gradient strategy, as displayed in Fig.3. As we can see, if the gradient strategy is removed, the generated images would meet a quality degradation in the visual perception.

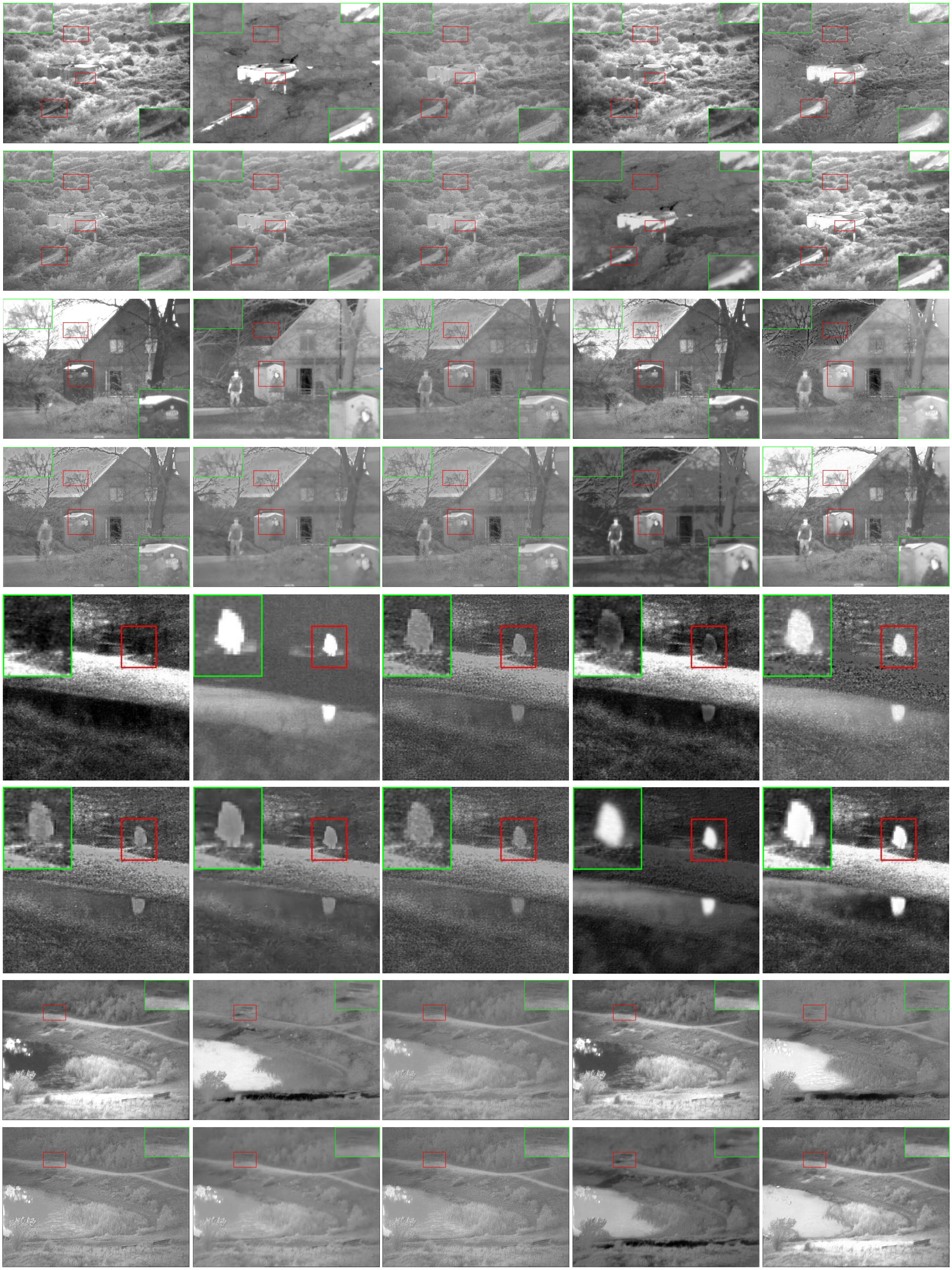

Figure 1: Results obtained by different methods. For each pair of images, the first two are visible and infrared source images; the subsequent images are obtained by CVT, GFF, GTF, ASR, CSR, FPDE, FusionGAN and D^{2}MDFN from left to right and top to down.

Table 1: Quantitative results on the fused images obtained by different methods. Red means the maximum and blue is the second largest. The sub-tables are associated with the images in Fig.1.

	First					Second				
Methods	EN	MI	$Q^{AB/F}$	AG	VIF	EN	MI	$Q^{AB/F}$	AG	VIF
CVT	6.7560	1.4041	0.4281	5.0515	0.2256	6.5019	1.6629	0.4245	3.9538	0.3043
GTF	6.7800	1.0904	0.5008	5.2657	0.1547	7.0941	2.4471	0.4375	3.8580	0.2172
ASR	6.8370	1.3922	0.5438	5.6593	0.2217	6.5639	1.5953	0.5268	4.2549	0.2809
CSR	6.7206	1.4787	0.2450	2.6303	0.1970	6.4179	1.8033	0.2792	2.1856	0.2643
FPDE	6.7601	1.4013	0.4897	5.2193	0.2161	6.4846	1.6917	0.4637	3.6299	0.2816
FusionGAN	6.4505	1.6789	0.1768	2.7768	0.0570	6.4955	1.9335	0.2357	2.5410	0.2019
D^{2}MDFN	**7.3965**	**3.8793**	**0.6441**	**5.5704**	**0.3308**	**7.0828**	**3.5970**	**0.5252**	**3.9774**	0.2803
	Third					Fourth				
Methods	EN	MI	$Q^{AB/F}$	AG	VIF	EN	MI	$Q^{AB/F}$	AG	VIF
CVT	6.7711	1.9407	0.5778	8.9009	0.1968	6.5834	1.7911	0.4486	3.6014	0.2730
GTF	6.7781	1.5464	0.5836	8.8087	0.1268	6.6217	2.0167	0.4866	3.5709	0.1731
ASR	6.8543	1.9155	0.6597	9.3405	0.2024	6.6239	1.7566	0.5469	3.8877	0.2827
CSR	6.5160	2.2816	0.3424	4.4886	0.1590	6.5309	1.9452	0.2145	1.6884	0.2079
FPDE	6.7133	2.1447	0.6109	8.0106	0.1738	6.5798	1.8106	0.5050	3.5761	0.2657
FusionGAN	6.4860	1.9262	0.2274	4.6864	0.1008	6.5194	2.1356	0.2452	2.2184	0.1506
D^{2}MDFN	**7.4041**	**3.9457**	**0.6406**	8.5894	**0.3134**	**7.3262**	**4.5828**	**0.5648**	3.4465	**0.2974**

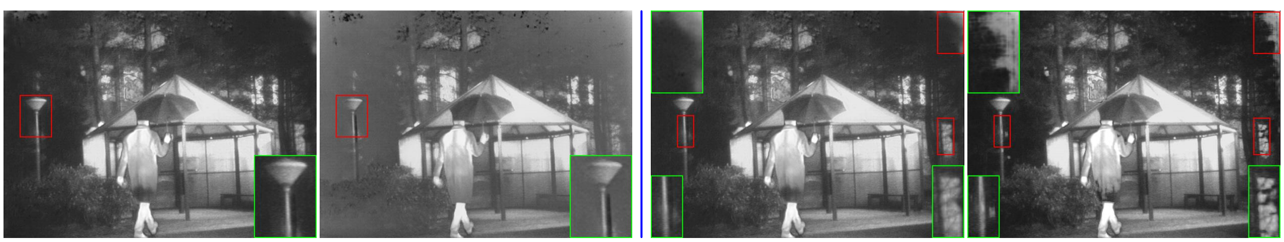
(a1) (b1) (a2) (b2)

Figure 2: (a1, a2) are obtained by the original D^{2}MDFN. (b1) are obtained by D^{2}MDFN when the mask is removed. (b2) are obtained by D^{2}MDFN when the SSIM loss is removed.

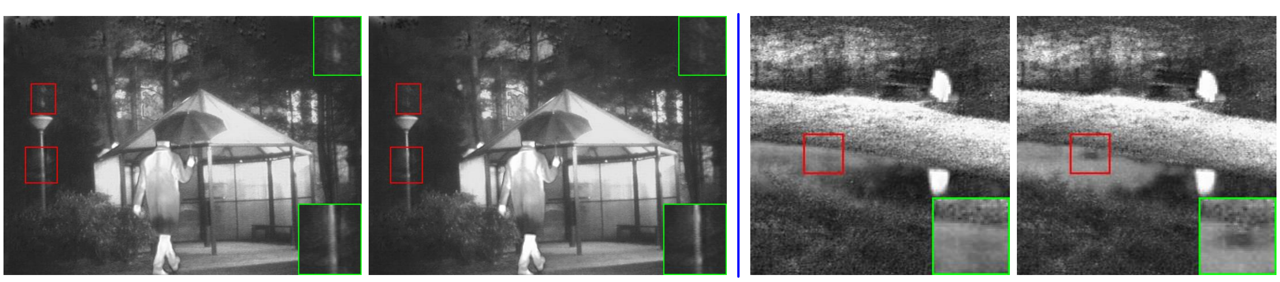
(a) (b) (c) (d)

Figure 3: (a) and (c) are obtained by the original D^{2}MDFN. (b) and (d) are obtained by D^{2}MDFN when the gradient strategy is removed.