Fractional Optimization Model for Infrared and Visible Image Fusion

Kang Zhang, Shiwei Wu, Zhiliang Wu, Xia Yuan*, Chunxia Zhao

南京理工大学 NANJING UNIVERSITY OF SCIENCE & TECHNOLOGY

BMVC 2022

Abstract

Infrared and visible image fusion is a fundamental task for image processing to enhance the image quality. To highlight target and retain effective details, different from previous methods using integer gradients, we use the fractional gradient to well represent image features and propose a novel fractional optimization model to fuse infrared and visible images. Specially, a fractional optimization function is designed with global contrast fidelity and fractional gradient constraint to obtain the pre-fused image. Then, the base layer of the pre-fused image obtained by multi-level decomposition latent low-rank representation is taken as the fused base layer, while for the fusion of detail layers, a fractional gradient energy function is designed to evaluate the importance of detail information to generate the fused detail layers. Compared with 10 state-of-the-art image fusion methods qualitatively and quantitatively on two public datasets (TNO and RoadScene), our method generally shows superior performance.

Motivation

- To alleviate the problem that the general base layer fusion rules easy to ignore the global contrast, we regard the fusion of the base layer as an optimization problem, which uses fractional gradient to better represent image features.
- To sufficiently extract the useful detail information, a fractional gradient energy function is designed to distribute the weight of detail information and generate the fused detail layers.

Model architecture

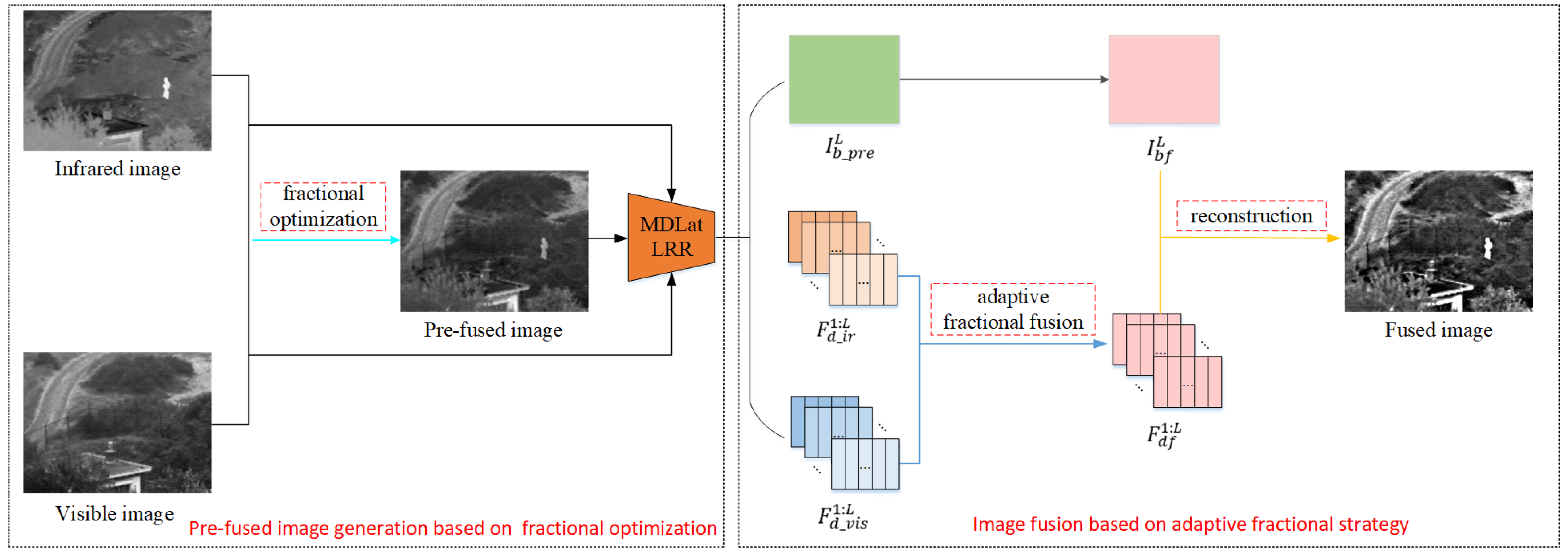

Fig. 1. The overview of the proposed method. Specially, a optimization function is designed with global contrast fidelity and fractional gradient constraint, and a Sylvester-based optimization method is used to obtain the pre-fused image. Then, the base layer of the pre-fused image obtained by MDLatLRR is taken as the fused base layer. For the fusion of detail layers, the fractional gradient energy function is introduced to evaluate the importance of detail information to generate the fused detail layers. Finally, the fused image is reconstructed by the fused base layer and detail layers.

The fusion of base layer

$$f_{pre} = \arg\min_f \frac{1}{2}\|f - f_{ir}\|_F^2 + \frac{\mu}{2}\left\|D^{v_f} f - D^{v_{f_{ir}}} f_{ir} - D^{v_{f_{vis}}} f_{vis}\right\|_F^2,$$

$$I_{bf} = \mathrm{MDLatLRR}(f_{pre}).$$

$D^v s$ denotes the v-order discrete fractional gradient of $s \in \mathbb{R}^{n\times n}$.

The fusion of detail layers

$$FGE^i_{dk} = \sum_{x=1}^{M-1}\sum_{y=1}^{N-1}\{[G_1 * F^i_{dk}(m,n)]^2 + [G_2 * F^i_{dk}(m,n)]^2\},$$

$$\omega^i_{dk}(m,n) = \frac{FGE^i_{dk}(m,n)}{\sum_{k=1}^{K} FGE^i_{dk}(m,n)},$$

$$F^i_{df}(m,n) = \sum_{k=1}^{K} \omega^i_{dk}(m,n) \times F^i_{dk}(m,n),$$

$$I^i_{df} = \mathrm{R}\,(F^i_{df}), i = [1,2,\cdots,L].$$

The fused image

$$I_f = I_{bf} + \sum_{i=1}^{L} I^i_{df}.$$

Ablation study

Table 3: Analysis on the key components of the proposed method on TNO dataset.

Methods	SD	CC	MS-SSIM	SCD
MDLatLRR [8]	83.7741	0.4393	0.8517	1.6332
MDLatLRR+*model-1*	88.9609	0.4626	0.8567	1.7399
MDLatLRR+*model-2*	91.3343	0.4225	0.7877	1.5675
Ours	93.1529	0.4880	0.9281	1.8291

Table 4: Fusion results on TNO dataset using different parameter combinations ($v_{f_{ir}}$,$v_{f_{vis}}$).

$v_{f_{ir}}$	$v_{f_{vis}}$	EN	SD	MI	MS-SSIM	SCD
0.3	0.3	6.8174	72.2746	13.6348	0.9281	1.8291
	0.6	6.9920	87.9021	13.9841	0.9196	1.7472
	0.9	7.0294	93.1529	14.0588	0.9065	1.5680
0.6	0.6	6.8527	74.1245	13.7054	0.9267	1.7948
	0.9	6.8359	77.1384	13.6718	0.9175	1.6886
	1.2	6.7987	76.3516	13.5973	0.9167	1.7118
0.9	0.9	6.8537	79.2122	13.7074	0.9231	1.7675
	1.2	6.8265	77.4651	13.6530	0.9222	1.7574
	1.5	6.8398	77.8568	13.6796	0.9216	1.7610
1.2	1.2	6.8243	77.0069	13.6486	0.9236	1.7516
	1.5	6.8292	78.3813	13.6584	0.9237	1.7552
	1.8	6.8427	78.5622	13.6854	0.9235	1.7604
1.5	1.5	6.8210	76.7004	13.6421	0.9234	1.7483
	1.8	6.8285	77.4367	13.6570	0.9233	1.7521
1.8	1.8	6.8216	76.9553	13.6433	0.9230	1.7494

Experiments

Performance comparison on TNO

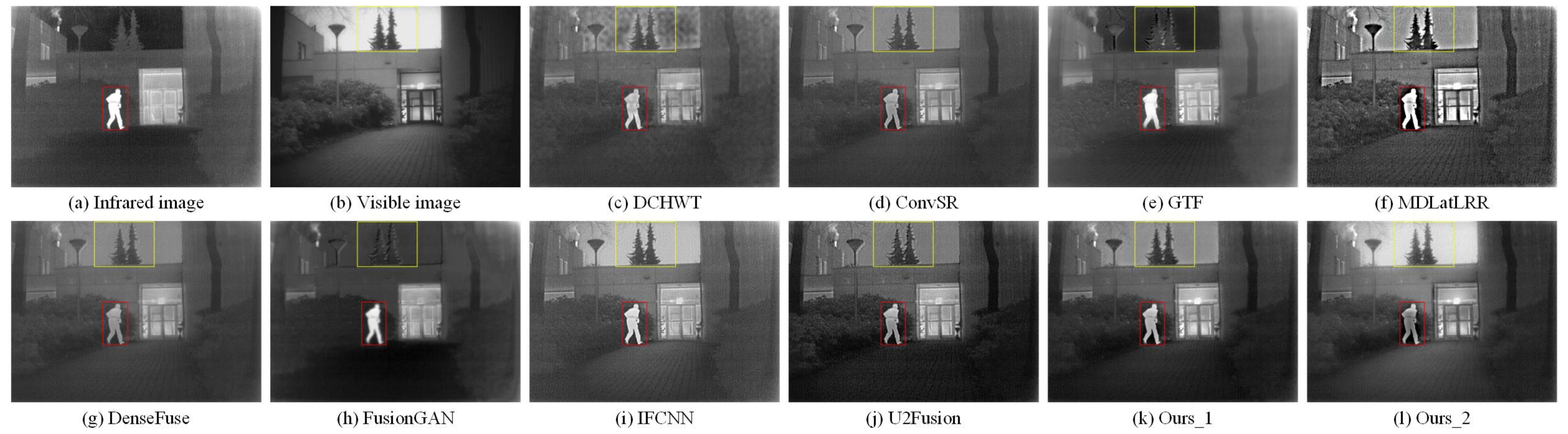

Table 1: Average quantitative results on TNO dataset (red: the best, blue: the second best).

Methods		EN	SD	MI	CC	MS-SSIM	SCD
DCHWT [5]		6.5678	64.9789	13.1355	0.4272	0.8433	1.6099
ConvSR [12]		6.2587	50.7437	12.5174	0.4922	0.9028	1.6482
GTF [13]		6.6343	67.5436	13.2687	0.3228	0.8084	1.0049
MDLatLRR [8]		6.9774	83.7741	13.9555	0.4393	0.8517	1.6332
DenseFuse [6]		6.6716	54.3575	13.3431	0.4994	0.9290	1.8350
FusionGAN [14]		6.3629	54.3575	12.7257	0.4257	0.7318	1.4569
IFCNN [25]		6.5955	66.8758	13.1909	0.4659	0.9053	1.7138
U2Fusion [23]		6.7571	64.9116	13.5142	0.5010	0.9253	1.7984
Ours	$v_{f_{ir}}$=0.3,$v_{f_{vis}}$=0.3	6.8174	72.2746	13.6348	0.4880	0.9281	1.8291
	$v_{f_{ir}}$=0.3,$v_{f_{vis}}$=0.9	7.0294	93.1529	14.0588	0.3935	0.9065	1.5680

Performance comparison on RoadScene

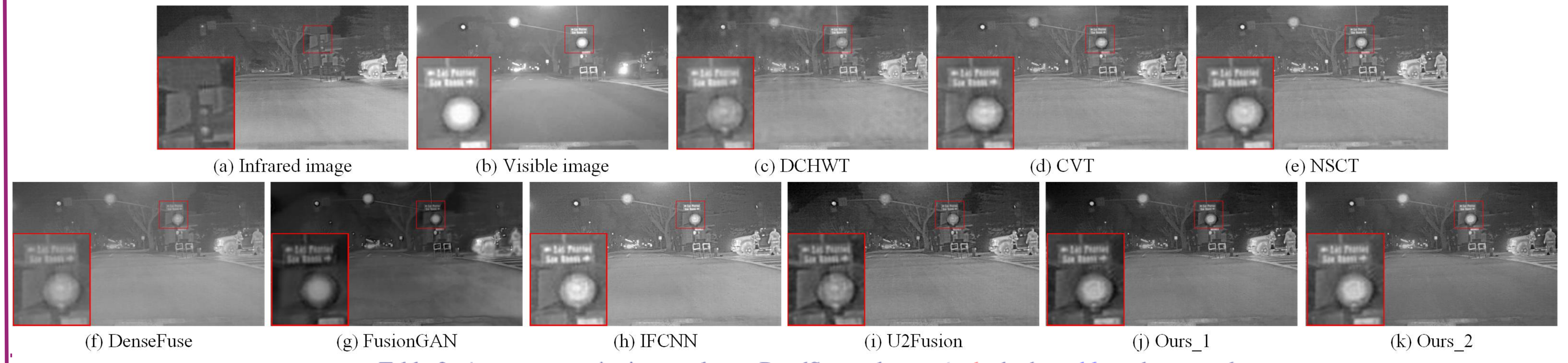

Table 2: Average quantitative results on RoadScene dataset (red: the best, blue: the second best).

Methods		EN	SD	MI	CC	MS-SSIM	SCD
DCHWT [5]		7.1710	63.0711	14.3420	0.4687	0.8283	1.4725
CVT [11]		7.0159	57.2032	14.0319	0.5255	0.8721	1.5710
NSCT [11]		6.9471	56.0915	13.8942	0.5387	0.9282	1.5965
DenseFuse [6]		6.6755	48.7363	13.3510	0.5583	0.8531	1.5607
FusionGAN [14]		7.1753	67.0645	14.3507	0.4416	0.7352	1.3753
IFCNN [25]		6.9730	56.8367	13.9461	0.5322	0.8798	1.5889
U2Fusion [23]		7.1969	68.0394	14.3938	0.5295	0.9250	1.7984
Ours	$v_{f_{ir}}$=0.3,$v_{f_{vis}}$=0.3	7.2357	70.8168	14.4714	0.5498	0.9058	1.8250
	$v_{f_{ir}}$=0.3,$v_{f_{vis}}$=0.9	7.2378	70.5262	14.4756	0.4825	0.8680	1.5497

Conclusion

- ✓ Propose a fractional gradient optimization function to obtain a pre-fused image, and the base layer of the pre-image obtained by using MDLatLRR is used as the fused base layer.
- ✓ Design a fractional gradient energy function to retain important detail features for the fusion of detail layers.
- ✓ Propose a novel fractional optimization model for infrared and visible image fusion. The experimental results show that the proposed method has good qualitative and quantitative performance on two public datasets (TNO and RoadScene).