1121

ACM multimedia
Chengdu, China OCT 20-24 2021

Feedback Network for Mutually Boosted Stereo Image Super-Resolution and Disparity Estimation

Qinyan Dai; Juncheng Li; Qiaosi Yi; Faming Fang*; Guixu Zhang
East China Normal University

Introduction

We propose a **S**tereo **S**uper-**R**esolution and **D**isparity **E**stimation **F**eedback Network (SSRDE-FNet) that jointly handles the two tasks in a unified framework and interact them for mutual boosting.

Besides leveraging the cross-view infomation explored by disparity estimation in LR space, we perform disparity estimation in the HR space to overcome the accuracy limitation of LR correspondence and better guide the stereo SR.

To achieve a more essential facilitation, we propose the **HR D**isparity **I**nformation **F**eedback (HRDIF) mechanism that feeds the aggragated HR features and the HR disparity back to previous layers for the refinement of low-level features in the SR process.

Method

SSRDE-FNet is essentially a recurrent network which contains two dual recursive sub-networks for left and right views. Each iteration involves two SR reconstruction steps and two disparity steps.

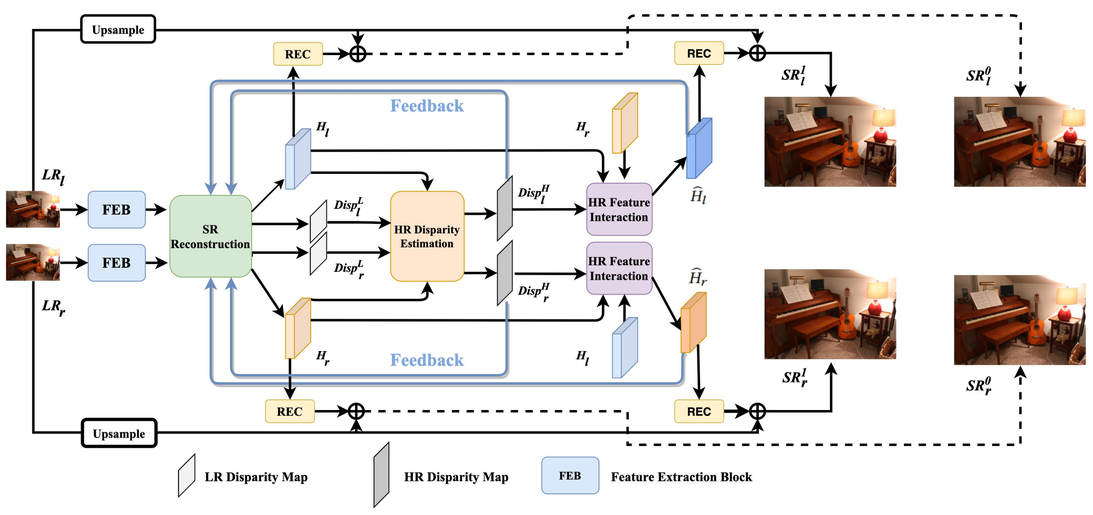

Figure 1: The framework of SSRDE-FNet.

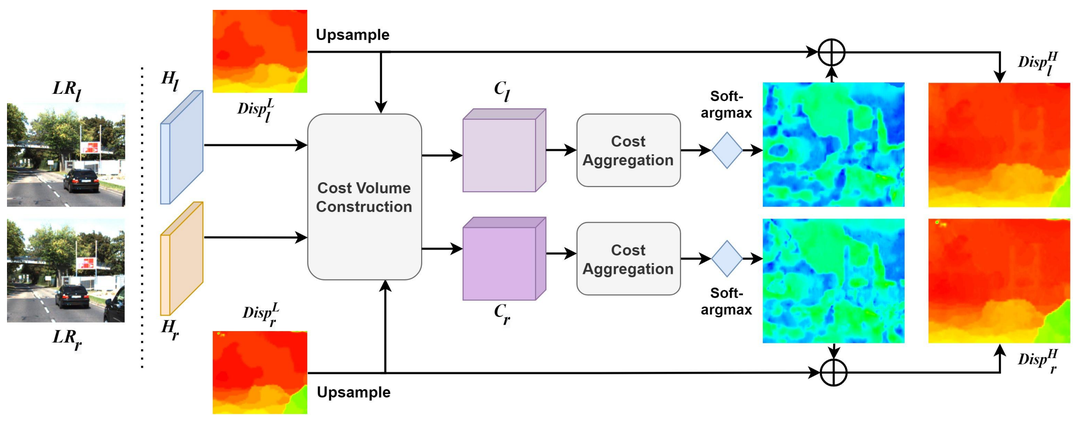

Figure 2: Illustration of HR disparity estimation module.

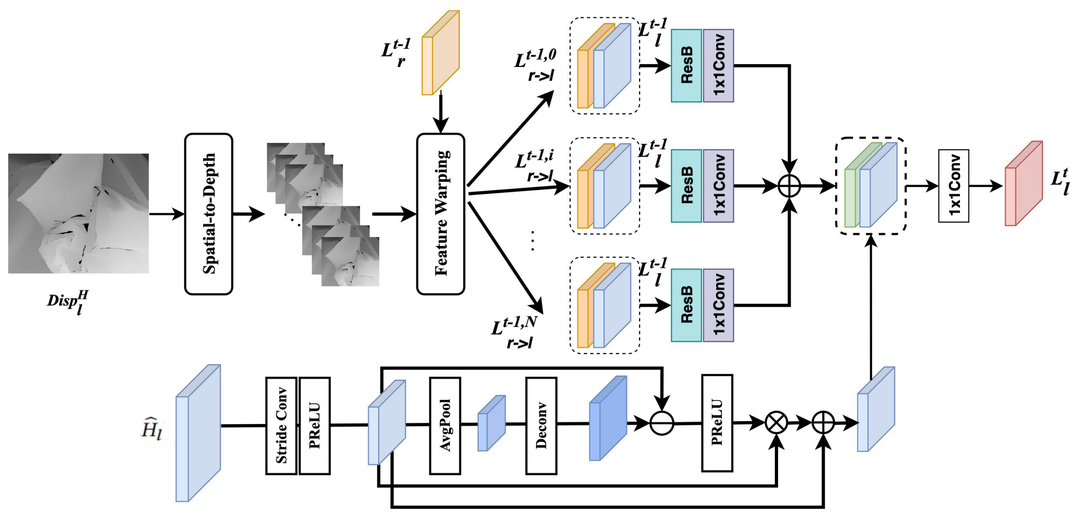

Figure 3: Illustration of our HR disparity information feedback(HRDIF) mechanism.

Stereo SR Results

Method	Scale	#P	*Left*			(*Left* + *Right*) /2			
			KITTI 2012	KITTI 2015	Middlebury	KITTI 2012	KITTI 2015	Middlebury	Flickr1024
VDSR	×2	0.66M	30.17/0.9062	28.99/0.9038	32.66/0.9101	30.30/0.9089	29.78/0.9150	32.77/0.9102	25.60/0.8534
EDSR	×2	38.6M	30.83/0.9199	29.94/0.9231	34.84/0.9489	30.96/0.9228	30.73/0.9335	34.95/0.9492	28.66/0.9087
RDN	×2	22.0M	30.81/0.9197	29.91/0.9224	34.85/0.9488	30.94/0.9227	30.70/0.9330	34.94/0.9491	28.64/0.9084
RCAN	×2	15.3M	30.88/0.9202	29.97/0.9231	34.80/0.9482	31.02/0.9232	30.77/0.9336	34.90/0.9486	28.63/0.9082
StereoSR	×2	1.08M	29.42/0.9040	28.53/0.9038	33.15/0.9343	29.51/0.9073	29.33/0.9168	33.23/0.9348	25.96/0.8599
PASSRnet	×2	1.37M	30.68/0.9159	29.81/0.9191	34.13/0.9421	30.81/0.9190	30.60/0.9300	34.23/0.9422	28.38/0.9038
iPASSR	×2	1.37M	30.97/0.9210	30.01/0.9234	34.41/0.9454	31.11/0.9240	30.81/0.9340	34.51/0.9454	28.60/0.9097
SSRDE-FNet (ours)	×2	2.10M	**31.08/0.9224**	**30.10/0.9245**	**35.02/0.9508**	**31.23/0.9254**	**30.90/0.9352**	**35.09/0.9511**	**28.85/0.9132**
VDSR	×4	0.66M	25.54/0.7662	24.68/0.7456	27.60/0.7933	25.60/0.7722	25.32/0.7703	27.69/0.7941	22.46/0.6718
EDSR	×4	38.9M	26.26/0.7954	25.38/0.7811	29.15/0.8383	26.35/0.8015	26.04/0.8039	29.23/0.8397	23.46/0.7285
RDN	×4	22.0M	26.23/0.7952	25.37/0.7813	29.15/0.8387	26.32/0.8014	26.04/0.8043	29.27/0.8404	23.47/0.7295
RCAN	×4	15.4M	26.36/0.7968	25.53/0.7836	29.20/0.8381	26.44/0.8029	26.22/0.8068	29.30/0.8397	23.48/0.7286
StereoSR	×4	1.42M	24.49/0.7502	23.67/0.7273	27.70/0.8036	24.53/0.7555	24.21/0.7511	27.64/0.8022	21.70/0.6460
PASSRnet	×4	1.42M	26.26/0.7919	25.41/0.7772	28.61/0.8232	26.34/0.7981	26.08/0.8002	28.72/0.8236	23.31/0.7195
SRRes+SAM	×4	1.73M	26.35/0.7957	25.55/0.7825	28.76/0.8287	26.44/0.8018	26.22/0.8054	28.83/0.8290	23.27/0.7233
iPASSR	×4	1.42M	26.47/0.7993	25.61/0.7850	29.07/0.8363	26.56/0.8053	26.32/0.8084	29.16/0.8367	23.44/0.7287
SSRDE-FNet (ours)	×4	2.24M	**26.61/0.8028**	**25.74/0.7884**	**29.29/0.8407**	**26.70/0.8082**	**26.45/0.8118**	**29.38/0.8411**	**23.59/0.7352**

Table 1: Quantitative results of different methods on KITTI 2012, KITTI 2015, Middlebury, and Flickr1024 datasets. #P represents the number of parameters of the networks.

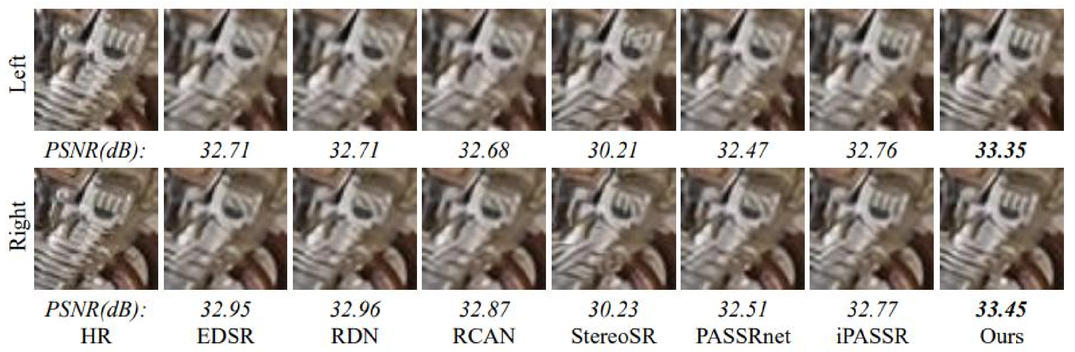

Figure 4: Qualitative results(×2) on image “motorcycle” from Middlebury dataset.

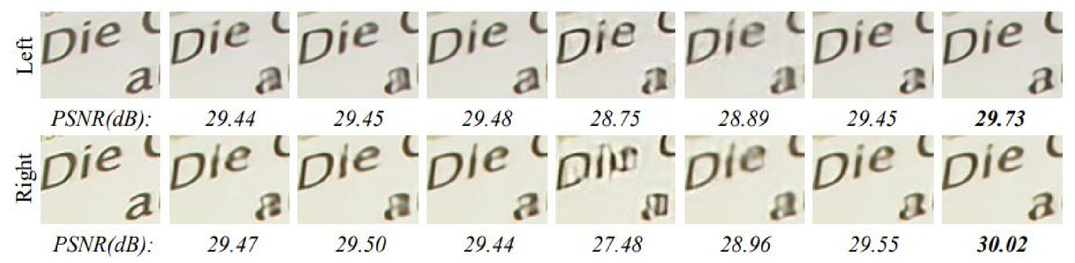

Figure 5: Qualitative results(×4) on image “testing 2” from Flickr1024 dataset.

Method	Disparity method		HRDIF		HFF	PSNR/SSIM
	Up disp	HR disp	AHFF	LRE		
baseline						29.16/0.8361
baseline + Up disp	✓					29.20/0.8370
baseline + HR disp		✓				29.27/0.8383
SSR-FNet					✓	29.27/0.8385
SSRDE-FNet w/o LRE		✓	✓			29.35/0.8407
SSRDE-FNet (Ours)		✓	✓	✓		**29.38/0.8411**

Table 2: Ablation study on different settings of SSRDE-FNet on Middlebury. The average PSNR and SSIM score of the SR left and right images are shown.

Disparity Estimation Results

		Baseline disparity	Estimated HR disparity	PASSRnet [30]	iPASSR [32]
KITTI 2012	Noc	6.72	**3.90**	11.33	7.88
	All	7.81	**5.12**	12.29	8.96
KITTI 2015	Noc	5.71	**3.52**	9.36	6.57
	All	6.38	**4.28**	9.91	7.20

Table 3: Average disparity EPE (lower is better) on KITTI 2012 and 2015 for 4x SR.

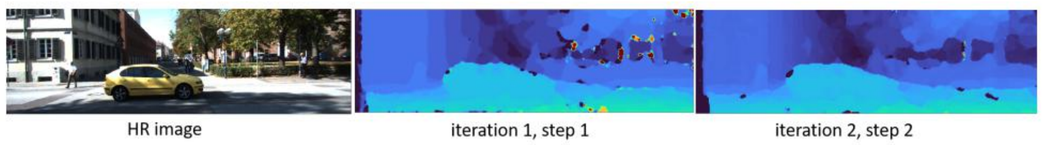

Figure 6: Visual result of the disparity map on KITTI 2015.

		Iteration 1		Iteration 2	
		Step 1	Step 2	Step 1	Step 2
KITTI 2012	Noc	7.13	6.50	4.59	3.90
	ALL	8.14	7.53	5.79	5.12
KITTI 2015	Noc	6.98	6.47	4.06	3.52
	ALL	7.60	7.11	4.81	4.28

Table 4: Disparity accuracy improvements during inference on KITTI 2012 and 2015.