I. SUPPLEMENTARY MATERIALS FOR METHODOLOGY

In the supplemental material, we describe the details in recurrent refinement module for better understanding. The recurrent refinement module consists of one RDB, one splat block and three bilateral filters.

A. RDB

To avoid the lost of information among forward propagation, we add the RDB block to reserve low frequency information. In the RDB, we use 6 convolutions with a growth rate of 16. Then we denote the output of the RDB as I^0_{res}, which is added to the newly generated I^m_{res} in each bilateral filter.

B. Splat Block

The splat block is the combined block described as "low-res coefficient prediction" in [20]. Different from [20], we modify the coefficient prediction part to make it adaptive to the task super-resolution. In experiments, we find the method achieve poor performance with two simple down-sampling operation. To protect the information for final result, we use a encoder-decoder structure in the splat block for a good trade-off between performance and efficiency. Then, we use the output of splat block F_{splat} to predict the bilateral grid Γ.

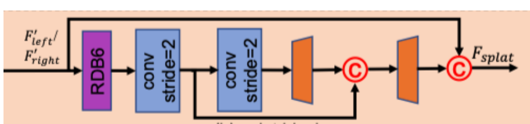

Fig. 6. The structure of the splat block.

C. Bilateral Grid

We predict the bilateral grid $\Gamma \in R^{w\times h\times d\times g}$ from the F_{splat} after a channel-fusion convolution in each bilateral filter. The image I^m is in HR space, whereas the bilateral grid is in LR space. $g = 81$ is the number of parameters in each grid, one for each coefficient of a $3\times3\times3\times3$ affine color transformation matrix K_c. We call K_c kernel in our paper, and denote the parameter in grid cell by a 81-dimensional vector $\hat{g}$. To predict the kernel for each pixel $p(x, y)$ in I_{in}, we use a 3×3 convolution to predict a guidance map Z.

$$Z(x, y) = Conv(I_{in}(x, y)) \tag{4}$$

Z is used to guide our filter to slice out an appropriate kernel K_c in the third dimension d. The slice out operation is obtained by trilinear interpolating the coefficients of Γ at locations $(x, y, Z(x, y))$:

$$K_c(x,y) = \sum_{i,j,k} \tau(s_x x - i)\tau(s_y y - j)\tau(d\cdot Z(x,y) - k))\Gamma[i,j,k] \tag{5}$$

The linear interpolation function $\tau(\cdot) = max(1 - |\cdot|, 0)$. s_x and s_y are the width and height ratios of the grids dimensions *w.r.t* dimensions of the HR image, $s_x = h/H, s_y = w/W$. We fix the spatial resolution of the grid to $h \times w$, and its depth to $d = 8$.

D. Bilateral Filter

After slicing out the kernel K_c, we apply it on the neighborhood patch $I_{patch} \in R^{3\times3\times3}$ instead of one single pixel $p(x, y)$. We add I^m_{res} and I^m to the resulting output to get I_{out}. Since our refinement module is designed in recurrent way, we use I_{out} to generate the residual image for the next bilateral filter. We put the I_{out} through a space-to-depth operation, to get a feature in LR space, and concatenate it with the initial residual feature I^0_{res} to generate the $m + 1^{th}$ residual feature I^{m+1}_{res}.