In this work, a flow-guided temporal-spatial network (FGTSN) is proposed to enhance the quality of HEVC compressed video. The framework of our proposed FGTSN method is shown in Fig. 1. Specially, we first employ a motion estimation subnet via trainable optical flow module to estimate the motion flow between current frame and its adjacent frames. Guiding by the predicted motion flow, the adjacent frames are aligned to current frame. Then, a temporal encoder is designed to discover the variations between current frame and its warped frames. Finally, the reconstruction frame is generated by training the model in a multi-supervised fashion. Our proposed method takes advantage of temporal-spatial information to enhance the Quality Of HEVC compressed video. Experimental results demonstrate the superior performance of our FGTSN method.

As shown in Fig. 2, a quality enhancement subnet (ENet) with multi-scale encoder-decoder structure is designed to explore the spatial information. The encoder consists of four convolutional layers with stride equal to 1 and four convolutional layers with stride equal to 2. Convolutions are directly followed down-sampling steps double the number of feature channels at each subsequent scale. The decoder consists of repeated application of a deconvolution that halves the number of feature channels, then concatenation with the cropped encoder feature map at corresponding resolution. Each set of decoder activations is passed through another depth-wise convolution layer to generate an intermediate prediction at its resolution.

The experimental results show that our proposed FGTSN method significantly improves the quality of HEVC compressed video at different QPs under LD configuration, much better than the state-of-the-art quality enhancement methods.

Table 1. Overall ΔPSNR (dB) results of different methods under LD