EvTexture: Event-driven Texture Enhancement for Video Super-Resolution

Dachun Kai[1], Jiayao Lu[1], Yueyi Zhang[1 2]✉, Xiaoyan Sun[1 2]

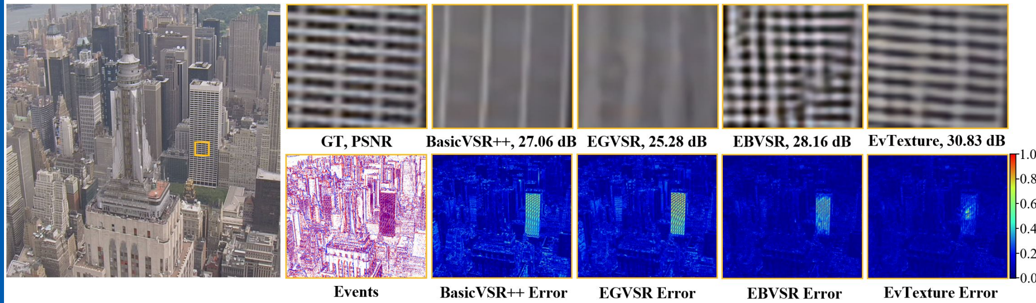
[1]University of Science and Technology of China, Hefei, China

[2]Institute of Artificial Intelligence, Hefei Comprehensive National Science Center

700+ star!

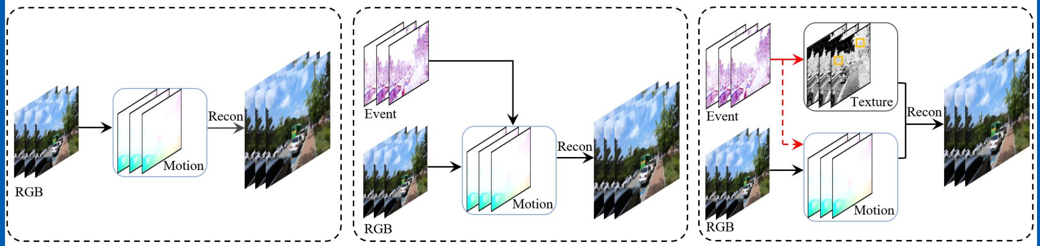
Code Page Project Page

Motivation



- Video super-resolution (VSR) aims at restoring high-resolution (HR) videos from their low-resolution (LR) counterparts.
- Due to the lack of texture details in LR videos, we notice that current VSR methods, with or without event signals, still suffer from blurry textures or jitter effects, resulting in large errors in texture regions.
- We observe that event signals are not only with high temporal resolution but also full of high-frequency dynamic details, which are desirable for texture restoration in VSR.

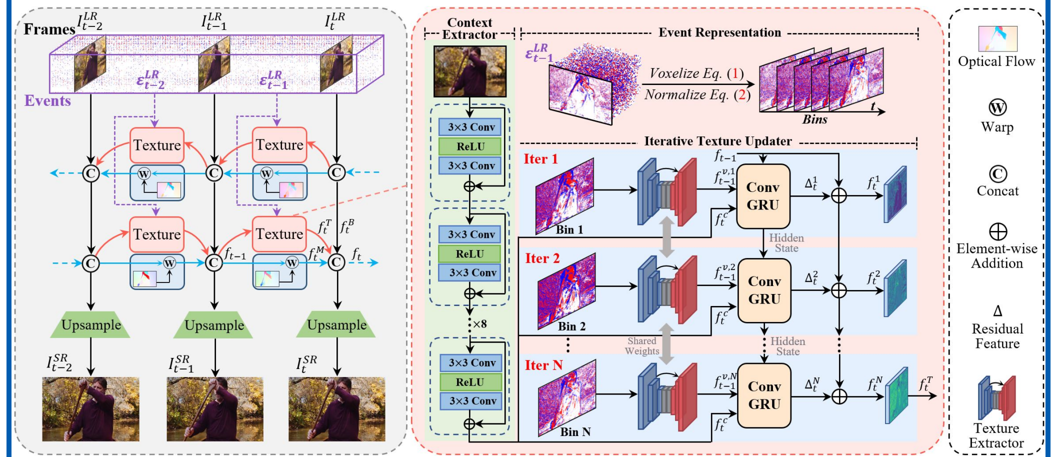
Contribution



(a) RGB-based VSR (b) Previous Event-based VSR (c) Our Method

- We propose the first event-driven scheme for texture restoration in VSR.
- We propose recovering high-frequency textural information progressively by our presented texture enhancement branch coupled with an Iterative Texture Enhancement (ITE) module.
- Our proposed texture restoration method achieves state-of-the-art performance on four VSR benchmarks and especially excels in restoring texture-rich clips.

Method



(a) The Overall Framework of **EvTexture** (b) **ITE**: Iterative Texture Enhancement

- **Two-branch Structure**: Our EvTexture adopts a bidirectional recurrent network, where features are propagated forward and backward. At each timestamp, it includes a motion branch and a parallel texture branch to explicitly enhance the restoration of texture regions.

$$F_{t\to t-1} = \mathcal{S}(I_t^{LR}, I_{t-1}^{LR}),\ f_t^M = \mathcal{W}(f_{t-1}, F_{t\to t-1}),$$

$$f_t^T = \mathcal{A}(f_{t-1}, \mathcal{E}_{t-1}^{LR}, I_t^{LR}),\ f_t = \mathcal{R}(I_t^{LR}, f_t^B, [f_t^M, f_t^T]),$$

where $\mathcal{S}$ is the optical flow model, $\mathcal{W}$ is the warping operation, $\mathcal{A}$ is the ITE module, and $\mathcal{R}$ are some residual blocks.

- **Iterative Texture Enhancement**: In the texture branch, we propose the ITE module to progressively enhance the propagation feature across multiple iterations. This module leverages high-frequency textural information from events, along with context information from the current frame.

$$f_t^c = \mathcal{C}(I_t^{LR}),\ f_{t-1}^{v,i} = \mathcal{T}\left(\widehat{\mathcal{V}}_{t-1}(i)\right),$$

$$h_t^i = \mathcal{G}\left(h_t^{i-1}, [f_t^c, f_{t-1}^{v,i}]\right),\ \Delta_t^i = \mathcal{R}(h_t^i),$$

$$f_t^i = f_t^{i-1} + \Delta_t^i,\ f_t^T = f_{t-1} + \sum_{i=1} \Delta_t^i,$$

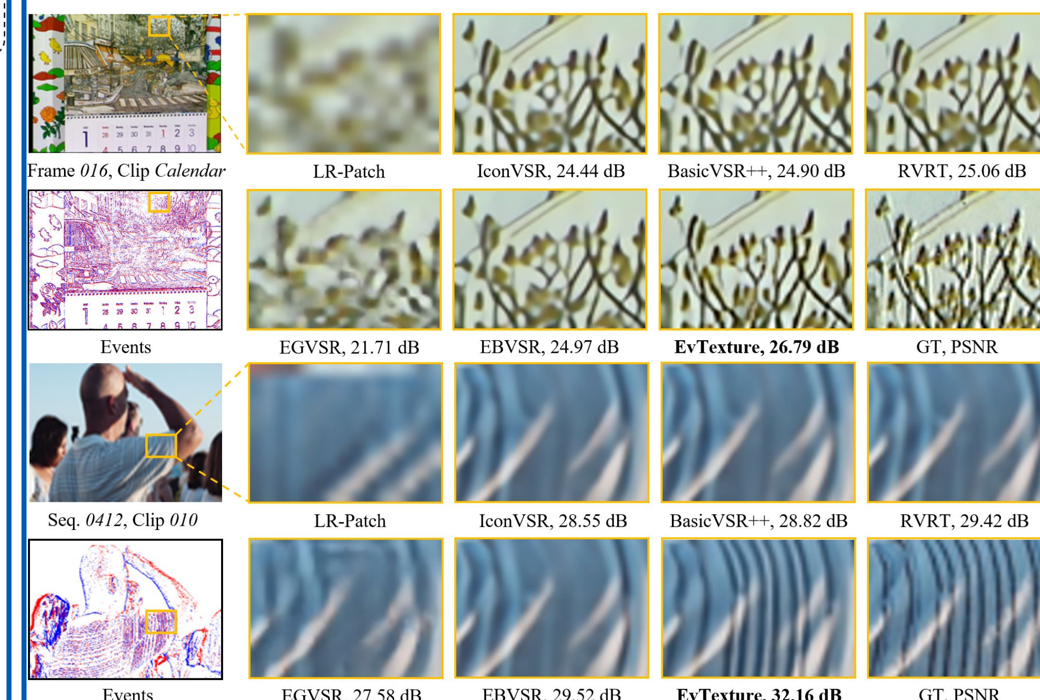
where $\mathcal{C}$ is the context extractor, $\mathcal{T}$ is the texture extractor, $\mathcal{G}$ is the GRU-based texture updater.

Experiment

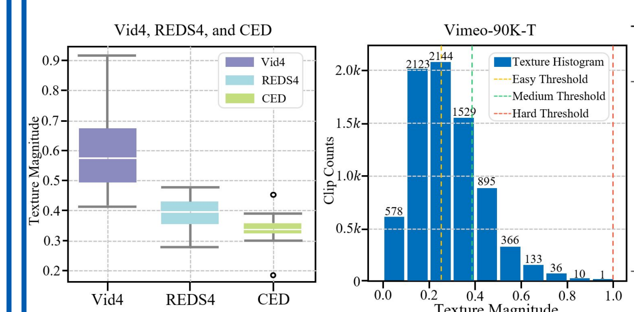
Quantitative Comparisons

Method	Input Type	Vid4 Calendar	Vid4 City	Vid4 Foliage	Vid4 Walk	Vid4 Average	REDS4	Vimeo-90K-T
EDVR (Wang et al., 2019)	I	23.98/0.8143	27.83/0.8112	26.34/0.7560	31.06/0.9153	27.30/0.8242	31.09/0.8800	37.61/0.9489
BasicVSR (Chan et al., 2021)	I	23.87/0.8094	27.66/0.8050	26.47/0.7710	30.96/0.9148	27.32/0.8265	31.42/0.8909	37.18/0.9450
IconVSR (Chan et al., 2021)	I	24.07/0.8143	27.86/0.8111	26.54/0.7705	31.08/0.9158	27.46/0.8290	31.67/0.8948	37.47/0.9476
RTVAR (Zhou et al., 2022)	I	24.65/0.8270	29.92/0.8428	26.41/0.7652	31.15/0.9167	27.90/0.8380	31.30/0.8850	37.84/0.9498
BasicVSR++ (Chan et al., 2022)	I	24.50/0.8288	28.05/0.8212	26.90/0.7868	31.71/0.9236	27.87/0.8413	32.39/0.9069	37.79/0.9500
RVRT (Liang et al., 2022)	I	24.55/0.8334	28.35/0.8363	26.98/0.7824	31.86/0.9251	27.94/0.8443	32.75/0.9113	38.15/0.9527
VRT (Liang et al., 2024)	I	24.52/0.8296	28.33/0.8308	26.78/0.7754	31.89/0.9258	27.88/0.8404	32.19/0.9006	38.20/0.9530
EGVSR (Lu et al., 2023)	I+E	21.53/0.6932	26.01/0.7068	24.33/0.6651	27.39/0.8574	24.84/0.7330	26.87/0.7790	34.62/0.9185
EBVSR (Kai et al., 2023)	I+E	25.17/0.8548	29.30/0.8846	27.31/0.8187	31.91/0.9265	28.46/0.8701	31.47/0.8919	37.56/0.9490
EvTexture	I+E	26.10/0.8756	31.24/0.9087	28.12/0.8475	32.67/0.9366	29.51/0.8909	32.79/0.9174	38.23/0.9544
EvTexture+	I+E	**26.44**/**0.8859**	**31.82**/**0.9217**	**28.21**/**0.8542**	**32.86**/**0.9381**	**29.78**/**0.8983**	**32.93**/**0.9195**	**38.32**/**0.9558**

Qualitative Comparisons



Discussion



Method	Vimeo-90K-T Easy	Vimeo-90K-T Medium	Vimeo-90K-T Hard
EDVR	41.98/0.975	35.10/0.942	30.40/0.894
BasicVSR	41.55/0.973	34.63/0.937	29.97/0.886
IconVSR	41.73/0.974	34.86/0.939	30.19/0.890
BasicVSR++	41.98/0.975	35.09/0.941	30.38/0.893
RVRT	42.43/0.976	35.69/0.946	30.73/0.902
VRT	**42.47**/0.976	35.73/0.947	30.74/0.902
EGVSR	38.75/0.959	32.15/0.905	27.90/0.836
EBVSR	41.55/0.973	35.09/0.941	30.49/0.896
EvTexture	42.45/**0.977**	**35.83**/**0.948**	**31.21**/**0.908**
# of clips	3,907	2,345	1,563
Avg. Tex. Mag.	0.16	0.32	0.49