# Real-time Controllable Denoising for Image and Video

Zhang Zhaoyang
CUHK

# Past Works

## Real-time Controllable Denoising for Image and Video

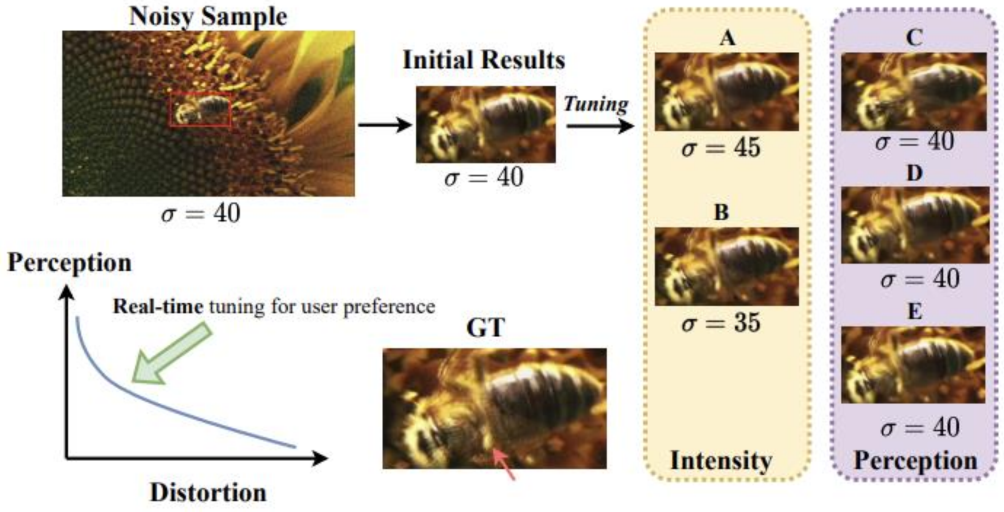



Figure 1. Real-time controllable denoising allows users further tuning the restored results to achieve Perception-Distortion trade-off. **A-B**: tuning with changing denoising intensity. **C-E**: tuning without changing denoising intensity.

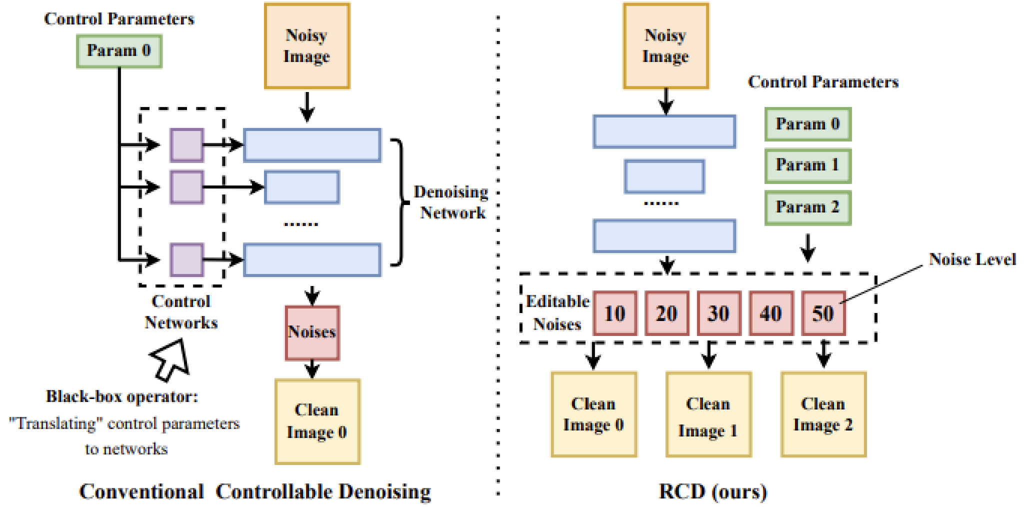



Figure 2. Comparison of pipelines between conventional controllable denoising and our RCD. RCD achieves real-time noise control by manipulating editable noises directly.

**Zhaoyang Zhang, Yitong Jiang, Wenqi Shao, Xiaogang Wang, Ping Luo, Kaimo Lin, Jinwei Gu**

# Backgrounds

## Why we need to control denoising ?

- The improvement in reconstruction accuracy (e.g., PSNR,SSIM) is not always accompanied by an improvement in visual quality, which is known as the **Perception-Distortion trade-off.**

- **Traditional denoising approaches**, we can easily adjust the denoising level by tuning related control parameters.

- **DNN-based approaches: we can only restore**
- **the degraded image or video to a fixed output with a predetermined restoration level.**

## How current methods attempt to control denoising ?

- **Interpolation-based:** use deep feature interpolation layers.
- **Condition-network-based:** import an extra condition network for denoising control.

## Drawbacks of current DNN-based Controllable denoising

- **Lack of explainability:** must use black-box operators (network layers) must to encode the relationship between the control parameters (how to modulate features) and the control operation (how the network outputs are changed)

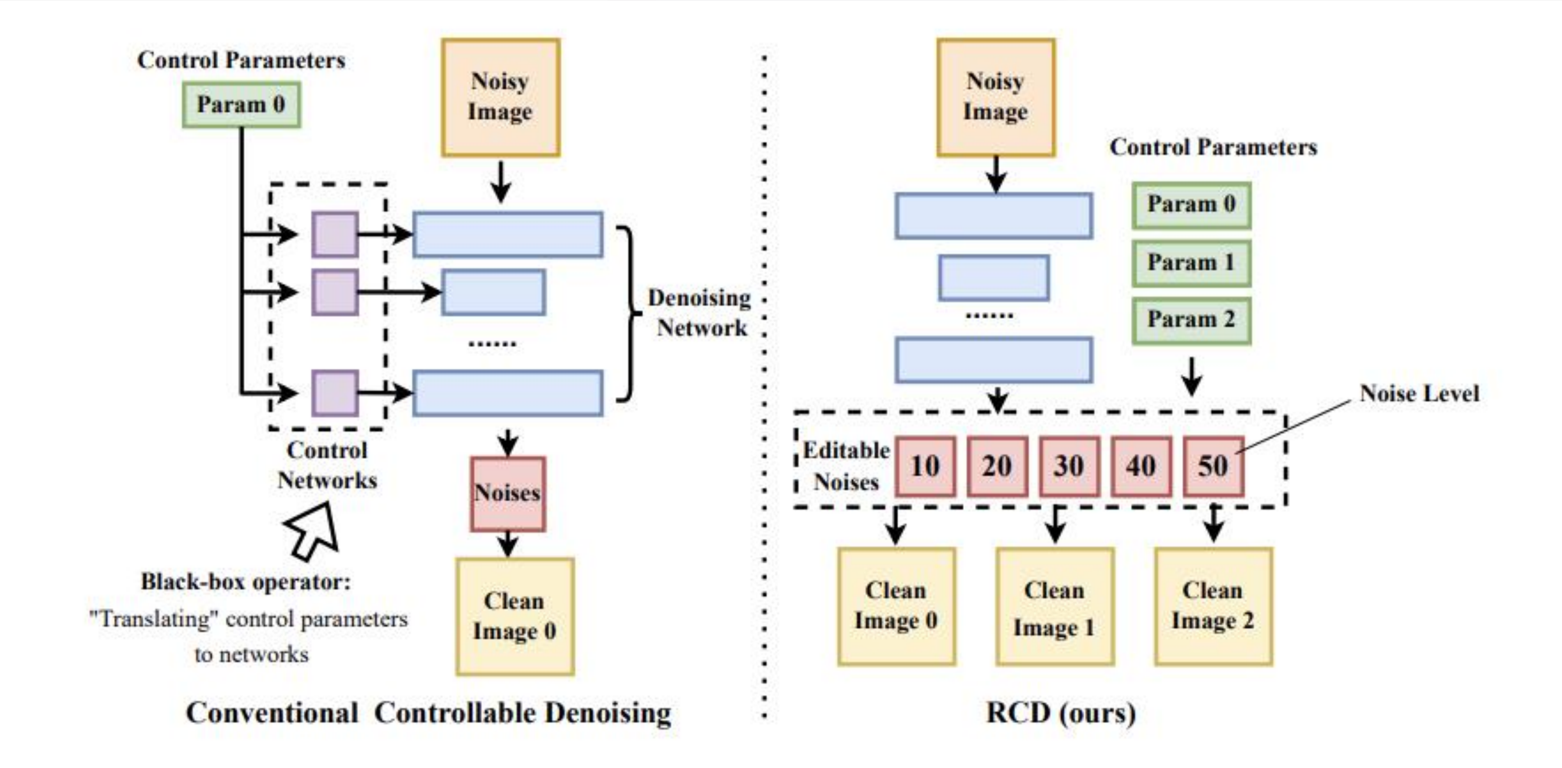
- **Lack of efficiency:** the use of control parameters as network inputs requires entire network propagation each time control parameters change

- **Not suitable for real-world samples:** current modulation methods often require an explicit degradation level during training, which is hard to obtain for real-world samples

# Our Goal:

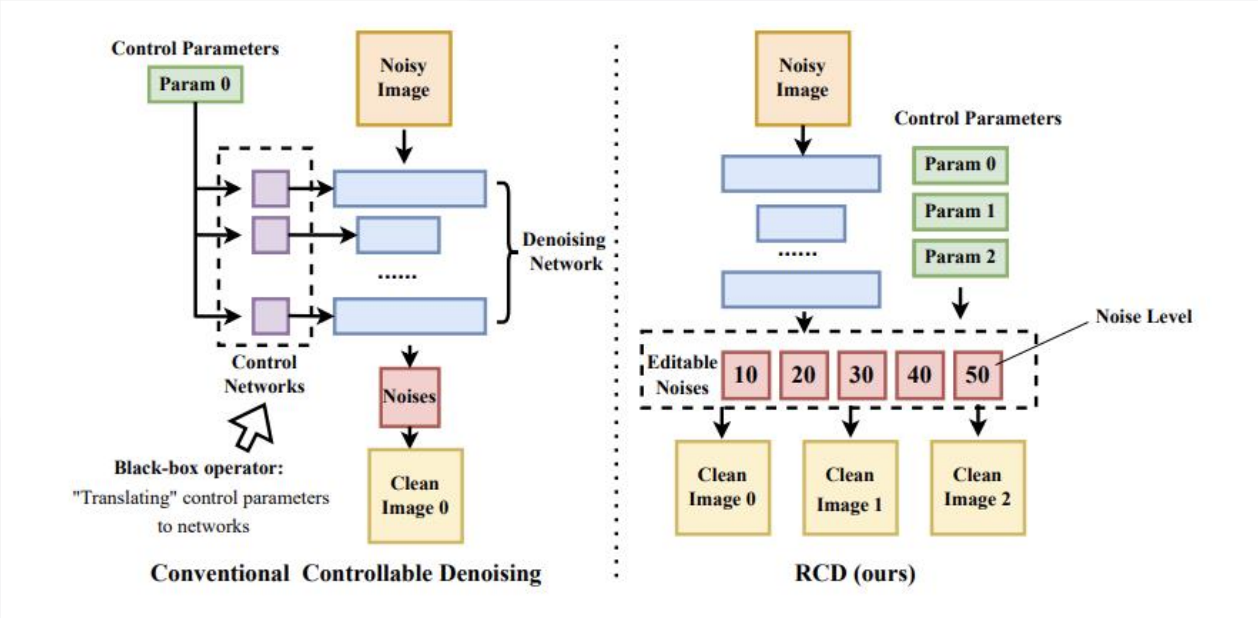
Can we achieve **real-time controllable denoising** that **abandons the auxiliary network** and requires **no network forward propagation** for changing restoration effects at test time?

# Real-time Controllable Denoising




In this work, We propose Real-time Controllable Denoising method (RCD), a lightweight pipeline for enabling rapid denoising control to achieve Perception Distortion Balance.

# Real-time Controllable Denoising




- The editing stage of RCD only involves image interpolation, users can edit their desired results on low-performance devices without the need for GPUs/DSPs.

- In contrast to traditional methods that rely on control networks, the RCD pipeline generates editable noises of varying intensities/levels, providing explicit control by external parameters and enabling network-free, real-time denoising editing.

- Real-time editing capabilities offered by RCD create new opportunities for numerous applications that were previously impossible using conventional techniques, such as online video denoising editing, even during playback (e.g., mobile phone camera video quality tuning for ISP tuning engineers), as well as deploying controllable denoising on edge devices and embedded systems.

# Real-time Controllable Denoising

## RCD’s Advantage:

- We propose RCD, a controllable denoising pipeline that firstly supports real-time denoising control (> 2000× speedup compared to conventional controllable methods) and larger control capacity (more than just intensity) without multiple training stages and auxiliary networks .
- RCD is the first method supporting controllable denoising on real-world benchmarks and video benchmarks.
- A general Noise Decorrelation technique to estimate editable noises, making RCD compatible to most NN-based denoising methods.
- We achieve comparable or better results on widely used real/synthetic image-denoising and video denoising datasets with minimal additional computational cost.

# Implementation Details

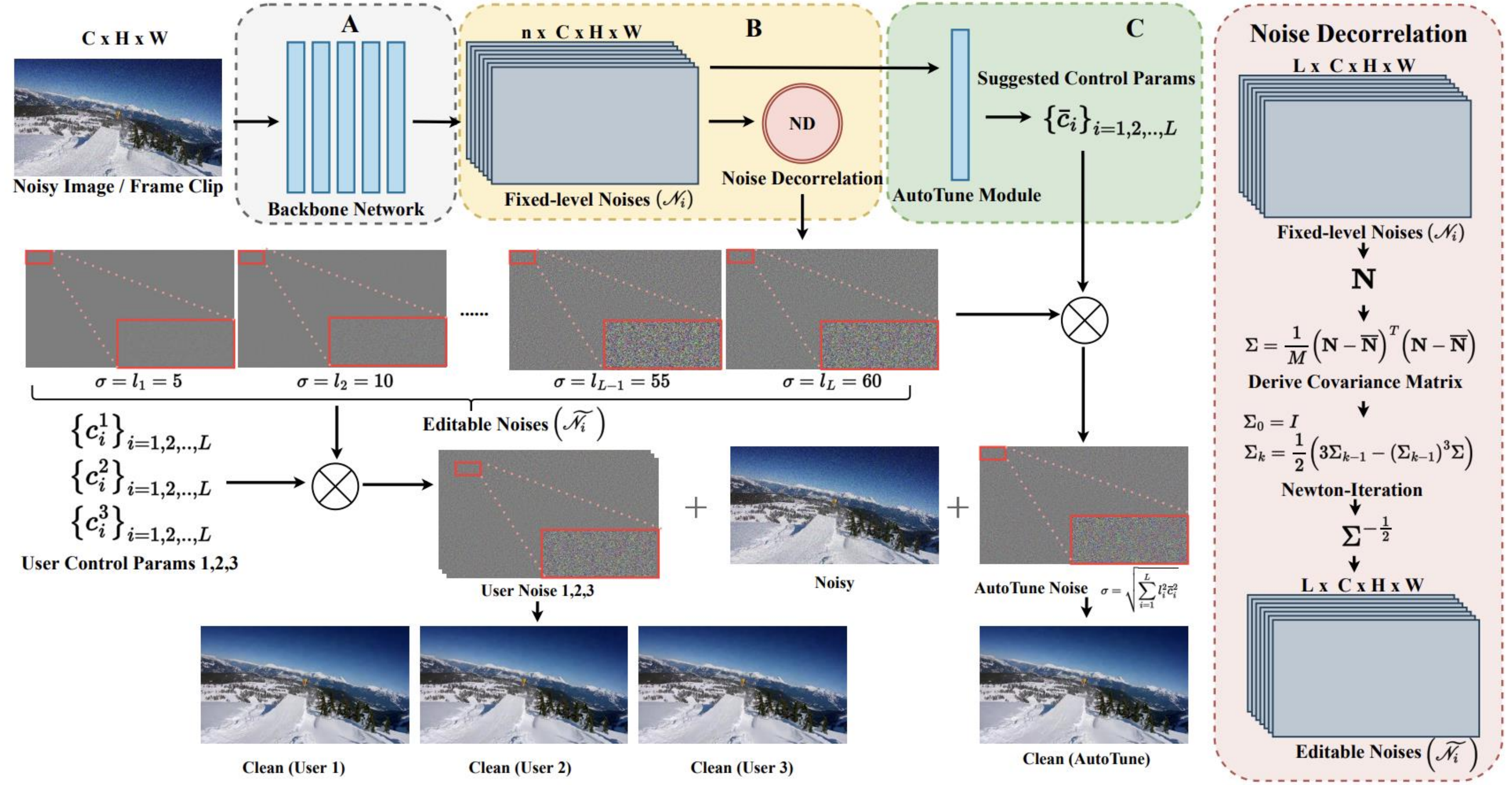



Figure 3. Pipeline overview of proposed RCD framework. **A:** Backbone network for generating multi-level noise maps. **B:** Noise Decorrelation module for editable noises. **C:** AutoTune module for providing reference control parameters for users.

# Implementation Details

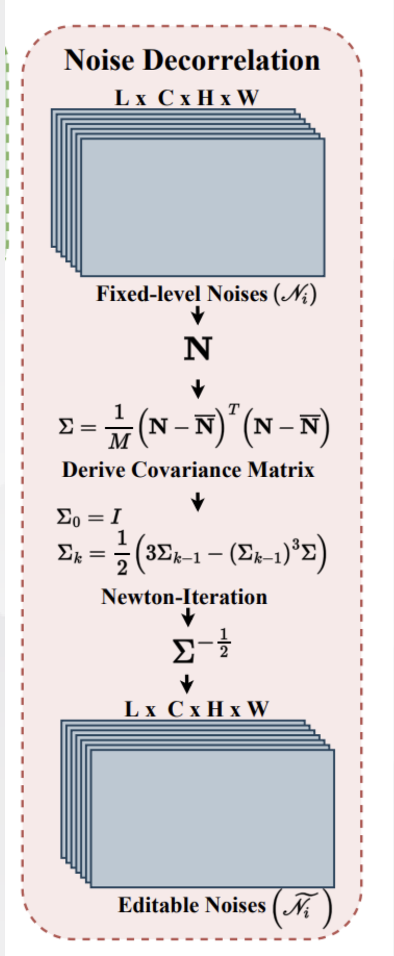
## Multi-level noise generation

$$\mathcal{N}_i = l_i \frac{\mathcal{M}_b(\mathbf{I_n})^{(i)}}{\sigma(\mathcal{M}_b(\mathbf{I_n})^{(i)})}, \forall i = 1, ..., L.$$

Generate estimated noise of multiple fixed levels.

# Implementation Details

## Noise Decorrelation




Transform fixed-level noises to editable noises by batch whitening

# Implementation Details

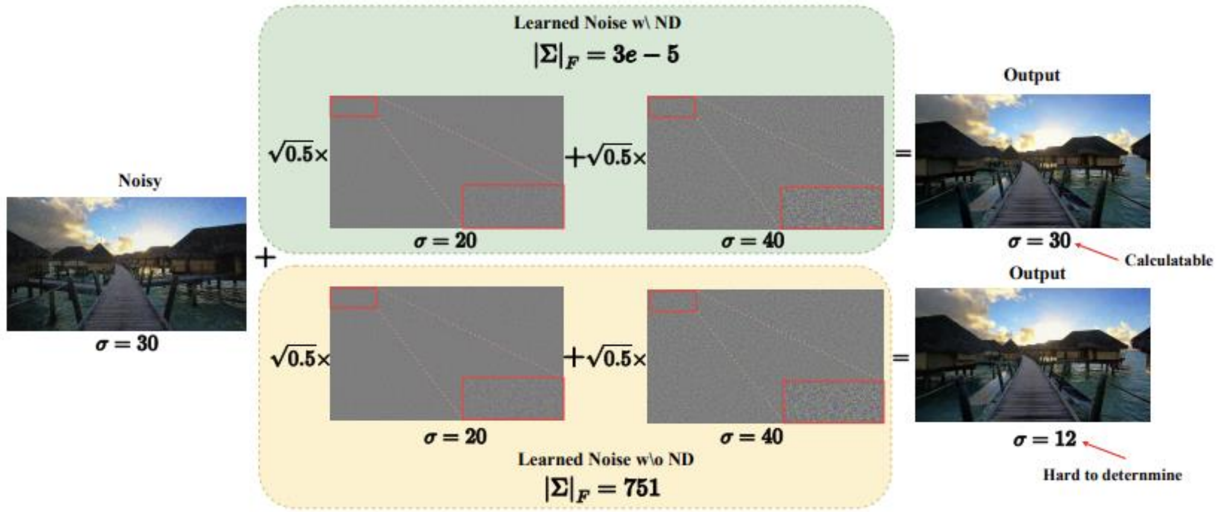



Figure 4. Demonstration of Noise Decorrelation's influence on noise editing. $|\Sigma|_F$ denotes norms of the covariance matrix for corresponding learned noises and $\sigma$ is noise intensity.

Decorrelated noises has calculatable noise levels when interpolated, allowing us to generate multiple noise maps on certain noise level by fusing multiple editable noises.

## Denoising Control

Since

$$\mathbf{Var}\left(\sum_{i=1}^{L} c_i \tilde{\mathcal{N}}_i\right) = \sum_{i=1}^{L} c_i^2 \mathbf{Var}(\tilde{\mathcal{N}}_i),$$

Users can control denoising results at certain noise level by changing control parameters {c_i}

# New Cardinality for Denoising Control.

Unlike existing methods that only modulate noise intensity, our RCD control scheme allows users to further optimize the denoising result to a given noise intensity by tuning {c_i}, as long as the weighted mean

$\sum_{i=1}^{L} c_i^2 \mathbf{Var}(\tilde{\mathcal{N}}_i) = \sum_{i=1}^{L} c_i^2 l_i^2$ remains the same.

# Implementation Details

## New Cardinality for Denoising Control.

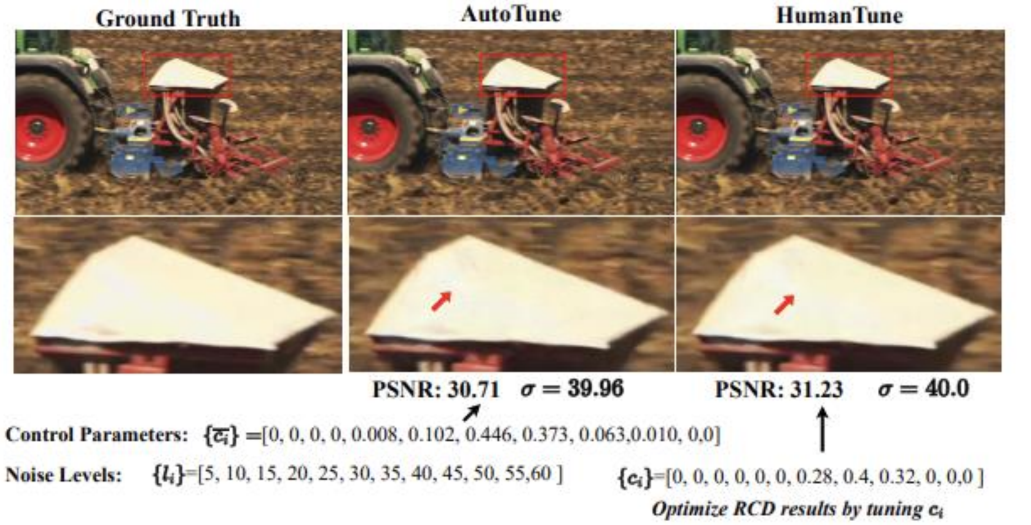



Figure 5. Example of RCD denoising results by AutoTune and HumanTune on Set8. AutoTune module provides reference control parameters, *i.e.*, $\{\bar{c}_i\}_{i=1}^{L}$), to generate the denoising result, and it can be further improved by fine-grained artificial tuning (HumanTune), *i.e.*, $\{c_i\}$, without changing the noise intensity (both $\sigma = 40$).

- Our AutoTune module can generate high-quality results using just the reference control parameters $\{\bar{c}_i\}$.
- Users can further improve the result by artificially tuning $\{c_i\}$ around $\{\bar{c}_i\}$, even at the same noise level.

# AutoTune Model

- Given the decorrelated noise maps from the Noise Decorrelation block, the AutoTune module will predict a set of model-suggested control parameters ($\{c_i\}$), to generate the default denoising result.
- Users can then use this set of parameters as a starting point to fine-tune their final desired denoising strength.
- Our AutoTune module is extremely lightweight, and is formulated as a single-layer module with temperature softmax activation

# Experimental Results

## Gaussian Single Image Denoising

Table 1. Gaussian single image denoising results (PSNR). RCD is evaluated with AutoTune results. "-": not reported

| Method | Controllable | CBSD68 $\sigma = 15$ | CBSD68 $\sigma = 25$ | CBSD68 $\sigma = 50$ | Kodak24 $\sigma = 15$ | Kodak24 $\sigma = 25$ | Kodak24 $\sigma = 50$ | McMaster $\sigma = 15$ | McMaster $\sigma = 25$ | McMaster $\sigma = 50$ | Urban100 $\sigma = 15$ | Urban100 $\sigma = 25$ | Urban100 $\sigma = 50$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| IRCNN [58] | ✗ | 33.86 | 31.16 | 27.86 | 34.69 | 32.18 | 28.93 | 34.58 | 32.18 | 28.91 | 33.78 | 31.20 | 27.70 |
| FFDNet [59] | ✗ | 33.87 | 31.21 | 27.96 | 34.63 | 32.13 | 28.98 | 34.66 | 32.35 | 29.18 | 33.83 | 31.40 | 28.05 |
| DnCNN [57] | ✗ | 33.90 | 31.24 | 27.95 | 34.60 | 32.14 | 28.95 | 33.45 | 31.52 | 28.62 | 32.98 | 30.81 | 27.59 |
| DSNet [41] | ✗ | 33.91 | 31.28 | 28.05 | 34.63 | 32.16 | 29.05 | 34.67 | 32.40 | 29.28 | | - | |
| CResMD [25] | ✓ | 33.97 | - | 28.06 | | - | | | - | | | - | |
| AdaFM-Net [24] | ✓ | 34.10 | 31.43 | 28.13 | | - | | | - | | | - | |
| NAFNet [11] | ✗ | 34.11 | 31.49 | 28.27 | 35.14 | 32.70 | 29.68 | 35.07 | 32.82 | 29.79 | 34.41 | 32.09 | 29.00 |
| **NAFNet-RCD (ours)** | ✓ | **34.13** | **31.49** | **28.26** | **35.15** | **32.72** | **29.69** | **35.11** | **32.84** | **29.81** | **34.45** | **32.12** | **29.02** |

# Experimental Results

## Gaussian Single Image Denoising

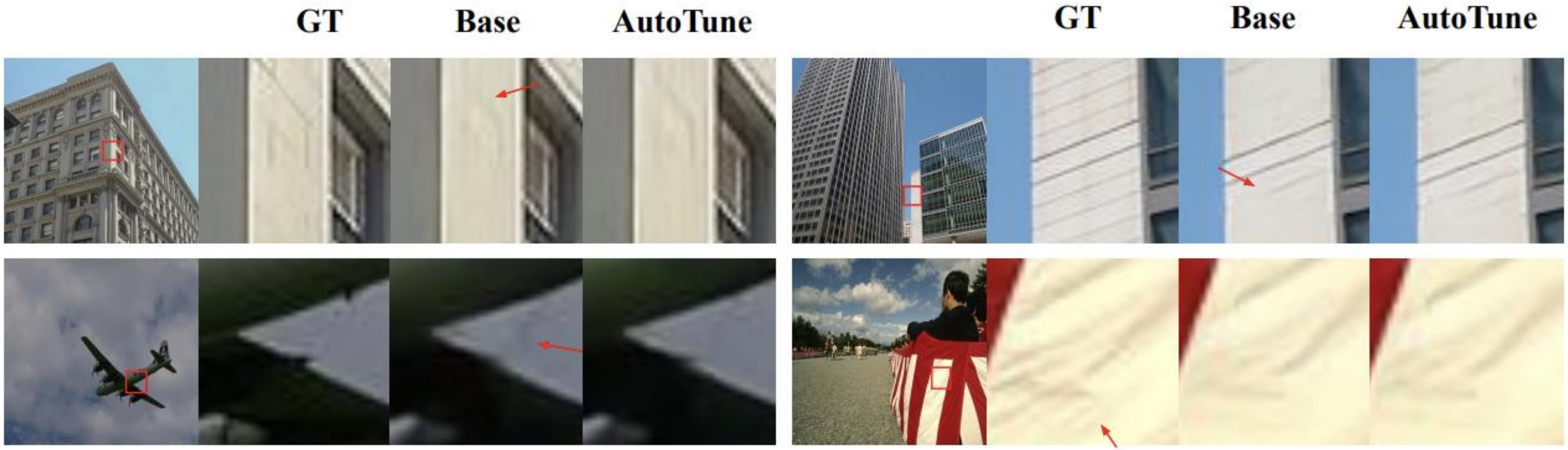



Figure 6. Visual comparison of RCD and their baseline results on $\sigma = 50$ denoising. **GT**: Ground truth. **Base**: Baseline model without RCD. **AutoTune**: RCD results by applying control parameters from AutoTune module.

# Experimental Results

Table 2. Running time comparison for RCD and other controllable methods during test time. **Full Pipeline** compares full pipeline latency for the model to infer 1000 images, and **Edit-only** compares latency for editing one image with 1000 different control parameters.

| Method | Multi-stage training | Full Pipeline | Edit-only |
|---|---|---|---|
| AdaFM-Net | required | 81.03s | 81.03s |
| CResMDNet | not required | 128.08 | 128.08s |
| **NAFNet-RCD** | not required | **64.84s** | **0.04s** |

- RCD enables Real-time noise control feature for DNN-based denoiser without performance drop, even when simply using results of AutoTune module.
- Compared to other controllable denoising methods, RCD achieve 2000x speed-up for denoising editing.

# Experimental Results

## Control Logic of RCD

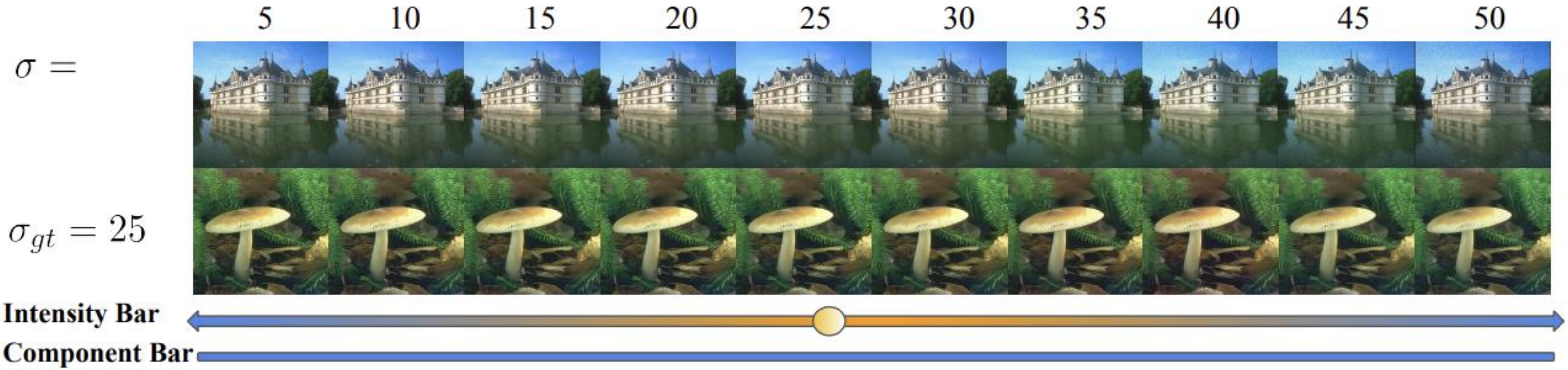



Figure 7. Illustration of RCD control logics. Users can retouch the **denoising level** by tuning the **Intensity** bar ($\sigma = \sqrt{\sum_{i=1}^{L} c_i^2 l_i^2}$ ) and setup their perceptual preference at **fixed level** by tuning **Component** bar (changing $\{c_i\}$ while keeping $\sigma$).

# Experimental Results

## Real Image Denoising and Video Denoising

Table 4. Image denoising results on SIDD. **Real noise**: results on real-world SIDD test sets. **Synthetic noise**: results on SIDD test set with additive Gaussian noise ($\sigma = 25$).

| **Method** | **Real noise** | | **Synthetic noise** | |
|---|---|---|---|---|
| | **PSNR** | **SSIM** | **PSNR** | **SSIM** |
| NAFNet-tiny | **42.19** | **0.9796** | 38.46 | 0.9551 |
| **NAFNet-RCD-tiny** | 41.86 | 0.9781 | **38.60** | **0.9558** |
| NAFNet | **43.22** | **0.9818** | 38.85 | 0.9481 |
| **NAFNet-RCD** | 42.91 | 0.9806 | **39.14** | **0.9580** |

Table 5. Video denoising results.

| Test set | $\sigma$ | 1 frame | | 5 frames | |
|---|---|---|---|---|---|
| | | FastDVD | FastDVD-RCD | FastDVD | FastDVD-RCD |
| DAVIS | 20 | 34.17 | **34.21** | **35.69** | 35.65 |
| | 30 | 32.45 | **32.69** | **34.06** | 34.04 |
| | 40 | 31.39 | **31.60** | **32.80** | 32.78 |
| | 50 | 30.26 | **30.57** | 31.83 | **31.85** |
| Set 8 | 20 | 31.99 | **32.01** | 33.43 | **33.46** |
| | 30 | 30.61 | **30.65** | 31.62 | **31.71** |
| | 40 | 29.62 | **29.83** | 30.36 | **30.42** |
| | 50 | 28.61 | **28.85** | 29.41 | **29.60** |

# Thanks !