South China University of Technology

Real-time Video Dehazing based on Spatio-temporal MRF

Bolun Cai, Xiangmin Xu, and Dacheng Tao

Contents

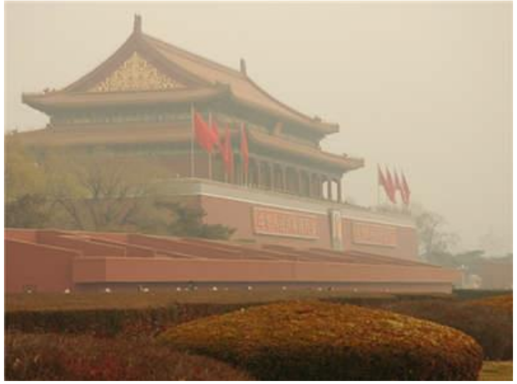
Haze & Video Dehazing

Introduction

1

What is Haze?

➢ The sky of Beijing



How can we do?

➢ Type of Haze

Haze is a traditional phenomenon where **fog**, **dust**, **smoke** and the others obscure in the clear sky.

Fog

Dust

Smoke

What is Haze?

➢ Application of Video Dehazing

Video dehazing has broader and broader application for real-time processing (e.g. **video surveillance**, **automobile recorder**, **automatic driving**).

Video Surveillance

Automobile Recorder

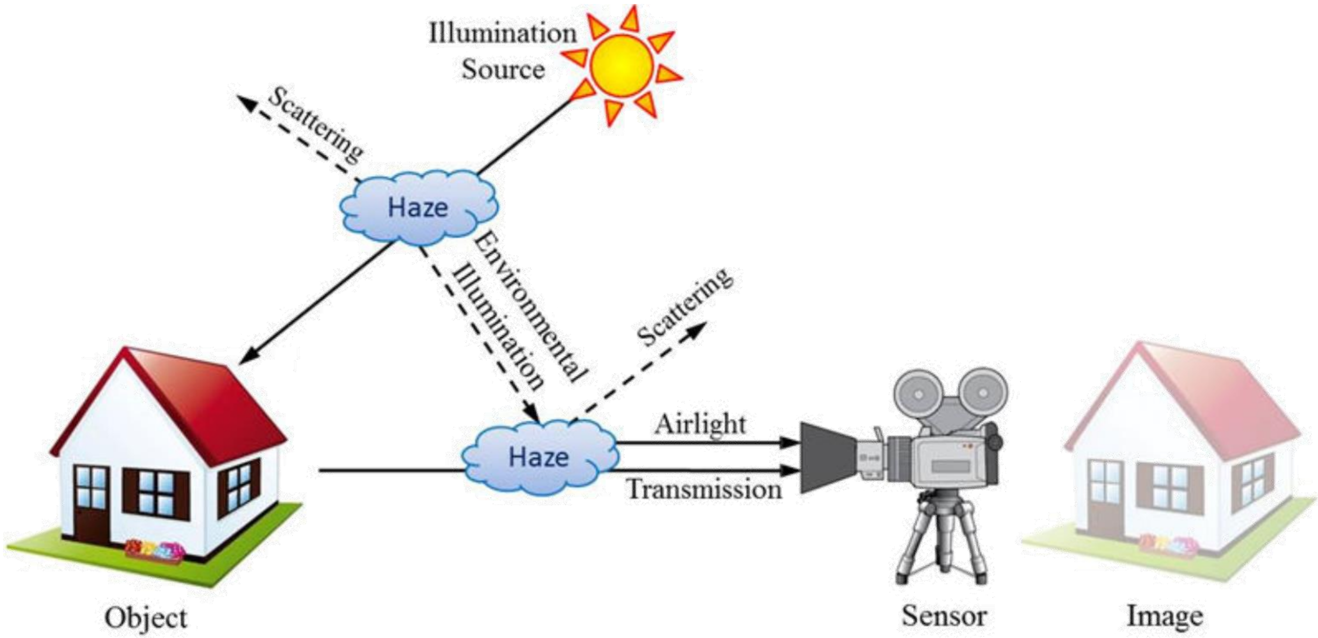
Automatic Driving

Video dehazing has
a wide range of real-time applications.

Challenge

➢ Imaging in hazy weather

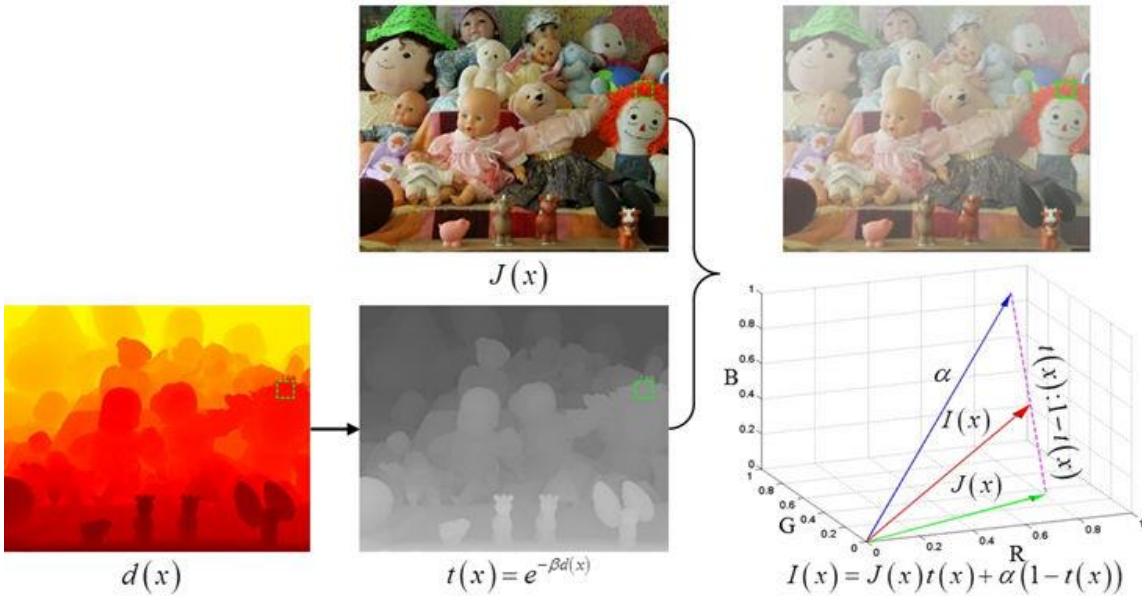
- The **transmission** caused by the reduction in **scattering** energy
- The airlight formed by the scattering of the **illumination**
- The **airlight** enhances the brightness and reduces the saturation



The Challenge is

- **Haze Transmission ∝ Unknown Depth**
- **Depth varies at different positions**

Atmospheric scattering model



$I(x)$ — hazy image

$J(x)$ — haze-free image

A — airlight

$T(x)$ — haze transmission

$D(x)$ — haze concentration

w — dark channel parameter

$$\begin{cases} I(x) = J(x)T(x) + A\left(1 - T(x)\right) \\ T(x) = 1 - wD(x) = 1 - w \min\limits_{c \in \{r,g,b\}} I^c(x) \end{cases}$$

It is the key to estimate an Accurate Haze Concentration $D(x)$

➢ Video Dehazing

- **Spatial consistency**

  —— Blocking artifacts

  The recovered video should be as natural as the original.

- **Temporal coherence**

  —— Fickering artifacts

  Frame-by-frame dehazing may break the temporal coherence.

- **Computational effciency**

  —— Real-time processing

  Video dehazing must be able to real-time process the large number of pixels in video sequences.

The challenges of video dehazing come from spatio-temporal coherence and computational effeciency.

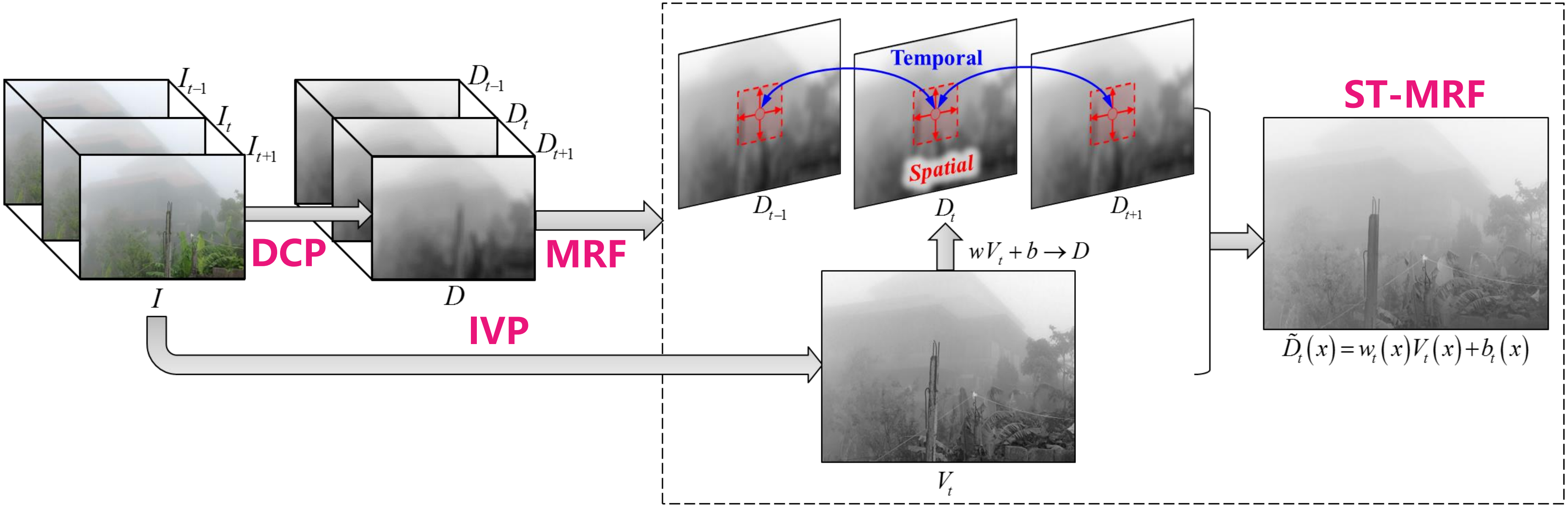
Real-time Video Dehazing

ST-MRF

2

华南理工大学
South China University of Technology

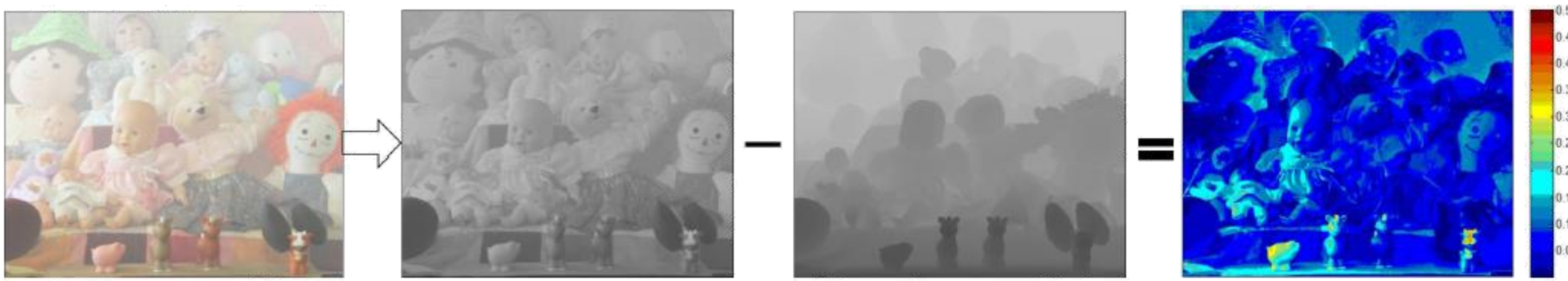
➢ Spatio-temporal MRF



- **DCP** is used to estimate the haze concentration
- **IVP** is used for the improvement of haze concentration
- **MRF** is built based on **IVP** between inner-frame (**Spatial**) and inter-frame (**Temporal**)

Intensity Value Prior (IVP)

- The intensity values of pixels in a hazy image vary sharply along with the change of the haze concentration. The residual between **intensity value** and **concentration** is close to **zero**.



(a)Synthetic image (b)Intensity value (c)Concentration (d)Residual error

The haze concentration is highly correlated with the intensity value.

ST-MRF

➢ Spatio-temporal MRF (ST-MRF)

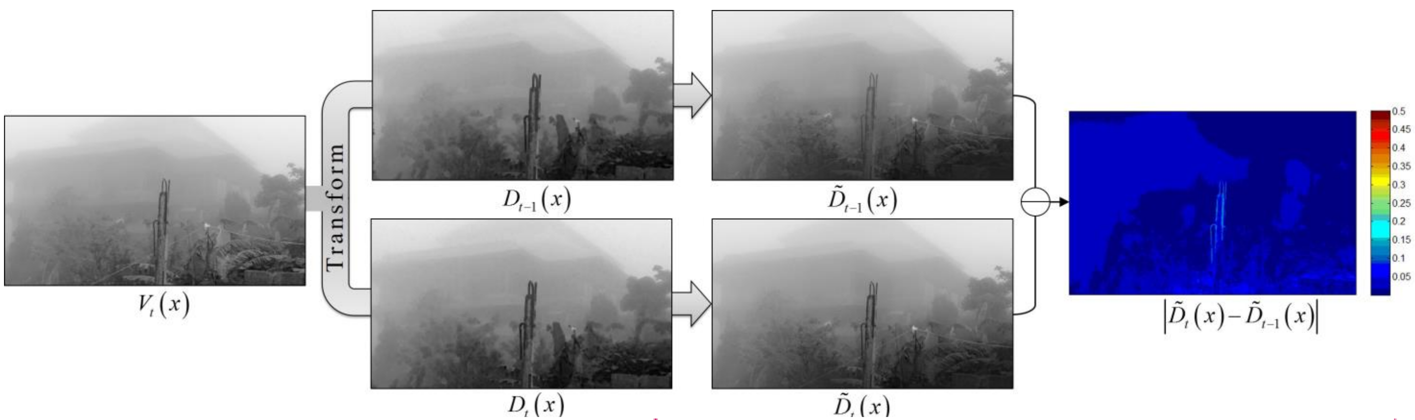
The intensity value $V(x)$ is transformed to the concentration $D(x)$, and the transformation fields $W=\{w\}$ and $B=\{b\}$ are only correlated with **the contextual information**.

- Spatial Consistency

$$P_s(w,b) \propto \prod_{y\in\Omega(x)} \exp\left(-\frac{\|w(x)V(y)+b(x)-D(y)\|_2^2}{\sigma_s^2}\right)$$

- Temporal Coherence

$$P_\tau(w_t,b_t) \propto \prod_{\tau\in[-f,+f]} \exp\left(-\frac{\|w_t(x)V_t(x)+b_t(x)-D_{t+\tau}(x)\|_2^2}{\sigma_\tau^2}\right)$$



$$P(w_t,b_t) = \prod_{\tau\in[-f,+f]} \prod_{y\in\Omega(x)} \exp\left(-\frac{\|w_t(x)V_t(y)+b_t(x)-D_{t+\tau}(y)\|_2^2}{\sigma_\tau^2}\right)$$

Maximum Likelihood Optimization

$$P(w_t, b_t) = \prod_{\tau \in [-f,+f]} \prod_{y \in \Omega(x)} \exp\left(-\frac{\|w_t(x)V_t(y) + b_t(x) - D_{t+\tau}(y)\|_2^2}{\sigma_\tau^2}\right)$$

Let $L(w_t, b_t) = \log P(w_t, b_t)$ and $\lambda_\tau = 1/\sigma_\tau^2$

$$L(w_t, b_t) = \sum_{\tau \in [-f,+f]} \sum_{y \in \Omega(x)} -\lambda_\tau \|w_t(x)V_t(y) + b_t(x) - D_{t+\tau}(y)\|_2^2$$

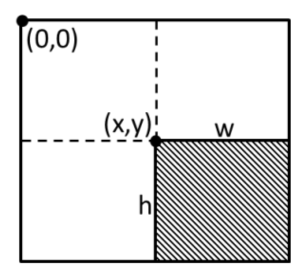
To find the W and B, we solve this linear system by max-likelihood

$$(w_t, b_t) = \arg\max P(w_t, b_t) = \arg\max L(w_t, b_t)$$

$$\begin{cases} \dfrac{\partial L(w_t, b_t)}{\partial w_t(x)} = 0 \\ \dfrac{\partial L(w_t, b_t)}{\partial b_t(x)} = 0 \end{cases} \Rightarrow \begin{cases} w_t(x) = \dfrac{\sum_\tau \lambda_\tau \left(\mathcal{U}_\Omega[V_t(x)D_{t+\tau}(x)] - \mathcal{U}_\Omega[V_t(x)]\mathcal{U}_\Omega[D_{t+\tau}(x)]\right)}{\mathcal{U}_\Omega[V_t^2(x)] - \mathcal{U}_\Omega^2[V_t(x)]} \\ b_t(x) = \sum_\tau \lambda_\tau \mathcal{U}_\Omega[D_{t+\tau}(x)] - w_t(x)\mathcal{U}_\Omega[V_t(x)] \end{cases}$$

Complexity Reduction

Mean Filter $\mathcal{U}_\Omega[F(x)] = \dfrac{1}{|\Omega|} \sum_{y \in \Omega(x)} F(y)$



Integral image $O(N)$

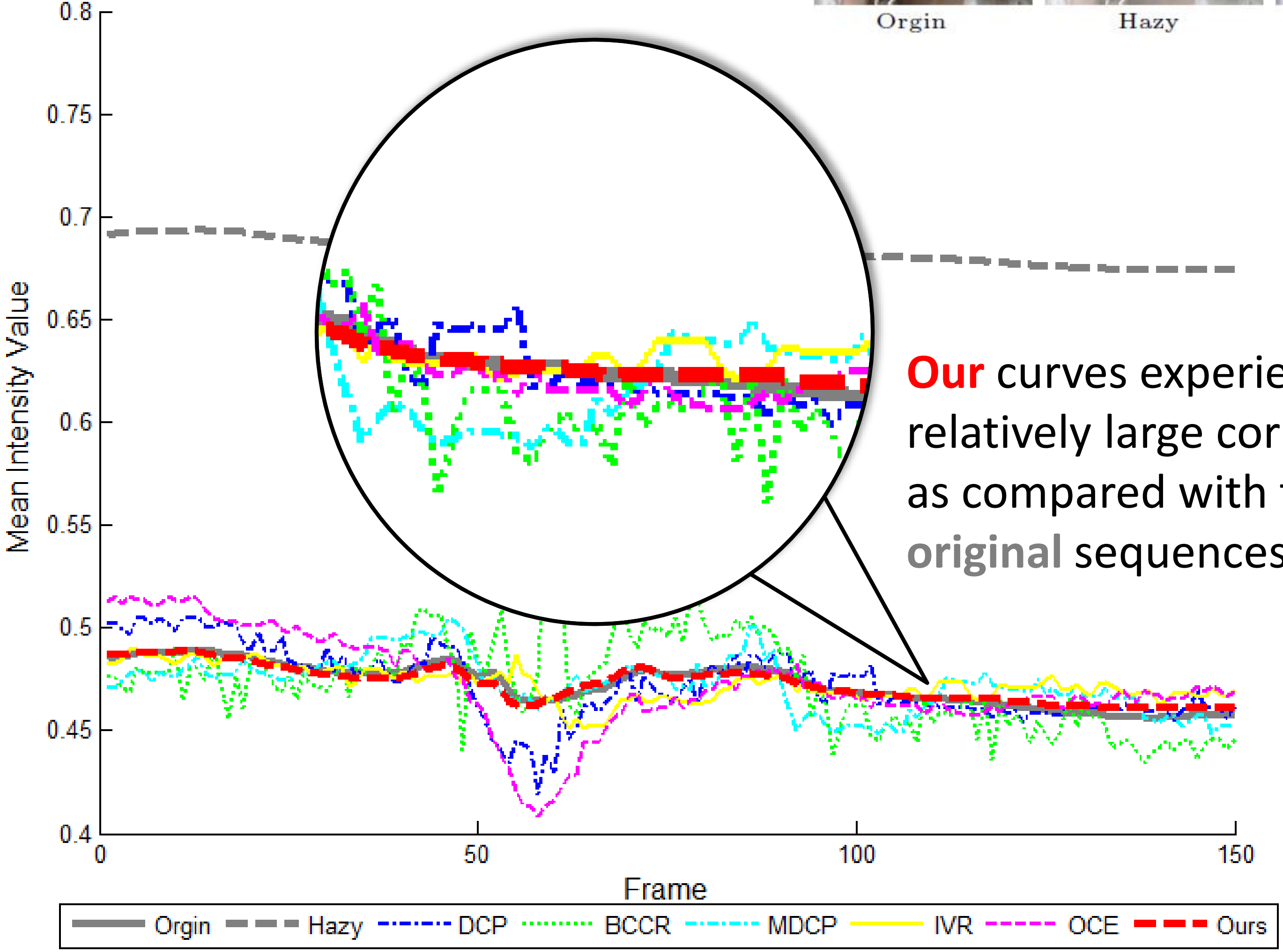
Real-time Video Dehazing

Experiments

3

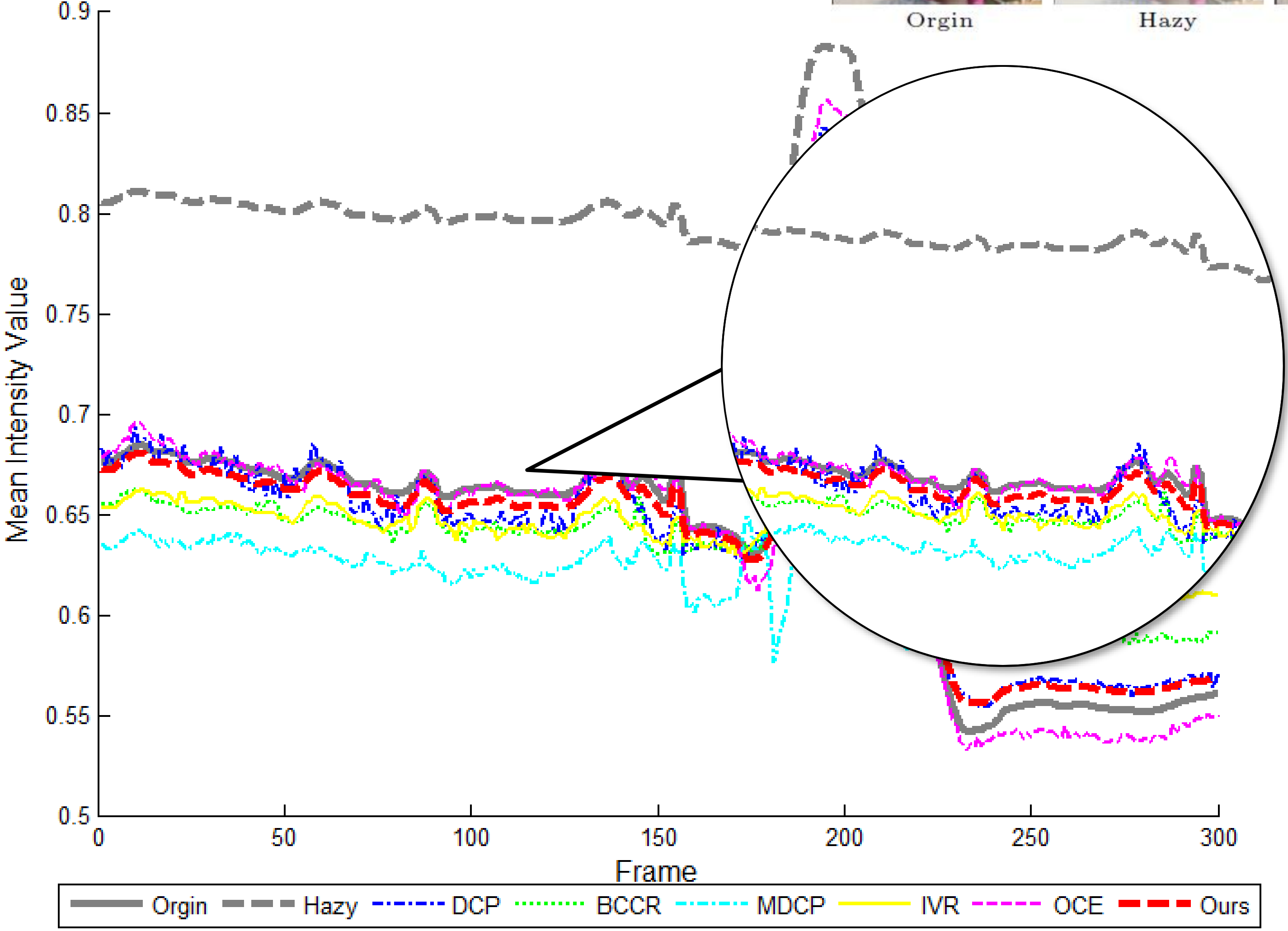
South China University of Technology

➢ Temporal Coherence Analysis



Our curves experience relatively large correlation as compared with the **original** sequences.

➢ Temporal Coherence Analysis



Temporal Coherence Analysis

We also quantify the fickering artifacts based on the correlation analysis of MIV between the dehazing result and the original video.

Table 1. The correlation coefficients of MIV between dehazing and original videos

	DCP [5]	BCCR [10]	MDCP [4]	IVR [13]	OCE [6]	Ours
Suzie	0.783	0.612	0.641	0.584	0.649	**0.976**
Foreman	0.980	0.920	0.015	0.949	0.994	**0.995**
Container	0.929	0.703	0.927	0.998	**1.000**	0.955
Hall	0.784	0.429	0.444	0.824	0.845	**0.991**
Silent	0.853	0.898	0.936	0.892	0.770	**0.990**
Avg.	0.866	0.712	0.592	0.849	0.851	**0.982**

In contrast, ST-MRF advoids the fluctuations and reduces the flickering artifacts.

Quantitative Results on Synthetic Videos

To verify the dehazing effectiveness, ST-MRF is tested on videos synthesized from **stereo videos** with a known depth map.

(a) Flower (b) Lawn (c) Road

Table 2. The dehazing results of MSE on the synthetic videos

	DCP [5]	BCCR [10]	MDCP [4]	IVR [13]	OCE [6]	Ours
Flower	0.0228	0.0240	0.0257	0.0479	0.0174	**0.0034**
Lawn	0.0198	0.0176	0.4902	**0.0141**	0.0408	0.0166
Road	0.0141	0.0191	0.0108	0.0364	0.0274	**0.0092**
Avg.	0.0189	0.0202	0.1756	0.0328	0.0285	**0.0097**

Real-time Analysis

Table 3. Comparision of the processing speeds in terms of frames per second (fps)

	DCP [5]	BCCR [10]	MDCP [4]	IVR [13]	OCE [6]	Ours
CIF (352 × 288)	1.485	1.322	7.343	1.205	97.076	**116.371**
VGA (640 × 480)	0.566	0.467	2.430	0.171	30.539	**36.609**
D1 (704 × 576)	0.414	0.358	1.830	0.102	22.930	**27.493**
XGA (1024 × 768)	0.216	0.197	0.842	0.028	12.106	**14.515**

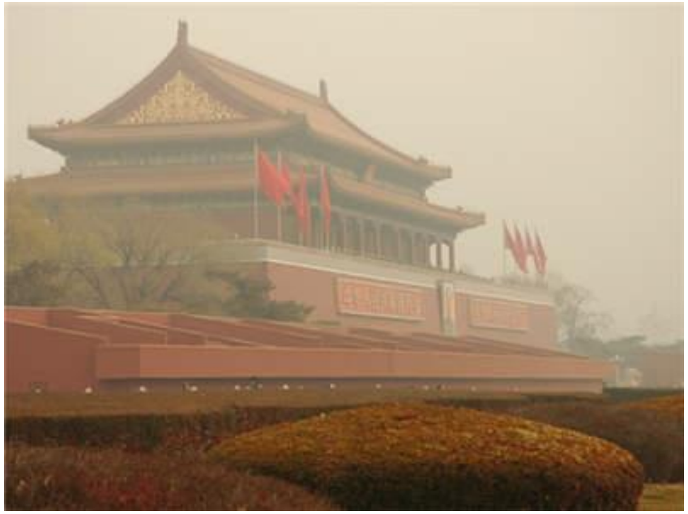
Run on a PC with Intel i7 3770 CPU (3.4GHz)

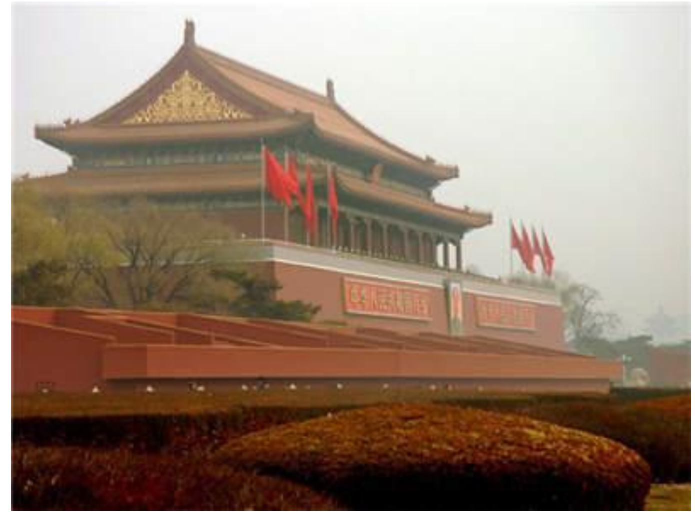
Typically, ST-MRF achieves the processing speed of about 120 fps on Common Intermediate Format (CIF, 352×288).

Qualitative Results on Real-world Videos

The codes and more comparisions can be found at http://caibolun.github.io/st-mrf/

Conclusion





Being Responsible for The Environment

保护环境，人人有责

Thank You

Q & A

caibolun@gmail.com

http://caibolun.github.io/