Turning an Urban Scene Video into a Cinemagraph

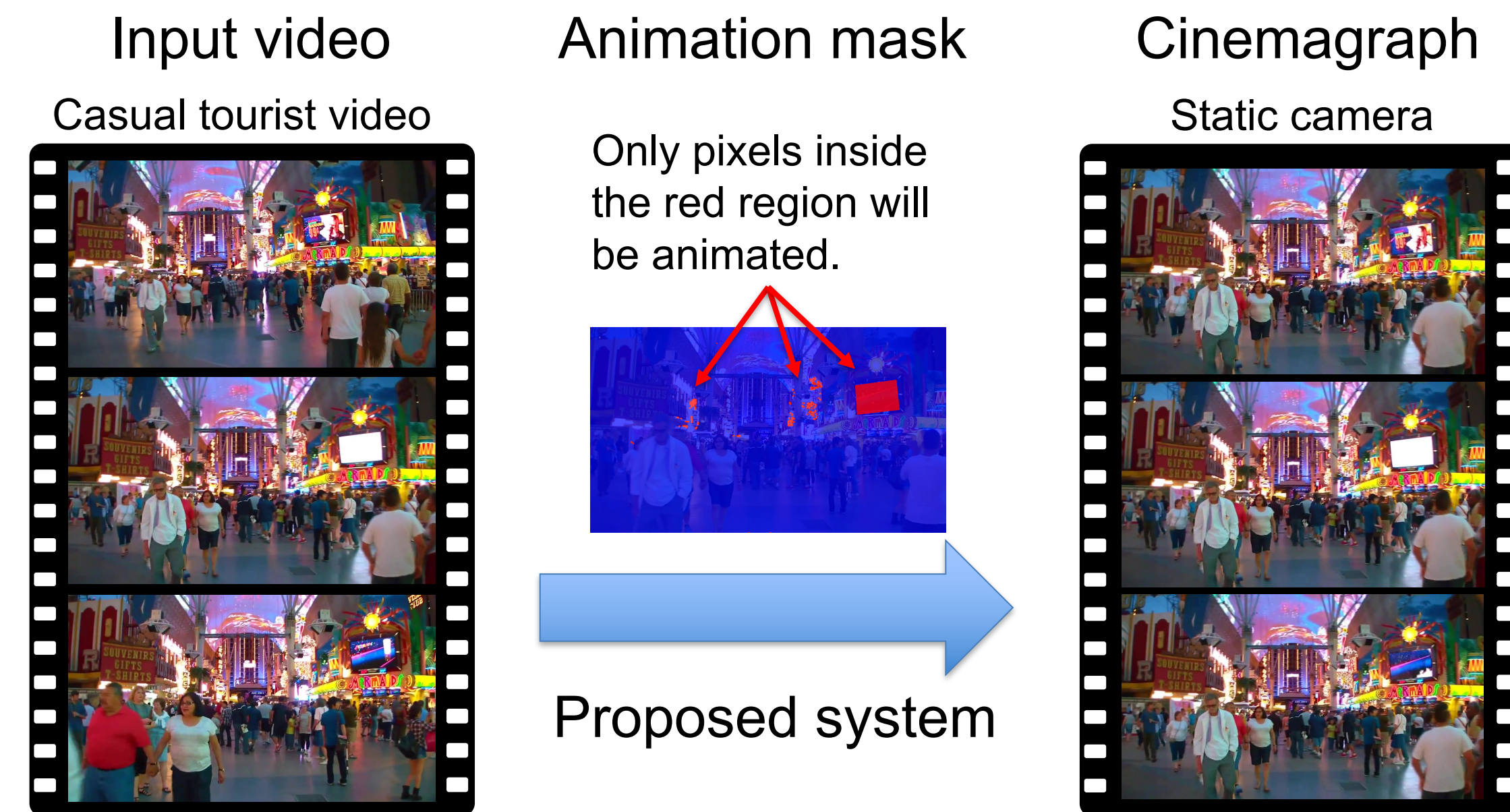Hang Yan[1]   Yebin Liu[2]   Yasutaka Furukawa[1]

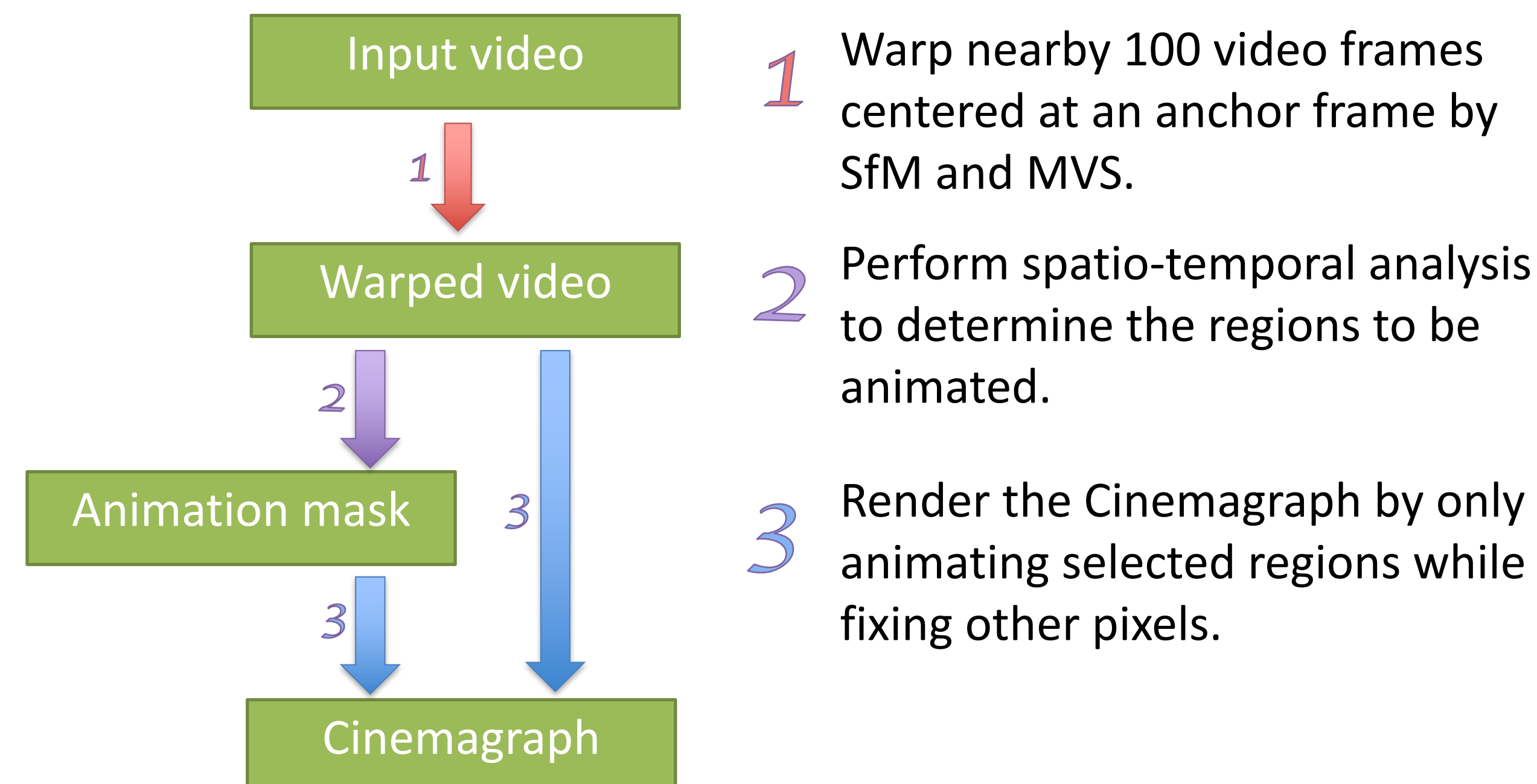[1]Washington University in St. Louis   [2]Tsinghua University

Introduction

➤ Cinemagraph: still photos with endless repeated animations in selective regions.

➤ Visualize subtle dynamics with minimum amount of data.

Input video
Casual tourist video

Animation mask
Only pixels inside the red region will be animated.

Cinemagraph
Static camera

Proposed system

System Overview

Input video

1

Warped video

2

Animation mask

3

3

Cinemagraph

1 Warp nearby 100 video frames centered at an anchor frame by SfM and MVS.

2 Perform spatio-temporal analysis to determine the regions to be animated.

3 Render the Cinemagraph by only animating selected regions while fixing other pixels.
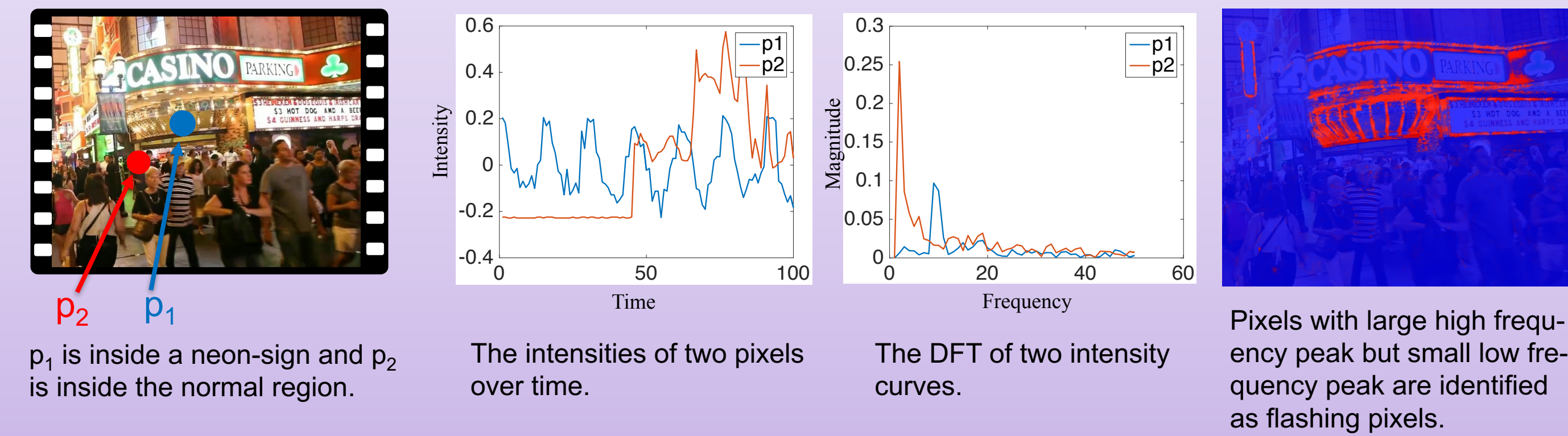
Create Animation Masks by Temporal Analysis

➤ Two types of contents in typical urban scenes: **flashing lights on neon-signs** and **advertisements on outdoor billboards.**
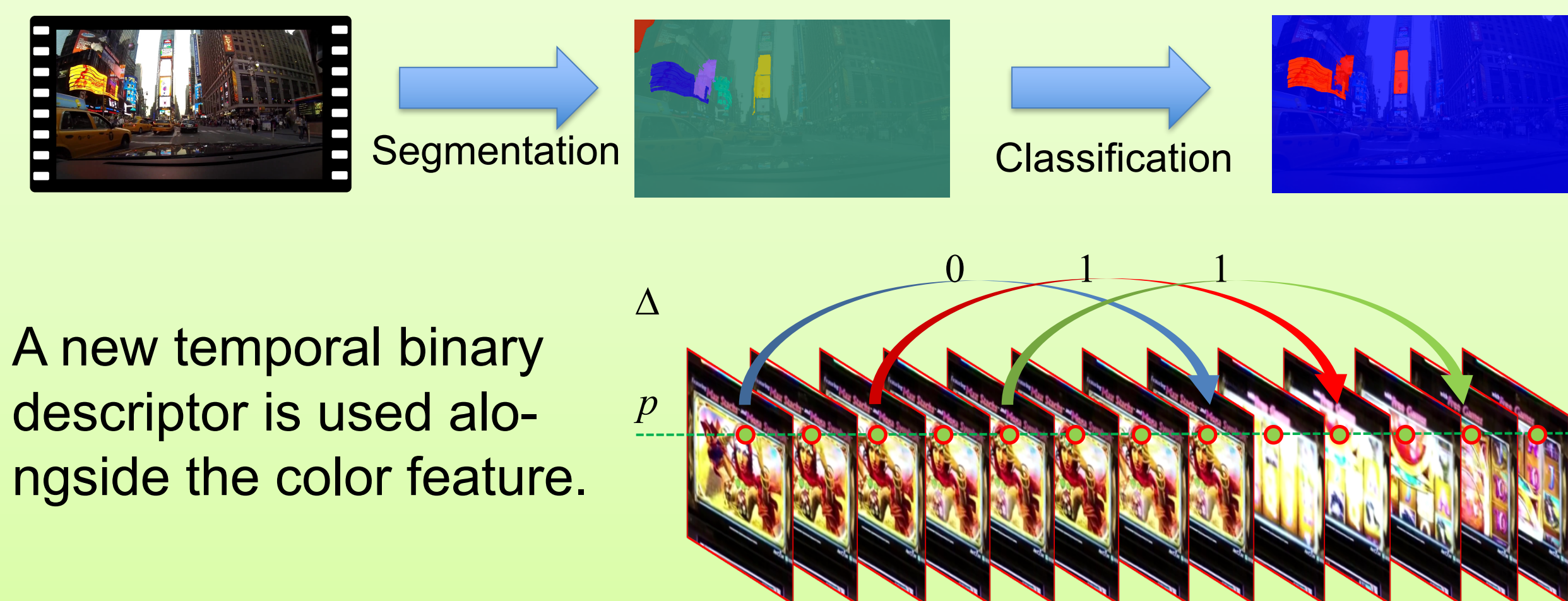
Identifying flashing pixels by frequency domain analysis

➤ Apply Discrete Fourier Transformation (DFT) to pixel intensities over time and analyze its frequency domain properties.

p_2 p_1

p_1 is inside a neon-sign and p_2 is inside the normal region.

The intensities of two pixels over time.

The DFT of two intensity curves.

Pixels with large high frequency peak but small low frequency peak are identified as flashing pixels.
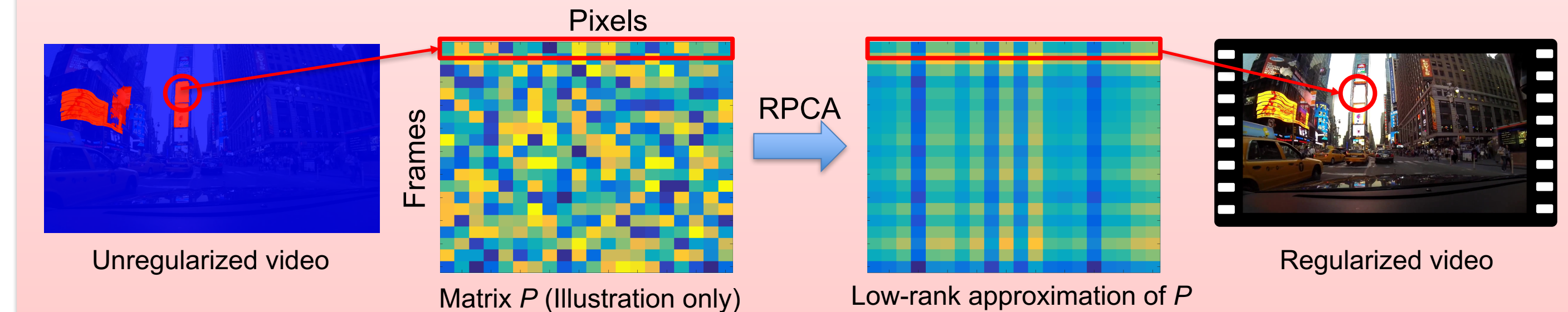
Identifying billboards by the novel spatio-temporal descriptor

➤ Segment the warped video into spatially and temporally coherent regions, then train a classifier to recognize billboard regions.

Segmentation   Classification

➤ A new temporal binary descriptor is used alongside the color feature.

Δ 0   1   1
p

Cinemagraph Rendering

➤ Cinemagraph is rendered by only playing the content inside masked regions while fixing other pixels to the anchor frame. Post-processing steps are applied to improve the visual quality.

Temporal regularization with RPCA

Unregularized video

Pixels

Frames

Matrix P (Illustration only)

RPCA

Low-rank approximation of P

Regularized video

➤ Arrange pixels of each masked region from one frame as a row and stack them across frames as matrix P.

➤ Compute a low-rank approximation of the matrix as the output pixel values.

Unregularized   Bennett et al.   Martin-Brualla et al.   Ours

➤ Out-performs previous regularization methods, which produce inconsistent results across pixels under fast motion.

Experimental results and more Information

➤ For experimental results, demos and more infomation, please visit our project page by scanning the QR code on the right.

Acknowledgement