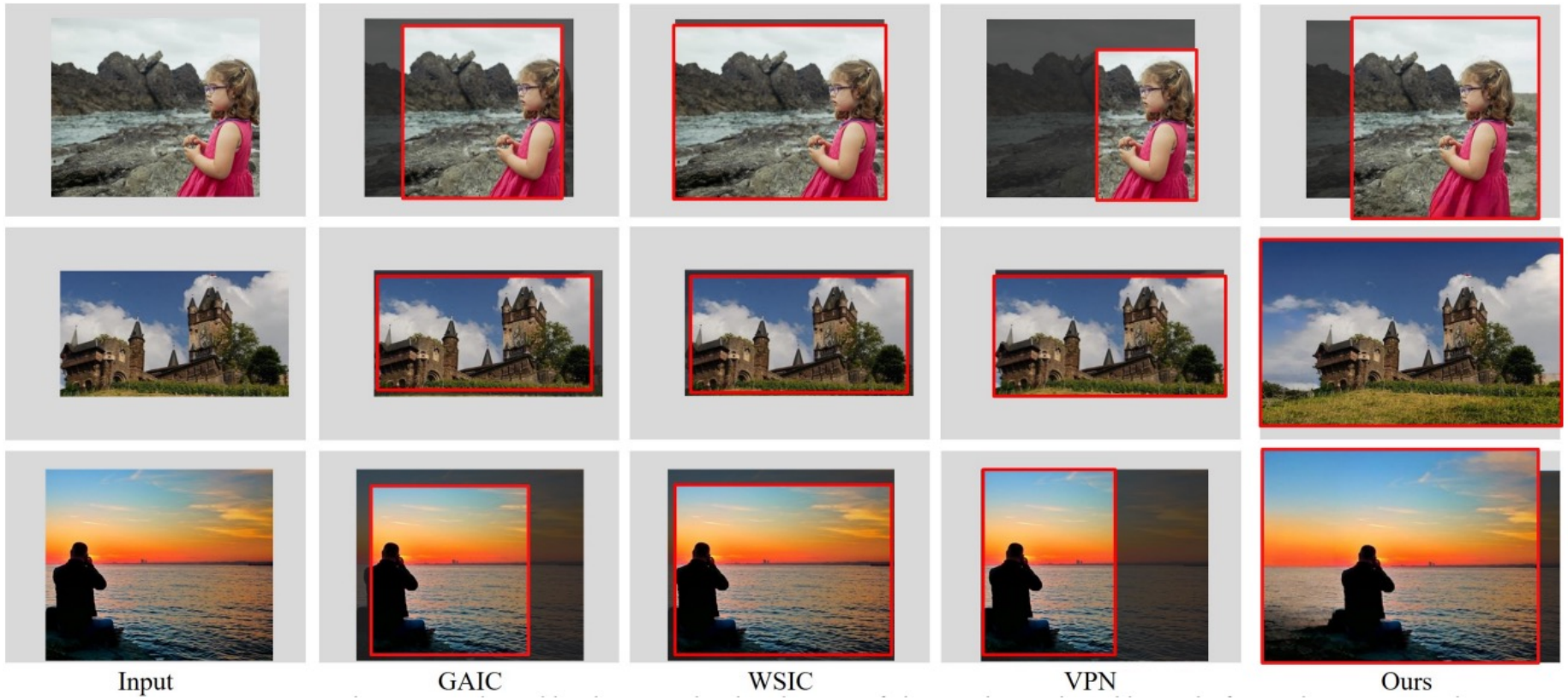
Aesthetic-guided Outward Image Cropping

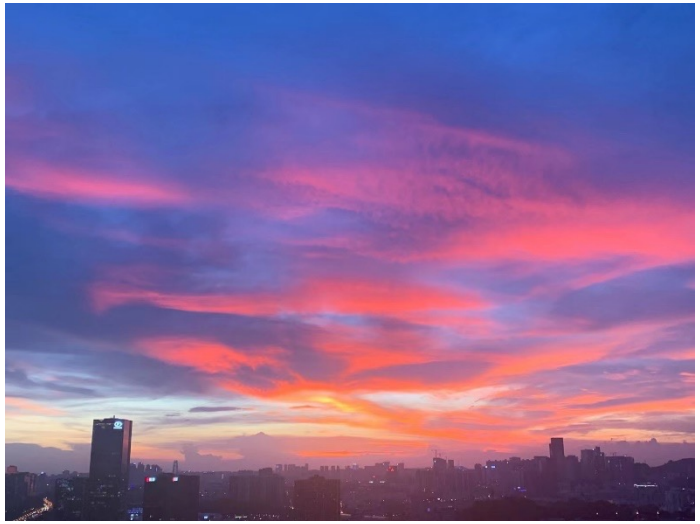
Lei Zhong*, Feng-Heng Li*, Hao-Zhi Huang, Yong Zhang, Shao-Ping Lu#, and Jue Wang.

SIGGRAPH ASIA, **2021.**



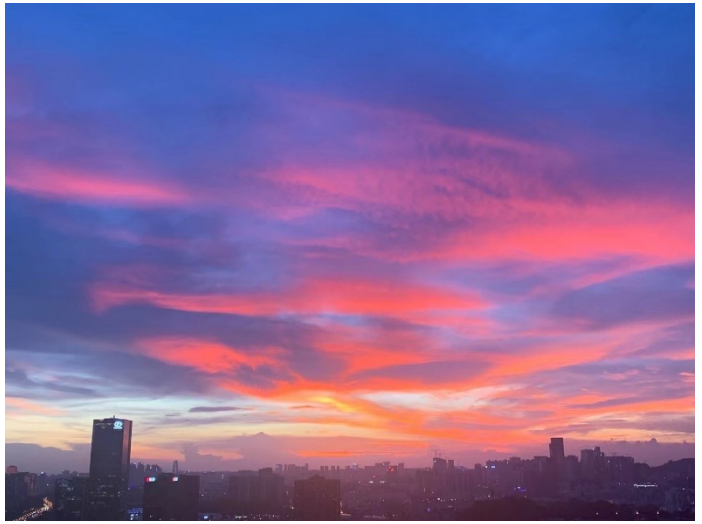
Motivation

We capture photos in memorable moments.



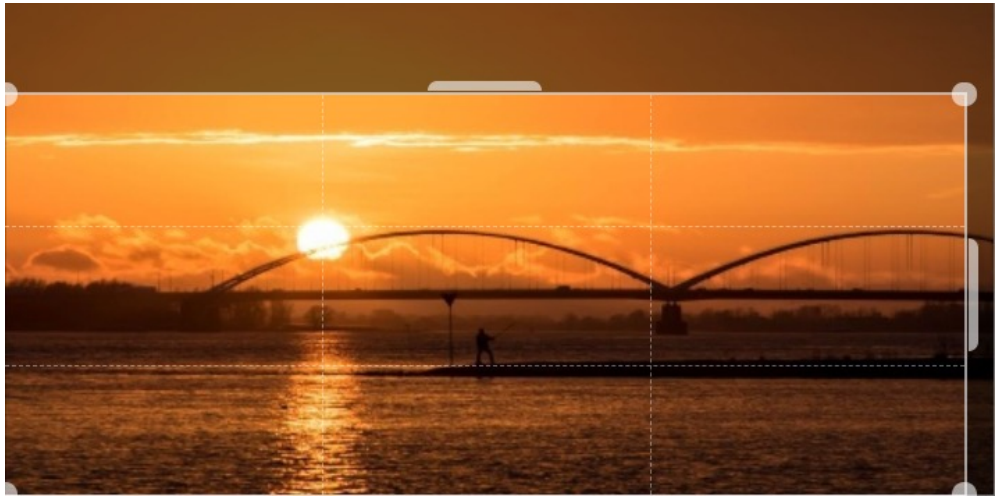
Motivation

Image composition is a crucial element that significantly impacts the aesthetic of an image.



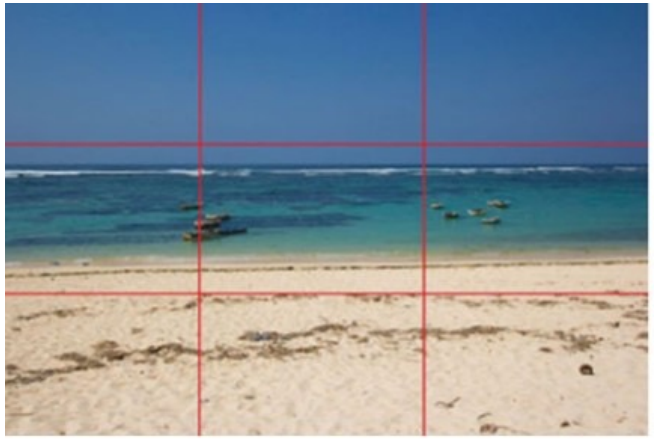
Motivation

Image cropping removes unwanted objects and re-position the main subject according to composition rules.

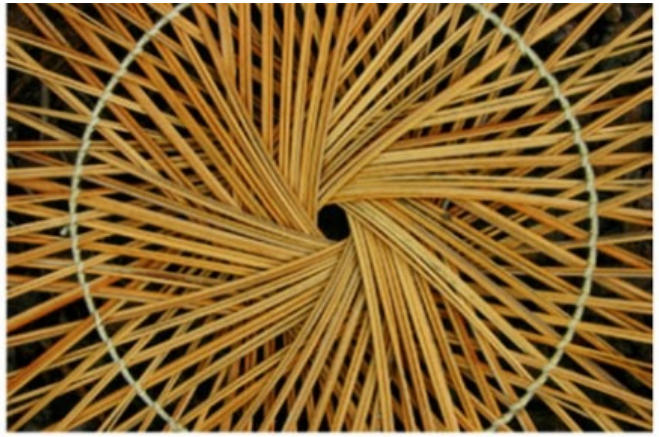


Motivation

(a) Rule of thirds,
(b) diagonal dominance,
(c) visual balance



(a)



(b)

(c)

Motivation

- Rule of thirds,
- diagonal dominance,
- visual balance

Could you find a good composition from these three photos ?

Image Cropping:

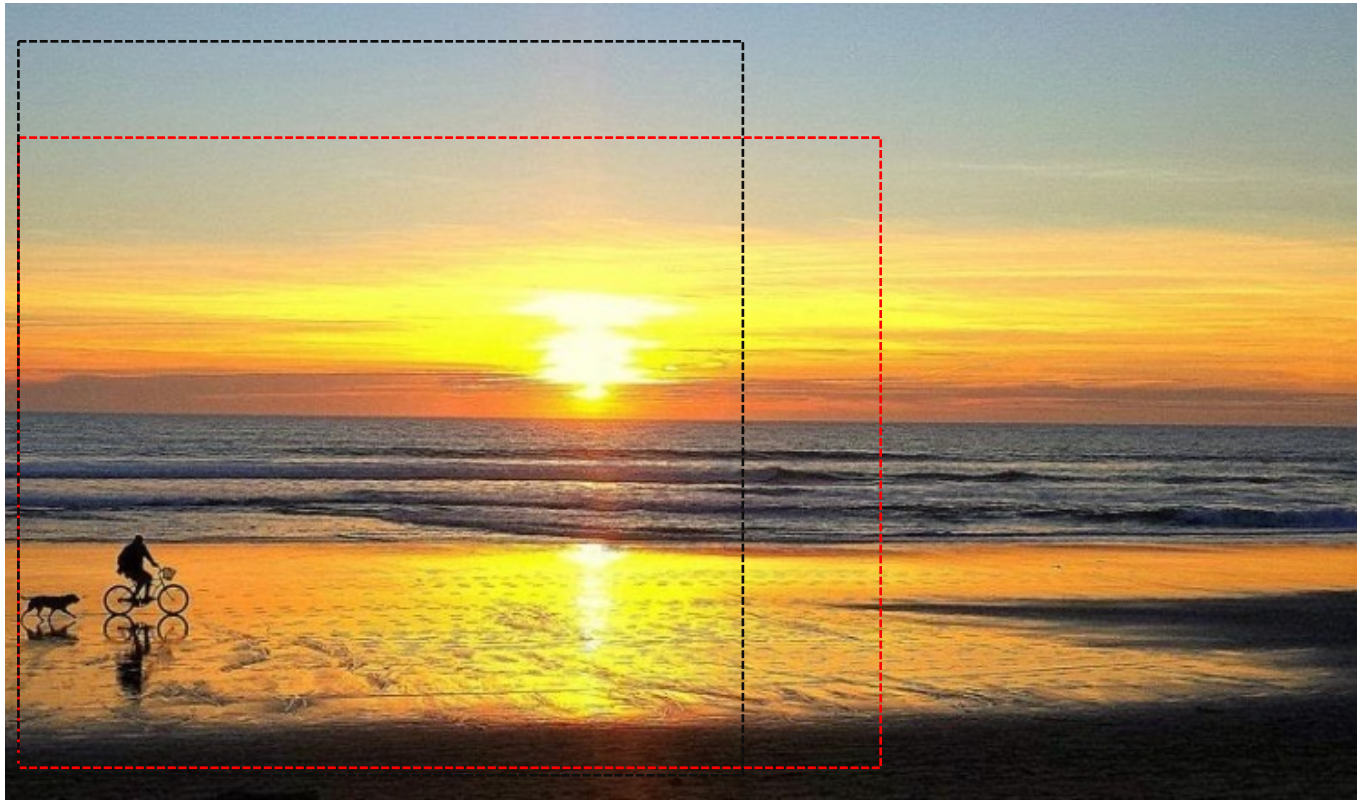
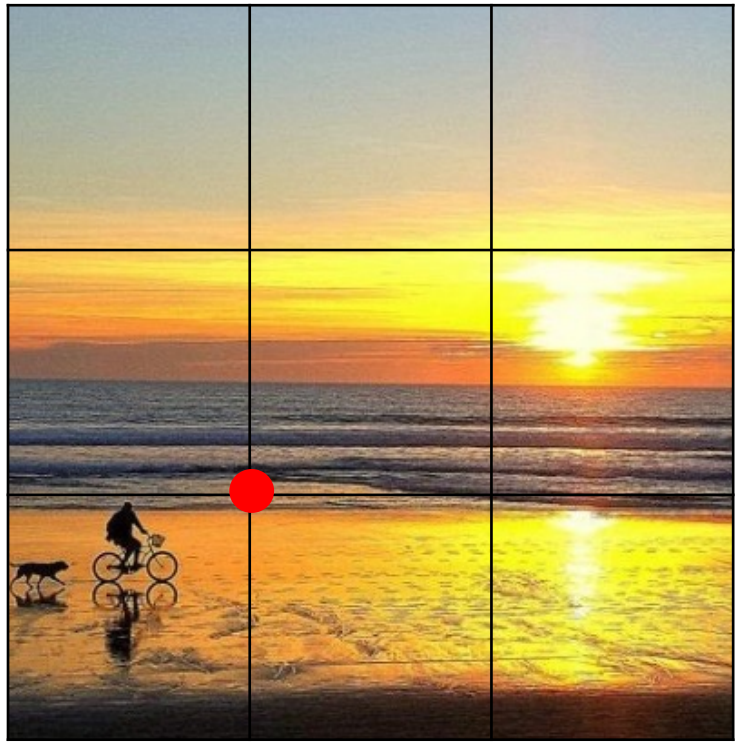


Image Cropping:

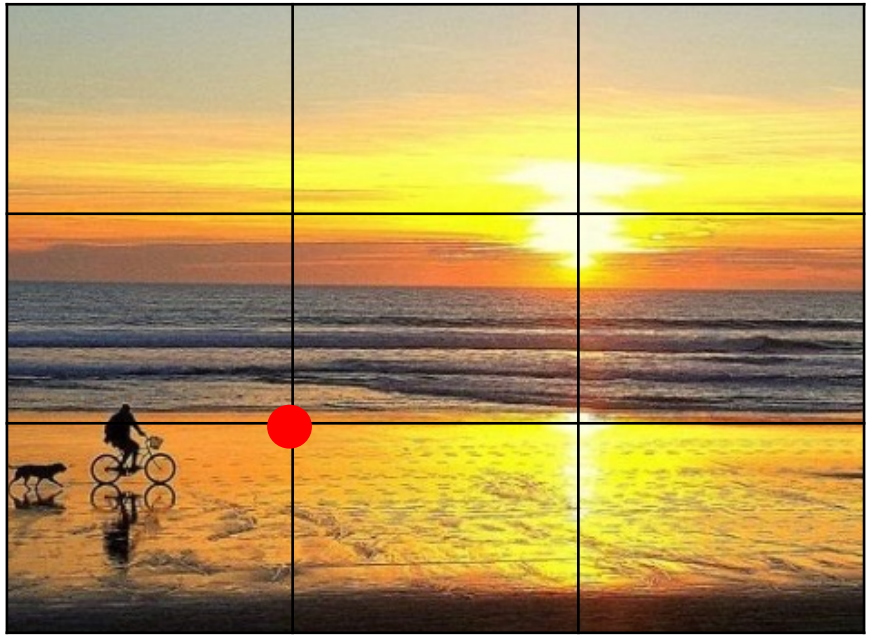
(1)

(2)

Rule of thirds



(1) ✗

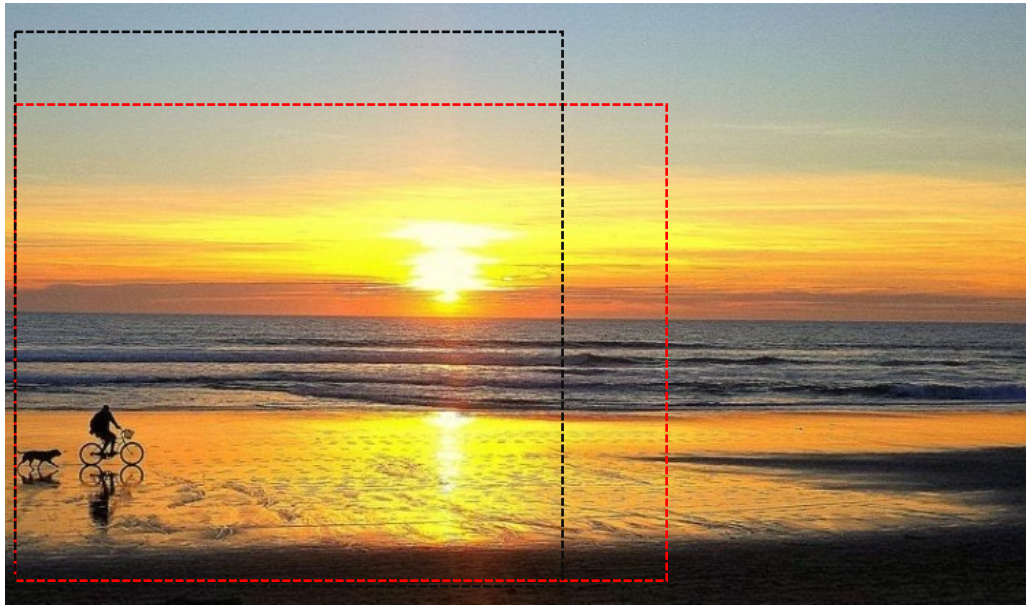


(2) ✗

Motivation

A good composition cannot be obtained by inward cropping when the main object either is (1) too close to the image border or (2) occupies a large portion of the image.

Inward Cropping:

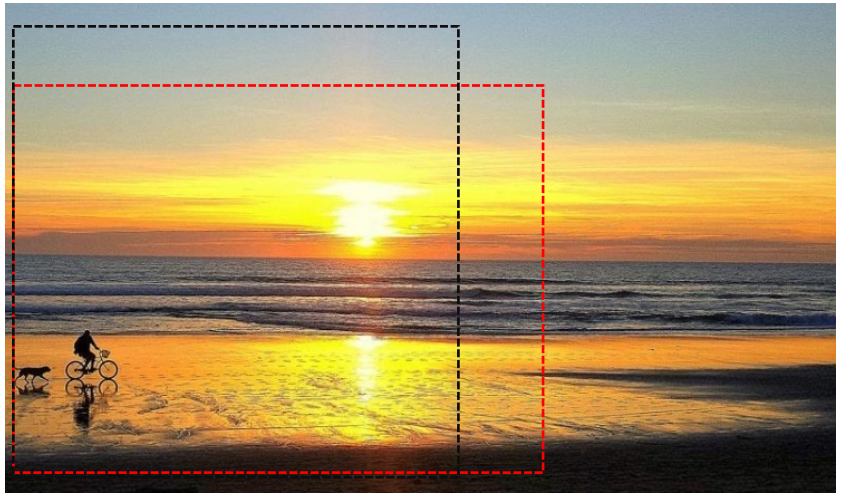


(1)

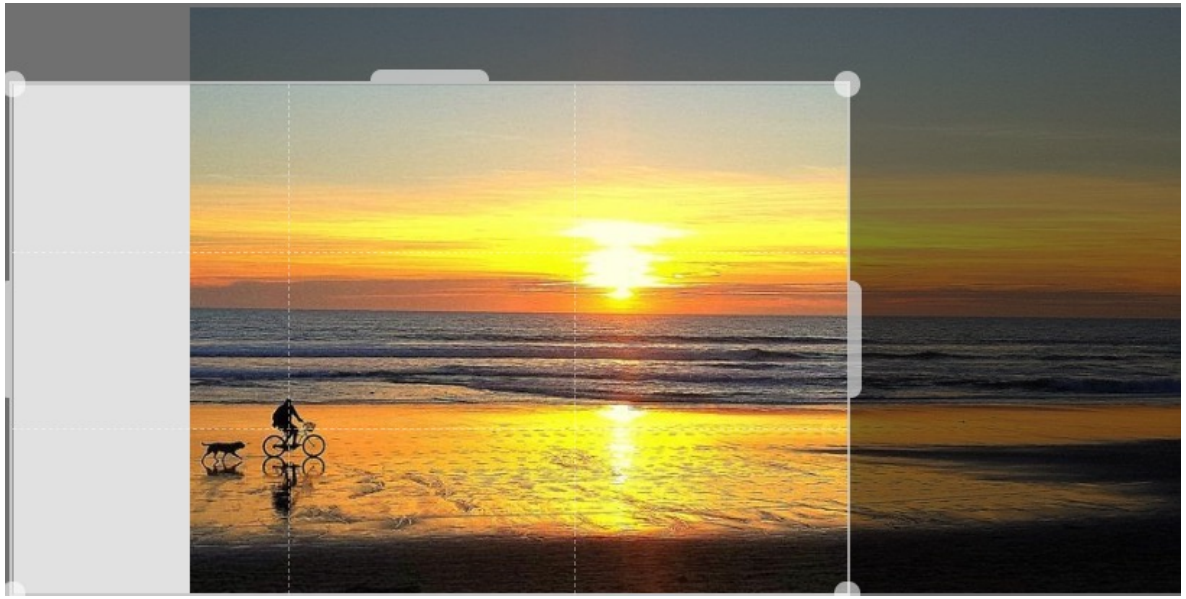
(2)

Motivation

The cropping window should not be limited inside the field of view (FOV) of the given image.



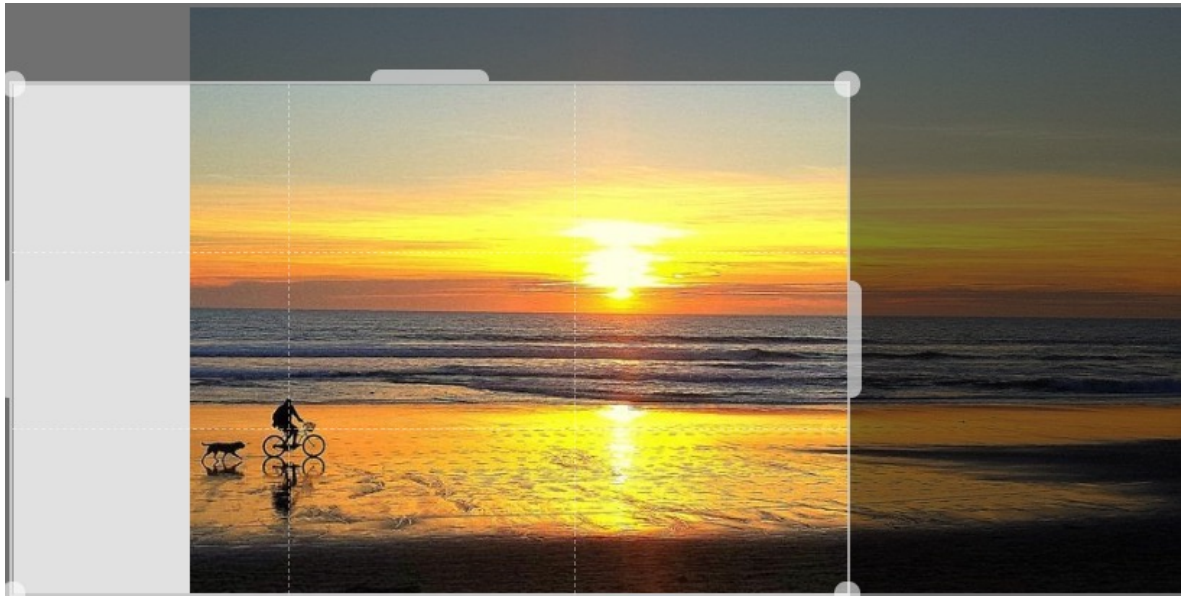
Inward Cropping



Outward Cropping

Outward Cropping

- The image is simply a fraction of that larger scene.
- An ideal cropping method should allow crossing the image boundaries to find the optimal view.



√

Related Work

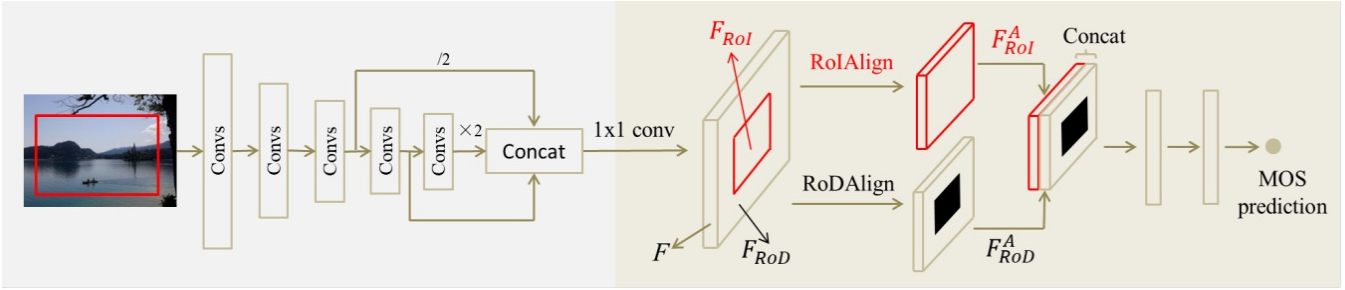
- **GAICD Dataset：**

1,236 images in total. Each Image is annotated with 80~90 anchor boxes, along with their corresponding aesthetic score.

- Image Cropping is a **Regression** problem:

$$M_N = F_\theta(I, V_N),$$

Where $M = m_1, \ldots, m_n$ represents a set of aesthetic scores, F_θ represents a network, I is input image, $\mathrm{V} = \mathrm{v}_1, \ldots, v_N$ is pre-defined anchor boxes.



(a) Multi-scale feature extraction (b) Image cropping modeling

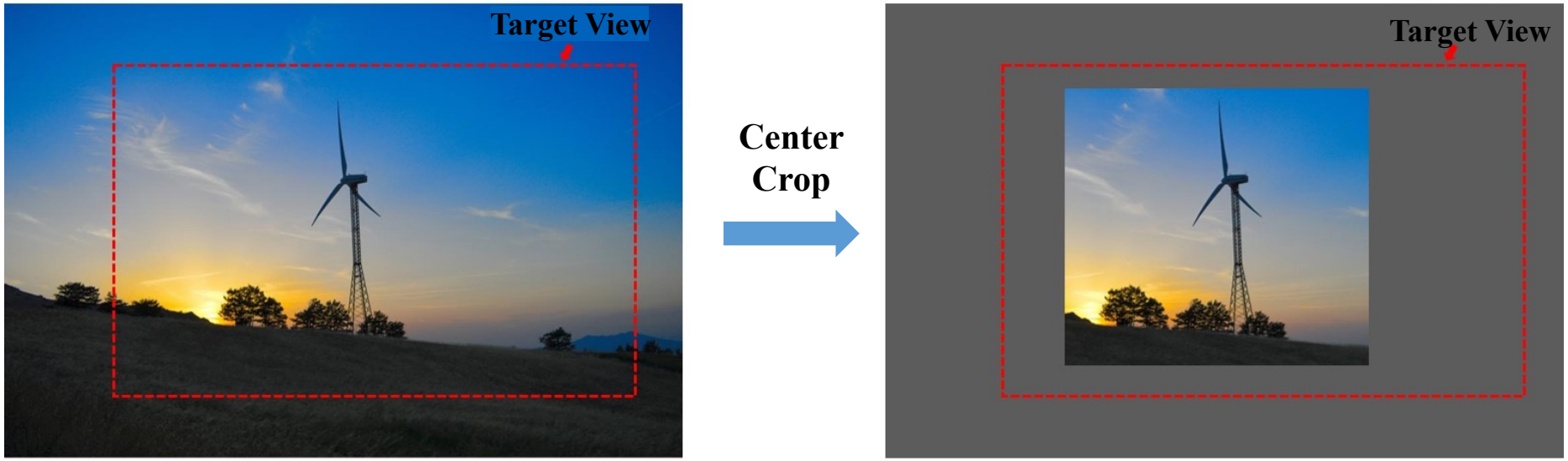
[1]

[1] Grid Anchor based Image Cropping: A New Benchmark and An Efficient Model.

南開大學

Outward Cropping Dataset

- **Randomly center-crop the original image from the GAICD dataset.**
- Input center-cropped image to find the original target view.



Outward Cropping Dataset

- Randomly center-crop the original image from the GAICD dataset.
- **Input center-cropped image to find the original target view.**

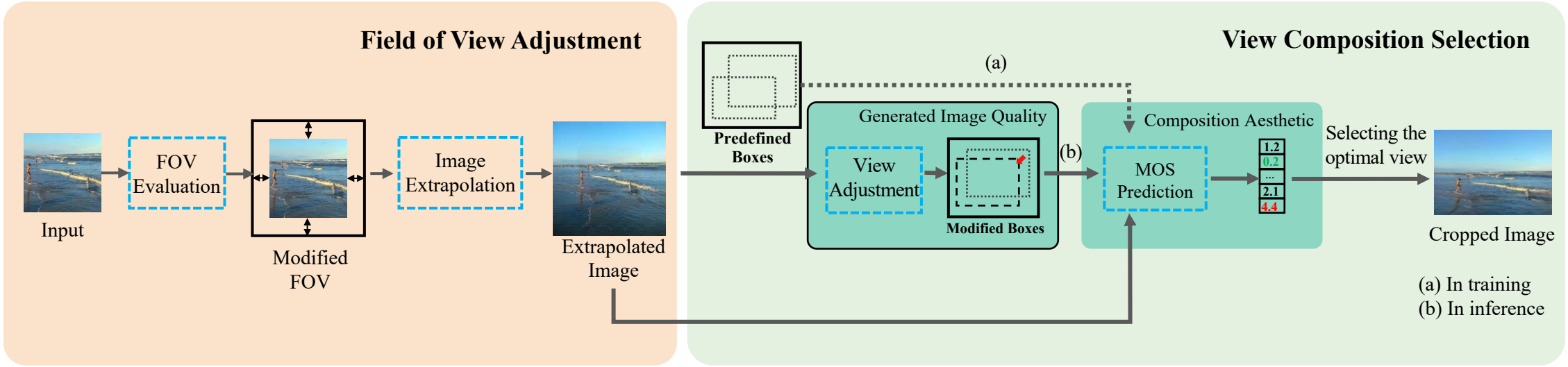
Input

F_θ

Output

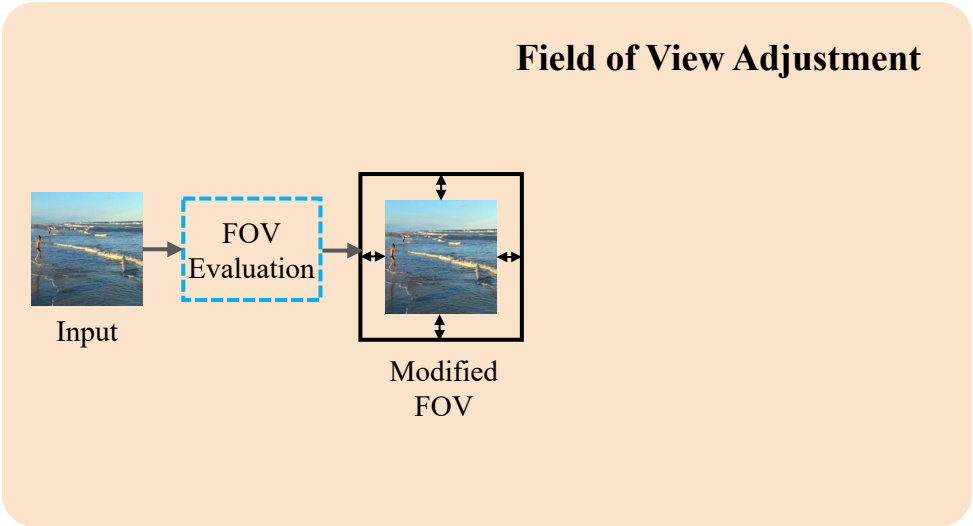
Pipeline

- FOV Evaluation
- Image Extrapolation
- View Composition Selection



Pipeline

- **FOV Evaluation**



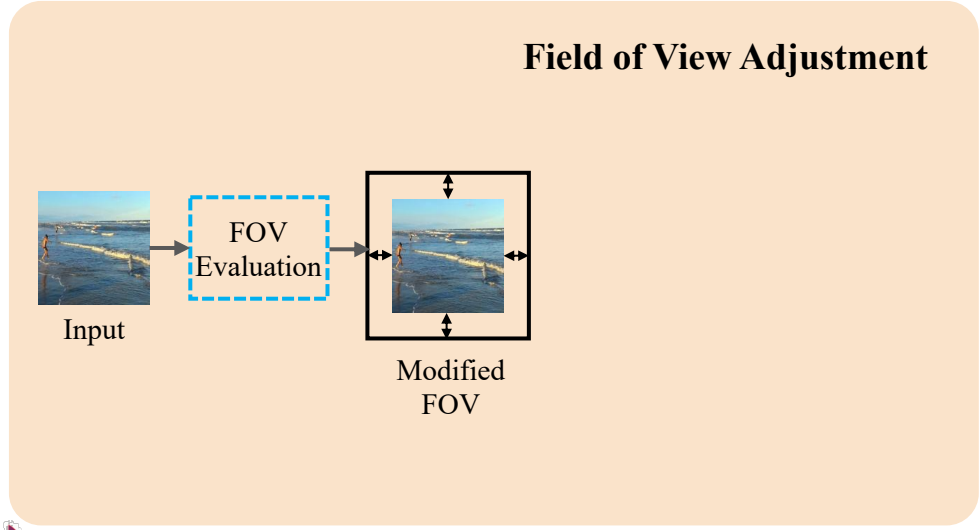
Pipeline

- **FOV Evaluation**

FOV Evaluation is a **multi-classification** task.

Regression is more straightforward but :

- It can be too subjective.
- It may be hard to achieve convergence.



Method

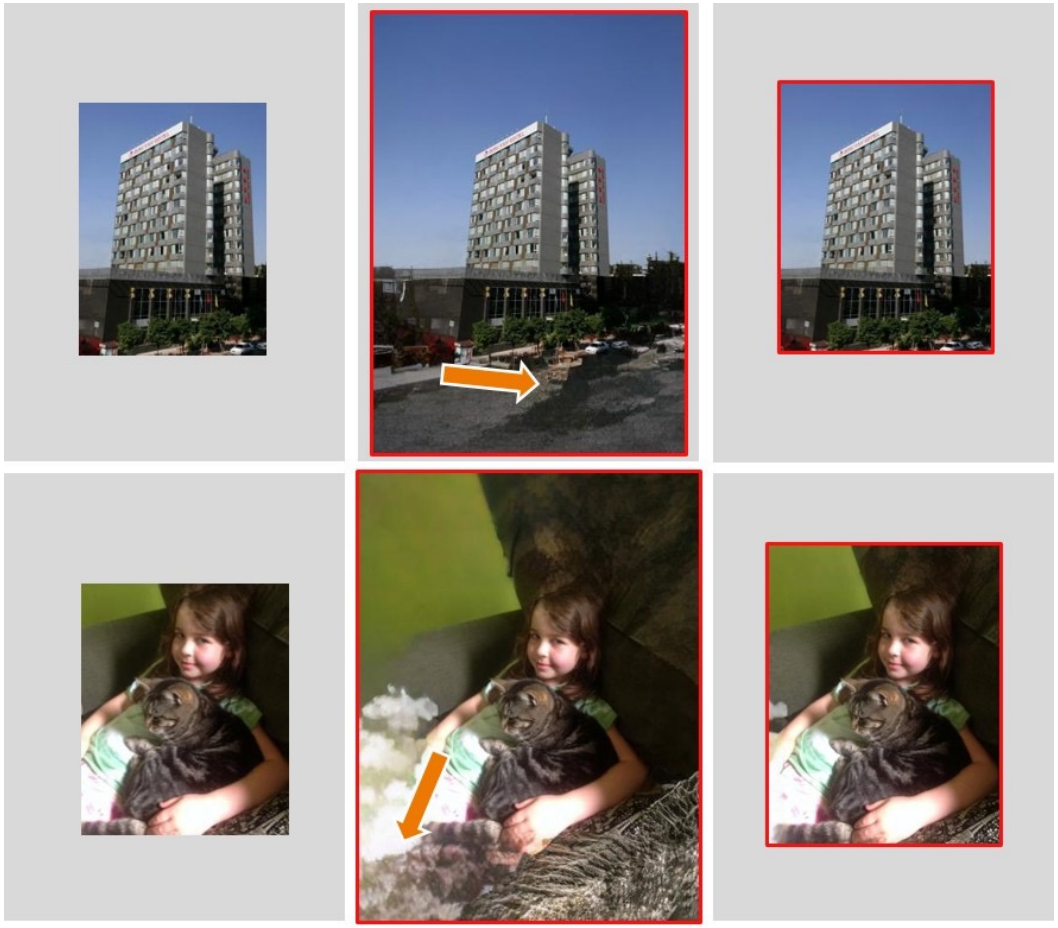
The simplistic method of FOV evaluation:

Maximum the FOV

Problem :

1 ）Unable to guarantee the image quality.

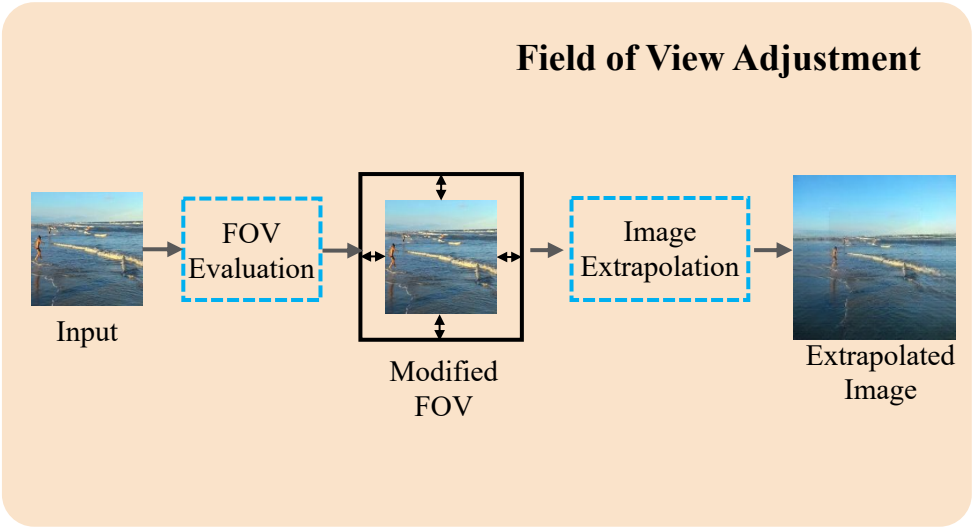
2 ） Increased the search space.



Input | Maximum the FOV | Ours

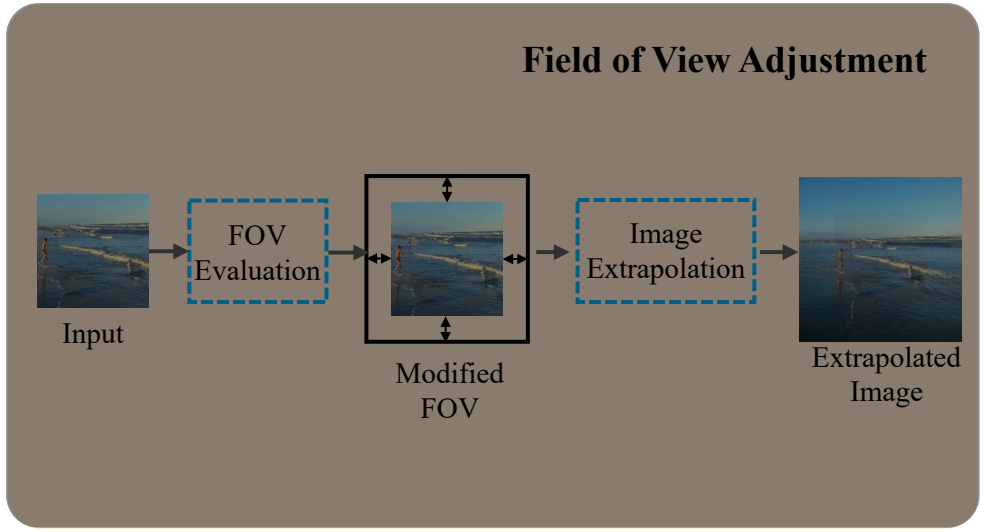
Pipeline

- FOV Evaluation
- **Image Extrapolation**



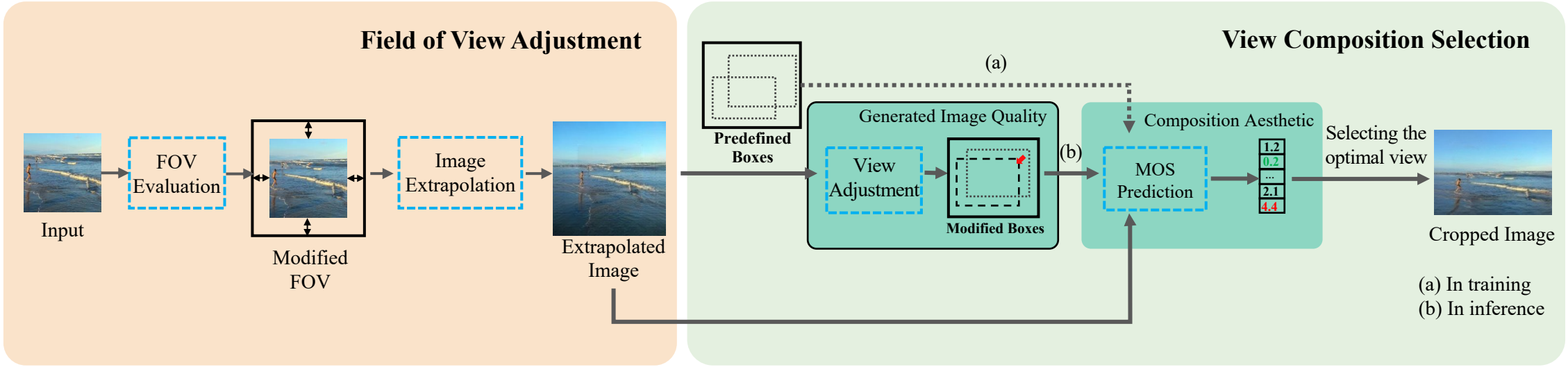
Pipeline

- FOV Evaluation
- **Image Extrapolation**



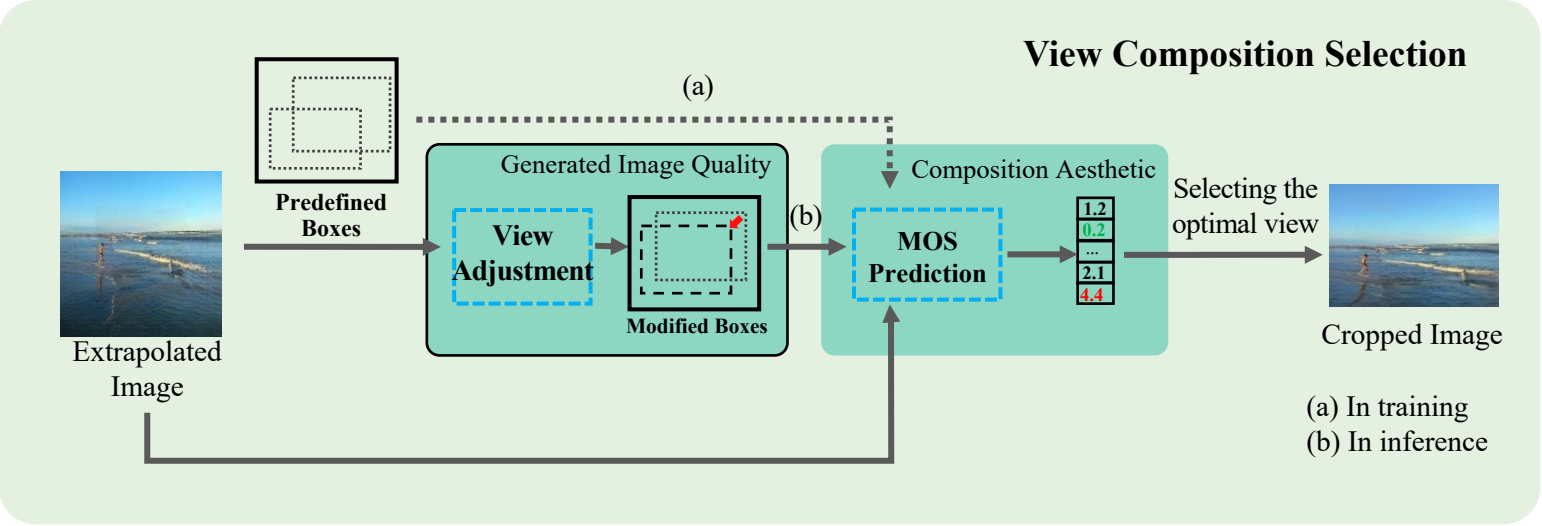
Pipeline

- FOV Evaluation
- Image Extrapolation
- **View Composition Selection**



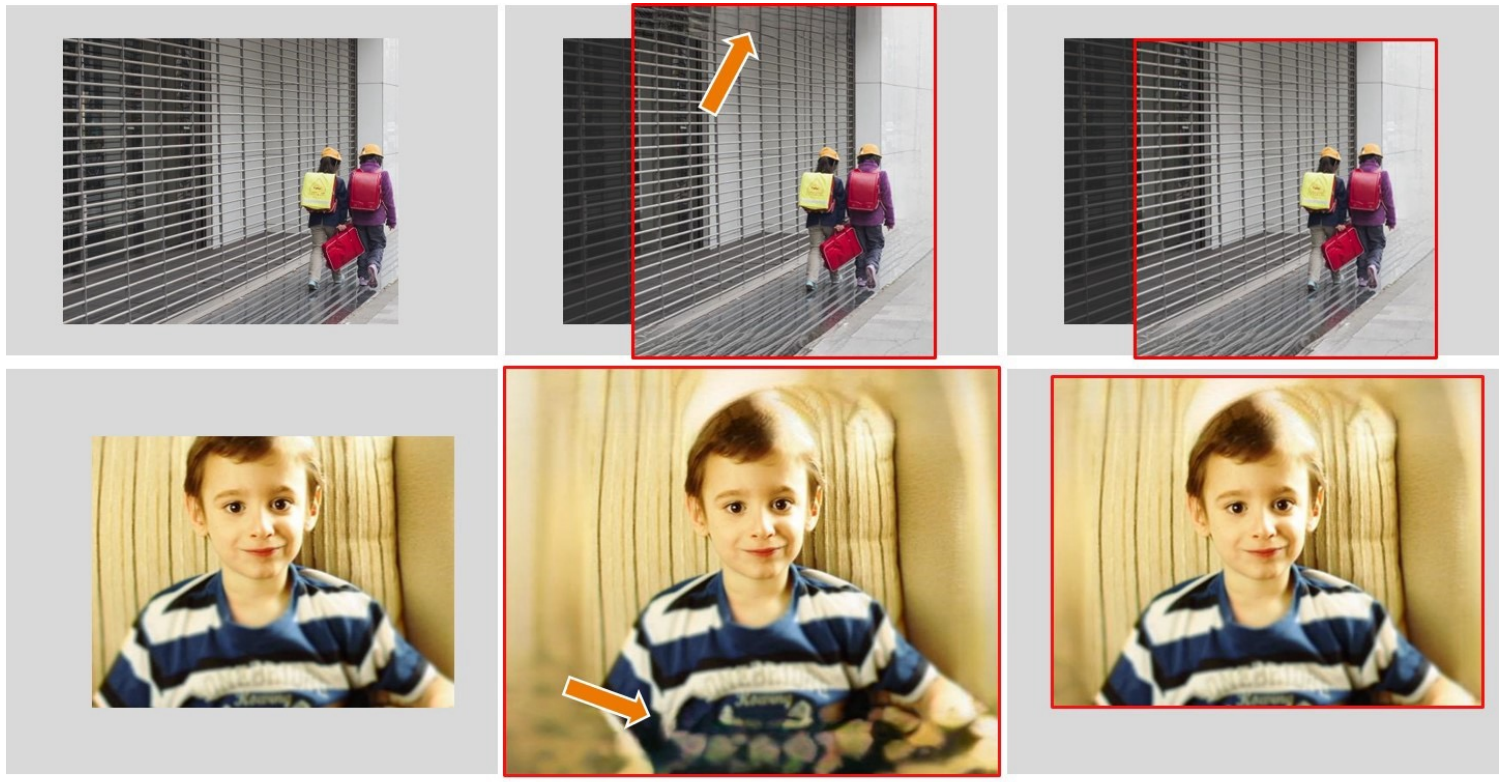
Pipeline

- FOV Evaluation
- Image Extrapolation
- **View Composition Selection**



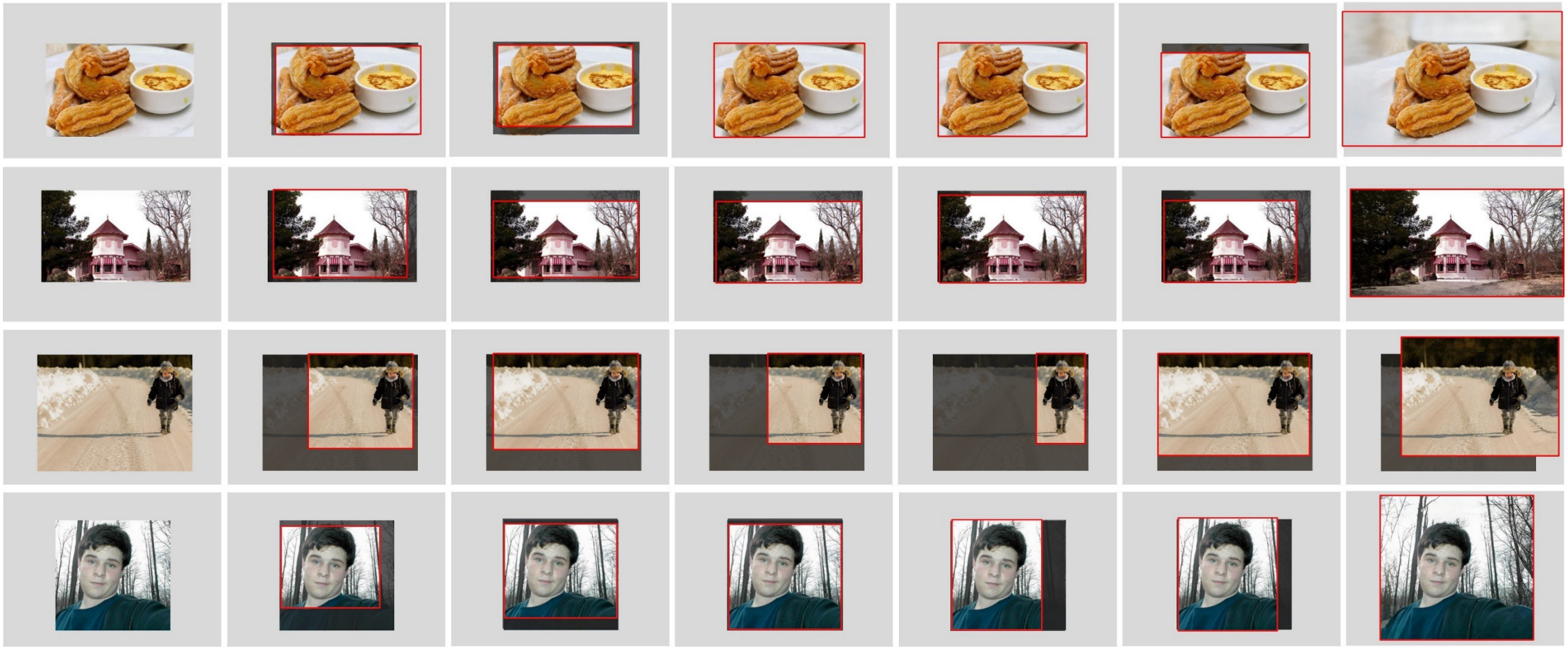
Method

The view adjustment can effectively improve the quality of the image within the candidate box.



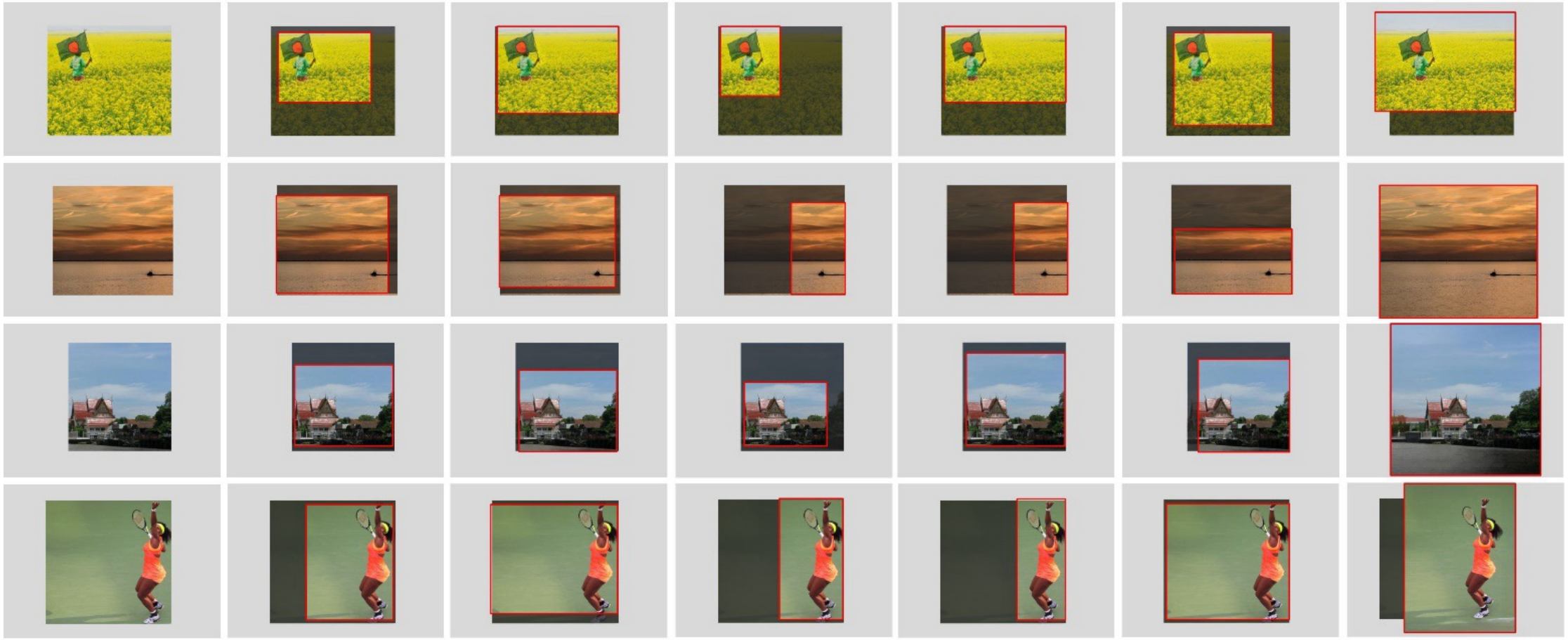
Result

Qualitative comparison (Outward Cropping)



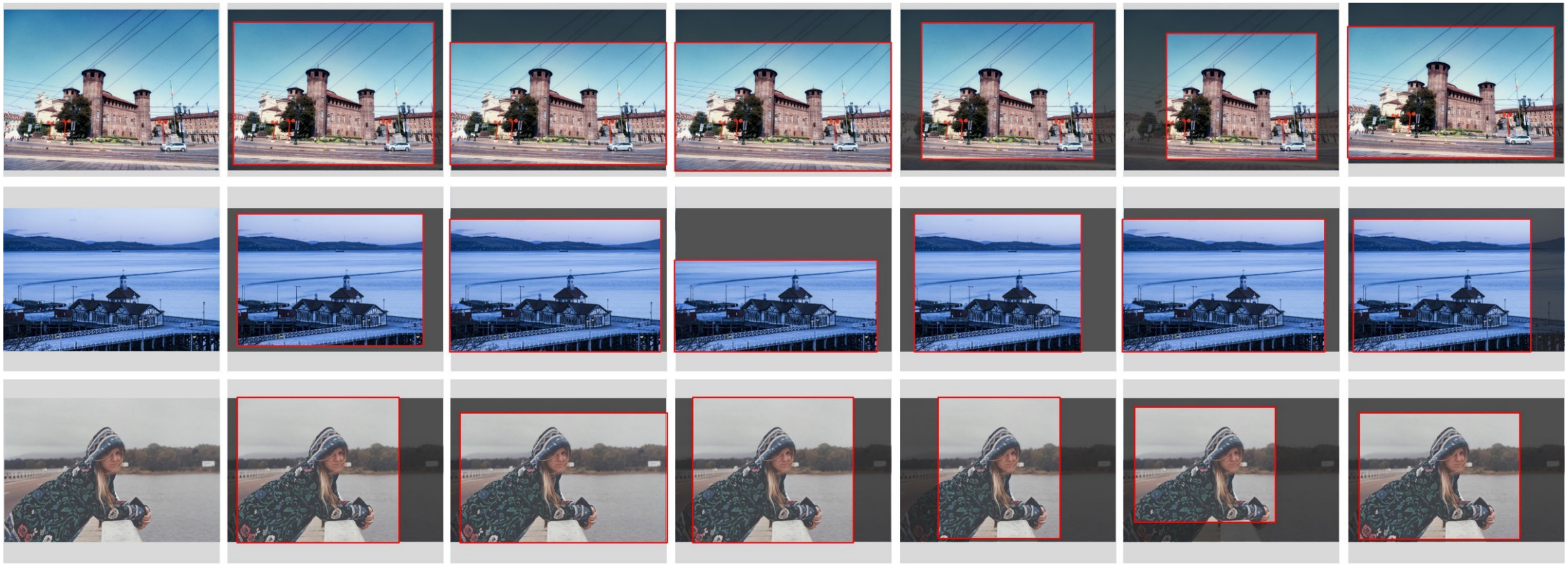
Result

Qualitative comparison (Outward Cropping)



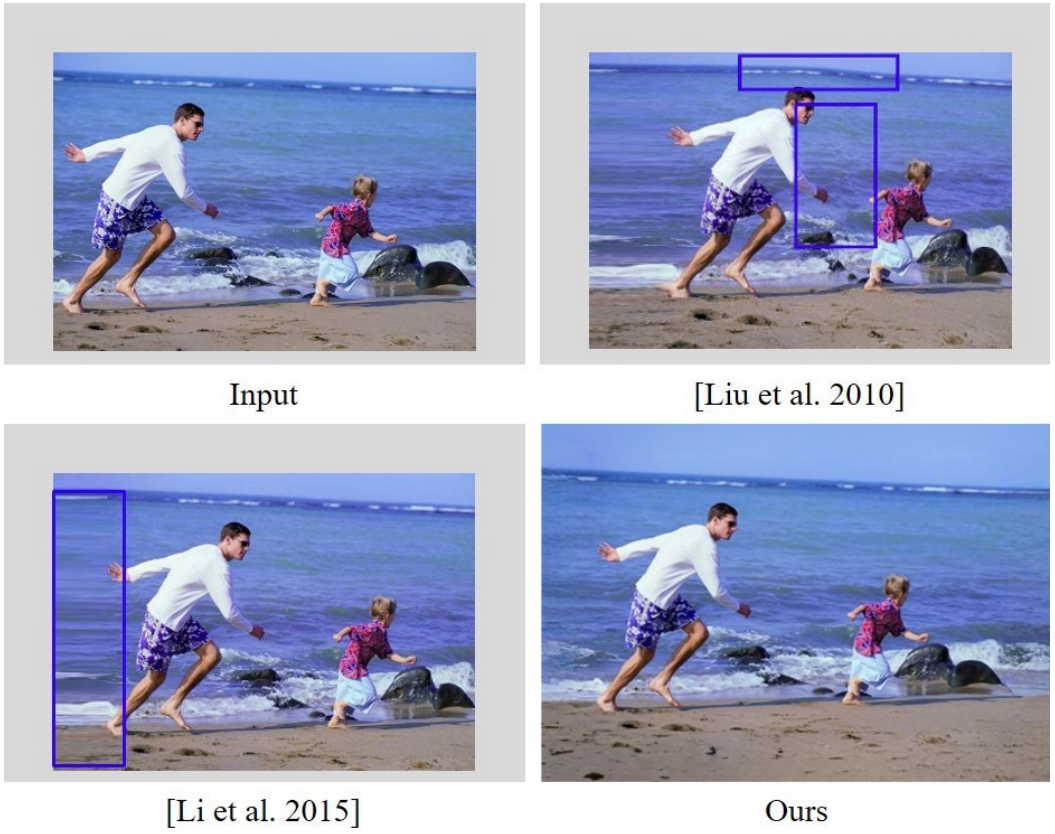
Result

Qualitative comparison (Inward Cropping)



Result

Compare with other Image composition methods.



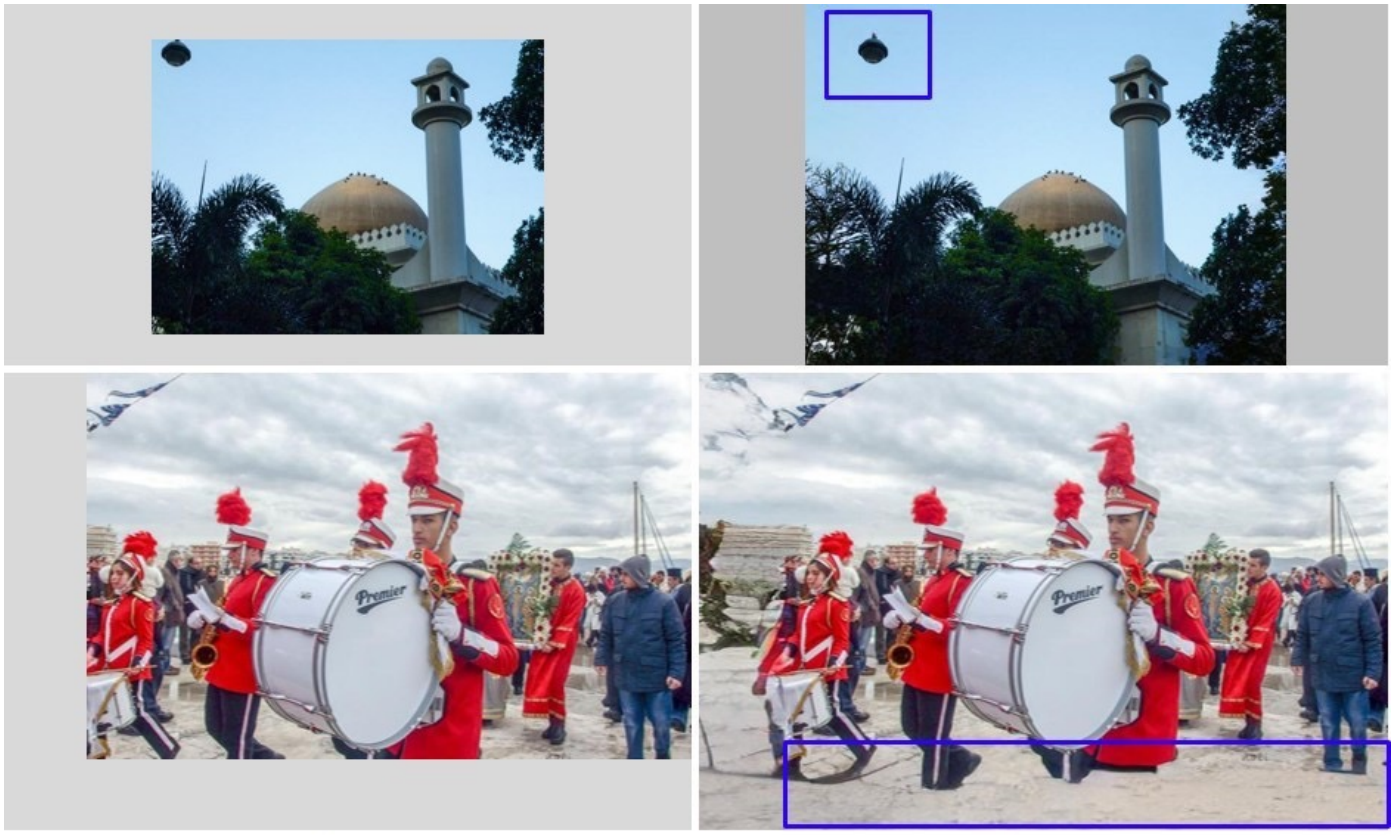
- [Liu et al.2010]: image warping based method.
- [Li et al. 2015]: seam carving based method.

Limitation

Fail Cases in:

- main object is missing essential parts or global context

- Locally realistic content √
- Semantically Wrong ×



Input Ours

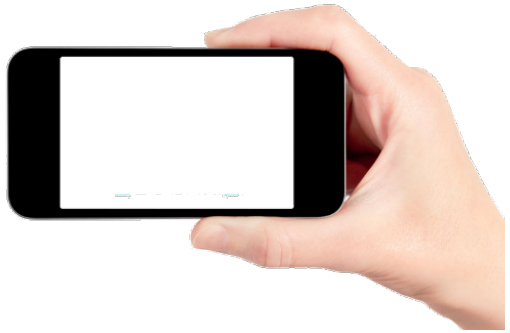
Conclusion

- We have presented a novel aesthetic-guided outward cropping.
- Our method achieves a good trade-off between composition aesthetics and image extrapolation quality.
- Extensive experiment results show that our method can generate a more visually pleasing composition than existing image cropping methods, especially when the original FOV lacks an aesthetic composition

Further Direction

1. Free View Point Image Cropping.

- More Freedom to choose the optimal cropping window.



2. Why for this Cropping.

- Providing the reason behind the choice.

F_θ

Window position.

Why choose this window:
<photograph subject>,
describe the scene, the reason.