Position-aware and Symmetry Enhanced GAN for Radial Distortion Correction

Yongjie Shi, Xin Tong, Jingsi Wen, He Zhao, Xianghua Ying and Hongbin Zha
Peking University
Outline

• Backgrounds
• Method
• Experiment
• Conclusion
Introduction

• Traditional computer vision algorithms, such as pose estimation and 3D reconstruction, usually depend critically on the assumption of the ideal pinhole camera model.

• However, most lenses in commonly used cameras suffer from lens distortion.

• Eliminating the radial lens distortion of an image is a crucial step for many computer vision applications.
Method overview

Generator

Input image → AIF Layer → Image after AIF → G-Encoder → SE Layer → G-Feature map → Decoder → Pixel flow

Warp&Concat

Warp → Warp & Concat → Curved fitting → Smoothed pixel flow → Warp → Output Image

Discriminator

Wasserstein Distance

Ground truth pair → Generated pair → D → Wasserstein Distance
Adaptive Inverted Foveal (AIF) Layer

- AIF can be described as
  \[ I' = W \odot I \]

  Where \( I \) is the input image. \( I' \) denotes image after intensity adjustment. \( W \) is the weight of AIF layer:
  \[ W(x, y) = \alpha \times \left( 1 - e^{-\frac{(x-w)^2 + (y-h)^2}{2\sigma^2}} \right) + \beta \]

- AIF is significant since AIF layer is able to transform the deformation to the intensity of the image.
Symmetry Enhanced Generator

- Symmetry enhanced CNN is proposed to solve the equivariance in CNN under a certain transformation.
- The operation of rotating convolution kernel motivates us to use G-CNN for distortion correction.

- The total loss can be defined as

\[ L_G = \frac{1}{N} \sum_{i=1}^{N} (-D(\hat{F}_i, \hat{T}_i) + \lambda_G|\hat{F}_i - F_i|) \]

Where \( \hat{F}_i \) and \( F_i \) are the output pixel flow map ground truth, respectively. \( \hat{T}_i \) are warped image. \( D \) is discriminator. The first term and the second term indicate adversarial loss and content loss, respectively.
Discriminator

• For discriminator, the loss function can be defined as

\[ L_D = \frac{1}{N} \sum_{i=1}^{N} \left( D(\hat{F}_i, I_i) - D(F_i, I_i) \right) + \lambda_D \frac{1}{N} \sum_{i=1}^{N} (\| \nabla_{\hat{F}} D(\hat{F}_i) \| - 1)^2 \]

The first term and the second term denote adversarial loss gradient penalty, respectively.
Experiment

We first compare our method with previous method in synthetic dataset.

<table>
<thead>
<tr>
<th>Method</th>
<th>#Params (M)</th>
<th>Dataset [25]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rong et al. [25]</td>
<td>70.66</td>
<td>6.71</td>
</tr>
<tr>
<td>Shi et al. [30]</td>
<td>11.21</td>
<td>4.98</td>
</tr>
<tr>
<td>Miguel et al. [1]</td>
<td>–</td>
<td>5.46</td>
</tr>
<tr>
<td>Pixel2pixel [20]</td>
<td>54.41</td>
<td>6.22</td>
</tr>
<tr>
<td>Unet [26]</td>
<td>54.41</td>
<td>6.27</td>
</tr>
<tr>
<td>Ours</td>
<td><strong>2.90</strong></td>
<td><strong>3.54</strong></td>
</tr>
</tbody>
</table>
Experiment

Then we compare our method with previous method in real image taken by ourself.
Conclusion

• In this paper, we present a novel method based on GAN for radial distortion correction.

• We develop a position-aware AIF layer to transform the deformation to the intensity of the image.

• Rotation symmetry enhanced convolution kernels are applied to learn geometric features from structured scenes explicitly.

• Further experiments show our method outperforms previous methods in both synthetic and real images.