| Title | パラメトリックな損失関数に基づく情景テキスト認識のための画像超解像 |
|---|---|
| Author(s) | SUPATTA, VIRIYAVISUTHISAKUL |
| Citation | |
| Issue Date | 2023-09 |
| Type | Thesis or Dissertation |
| Text version | ETD |
| URL | http://hdl.handle.net/10119/18779 |
| Rights | |
| Description | Supervisor: NGUYEN, Minh Le, 先端科学技術研究科, 博士 |

JAIST
JAPAN
ADVANCED INSTITUTE OF
SCIENCE AND TECHNOLOGY

Japan Advanced Institute of Science and Technology

# Parametric Loss Based Super-Resolution for Scene Text Recognition

Supatta Viriyavisuthisakul
2020420

Abstract

In this study, we propose the application of multiple parametric regularizations and parametric weight parameters to the loss function of the scene text image super-resolution (STISR) method to improve scene text image quality and text recognition accuracy. STISR is regarded as the process of improving the image quality of low-resolution (LR) scene text images to improve text recognition accuracy. In a previous study, a text attention network (TATT) was introduced to reconstruct high-resolution scene text images; the backbone method involved the convolutional neural network (CNN)-based and transformer-based architecture. Although it can deal with rotated and curved-shaped texts, it still cannot properly handle images containing improper-shaped texts and blurred text regions. This can lead to incorrect text predictions during the text recognition step. Parametric regularization in the single-image super-resolution (SISR) model has recently been proposed to deal with artifacts and restore the unseen texture in the natural image domain. However, unlike STISR, SISR does not focus only on text information. Here, we design and extend it into three types of methods: adding multiple parametric regularizations, modifying parametric weight parameters, and combining parametric weights and multiple parametric regularizations. Experiments were conducted and compared with state-of-the-art of STISR models. The results showed a significant improvement for every proposed method. Our methods achieved the best text recognition accuracy of 80.4% for the easy set, 64.1% for the medium set, and 46.5% for the hard set of Textzoom. Moreover, our methods generated clearer and sharper edges than the baseline with a better-quality image score.

Keywords: Scene Text, Image Reconstruction, Trainable parameter, Parametric, Regularization, Loss function