論 文 の 内 容 の 要 旨

論文題目　Image-to-Image Translation for Content Creation Support
（コンテンツ生成支援のための画像変換技術）

氏　　名　井上　直人

The emergence of digital tools has opened up an opportunity for ordinary people to create visual content digitally, such as images and videos. However, many steps still have to be done manually and are time-consuming in the creative workflow. Many methods that go beyond classical low-level filters have been introduced, such as image segmentation, owing to the recent evolution of machine learning methods.

Image-to-image translation (I2IT) is a task to convert an image in one domain to a corresponding image in another domain. I2IT is essential since many tasks in visual content creation are I2IT. Recent works have established a unified solution for supervised I2IT using convolutional neural networks (CNN). All one has to do is to prepare many pixel-aligned input-target image pairs for supervised learning. However, sometimes it is impossible or significantly hard in terms of time amd cost to collect the paired data. In this thesis, we have tackled several I2IT tasks that suffer from the aforementioned issue by exploring underlying priors to form an image.

First, we address the task of tracing photographs to create line drawings. We propose a two-stage extract-and-refine model to immitate the tracing workflow and present a self-supervised loss to learn the refinement process. We confirm that our model produces much more clean and expressive line drawings and illustrate two practical applications.

Second, we propose a novel task of adding a non-directional shading effect on an image. Inspired by approaches for a real-time approximation of global illumination in rendering, we propose to generate ambient occlusion (AO) map from an input image as a proxy task. We show that our model trained on synthetic RGB-AO pairs dataset performs well on a broad range of images and illustrate image editing applications: image composition and geometry-aware contrast enhancement.

Third, we address the task of shadow removal and detection. We derive a pipeline to synthesize the dataset from a classical physically-grounded shadow illumination model. We

demonstrate that shadow removal and detection models trained on the dataset from our pipeline generalizes well to a broad range of datasets, thanks to the synthesized data's fidelity and diversity.