EE 367 Project Proposal

Yichen Jiang, ycjiang@stanford.edu

Feb 2024

1 Motivation

As a common image or video editing schemes, image harmonization (seen as a subarea in image composition tech) aims to harmonize the color, texture, or style between foreground objects and background, intending to erase the inconsistency between foreground and background. There are multiple types of inconsistency including appearance inconsistency (e.g different resolutions), geometry inconsistency (e.g., unrealistic size or shape), and semantic inconsistency (e.g., unreasonable semantic contexts), among which image harmonization aims to adjust the appearance of composite foreground according to composite background to make it compatible with the composite background. In this project, I plan to investigate into state-of-art technologies used in image harmonization, with an interest of utilizing deep learning network into constructing realistic images.

2 Related Work

Conventional image relighting is designed to adjust an image or an object under different illumination conditions, which resembles image harmonization, but such relighting techs usually requires explicit illumination conditions, 3D shape information, and texture information, making the whole algorithm relatively sophisticated and hard to implement. Such requirement also makes it not very applicable in complex real-world scenarios.

Later, non-rendering based methods flourished in image harmonization domains, and these methods mainly focus on extracting different types of

features from background and adjust corresponding features in foreground to achieve a good match between background and foreground.

Deep learning methods target at making image the harmonized images indistinguishable from real-world images. They achieve so by utilizing machine learning schemes like CNN classifier or adversarial learning. The recent emergence of abundant image harmonization datasets also facilitate greatly the utilization of deep learning network in outputting realistic composited images.

3 Dataset

There are multiple usable datasets that could be used in training neural network, including RealHM, HFlickr, and HVIDIT dataset. Each training (or testing) sample consists of a composited input image, the mask of foreground object, and the corresponding ground-truth image (realistically composited image). Different datasets generate such sample using different methods, among which the most frequently seen are forward adjustment, backward adjustment, and replacement, as illustrated below.

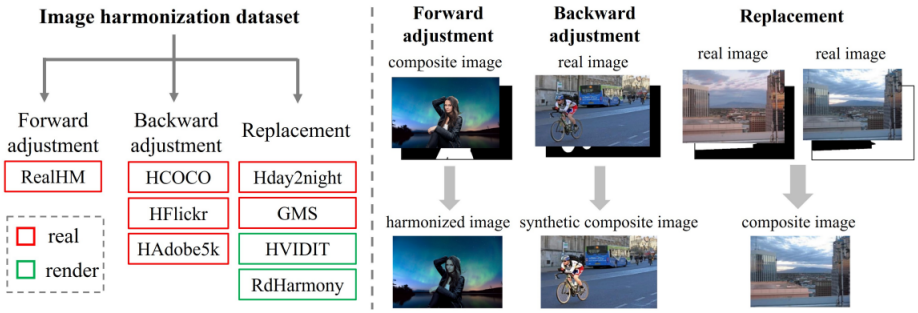

Figure 1: three most common methods to construct image harmonization dataset and corresponding samples from datasets adopting aforementioned three methods

4 Timeline

Week 7: Fix the method and dataset that the project will be built on; Search for more related research and work; Start the project officially
Week 8: Implement image harmonization; Assemble dataset
Week 9: Test the model on the dataset; Adjust model to improve the results
Week 10: Write project report and poster; Submit the project

5 References

[1] Li N, Wenyan C, Liu L, Yan H, Bo Z, Jing L, Liqing Z: Making Images Real Again: A Comprehensive Survey on Deep Image Composition, 7 Aug 2023
[2] Rohit Kumar Pandey, Sergio Orts Escolano, Chloe LeGendre, Christian Haene, Sofien Bouaziz, Christoph Rhemann, Paul Debevec, and Sean Fanello. Total relighting: Learning to relight portraits for background replacement. In SIGGRAPH, 2021
[3] Jun-Yan Zhu, Philipp Krahenbuhl, Eli Shechtman, and Alexei A Efros. Learning a discriminative model for the perception of realism in composite images. In ICCV, 2015.
[4] Fangneng Zhan, Shijian Lu, Changgong Zhang, FeiyingMa, and Xuansong Xie. Adversarial image composition with auxiliary illumination. In ACCV, 2020.
[5] Devin Schumacher. Image Harmonization. From SERP AI