Proposal of MSc projects

Sirui Wu

June 24, 2022

---

1 Basic information

- Project title: Infrared and visible image fusion under poor image quality.

- Student name: Sirui Wu

- Student ID: 2274668

- Supervisor name: Felipe Orihuela-Espina

2 Project aim

This project is aiming to merge a low-quality visible and far infrared photo into one picture by a neural network. Since the wavelength of visible light only covers 400-760mm, the information covered by this waveband only accounts for a very limited part of the object's spectrum, and infrared has the advantage of strong penetration, not affected by the weather, so it can be used to detect the temperature of the target object. I will combine convolutional neural network(CNN) and Laplacian pyramid method to complete this task.

3 Related work

Actually, there has been numerous thesis related to this field. Image fusion has been around over 70 years, but it is still a challenging problem. The currently available algorithms are divided into traditional methods and deep learning methods. Now there are about 20 algorithms exists to solve this problem. Before 2019, there were basically supervised models or direct application of pre-trained models, and in 2020, self-supervised and GAN models were applied to this model. And in 2021, many new ideas emerged instead of rote modify the network structure and loss function. I will discuss it elaborately in my thesis's introduction section.

4 Project Objectives

This project can be decomposed into these four sub objectives:

- Implementing a neural network by pytorch according to the original paper and correctly get the desired results;

- Determine a effective metric to measure the quality of the image, and the lost function;

- Modify the Gaussian kernel size, adjust the network layer and optimize some hyperparameters of this network and then calculate the quality of the image;

- Repeat (3) until the satisfied result is obtained.

5 Methodology

In order to get a better fusion effect, I will combine both traditional algorithm(MST decomposition) with their deep learning. The traditional algorithm will decompose the image into multiple layers with different frequencies:

Algorithm 1 Laplacian decomposition

 1: **repeat**
 2:     Applying Gaussian blurring Kernel on the image
 3:     Subtracted the blurred image with the original Image
 4: **until** Decomposition has reached the required number of layers

---

The DC part includes most essential part of the image and most of the energy. In this frequency band, the information from infrared and visible images will be very different. And the high frequency part contains edges of a object,which would be similar in the infrared band.

After the Laplacian pyramid has been constructed, then a Siamese Neural Network is applied to infrared photo and visible light photo, it would contain the weight of similarity of this two photos.

I would add a schematic network ,input visible and infrared 16x16 pixels patch simultaneously into it. Then, these input pixels will be processed by three convolutional layers, one max pooling layer, and eventually a fully connected layer to only two pixels. Slide this patch from the top to the bottom, then the similarity matrix is obtained as it size would be (x-8+1)*(y-8+1).Once the weight values are determined, we move the image perceptual fields one by one to obtain the weight matrix A. Eventually, the infrared and visible light images will be fused based on these rules:

$$C_l = \frac{\arctan \omega P}{\arctan P}$$

and

$$LF_n = C_l LA_n + (1 - C_l)LB_n$$

, while LF_n represents the nth layer of the Laplacian pyramid, and ω represents the weight. Reconstruct each Laplcaian layer, then the fused image is obtained.

6 Project Plan

I would want to start my project in this sort:

Table 1: 'PLAYER' Raw Data Format	
June 27th – July 2th	Reading related paper and code, learning Pytorch and OpenCV
June 27th – July 2th	Rudimentary network designing and programming
July 4th – July 8th	Code debugging, testing, result discussion
July 11th – July 15th	Complete the overview and experiment session
July 18th	Complete the fist edition of the thesis

7 Risks and contingency plan

Actually, this is a very difficult task that may encounter many difficulties, including:

- The code does not work even after multiple debugging attempts

- Failed to get an expected fusion image and got no idea to optimize the algorithm

- Failed to implement the neural network by the

- Can not find sufficient documentation to inspire my work

To overcome these difficulties, I have to communicate timely with my tutor, endeavor to analyze and debug the trickly and frustrated error, and most importantly,

8 Hardware Resources

To implement this task, a GPU is necessary since the graphics card of a personal computer is not capable to handle a large number of matrix computations. I am likely to run my program on Kaggle's notebook's TPU session(20 hours per week).

9 Data resources

The Github link of our dataset is https://github.com/Linfeng-Tang/MSRS.git. It contains 1,569 image pairs (820 taken at the daytime and 749 taken at nighttime) with spatial resolution is 480×640. However, There are many misaligned image pairs in the MFNet dataset and most infrared images are low signal-to-noise and low contrast. It is a public dataset on Github so both my supervisor and me have permission to this dataset.