HDR+

Chen Qian
Stanford University
@stanford.edu

Tong Yang
Stanford University
@stanford.edu

Wanzi Zhou
Stanford University
wanziz@stanford.edu

1. Motivation

There are mainly two motivations for this project. We want to do a project that are both highly related to the topics and theories we covered in class and are also up-to-date, interesting and influential.

One of the most common problems that we encounter when we shoot photos with smart phones is that there is not enough light, leading to a lot of noise and low dynamic range in the image. Usually there are two standard solutions to this problem: applying analog or digital gain and lengthening exposure time. The former would amplify the noise even more and the latter might lead to motion blur due to camera shake and subject motion[3]. We first find this paper by Google [3] published in 2016 which tries to solve the problem, and especially targeted at smartphone, by using a newly proposed computational photography pipeline called HDR plus (we will cover that more in the relational work), which excels in low-light and high dynamic range scenes and is computationally efficient and reliably artifact-free. We find this really interesting in that the original input photos of this pipeline are a burst of photos with the same low exposure time, while in class we only explored fusing photos with a burst of photos with different exposure time to provide a high dynamic range (from the well-known Debevec paper[1]). And according to this paper the HDR+ pipeline is very computationally efficient and has received very good user and press response. This topic is highly relevant to what we have learned in class and also has very high application values.

And thus, we want to start our project by understanding and implementing the ideas from this paper and make further explorations.

2. Related Work

In 1997, Debevec and Malik proposed a method for recovering high dynamic range radiance maps with burst photography[1]. The idea is to take a series of photographs with different exposures, and fuse into one high dynamic range photograph with the response function recovered using the assumption of reciprocity. The ideas works well with the photographs taken with conventional imaging equipment like digital camera, because they usually have relatively large sensor size and easily controlled ISO and exposure time. However, it is much more difficult in the case for cell phone cameras. Cameras on the phone usually have small apertures and sensor pixel size, which means that both the number of photons and the number of electrons each pixels can store are limited. In addition, different exposure time introduces challenges in aligning images because of the different levels of noises and motion blur, leading people to explore new approaches.

In [3], researchers at Google proposed a complete system called HDR plus in 2016. It captures a burst of underexposed frames, aligns and merges these frames to produce a single intermediate image of high bit depth, and tone maps this image to produce a high-resolution photograph. This paper tries to present this commercialized pipeline into an academia-stylized paper, and several main parts in this system include: it is composed of a hardware pipeline and a software pipeline. The input of both pipelines is Bayer mosaic image. When the camera app is launched, the hardware pipeline launches and displays a de-mosaicked low resolution on the phone screen. As soon as the shutter is pressed, a burst of frames is captured at constant exposure and stored in the main memory, then the software pipeline starts, where it aligns and merges the frames in a burst, produces a single intermediate image of high bit depth, and tone maps this image to produce a single full-resolution 8-bit output photograph for compression and storage in flash memor[3]. Several important features of system include: using constant exposure time in the burst of photos (as mentioned in Motivation part different exposure time leads to misalignment issue especially in phone cameras), using Bayer raw image as input instead of demosaicked RGB frame (providing more bits per pixel) and utilizing a novel FFT-based alignment algorithm and a hybrid 2D/3D Wiener filter to denoise and merge the frame. According to the paper, the entire system only takes 4 second on an Android phone, which is pretty efficient. They also put up an HDR+ dataset[4], which we think can be utilized in our project as an oracle baseline.

3. Project Overview

This project is built on top of the HDR+ dataset, with the target of studying and exploring modern techniques of rendering HDR image from LDR images taken from cellphone cameras. One major difference between this project and what we learned from lecture and implemented in the homework 4 is that all images in the burst are taken by the same exposure time and digital gain, see Fig. 1 for details.

Figure 1. First row: a burst of image for conventional HDR converting; Second row: a burst of image in HDR+ dataset

There are two major parts of this project: the first is to find an appropriate method to merge different images in the burst, aiming to capture as many details as possible on the scene. By combining multiple observations of the scene over time, we can achieve high dynamic range and noise deduction. The denoising idea is inspired by Kokaram[5], who proposed the frequency-domain video denoising techniques by applying Wiener filtering in the 3D DFT domain. Another variation of the denoising method we want to explore was proposed by Delbracio and Sapiro in 2015[2]. We can view the captured burst images as nearby frames in the video and apply the same denoising method proposed for video denoising above. Based on the computed weighted average in Fourier domain across the burst, the image is reconstructed by combining the least attenuated frequencies in each image.

The other major part of the project is focused on post-processing of the HDR image, including correction, demosaicking, denoising and tone-mapping. Here we would also like to separately explore removing the Gaussian and Poisson noise during the process.

4. milestones, timeline and goals

Milestone: We break our project into 3 stages:

a. Implement the merging method proposed in [3], including the denoising methods that operate on 3D stacks of matching images patches.

b. Implement the post-processing method proposed in [3], explore both Gaussian and Poisson noise removal.

c. Refining algorithms to improve the performance of our pipeline.

Timeline: We plan to focus on our project from Week 7, then conquer each stage as listed above in one week, and finish poster and report in the last week.

Goals: Our basic goal is to implement the HDR image converting pipeline while assuming the images are already aligned. Furthermore, we set the advanced goal as exploring modern denoising techniques to make the darker parts in our processed HDR more natural and also make the pipeline's running time shorter.

References

[1] P. E. Debevec and J. Malik. Recovering high dynamic range radiance maps from photographs. 1997.

[2] M. Delbracio and G. Sapiro. Hand-held video deblurring via efficient fourier aggregation. 2015.

[3] S. Hasinoff, D. Sharlet, R. Geiss, A. Adams, J. T. Barron, F. Kainz, J. Chen, and M. Levoy. Burst photography for high dynamic range and low-light imaging on mobile cameras. *ACM Transactions on Graphics (Proc. SIGGRAPH Asia 2016)*, 2016.

[4] G. INC. Hdr+ burst photography dataset. 2016.

[5] A. C. Kokaram. Motion picture restoration (doctoral thesis). 1993.