Light Field Enhancement in Low Light Using a Lens-Based Depth

Paper ID XXXX

Abstract—Outside of the laboratory, light and time are rarely things that can be counted on for consistency. Being able to collect a sample of data quickly and under uncertain conditions is an advantage for a computer vision system. Much interest has been shown in using light fields that can provide high quality data in areas where environment variables are uncertain. In this paper, we investigate the problem of enhancing light fields with state of the art enhancement techniques. We focus on using deep learning to enhance the low light light fields, and add our own loss and gradient to improve it. the qualities that make light fields unique, such as perspective shifts, depth, and 2D synthetic views. We show that by adding depth as loss in a neural network, we are able to improve our light field.

Index Terms—Computational Photography

1 Introduction

Low light image enhancement.

Passive sensor is required in the case of filming animals, for archival purposes, quality control, space, or for situations when lighting simply isn't optimal.

Depth might also be a requirement, and for a light field, depth and 2D reconstruction are tied together (Georgiev).

There's lots of research regarding low light enhancement of 2D images, but nothing for Light Fields.

The total spatial resolution for the light field camera is comparable to that of traditional cameras, but the micro lens array (MLA) splits the main image into smaller versions (Perwass), introducing an angular component

The light field can yield Sub aperture images, depth, and synthesized 2D images.

Sub aperture images (SAIs) are generally extracted from the light field at one pixel per lens, followed by a 1D interpolation to de-hex the image. The resulting resolution for each SAI is a fraction of the resolution of the center.

Most research relies on the size of the image sensor for adequately sized SAIs, and perform

Keeping the MLA structure intact let's us utilize sampling techniques that allow us to get greater resolution images per light field than SAI sampling.

(remember the Brown paper regarding light fields for robotics) and remember the metric values, and how does this compare with the stereo system? Do we even want to open that up?

The U-Net architecture has good results with 2D low light image enhancement (cite learning to see in the dark).

SAIs and the processed light field are improved according to our metrics, but depth and depth-based results remain degraded

2 Related Work

2.1 Low Light Enhancement

There has not been research done on low light light fields, or light field SAIs.

figures/proc_psnr.pdf

Fig. 1. Would like this to show that our method is the best overall. Need to think of the best way to represent this.

the work focused on 2D image enhancement. What work has been done on light fields has been focused on denoising, and then, largely on the SAIs rather than the raw light field.

2.2 Denoising

Light field denoising is generally carried out on the SAIs.

LFBM5D utilizes the angular data within the SAIs.

3 Method

We combine multiple methods, combining the U-Net, ResNet, GC-Net and light field multi-view stereo by Palmieri.

- *This paper is under review for ICCP 2020 and the PAMI special issue on computational photography. Do not distribute.*

figures/savefig.png

Fig. 2. Side by side comparison showing the enhanced processed image, but depth remaining the same, along with the plots that support this

figures/depth_psnr.pdf

Fig. 3. Would like this plot to show the problem, that while overall quality is improved, clearly it doesn't translate over to depth estimate becoming significantly better. Perhaps a badpix representation would be more suitable

figures/BadPix.pdf

Fig. 4. BadPix for each method as an average of all light field depths. This shows that for each threshold, our method has the fewest number of incorrect depths.

The U-Net architecture provides the overall enhancement of the light field. Patches of the light field are fed through it and compared with a ground truth.

The patches are upsampled to allow for network structure (cite low resolution deep learning paper)

Lenses are extracted from this patch and features are extracted from each via a Siamese ResNet architecture with as many towers as lenses extracted.

The lens features are combined into a cost volume. Similar to GC-Net, the volume is made up of concatenated features and a 3D convolution process regularizes the depth map in place of the traditional stereo regularization pipeline.

The ground truth is based on reference, high light light fields taken of a static scene.

The depth reference is generated by the PlenopticToolbox2.0.

In using the lenslets, we increase the total number of depth maps. For each light field, depth is calculated for each lens, effectively yielding 8700 depth maps per light field.

4 Experiments and Results

We chose to use the metrics seen here as little research has been done regarding the best way to qaulitatively assess a light field, especially given the number of medium niches a light field can occupy. What work there is has yielded inconclusive results. (see study on impact of vis techniques).

5 Conclusion

Look at "Towards a Quality Metric for Dense Light Fields" as a counter to using viewers. It seems to validate the idea that the metrics they introduce line up nicely with the subjective results.

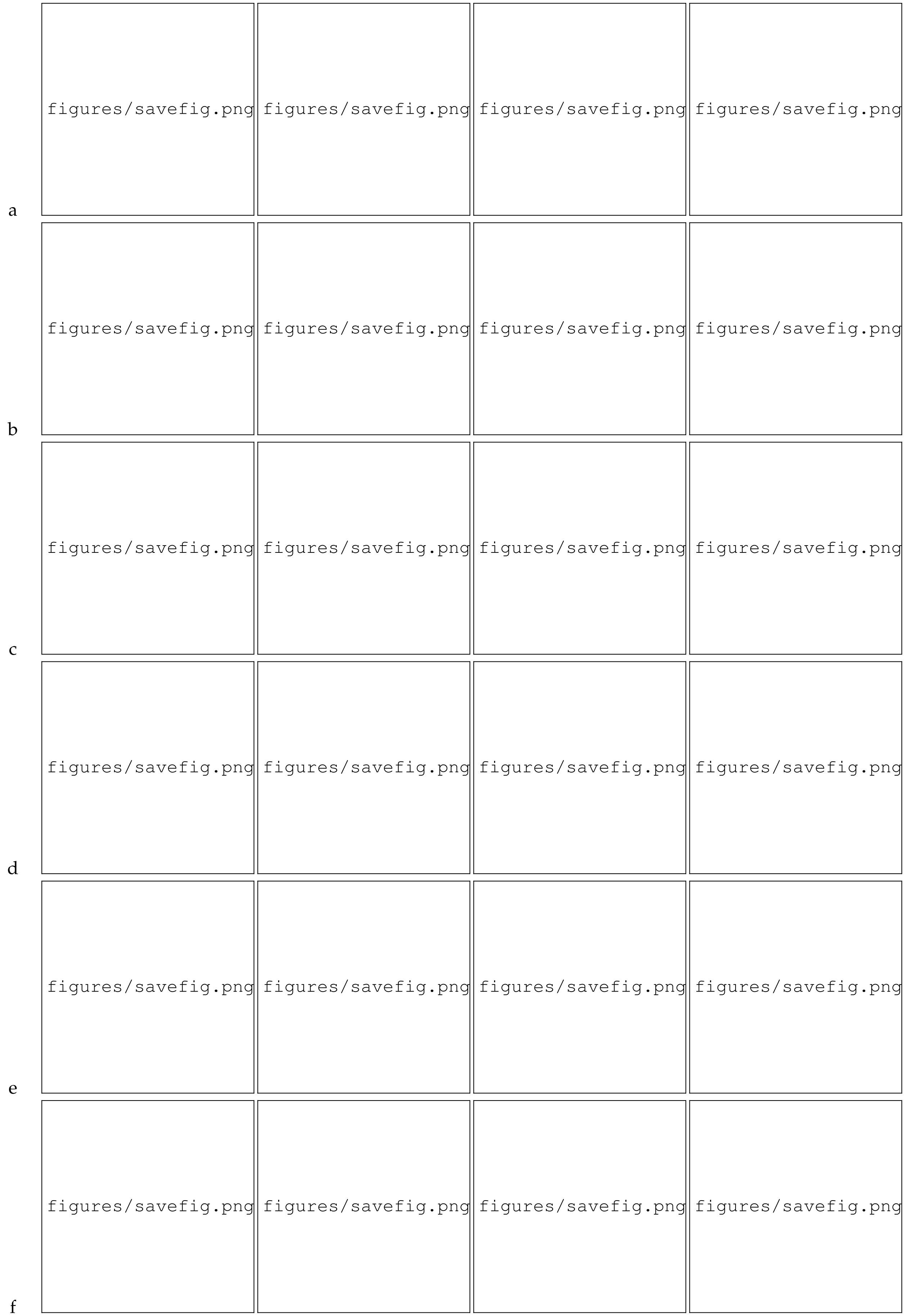

Fig. 5. A comparison of enhancement techniques applied to our test, low-light lightfield dataset.

Fig. 6. Example two-column figure.

Since "Towards a Quality.." is largely based on the quality of viewing (hence the stereo etc comparators), we're interested in more than just viewing, so we can look to epi-net and epi-shift for BadPix (0.7, 0.3, 0.1) metrics