Single Shot HDR Imaging via Compressed Sensing

EE367 Final Project Proposal

Matthew McCready

1 Motivation

Many scenes of interest exhibit a dynamic range beyond the capabilities of typical camera hardware to fully measure. With high dynamic range (HDR) imaging techniques, one can expand the measurement bit depth to fully encapsulate the details of the scene through modifications to hardware and/or software. The classic implementation of HDR imaging is multiple exposure fusion (MEF) where a set of low dynamic range (LDR) images of the same scene are acquired with different exposures and systematically combined to produce an HDR image. However, MEF often requires multiple time consuming long-exposure images for poorly lit details, and also suffers from ghosting and misalignment artifacts should the camera or scene move at all between LDR images.

A better solution would allow for capture of HDR images in a single camera shot. This has been implemented in the past, often by introducing a spatially varying pixel exposure (SVE). In this work we propose to decode a grayscale SVE masked image using compressed sensing to produce multiple exposure images from the single LDR image for input to traditional MEF.

2 Related Work

2.1 Single-shot HDR imaging

Single shot HDR often relies upon the use of images taken with SVE. To implement this one typically uses either a physical neutral density filter mask, spatially varying ISO, or direct spatial variation of exposure. These approaches use a variety of decoding techniques, some depending on the choice of mask. A previous EE 367 final project optimized their SVE encoding for the specific scene to be imaged, however this is not true single-shot HDR as it relies on some prior knowledge of the scene from metering or a previous frame in HDR video (1). Neural networks have been used for hallucination of HDR images from single saturated LDR images, however this is not always reliable to faithfully recreate the true scene. Neural networks have also been used to decode images taken with SVE or other optical filters (2,3). These methods can be end-to-end optimized and perform well but are specific to a chosen mask/filter and as with all neural networks, could be unpredictable when presented with unfamiliar data. Classic single-shot HDR imaging using SVE implements interpolation of unsaturated significant pixels (4), however interpolation can lose details, in particular if resulting sampling is below the Nyquist rate.

2.2 Compressed sensing

It has been shown that compressed sensing outperforms classical interpolation of randomly undersampled signals in regions of rapid change (5). In images this would correspond to regions with sharp edges in the scene. This is possible if a domain exists where the data can be represented sparsely, and the measurement and transformation matrices are incoherent with respect to one another.

3 Project Overview

This work seeks to accomplish single shot HDR imaging by formulating N compressed sensing problems from an image taken with N SVE values. The resulting N LDR images are then treated as a typical MEF input image set to produce the final HDR image. The project will be implemented entirely through simulation, with sets of true MEF input images sampled via the chosen grayscale mask and patchworked together into the “single shot” image.

3.1 Formulation of compressed sensing problem

Given our known SVE mask of *N* unique exposures, we will separate the input image into *N* under sampled images. For a given under sampled image $\tilde{x}$, we will attempt to reconstruct the full image by solving the underdetermined inverse problem:

$$\begin{gathered}\tilde{x} = C\hat{\mathcal{F}}s \\ x = \hat{\mathcal{F}}s\end{gathered} \tag{1}$$

for its Fourier transform s, where $\hat{\mathcal{F}}$ is the discrete Fourier transform matrix, C is our measurement matrix, and x is the full LDR image. Given that the Fourier transform of natural images is known to be sparse, we can solve this problem for the sparsest s by minimizing:

$$\min_{s} \|C\hat{\mathcal{F}}s - \tilde{x}\|_2 + \lambda\|s\|_1 \tag{2}$$

The resulting s can then be inverse Fourier transformed to return an approximation of the fully sampled LDR image. The process is carried out on each of the *N* under sampled images. We then combine the *N* fully sampled approximations through MEF to produce an approximate HDR image.

4 Timeline and Milestones

4.1 Implement compressed sensing problem (*Week 7*)

Week 7 will focus on the implementation of the *N* compressed sensing problems. This will use an entirely stochastic mask. A choice of regularizer will be made, though the current plan is to use the L_1-norm as in equation 2. A solver for the inverse problem will then be chosen and/or implemented (most likely ADMM).

4.2 Tuning of compressed sensing (*Week 8*)

Week 8 will focus on tuning hyper parameters of the compressed sensing problems such as λ in equation 2 and learning rates if a gradient descent method is used. The hyper parameters to be tuned will depend on the implementation from Week 7. Initialization of s will also be contrasted between randomly generated noise and a Fourier transformed zero-order hold or interpolated copy of $\tilde{x}$ to see if calculation efficiency can be sped up.

4.3 Denoising of LDR images and alternate formulation (*Weeks 9/10*)

Weeks 9 and 10 will see the implementation of a denoising method for the recovered compressed sensed LDR images of different exposures. The choice of denoising method will depend on the results of weeks 7 and 8 and the type of resulting noise in the recovered images. It’s possible that denoising the individual LDR images will reintroduce ghosting artifacts under fusion, in which case this could be altered to perform denoising on the final HDR image. In these weeks an investigation could also be made into reformulating the entire problem as a single three-dimensional compressed sensing problem with the third dimension given by the various mask gray levels (exposures) rather than *N* individual two-dimensional problems. The report and poster will also be made during this time.

5 References

1. Wang H, Jiang J, Lyu Z. One Shot High Dynamic Range Photography Based on Hardware-level Tunable Mask. EE367 Comput. Imaging 2018;22.

2. Metzler CA, Ikoma H, Peng Y, Wetzstein G. Deep Optics for Single-Shot High-Dynamic-Range Imaging. Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. 2020:1372–1382 doi: 10.1109/CVPR42600.2020.00145.

3. Martel JNP, Muller LK, Carey SJ, Dudek P, Wetzstein G. Neural Sensors: Learning Pixel Exposures for HDR Imaging and Video Compressive Sensing with Programmable Sensors. IEEE Trans. Pattern Anal. Mach. Intell. 2020;42:1642–1653 doi: 10.1109/TPAMI.2020.2986944.

4. Nayar SK, Mitsunaga T. High dynamic range imaging: spatially varying pixel exposures. In: Proceedings IEEE Conference on Computer Vision and Pattern Recognition. CVPR 2000 (Cat. No.PR00662). Vol. 1. IEEE Comput. Soc; pp. 472–479. doi: 10.1109/CVPR.2000.855857.

5. Usman K, Ramdhani M. Comparison of Classical Interpolation Methods and Compressive Sensing for Missing Data Reconstruction. In: 2019 IEEE International Conference on Signals and Systems (ICSigSys). IEEE; 2019. pp. 29–33. doi: 10.1109/ICSIGSYS.2019.8811057.