

EE 367 Project Proposal

Model-Agnostic Meta-Learning for Single Shot HDR Reconstruction

Anthony Vento Edwin Pan

{avento, edwinpan}@stanford.edu

1 Motivation

Meta-learning is a powerful & novel framework for computing a set of meta-parameters that can quickly be optimized for a specific task given task-specific information. In Low Dynamic Range (LDR) to High Dynamic Range (HDR) conversion, there exists a non-linear mapping between the radiance in a scene and the actual recorded pixel values. While existing deep learning models attempt to learn this nonlinear mapping through extensive datasets, they all rely on the assumption that the nonlinear mapping is fixed despite changes in scene, hardware, etc. Through meta-learning, we can learn a set of meta-parameters that captures the common patterns between separate nonlinear maps. This way, given a few specific examples of LDR images, we can quickly adjust the non-linear mapping to accommodate the specific LDR-to-HDR conversion.

2 Related Work

Conversion from LDR-to-HDR has been a long-standing task. As we saw in lecture and the homework, Debevec and Malik fuse a series of images captured at different exposures to create a HDR image [1]. Later work then tries to pass a single LDR image to create a HDR image (i.e. single shot LDR-to-HDR). Bist et al applies a variable gamma correction to the image depending on the determined light source [2].

Recent research attempts to pass one LDR image via some deep learning pipeline to hallucinate a HDR image. In Eilertsen et al, an input LDR image is passed through a U-Net to output a HDR image [3]. Alternatively, it's possible to utilize optical hardware to redistribute the luminance information. In Metzler et al, an optical encoder parameterized by the PSF and CNN decoder are used to create an LDR-to-HDR pipeline [4] that reconstructs HDR images without hallucination.

3 Objective

Our objective is to train a set of generalizable meta-parameters (θ i.e. weights) that will allow our model to quickly optimize towards a set of task-specific parameters (ϕ_i) when presented with a specific LDR-to-HDR conversion task ($\mathcal{T}_i$). In order to facilitate the meta-learning framework, we choose to interpret each LDR-to-HDR image pair as *separate* tasks as opposed to samples drawn from the same task distribution. Different exposures of the same scene will be treated as samples from the same task distribution. This interpretation is valid because the nonlinear mapping our model will attempt to deconstruct is distinct between separate images. Therefore, we can rest assured that the LDR-to-HDR image pairs exhibit the mutual exclusivity necessary to facilitate meta-learning. The datasets ($\mathcal{T}$) that we plan to use to train our model, as well as the model architectures, are described in more detail below.

3.1 Models

For the task-specific LDR-to-HDR conversion, we plan to follow the approach outlined in Eilertsen et al. and utilize a U-Net architecture to produce the HDR images from LDR inputs. In order to achieve task-generalization, we plan to use the model-agnostic meta-learning (MAML) model proposed by Finn et al [5] as our meta-learning framework. For a visual depiction of the desired outcome, see Figure 1. For each task $\mathcal{T}_i$ within our dataset, we split the samples within $\mathcal{T}_i$ into train ($\mathcal{D}_i^{tr}$) and test ($\mathcal{D}_i^{ts}$). The model objective function, task-specific parameter ϕ_i update step equation, and meta-parameter θ update set equation are written out below.

$$\begin{aligned} \min_{\theta} \sum_{\text{task } i} \mathcal{L}(\theta - \alpha \nabla_{\theta} \mathcal{L}(\theta, \mathcal{D}_i^{tr}), \mathcal{D}_i^{ts}) \\ \phi_i \leftarrow \theta - \alpha \nabla_{\theta} \mathcal{L}(\theta, \mathcal{D}_i^{tr}) \\ \theta \leftarrow \nabla_{\theta} \mathcal{L}(\phi_i, \mathcal{D}_i^{ts}) \end{aligned}$$

Losses are computed between the output of the U-Net and ground-truth. Choosing to use MAML for our meta-learning framework allows us to transplant the U-Net architecture into the inner-loop of the training cycles.

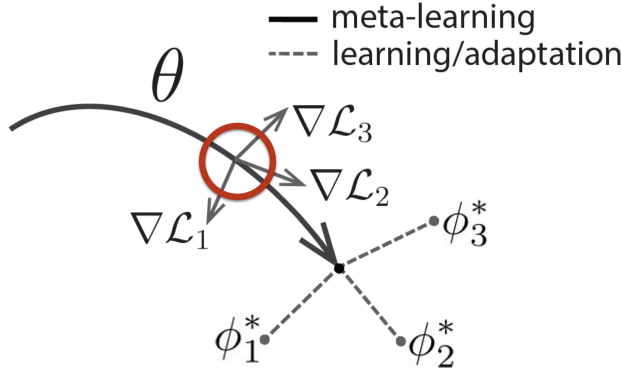


Figure 1: Meta-learning Process [5].

Time permitting, there are other models that might be interesting to explore in this space. For instance, the introduction of a bi-level optimization problem in MAML can lead to instability and difficulty in training [5]. One possible extension could be to use a first-order gradient method like REPTILE [6]. Additionally, we can explore other CNN architectures such as ResNet.

3.2 Datasets

Several sources of LDR and HDR images exist [7–9]. These datasets include 1350, 5000, and 120 LDR-to-HDR images respectively. The 1350 image-pair dataset contains scenes captured at 3 different exposures. The 120 image-pair dataset contains scenes captured at 8 different exposures. We plan to start out by only using one dataset and if need be, extend to use more.

4 Timeline

- 2/17-2/26. Prepare and process data, continue literature review, and begin implementation.
- 2/27-3/5. Complete minimum-viable-product (MVP) and train prototype architecture.
- 3/6-3/12. Tune parameters and extend if time allows. Run comparison models.
- 3/13-3/19. Collect findings, create poster/video, and write report. Mentally prepare for Spring Break.

References

- [1] Paul E Debevec and Jitendra Malik. Recovering high dynamic range radiance maps from photographs. In *ACM SIGGRAPH 2008 classes*, pages 1–10. 2008.
- [2] Cambodge Bist, Rémi Cozot, Gérard Madec, and Xavier Ducloux. Tone expansion using lighting style aesthetics. *Computers & Graphics*, 62:77–86, 2017.
- [3] Gabriel Eilertsen, Joel Kronander, Gyorgy Denes, Rafał Mantiuk, and Jonas Unger. Hdr image reconstruction from a single exposure using deep cnns. *ACM Transactions on Graphics (TOG)*, 36(6), 2017.
- [4] Christopher A Metzler, Hayato Ikoma, Yifan Peng, and Gordon Wetzstein. Deep optics for single-shot high-dynamic-range imaging. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 1375–1385, 2020.
- [5] Chelsea Finn, Pieter Abbeel, and Sergey Levine. Model-agnostic meta-learning for fast adaptation of deep networks. In *International Conference on Machine Learning*, pages 1126–1135. PMLR, 2017.
- [6] Alex Nichol, Joshua Achiam, and John Schulman. On first-order meta-learning algorithms, 2018.
- [7] Hanbyol Jang, Kihun Bang, Jinseong Jang, and Dosik Hwang. Dynamic range expansion using cumulative histogram learning for high dynamic range image generation. *IEEE Access*, 8:38554–38567, 2020.
- [8] Omar Al Shaya, Pengpeng Yang, Rongrong Ni, Yao Zhao, and Alessandro Piva. A new dataset for source identification of high dynamic range images. *Sensors*, 18(11):3801, 2018.
- [9] Hojatollah Yeganeh and Zhou Wang. Objective quality assessment of tone-mapped images. *IEEE Transactions on Image Processing*, 22(2):657–667, 2012.