HDR image reconstruction from a single exposure using deep CNNs, Supplementary material

GABRIEL EILERTSEN*, JOEL KRONANDER*, GYORGY DENES+, RAFAŁ K. MANTIUK+, AND JONAS UNGER*,
*Linköping University, Sweden, +University of Cambridge, UK

1 INTRODUCTION

The supplementary material of the paper "HDR image reconstruction from a single exposure using deep CNNs" consists of the following components:

(1) This document provides a set of additional examples of HDR reconstructions using the proposed method (Section 3). Also, a list with the HDR data resources is given (Section 2).
(2) The supplementary video demonstrates the basic principles of the reconstruction framework, as well as a range of different examples of reconstructions, including video sequences.
(3) The accompanying folder contains all the images from the test set used in the paper, together with simulated LDR images and corresponding reconstructions. The images can be viewed with the provided HTML gallery, which enables easy comparisons of input, reconstruction and ground truth.
(4) The model used for reconstruction is available at: *https://github.com/gabrieleilertsen/hdrcnn*. This repository shares a TensorFlow implementation of the CNN, and together with the provided trained parameters reconstructions can be made from arbitrary LDR input.

2 HDR DATA

The set of HDR images used for training are gathered from the resources listed in Table 1. There are also a few other potential online resources that are available at a high bit depth, *e.g.* the SJUTV[1] and Big Buck Bunny[2] datasets. However, from a visual inspection these contain many saturated pixels and are thus not suitable for training the HDR reconstruction CNN.

3 SUPPLEMENTARY RESULTS

We provide a set of additional predictions in order to complement some of the figures in the paper with more examples:

Blending components: In Section 3.1 in the paper the blending of the CNN output with the input image is described through Equation 1. The components of this blending are illustrated in Figure 2 in the paper. In order to complement this figure with the full image, and with additional examples, Figure 2 in this document shows the different components for two images from the test set. Furthermore, the blending equation is illustrated in the overview of the reconstruction pipeline in Figure 1.

Reconstruction with ground truth: A set of reconstructions are presented in Figure 3. The inputs have been constructed by simulating LDR images from HDR sources, in order to provide reconstruction with ground truth comparison, in the same manner as in Figure 9 in the paper. Around 5% of the pixels are saturated in each input image.

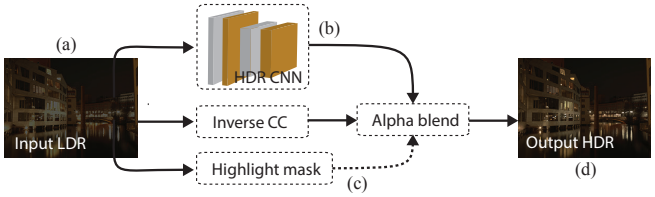


Fig. 1. Highlight reconstruction is made by the HDR CNN. Subsequently, the predicted linear highlights are blended with the linearized input to create the final HDR reconstruction. The labels correspond to the respective sub-figures in Figure 2.

Reconstruction from smart-phone images: The HDR reconstruction also generalizes to more realistic scenarios, as exemplified in Figure 4. This shows a range of different reconstructions using iPhone images as input, in the same manner as in Figure 11 in the paper. In contrast to the artificial LDR images in Figure 3, these images have undergone unknown camera transformation and post-processing. Still, reconstruction can be convincingly performed in a wide range of situations.

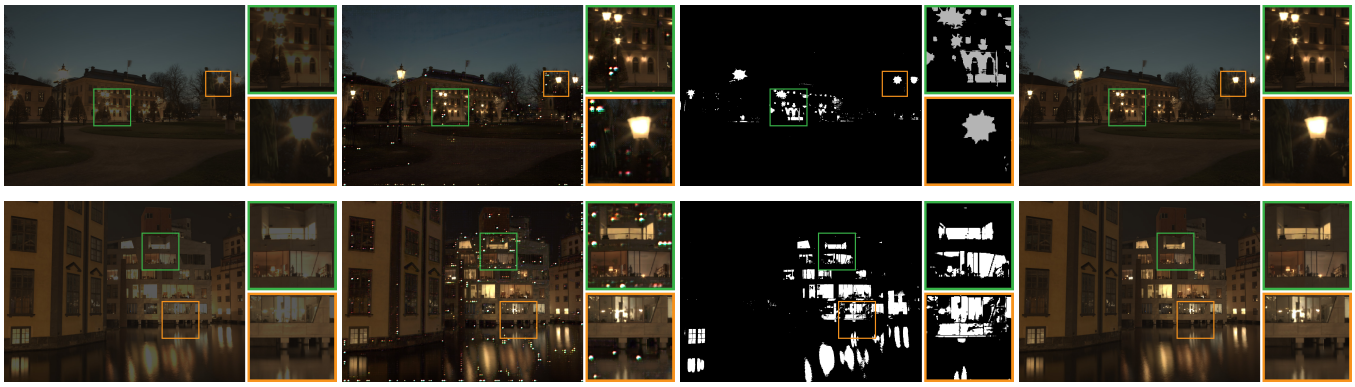
Image based lighting: Figure 5 complements the IBL application example in the paper (Figure 15) with an additional reconstructed panorama. Although large areas are saturated, so that all details cannot be reconstructed convincingly, the recovered illuminance enables a rendering that is close to using the ground truth panorama. This makes for a substantial improvement as compared to existing inverse tone-mapping methods.

Reconstruction with different exposures: The amount of lost information that can be recovered by the CNN is highly content dependent. It depends on the type of saturated region as well as if there is information left in any of the color channels. In order to demonstrate the difference in reconstruction with different amount of lost information, Figure 6 shows predictions with the same scenes that has been simulated to be captured with different exposure times. This is done in the same way as demonstrated in Figure 12 in the paper, and illustrates that in some situations the results are very similar, and in other situations there is a clearly visible difference due to the different amount of information that is available depending on the exposure time.

[1] http://medialab.sjtu.edu.cn/HDR/index.html
[2] https://peach.blender.org

Name	Source	Size
EMPA	http://www.empamedia.ethz.ch/hdrdatabase/index.php	33
HDReye	http://mmspg.epfl.ch/hdr-eye	46
Fairchild	http://rit-mcsl.org/fairchild//HDRPS/HDRthumbs.html	106
Ward	http://www.anyhere.com/gward/hdrenc/pages/originals.html	33
Stanford	http://scarlet.stanford.edu/~brian/hdr/hdr.html	91
MCSL	http://www.cis.rit.edu/research/mcsl2/icam/hdr/rit_hdr/	74
Funt	http://www.cs.sfu.ca/~colour/data/funt_hdr/#DATA	112
Boitard	https://people.irisa.fr/Ronan.Boitard/	7 sequences
MPI	http://resources.mpi-inf.mpg.de/hdr/video/	2 sequences
DML-HDR	http://dml.ece.ubc.ca/data/DML-HDR/	5 sequences
HDR book	Images accompanying the HDR book by Reinhard *et al.* [2005].	327
JPEG-XT	Images used in the evaluation by Mantiuk *et al.* [2016].	174
Stuttgart	https://hdr-2014.hdm-stuttgart.de/	33 sequences
LiU HDRV	http://hdrv.org	10 sequences
Sequences	Miscellaneous sequences	10 sequences
Probes	Miscellaneous lightprobes/panoramas	12
Images	Miscellaneous images	113
		1121 images, 67 sequences

Table 1. List of HDR sources used for creating the training and testing database. Every 10th frame is taken from each sequence, which makes for a total of around 3700 high resolution HDR images. Augmentation of this dataset is described in Appendix A of the paper.



(a) Input LDR (b) CNN output (c) Highlight mask (d) Final HDR reconstruction

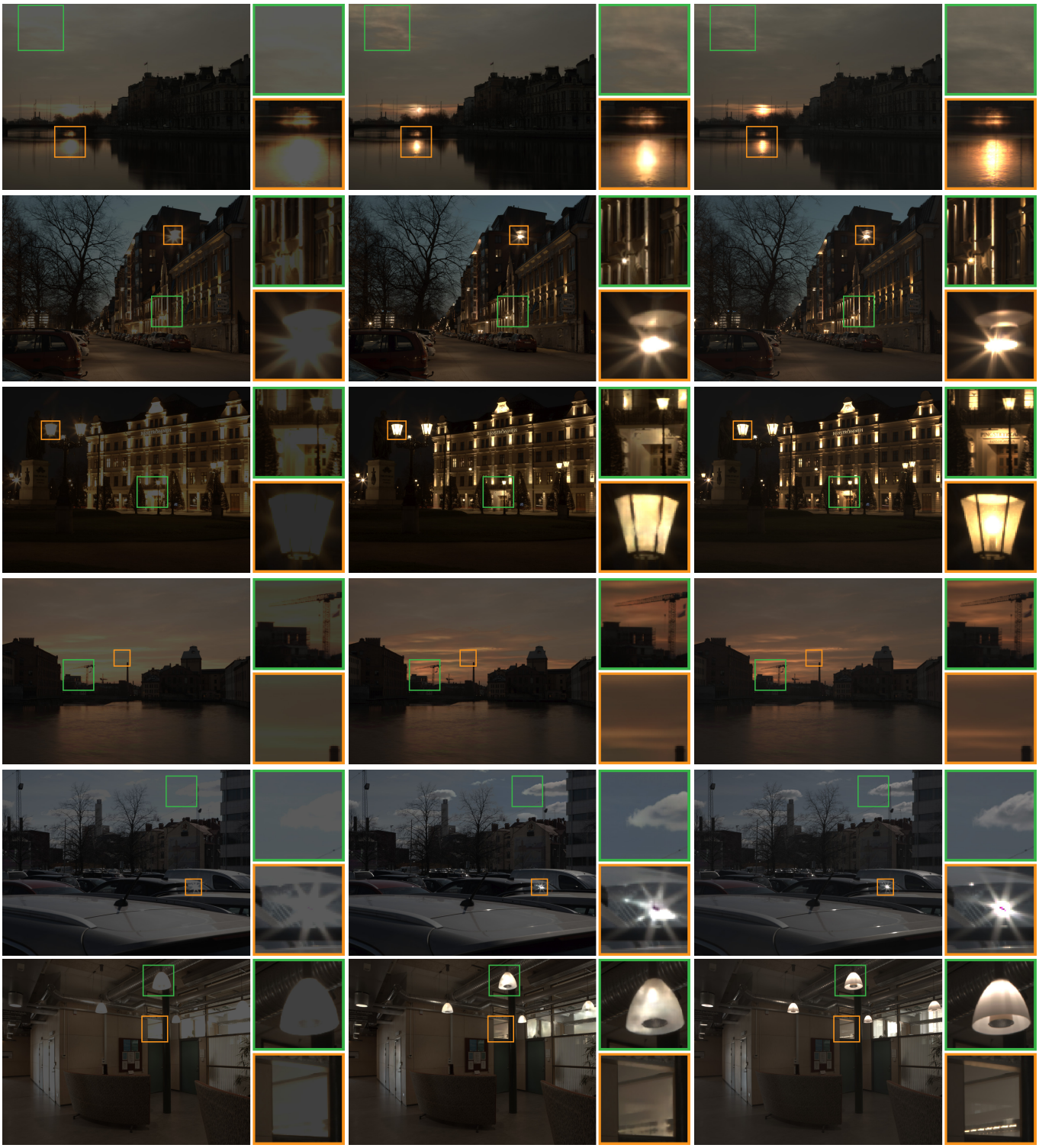
Fig. 2. Examples to complement Figure 2 in the paper. The components of the blending operation, as illustrated here in Figure 1 and detailed in Equation 1 in the paper. Gamma correction has been applied to the images, for display purpose.

REFERENCES

F. Banterle, P. Ledda, K. Debattista, and A. Chalmers. 2008. Expanding Low Dynamic Range Videos for High Dynamic Range Applications. In *Proceedings of the 24th Spring Conference on Computer Graphics (SCCG '08)*. ACM, 33–41.

R. K. Mantiuk, T. Richter, and A. Artusi. 2016. Fine-tuning JPEG-XT compression performance using large-scale objective quality testing. In *2016 IEEE International Conference on Image Processing (ICIP '16)*. 2152–2156.

E. Reinhard, G. Ward, S. Pattanaik, and P. Debevec. 2005. *High Dynamic Range Imaging: Acquisition, Display, and Image-Based Lighting (The Morgan Kaufmann Series in Computer Graphics)*. Morgan Kaufmann Publishers Inc.



(a) Input (b) Reconstruction (c) Ground truth

Fig. 3. Examples to complement Figure 9 in the paper. The input images (left column) have been exposure corrected, followed by camera transformation, quantization and clipping. 5% of the pixels are saturated and contain no information. Visually convincing reconstructions (middle column) can be made in a wide range of situations. The reconstructions correspond well to the ground truth HDR images (right column). The exposures of the images have been reduced to show the differences, and all images have been gamma corrected for display.

(a) Input (b) Reconstruction (c) Input (d) Reconstruction

Fig. 4. Examples to complement Figure 11 in the paper. Predictions on iPhone camera images, demonstrating that plausible reconstructions can be made in a wide range of situations using input images with unknown camera transformations applied.

(a) Panorama (b) IBL

Fig. 5. Example to complement Figure 15 in the paper. IBL using reconstructed highlights. The left column shows the panoramas that are used for IBL in the right column. Rendering with the LDR input gives a dark and undynamic result. The iTMO boosts brightness to alleviate the problems. With our reconstruction, although all details cannot be recovered in the large saturated areas, the estimated luminance enables a visually convincing rendering that is much closer to the ground truth..

(a) 4% saturated pixels (b) 6% saturated pixels (c) 10% saturated pixels

(d) 4% saturated pixels (e) 6% saturated pixels (f) 10% saturated pixels

(g) Ground truth, example 1 (h) Ground truth, example 2

Fig. 6. Examples to complement Figure 12 in the paper. Predictions at different exposure settings of the input. The numbers indicate how large fraction of the total number of pixels are saturated in the input. The images have then been scaled to have the same exposure after clipping has been applied. The results are visually similar in many situations, but more details are available in reconstruction of the shorter exposure images, which allows for better reconstruction.