# Learning a Reinforced Agent for Flexible Exposure Bracketing Selection
## Supplementary Material

Zhouxia Wang [1,2], Jiawei Zhang[1], Mude Lin[1],Jiong Wang[1],Ping Luo [2],Jimmy Ren[1]

[1]SenseTime Research, [2]The University of Hong Kong

{zhouzi1212, zhjw1988, jimmy.sj.ren}@gmail.com, {linmude, wangjiong}@sensetime.com, pluo@cs.hku.hk

## 1. Determination of the Candidate Exposure Bracketing $\mathbf{A}'$

In this paper, an exposure bracketing contains $K$ images and there are $J = 10$ images ($\{\mathbf{z}_0, \mathbf{z}_1, ..., \mathbf{z}_9\}$) captured under different exposures in our dataset. As a result, there are total $N$=$C_J^K$ possible exposure bracketings in a scene and the candidate exposure bracketing is denoted as $\mathbf{A} = \{\mathbf{Y}_0, \mathbf{Y}_1, \cdots, \mathbf{Y}_{N-1}\}$. However, some exposure bracketings, such as $\{\mathbf{z}_7, \mathbf{z}_8, \mathbf{z}_9\}$ which is too dark to capture any details in the dark areas, are almost impossible to yield an acceptable HDR image. We tend to remove these bad candidates and get a smaller set $\mathbf{A}'$ which contains $N'$ ($N' \leq N$) possibilities.

We determine the candidate exposure bracketing $\mathbf{A}'$ in the following four steps. First, we train MEFNet with $K$ images that randomly selected from the candidate exposure bracketing $\mathbf{A}$. Second, for each sample in the training set, we send all the candidate exposure bracketings into the trained MEFNet for generating the corresponding HDR images and calculate the PSNR with the ground truth HDR image. Third, for every candidate exposure bracketing, we count the number of samples that achieve the highest PSNR in that candidate exposure bracketing. Forth, we remove the candidate exposure bracketing whose count is less than $\varepsilon$ and we attain $\mathbf{A}'$ finally.

Our proposed dataset contains night scenes and daytime scenes. However, since there exists a large difference between night data and daytime data in illumination and scene objects, it is hard to well process these two kinds of scenes at the same time with one model. So we train night scenes and daytime scenes separately.

When $K = 3$, $N = C_{10}^3 = 120$. For night scenes, the count of all the candidate exposure bracketings in $\mathbf{A}$ is shown in Figure 1. When $\varepsilon$ is set to 5, we get $\mathbf{A}'$ contains $N' = 36$ elements (shown in the top rows of Table 1).

Obviously, the exposure bracketings that contain long-exposure and short-exposure images at the same time are more likely to get the best HDR image, such as $\{\mathbf{z}_1, \mathbf{z}_3, \mathbf{z}_4\}$ and $\{\mathbf{z}_1, \mathbf{z}_3, \mathbf{z}_7\}$, because they cover a larger dynamic range. However, the exposure bracketings that only contain long-exposure or short-exposure images are rarely to get the best HDR image, such as $\{\mathbf{z}_1, \mathbf{z}_2, \mathbf{z}_3\}$ and $\{\mathbf{z}_7, \mathbf{z}_8, \mathbf{z}_9\}$, since they only focus on a narrow dynamic range. Therefore, it is reasonable to remove the bad candidate exposure bracketings for training a lighter EBSNet.

With the same process, we get $\mathbf{A}'$ contains $N' = 30$ elements for daytime scenes (shown in the bottom rows of Table 1).

| | | | | | | |
|---|---|---|---|---|---|---|
| Night | $\{\mathbf{z}_0, \mathbf{z}_2, \mathbf{z}_4\}$ | $\{\mathbf{z}_0, \mathbf{z}_3, \mathbf{z}_4\}$ | $\{\mathbf{z}_0, \mathbf{z}_3, \mathbf{z}_5\}$ | $\{\mathbf{z}_0, \mathbf{z}_3, \mathbf{z}_7\}$ | $\{\mathbf{z}_0, \mathbf{z}_3, \mathbf{z}_8\}$ | $\{\mathbf{z}_0, \mathbf{z}_3, \mathbf{z}_9\}$ |
| | $\{\mathbf{z}_1, \mathbf{z}_2, \mathbf{z}_4\}$ | $\{\mathbf{z}_1, \mathbf{z}_2, \mathbf{z}_5\}$ | $\{\mathbf{z}_1, \mathbf{z}_2, \mathbf{z}_6\}$ | $\{\mathbf{z}_1, \mathbf{z}_2, \mathbf{z}_7\}$ | $\{\mathbf{z}_1, \mathbf{z}_2, \mathbf{z}_8\}$ | $\{\mathbf{z}_1, \mathbf{z}_2, \mathbf{z}_9\}$ |
| | $\{\mathbf{z}_1, \mathbf{z}_3, \mathbf{z}_4\}$ | $\{\mathbf{z}_1, \mathbf{z}_3, \mathbf{z}_5\}$ | $\{\mathbf{z}_1, \mathbf{z}_3, \mathbf{z}_6\}$ | $\{\mathbf{z}_1, \mathbf{z}_3, \mathbf{z}_7\}$ | $\{\mathbf{z}_1, \mathbf{z}_3, \mathbf{z}_8\}$ | $\{\mathbf{z}_1, \mathbf{z}_3, \mathbf{z}_9\}$ |
| | $\{\mathbf{z}_1, \mathbf{z}_4, \mathbf{z}_5\}$ | $\{\mathbf{z}_1, \mathbf{z}_4, \mathbf{z}_6\}$ | $\{\mathbf{z}_1, \mathbf{z}_4, \mathbf{z}_7\}$ | $\{\mathbf{z}_1, \mathbf{z}_4, \mathbf{z}_8\}$ | $\{\mathbf{z}_1, \mathbf{z}_4, \mathbf{z}_9\}$ | $\{\mathbf{z}_2, \mathbf{z}_3, \mathbf{z}_4\}$ |
| | $\{\mathbf{z}_2, \mathbf{z}_3, \mathbf{z}_5\}$ | $\{\mathbf{z}_2, \mathbf{z}_3, \mathbf{z}_6\}$ | $\{\mathbf{z}_2, \mathbf{z}_3, \mathbf{z}_7\}$ | $\{\mathbf{z}_2, \mathbf{z}_3, \mathbf{z}_8\}$ | $\{\mathbf{z}_2, \mathbf{z}_3, \mathbf{z}_9\}$ | $\{\mathbf{z}_2, \mathbf{z}_4, \mathbf{z}_5\}$ |
| | $\{\mathbf{z}_2, \mathbf{z}_4, \mathbf{z}_5\}$ | $\{\mathbf{z}_2, \mathbf{z}_4, \mathbf{z}_6\}$ | $\{\mathbf{z}_2, \mathbf{z}_4, \mathbf{z}_9\}$ | $\{\mathbf{z}_2, \mathbf{z}_6, \mathbf{z}_7\}$ | $\{\mathbf{z}_2, \mathbf{z}_6, \mathbf{z}_9\}$ | $\{\mathbf{z}_2, \mathbf{z}_7, \mathbf{z}_8\}$ |
| Daytime | $\{\mathbf{z}_0, \mathbf{z}_1, \mathbf{z}_4\}$ | $\{\mathbf{z}_0, \mathbf{z}_1, \mathbf{z}_5\}$ | $\{\mathbf{z}_0, \mathbf{z}_1, \mathbf{z}_6\}$ | $\{\mathbf{z}_0, \mathbf{z}_1, \mathbf{z}_7\}$ | $\{\mathbf{z}_0, \mathbf{z}_1, \mathbf{z}_8\}$ | $\{\mathbf{z}_0, \mathbf{z}_2, \mathbf{z}_4\}$ |
| | $\{\mathbf{z}_0, \mathbf{z}_2, \mathbf{z}_5\}$ | $\{\mathbf{z}_0, \mathbf{z}_2, \mathbf{z}_6\}$ | $\{\mathbf{z}_0, \mathbf{z}_2, \mathbf{z}_7\}$ | $\{\mathbf{z}_0, \mathbf{z}_2, \mathbf{z}_8\}$ | $\{\mathbf{z}_0, \mathbf{z}_2, \mathbf{z}_9\}$ | $\{\mathbf{z}_0, \mathbf{z}_3, \mathbf{z}_5\}$ |
| | $\{\mathbf{z}_0, \mathbf{z}_3, \mathbf{z}_6\}$ | $\{\mathbf{z}_0, \mathbf{z}_3, \mathbf{z}_7\}$ | $\{\mathbf{z}_0, \mathbf{z}_3, \mathbf{z}_8\}$ | $\{\mathbf{z}_0, \mathbf{z}_3, \mathbf{z}_9\}$ | $\{\mathbf{z}_0, \mathbf{z}_4, \mathbf{z}_6\}$ | $\{\mathbf{z}_0, \mathbf{z}_4, \mathbf{z}_8\}$ |
| | $\{\mathbf{z}_0, \mathbf{z}_4, \mathbf{z}_9\}$ | $\{\mathbf{z}_0, \mathbf{z}_5, \mathbf{z}_6\}$ | $\{\mathbf{z}_0, \mathbf{z}_5, \mathbf{z}_7\}$ | $\{\mathbf{z}_0, \mathbf{z}_5, \mathbf{z}_9\}$ | $\{\mathbf{z}_0, \mathbf{z}_6, \mathbf{z}_9\}$ | $\{\mathbf{z}_0, \mathbf{z}_8, \mathbf{z}_9\}$ |
| | $\{\mathbf{z}_1, \mathbf{z}_2, \mathbf{z}_6\}$ | $\{\mathbf{z}_1, \mathbf{z}_2, \mathbf{z}_7\}$ | $\{\mathbf{z}_1, \mathbf{z}_2, \mathbf{z}_9\}$ | $\{\mathbf{z}_1, \mathbf{z}_3, \mathbf{z}_7\}$ | $\{\mathbf{z}_1, \mathbf{z}_3, \mathbf{z}_8\}$ | $\{\mathbf{z}_1, \mathbf{z}_3, \mathbf{z}_9\}$ |

Table 1: The elements of $\mathbf{A}'$ while $K = 3$ for night scenes and daytime scenes. They tend to contain long-exposure and short-exposure images at the same time which may lead to a larger dynamic range.

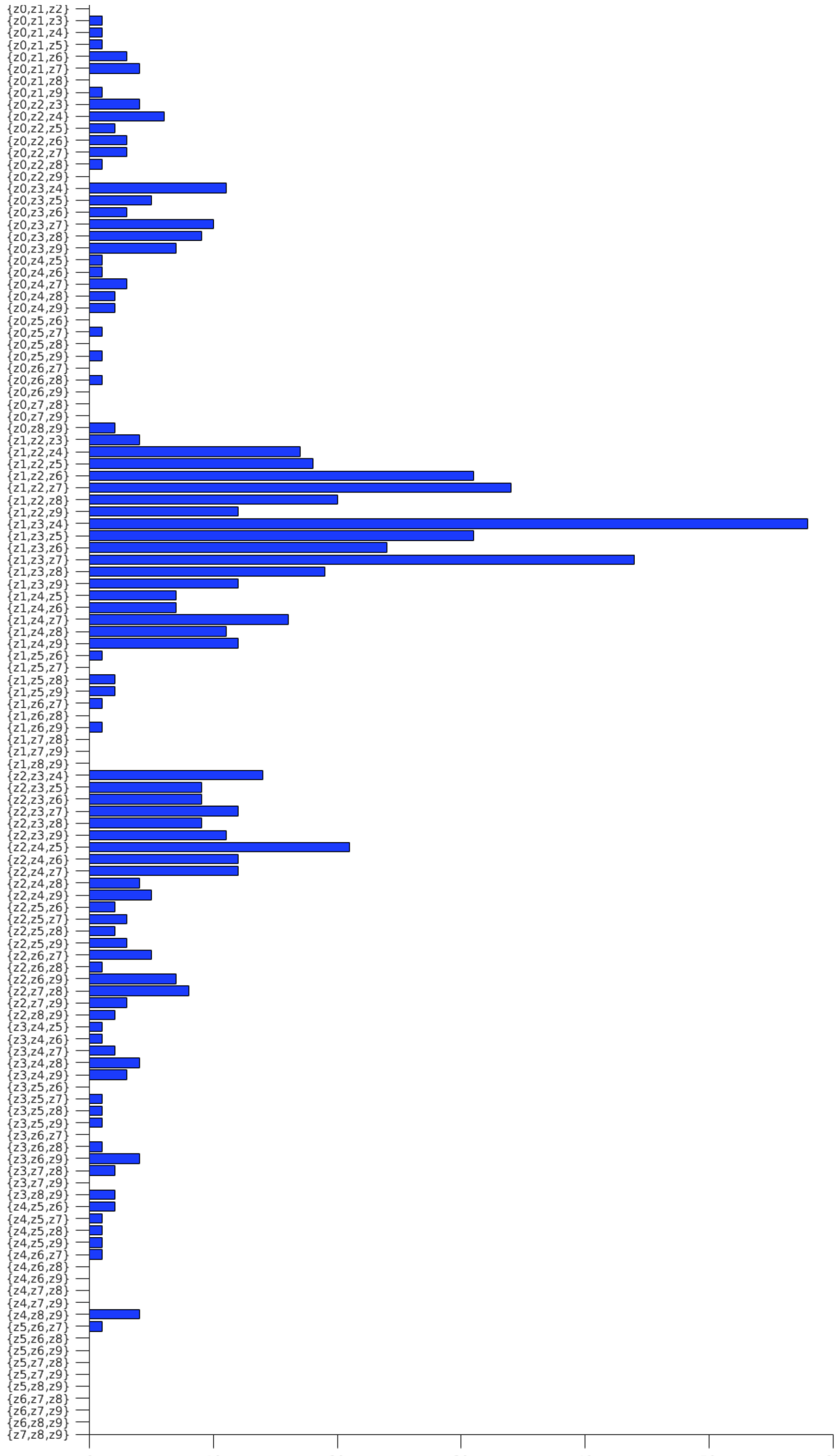


Figure 1: The count of all the candidate exposure bracketings in $\mathbf{A}$ when $K = 3$. The vertical ordinate presents all the $N = C_{10}^{3} = 120$ candidate exposure bracketings while the horizontal ordinate presents the count of each candidate exposure bracketing. It is obvious that most of the samples focus on the exposure bracketings that contain long-exposure and short-exposure images at the same time since they can cover a larger dynamic range.