

# 数据集说明文档

## ● 数据集说明:

我们采用开源 GoPro 数据集的制作方法以及我们的基于三维重建的方法制作数据集。数据集由 66 个序列共 10000 余对模糊、清晰图像组成。

基于 GoPro 方法制作的数据有 30 个序列共 3000 对 1280x720 分辨率的模糊和清晰图像，分 normal\_light(正常光照)及 low\_light(低光照)场景下采集的数据，命名规则为“视频序列号\_合成一张模糊图所用的帧数”。

基于三维重建方法制作的数据有 36 个序列共 7017 对 1000x720 分辨率的模糊和清晰图像，分室内白天(day)和室内夜晚(night)场景下采集的数据，命名规则为：“sharp”文件中为清晰图像，2~10 数字命名的文件夹中为不同模糊程度的模糊图像，模糊、清晰图像被命名为相同的名字。

## ● 数据集下载:

数据集通过百度网盘下载，两种方法生成的数据分别保存在:

链接: <https://pan.baidu.com/s/111YzsF-KDu3FdUt1oZUIYA> 提取码: tszp

链接: <https://pan.baidu.com/s/1VEpwlIKPjqKUTfFoEuIEog> 提取码: x6lh

## ● 数据集制作方法说明:

我们采用开源 GoPro 数据集的制作方法以及基于三维重建的方法制作数据集。

### ■ 模糊图像生成原理:

相机传感器在曝光时间内接收到了不同颜色的光，从而累积产生了图像模糊。其过程可建模如下:

$$B = g\left(\frac{1}{T} \int_{t=0}^T S(t) dt\right)$$

其中 $T$ 表示曝光时间， $S(t)$ 表示在 $t$ 时刻的传感器接收到的光度值(颜色)。 $g$ 是非线性 CRF(相机响应函数)，它将积分的电信号映射到观察到的图像 $\hat{S}(t)$ 中，使得 $\hat{S}(t) = g(S(t))$ 或 $\hat{S}(i) = g(S(i))$ 。在这里，我们不考虑积分信号通过非线性 CRF(相机响应函数)转换为像素值的过程。于是我们可以通过累加相机曝光时间内的潜在图像(如来自高速相机的视频帧)来近似该过程:

$$B \cong \frac{1}{M} \sum_{i=0}^{M-1} S[i],$$

其中 $M$ ， $S[i]$ 分别是在曝光时间内捕获的采样帧数和第 $i$ 个清晰帧信号。

### ■ 基于 GoPro 方法[1]的制作步骤:

### 1. 使用 GOPRO8 Hero Black 相机采集视频

采集视频中不重复拍摄同一场景，保证拍摄目标多样化。

**拍摄正常光照情况下：**我们使用 GOPRO 相机拍摄 240 fps 的视频。

**拍摄低光照情况下：**由于 GOPRO 相机拍摄 240fps 视频时进光量少，加上拍摄暗光时噪点大，导致采集到的视频效果不佳，拍摄目标不够清晰。因此在低光照数据集的采集中，我们采集 120fps 的视频，通过帧插值算法[2]，对 120fps 的视频进行插帧，生成高帧率视频。

### 2. 用 Adobe Premiere 对视频转换帧率、剪辑、去鱼眼处理。

### 3. 用剪好的视频合成模糊图。

**模糊图：**通过平均不同数量的连续帧来合成不同模糊强度的模糊图，例如，平均 15 帧模拟以 1/16 快门速度拍摄的图片。同一序列（对应一个视频）采用相同的合成帧数来生成模糊图，由于视频不是匀速运动拍摄，因此，对于同一拍摄场景能够产生不同模糊强度的模糊图。此外，根据不同的视频序列移动速度快慢来决定该序列模糊图的合成帧数，可以保证对于不同场景都能生成由小到大不同模糊程度的模糊图。

**清晰图：**将用于合成一张模糊图的中间帧作为该模糊图对应的清晰图。

### 4. 最终合成 3000 对 1280x720 分辨率的模糊和清晰图像，包括 15 个正常光照场景采集的序列和 15 个暗光场景采集的序列。

## ■ 基于 Zhu et al. 方法[3]的制作步骤：

### 1. 获取带彩色及深度图像的慢速相机运动视频

使用深度传感器采集对齐的彩色及深度图像序列，相机为慢速运动。受限于深度相机的采集条件，采集场景为室内，帧率为 30fps。该方法不区分正常光和暗光数据的制作。

### 2. 相机定位及三维场景模型重建

将事先对齐的彩色及深度图像序列输入三维重建系统，通过同时建图及相机运动估计建立视频所拍摄到的场景的三维几何模型，并输出视频中相机的运动轨迹。

### 3. 采样快速运动时的清晰图像序列

通过参数 $\lambda$ 控制采样，如慢速运动视频为 1000 帧，当 $\lambda = 5$ 时，对应的 $\lambda$ 倍速度的快速运动序列应为 200 帧左右，我们通过间隔采样从慢速视频中获取对应的 200 张清

晰图像。

#### 4. 模糊图合成

对于清晰图像对应的模糊图像，我们设定曝光时间为 20ms，清晰图像为其曝光时间中间时刻相机所在位置观察到的图像，并通过相机运动轨迹、场景三维模型以及原始彩色清晰图像序列合成曝光时间内的潜在清晰图像，合成数量为 $6 * \lambda + 1$ ，图像像素的颜色值通过投影的方式从清晰图像序列中获取。最后，通过平均潜在清晰图像生成模糊图。针对一个慢速运动视频，我们控制 $\lambda$ 生成其对应不同运动速度的模糊视频（ $\lambda$ 取 2~10）。

#### 参考文献：

- [1] Seungjun Nah, Tae Hyun Kim, and Kyoung Mu Lee. Deep multi-scale convolutional neural network for dynamic scene deblurring. In Proc. IEEE Conf. Comp. Vis. Patt. Recogn., pages 257 – 265, 2017.
- [2] Simon Niklaus, Long Mai, and Feng Liu. Video frame interpolation via adaptive separable convolution. In Proceedings of the IEEE International Conference on Computer Vision, pages 261–270, 2017.
- [3] Zunjie Zhu, Feng Xu, Mingzhu Li, Zheng Wang, and Chenggang Yan. Challenges from Fast Camera Motion and Image Blur: Dataset and Evaluation. RLQ-TOD Workshop on European Conference on Computer Vision, 2020.