# 基于图像传感器物理噪声模型和扩散模型的低光/极低光RAW域图像增强方法

## 技术领域

本发明属于图像处理领域，具体涉及一种通过扩散模型、针对低光/极低光RAW域图像的增强方法。

## 背景技术

在低光/极低光等光照不充分的开放场景下，在基于光电转换原理基础上制作的现有CMOS或CCD图像传感器，都会由于“有效”入射光能量不足，导致其最终采集的图像与视频的RAW数据中包含了大量与后续计算机视觉任务所需的“真实信息”相关或无关的“噪声”信号，即图像/视频数据的信噪比（Signal Noise Ratio,SNR）较低，直接限制了后续计算机视觉任务的性能，降低了观察者的感受。同时，经过图像信号处理器(Image Signal Processor,ISP)将RAW域转换为sRGB、YUV或RGB等色彩空间域后，色彩空间域图像含有的原始噪声变得极为复杂，基本无法建模。现有ISP采用传统算法来实现RAW域图像增强，包含去噪和亮度矫正两个步骤。其中去噪包括空间域上的均值滤波、中值滤波、高斯滤波、双边滤波、NLM等；变换域上的低通滤波、高通滤波、带通滤波、基于小波系数阈值、基于系数相关性、DCT(离散余弦)变换等；时域上的多帧匹配、动静判断、前后景判断等。亮度矫正包含黑电平矫正、白平衡、Gamma矫正等。虽然通过现有ISP可以实现部分RAW域图像增强，但对于低光/极低光RAW域图像增强效果并不理想。

在各项计算机视觉任务中，深度学习技术展现了强大的拟合能力和泛化性能，其在低光/极低光RAW域图像增强上也获得了较好的效果，且相较于传统算法有着更好的泛化性。目前基于深度学习的低光/极低光RAW域增强主要采用基于数据驱动的方法，该方法主要采用深度卷积神经网络（Deep Convolutional Neural Networks, DCNN）对正常光-低光/极低光RAW域图像对进行映射关系学习，从而建立低光/极低光RAW域图像与正常光RAW域图像的拟合或映射关系，通过最小化像素损失函数（如L1、L2等计算数字图像像素间的距离的损失函数）进行模型参数优化，最终利用模型进行前向推理实现低光/极低光RAW域图像的增强功能。基于数据驱动的低光/极低光RAW域图像增强方法虽然取得了较好的增强效果，但需要采集大量的正常光-低光/极低光RAW域图像对，需要消耗大量的人力、物力，从而在现实应用中受到极大限制。

扩散模型（diffusion model）依赖较强的数学推理、具有较强的可解释性。扩散模型包含两个过程：前向过程和反向过程。前向过程对输入数据不断地加高斯噪声最终生成与输入数据相同维度的标准高斯噪声，反向过程训练深度神经网络模型负责将标准高斯噪声恢复回原始的输入数据。扩散模型在多项计算机视觉、自然语言处理等领域任务中展现出较好的性能。Cold Diffusion[1]将扩散模型的前向过程进行进一步推广，使得前向过程可对输入数据执行不同类型的降质，从而将扩散模型应用于去噪、去模糊、修复等不同的底层计算机视觉任务。

图像传感器的本质是一种测量仪器，一定会带有测量误差与测量噪声。基于光电转换原理为基础制造的CMOS或CCD图像传感器，其在低光/极低光等光照不足的情况下出现的SNR降低，其本质是因为含有“真实信息”的“有效入射光”能量不足，导致测量噪声过大。而由于噪声的随机性，使得图像增强这一过程本质是一个“一对多”的求解过程。而基于正常光-低光/极低光图像对进行训练的基于数据驱动的深度学习方法，是一种“一对一”的学习过程，从本质上就不适合这种“一对多”的求解过程；而扩散模型的前向过程对输入数据进行多步降质，每步所执行降质加入的噪声都是服从某个分布的随机噪声，因此获取的降质数据是随机的，将该随机数据输入反向过程的深度神经网络模型，通过最小化模型输出与目标数据间的像素损失函数对网络参数进行优化，从本质上适用于“一对多”的任务需求。然而，目前基于扩散模型的低光/极低光图像增强方法并未考虑低光/极低光图像形成的实际物理成因，因而在实际应用中存在泛化能力不足的问题，导致增强方法性能下降甚至失效的情况。

## 发明创造目的

为提高低光/极低光RAW域图像增强的泛化性问题，拟在Cold Diffusion扩散模型的基础上引入低光/极低光RAW域图像的实际物理成因，将正常光RAW域图像到低光/极低光RAW域图像的转换变为一个基于真实物理噪声模型和扩散模型的生成过程，继而实际应用中获得低光/极低光RAW域图像到正常光RAW域图像的重建，在保证性能的前提下，不仅实现了整个网络的可解释性，也大幅度提高了整个网络模型对低光/极低光RAW域图像增强的泛化性。

## 技术方案

本发明设计了一种基于图像传感器物理噪声模型和扩散模型的低光/极低光RAW域图像增强方法，该方法基于标定的图像传感器物理噪声模型构建扩散模型的前向过程，将正常光RAW域图像逐步降质为低光/极低光RAW域图像，并将前向过程任意步降质图像与前向步数作为深度神经网络模型的输入预测对应的正常光RAW域图像，最终在不同典型条件下大幅提高低光/极低光RAW域图像增强性能和泛化性。

基于图像传感器物理噪声模型和扩散模型的低光/极低光RAW域图像增强方法技术方案主要分为两个部分：①基于物理噪声模型和扩散模型的网络设计;②基于物理噪声模型和扩散模型的训练与推理。

### 基于物理噪声模型和扩散模型的网络设计

为保证基于扩散模型的低光/极低光RAW域图像增强网络可引入低光/极低光RAW域图像真实物理噪声及相应降质过程，由此设计了如下基于物理噪声模型和扩散模型的网络，该网络包括前向降质过程和反向逆降质过程，其中前向降质过程根据低光/极低光RAW域图像的真实物理建模过程将正常光RAW域图像逐步降质为低光/极低光RAW域图像，反向逆降质过程通过深度神经网络实现，将输入的前向降质过程中任意步降质图像和对应前向降质步数转变为相应的正常光RAW域图像。

#### 前向降质过程

基于低光/极低光RAW域图像的真实物理建模过程涉及扩散模型的前向降质过程，该前向降质过程对输入的正常光RAW域图像进行T步前向降质：第一步除以低光因子，后续T-1基于标定的物理噪声参数及分布添加不同的T-1种噪声，且该T-1种噪声应按照真实的物理建模中噪声存在的顺序进行逐步加入。基于真实物理建模过程的前向降质过程可以模拟低光/极低光RAW域图像产生的真实过程，从而使得扩散模型的反向逆降质过程能够更好地从每一步降质图像中还原出相应的正常光RAW域图像。

#### 反向逆降质过程

如前所述，前向降质过程对正常光RAW域图像进行T步降质，对于一张正常光RAW域图像，通过T步前向降质过程，可产生相应的T张降质图像，反向逆降质过程的作用是通过输入前向降质过程中产生的任意步降质图像和相应的前向降质步数重建出原始的正常光RAW域图像，即能够对前向降质过程的前t（t=1,...,T）步降质进行逆操作。

扩散模型的反向逆降质过程通过训练深度神经网络实现，其中深度神经网络包含两个输入：前向降质过程中任意步降质图像和相应的前向降质步数t，输出为重建的正常光RAW域图像，因此此深度神经网络的输出与输入应具有相同的维度。方向逆降质过程的深度神经网络包含步数编码、通道转换、信息融合、特征提取和重建五个模块。

##### 步数编码模块

该模块将输入的标量前向降质步数t编码为一维向量，从而能够通过特征融合模块将前向降质步数信息与降质图像信息进行融合。

##### 通道转换模块

该模块通过一个卷积层对输入的降质图像进行通道转换，在保持长宽不变的情况下，将通道数转变为步数编码模块输出的一维向量的维度，从而能够通过信息融合模块对前向降质步数信息与降质图像信息进行融合。

##### 信息融合模块

该模块能够将前向降质步数t和降质图像进行信息融合，获取前向降质步数t和降质图像的融合特征，从而使得特征提取模块能够对前向降质步数t和降质图像的融合特征提取出高级融合特征。

##### 特征提取模块

该模块对信息融合模块输出的融合特征通过深层卷积网络进行高级融合特征提取，进一步获取融合特征的高级融合特征，进而作为重建模块的输入，重建出相应的正常光RAW域图像。

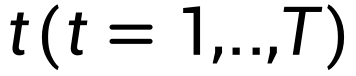
##### 重建模块

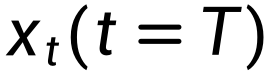
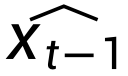
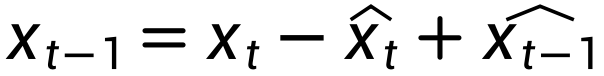
该模块将特征提取层提取的高级融合特征作为输入，重建出相应的正常光RAW域图像。

#### 损失函数的设计

为使得重建的正常光RAW域图像和真实的正常光RAW域图像尽可能相近，通过最小化重建的正常光RAW域图像与真实的正常光RAW域图像间的像素损失实现对反向逆降质过程的深度神经网络优化。

### 基于物理噪声模型和扩散模型的训练与推理

网络训练阶段，首先对采集的真实正常光RAW域图像执行前向降质过程，基于物理噪声建模过程及标定出的相应参数随机执行步前向降质，从而获得对应的降质图像latexmath，将降质图像latexmath和相应的前向降质步数t输入到反向逆降质过程的深度神经网络中，通过最小化深度神经网络重建的正常光图像和真实正常光图像间的像素损失对网络参数进行优化更新。

网络推理阶段，真实的低光/极低光RAW域图像作为执行T步降质降质图像，执行以下操作：步骤1：将降质图像latexmath和前向降质步数t输入反向逆降质过程深度神经网络得到重建的正常光RAW图像；步骤2：对执行t-1步和t步前向降质分别得到降质图像和；步骤3：利用公式计算t-1步推理图像；步骤4：更新latexmath，latexmath；重复步骤1至步骤4过程，直至步骤4执行后latexmath结束，此时步骤3获取的推理图像latexmath即为推理出的正常光RAW域图像。

## 有益效果

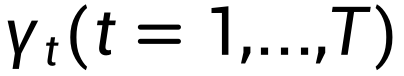
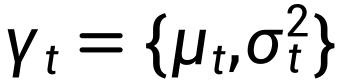
本发明基于低光/极低光RAW域真实物理噪声模型作为扩散模型的前向降质过程，扩散模型的反向逆降质过程输入任意步降质图像和相应的前向降质部署重建相应的正常光RAW域图像，从而实现低光/极低光RAW域图像增强。

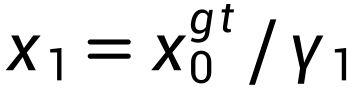
基于真实物理噪声模型和扩散模型的低光/极低光RAW域图像增强数学理论清晰、可解释性强，因此对低光/极低光RAW域图像具有突出的增强性能以及良好的鲁棒性。

## 具体实施方法

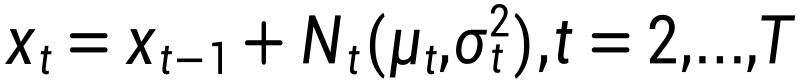
### 基于物理噪声模型和扩散模型的网络设计

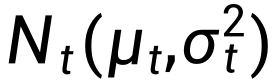
基于物理噪声模型的扩散模型包含前向降质过程和反向逆降质过程，其中前向降质过程根据低光/极低光RAW域图像的真实物理建模过程将正常光RAW域图像逐步降质为低光/极低光RAW域图像，真实物理建模过程中所涉及的降质参数由真实的图像传感器物理噪声模型标定得到。

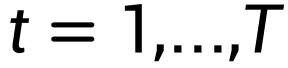
真实物理建模过程通过T步将正常光RAW域图像建模为低光/极低光RAW域图像，其中每步所标定的参数为，其中latexmath表示低光系数，表示T-1种服从分布的噪声均值latexmath和方差。对于一张正常光RAW域图像，前向降质过程包含T步，其中第一步表示如下：



降质过程的第2至T步表示如下：



其中表示与latexmath具有相同维度的服从分布均值为latexmath和为方差的随机噪声。

基于物理噪声模型的扩散模型的反向逆降质过程通过训练深度神经网络实现，其中深度神经网络包含两个输入：前向降质过程中的任意步降质图像latexmath和对应的前向降质步数t，其中，通过深度神经网络重建出相应的正常光RAW域图像。反向逆降质过程的深度神经网络包含步数编码、通道转换、信息融合、特征提取和重建五个模块。

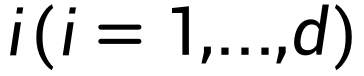
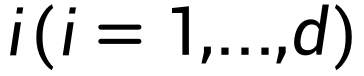
##### 步数编码模块

该模块将标量步数编码为一维向量，采用transformer[2]对位置嵌入的编码方式，通过sin和cos的线性变换编码前向降质步数latexmath，从而将每个标量前向降质步数t编码为一个维度为latexmath的一维向量，其中维度latexmath一般取128或256。

##### 通道转换模块

该模块是将降质图像转换为通道数与步数编码模块输出的一维向量维数相同的浅层特征图。通过一个卷积层对输入的降质图像latexmath进行通道转换，卷积核大小为latexmath，卷积步长为1，卷积核数量为latexmath。因此通过该通道转换模块，输出浅层特征图长宽与输入降质图像latexmath一致，通道数变为d。

##### 信息融合模块

该模块对降质图像和前向降质步数信息进行融合。对于通道转换模块输出的浅层特征图，对第个通道对应的二维特征图的每个像素值加上步数编码模块输出的一维向量的第个元素，进而将降质图像与前向降质步数进行信息融合，生成融合特征。

##### 特征提取模块

该模块对信息融合模块得到的融合特征进行深层特征提取，得到融合特征的高级融合特征。该模块多采用基于UNet结构的深层卷积网络实现，从而对融合特征提取高级融合特征。

##### 重建模块

该模块将输入的高级融合特征重建为正常光RAW域图像。通过一层卷积层实现，卷积核大小为latexmath，卷积步长为1，卷积核数量与输入降质图像的通道数相同，从而通过输入的高级融合特征重建出相应的正常光RAW域图像。

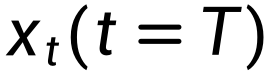
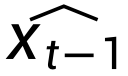
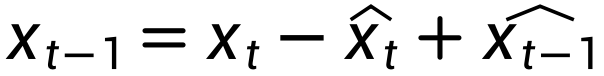
##### 损失函数的设计

为使得重建的正常光RAW域图像和真实的正常光RAW域图像尽可能相近，通过最小化重建的正常光RAW域图像和真实的正常光RAW域图像间的像素损失来优化反向逆降质过程深度神经网络，和间像素损失可选择L1、L2或MSE损失中的任意一种实现。

### 基于物理噪声模型的扩散网络模型的训练与推理

通过数据采集和标定可得到正常光RAW域图像、以及其对应相机设置下的物理降质过程和相应的参数。

网络训练阶段，首先对采集的真实正常光RAW域图像，基于物理降质建模过程及标定出的参数随机执行t步前向降质，从而获得对应的降质图像latexmath,将降质图像latexmath和降质步数t输入到反向逆降质过程的深度神经网络中，通过最小化深度神经网络重建的正常光RAW域图像和相应的真实正常光RAW域图像间的像素损失对网络参数进行优化更新。

网络推理阶段，真实的低光/极低光RAW域图像作为执行T步降质降质图像，执行以下操作：步骤1：将降质图像latexmath和前向降质步数t输入反向逆降质过程深度神经网络得到重建的正常光RAW图像；步骤2：对执行t-1步和t步前向降质分别得到降质图像和；步骤3：利用公式计算t-1步推理图像；步骤4：更新latexmath，latexmath；重复步骤1至步骤4过程，直至步骤4执行后latexmath结束，此时步骤3获取的推理图像latexmath即为推理出的正常光RAW域图像。，其算法流程如下所示：

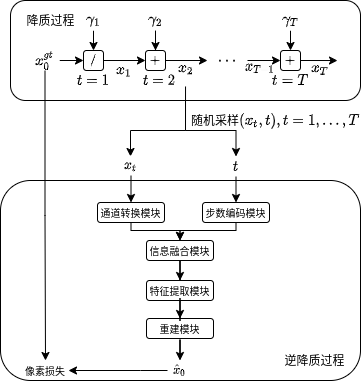
|  |
| --- |
| 输入：真实低光/极低光RAW域图像latexmath，初始化latexmath、前向降质步数latexmath; |
| 如果latexmath执行以下步骤1-4过程： |
| 步骤1：将latexmath输入反向逆降质过程深度神经网络重建出正常光RAW域图像latexmath; |
| 步骤2： 对latexmath分别执行前向降质过程t-1步和t步分别得到降质图像latexmath和latexmath; |
| 步骤3：利用公式latexmath计算t-1步的推理图像； |
| 步骤4：更新latexmath，latexmath; |
| 循环结束 |
| 推理的正常光RAW域图像当t取1时步骤3计算获得的latexmath。 |

## 总结

提出了一种基于物理降质建模过程的扩散模型，通过将扩散网络模型的前向降质过程应用低光/极低光RAW域图像真实物理降质过程，并通过反向逆降质过程的深度神经网络对任意步降质图像和前向降质步数重建出相应的正常光RAW域图像，使得模型具有良好的可解释性，并能够在给定低光/极低光RAW域图像的情况下得到正常光RAW域图像。

## 附图

### 附图1



## 参考文献

1. Bansal A, Borgnia E, Chu H M, et al. Cold diffusion: Inverting arbitrary image transforms without noise[J]. arXiv preprint arXiv:2208.09392, 2022.
2. Vaswani A , Shazeer N , Parmar N , et al. Attention Is All You Need[C]// arXiv. arXiv, 2017.