1. 在上次汇报中，我简单介绍了Diffusion model，也就是扩散模型，它的核心思想是先对图像进行逐步加噪，再逐步去噪，从而生成我们想要的目标图像。本次我要汇报的论文是Diffusion Depth，它将扩散模型与深度估计结合了起来，并且在估计结果上实现了不错的效果。
2. 这里呢是扩散模型的示意图，就是逐步加噪到全损的图像，再逐步去噪生成目标图像
3. 而Diffusion Depth就运用了扩散模型的思想，它把深度估计的过程转为了这样一个从随机深度图中不断精细化的过程。在这个过程中，它会根据我们传入的图像C作为视觉引导条件，从而进行逐步细化（就像给你一个参照物让你照着样子画画）。
4. Diffusion Depth的核心思想就是把降噪的过程作为深度细化的过程，然后这些就是它的数学表达式；这里要提到一个叫DDIM策略的东西， DDIM策略就是将去噪过程中的方差近似为0，使得扩散过程不是纯随机化，而是有迹可循的；
5. 这一页是该项目的整体架构，我们来分块看
6. 首先是对输入的图像进行预处理，将其分割、展平后进行位置编码，这个处理策略和Vit很相似
7. 将处理后的Patch输入进swin-transformer内进行多尺度特征提取，再将提取到的特征进行聚合，作为视觉引导条件传入潜空间
8. 在潜空间内再通过MCDB模块对深度噪声图进行逐步的细化，MCDB模块后续会说到；这里需要注意的是采用了LDM策略
9. 这个LDM策略就是它会对加噪去噪过程重复很多轮，每次生成深度图后，会通过一个掩码来筛选有用的深度信息，然后再根据损失函数进行多次优化，在传回潜空间中进行重复
10. 这就是它整体的一个工作流程
11. 这一页讲的是MCDB，也就是它的去噪部分的实现，视觉引导条件与噪声图的融合部分我就不细说了，我们重点来讲述DDIM数学模型：在标准的DDPM中，我们定义前向和反向过程为这两个公式，而我们训练的模型要做的工作就是学习均值和来逐步降噪恢复。