Seeds of Structure: Patch PCA Reveals Universal Compositional Cues in Diffusion Models

Qingsong Wang[1], Zhengchao Wan[2], Mikhail Belkin[1], Yusu Wang[1]
[1]University of California, San Diego, [2]University of Missouri

(1) The Noise-to-Image Map

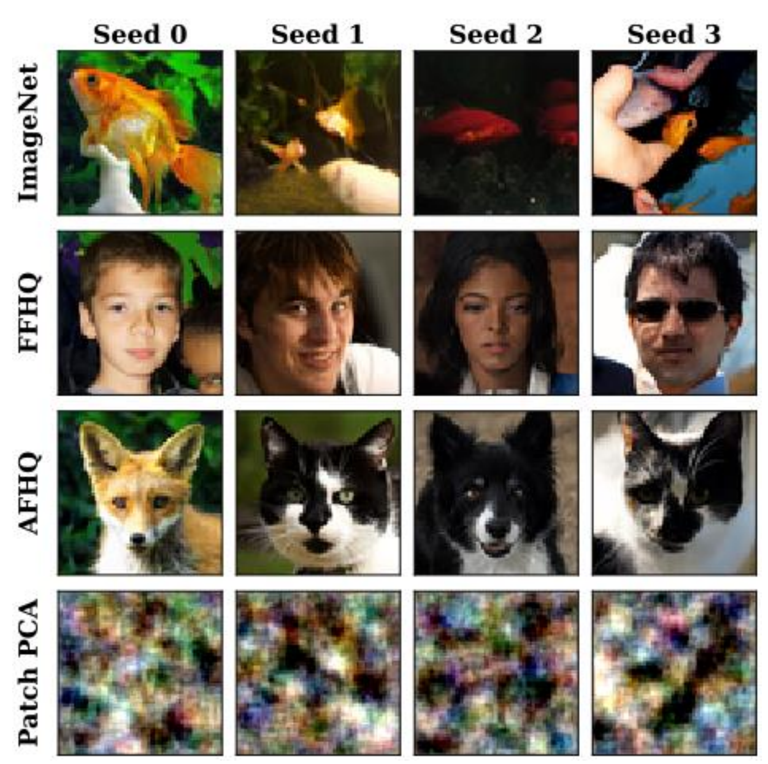
Diffusion model with an ODE sampler deterministically transforms noise into an image.

Question: What are the properties of this noise-to-image map?

Our findings:
- **Noise is not just random:** compositional structure of a generated image (e.g., layout, lighting) predominantly determined by **the low frequency part of the initial noise**.
- **Patch PCA denoiser** effectively extracts the shared compositional cue in noise.
- **Enable compositional control through noise editing.**

Similar Layout from the Same Seed

Same noise seed generates visually similar images with diffusion models trained on **different datasets**



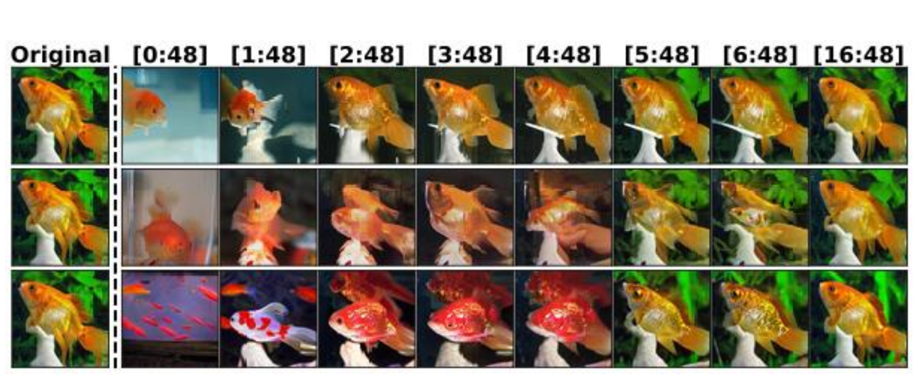
Images from identical seeds on ImageNet, FFHQ, AFHQ, and Patch PCA (to be introduced later)

Quantitative validation: high similarity score for the generated image from the same noise seed compared with randomly paired.

	SSIM (*higher is better*)		MSE (*lower is better*)	
Network Pair	Same Seed	Random	Same Seed	Random
ImageNet vs FFHQ	0.423 ± 0.087	0.065 ± 0.040	0.041 ± 0.017	0.136 ± 0.055
ImageNet vs AFHQ	0.447 ± 0.097	0.062 ± 0.039	0.038 ± 0.018	0.130 ± 0.055
FFHQ vs AFHQ	0.469 ± 0.074	0.054 ± 0.041	0.032 ± 0.012	0.128 ± 0.046

Sensitivity of the Noise-to-Image Map

Perturbing low-frequency Patch-PCA components in initial noise dramatically changes generated image, while high-frequency perturbation leaves layout and semantics nearly unchanged.



(2) The Patch PCA Denoiser

Extracting the compositional cue with *Patch PCA denoiser*:

1. **Patchify** the image into overlapping patches.
2. **Apply PCA denoiser for each patch** with PCA computed from a generic set of patches (μ is the mean and λ_i are eigenvalues and u_i are eigenvectors).

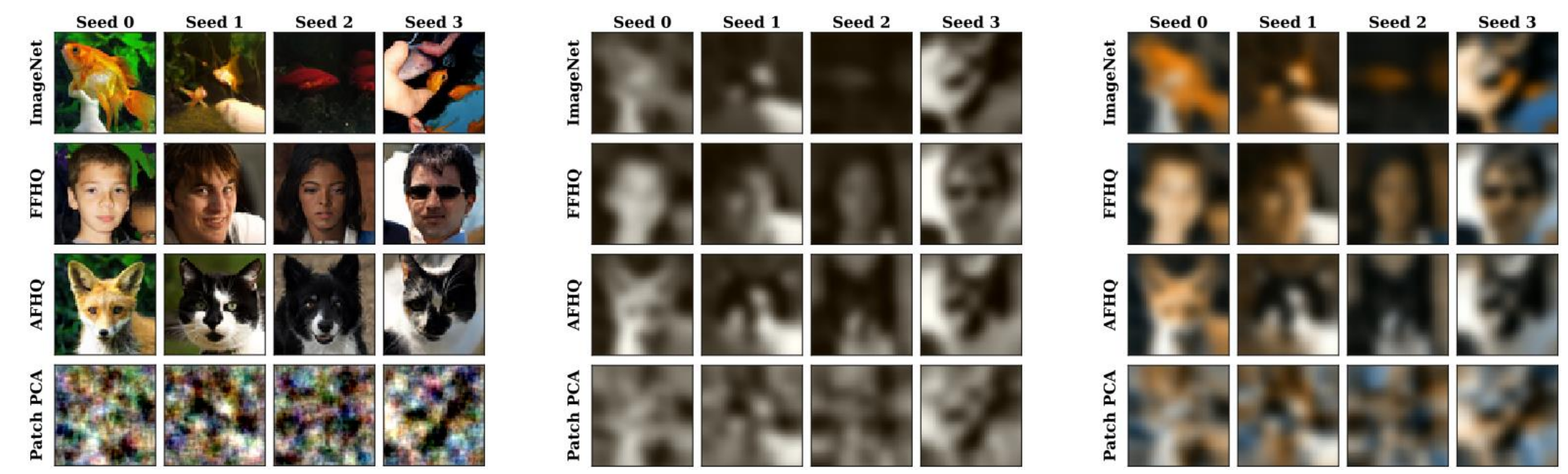
$$D_{\mathrm{PCA}}(\mathbf{p}_i, \sigma) := \boldsymbol{\mu} + \sum_{j=1}^{p^2 c} \frac{\lambda_j}{\lambda_j + \sigma^2} \langle \mathbf{p}_i - \boldsymbol{\mu}, \mathbf{u}_j \rangle \mathbf{u}_j$$

3. **Unpatchify** to reconstruct the image.

Theorem 5.1 (informal) Under a Gaussian patch assumption, the Patch PCA denoiser is optimal among all "patch-local" denoisers.

Qualitative validation of Patch PCA Denoiser:

Images (left) from identical seeds on ImageNet, FFHQ, AFHQ, and Patch PCA denoiser and their patch-wise projection to top-1 (middle) and top-2 (right) eigenspaces.



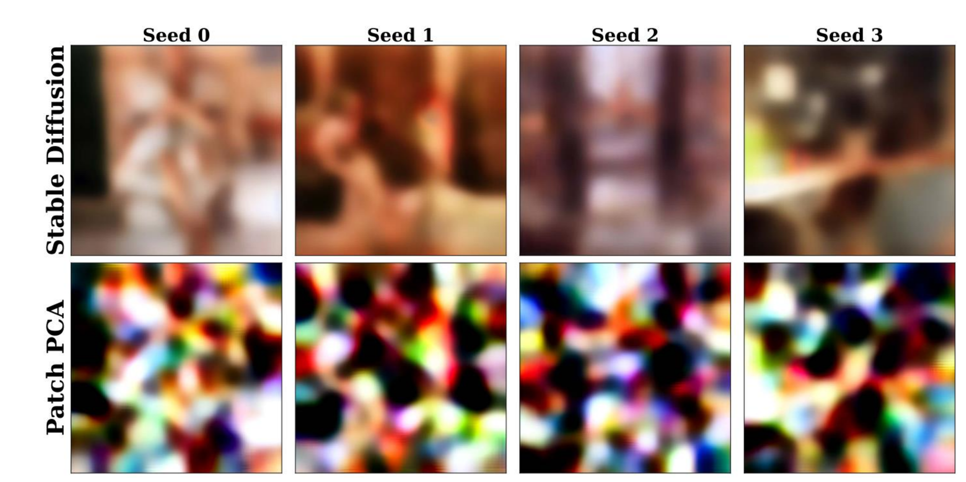
Quantitative validation of Patch PCA Denoiser

	SSIM (*higher is better*)		MSE (*lower is better*)	
Network Pair	Same Seed	Random	Same Seed	Random
Patch PCA vs ImageNet	0.474 ± 0.137	0.031 ± 0.027	0.049 ± 0.025	0.115 ± 0.037
Patch PCA vs FFHQ	0.473 ± 0.077	0.029 ± 0.033	0.045 ± 0.014	0.114 ± 0.027
Patch PCA vs AFHQ	0.548 ± 0.084	0.029 ± 0.034	0.036 ± 0.014	0.104 ± 0.029

Extension to Latent Diffusion model

Our method extends to latent diffusion models. A PCA denoiser built from patch statistics of encoded images generates images with a similar layout to neural network outputs, as illustrated below and quantified in our paper.

Stable Diffusion Generated Images



Patch-wise projection to the top 2 eigenspaces

(3) Training-Free Image Editing by Noise Editing

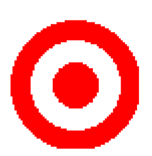
Editing with CIFAR-10 diffusion model

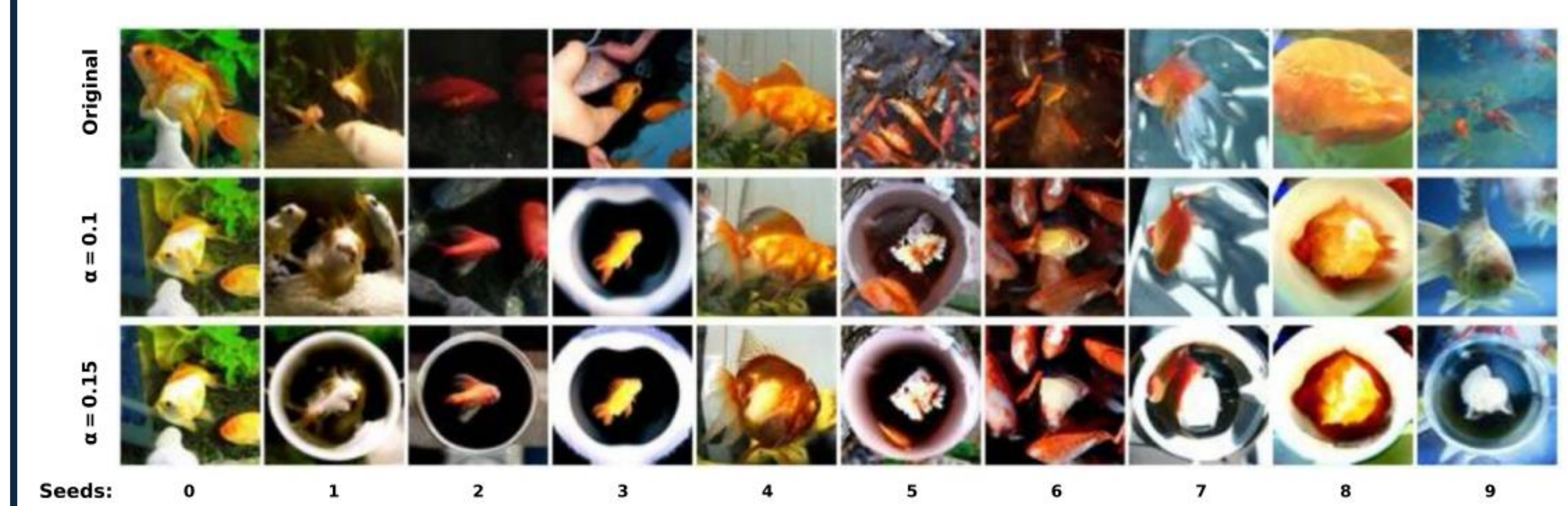
Original generation

Reference

Aligning the low-frequency part in initial noise with the reference image.

Edited generation

Editing with ImageNet diffusion model (still use  as reference) at different aligning strengths.



Takeaways

Noise is not just random

- **Compositional cues in noise are universal across models**
- **Patch PCA captures the compositional cue**
- **Training-free control via noise editing**

Acknowledgements

This material is based upon work supported by NSF via grants CCF-2112665, DMS-2502083, and DMS-2502084, by the Office of Naval Research (ONR N000142412631), as well as by the Defense Advanced Research Projects Agency (DARPA) under Contract No. HR001125CE020.