Manipulating Embeddings of Stable Diffusion Prompts

Niklas Deckers, Julia Peters and Martin Potthast

https://www.ijcai.org/proceedings/2024/845

Abstract

- Prompt engineering is used to manipulate images generated with text-to-image models
- We propose a new method to directly manipulate the embedding of a prompt instead of the prompt text
- Treating the model as a continuous function and passing gradients between the image space and the prompt embedding space
- We derive three practical interaction tools:

(1) Optimization of a metric defined in the image space

(2) Supporting a user in creative tasks

(3) Changing the embedding of the prompt to include information that is seen in a particular seed but difficult to describe.

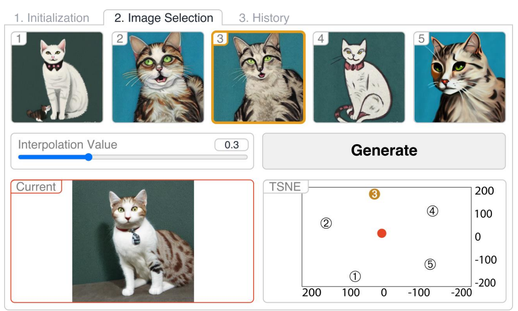
The prompt embedding space is continuous, allowing manipulations like interpolation.

Proposed Methods

Our three techniques for manipulating prompt embeddings enable a user to

(1) optimize an image quality metric,

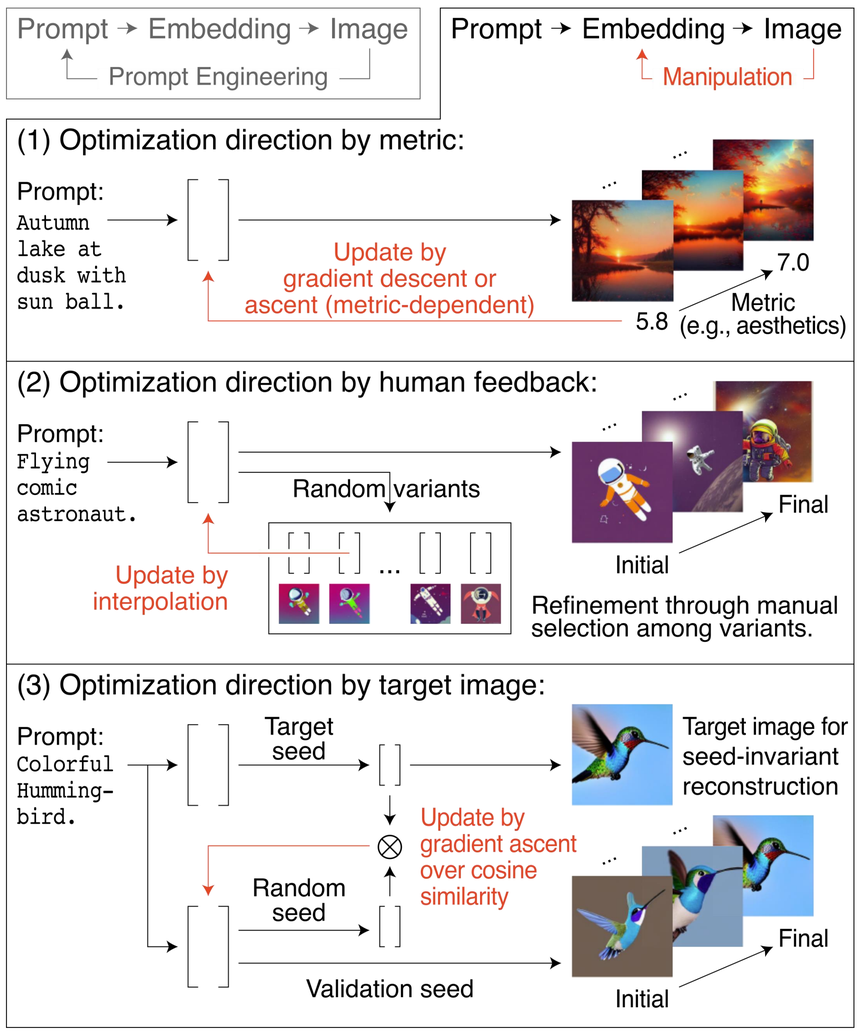
(2) navigate the prompt embedding space towards nearby variants,



(3) reconstruct a preferred image by introducing seed invariance.

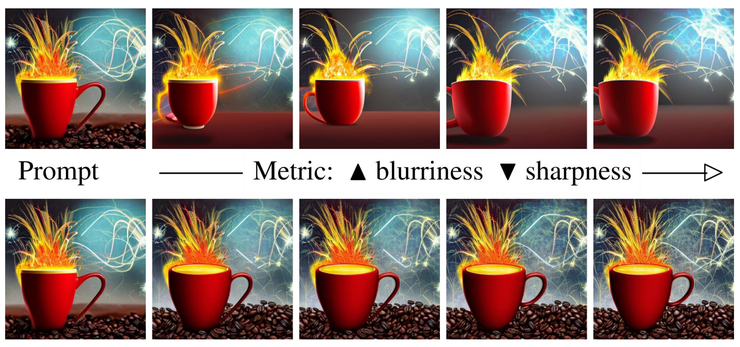
Algorithm 1: Seed-Invariant Prompt Embeddings

$\mathcal{I} \leftarrow \mathrm{LDM}(\psi(\mathcal{P}), z)$
$\mathcal{C} \leftarrow \psi(\mathcal{P})$
for $\alpha \leftarrow \frac{1}{n}, \ldots, \frac{n}{n}$ **do**
  Sample $\tilde{z}$ as a batch of random initial latents
  $L \leftarrow \|\mathcal{I} - \mathrm{LDM}(\mathcal{C}, \mathrm{SLERP}(z, \tilde{z}, \alpha))\|_2^2$
  $\mathcal{C} \leftarrow \mathcal{C} - \eta \nabla_{\mathcal{C}} L$
end for
return $\mathcal{C}$

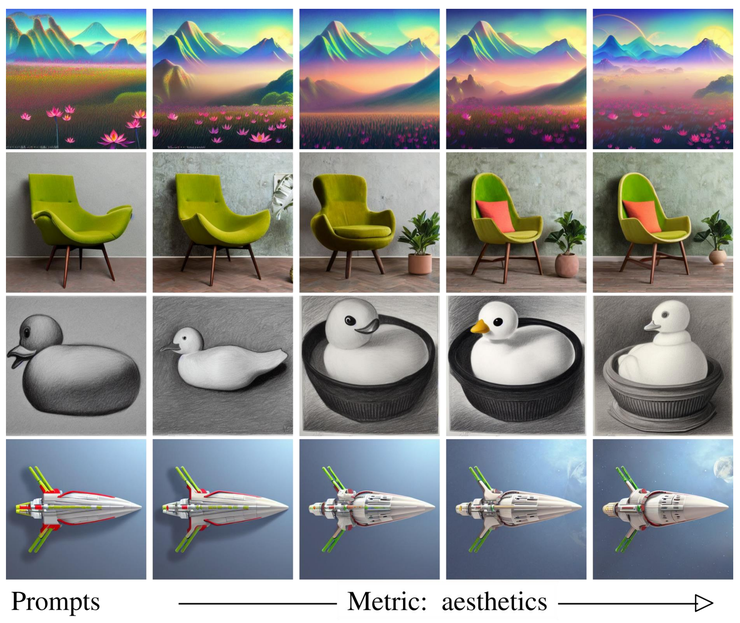


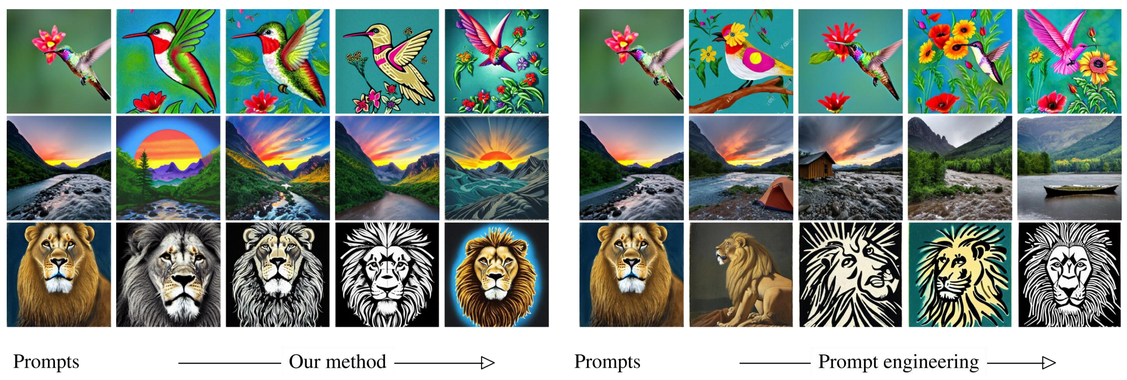
Results

(1) Images for prompt embeddings optimized w.r.t. a blurriness and sharpness metric,

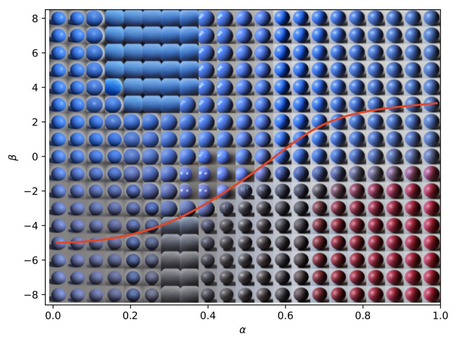


and optimized w.r.t. an aesthetics metric based on human feedback.

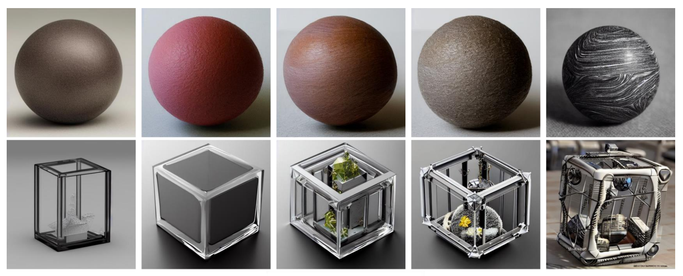




(2) Images created in our user study using iterative human feedback vs. prompt engineering.



(3) Optimizing prompt embeddings (β-axis) starting from a given image (bottom left) to resemble the original image when using a different seed (right).



Results without restricting the prompt embedding adjustments to a given dimension.

CONTACT

Niklas Deckers
Leipzig University & ScaDS.AI
niklas.deckers@uni-leipzig.de

Martin Potthast
University of Kassel & hessian.AI
martin.potthast@uni-kassel.de

SPONSORED BY

Bundesministerium für Bildung und Forschung
FKZ: ScaDS.AI

SACHSEN
Diese Maßnahme wird gefördert durch die Bundesregierung aufgrund eines Beschlusses des Deutschen Bundestages. Diese Maßnahme wird mitfinanziert durch Steuermittel auf der Grundlage des von den Abgeordneten des Sächsischen Landtags beschlossenen Haushaltes.

The authors gratefully acknowledge the GWK support for funding this project by providing computing time through the Center for Information Services and HPC (ZIH) at TU Dresden on HRSK-II.