Tailored Visions: Enhancing Text-to-Image Generation with Personalized Prompt Rewriting

Supplementary Material

7. Details of User Preference Summarization in PIP Dataset

Here, we provide details about how the user preference P_u in PIP dataset is summarized. As mentioned in Section 3.3, we use ChatGPT to summarize P_u based on the user u's historical prompts $\mathcal{Q}_t$. Considering ChatGPT's context length, we randomly select 50 historical prompts for summarization, assuming users have more than 50 different historical prompts. The input template for ChatGPT is shown in Table 6.

ChatGPT Input Template for Summarizing User Preference

Your task is to use no more than five phrases to summarize user's preference based on history text prompts he uses in text to image generation.

A user's preference reveals what kind of image he might prefer or the image style he likes. Do not include objects that appear in history prompts in your answer. Just answer the phrases in sequential order, separate using a comma.

Please summarize the preference for the following user:

The history prompts of a user: $\mathcal{Q}_t$

The keywords of the user's preference:

Table 6. Prompt Template we use for summarizing user preference

8. Further Details of In-Context Personalized PR and General PR

As mentioned in 5.1, our prompt manner for in-context rewriting is as similar as for context-independent, both in Personalized PR and General PR. Here we show the in-context personalized PR input template in Table 7, and we show the general PR input template in Table 8, respectively.

9. Additional Qualitative Results

We provide additional qualitative results of our Personalized PR method to demonstrate the effectiveness of our method.

In Figure 10 and Figure 11, we show comparisons of our method against baseline methods, i.e. Promptist [7] and General PR, in terms of the rewritten prompt as well as the corresponding images generated using SD v1-5. All methods use the "Shortened Prompt" as input x_t. To evaluate the alignment with the user preference of the compared methods, we show the user preference P_u for several cases of the exemplar. By leveraging historical prompts, our method

---

ChatGPT Input Template for In-context Rewriting

Prompt in text-to-image generation describes the detailed attributes of the object user plans to draw. User's preference in text-to-image generation is shown in history prompts.

Given 3 history prompts, your task is to rewrite the current prompt so that it matches the user's preference. The rewritten prompt should retain primary objects in the original prompt and conform to the user's preference. Please avoid being too diffused and restrict your output within 70 words.

Examples:$\{\mathcal{E}\}$

The history prompts are: $\{\mathcal{R}_t\}$

The current prompt is: $\{x_t\}$

The rewritten prompt (one sentence less than 70 words) is:

Table 7. ChatGPT input template for in-context personalized prompt rewriting.

ChatGPT Input Template for General Prompt Rewriting

Prompt in text-to-image generation describes the detailed attributes of the object he plans to draw. A nice prompt for text-to-image generation usually includes various aspects of the image including descriptions of the scene, mood, style, lighting, and more.

Given an input prompt, your task is to rewrite the prompt to a better one. Your rewritten prompt is supposed to describe the image better, and less than 70 words.

The input prompt is: $\{x_t\}$

The rewritten prompt (one sentence less than 70 words) is :

Table 8. Input template for General PR.

can rewrite the prompt in accordance with the user's preference in each case. However, general prompt rewriting methods either insert just a few tags (such as "digital art" or "by greg rutkowski" etc.) arbitrarily. These generations either produce something resembling a carefully crafted text-to-image prompt or continually add tedious details, resulting in a lengthy text output. Lost histories, these methods fail to meet the user's preference.

As mentioned in 5.3 and Table 3, we have conducted experiments to demonstrate the performance of our method in terms of two shorter types of input prompt x_t. The quantitative results in 3 showcase that our method performs robustly even on conditions of very short prompt. Here, we provide two qualitative examples along with user ground truth generated images in Figure 12. These results showcase that, by leveraging historical prompts, our method is capable of
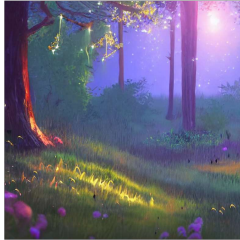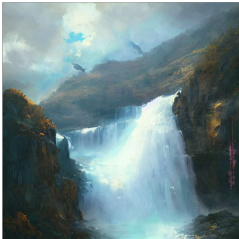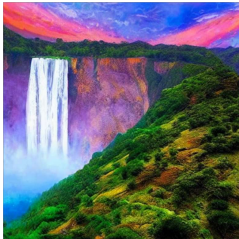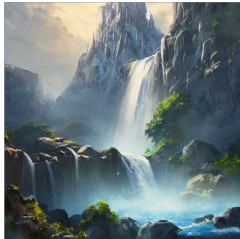
Shortened Prompt	Promptist	General PR	Personalized PR (Ours)	User Preference

Magical forest with dramatic lighting.

Magical forest, dramatic lighting, by greg rutkowski, digital art, trending on artstation

Immerse yourself in the enchanting depths of a mystical forest as vibrant rays of golden light pierce through the ancient trees, casting an ethereal glow that whispers of untold secrets and hidden wonders.

Magical forest with fireflies, wild flowers, vines, and dramatic lighting, fantastic mysterious style, trending in artstation.

night scenes, nature, fantasy landscapes, dream-like atmosphere, surrealism

Sky Waterfall

Sky Waterfall, by WLOP,

Immerse yourself in the breathtaking beauty of a sky waterfall, where the boundless expanse of the sky seamlessly cascades down like a majestic waterfall, painting the horizon with a mesmerizing display of colors and patterns that evoke a sense of tranquility and wonder.

A majestic sky waterfall cascades down against the backdrop of a towering snow mountain, creating a visual masterpiece that is currently trending in artstation.

scenary, peaceful, trending in artstation, snowy landscape, ancient and mysterious

Moon and boat

Moon and boat. intricate, elegant, highly detailed, digital painting, art station, concept art, sharp focus, illustration, by justin gerard and artgerm, 8k

A serene scene of a crescent moon shining brightly over a small wooden fishing boat, peacefully gliding across the calm waters, with the moon's reflection gently dancing on its surface, creating a tranquil and surreal atmosphere. The soft rays of moonlight illuminate the boat, casting shadows and highlighting its intricate details, while the distant stars twinkle in the night sky, adding a touch of enchantment to the captivating image.

A small boat floats on a river stream illuminated by moonlight, creating a serene and captivating scene that is trending on Artstation.

moonlight, river, mountains, fantasy, romantic

Figure 10. Qualitative comparison of our method Personalized PR against baseline methods. Compared with Promptist, General PR, our personalized method aligns with user preference better. Please zoom in to see details.

rewriting input prompt x_t properly and further generating images that are very close to the users' true intentions, either in terms of style or objects. Even in scenarios where the prompt only comprises nouns without attributes, our method adeptly discerns the implicit taste of users toward these objects, resulting in image generation much closer to the users' actual preferences.
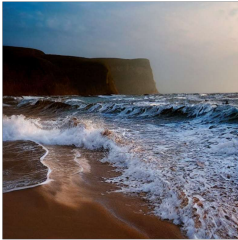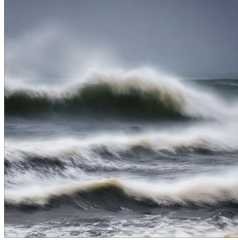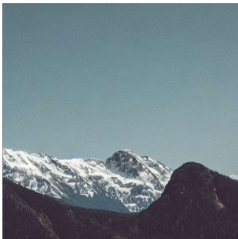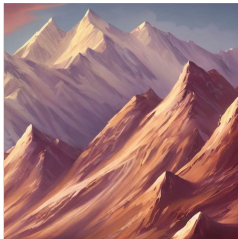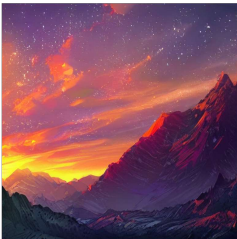
Shortened Prompt	Promptist	General PR	Personalized PR (Ours)	User Preference

Angry sea wave

angry sea waves, intricate, elegant, highly detailed, digital painting, artstation, concept art, smooth, sharp focus, illustration,

Furious waves crash against the shoreline, swirling with uncontrol-llable rage, sending torrents of water spraying into the air, creating an intense and tempestuous scene.

The sea is filled with raging waves, crashing against the shore with a nger and intensity, creating a tumultuous sight amidst the vastness of the ocean, trending in artstation.

oceanic landscapes, dusty scenery, morning scenes, powerful,

Strange dreams, trending in artstation.

strange dreams, trending in artstation.

Enter the realm of Strange Dreams, where ethereal and otherworldly visions captivate the senses in the latest art sensation taking over ArtStation. Immerse yourself in a scene that seamlessly blends reality and fantasy, creating a mood of intrigue and mystery. Marvel at the impeccable style and attention to detail, enhanced by the play of light and shadows, as each stroke of the artist's brush brings these dreams to life on canvas.

In the dreamy distance, a surreal world is inhabited by a charming group of people. As the train arrives, the infinite beauty of the sunset unfolds before their eyes. This enchanting scene captures the essence of strange and captivating dreams. It is a trending topic on Artstation, captivating the imagination of art enthusiasts.

magical scenes, ancient architecture, dreamy atmosphere, nature, beautiful

Mountains

Mountains, intricate, elegant, highly detailed, digital painting, artstation, concept art, smooth, sharp focus, illustration,

A majestic mountain range rises against a clear blue sky, displaying its rugged peaks and deep valleys, creating a sense of awe and wonder. The sunlight illuminates the scene with a warm and golden glow, casting long shadows that highlight the sublime beauty of nature. The serene and peaceful atmosphere invites one to explore and experience the tranquility of these breathtaking mountains.

In the vastness of mountains, the sunset paints the sky with vibrant colors, while the night sky is adorned with a canopy of countless stars, trending in artstation.

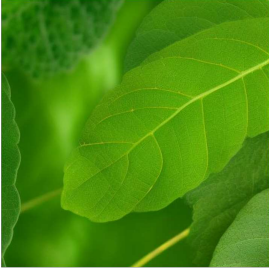night scenes, natural landscapes, romantic atmosphere, dreamy elements, seasonal beauty.
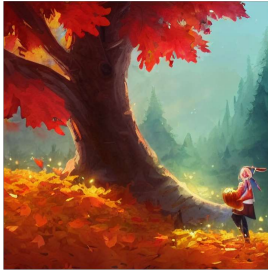
Figure 11. More comparison results. Please zoom in to see details.
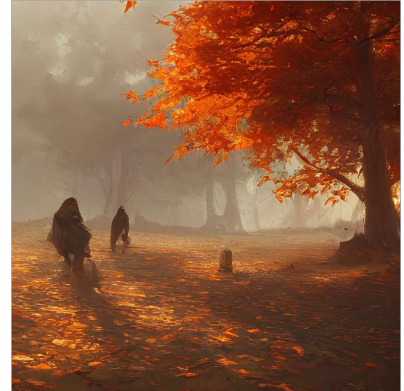
10. Failure Cases

Our Personalized PR is a first trail of personalized text-to-image prompt rewriting, where we evaluate our methods using naive yet effective methods and models. Nevertheless, we would like to present failure cases of our Personalized PR method. While such failures do occur for some reasons, they are exceptionally rare. From which, we expect that our method's failure could help promote further studies on personalized prompt rewriting.

Figure 13 illustrate a failure case where irrelevant objects are involved into the image. We attribute the results

Figure 12. Qualitative results of our method under different type (length) of input prompts. Our personalized PR can generate more expressive user-preferred images as similar as the user groundtruth compared with the original prompt. Please zoom in to see details.
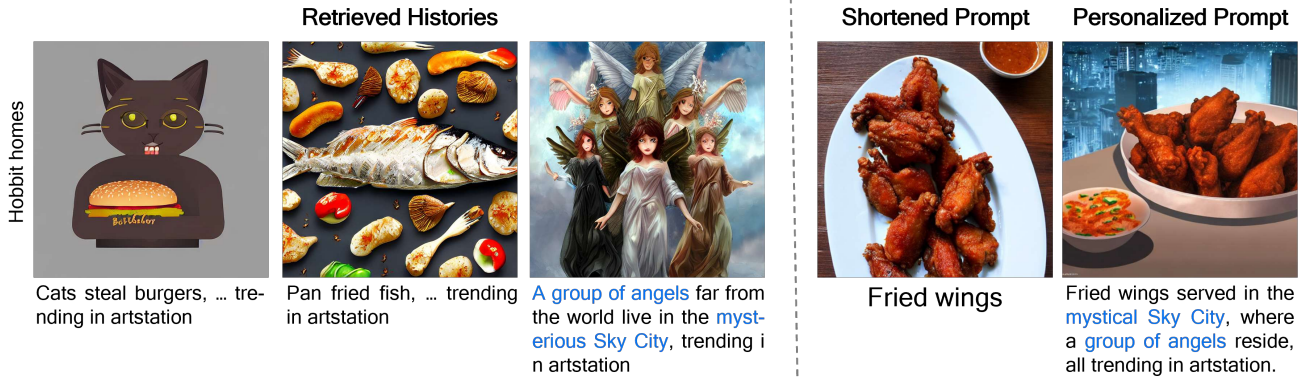
Retrieved Histories

Hobbit homes

Cats steal burgers, … trending in artstation

Pan fried fish, … trending in artstation

A group of angels far from the world live in the mysterious Sky City, trending in artstation

Shortened Prompt — **Personalized Prompt**

Fried wings

Fried wings served in the mystical Sky City, where a group of angels reside, all trending in artstation.

Figure 13. A failure case of our method. The words in blue indicate irrelevant objects are added to the rewritten prompt, resulting bad generation.
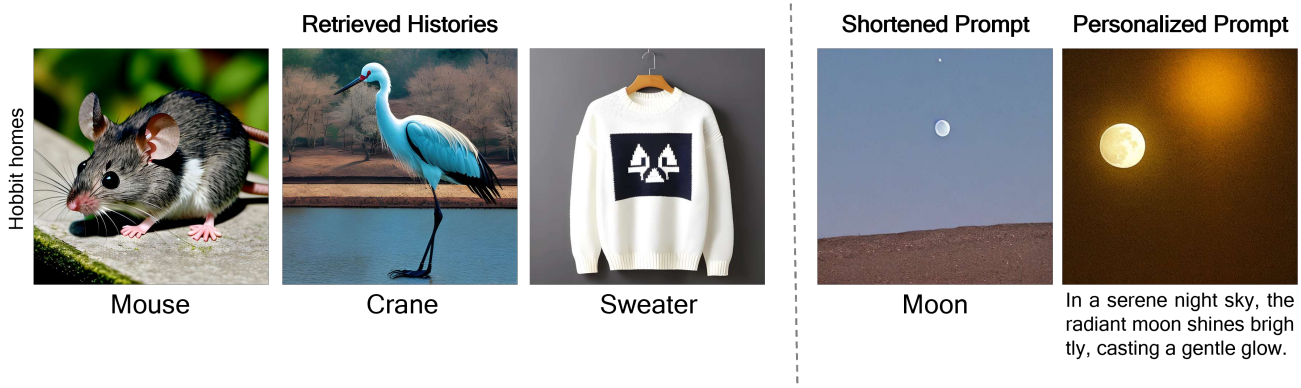

Retrieved Histories

Hobbit homes

Mouse

Crane

Sweater

Shortened Prompt — **Personalized Prompt**

Moon

In a serene night sky, the radiant moon shines brightly, casting a gentle glow.

Figure 14. A failure case of our method.

into two reasons: the irrelevant retrieval result $\mathcal{R}_t$, and the misconception of Rewriter. For the irrelevant retrieval result, the retrieved historical prompts describe some totally different scenes, such as "group of angels" and "Sky City" in the third retrieved history shown in Figure 13. For the misconception of Rewriter, the Rewriter ChatGPT [15] captures some objects irrelevant to "fried wings" which is the user's preference, thus leading to this failure.

Furthermore, we visualize another failure case in Figure 14. It happens when the Rewriter fails to acquire adequate personalized information from historical prompts. In this case, the user prefer to generate various objects without sequential connections in the historical prompts. Therefore, the Rewriter fails to capture useful information for prompt rewriting. As a result, the rewritten prompt can be likely the generation result sourcing from the ChatGPT's hallucination.

Even though there might be instances of failure, our Personalized PR poses a foundation method to align text-to-image prompts with user preferences. With our constructed PIP dataset, proposed Personalized PR method and the corresponding standard evaluation, we have paved a pathway

of personalization in text-to-image generation and we expect to foster further research towards better user preference extraction.