

GENERATIVE
RENAISSANCE

[ASIA.SIGGRAPH.ORG/2025](https://asia.siggraph.org/2025)



SIGGRAPH 香港
ASIA 2025
HONG KONG 港

Conference 15 – 18 December 2025

Exhibition 16 – 18 December 2025

Venue **Hong Kong Convention
and Exhibition Centre**

**In-Context Brush:
Zero-Shot Customized Subject Insertion
with Context-Aware Latent Space
Manipulation**

Sponsored by



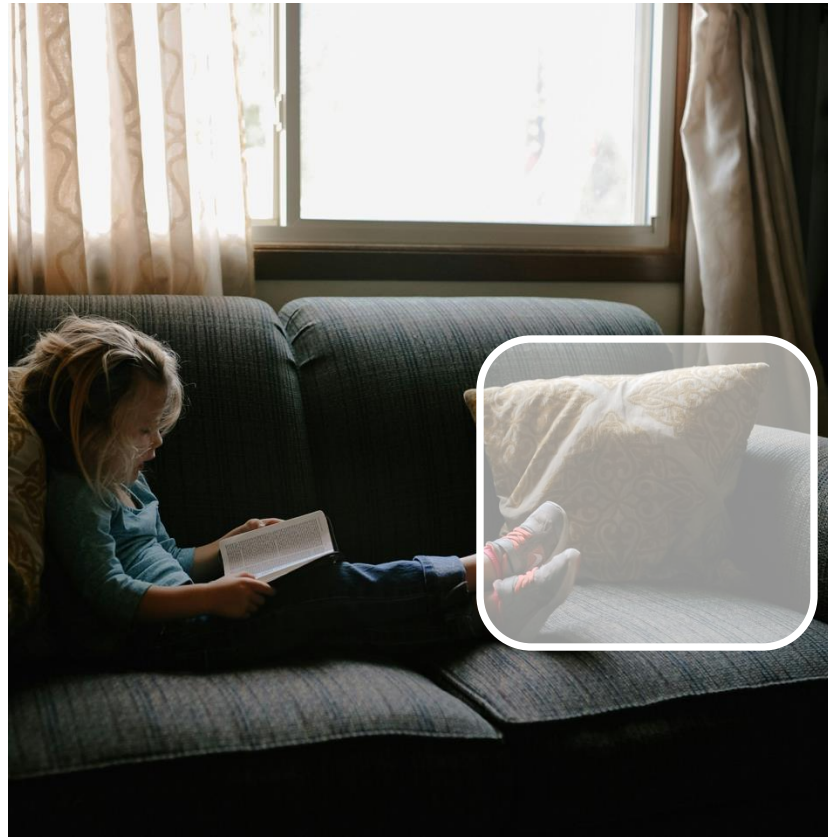
Organized by



In-Context Brush: Zero-Shot Customized Subject Insertion with Context-Aware Latent Space Manipulation

Yu Xu, Fan Tang, You Wu, Lin Gao, Oliver Deussen,
Hongbin Yan, Jintao Li, Juan Cao, Tong-Yee Lee

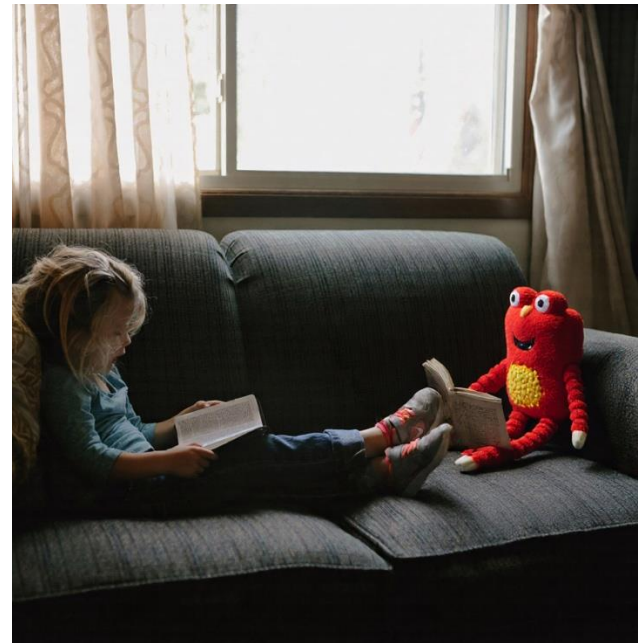
Institutes of Computing Technology, Chinese Academy of Sciences
University of Konstanz
National Cheng Kung University



Insert subject into a specified region of a given image



Playing a guitar



Reading a book



working on a laptop

Motivation



⊗ Inconsistent subject identity

⊗ Unnatural subject-background interaction

In Context learning in image generation

task prompt

“Landscape Art
Print in the style
of a comic book”

task input image



task output image



query image

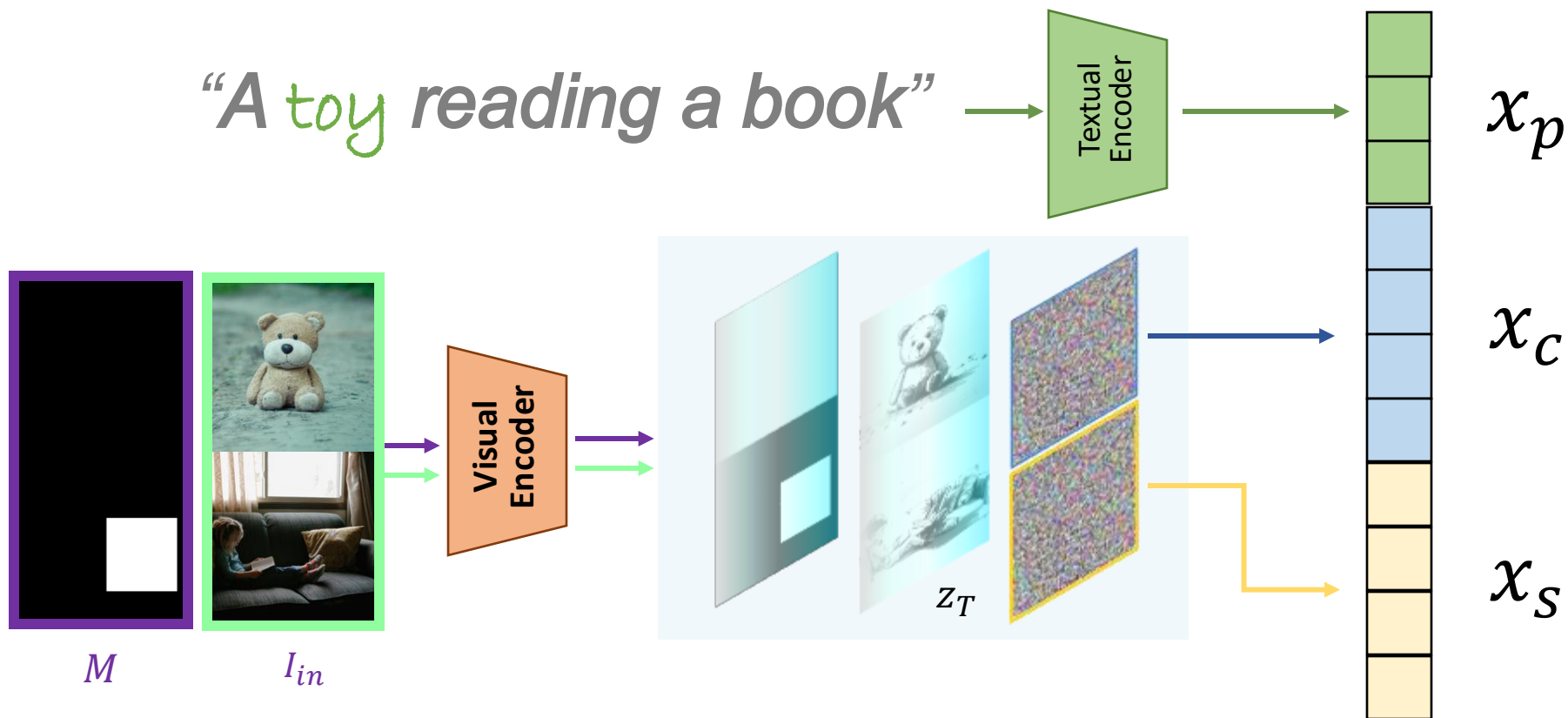


output image

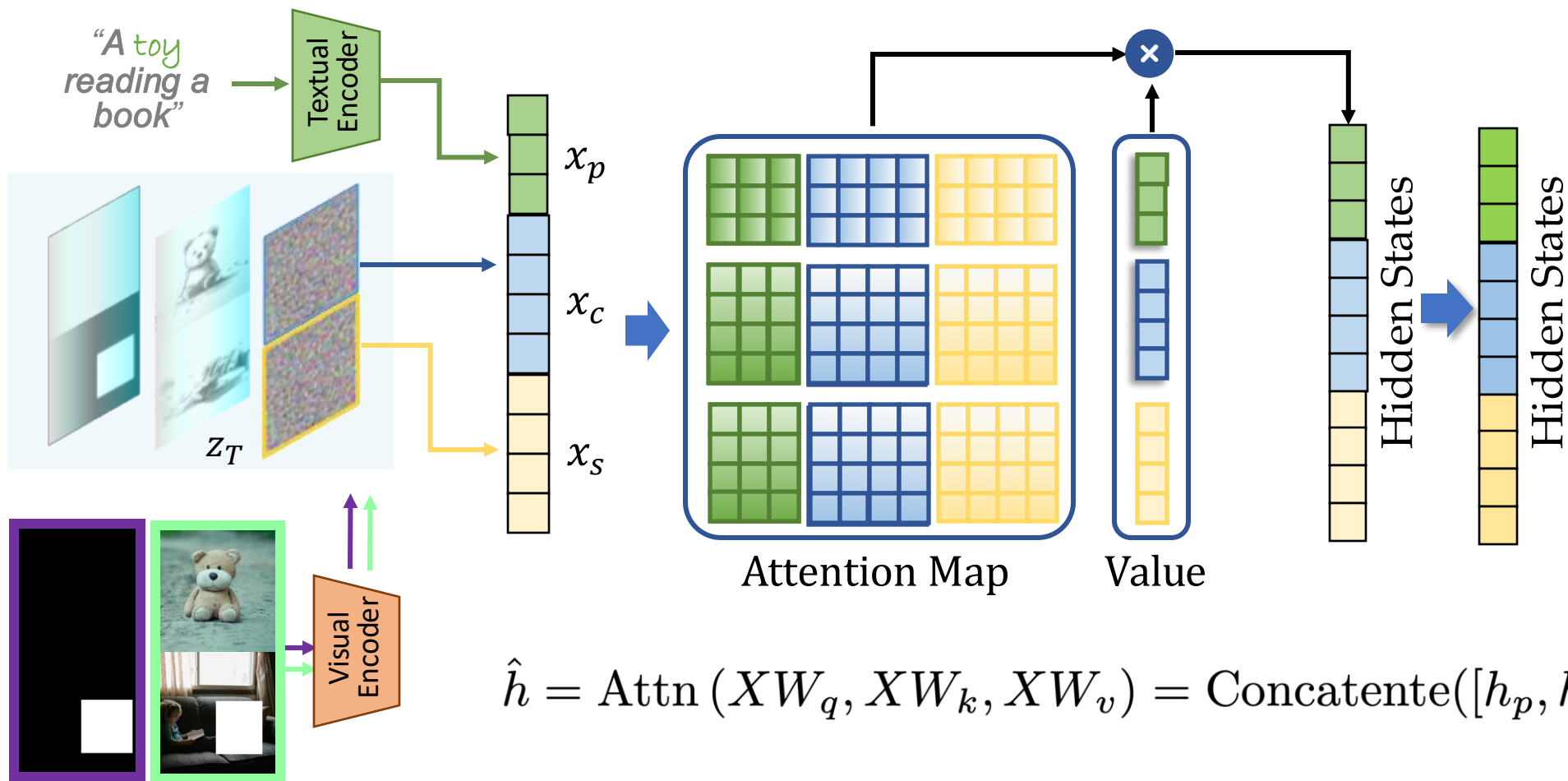




Insight: Represent the subject with in-context learning!

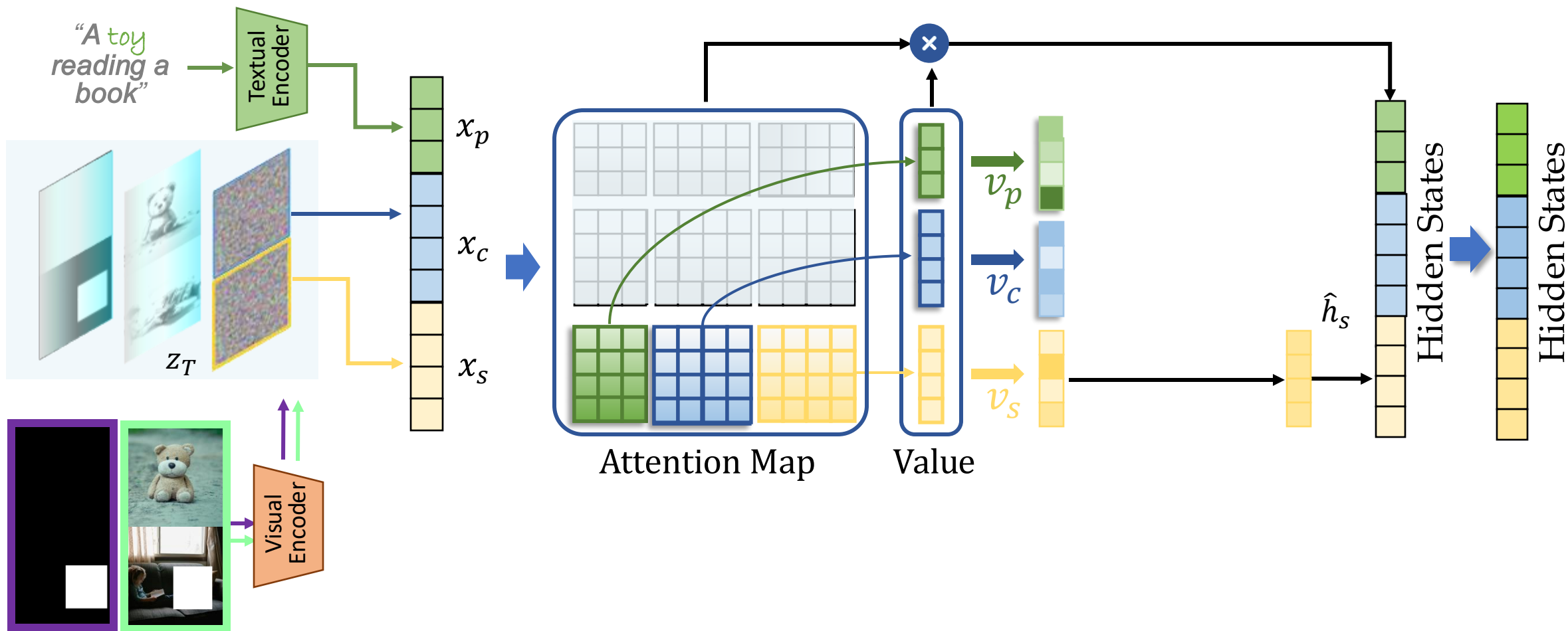


Latent feature shifting for subject injection

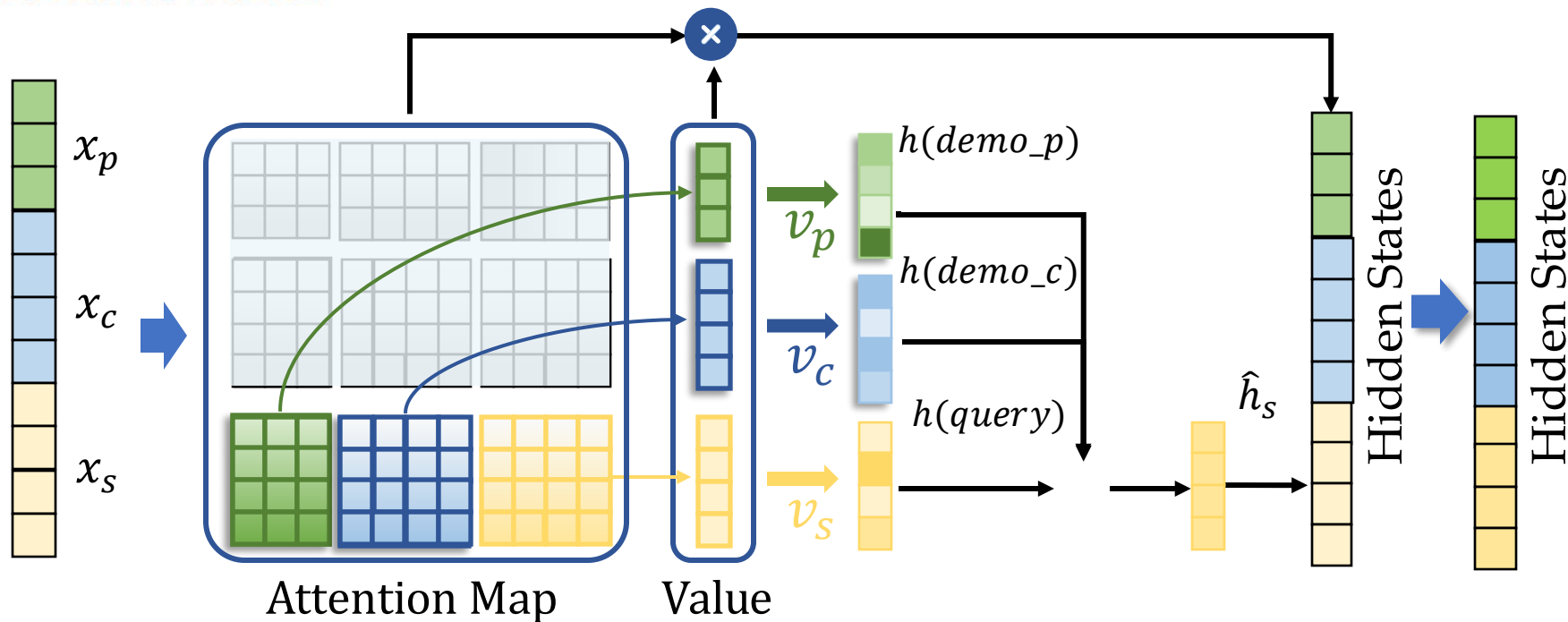


$$\hat{h} = \text{Attn}(XW_q, XW_k, XW_v) = \text{Concatente}([h_p, h_c, h_s])$$

Latent feature shifting for subject injection

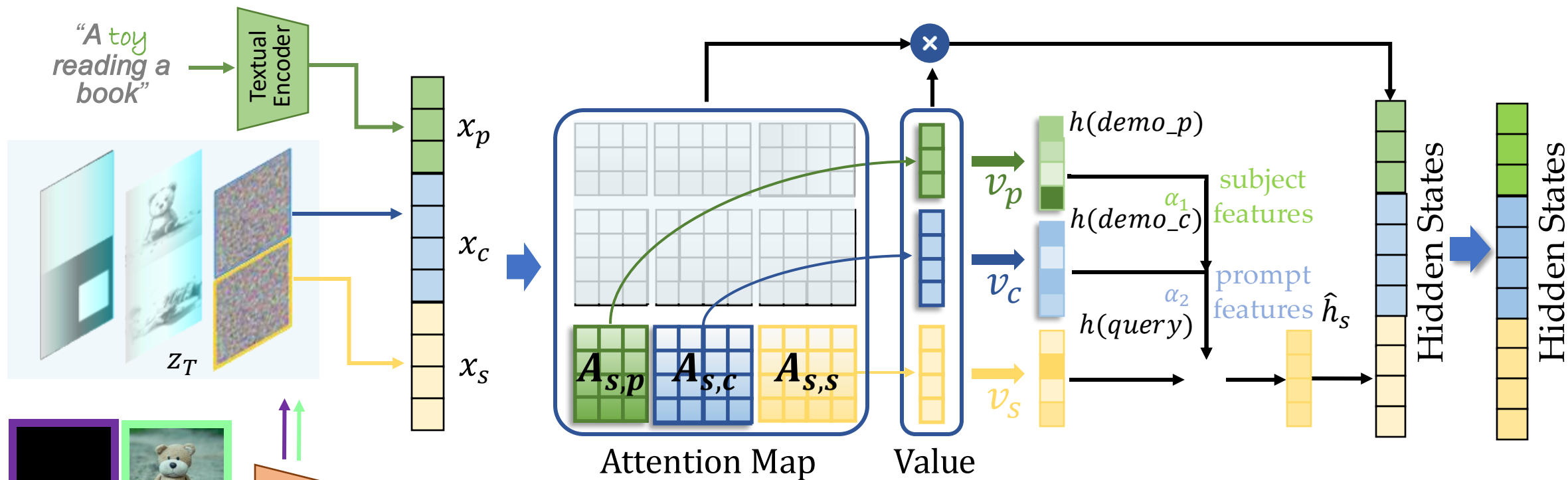


Latent feature shifting for subject injection



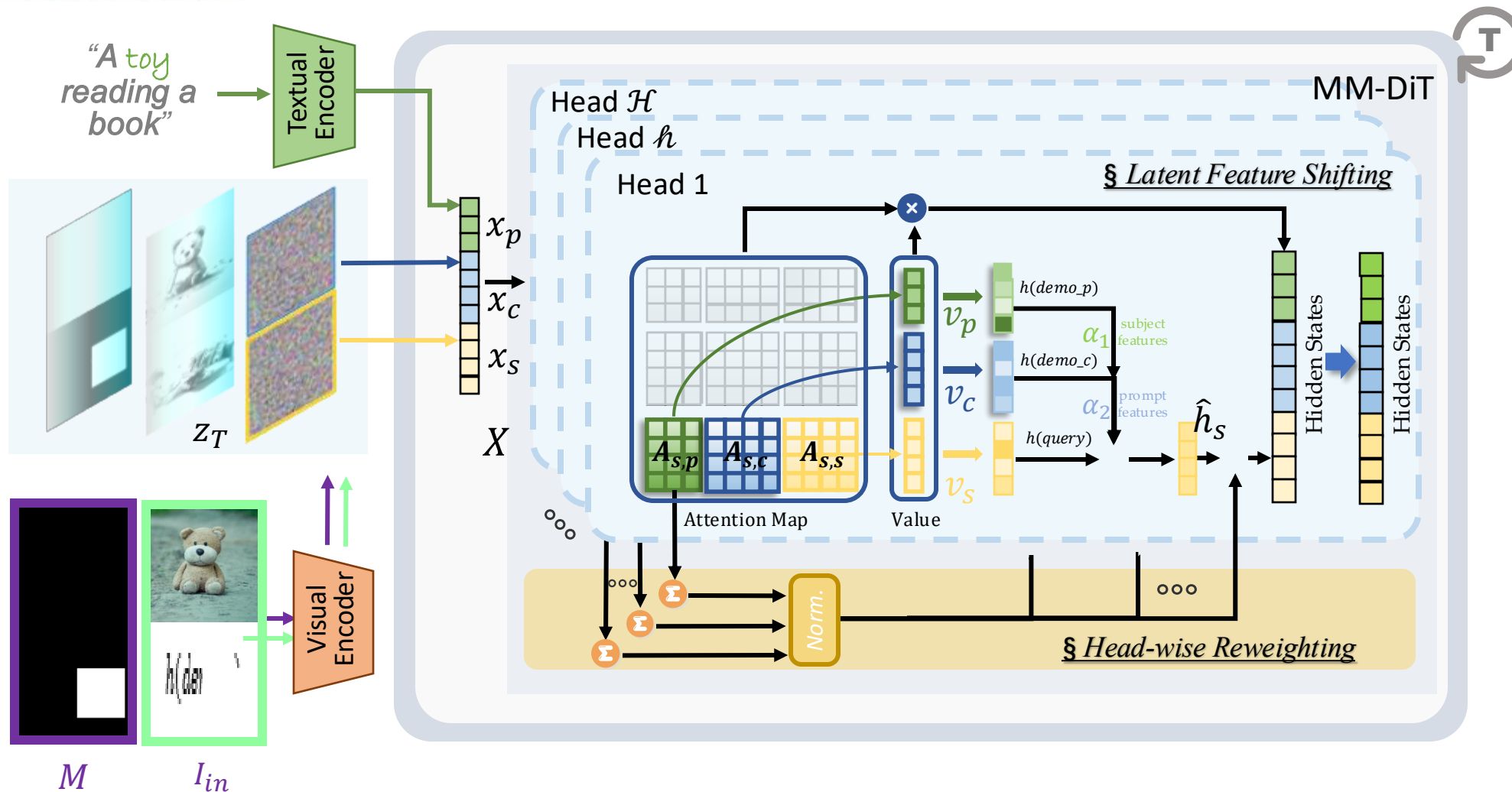
$$\begin{aligned}
 h_s &= \text{Softmax} \left(\begin{bmatrix} x_s W_{qk} x_p^\top & x_s W_{qk} x_c^\top & x_s W_{qk} x_s^\top \end{bmatrix} \right) \begin{bmatrix} x_p W_v \\ x_c W_v \\ x_s W_v \end{bmatrix}, \\
 &= \alpha_p \cdot \text{Attn}(x_s W_q, x_p W_k, x_p W_v) + \alpha_c \cdot \text{Attn}(x_s W_q, x_c W_k, x_c W_v) + \alpha_s \cdot \text{Attn}(x_s W_q, x_s W_k, x_s W_v), \\
 &= \alpha_p \cdot h(\text{prompt}) + \alpha_c \cdot h(\text{content}) + \alpha_s \cdot h(\text{output}), \\
 &= \alpha_p \cdot h(\text{demo}_p) + \alpha_c \cdot h(\text{demo}_c) + \alpha_s \cdot h(\text{query}),
 \end{aligned}$$

Latent feature shifting for subject injection

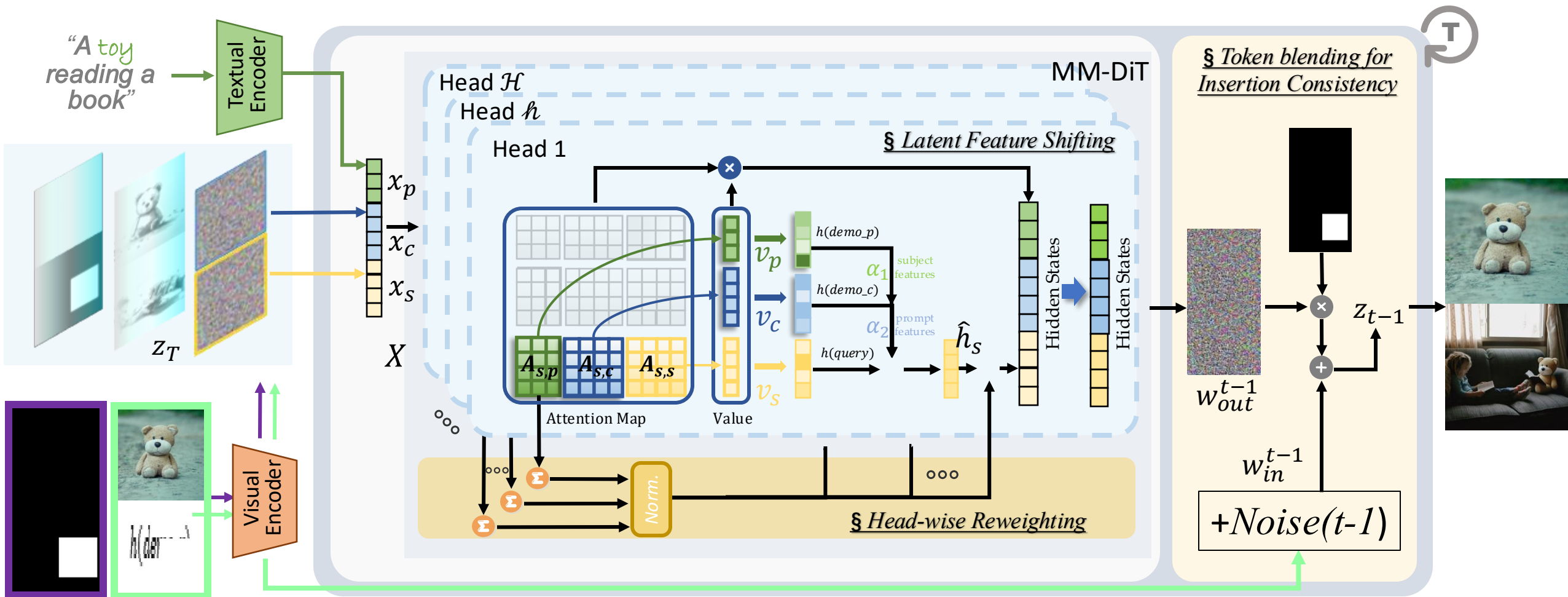


$$\hat{h}_s = h_s + \alpha_1 A_{s,p} v_p + \alpha_2 A_{s,c} v_c$$

Head-wise reweighting for textual control injection

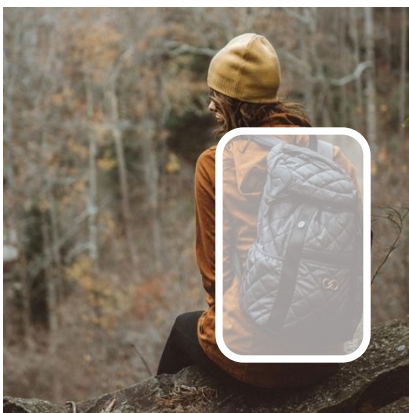


Token blending for insertion consistency



Comparison

Ours



Break-a-scene



Swap-anything



DreamEdit



IC-LoRA



TF-ICON



TIGIC



PrimeComposer



Diptych



Flux-Fill



“A woman with the backpack”

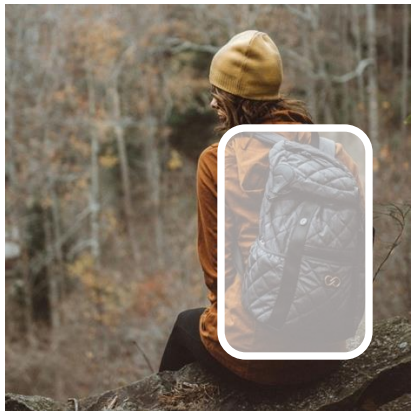


SIGGRAPH ASIA 2025 HONG KONG

香港

Comparison

Ours



Break-a-scene



Swap-anything



DreamEdit



IC-LoRA



TF-ICON



TIGIC



PrimeComposer



Diptych

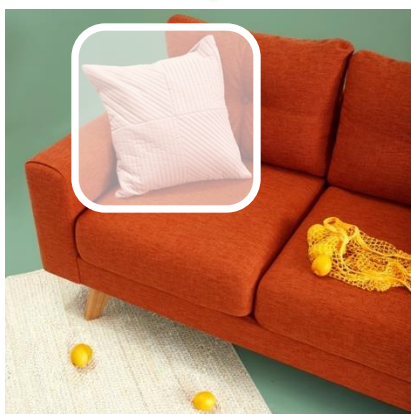


Flux-Fill



“A woman with blue backpack”

Comparison



Ours



Break-a-scene



Swap-anything



DreamEdit



IC-LoRA



TF-ICON



TIGIC



PrimeComposer



Diptych

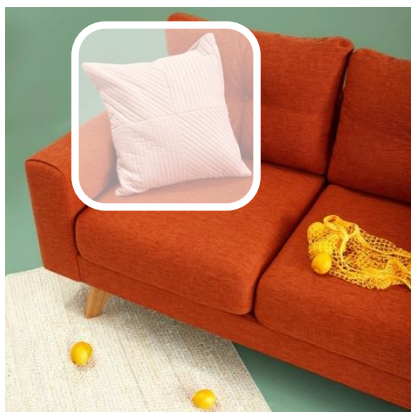


Flux-Fill



“A toy on the sofa”

Comparison



Ours



Break-a-scene



Swap-anything



DreamEdit



IC-LoRA



TF-ICON



TIGIC



PrimeComposer



Diptych



Flux-Fill



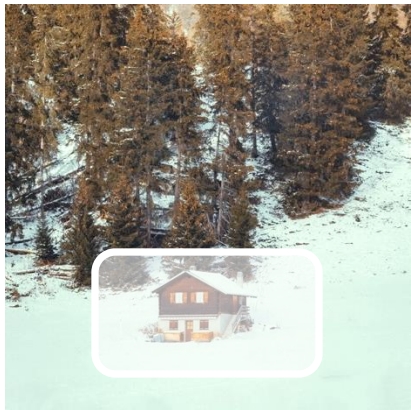
“A toy is playing a guitar”



SIGGRAPH
ASIA 2025
HONG KONG 香港

Comparison

Ours



Break-a-scene



Swap-anything



DreamEdit



IC-LoRA



TF-ICON

TIGIC



PrimeComposer



Diptych



Flux-Fill



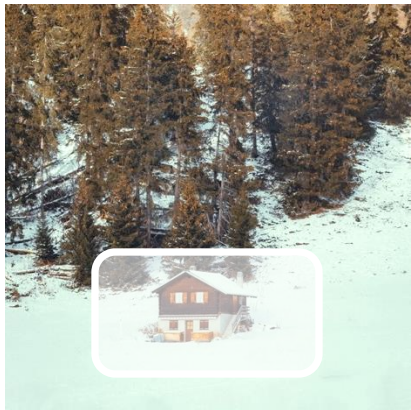
“A barn”



SIGGRAPH
ASIA 2025
HONG KONG 香港

Comparison

Ours



Break-a-scene



Swap-anything



DreamEdit



IC-LoRA



TF-ICON

TIGIC



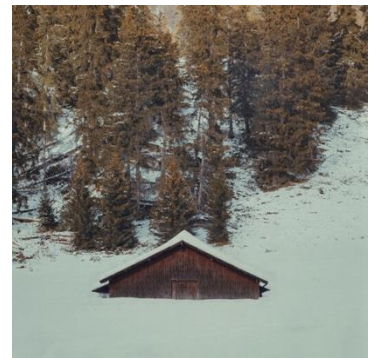
PrimeComposer



Diptych



Flux-Fill



“A barn covered by snow”

Quantitative comparison with baseline methods

Methods	DINO(↑)		CLIP-I(↑)		CLIP-T(↑)		FID(↓)
	Injection	Editing	Injection	Editing	Injection	Editing	
Break-a-scene	0.7041 ± 0.0659	0.7087 ± 0.0546	0.7128 ± 0.2076	0.6640 ± 0.1855	0.2570 ± 0.0230	0.2820 ± 0.0659	217.81
Swap-anything	0.7058 ± 0.0625	0.7035 ± 0.0556	0.7275 ± 0.1821	0.7296 ± 0.1602	0.2493 ± 0.0181	0.2296 ± 0.0354	190.69
TF-ICON	0.7053 ± 0.0280	0.7061 ± 0.0225	0.7235 ± 0.1188	0.7018 ± 0.1542	0.2334 ± 0.0270	0.2234 ± 0.0297	169.04
TIGIC	0.6901 ± 0.0231	0.6906 ± 0.0302	0.7145 ± 0.1744	0.6789 ± 0.2012	0.2656 ± 0.0313	0.2272 ± 0.0631	179.07
PrimeComposer	0.7029 ± 0.0510	0.6931 ± 0.0476	0.7124 ± 0.1128	0.7426 ± 0.0619	0.2609 ± 0.0163	0.2300 ± 0.0532	166.85
Diptych	0.6559 ± 0.0679	0.6531 ± 0.0691	0.7225 ± 0.1287	0.7145 ± 0.1140	0.2321 ± 0.0569	0.2287 ± 0.0403	179.28
Flux-Fill	0.6798 ± 0.0602	0.6761 ± 0.0497	0.7668 ± 0.1274	0.7843 ± 0.1124	0.2634 ± 0.0213	0.2667 ± 0.0288	127.18
Ours w/o head	0.7115 ± 0.0564	0.6891 ± 0.0484	0.7882 ± 0.1201	0.7889 ± 0.1121	0.2621 ± 0.186	0.2682 ± 0.331	123.82
Ours w/o blend	0.7116 ± 0.0591	0.6902 ± 0.0519	0.7901 ± 0.1192	0.7869 ± 0.1133	0.2633 ± 0.171	0.2703 ± 0.294	129.47
Ours	0.7121 ± 0.0542	0.6945 ± 0.0563	0.7957 ± 0.1138	0.7924 ± 0.1113	0.2685 ± 0.0173	0.2834 ± 0.0365	122.61

Identity alignment  + **7.35 %**

Instruction alignment  + **4.11 %**

image quality  + **3.73 %**

Ablation study on attention heads activation

References



Ours



W/o Head Activation



... metal material.



... with white T-shirts.

Ablation study on token blending

References



Ours



W/o Blend



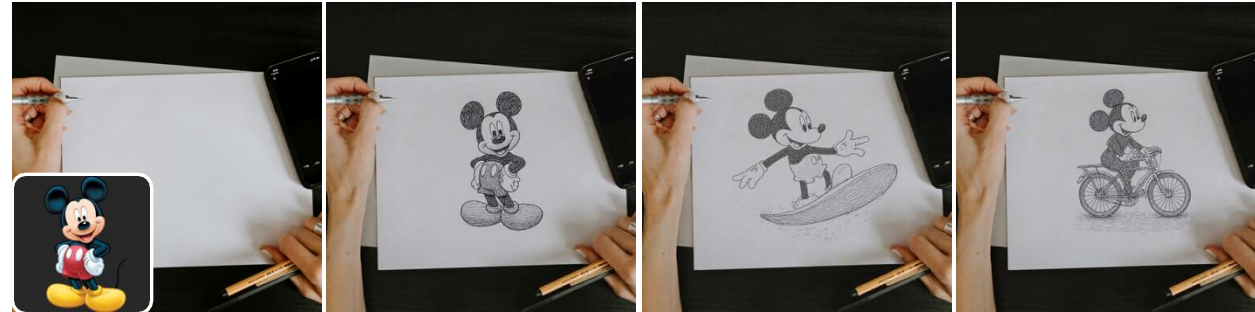
... graduated.



... sitting on the broom.

More results of our method

Generation with diverse prompts



Insert into diverse scenes



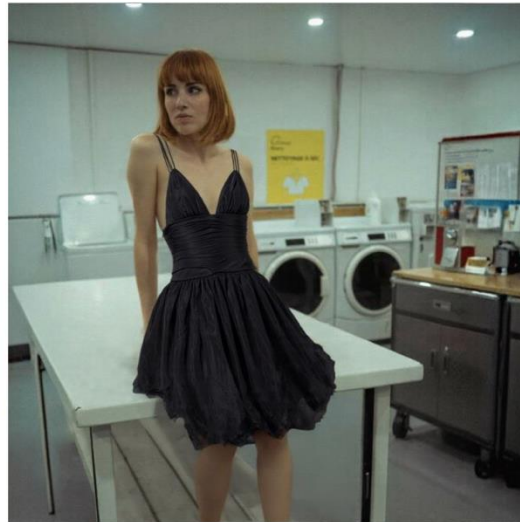
Generation in cross domains



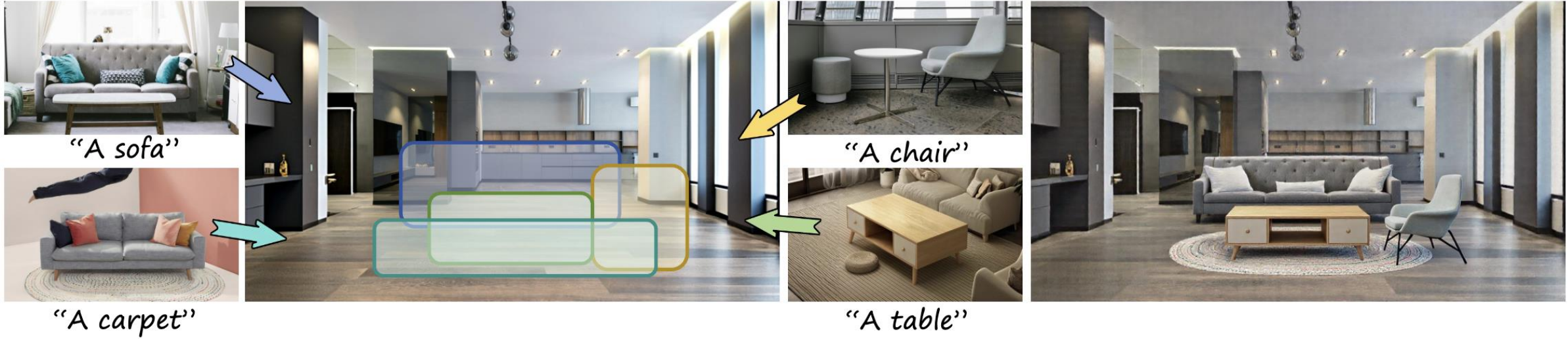


SIGGRAPH
ASIA 2025
HONG KONG 香港

Application-I: Virtual try-on



Application-II: Compositional Generation



Application-III: Partly insertion



Car wheels



Legs of the chair



The dome of the temple

Thanks.

In-Context Brush: Zero-Shot Customized Subject Insertion with Context-Aware Latent Space Manipulation

Project page: <https://yuci-gpt.github.io/In-Context-Brush>

Code: <https://github.com/ICTMCG/In-Context-Brush>



GENERATIVE

RENAISSANCE



SIGGRAPH 香港
ASIA 2025
HONG KONG 港