



Training-Free Text-Guided Color Editing with Multi-Modal Diffusion Transformer

Zixin Yin, Xili Dai, Ling-Hao Chen, Deyu Zhou, Jianan Wang, Duomin Wang
Gang Yu, Lionel M. Ni, Lei Zhang, Heung-Yeung Shum



PDF

Website

Text-Conditioned Color Editing



Benchmark Result

Method	SD3					FLUX.1-dev				
	Canny SSIM ↑	BG Preservation PSNR ↑	Clip Similarity SSIM ↑	Whole	Edited	Canny SSIM ↑	BG Preservation PSNR ↑	Clip Similarity SSIM ↑	Whole	Edited
Fix seed	0.5787	20.44	0.8411	29.17	27.54	0.7180	22.32	0.8877	27.72	25.76
FireFlow (Deng et al., 2025)	0.6078	19.19	0.8461	28.63	27.24	0.8322	35.87	0.9717	25.84	23.56
RF-Solver (Wang et al., 2025)	0.6711	23.30	0.8906	27.43	25.96	0.8394	36.25	0.9715	25.83	23.58
SDEdit (Meng et al., 2022)	0.6353	27.78	0.8699	25.98	24.47	0.8285	32.62	0.9605	25.54	23.17
DiTCtrl (Cai et al., 2025)	0.8119	35.40	0.9812	26.21	24.67	0.8306	34.48	0.9791	25.89	23.58
FlowEdit (Kulikov et al., 2024)	0.7852	34.24	0.9704	26.13	24.97	0.8639	32.64	0.9774	25.95	23.63
UniEdit-Flow (Jiao et al., 2025)	0.8016	36.31	0.9774	26.08	24.67	0.8498	37.57	0.9777	25.78	23.44
Ours	0.8473	42.93	0.9960	28.32	26.96	0.9196	39.49	0.9936	27.34	24.90

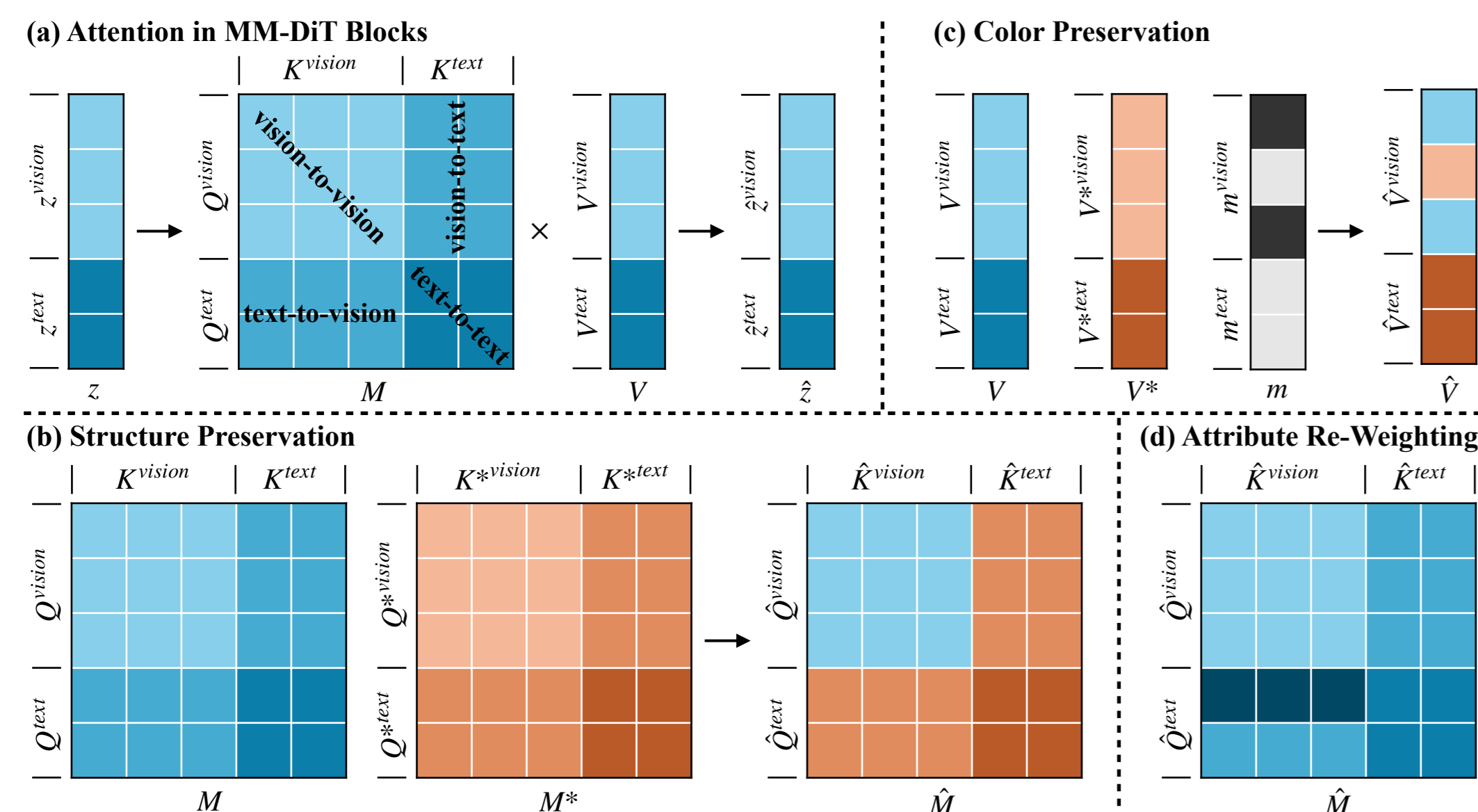
Method	SD3					FLUX.1-dev				
	Canny SSIM ↑	BG Preservation PSNR ↑	Clip Similarity SSIM ↑	Whole	Edited	Canny SSIM ↑	BG Preservation PSNR ↑	Clip Similarity SSIM ↑	Whole	Edited
FLUX.1 Kontext Max (Black Forest Labs, 2025)	0.7305	31.97	0.9254	28.93	27.30	0.7607	26.77	0.9165	28.36	26.10
GPT-4o Image Generation (OpenAI, 2025)	0.6240	23.69	0.8134	29.51	28.08	0.7431	23.71	0.8755	28.84	26.46
Ours	0.8473	42.93	0.9960	28.32	26.96	0.9196	39.49	0.9936	27.34	24.90

Method	PIE-Bench (Ju et al., 2023)					ColorCtrl-Bench				
	Canny SSIM ↑	BG Preservation PSNR ↑	Clip Similarity SSIM ↑	Whole	Edited	Canny SSIM ↑	BG Preservation PSNR ↑	Clip Similarity SSIM ↑	Whole	Edited
Fix seed	0.6228	20.73	0.8912	27.97	26.63	0.6671	19.49	0.8489	28.01	26.66
FireFlow (Deng et al., 2024)	0.7517	30.74	0.9605	25.42	24.18	0.7957	28.78	0.9323	25.26	24.12
RF-Solver (Wang et al., 2024)	0.7677	35.41	0.9730	25.40	24.08	0.8102	32.64	0.9562	25.36	24.16
SDEdit (Meng et al., 2021)	0.7880	35.46	0.9689	24.64	23.21	0.8281	34.30	0.9626	24.74	23.46
DiTCtrl (Cai et al., 2024)	0.6912	27.70	0.9435	27.02	25.57	0.7270	26.14	0.9165	26.54	25.14
UniEdit-Flow (Jiao et al., 2025)	0.7645	38.00	0.9772	25.35	24.13	0.8002	35.33	0.9613	25.39	24.14
Ours	0.8651	38.98	0.9916	27.12	25.96	0.8885	38.68	0.9893	27.32	26.16

Generalization to Other Models



Pipeline of ColorCtrl



Attribute Re-Weighting

