Perceptual-GS: Scene-adaptive Perceptual Densification for Gaussian Splatting

Hongbi Zhou[1], Zhangkai Ni[1]

[1]Tongji University

Background

■ 3DGS-based novel view synthesis:

- ☐ Densification operation in adaptive density control fails to distribute Gaussian primitives effectively in certain regions.

(a) Stump

(b) Flowers

(c) Treehill

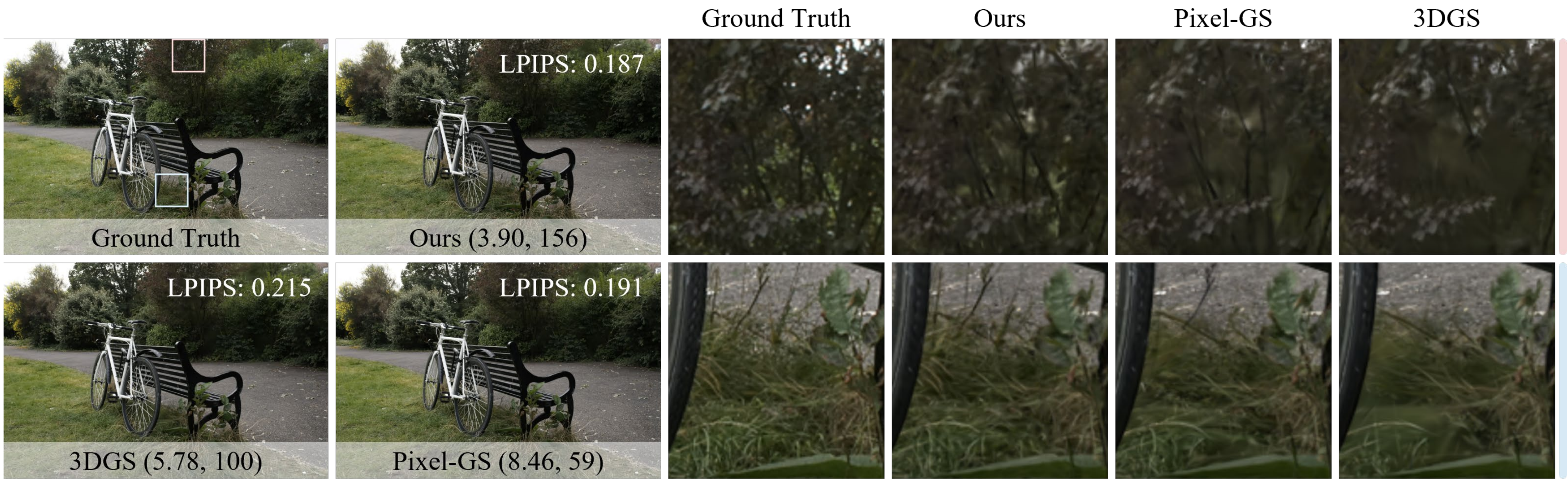
■ Attempts on optimizing densification operation of 3DGS:

- ☐ Enhancing the **calculation of position gradient** of Gaussian primitives.
- ☐ Proposing **additional metrics** to select primitives to be densified.

[1] Kerbl, Bernhard, et al. "3D Gaussian Splatting for Real-time Radiance Field Rendering." (TOG 2023).
[2] Barron, Jonathan T., et al. "Mip-NeRF 360: Unbounded Anti-aliased Neural Radiance Fields." (CVPR 2022).

Motivation

■ **Achieving high-fidelity reconstruction without largely increasing rendering overhead is challenging:**



[3] Zhang, Zheng, et al. "Pixel-GS: Density Control with Pixel-aware Gradient for 3D Gaussian Splatting." (ECCV 2024).

Motivation

■ **Limited utilization of human perception makes subtle structures overlooked and reduces perceptual quality:**

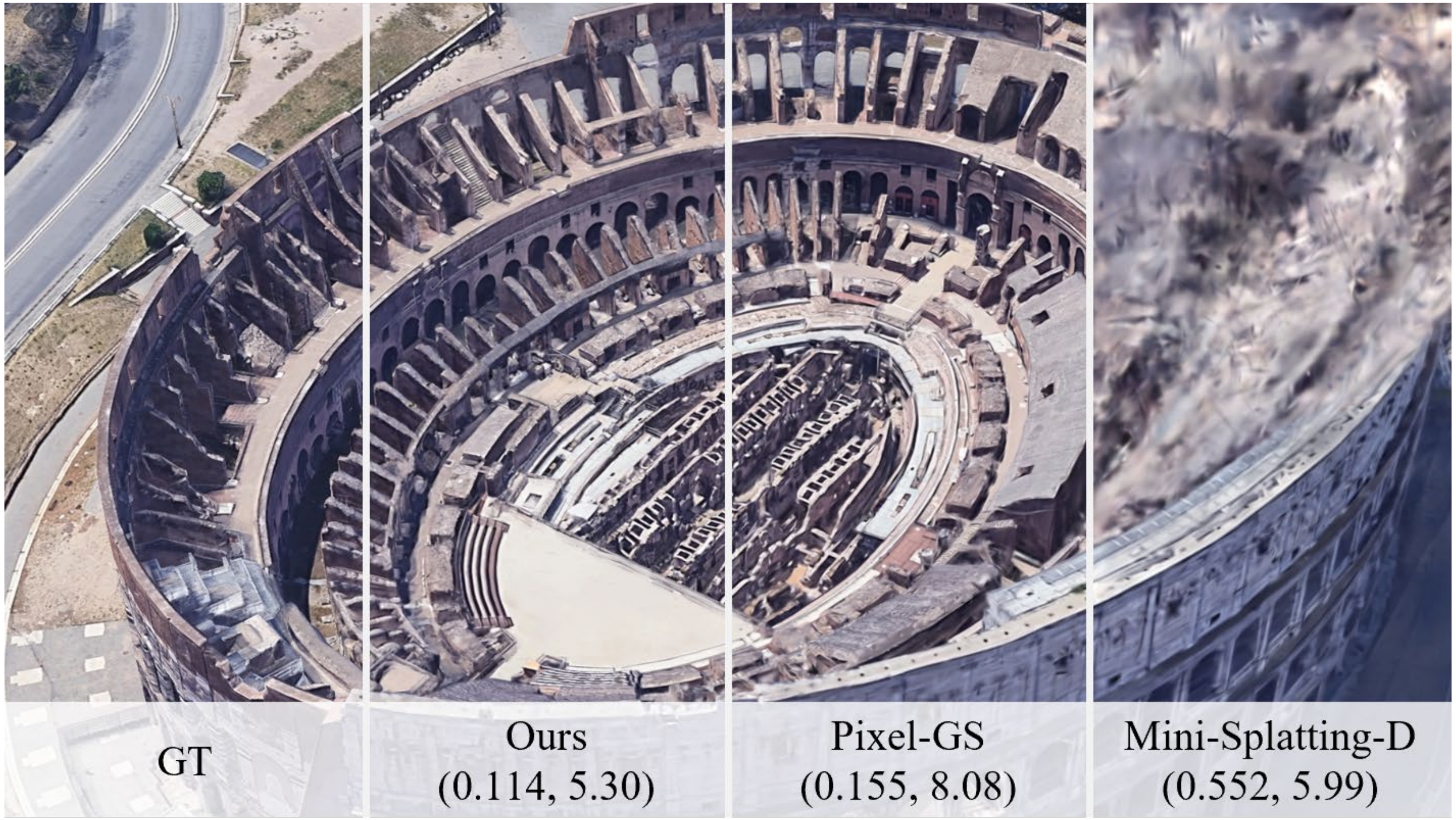
(a) 3DGS

(b) 3DGS+Sobel

(c) Ours

Motivation

■ The inability to adapt densification to scene-specific properties makes current approaches fail on certain scenes:



[4] Fang, Guangchi, and Bing Wang. "Mini-Splatting: Representing Scenes with a Constrained Number of Gaussians." (ECCV 2024).

Contribution

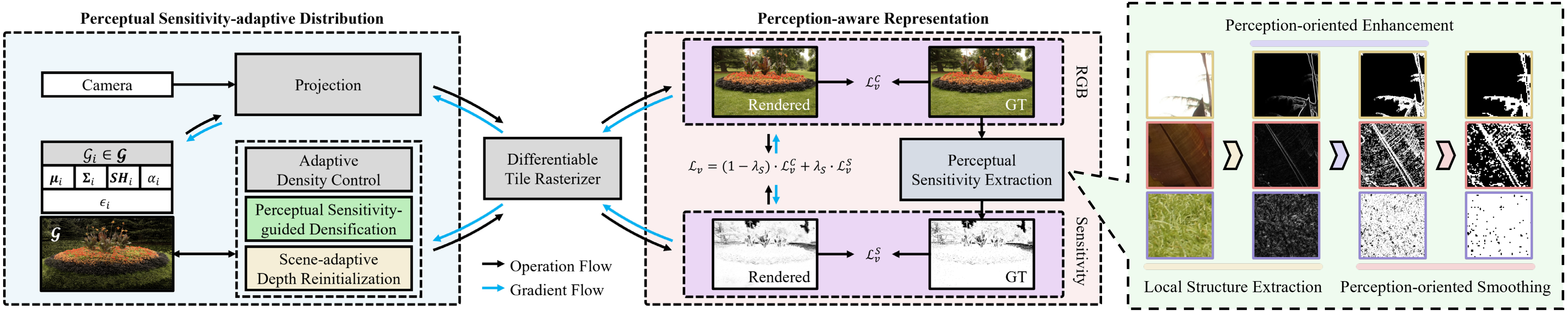
- **Perception-aware representation:** Allowing each Gaussian primitive to **adapt to** perceptual sensitivity across different spatial regions through **Perceptual Sensitivity Extraction** and **Dual-branch Rendering**.

- **Perceptual sensitivity-adaptive distribution:** Allocating Gaussian primitives **dynamically** based on perceptual sensitivity in different areas through **Perceptual Sensitivity-guided Densification** and **Scene-adaptive Depth Reinitialization**.

- **State-of-the-art performance and generalizability:** Perceptual-GS achieves **state-of-the-art performance** with **fewer** Gaussian primitives and can be integrated with **other** 3DGS-based methods.

Method

■ **Perceptual-GS: Integrating multi-view perceptual sensitivity into the training process to optimize the distribution of Gaussian primitives.**



□ Gaussians to be densified in Perceptual Sensitivity-guided densification:

$$\mathcal{G}_D = \{\mathcal{G}_i \mid \omega_i^{max} > \tau^{\omega} \wedge i \in [1,N]\} \cap (\mathcal{G}_h \cup \mathcal{G}_m),$$

$$\mathcal{G}_h = \{\mathcal{G}_i \mid \epsilon_i > \tau_h \wedge i \in [1,N]\},$$

$$\mathcal{G}_m = \{\mathcal{G}_i \mid \epsilon_i \in [\tau_l, \tau_h] \wedge i \in [1,N]\},$$

$$\omega_i^{max} = MAX(\{\textstyle\sum_{\mathbf{u}\in \boldsymbol{pix}_v} \omega_i^v(\mathbf{u}) \mid v \in \mathbf{V}\}).$$

□ Sensitivity map rendering:

$$\mathcal{R}_v^S(\boldsymbol{u}) = \sum_i^N \omega_i^v(\boldsymbol{u})\sigma(\epsilon_i)$$

□ Opacity decline for clone operation:

$$\hat{\alpha} = 1 - \sqrt{1 - OD(\alpha)}$$

同济大学
TONGJI UNIVERSITY

Experiments

■ Results on reconstruction quality:

a. Quantitative results in reconstruction quality.

Method	Mip-NeRF 360			Tanks & Temples			Deep Blending			BungeeNeRF		
	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓
3DGS*	27.71	0.826	0.202	23.61	0.845	0.178	29.54	0.900	0.247	27.64	0.912	0.100
Pixel-GS*	27.85	0.834	0.176	23.71	0.853	0.152	28.92	0.893	0.250	OOM in 1 scene		
Mini-Splatting-D	27.51	0.831	0.176	23.23	0.853	0.140	29.88	0.906	0.211	25.58	0.861	0.149
Taming-3DGS	27.79	0.822	0.205	24.04	0.851	0.170	30.14	0.907	0.235	OOM in 2 scenes		
Ours	28.01	0.839	0.172	23.90	0.857	0.151	29.94	0.907	0.231	27.86	0.918	0.095

b. Qualitative comparison results on BungeeNeRF.

Ground Truth   Ours   Taming-3DGS

Pixel-GS   Mini-Splatting-D   3DGS

[5] Mallick, Saswat Subhajyoti, et al. "Taming 3DGS: High-quality Radiance Fields with Limited Resources." (SIGGRAPH Asia 2024).
[6] Wang, Zhou, et al. "Image Quality Assessment: from Error Visibility to Structural Similarity." (TIP 2004).
[7] Zhang, Richard, et al. "The Unreasonable Effectiveness of Deep Features as a Perceptual Metric." (CVPR 2018).
[8] Knapitsch, Arno, et al. "Tanks and Temples: Benchmarking Large-scale Scene Reconstruction." (TOG 2017).
[9] Hedman, Peter, et al. "Deep Blending for Free-viewpoint Image-based Rendering." (TOG 2018).
[10] Xiangli, Yuanbo, et al. "BungeeNeRF: Progressive Neural Radiance Field for Extreme Multi-scale Scene Rendering." (ECCV 2022).

Experiments

■ Results on reconstruction efficiency:

c. Quantitative results in reconstruction efficiency.

Method	Mip-NeRF 360		Tanks & Temples		Deep Blending		BungeeNeRF	
	#G↓	FPS↑	#G↓	FPS↑	#G↓	FPS↑	#G↓	FPS↑
3DGS*	3.14M	193	1.83M	247	2.81M	194	6.92M	69
Pixel-GS*	5.23M	105	4.49M	101	4.63M	114	OOM in 1 scene	
Mini-Splatting-D	4.69M	120	4.28M	115	4.63M	159	6.08M	86
Taming-3DGS	3.31M	122	1.84M	149	2.81M	130	OOM in 2 scenes	
Ours	2.69M	166	1.72M	218	2.86M	178	4.97M	89

d. Effect of Opacity Decline and Dual-branch Rendering.

	PSNR↑	SSIM↑	LPIPS↓	#G↓
3DGS*	27.71	0.826	0.202	3.14M
+OD	27.74	0.825	0.207	2.22M
+OD +DBR	27.69	0.822	0.212	1.94M

Experiments

■ Results of the proposed method integrating with different models:

e. Quantitative results on Mip-NeRF 360, Tanks & Temples and Deep Blending.

Method	Mip-NeRF 360				Tanks & Temples				Deep Blending			
	PSNR↑	SSIM↑	LPIPS↓	#G↓	PSNR↑	SSIM↑	LPIPS↓	#G↓	PSNR↑	SSIM↑	LPIPS↓	#G↓
3DGS*	27.71	0.826	0.202	3.14M	23.61	0.845	0.178	1.83M	29.54	0.900	0.247	2.81M
w/ Ours	28.01	0.839	0.172	2.69M	23.90	0.857	0.151	1.72M	29.94	0.907	0.231	2.86M
Δ	+0.30	+0.013	-0.030	-0.45M	+0.29	+0.012	-0.027	-0.11M	+0.40	+0.007	-0.016	+0.05M
Pixel-GS*	27.85	0.834	0.176	5.23M	23.71	0.853	0.152	4.49M	28.92	0.893	0.250	4.63M
w/ Ours	28.01	0.841	0.167	3.37M	23.95	0.859	0.142	2.96M	29.71	0.901	0.233	3.59M
Δ	+0.16	+0.007	-0.009	-1.86M	+0.24	+0.006	-0.010	-1.53M	+0.79	+0.008	-0.017	-1.04M

g. Quantitative results on 24-view Mip-NeRF 360.

	PSNR↑	SSIM↑	LPIPS↓
CoR-GS*	22.26	0.664	0.341
w/Ours	22.42	0.681	0.281
Δ	+0.16	+0.017	-0.060

f. Quantitative results on BungeeNeRF.

Method	BungeeNeRF				Pompidou				Chicago				Amsterdam			
	PSNR↑	SSIM↑	LPIPS↓	#G↓	PSNR↑	SSIM↑	LPIPS↓	#G↓	PSNR↑	SSIM↑	LPIPS↓	#G↓	PSNR↑	SSIM↑	LPIPS↓	#G↓
3DGS*	27.64	0.912	0.100	6.92M	27.00	0.916	0.095	9.11M	27.97	0.927	0.086	6.32M	27.60	0.913	0.100	6.19M
w/ Ours	27.86	0.918	0.095	4.97M	27.18	0.922	0.089	6.12M	28.39	0.933	0.081	4.48M	27.89	0.922	0.087	4.96M
Δ	+0.22	+0.006	-0.005	-1.95M	+0.18	+0.006	-0.006	-2.99M	+0.42	+0.006	-0.005	-1.84M	+0.29	+0.009	-0.013	-1.23M
Pixel-GS*	OOM in 1 scene				OOM				27.52	0.921	0.090	9.76M	27.76	0.916	0.095	10.26M
w/ Ours	27.64	0.913	0.100	5.92M	27.01	0.918	0.092	7.39M	28.36	0.930	0.081	5.58M	27.98	0.922	0.085	6.60M
Δ	—	—	—	—	—	—	—	—	+0.84	+0.009	-0.009	-4.18M	+0.22	+0.006	-0.010	-3.66M

[11] Zhang, Jiawei, et al. "CoR-GS: Sparse-view 3D Gaussian Splatting via Co-regularization." (ECCV 2024).

Experiments

Ablation study:

h. Ablation studies on different modules.

	PSNR↑	SSIM↑	LPIPS↓	#G↓
FULL	28.01	0.839	0.172	2.69M
3DGS*	27.71	0.826	0.202	3.14M
w/o PE	27.74	0.825	0.204	2.09M
w/o HD	27.74	0.826	0.204	2.02M
w/o MD	27.86	0.831	0.179	2.56M
w/o SDR	27.93	0.832	0.176	2.68M
w/o OD	27.99	0.839	0.172	3.25M

i. Ablation studies on hyperparameters.

H.P.	Value	PSNR↑	SSIM↑	LPIPS↓	#G↓
λ_S	**0.1**	28.01	0.839	0.172	2.69M
	0.3	27.82	0.835	0.181	2.10M
	0.5	27.48	0.823	0.196	1.92M
τ_h^ω	10	28.05	0.841	0.166	3.61M
	15	28.00	0.840	0.169	3.09M
	25	28.01	0.839	0.172	2.69M
τ_m^ω	**10**	28.01	0.839	0.172	2.69M
	15	27.98	0.838	0.173	2.65M
	25	27.97	0.838	0.174	2.63M
$Iter_h$	**1000**	28.01	0.839	0.172	2.69M
	1500	27.95	0.838	0.174	2.57M
	2000	27.93	0.837	0.175	2.52M
$Iter_m$	1000	27.92	0.839	0.172	2.70M
	1500	28.01	0.839	0.172	2.69M
	2000	27.98	0.839	0.173	2.66M

同濟大學
TONGJI UNIVERSITY

Code: https://github.com/eezkni/Perceptual-GS

ICML
International Conference
On Machine Learning