Splat Feature Solver

Butian Xiong's Website

Splat Feature Solver

Slide

Butian Xiong

Content Summary

- Splats Based Rendering

- Recent Progress in Feature Lifting (Unsolved Problem and Trend)
    - Training Based Method
    - Grouping Based Method
    - Heuristic Solver

- Splat Feature Solver
    - Problem Definition
    - Intuition
    - Mathematical Foundation
    - Inverse Problem (How originally figure out the solution)

- How far our solver can go, and what is the possible future work

Splats Method Reviewing

$$\omega_p = \sigma_p \prod_{j=1}^{p-1}(1-\sigma_j), \sigma_p = \alpha_p \delta_{pr}, \alpha_p = \frac{1}{1+e^{-\theta}} \; C_r = \sum_p \omega_p c_p + \left(1 - \sum_p \omega_p\right) C_b$$

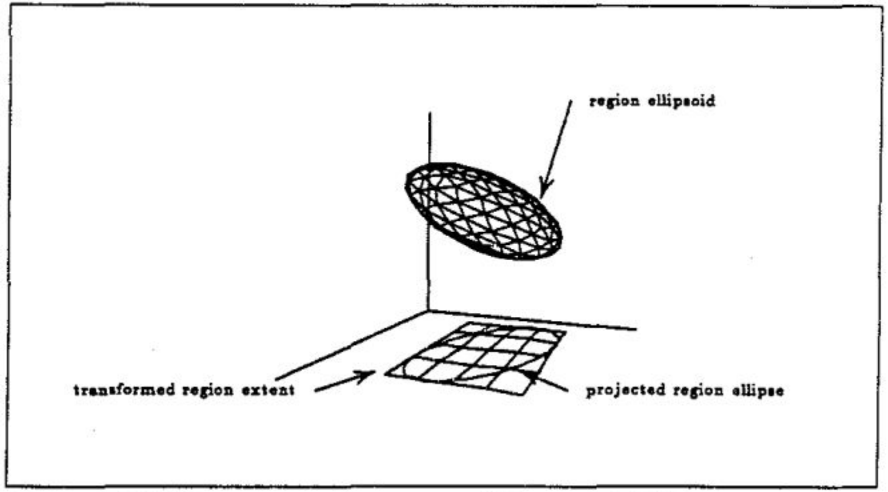


Figure 3.10: View-Transformed Kernel.

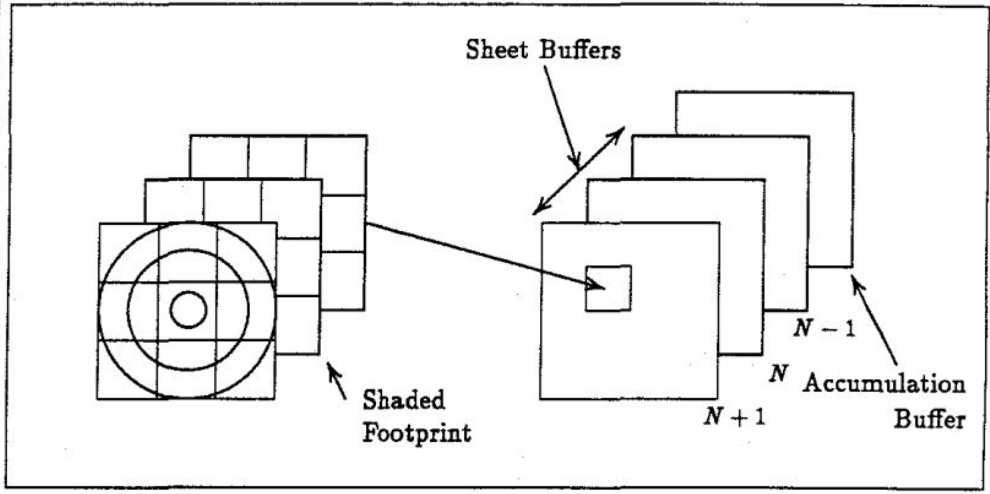


Figure 5.01: Ideal Splatting Method.

Westover, Lee Alan. *Splatting: a parallel, feed-forward volume rendering algorithm*. The University of North Carolina at Chapel Hill, 1991.

Kernel Evaluation Back in 1991 and 2025

1991

1. **Box Kernel (nearest neighbor / piecewise constant)**
2. **Cone / Pyramid Kernel (tent function)**
3. **Gaussian Kernel (radially symmetric)**
4. **Ellipsoidal Footprints (anisotropic versions)**

2025

1. **3D Gaussian Splats**
2. **2D Gaussian Splats**
3. **Deformable Beta Splats**
4. **Deformable Radial Kernel**

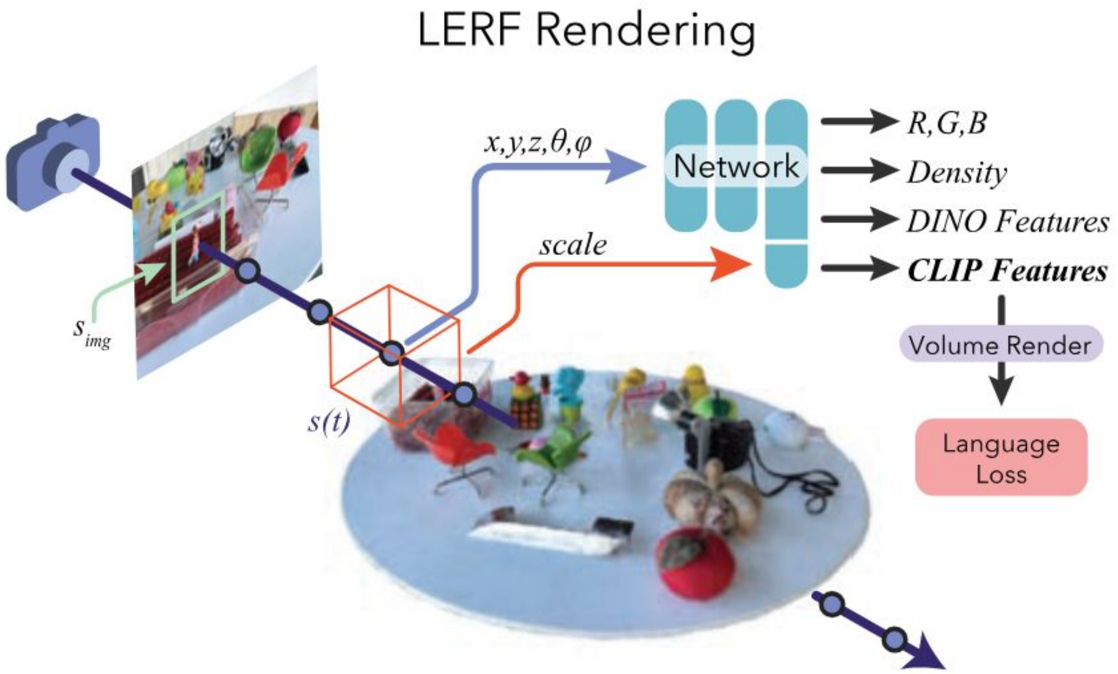
Liu, Rong, et al. "Deformable beta splatting." *Proceedings of the Special Interest Group on Computer Graphics and Interactive Techniques Conference Conference Papers*. 2025.

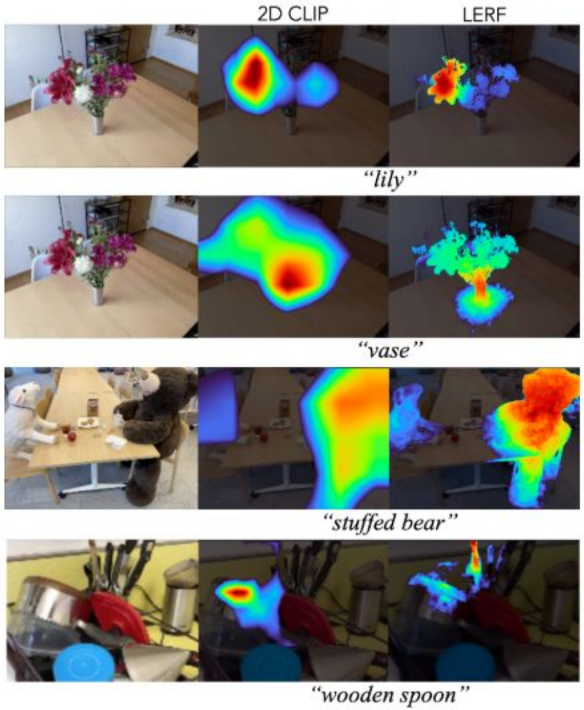
Huang, Yi-Hua, et al. "Deformable Radial Kernel Splatting." *Proceedings of the Computer Vision and Pattern Recognition Conference*. 2025.

Feature Lifting

- Training Based Method
  - NeRF Based
  - Splats Based
  - Feature Focused?!
- Grouping Based
- Heuristic Based

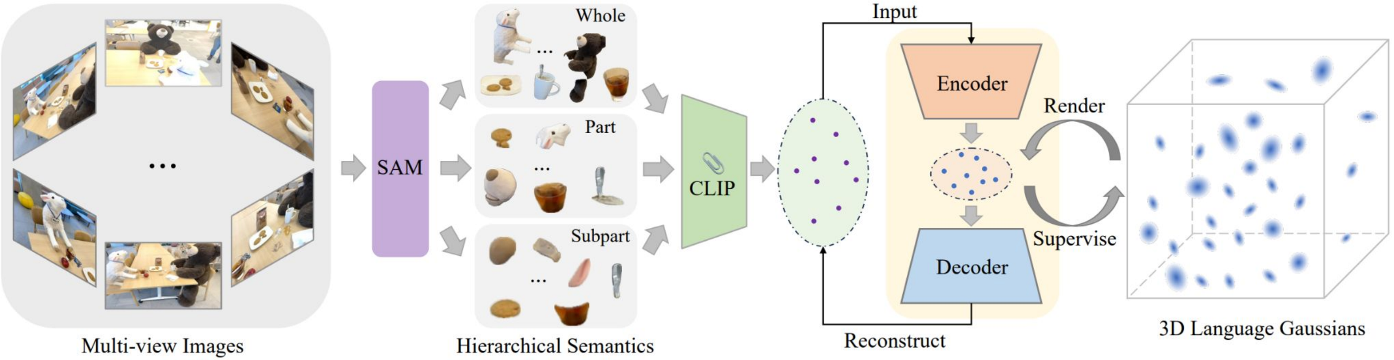
Training Based (Pre SAM and GS Era)

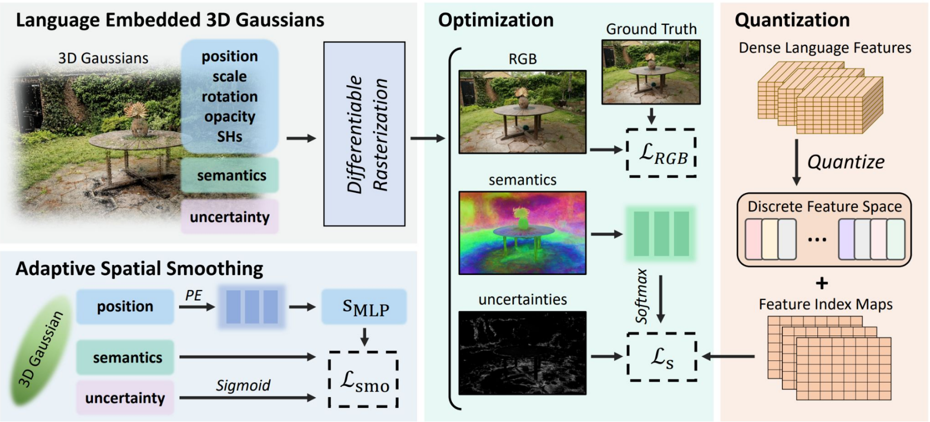




Kerr, Justin, et al. "Lerf: Language embedded radiance fields." *Proceedings of the IEEE/CVF international conference on computer vision*. 2023.

Training Based (GS Based)





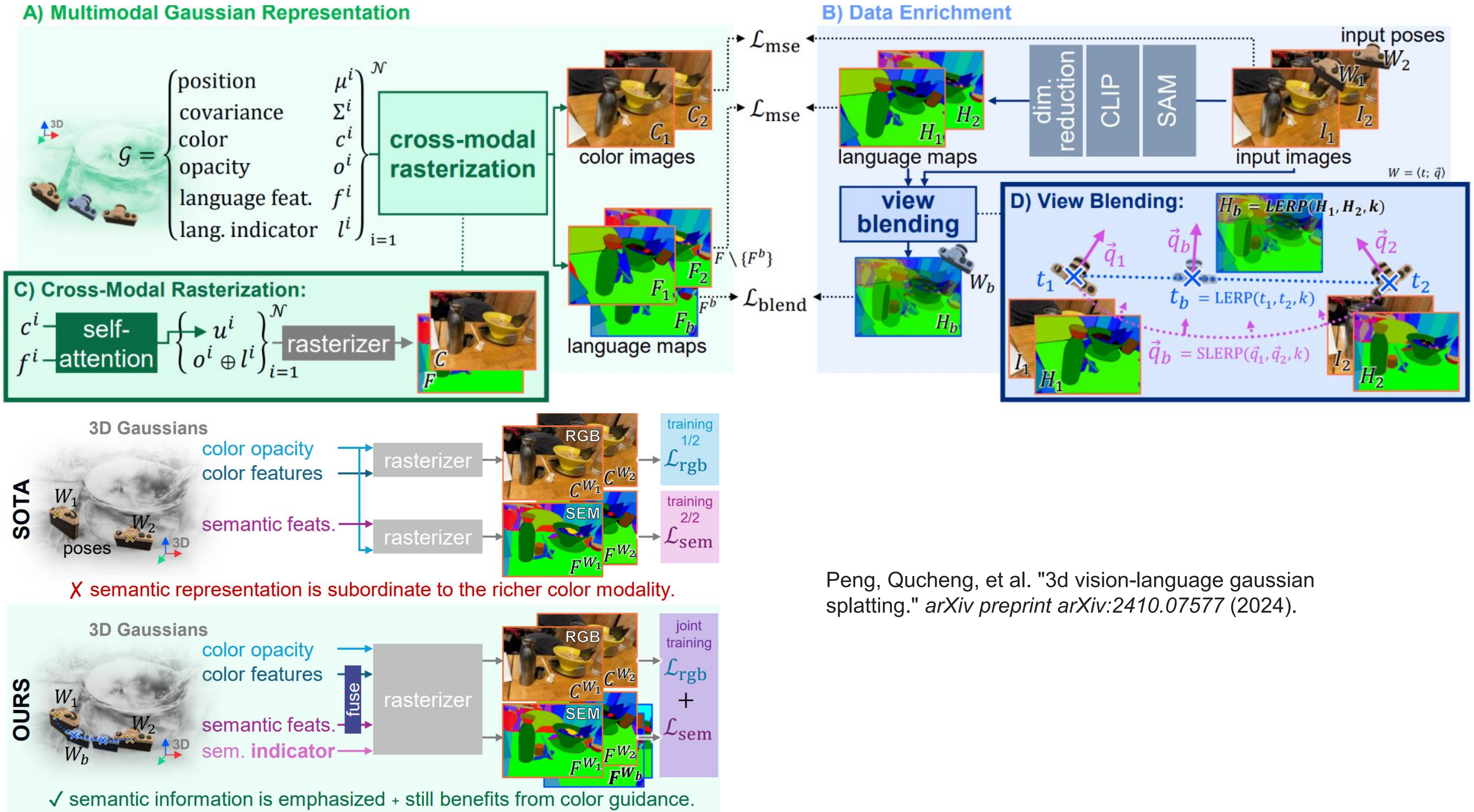
Zhou, Shijie, et al. "Feature 3dgs: Supercharging 3d gaussian splatting to enable distilled feature fields." *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2024.

Zuo, Xingxing, et al. "Fmgs: Foundation model embedded 3d gaussian splatting for holistic 3d scene understanding." *International Journal of Computer Vision* 133.2 (2025): 611-627.

Shi, Jin-Chuan, et al. "Language embedded 3d gaussians for open-vocabulary scene understanding." *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2024.

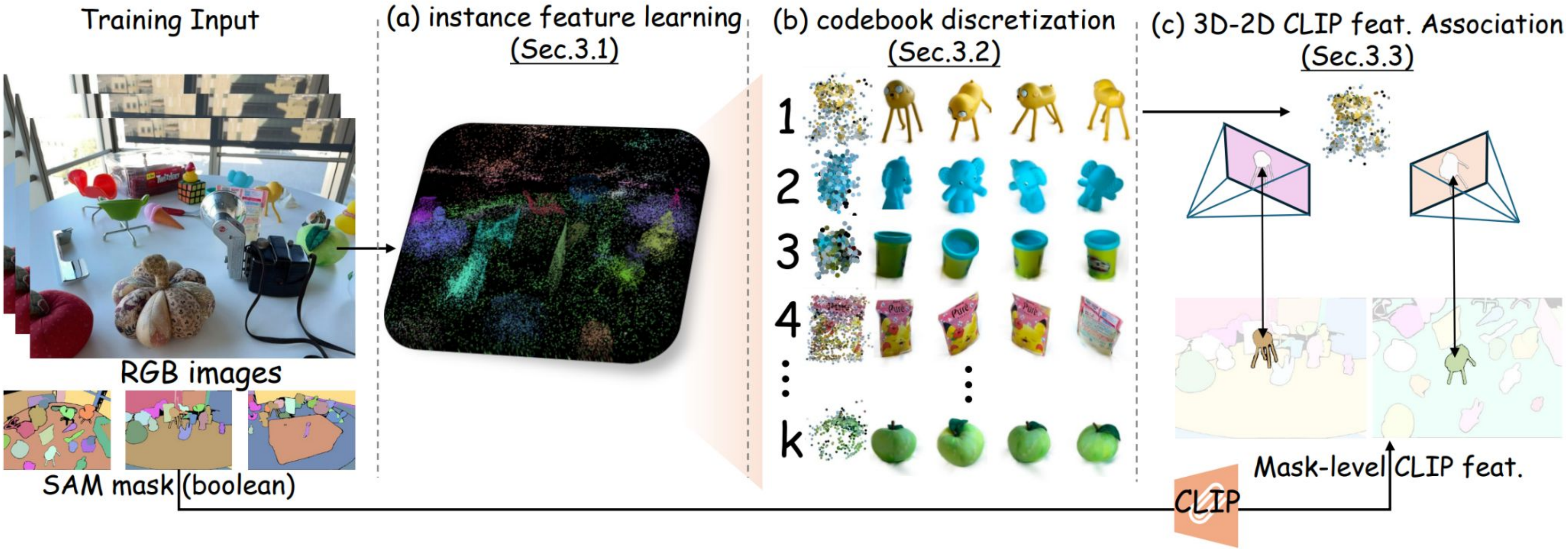
Qin, Minghan, et al. "Langsplat: 3d language gaussian splatting." *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2024.

Feature Focus Method



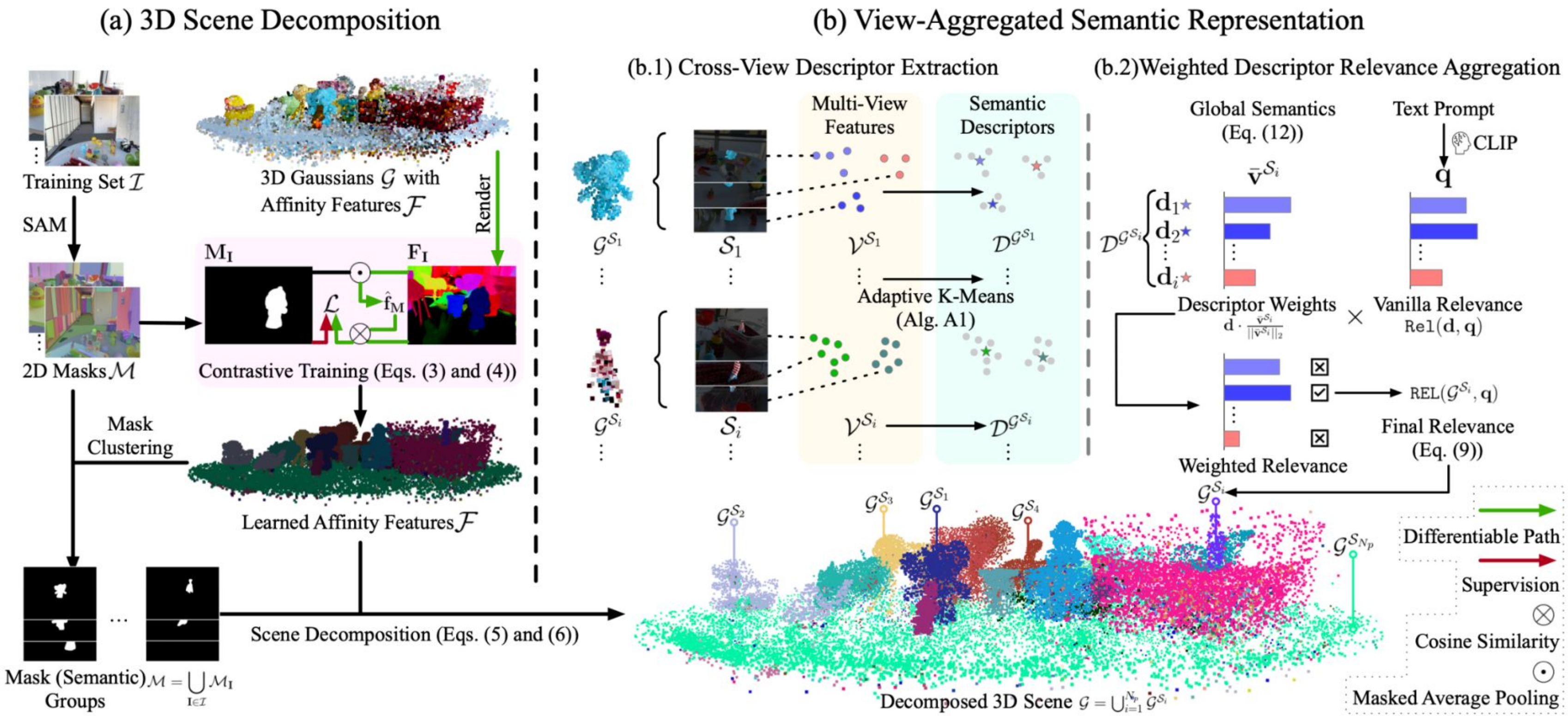
Peng, Qucheng, et al. "3d vision-language gaussian splatting." *arXiv preprint arXiv:2410.07577* (2024).

Grouping Based Method - OpenGaussian



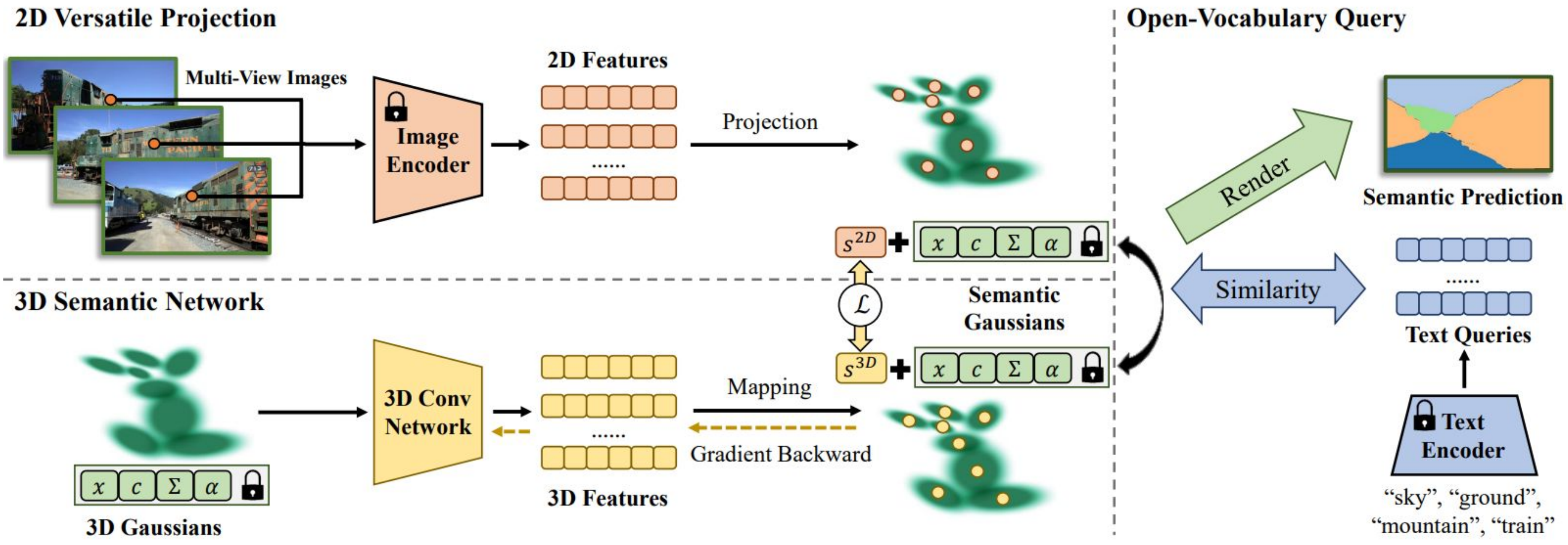
Wu, Yanmin, et al. "Opengaussian: Towards point-level 3d gaussian-based open vocabulary understanding." *Advances in Neural Information Processing Systems* 37 (2024): 19114-19138.

Grouping Based Method - LAGA



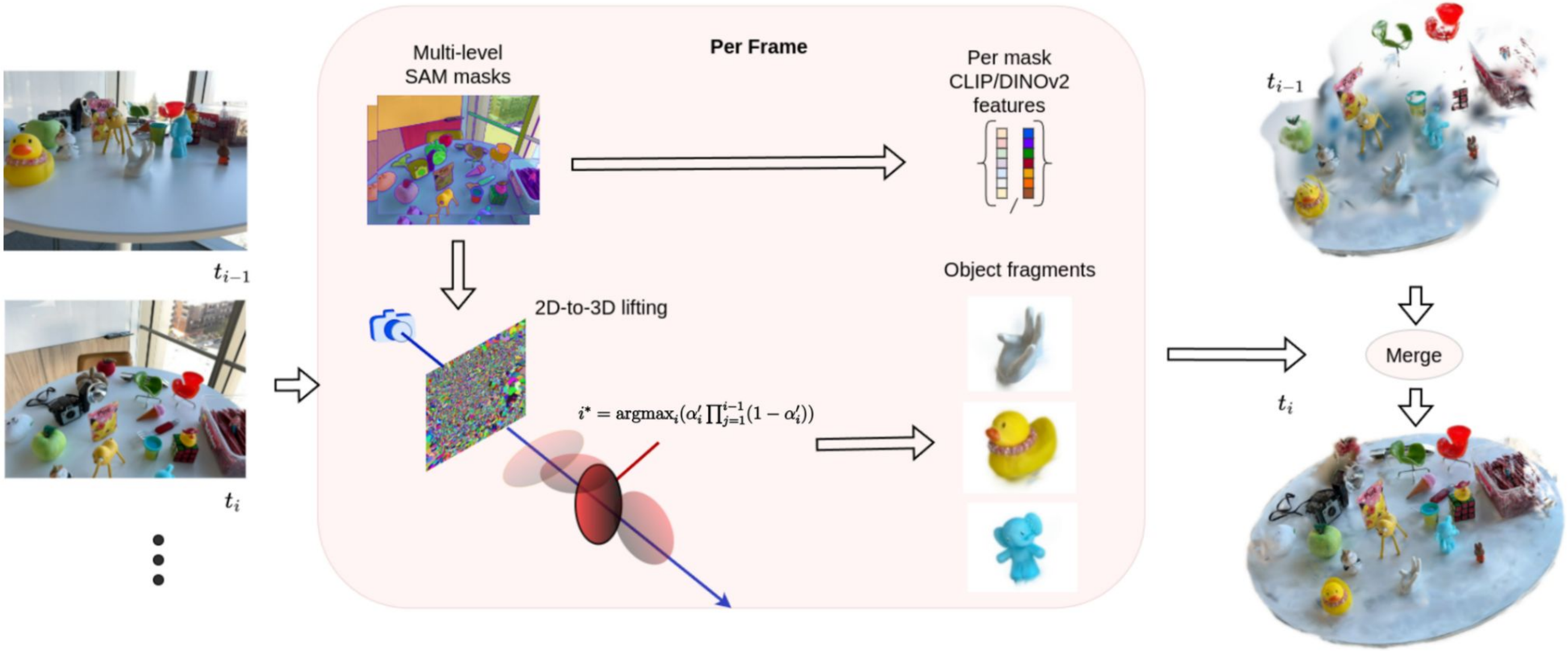
Cen, Jiazhong, et al. "Tackling View-Dependent Semantics in 3D Language Gaussian Splatting." *arXiv preprint arXiv:2505.24746* (2025).

Heuristic Method (AVG Pooling)



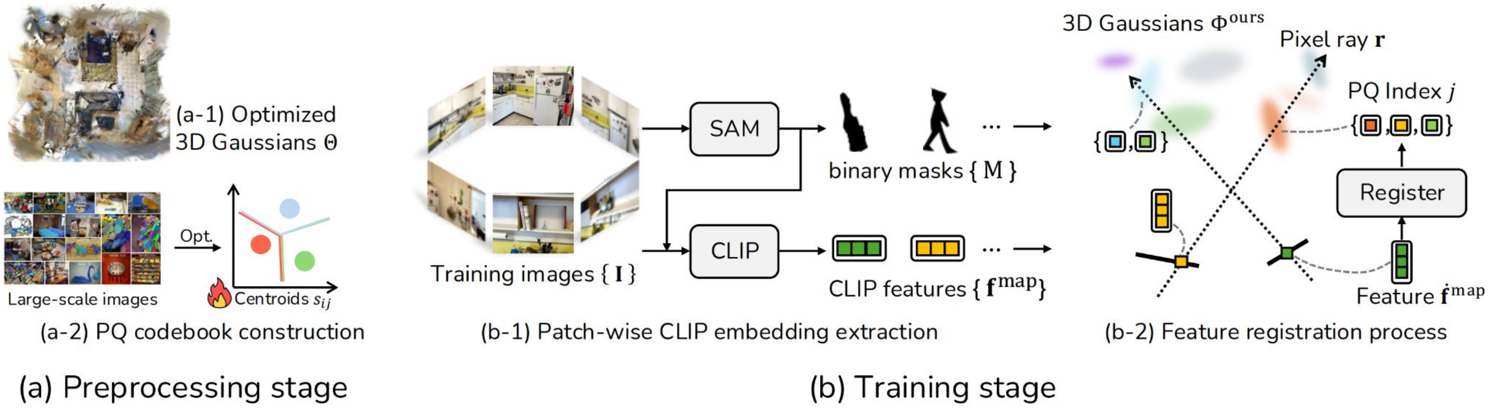
Guo, Jun, et al. "Semantic gaussians: Open-vocabulary scene understanding with 3d gaussian splatting." *arXiv preprint arXiv:2403.15624* (2024).

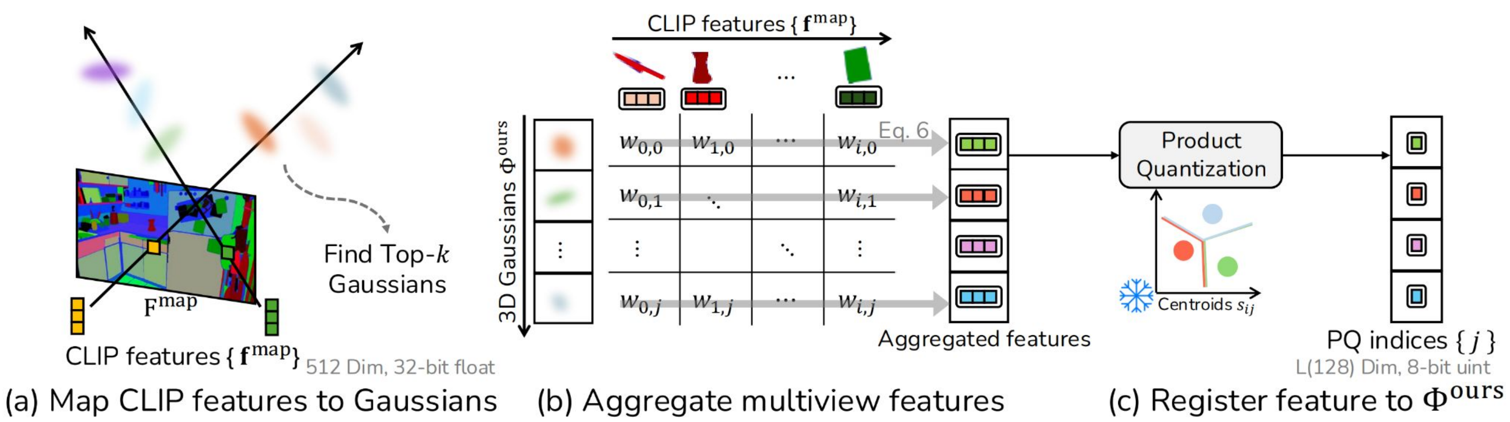
Heuristic Method (ArgMax)



Chacko, Rohan, et al. "Lifting by gaussians: A simple, fast and flexible method for 3d instance segmentation." *2025 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*. IEEE, 2025.

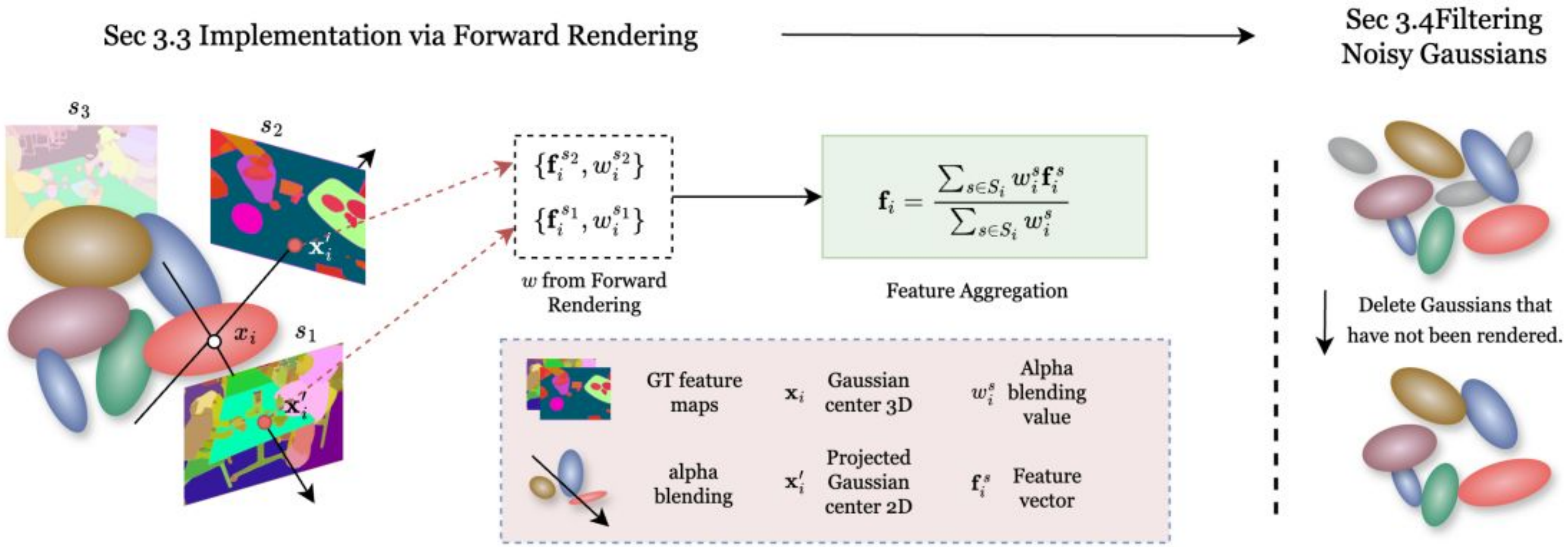
Heuristic Method (Nearest K weighted SUM)





Jun-Seong, Kim, et al. "Dr. splat: Directly referring 3d gaussian splatting via direct language embedding registration." *Proceedings of the Computer Vision and Pattern Recognition Conference*. 2025.

Heuristic Method (Weighted Average)



Cheng, Jiahuan, et al. "Occam's LGS: A Simple Approach for Language Gaussian Splatting." *arXiv e-prints* (2024): arXiv-2412.

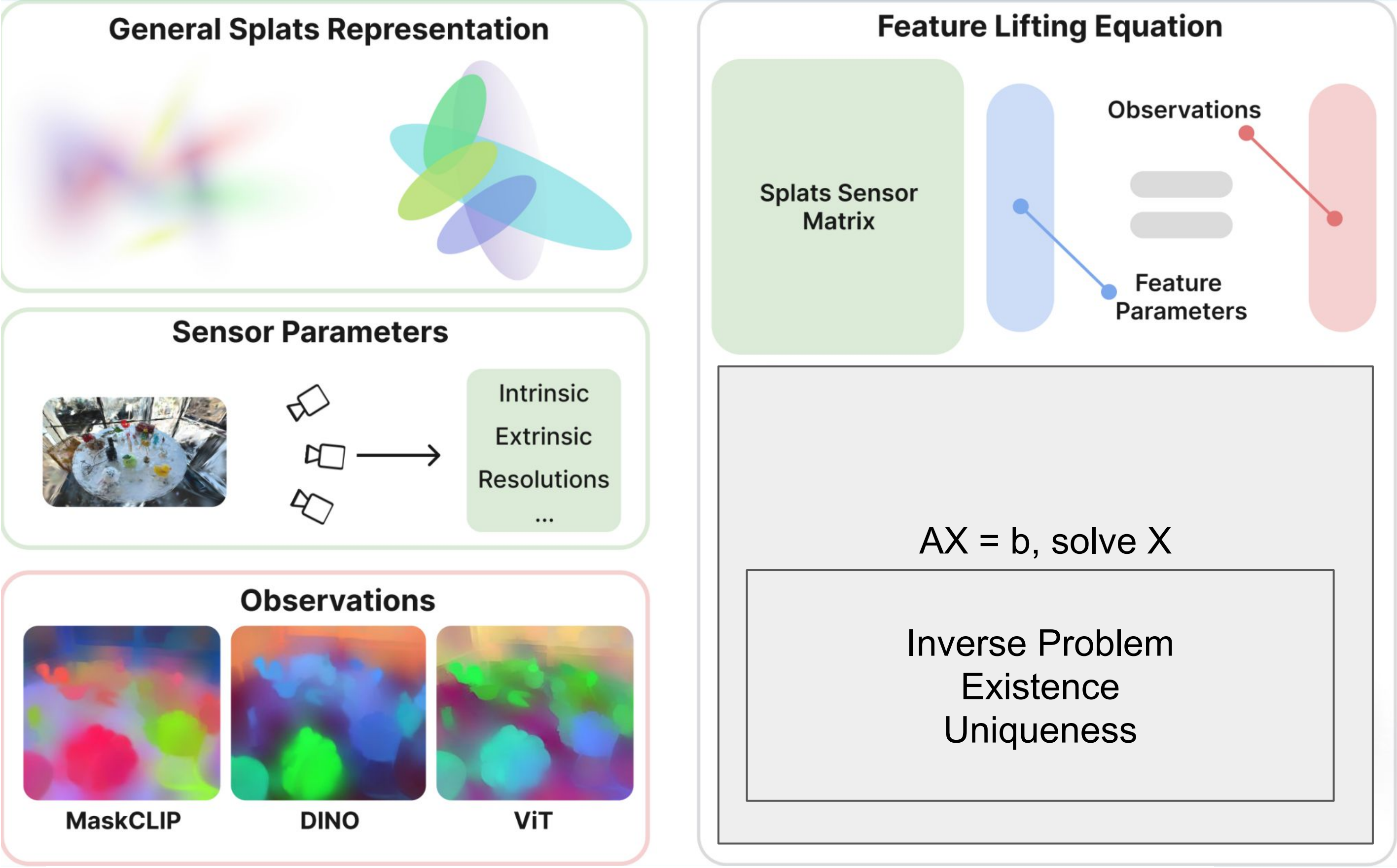
Potential Problem of Heuristic Based Method

- No Theory Foundation (Grounded Mathematical Description)
- No analysis on how good the method is
- Does not know how general this method can be applied
    - Feature Agnostic?
    - Kernel Agnostic?

Problem Definition: Feature Lifting

3DGS
2DGS
DBS

Weight



AX = b, solve X

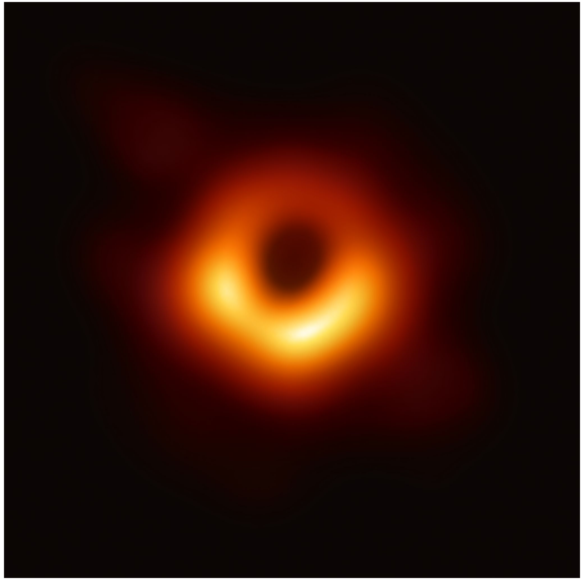
Inverse Problem
Existence
Uniqueness

Least Square

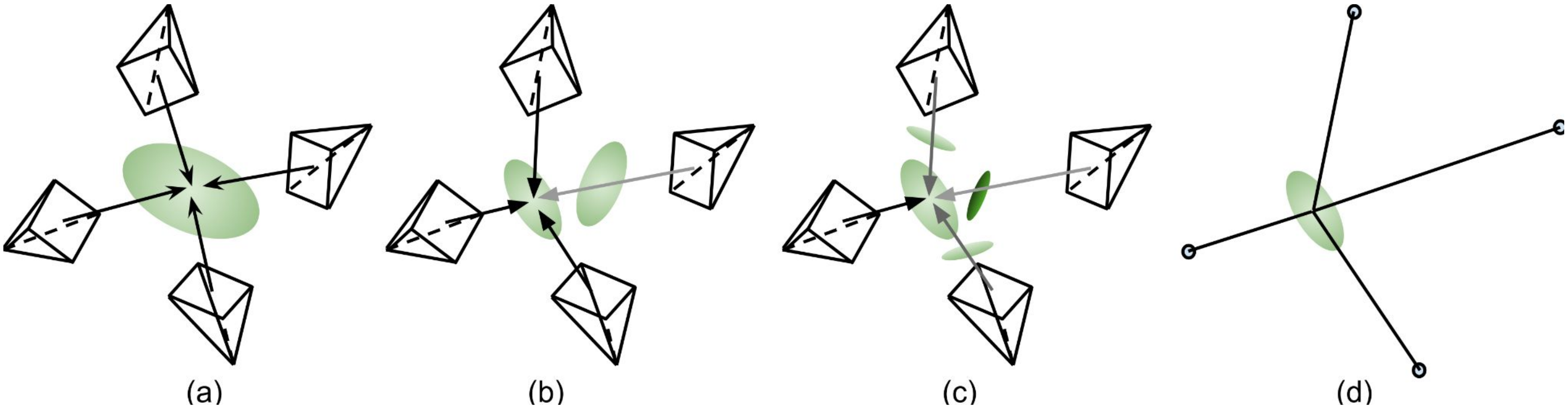
$$A_{ij} = \omega_{rp}, \quad x_j = c_p, \quad B_i = \hat{C}_r, \quad (i = r, j = p) \quad (4)$$

Inverse Problem Examples and Definition

1. Existence
2. Uniqueness



Intuition



Equal Weight

Different Weight

$$D^{\frac{1}{2}}(A^T A) = \sqrt{D(A^T A)}$$

$$x = D^{-\frac{1}{2}}(A^T A)e \times D^{\frac{1}{2}}(A^T A)B \quad x_j = \frac{\sum_i A_{ij} B_i}{\sum_i A_{ij}}$$

Mathematical Proof of Bounding (Well Posedness)

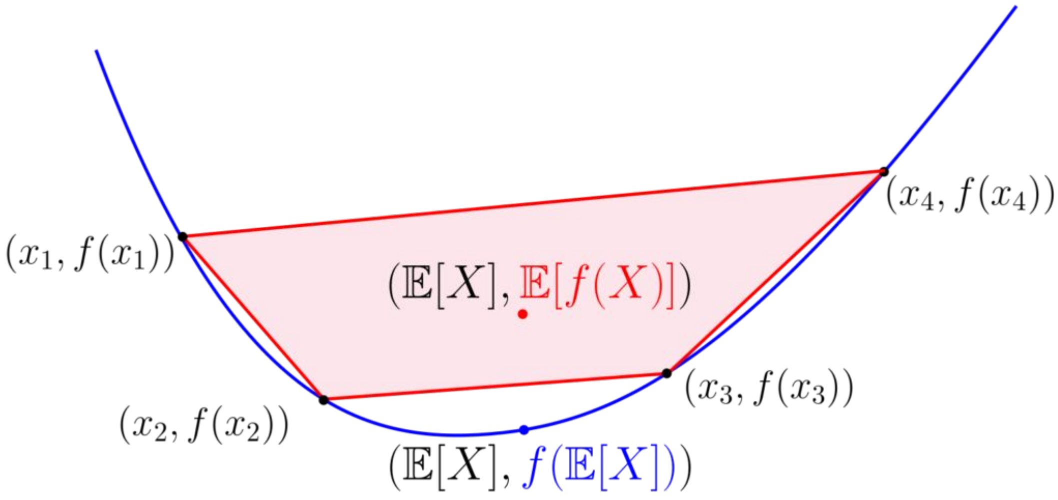
$$\mathcal{L}(x) = \sum_{i=1}^{R} \left\| \sum_{j=1}^{P} A_{ij} x_j - B_i \right\|$$

$$\mathcal{J}(x) = \sum_{i} \sum_{j} A_{ij} \left\| x_j - B_i \right\|$$

$$\sum_j A_{ij} = 1 \quad \Rightarrow \quad \mathcal{L}(x) \overset{\text{Jensen}}{\leq} \mathcal{J}(x)$$

$$\frac{\partial \mathcal{J}}{\partial x_j} = \sum_i A_{ij} \left(x_j - B_i \right)$$

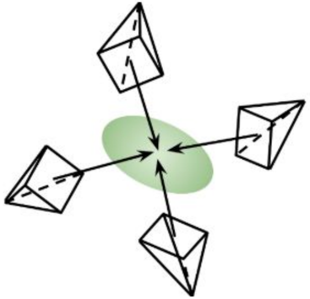
$$\frac{\partial \mathcal{J}}{\partial x_j} = 0 \quad \Rightarrow \quad x'_j = \frac{\sum_i A_{ij} B_i}{\sum_i A_{ij}}$$



!?

What is Beta?

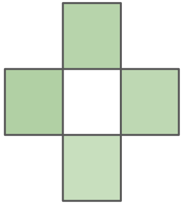
Optimal

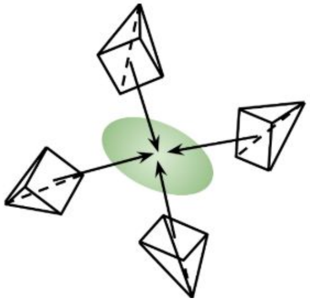


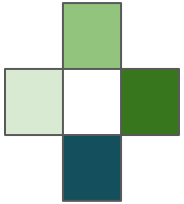
$$\Delta_{ij} = \|\hat{x}_j - B_i\|, \quad \mu_i = \sum_j A_{ij}\Delta_{ij}$$

$$\sigma_i^2 = \sum_j A_{ij}\left(\Delta_{ij}^2 - \mu_i^2\right), \beta_i = \frac{\sigma_i^2}{\mu_i^2}, \beta = \max_i(\beta_i)$$

β_i







β_i

Mathematical Proof of Bounding (Well Posedness)

$$\mathcal{L}(x') \leq \mathcal{J}(x') \leq \mathcal{J}(\hat{x})$$

$$\mathcal{J}(\hat{x}) = \sum_i \sum_j A_{ij} \Delta_{ij}^2$$

$$= \sum (1 + \beta_i) \mu_i^2 \leq (1 + \beta) \mathcal{L}(\hat{x})$$

$$\Rightarrow \mathcal{L}(x') \leq (1 + \beta) \mathcal{L}(\hat{x})$$

Mathematical Proof of Bounding (Well Posedness)

$$\mathcal{L}(x') \leq \mathcal{J}(x') \leq \mathcal{J}(\hat{x})$$

$$\mathcal{J}(\hat{x}) = \sum_i \sum_j A_{ij} \Delta_{ij}^2$$

$$= \sum (1 + \beta_i)\mu_i^2 \leq (1 + \beta)\mathcal{L}(\hat{x})$$

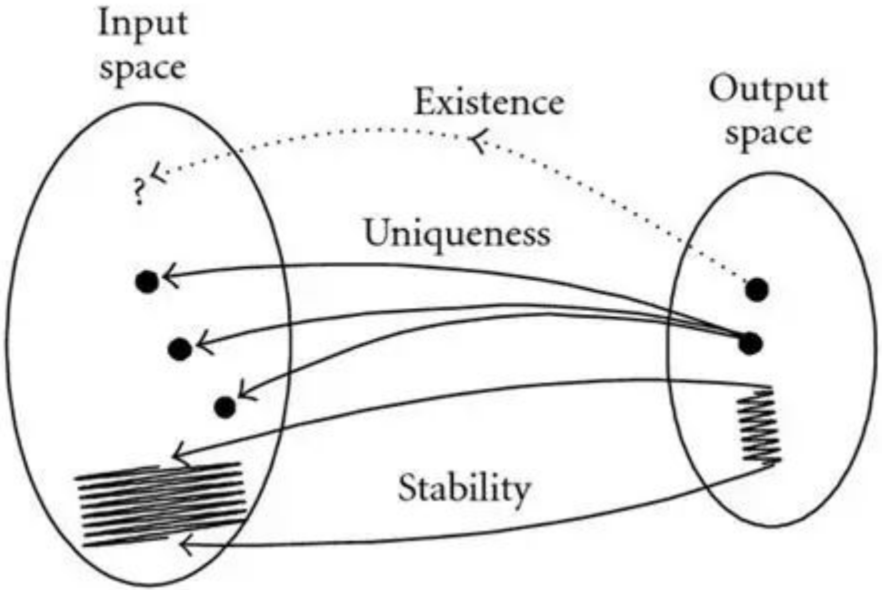
$$\Rightarrow \mathcal{L}(x') \leq (1 + \beta)\mathcal{L}(\hat{x})$$

Heuristic Method (AVG Pooling)

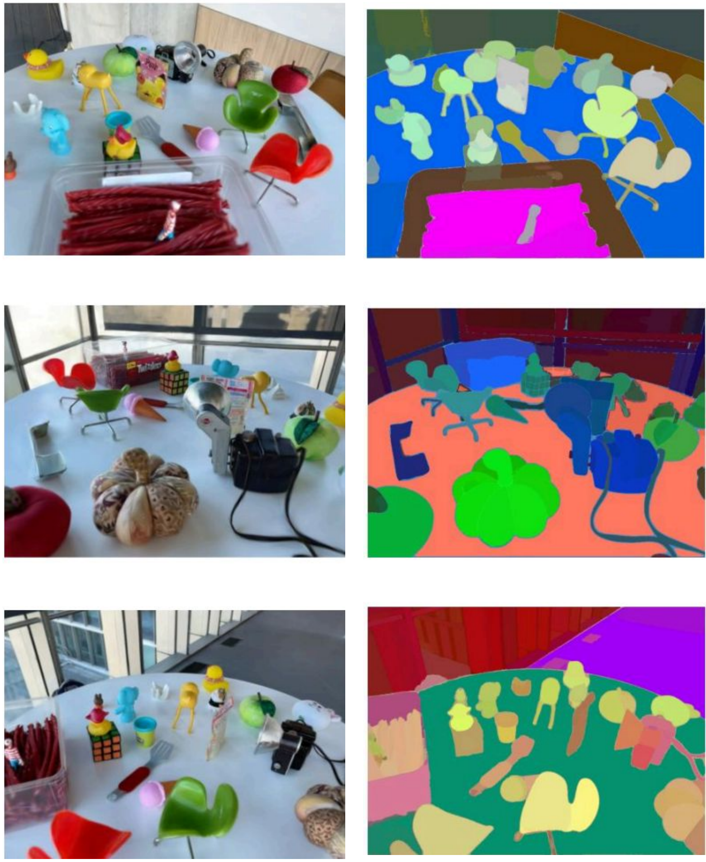
Heuristic Method (ArgMax)

Heuristic Method (Nearest K weighted SUM)

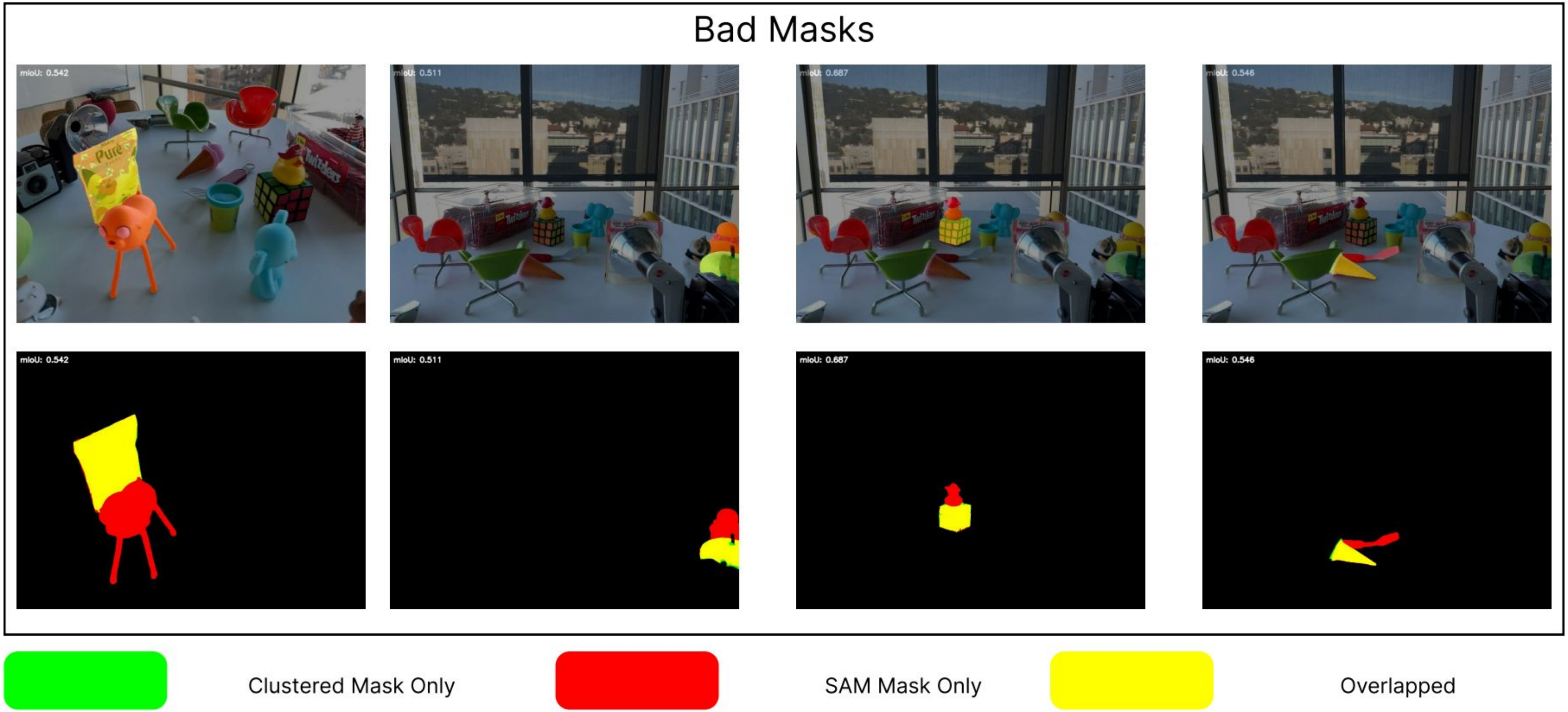
Feature Lifting is an ill-posedness problem



Non-Trivial Noise Exist

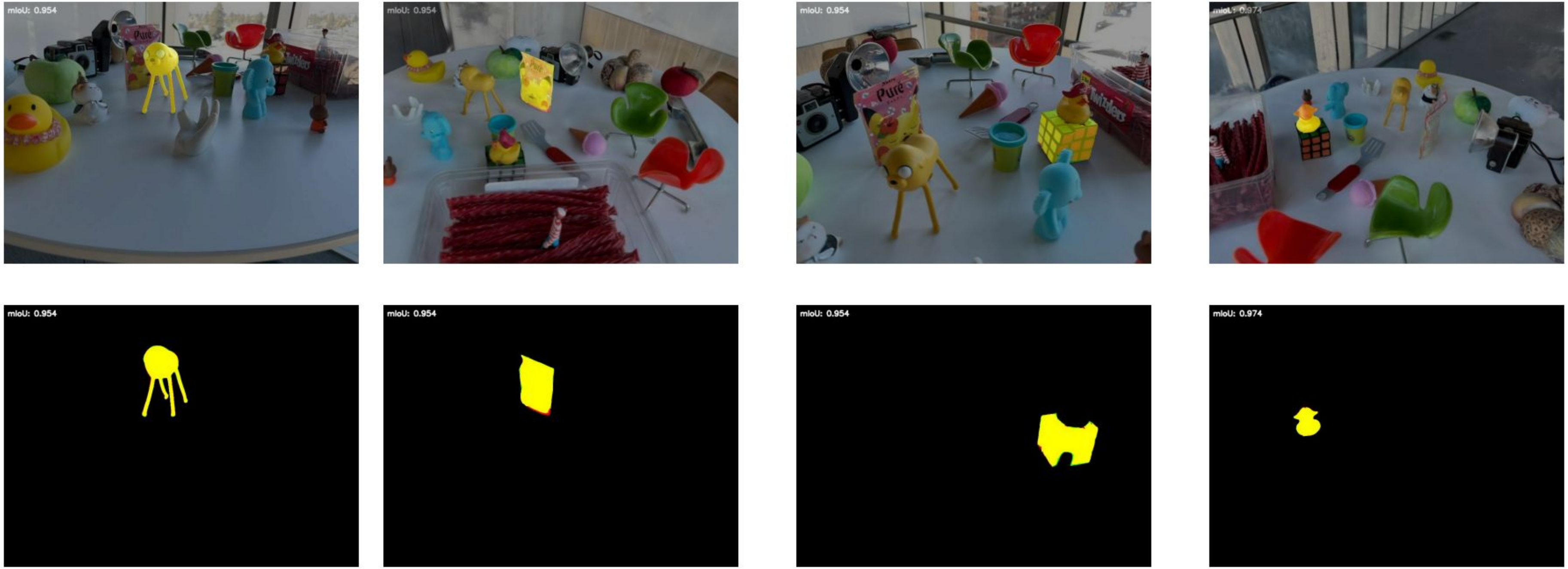


Noise Example



Expected Mask

Good Masks



Clustered Mask Only

SAM Mask Only

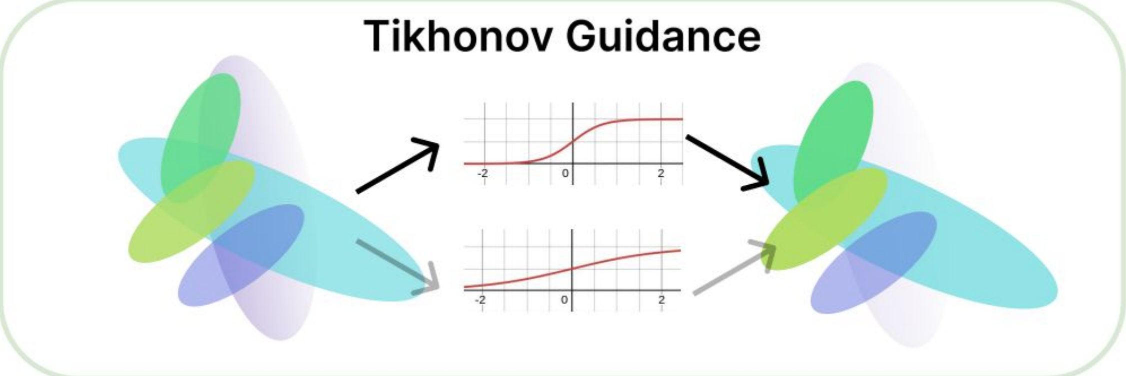
Overlapped

Stabilizing (Tikhonov Regularization)

$$\min\left(||A\tilde{x}-b||^2+||\lambda I||^2\right)$$

$$\tilde{A_{ij}}=\tilde{\omega_{ij}},\quad \tilde{\alpha_p}=\frac{1}{1+e^{-\lambda\theta}}$$

$$x=D^{-1}(\tilde{A^T}\tilde{A})e\times D(\tilde{A^T}\tilde{A})B\quad x_j=\frac{\sum_i \tilde{A}_{ij}^2 B_i}{\sum_i \tilde{A}_{ij}^2}$$



Denoising (Post Lifting Aggregation)

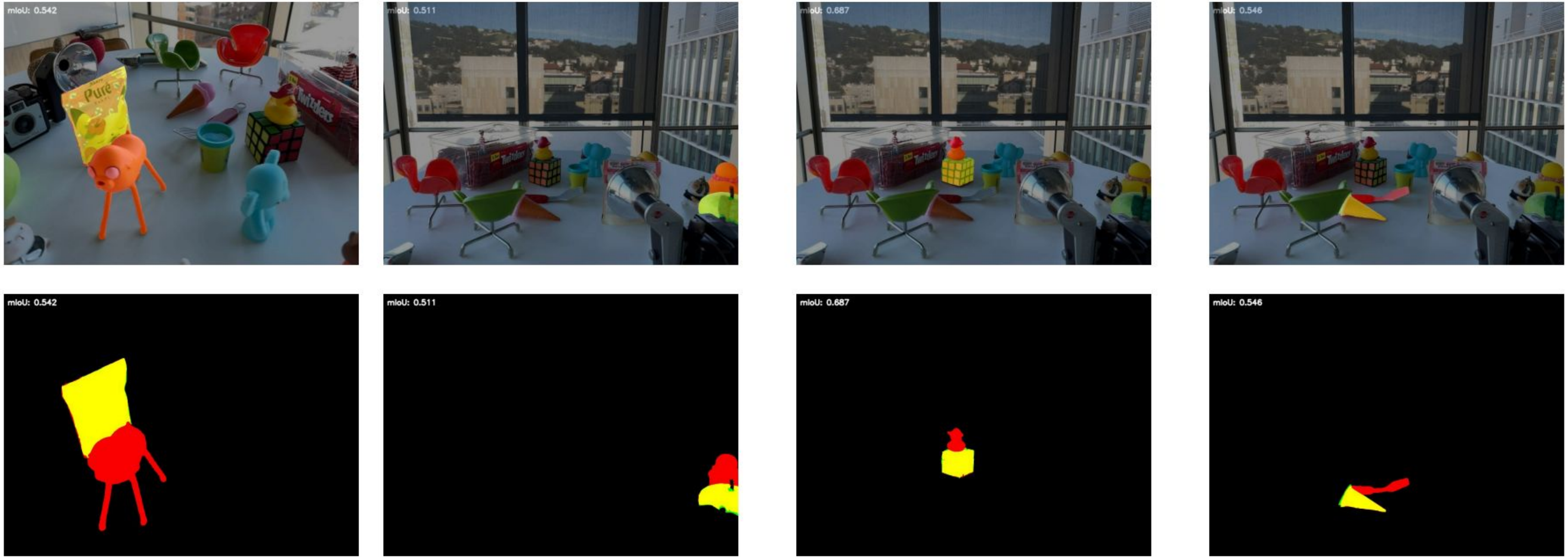
$$\gamma = \mathrm{Agg}(x) \ \in \{-1, \ldots, K\}^P,$$
$$\Gamma = \mathrm{onehot}(\gamma) \ \in \{0, 1\}^{P \times (K+1)},$$
$$\kappa = \arg\max(A\,\Gamma) - 1 \ \in \{-1, \ldots, K\}^R,$$

$$M(j, B) = m_j \in \mathbb{R}^{H \times W}, M(j, \kappa) = m'_j \in \mathbb{R}^{H \times W} \quad (21)$$

$$B'_j = \begin{cases} B_j, & \textbf{IoU}\,(m_j, m'_j) \geq \tau, \\ \emptyset, & \text{otherwise.} \end{cases} \quad (22)$$

Denoising (Post Lifting Aggregation)

Bad Masks



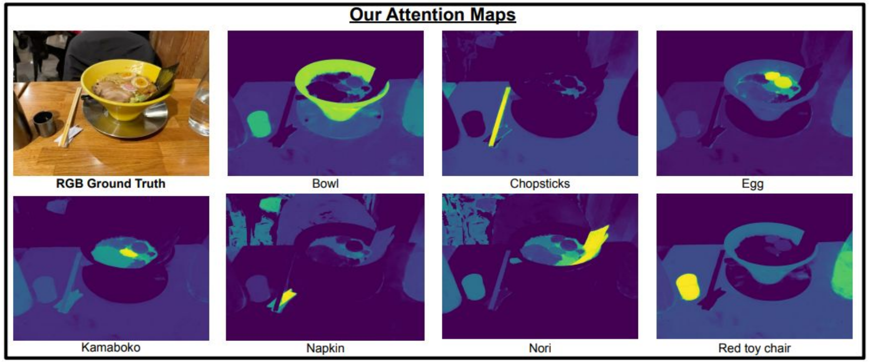
Clustered Mask Only

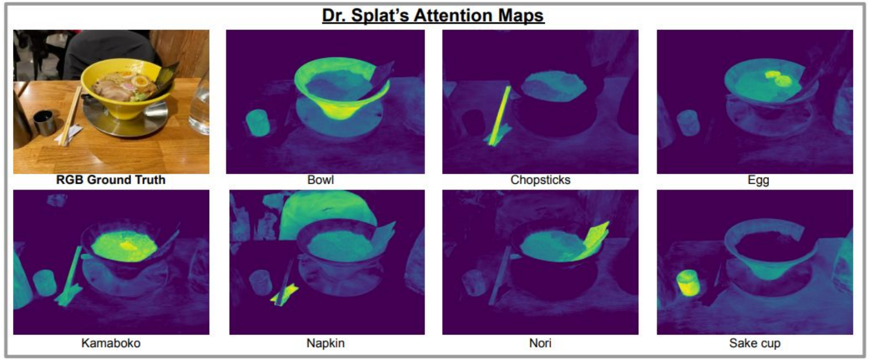
SAM Mask Only

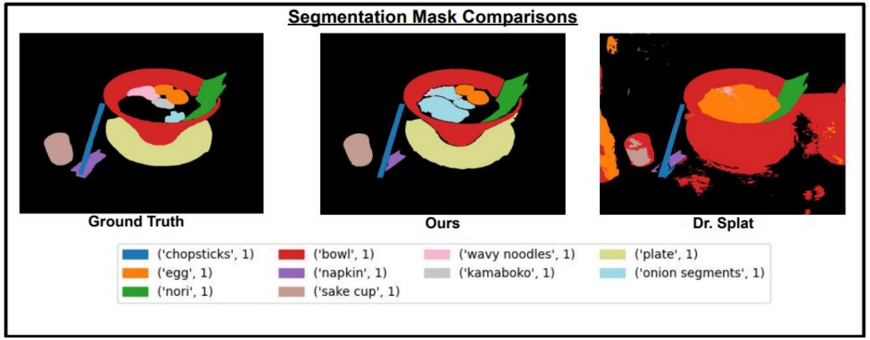
Overlapped

Comparison

Method/Scene	F.	T.	R.	W.	Mean
LSeg† (Li et al. 2022)	7.6	21.7	7.0	29.9	16.6
LeRF† (Kerr et al. 2023)	38.6	45.0	28.2	37.9	37.4
N2F2† (Bhalgat et al. 2024)	47.0	69.2	56.6	47.9	54.4
LangSplat† (Qin et al. 2024)	25.9	35.6	29.3	33.5	31.1
LeGaussian† (Shi et al. 2024)	31.2	34.5	17.6	17.3	25.2
SuperGseg† (Liang et al. 2024)	43.7	55.3	18.1	26.7	35.9
VLGS† (Peng et al. 2024a)	58.1	73.5	61.4	54.8	62.0
SAGA† (Cen et al. 2023)	36.2	19.3	53.1	14.4	30.7
OpenGaussian† (Wu et al. 2025)	61.1	59.1	29.2	31.9	45.3
GS Grouping† (Ye et al. 2023)	60.9	40.0	45.5	38.7	46.3
LAGA‡ (Cen et al. 2025)	56.1	68.9	57.4	64.6	61.7
LAGA† (Cen et al. 2025)	64.1	70.9	55.6	65.6	64.0
DrSplat‡ (Jun-Seong et al. 2025)	47.5	66.2	36.7	47.5	49.5
DrSplat† (Jun-Seong et al. 2025)	53.4	57.2	24.7	39.1	43.6
OccamLGS‡ (Cheng et al. 2024)	60.1	68.3	55.3	47.7	57.8
OccamLGS† (Cheng et al. 2024)	58.6	70.2	51.0	65.3	61.3
Ours	67.6	68.5	62.3	62.1	65.1







Performance and Ablation Study

[Splat Feature Solver]

Features/scene	F.	T	R	W.	Mean
Ours w/o (T_iP)	00:03:29	00:02:06	00:01:28	00:02:13	00:02:12
Ours w/o (P)	00:03:14	00:02:05	00:01:12	00:01:47	00:02:05
Ours full	00:05:22	00:02:37	00:01:59	00:03:02	00:03:15
DrSplat	00:02:55	00:01:19	00:01:03	00:01:33	00:01:43
LAGA	01:43:33	01:23:22	01:20:48	01:30:16	01:29:30

Features/scene	F.	T	R	W.	Mean
Ours w/o (T_iPA)	60.1	68.3	55.3	47.7	57.8
Ours w/o (PA)	61.7	67.8	53.8	49.6	58.2
Ours w/o (T_iA)	65.5	72.0	58.6	50.4	61.6
Ours w/o (A)	64.8	71.6	61.7	54.7	63.2
Ours	67.6	62.3	68.5	62.1	65.1

Features/scene	F.	T	R	W.	Mean
SAMOpenCLIP (Kirillov et al. 2023)	89.3	90.7	90.6	91.0	90.4
MaskCLIP (Dong et al. 2023)	92.4	94.0	94.5	94.7	93.9
CLIP (Radford et al. 2021)	91.9	93.7	94.6	94.9	93.8
DINO (Caron et al. 2021)	78.7	80.2	83.0	82.0	81.0
DINOv2 (Oquab et al. 2023)	83.5	85.6	89.6	89.2	87.0
ViT (Dosovitskiy et al. 2020)	84.6	83.7	86.4	88.0	85.7
ResNet (He et al. 2016)	95.8	94.8	96.2	97.0	96.0

Method	F.	T	R	W.	Mean
DBS (Liu et al. 2025)	49.5	50.8	51.2	**61.3**	53.2
3DGS* (Tancik et al. 2023)	55.3	63.5	47.8	49.8	54.1
3DGS# (Kerbl et al. 2023)	60.1	**68.3**	55.3	47.7	57.8
2DGS* (Huang et al. 2024)	**62.0**	66.3	**56.0**	51.1	**58.9**

Full Ablation

Method	Figurines	Ramen	Teatime	Waldo Kitchen	Means
DrSplat (3D Query)	47.48	36.66	66.16	47.48	49.45
Gsplat (3DGS) W/o P W/o T (Threshold = 0.65)	55.30	47.81	63.48	49.78	54.09
Gsplat 2DGS Backbone W/o P W/o (Threshold = 0.65)	62.00	56.04	66.34	51.07	58.86
Gsplat (3DGS) W/P w/o T (Threshold = 0.65)	56.55	62.24	68.04	53.01	59.96
Gsplat (2DGS) W/P W/o T (Threshold = 0.65)	67.83	60.20	66.96	47.44	60.62
Inria Trained Result W/o P, Naive (W/o Tikhonov) (Threshold)	60.06	55.30	68.33	47.73	57.85
Inria Trained Result W/o P, (Threshold=0.65)	61.30	54.20	67.60	45.01	57.03
Inria Trained W/p W/ Tikhonov (1.2 / 1.2)	61.70	53.75	67.80	49.60	58.21
Inria Trained Result W/p w/ (Tikhonov square, no sigmoid)	71.96	58.16	69.29	52.73	63.03
Inria Trained Result W/p (Threshold 0.65) + W/T	64.76	61.70	71.63	54.71	63.20
Inria Trained W/p W/o Tikhonov Square, w/o sigmoid	65.45	58.55	72.04	50.38	61.61
Inria Trained W/p W/o (Tikhonov λ w/o squeer) W/ hist selection	72.19	64.28	66.08	56.69	64.81
Inria Trained W/o T W/p W/ Dynamic Threshold	69.42	63.89	69.53	61.55	66.10
Inria Trained W/ T W/ P W/ Dynamic Threshold	67.64	62.34	68.48	62.11	65.14

Compression Strategy: Feature Map Compression

Precision	Format	Time	Relative Err	Ratio
INT 16	PNG	0.50 Sec/Frame	2.90*e-5	2.4
INT 8	PNG	0.25 Sec/frame	7.46*e-3	9.7
INT 8	JPG	0.2 Sec/Frame	2.31*e-2	47.0

Future Work

- Sparse Feature
- Would it be possible to solve further attributes? Say (Scale, Quaternions …) by fixing means
- Would it be possible to denoising during run time?
- Compression of the result
- Scene Decomposition
- Visual Localization

Thank You

Meida Chen

Andrew Feng

Rong Liu
Research Engineer
USC Institute for Creative Technologies

Rong Liu