3D-SpLineNet: 3D Traffic Line Detection using Parametric Spline Representations

Maximilian Pittner[1,2], Alexandru Condurache[1,2], Joel Janai[1]

[1]Bosch Mobility Solutions, Robert Bosch GmbH

[2]Institute for Signal Processing, University of Lübeck

{Maximilian.Pittner, AlexandruPaul.Condurache, Joel.Janai}@de.bosch.com

BOSCH Invented for life

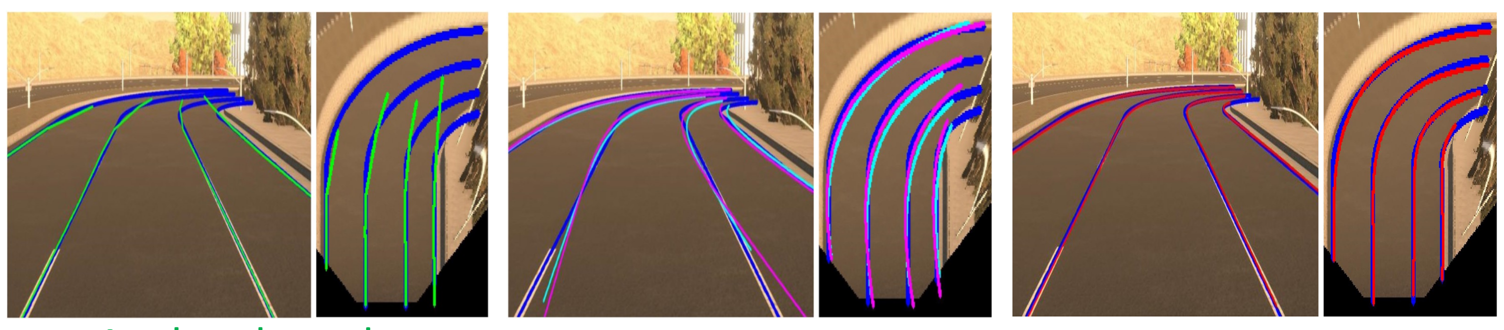
UNIVERSITÄT ZU LÜBECK

WACV WAIKOLOA HAWAII JAN 3-7 • 2023

I. Motivation and Contributions

Abstract. *We present a method to detect 3D traffic lines from single RGB images. To overcome limitations of discrete anchor- or grid-based approaches we propose a parametric lane representation based on sophisticated B-Splines.*

Lane detections using different representations for typical road shape

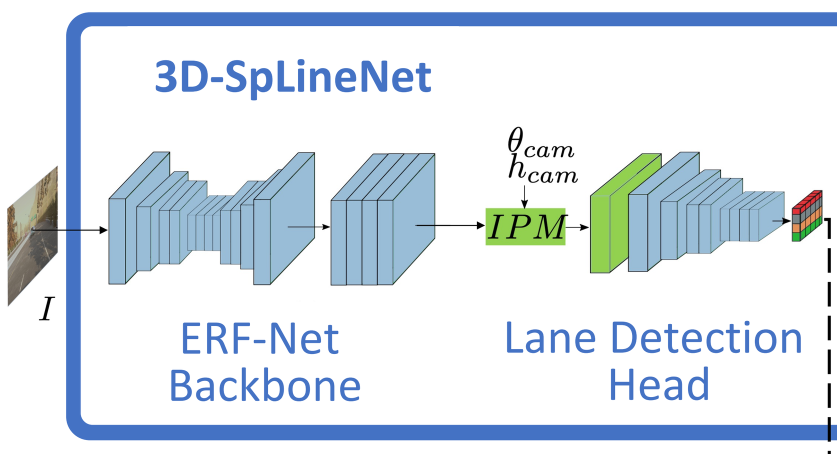


Anchor-based (Gen-LaneNet [2])	Polynomials of 3rd and 5th degree	B-Splines of 3rd degree
• Discrete / not smooth • Fail for strong deviations from anchor	• Requires high degrees for simple road shapes • Not suited for lanes	• Benefits of parametric representations • Suited for typical roads

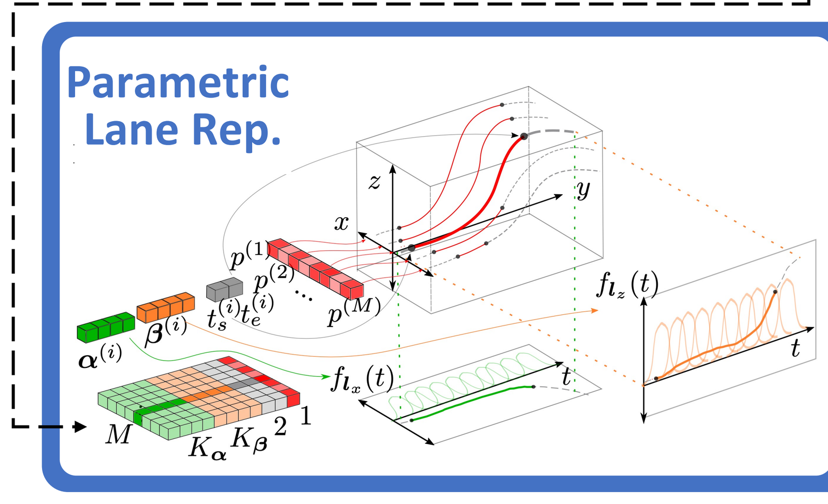
Contributions

- A **new representation** to model 3D lanes with parametric B-Splines
- An end-to-end-trainable detection network called **3D-SpLineNet**
- **SOTA performance** on Apollo 3D Lanes Synthetic benchmark

II. Overview of Method



- Modified ERF-Net Network outputs multi-channel feature maps instead of seg. map
- IPM Layer transforming feature maps to top-view
- Fully-Conv. Detection Head
- Trained end-to-end



- Detection Head provides parameters for M line proposals
- Each proposal i consisting of:
  - Control points $\boldsymbol{\alpha}^{(i)}, \boldsymbol{\beta}^{(i)}$ for x- and z-component
  - Start and end point $t_s^{(i)}, t_e^{(i)}$
  - Existence probability $p^{(i)}$

III. Parametric Representation

3D Lane Representation

$$\boldsymbol{l}(t) = \begin{pmatrix} x(t) \\ y(t) \\ z(t) \end{pmatrix} = \boldsymbol{\eta} \odot \begin{pmatrix} f_{l_x}(t) \\ t \\ f_{l_z}(t) \end{pmatrix}$$

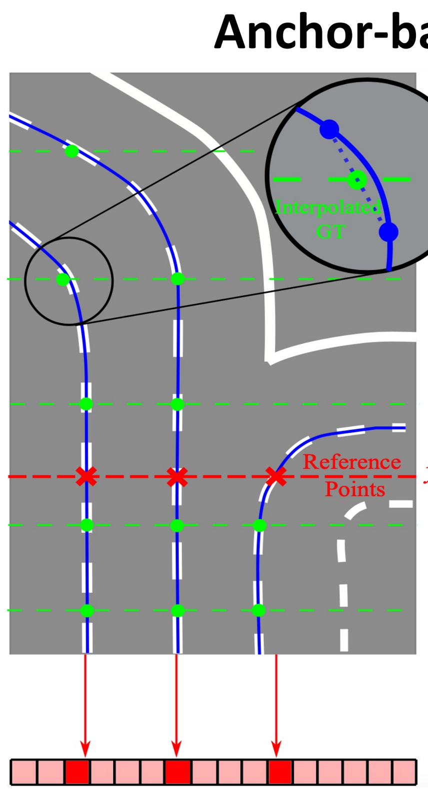
x-comp: $$f_{l_x}(t) = \sum_{k=1}^{K_B} \alpha_k \cdot B_{k,d}(t)$$

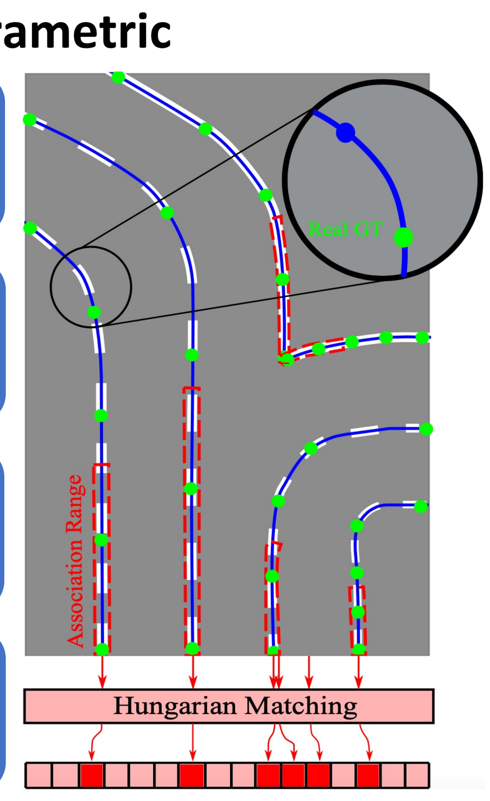
z-comp: $$f_{l_z}(t) = \sum_{k=1}^{K_B} \beta_k \cdot B_{k,d}(t)$$

Normalization vector: $$\boldsymbol{\eta} = \begin{pmatrix} x_{max} \\ y_{max} \\ z_{max} \end{pmatrix}$$

Start- / end-point: $t \in [t_s, t_e]$

Anchor-based	Parametric
⊖ Requires GT interpolation	⊕ Uses real GT
⊖ Regression at fix positions	⊕ Reg. over entire line
⊖ Neglects lines missing ref.	⊕ Considers all lines
⊖ Static matching	⊕ Optimized matching





IV. Training Objective

Classification loss: $$\mathcal{L}_c = -\sum_{i=i}^{M} \hat{p}^{(i)} \log p^{(i)} + (1 - \hat{p}^{(i)}) \log(1 - p^{(i)})$$

Shape matching loss: $$\mathcal{L}_s = \int_{\hat{t}_s}^{\hat{t}_e} \left\| \boldsymbol{w}(t) \odot \left(\boldsymbol{f}_l(t) - \boldsymbol{\eta}^{-1} \odot \hat{\boldsymbol{l}}(t) \right) \right\|_1 dt$$

Range loss: $$\mathcal{L}_r = |t_s - \hat{t}_s| + |t_e - \hat{t}_e|, \quad \text{with } t_s < t_e$$

Final loss function: $$\mathcal{L} = \lambda_c \cdot \mathcal{L}_c + \sum_{i=1}^{M} \hat{p}^{(i)} \cdot \left(\lambda_s \cdot \mathcal{L}_s^{(i)} + \lambda_r \cdot \mathcal{L}_r^{(i)} \right)$$

V. Comparison to State-of-the-art

Quantitative Evaluation on Apollo 3D Lanes Synthetic dataset [2]

Detection Perf. / Geometric Acc.

- Outperforming anchor-based on all test sets
- Strong gap in detection rate even on complex Rare Scenes set
- Lower geometric errors wrt. almost all criteria
- Room for improvement for far-range z-error → Architecture / IPM

Method	F	AP	x-error near	x-error far	z-error near	z-error far
3D-L.[1]	86.4	89.3	6.8	47.7	1.5	**20.2**
Gen-L.[2]	88.1	90.1	6.1	49.6	1.2	21.4
Ours	**96.3**	**98.1**	**3.7**	**32.4**	**0.9**	21.3

Standard test set

Method	F	AP	x-error near	x-error far	z-error near	z-error far
3D-L.	72.0	74.6	16.6	85.5	3.9	**52.1**
Gen-L.	78.0	79.0	13.9	90.3	3.0	53.9
Ours	**92.9**	**94.8**	**7.7**	**69.9**	**2.1**	56.2

Rare Scenes test set

Method	F	AP	x-error near	x-error far	z-error near	z-error far
3D-L.	72.5	74.9	11.5	60.1	3.2	**23.0**
Gen-L.	85.3	87.2	7.4	53.8	1.5	23.2
Ours f.b.	**91.3**	**93.1**	**6.9**	**46.8**	**1.3**	24.8

Visual Variations test set

Runtime

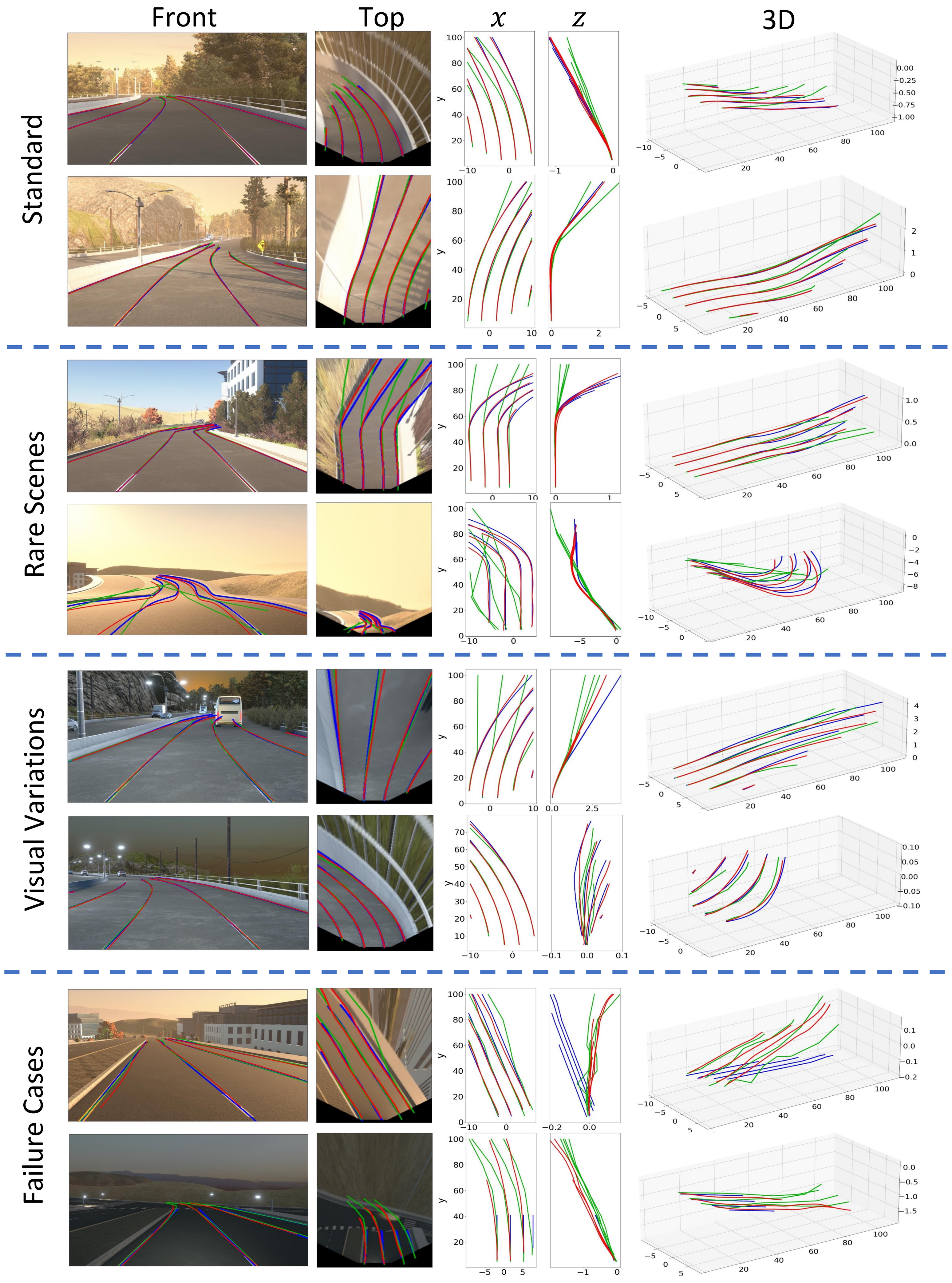
- No time consuming post-processing required
- Faster inference than discrete representations

Method	3D-L.	Gen-L.	Ours
Runtime	41.9 fps	36.3 fps	**74.3 fps**

All methods evaluated on NVIDIA GeForce Titan X

Qualitative Evaluation

- Ground truth - Gen-LaneNet - 3D-SpLineNet



VI. Ablation Studies

Association Criterion

- Association range matching surpassing fixed ref. point
- Best results for 40 %

Ref.	20 m	First 20 %	**First 40 %**	100 %
F	90.7 %	92.1 %	**92.9 %**	91.2 %
AP	92.5 %	94.1 %	**94.8 %**	93.4 %

Analysis of Representations

- Polynomials of degree 2-3 not sufficient
- 3rd degree splines sufficient even for few knots
- Impact of knot number decreases for higher values → 5 knots sufficient

Rep.	d	N	F	x-error near	x-error far	z-error near	z-error far
Poly.	2	–	88.0	14.0	83.1	2.4	58.1
	3	–	90.4	13.6	75.4	2.6	57.3
	5	–	91.6	9.6	74.8	2.4	58.2
B-Sp.	1	3	81.3	26.3	102.3	3.0	58.1
	3	3	90.8	9.9	69.7	2.4	56.4
	3	5	92.5	8.2	**68.3**	2.2	56.
	1	10	91.5	9.1	71.3	2.2	56.3
	3	10	92.1	8.3	70.8	**1.8**	56.3
	1	15	91.6	9.1	69.1	1.9	**54.9**
	3	**15**	**92.9**	**7.7**	69.9	2.1	56.2

VII. Conclusion and Future Work

In summary:

- Parametric representations show benefits over discrete anchor-based
- B-Splines of 3rd degree sufficient to model complex road shapes
- SOTA performance of 3D-SpLineNet

In future we investigate:

- Ways to improve height estimation
- Better initialization strategies with 3-component representation
- Weakly supervised approaches using 2D ground truth

[1] Noa Garnett, Rafi Cohen, Tomer Pe'er, Roee Lahav, and Dan Levi. 3d-lanenet: End-to-end 3d multiple lane detection. In *Proc. of the IEEE International Conf. on Computer Vision (ICCV)*, 2019.

[2] Yuliang Guo, Guang Chen, Peitao Zhao, Weide Zhang, Jing- hao Miao, Jingao Wang, and Tae Eun Choe. Gen-lanenet: A generalized and scalable approach for 3d lane detection. In *Proc. of the European Conf. on Computer Vision (ECCV)*, 2020.