Supplementary Material: Learning to Localize through Compressed Binary Maps

Xinkai Wei[1,2,*] Ioan Andrei Bârsan[1,3,*] Shenlong Wang[1,3,*]
Julieta Martinez[1] Raquel Urtasun[1,3]
[1]Uber Advanced Technologies Group [2]University of Waterloo [3]University of Toronto
{xinkai.wei,andreib,slwang,julieta,urtasun}@uber.com

1. Additional Experimental Results

1.1. Quantitative Results

Comparison to smaller map resolutions: We also performed experiments with reduced map resolutions on our urban dataset to investigate the impact on storage requirements and localization accuracy. As shown in Table 1, We note that unlike the tables in the paper, here we measure the storage requirements in bits / m^2, in order to account for the different map resolutions. magnitude of the storage required by our approach. However, the localization performance is substantially reduced (16.28% failure rate, as opposed to 2.56% for our binary coding).

Ablation Study on Highway Dataset

1.2. Qualitative Results

Figures 1–7 contain samples from our localization application. Note that the compression happens independently of the online observations. Here, we show the online observations and their embedding for reference, and to highlight that our matching is robust to any traffic conditions.

We note that in all our visualizations the original map is shown for illustration purposes only. In practice, the original map does not need to be saved onboard, as the compressed embedding is enough for performing online localization.

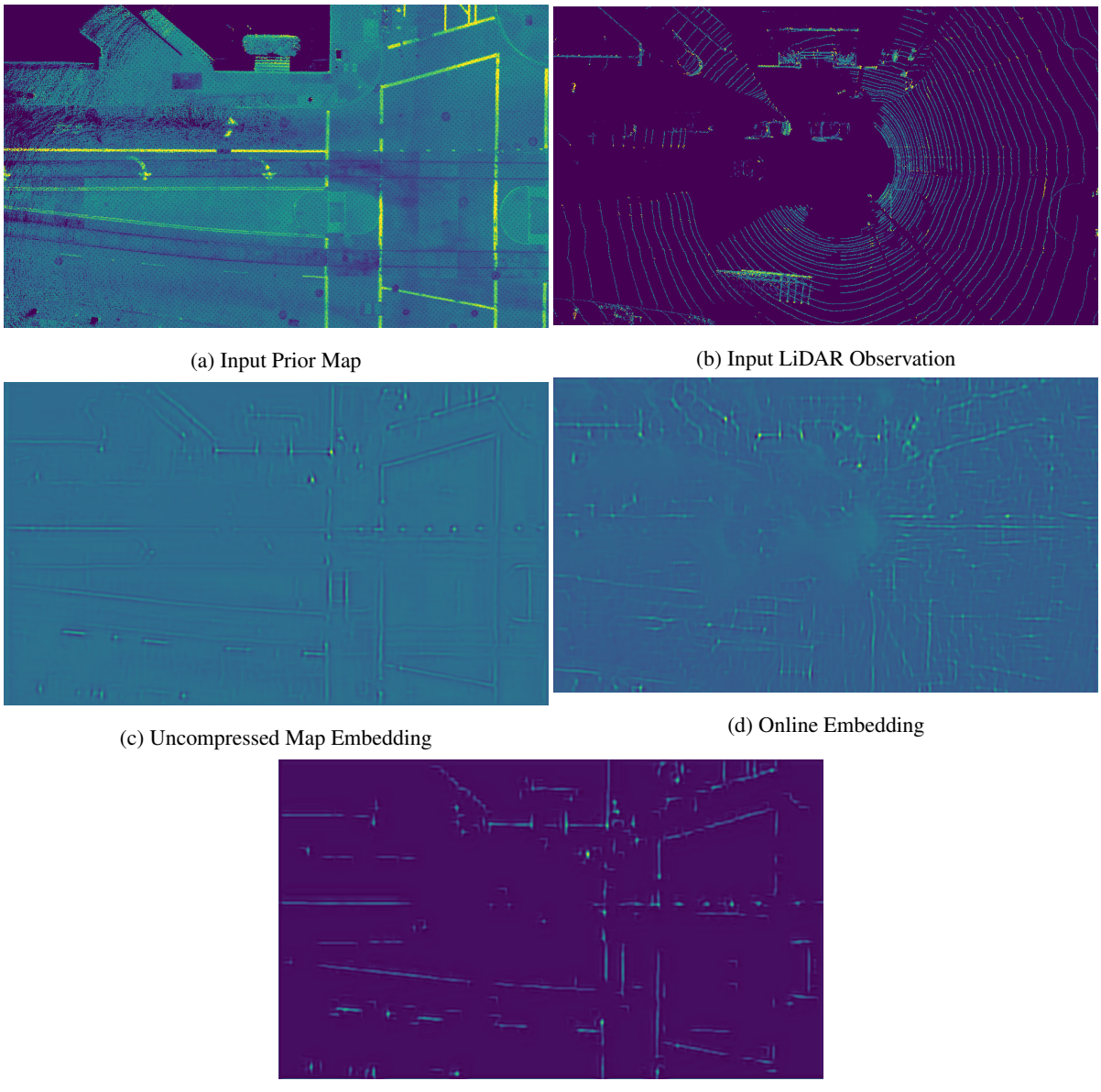
Figure 8 shows an example of the binary codes generated by our compression module. This example shows the 64-way binarized embedding that is the output of our grouped softmax compression module. We note that only some of the channels are activated, and thus most of the binary bits correspond to a small subset of the binary embedding channels. This is a result of having an optimization objective that consists of entropy-based losses, and these sparse results show that we successfully learned to discard unused channels, keeping channels only if they capture information important for our localization task. We observe that the compression scheme learns to dedicate channels to represent important geometric features such as road boundaries, and lane markings.

Method	**Median Err (cm)**			**Failure Rate (%)**			**b/m^2**
	Lat	Lon	Total	$\leq$ 100m	$\leq$500m	End	
PNG, 5cm/px	**1.55**	**2.05**	**3.09**	**0.00**	<u>1.09</u>	**2.44**	1948.55
PNG, 10cm/px	4.37	6.68	9.50	3.19	3.26	4.00	402.84
JPG@50, 10cm/px	4.51	5.78	8.95	0.00	<u>1.09</u>	10.64	63.42
PNG, 15cm/px	15.73	23.66	31.73	10.31	20.65	22.03	173.97
JPG@50, 15cm/px	11.67	18.20	25.14	9.28	13.04	16.28	<u>29.00</u>
Ours (16$\times$)	<u>1.76</u>	<u>2.48</u>	<u>3.62</u>	**0.00**	**0.00**	<u>2.56</u>	**2.87**

Table 1: Localization performance on our urban dataset using reduced resolution maps. We used 5cm/px in the submission. Map storage is measured in bits/m^2 in order to account for different resolutions (bits-per-pixel (bpp) are no longer meaningful if the area of a pixel can change). *Ours* refers to our 16$\times$ downsampling method. JPG quality is 50.

Method	Median error (cm)			Failure rate (%)			Bit per pixel
	Lat	Lon	Total	$\leq$ 100m	$\leq$ 500m	End	
Lossless (PNG)	**3.62**	**4.53**	**7.06**	**0.00**	0.35	0.72	4.97
Ours (recon, 8x)	3.77	4.72	7.32	**0.00**	**0.00**	0.72	0.021
Ours (recon, 16x)	3.84	4.61	7.25	0.35	1.06	1.43	0.016
Ours (recon + match, 8x)	3.33	4.73	6.99	**0.00**	**0.00**	**0.36**	0.0074
Ours (recon + match, 16x)	**3.62**	4.77	7.19	0.35	0.35	0.72	**0.0072**
	0.35	0.71	**0.0072**				

Table 2: **Ablation studies on the highway dataset.**



(a) Input Prior Map

(b) Input LiDAR Observation

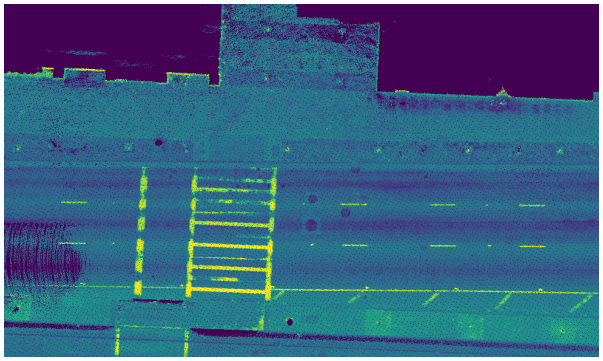
(c) Uncompressed Map Embedding

(d) Online Embedding

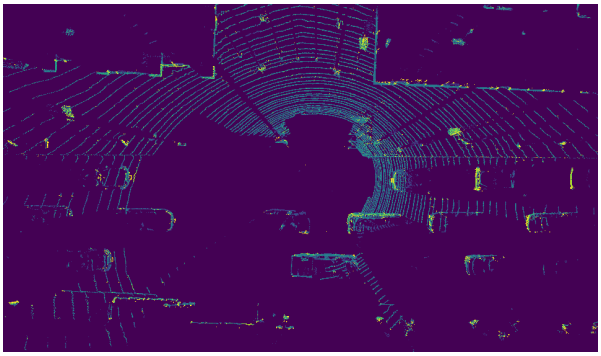
(e) Compressed Map Embedding

Figure 1: Preview of the localizer operating in a regular intersection.

For further results, we would like to refer the reader to the video associated with this submission, which shows our



(a) Input Prior Map



(b) Input LiDAR Observation

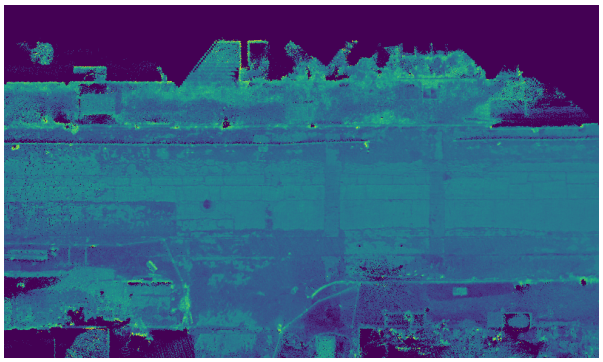
(c) Uncompressed Map Embedding

(d) Online Embedding

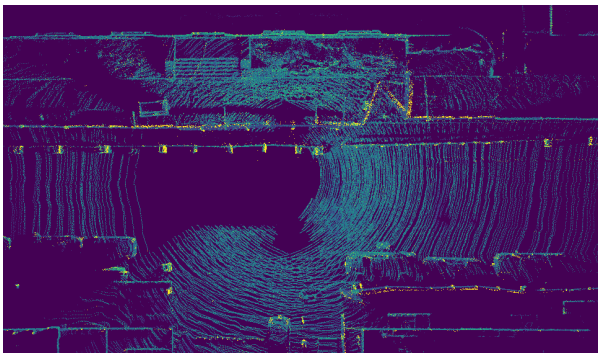
(e) Compressed Map Embedding

Figure 2: Example where the online localization successfully deals with heavy traffic.

probabilistic localizer running online using a compressed map embedding.

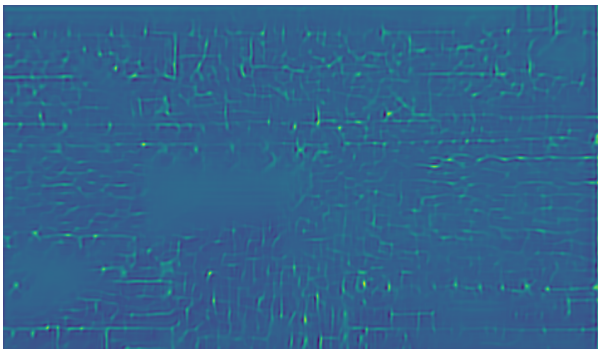


(a) Input Prior Map



(b) Input LiDAR Observation

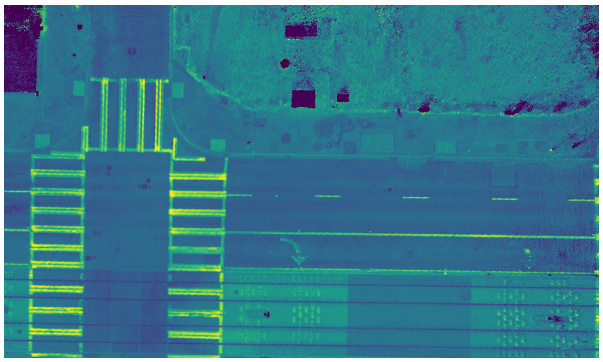
(c) Uncompressed Map Embedding



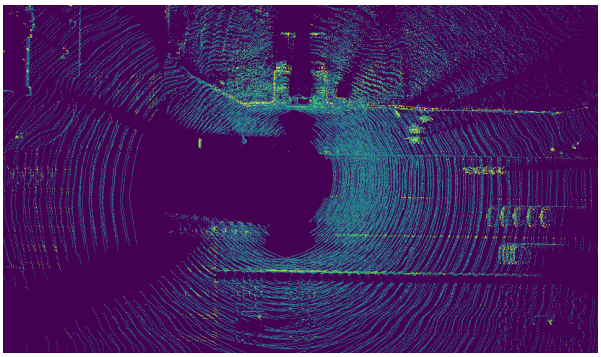
(d) Online Embedding

(e) Compressed Map Embedding

Figure 3: Example of a side road with no lane markings.

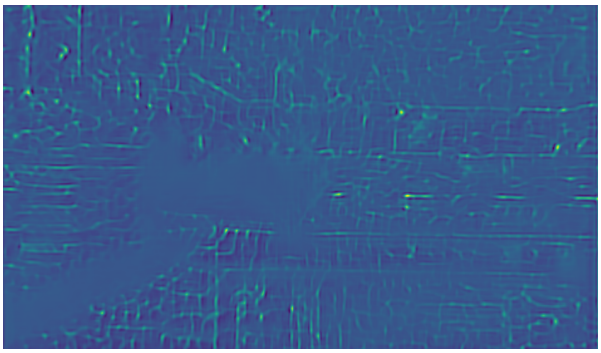


(a) Input Prior Map



(b) Input LiDAR Observation

(c) Uncompressed Map Embedding

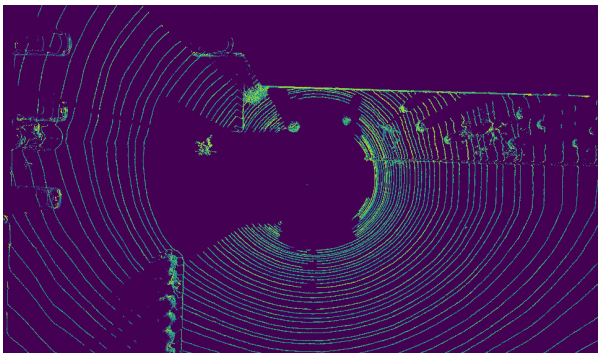


(d) Online Embedding

(e) Compressed Map Embedding

Figure 4: A section of the map with tram lines.

(a) Input Prior Map



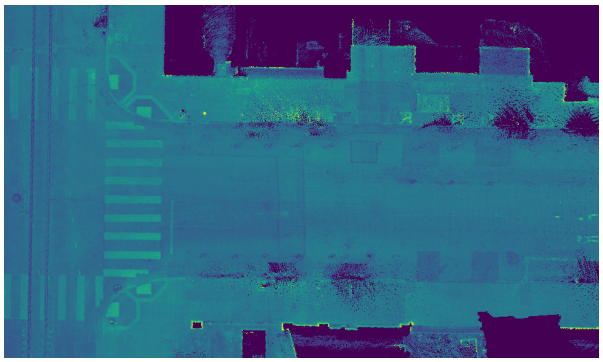
(b) Input LiDAR Observation

(c) Uncompressed Map Embedding

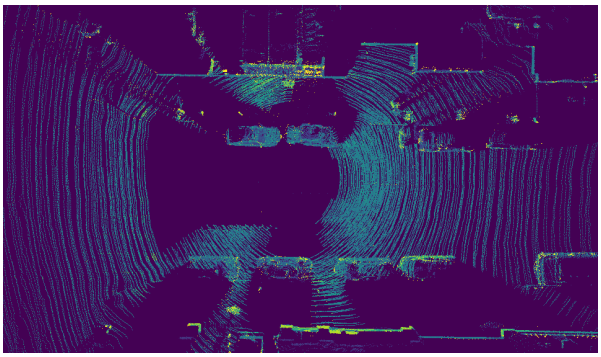
(d) Online Embedding

(e) Compressed Map Embedding

Figure 5: A parking lot.

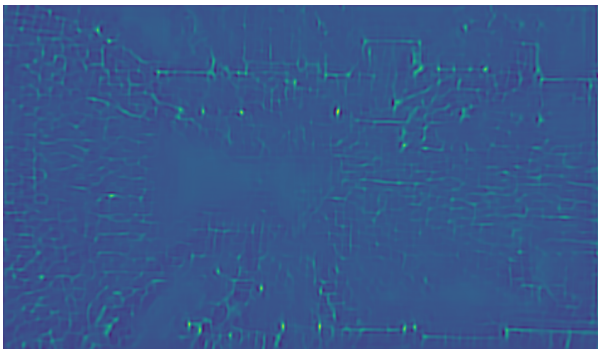


(a) Input Prior Map



(b) Input LiDAR Observation

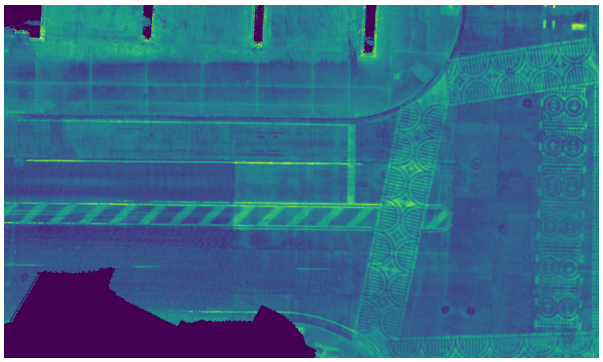
(c) Uncompressed Map Embedding



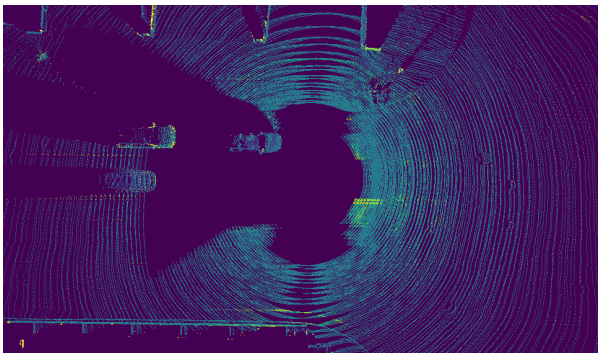
(d) Online Embedding

(e) Compressed Map Embedding

Figure 6: An intersection with fainter-than-usual crosswalks.

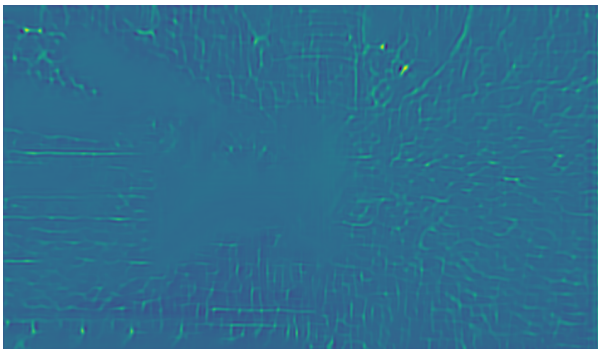


(a) Input Prior Map

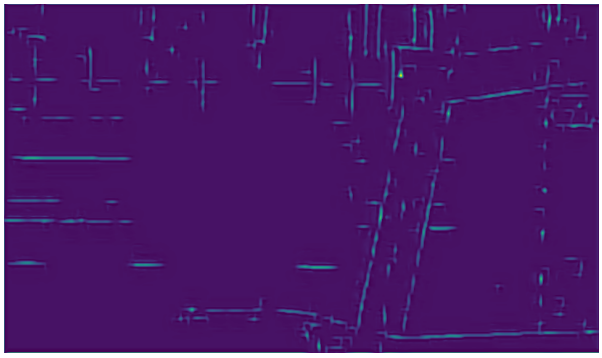


(b) Input LiDAR Observation

(c) Uncompressed Map Embedding

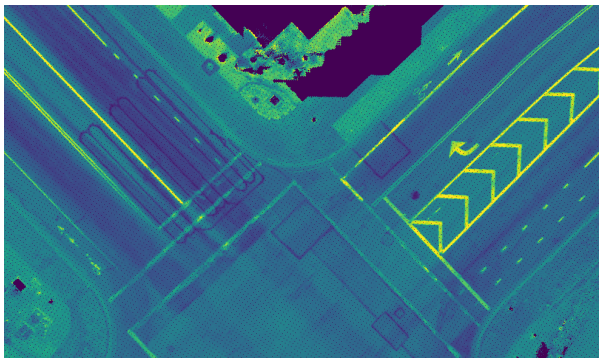


(d) Online Embedding



(e) Compressed Map Embedding

Figure 7: Unusual crosswalks.



(a) Input Prior Map

(b) Uncompressed Map Embedding

(c) Compressed Map Embedding (i.e., reconstructed from the binary codes)

(d) Binary codes learned by our system.

Figure 8: Examples of inputs to our system, together with the computed binary codes used to represent the learned map embedding in a compact way. Recall that the binary code maps are lower resolution than the inputs (this example uses the $8\times$ downsampling, so each code has $1/8$ the resolution of the input). Moreover, the neural network learns to only use a limited subset of the possible binary codes, leading to the reduced storage requirements described in the experimental section.