The Expectation Maximization Algorithm: Motivating Example I – WAMI to Roadmap Registration

Gaurav Sharma

University of Rochester

EM Application Example I

- Accurate Alignment of Wide Area Motion Imagery (WAMI) to Roadmaps Using Expectation Maximization
- Acknowledgment
  - Ahmed Elliethy, PhD Student University of Rochester

A. Elliethy and G. Sharma, “Automatic Registration of Wide Area Motion Imagery to Vector Road Maps by Exploiting Vehicle Detections,” IEEE Trans. Image Proc., accepted for publication August 2016.

Agenda

- WAMI Introduction
- Roadmap Alignment Exploiting Vehicle Detections
  - EM Formulation
- Results
- Benefits of EM Formulation

Wide Area Motion Imagery (WAMI)

- Aerial imaging platform
- Array of cameras - capture large areas = substantial parts of a city
- Sequence of image frames: high spatial & low (1–2 fps) temporal resolution

Images taken from

[1] E. Blasch, G. Seetharaman, S. Suddarth, K. Palaniappan, G. Chen, H. Ling, and A. Basharat, “Summary of methods in wide-area motion imagery (WAMI),” in Proc. SPIE, vol. 9089, 2014, pp. 90 890C–90 890C–10.
[2] http://www.globalsecurity.org/intell/systems/angel-fire.htm

WAMI Frame Detail

- High spatial resolution
- Can resolve
  - Building features
  - Individual vehicles
  - Roads
  - Trees
  - ...

WAMI Applications

- WAMI
  - Large scale
  - Can be near real-time, potentially interactive
- Enables new visual analytics applications:
  - Security and surveillance
    - Vehicle detection and tracking
    - Abnormal activity detection
  - Resource planning/allocation
    - City (3D) reconstruction
    - Traffic pattern analysis
  - …

WAMI Analytics Challenges

- Focus: Vehicle Detection and Tracking

- Challenges
  - Scale: several 100s of vehicles
  - Data rates ≈ 1.25 Gbps (in compressed format)
  - Low frame rate 1-2 fps
    - FMV is 30 fps
  - Vehicular spatial detail is limited
  - Multiple image modalities
    - Night-time IR

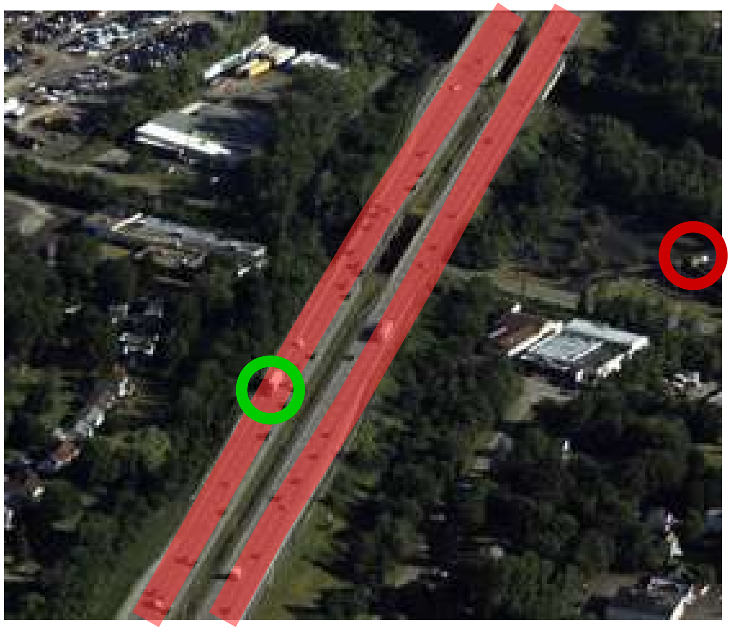
Agenda

- WAMI Introduction
- Roadmap Alignment Exploiting Vehicle Detections
  - Motivation, Challenges, Problem Statement
  - Prior Approaches and Challenges
  - Proposed Approach Exploiting Vehicle Detections
    - EM formulation
  - Results
- Benefits of EM Formulation

WAMI to Vector Road Map Alignment : Motivations

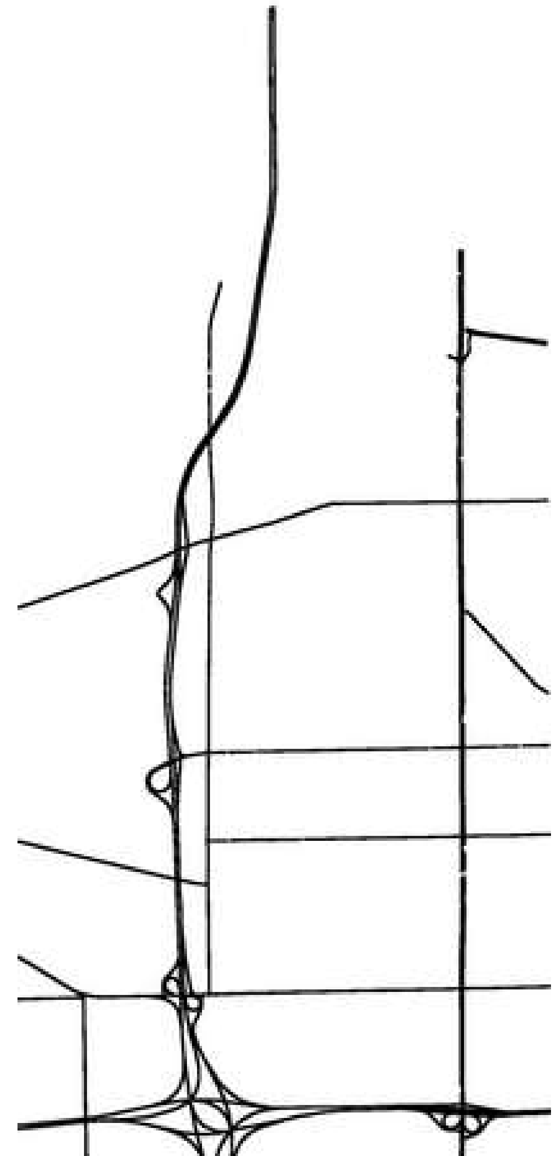
- Objective: exploit rich GIS information sources with WAMI for enhanced analysis
- Enabled by aligning WAMI frames to geo-referenced road maps

Vehicle detection



WAMI to Vector Road Map Alignment : Challenges

- Different data representations
  - WAMI frame as raster (pixels)
  - Vector representation of the road map.

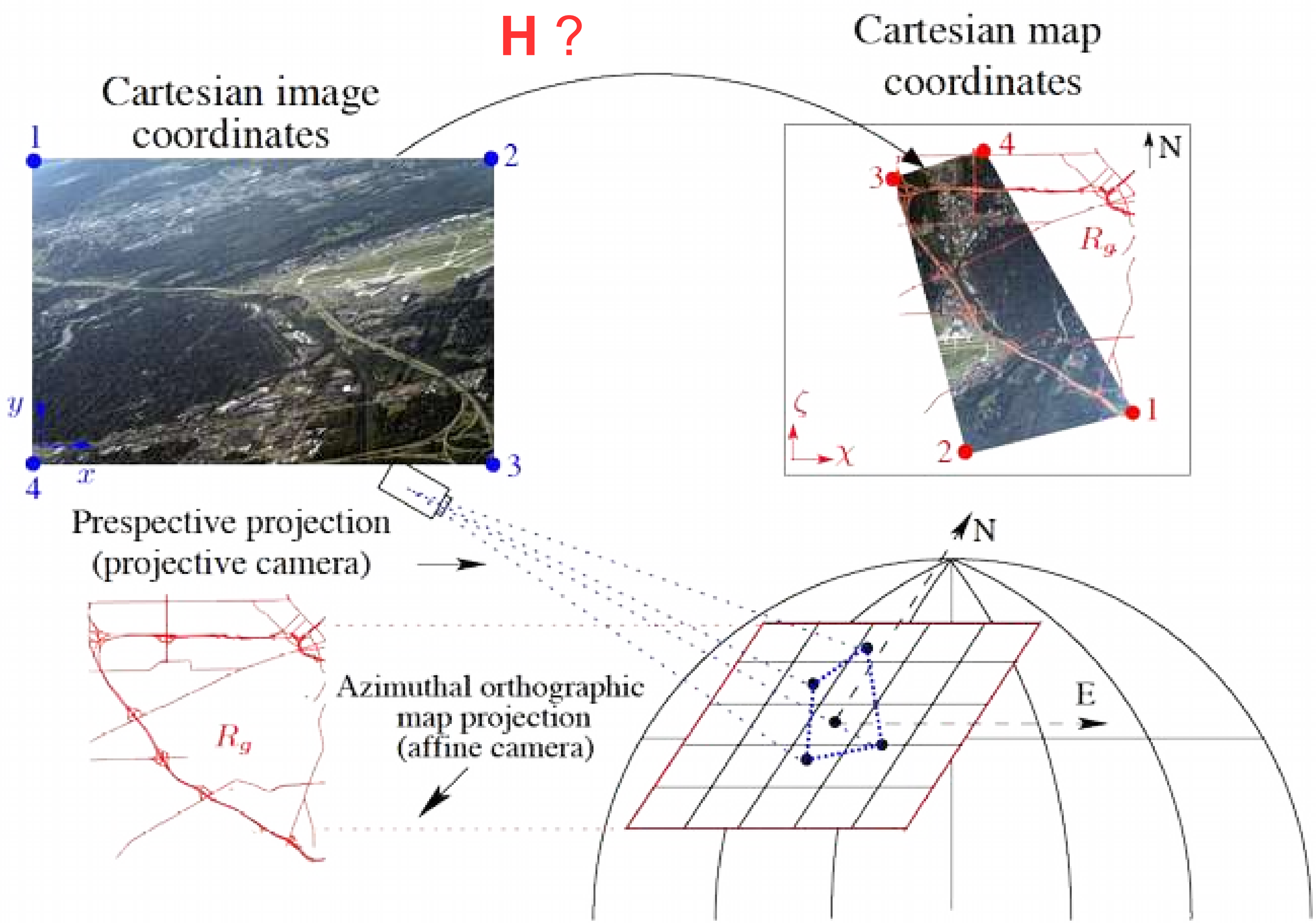


WAMI to Vector Road Map Alignment : Problem Statement

Given :

- Captured WAMI frame (raster)
- Vector road network (vector)

Goal: Estimate geometric transformation **H**



Agenda

- WAMI Introduction
- Roadmap Alignment Exploiting Vehicle Detections
  - Motivation, Challenges, Problem Statement
  - Prior Approaches and Challenges
  - Proposed Approach Exploiting Vehicle Detections
    - EM formulation
  - Results
- Benefits of EM Formulation

Agenda

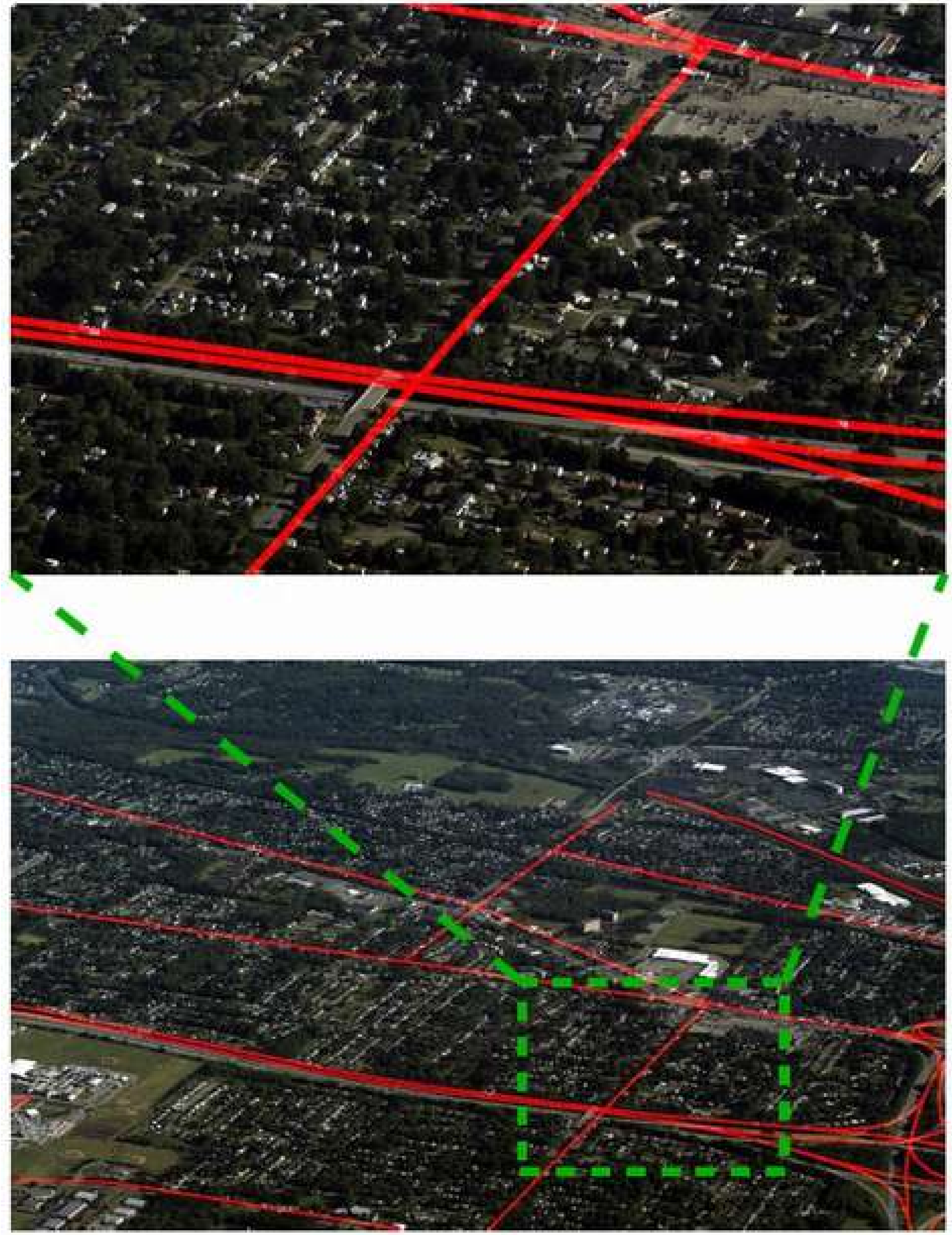
- WAMI Introduction
- Roadmap Alignment Exploiting Vehicle Detections
  - Motivation, Challenges, Problem Statement
  - Prior Approaches and Challenges
  - Proposed Approach Exploiting Vehicle Detections
    - EM formulation
  - Results
- Benefits of EM Formulation

Prior Approaches

- Use GPS/INS for geo-location

    + Often available as metadata

    - Limited accuracy (only approximate alignment)



Prior Approaches of Vector Road Map Registration

(2) Road segmentation + road intersections matching in both representations

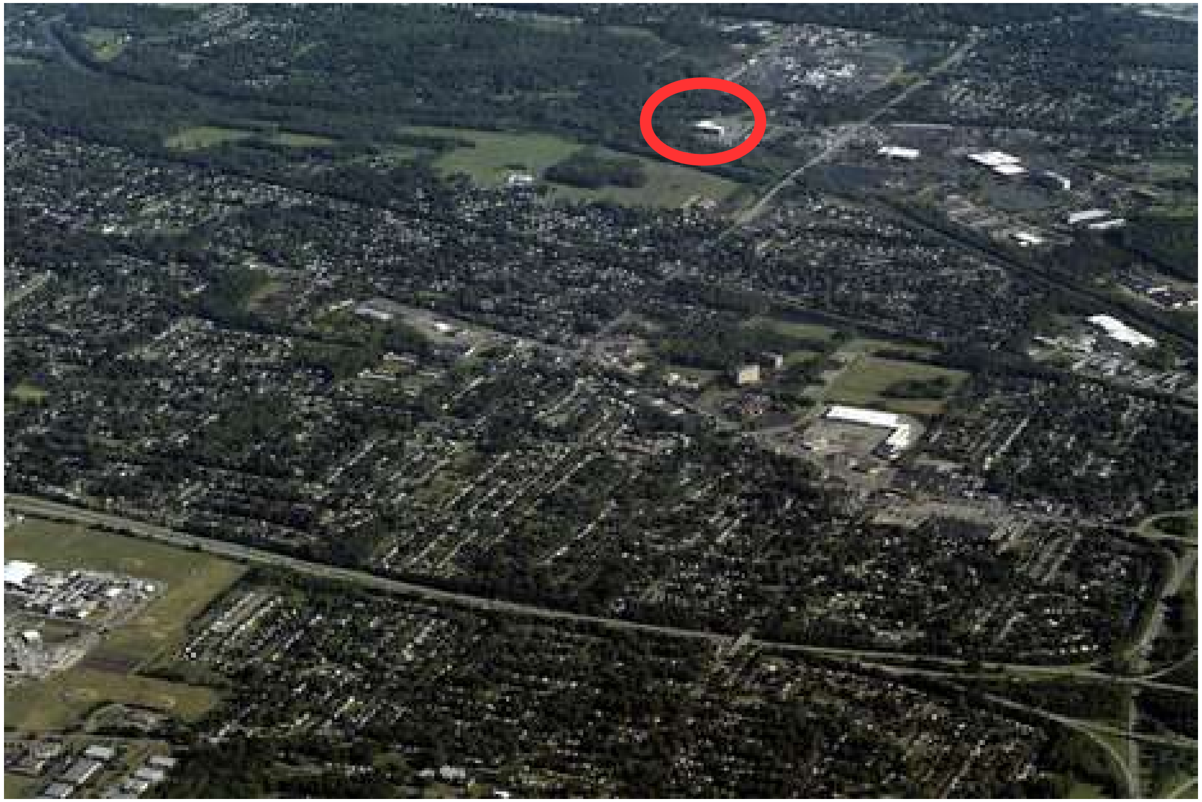
- Not suitable for our WAMI settings
  - Require manual labeling (not real time).
  - HS data is not available.

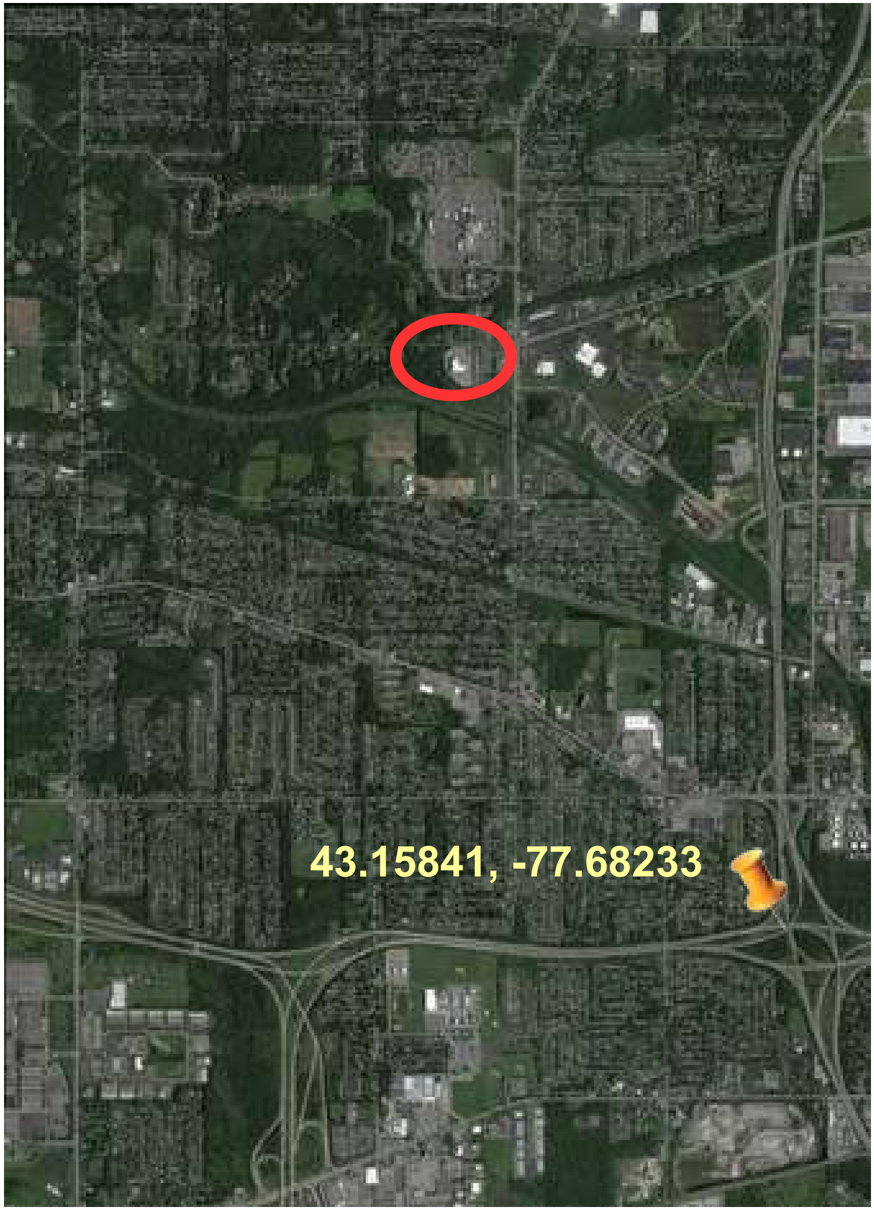
C.-C. Chen, C. A. Knoblock, and C. Shahabi, “Automatically conflating road vector data with orthoimagery,” GeoInformatica, vol. 10, no. 4, pp. 495–530, 2006.

W. Song, J. Keller, T. Haithcoat, and C. Davis, “Automated geospatial conflation of vector road maps to high resolution imagery,” IEEE Trans. Image Proc., vol. 18, no. 2, pp. 388–400, Feb 2009.

Prior Approaches

- Image-feature based matching (e.g. SIFT, SURF, ...) with auxiliary geo-referenced image (e.g. Google Maps) + parametric transform estimation
- Automated feature matching is challenging
  - Different (Illumination, capturing times, view point), and occlusion
  - Different imaging modality (e.g. infra-red)





Prior Approaches

- Image-feature based matching (e.g. SIFT, SURF, ...) with auxiliary geo-referenced image (e.g. Google Maps) + parametric transform estimation
- Automated feature matching is challenging
  - Different (Illumination, capturing times, view point), and occlusion
  - Different imaging modality (e.g. infra-red)

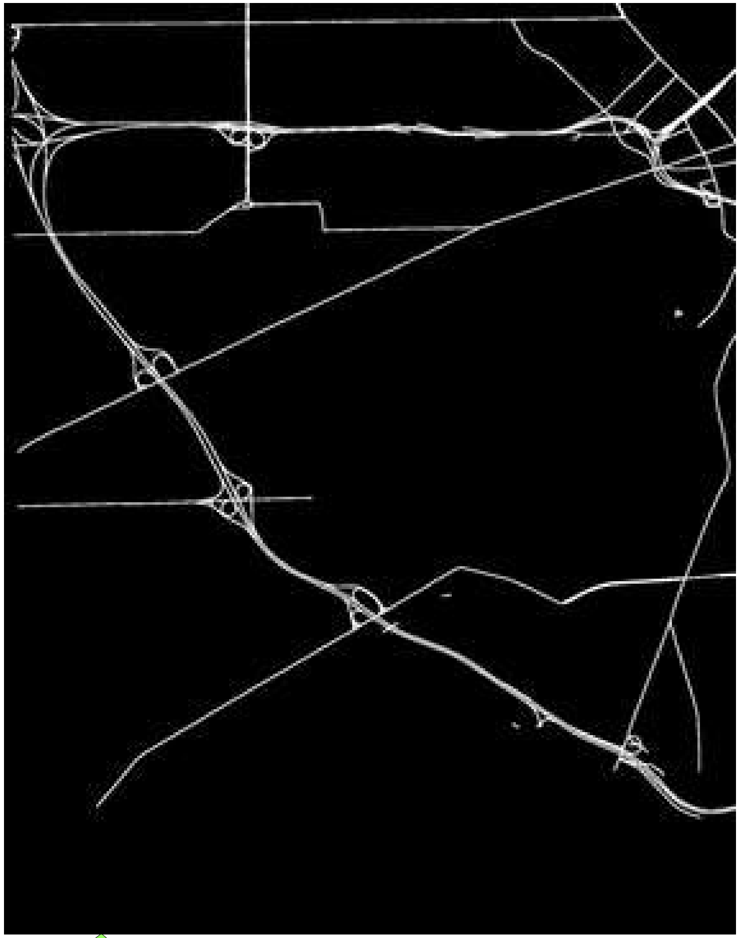
Agenda

- WAMI Introduction
- Roadmap Alignment Exploiting Vehicle Detections
  - Motivation, Challenges, Problem Statement
  - Prior Approaches and Challenges
  - Proposed Approach Exploiting Vehicle Detections
    - EM formulation
  - Results
- Benefits of EM Formulation

Proposed Approach Exploiting Vehicle Detections (Main Idea)

WAMI frame

Road network



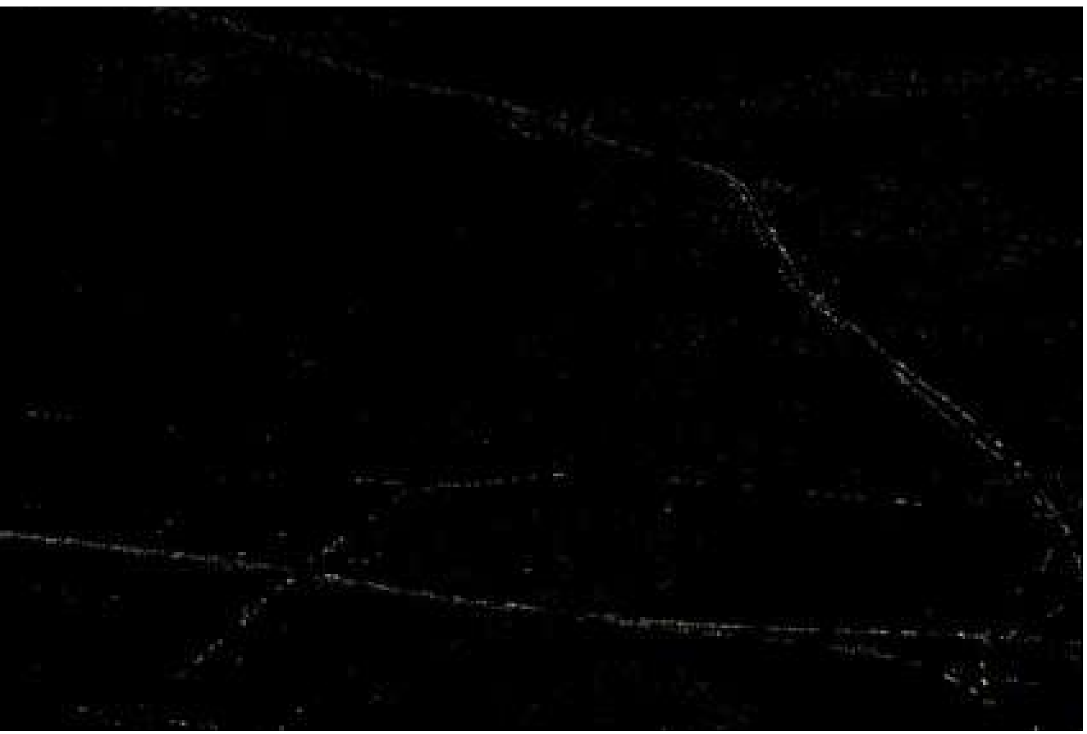
H ?

Proposed Approach Exploiting Vehicle Detections (Main Idea)

- Use vehicles' locations instead of WAMI frame.

Road network

Vehicle detections



Binary

H ?

Binary

Vehicle Detection

- Compensated frame difference I_1 (reference)

Vehicle Detection

- Compensated frame difference

$\mathbf{I}_2$ Registered to $\mathbf{I}_1$

Vehicle Detection

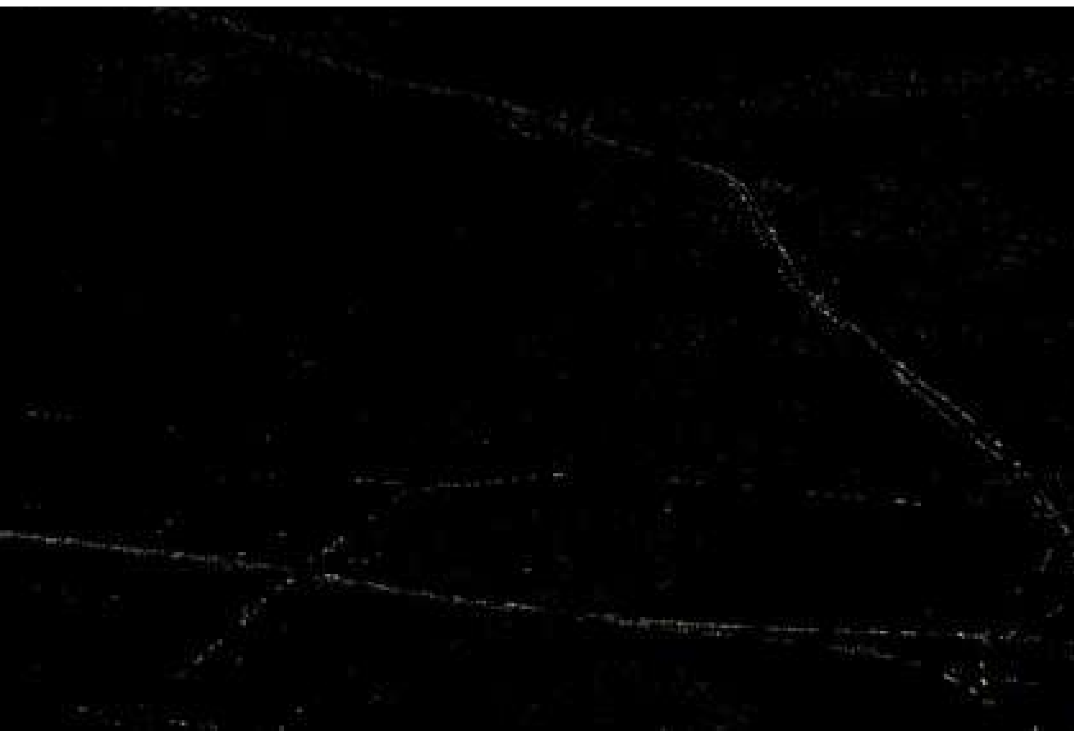
- Compensated frame difference $\mathbf{I_d}$

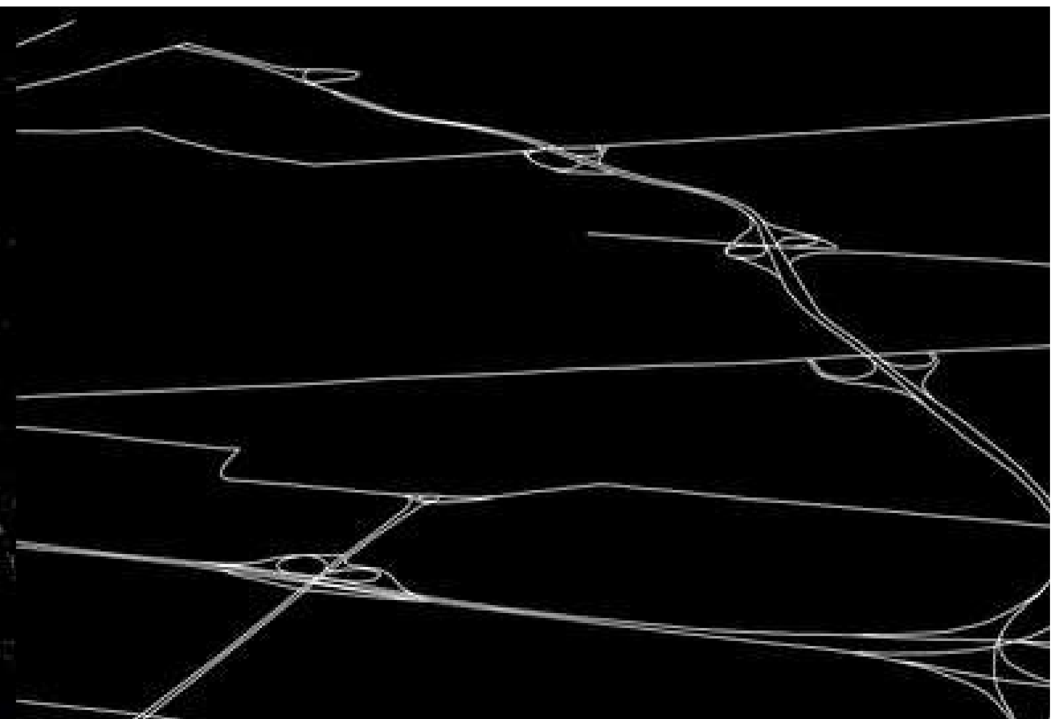
Proposed Approach Exploiting Vehicle Detections (Main Idea)

- Use vehicles' locations instead of WAMI frame.

Vehicle detections

Initial aligned road network with GPS/INS information





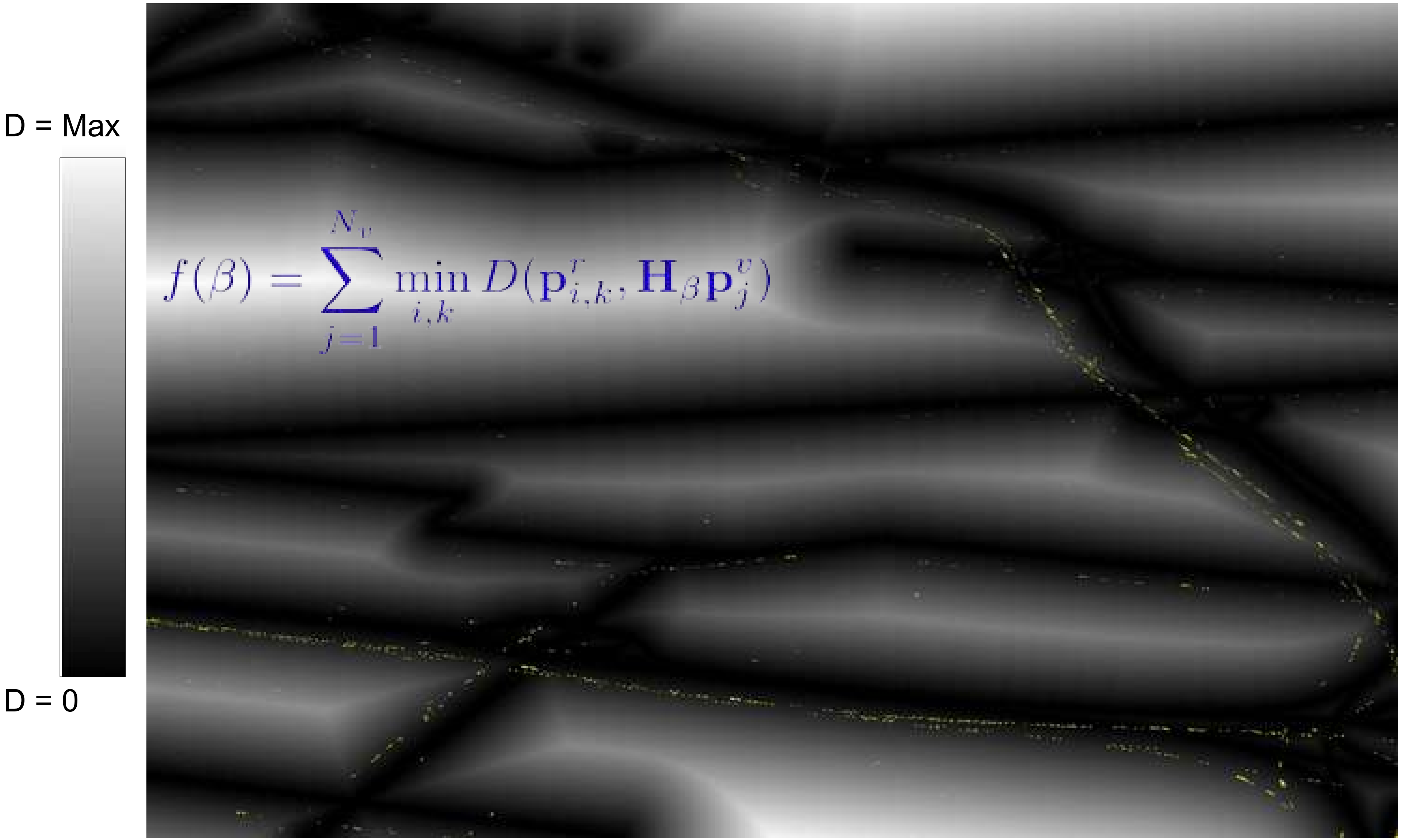
Binary

H ?

Binary

Proposed Approach Exploiting Vehicle Detections (Main Idea)

- Optimal alignment: minimize the chamfer distance between detected vehicle locations and the vector road network.
- Chamfer distances efficiently calculated by distance transform.

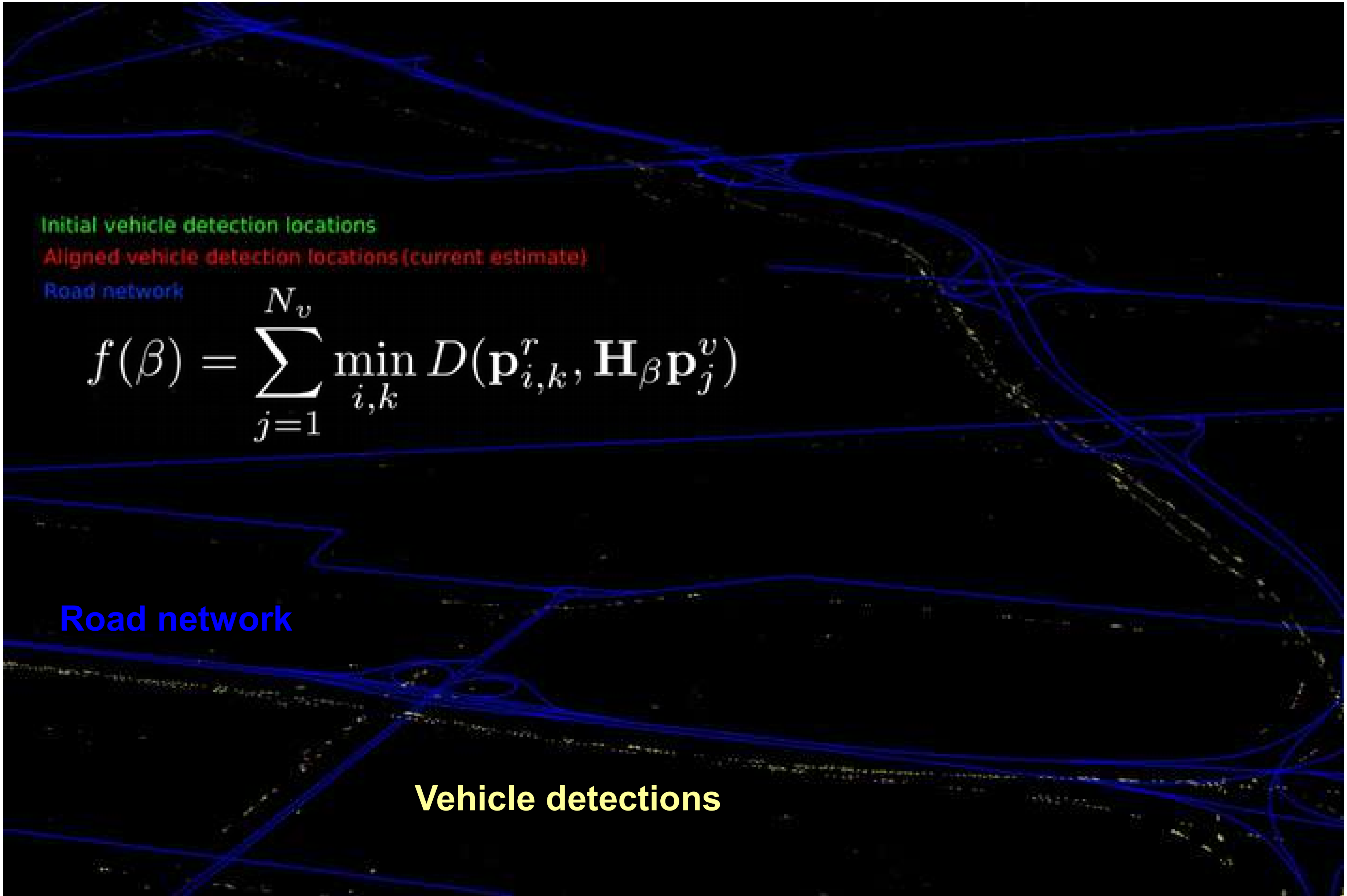


$$f(\beta) = \sum_{j=1}^{N_v} \min_{i,k} D(\mathbf{p}^r_{i,k}, \mathbf{H}_\beta \mathbf{p}^v_j)$$

Alignment Estimation by Chamfer Distance Minimization

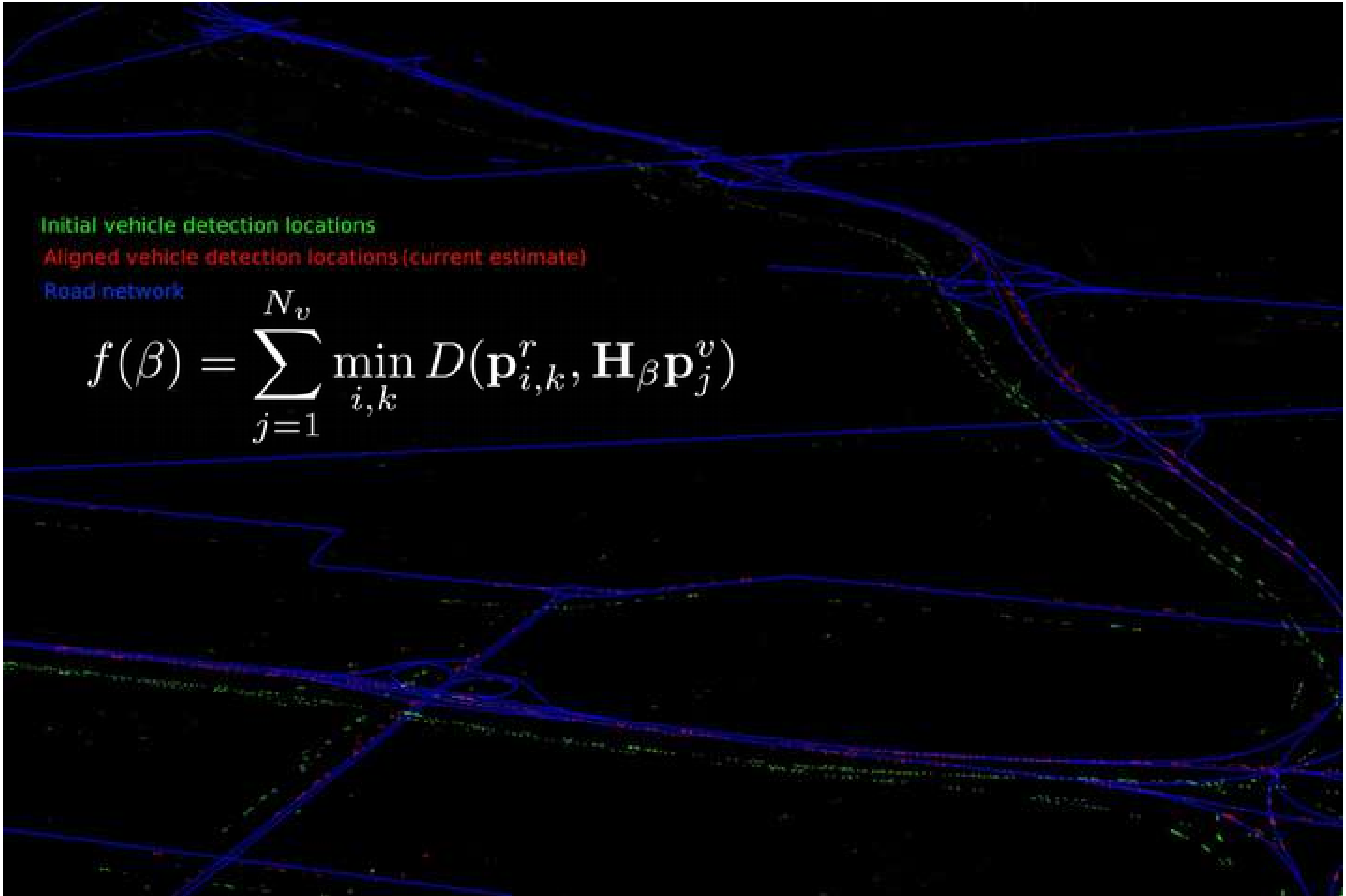
$\beta^* = \arg\min_{\beta} f(\beta)$

β^*: Levenberg-Marquardt (LM) optimization



Alignment Estimation by Chamfer Distance Minimization

β*



Final Alignment Result

What is the effect of the vehicle detection quality on alignment result?

Vehicle Detection

- Compensated frame difference I_1 (reference)

Vehicle Detection

- Compensated frame difference

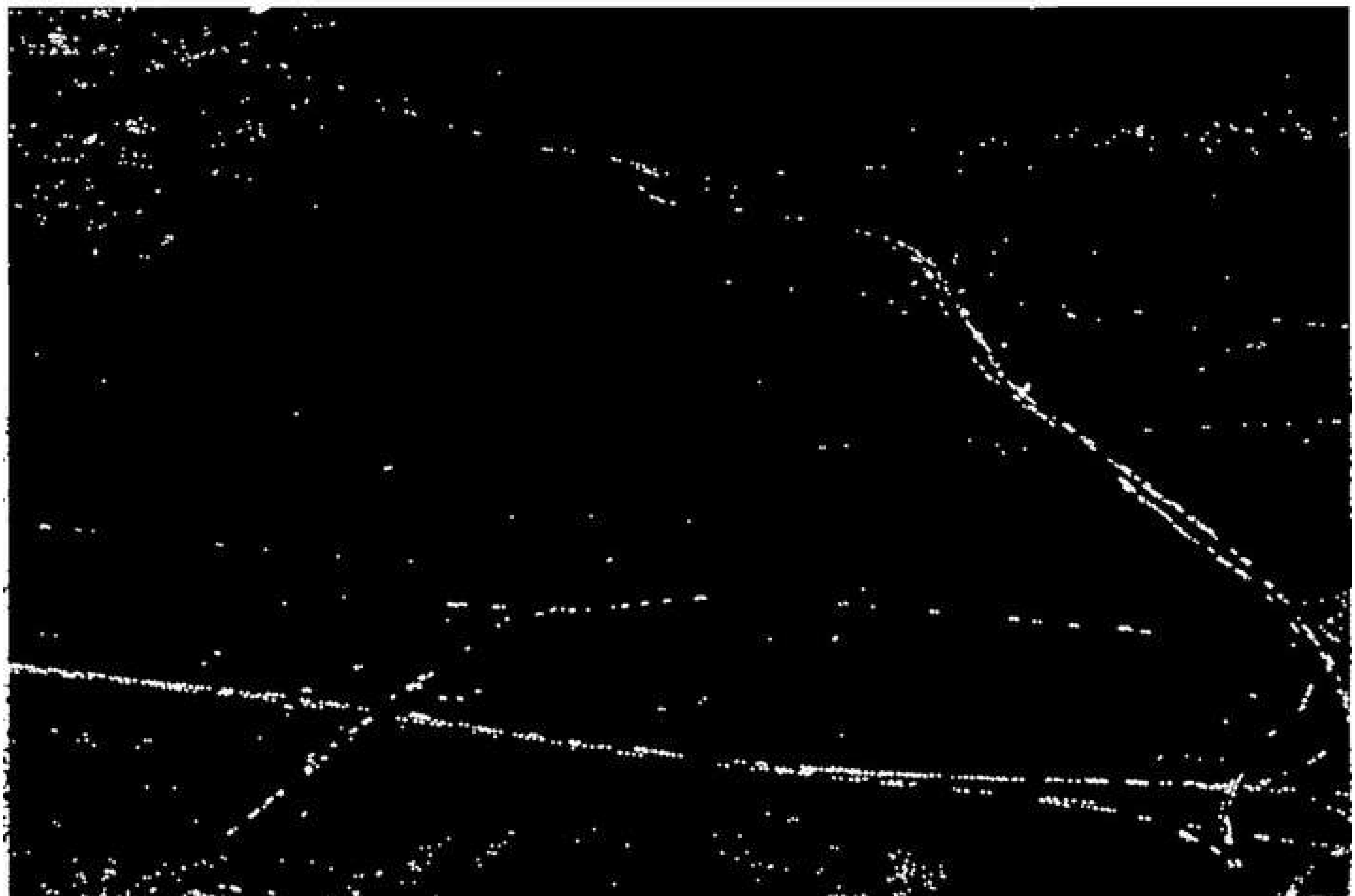
$\mathbf{I}_2$ Registered to $\mathbf{I}_1$

Vehicle Detection

- Compensated frame difference $\mathbf{I_d} > \tau$

Vehicle Detection

- Compensated frame difference $\mathbf{I}_d > \tau$



Final Alignment Result

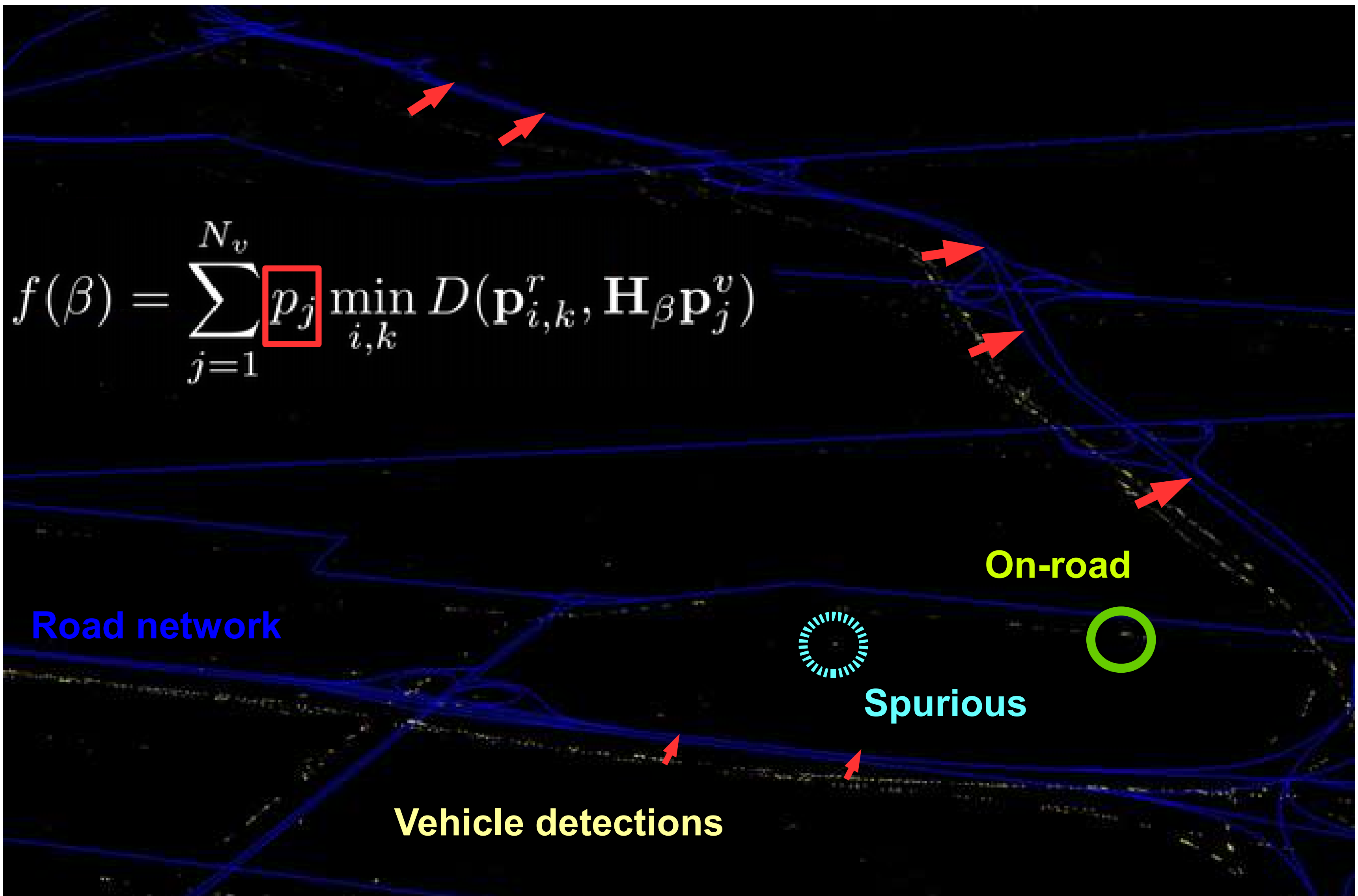
For low quality vehicle detections

EM Formulation

- Model in a probabilistic framework and use ML estimation
- Expect on-road vehicles to be close to road
  - Can model via appropriate distribution
- Challenge: detection may correspond to an “on-road vehicle” or may be “spurious”
  - Cannot directly model distribution of all vehicles as concentrated on roads
- Introduce latent random variables (EM)
  - One per detection, 1 if on road, 0 if spurious
  - IID distribution with unknown parameters

Second Phase: Handling Spurious Detections

- Optimal alignment: minimize the **weighted** chamfer distance between detected vehicle locations and the road network.
- Weight=posterior probability that a detection is on-road vehicle

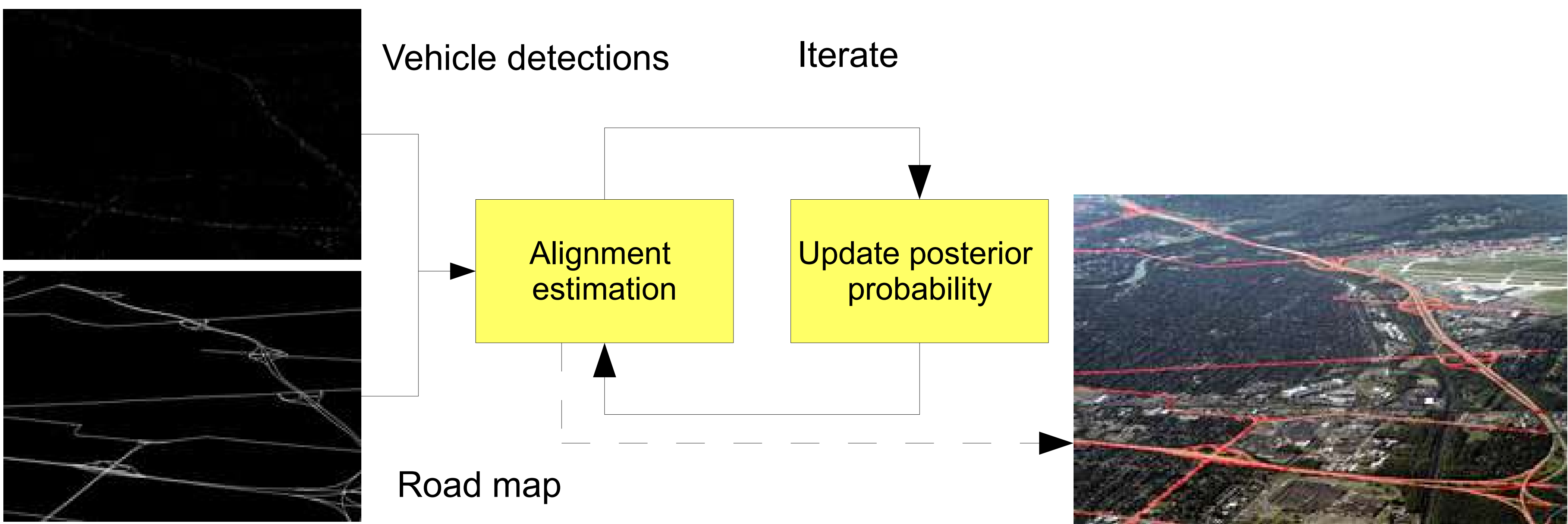


Challenge of Spurious Detections

- Classic “chicken and egg” problem:
  - If we knew which detections were spurious, we could exclude them from computing the alignment between the WAMI frame and roadmap and obtain good alignment
  - If we knew the alignment between the WAMI frame and the roadmap, we could identify which detections are spurious
- How do you address both simultaneously?
  - Iteratively, using a probabilistic framework (EM)

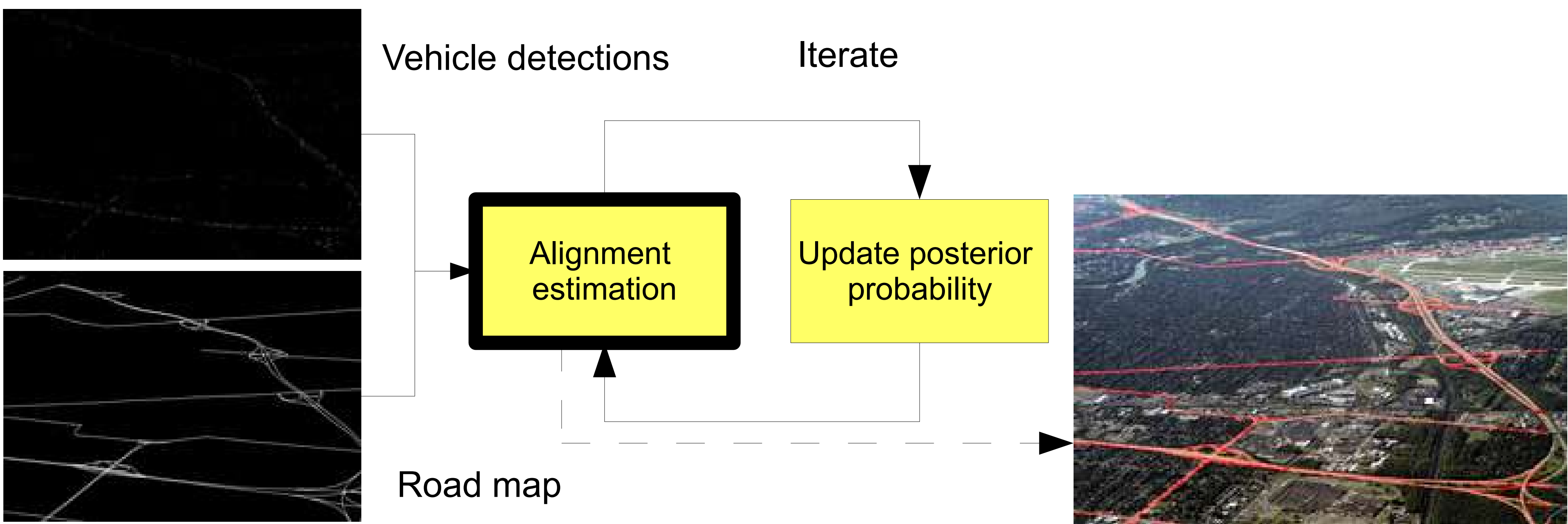
Second Phase: Handling Spurious Detections

- Optimal alignment: minimize the **weighted chamfer** distance between detected vehicle locations and the road network.
- Weight=posterior probability of detection to be on-road vehicle
- Use EM framework
  - E step: Update posterior probabilities.
  - M step: Re-estimate the alignment by minimize weighted distance



Second Phase: Handling Spurious Detections

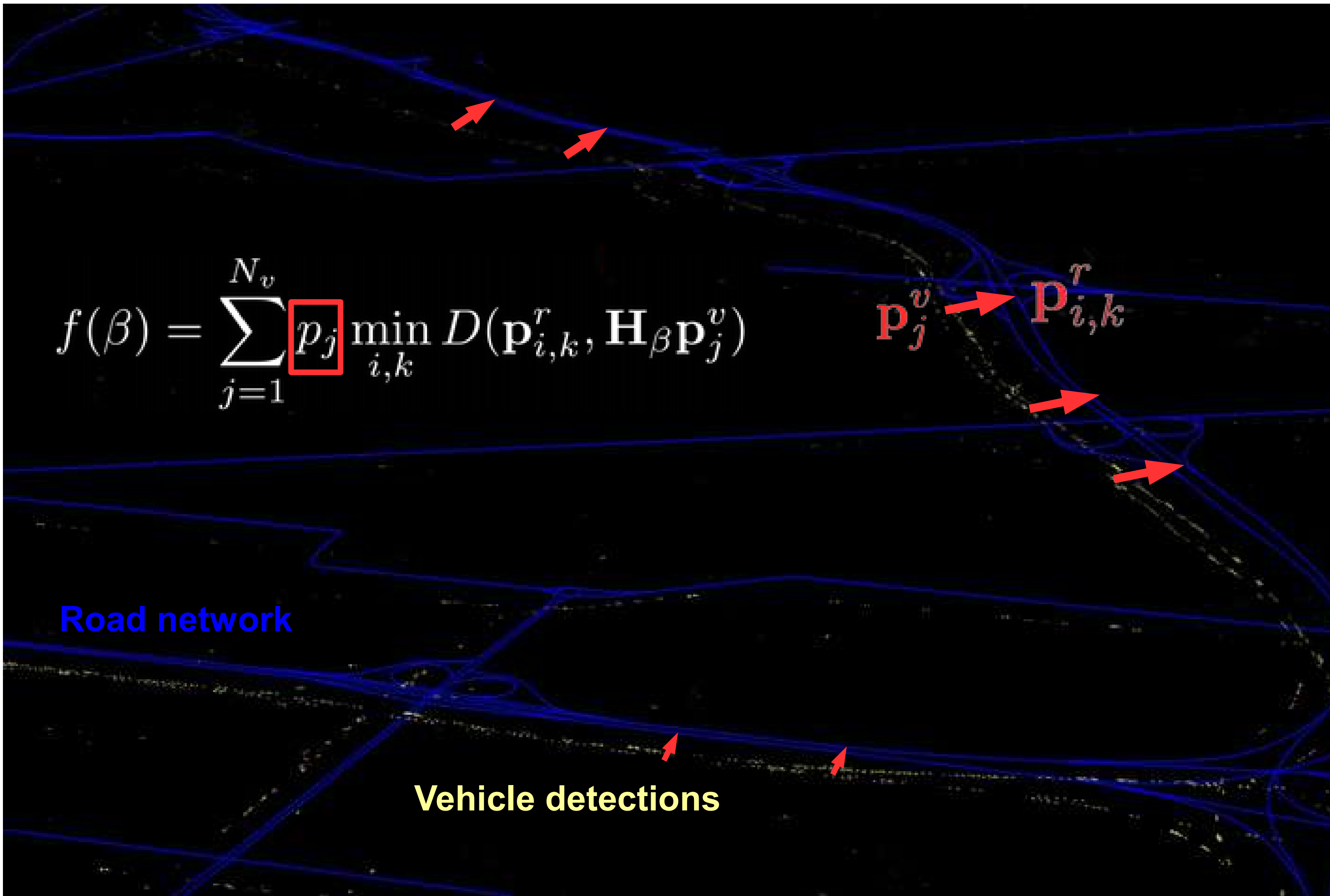
- Optimal alignment: minimize the **weighted chamfer** distance between detected vehicle locations and the road network.
- Weight=posterior probability of detection to be on-road vehicle
- Use EM framework
  - E step: Update posterior probabilities.
  - M step: Re-estimate the alignment by minimize weighted distance



Alignment Estimation by Weighted Chamfer Distance Minimization (M step)

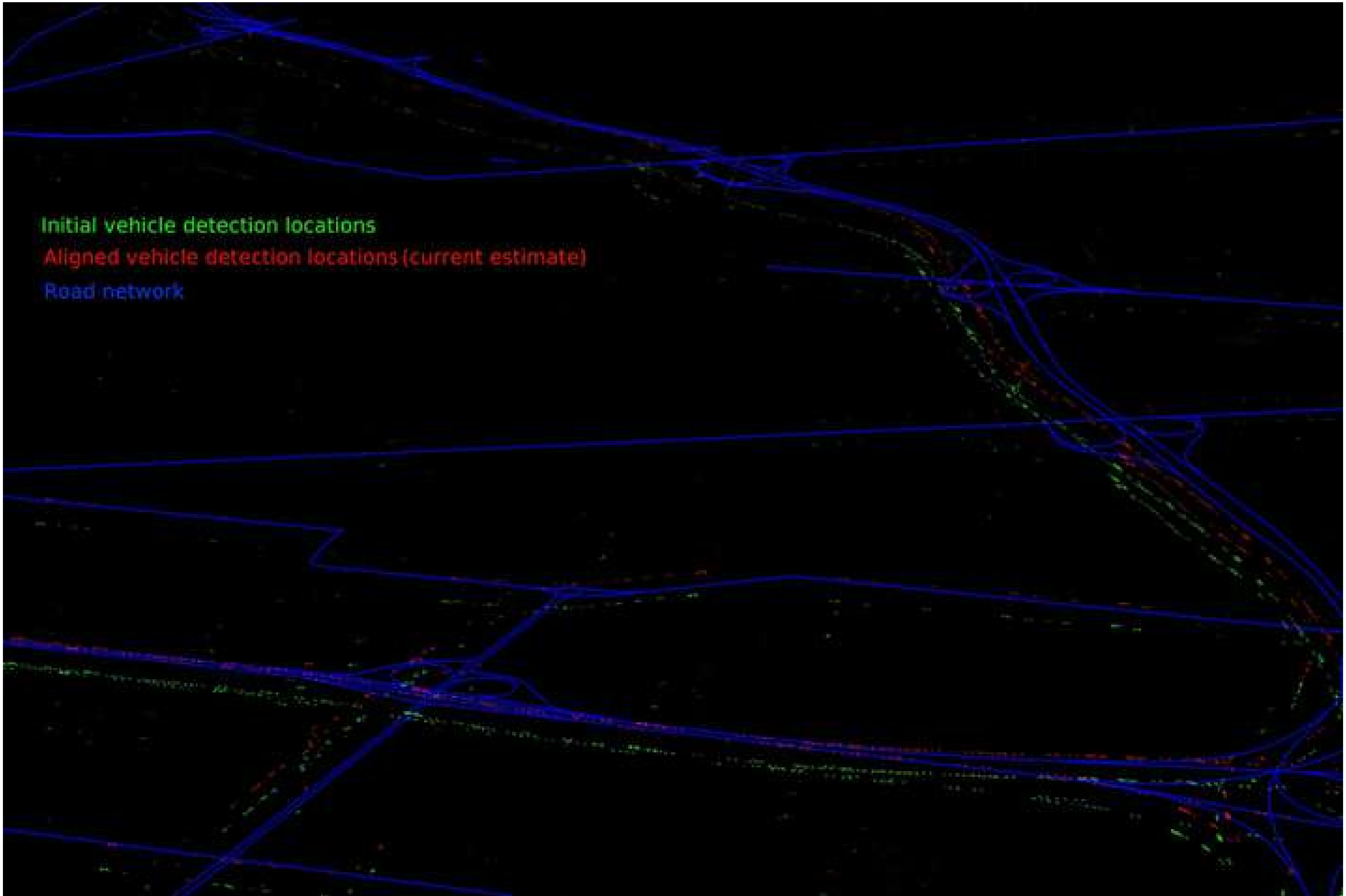
$\beta^* = \arg\min_{\beta} f(\beta)$

β^*: Levenberg-Marquardt (LM) optimization



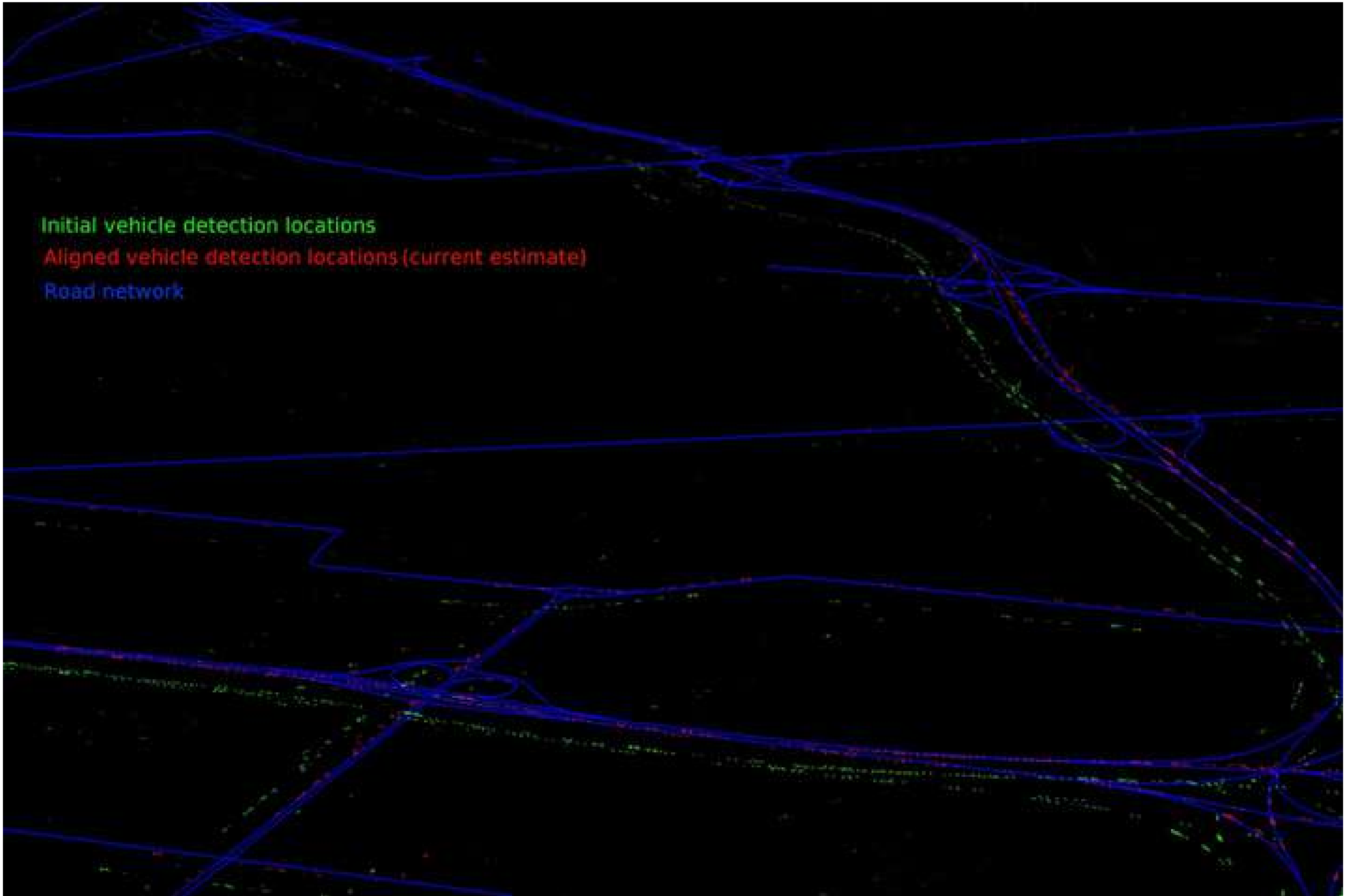
Alignment Estimation by Weighted Chamfer Distance Minimization (M step)

$\beta^{(1)}$



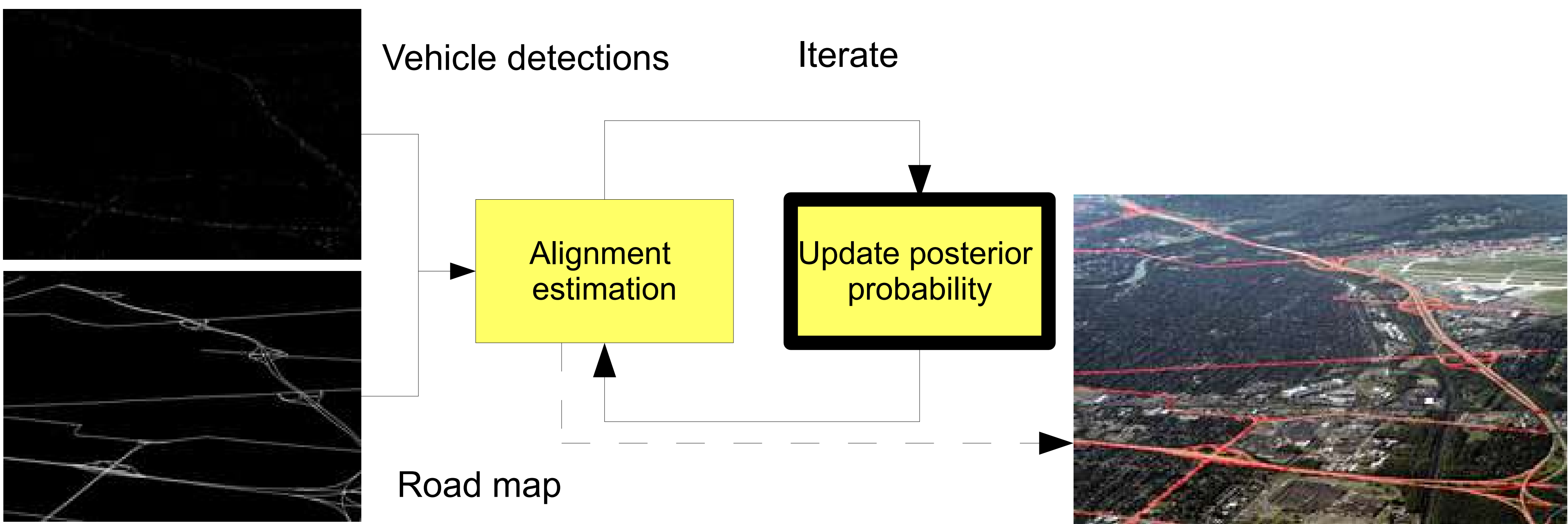
Alignment Estimation by Weighted Chamfer Distance Minimization (M step)

β^*

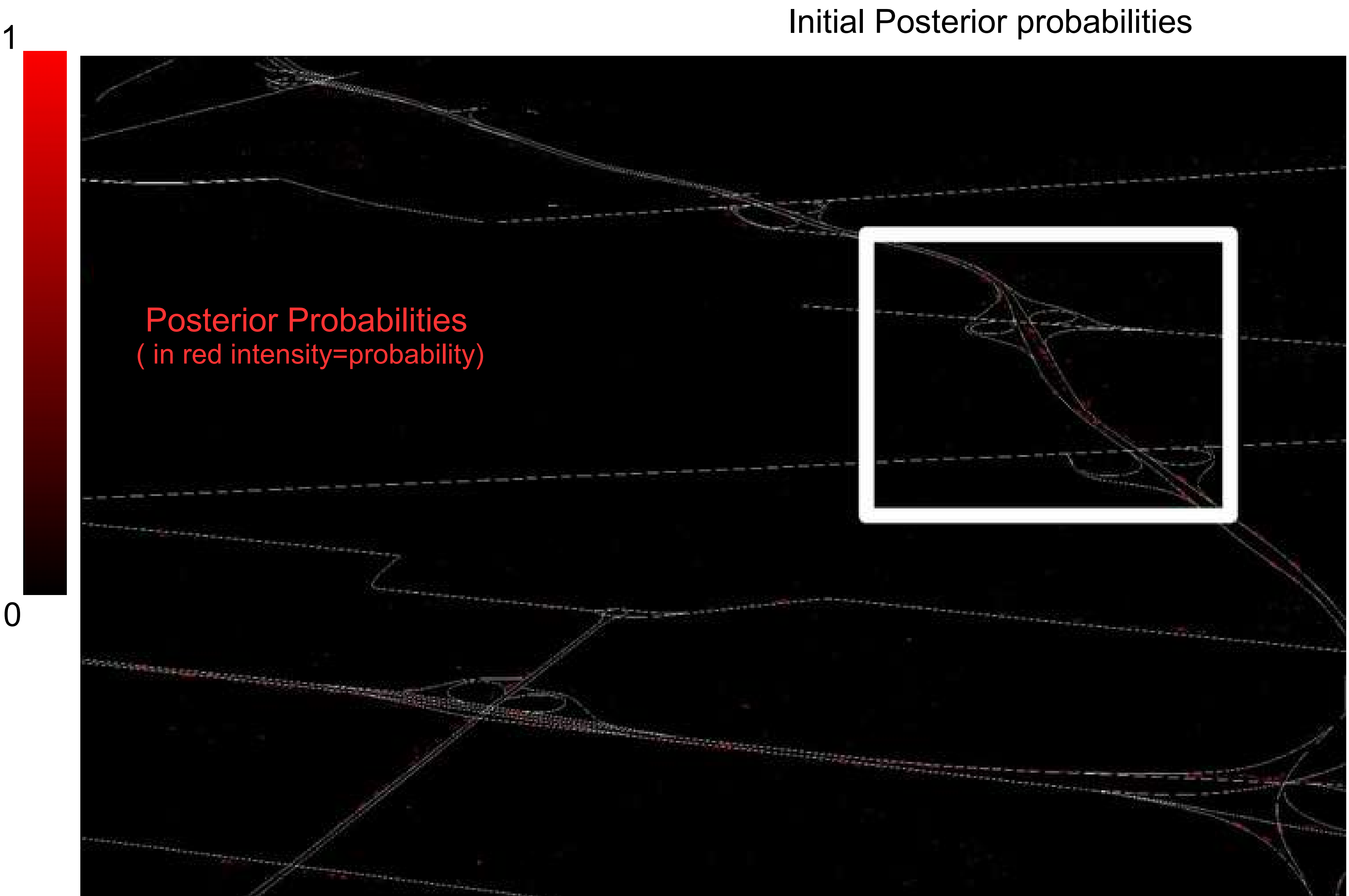


Proposed Approach Exploiting Vehicle Detections

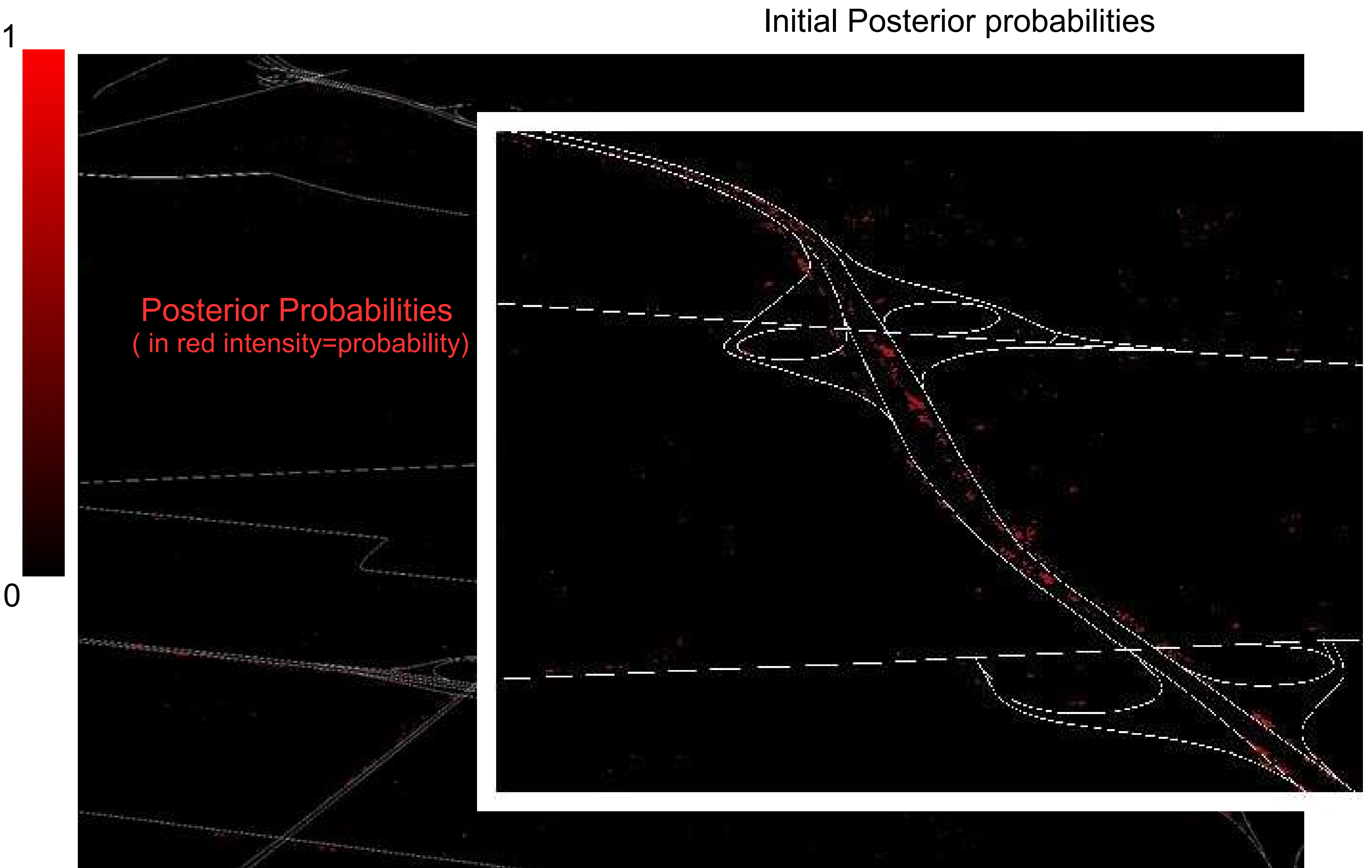
- Optimal alignment: minimize the **weighted chamfer** distance between detected vehicle locations and the road network.
- Weight=posterior probability of detection to be on-road vehicle
- Use EM framework
  - E step: Update posterior probabilities.
  - M step: Re-estimate the alignment by minimize weighted distance



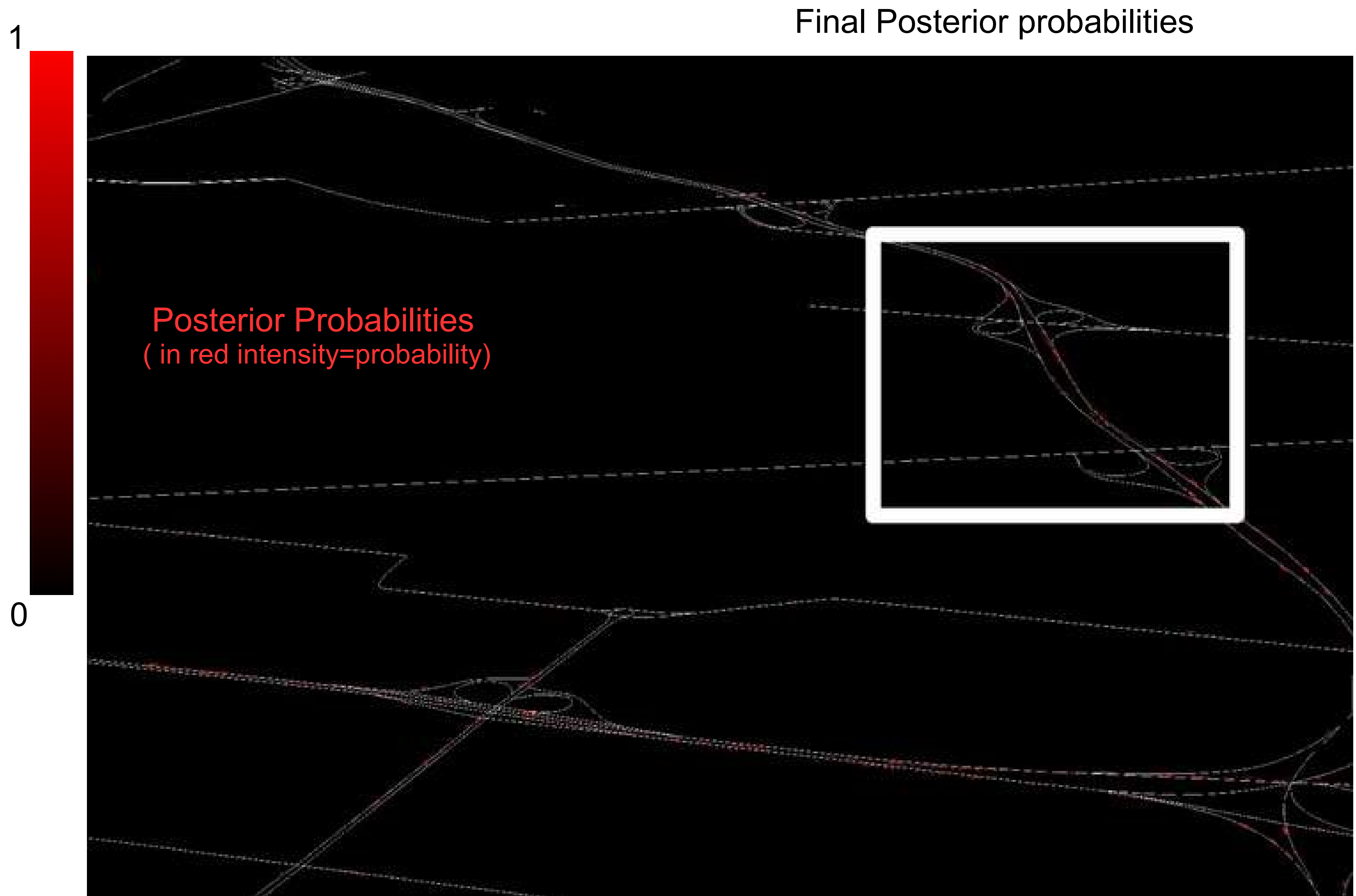
Update Posterior Probabilities
(E step)



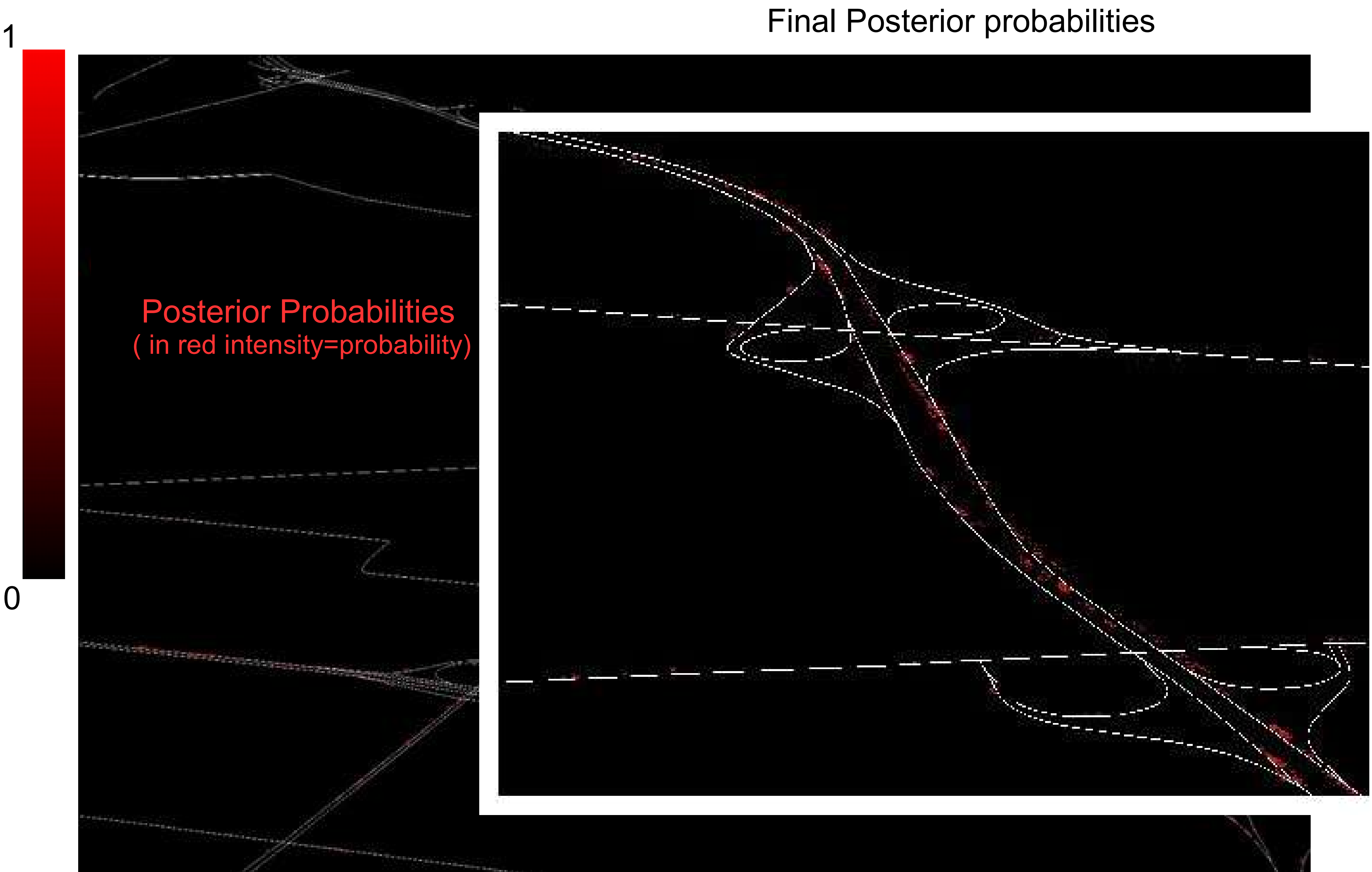
Update Posterior Probabilities (E step)



Update Posterior Probabilities (E step)



Update Posterior Probabilities
(E step)



Final Alignment Result

Agenda

- WAMI Introduction
- Roadmap Alignment Exploiting Vehicle Detections
  - Motivation, Challenges, Problem Statement
  - Prior Approaches and Challenges
  - Proposed Approach Exploiting Vehicle Detections
    - EM formulation
  - Results
- Benefits of EM Formulation

Experimental Results: Data

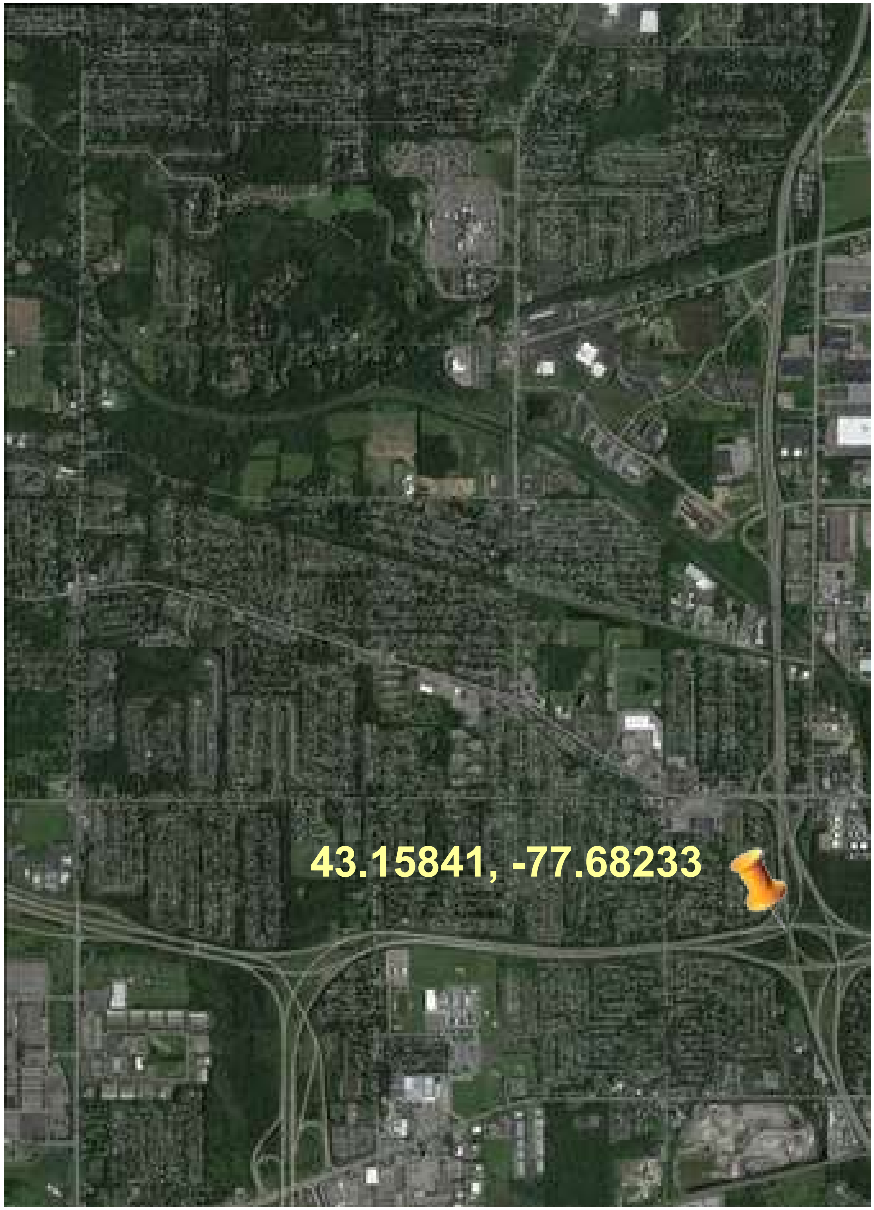
- Vector road map from OpenStreetMap (OSM)
  - Road network + associated meta-data
    - Direction of travel, speed limits, etc
- WAMI data set recorded using CorvusEye 1500 Wide-Area Airborne System.
  - Individual WAMI frames as NITF 2.1 format files
    - JPEG 2000 image + associated meta-data
  - 4400x6600 pixels
  - 2 fps

Experimental Results: Methods

- Proposed method based on vehicle movements (Proposed)
- Alternative methods (for comparison):
  - Meta-data Based Alignment (MBA)
    - Equivalent to initial alignment
  - SIFT matching with auxiliary geo-referenced image (SBA)
  - Preliminary version of this work [1] : perform alignment without EM (assume no spurious detections)

SIFT based matching with auxiliary geo-referenced image (SBA)

- WAMI Frame
- Geo-referenced Image

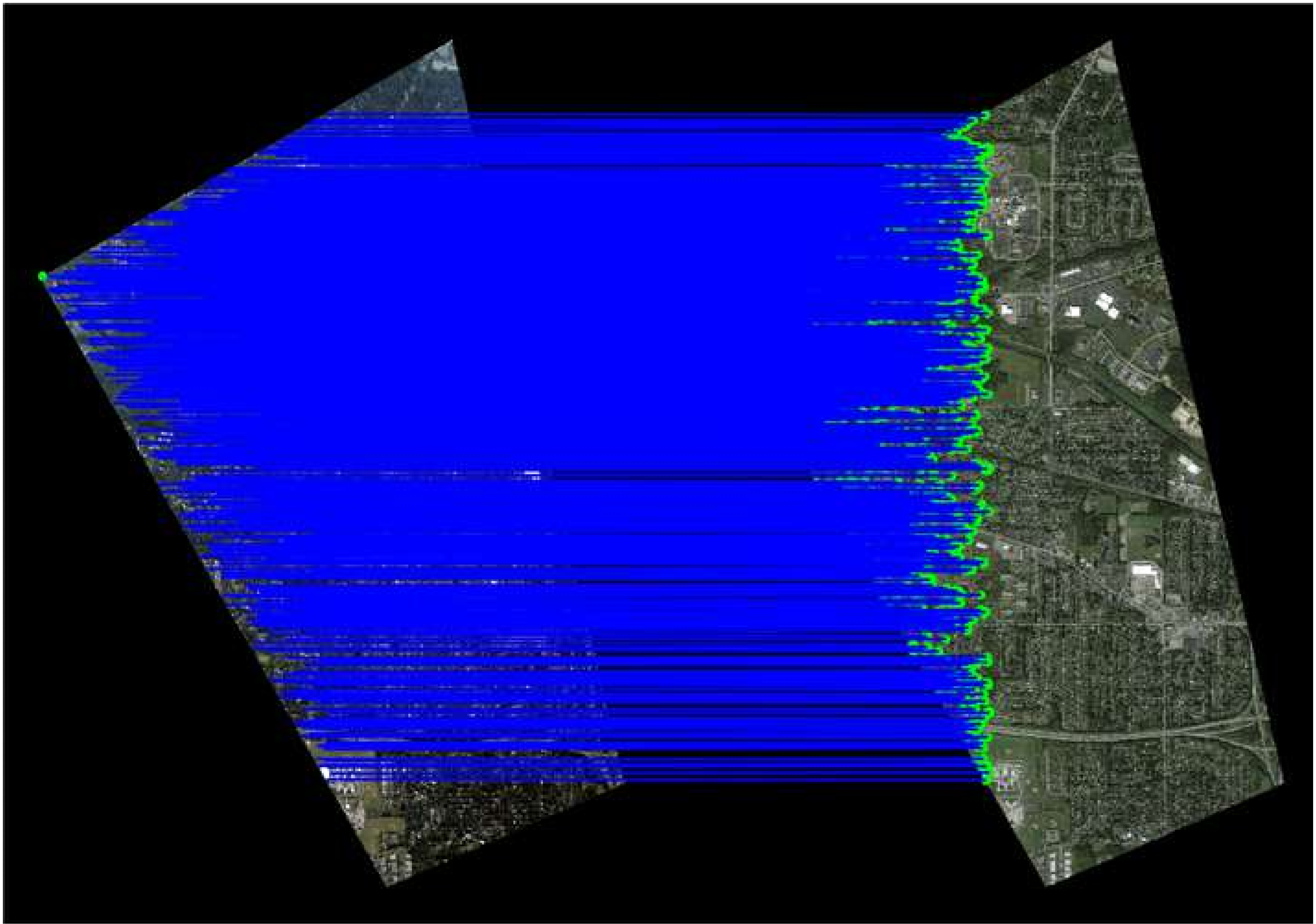


SIFT Based Alignment (SBA)

- SIFT with auxiliary image.

SIFT Based Alignment (SBA)

- SIFT with auxiliary image.



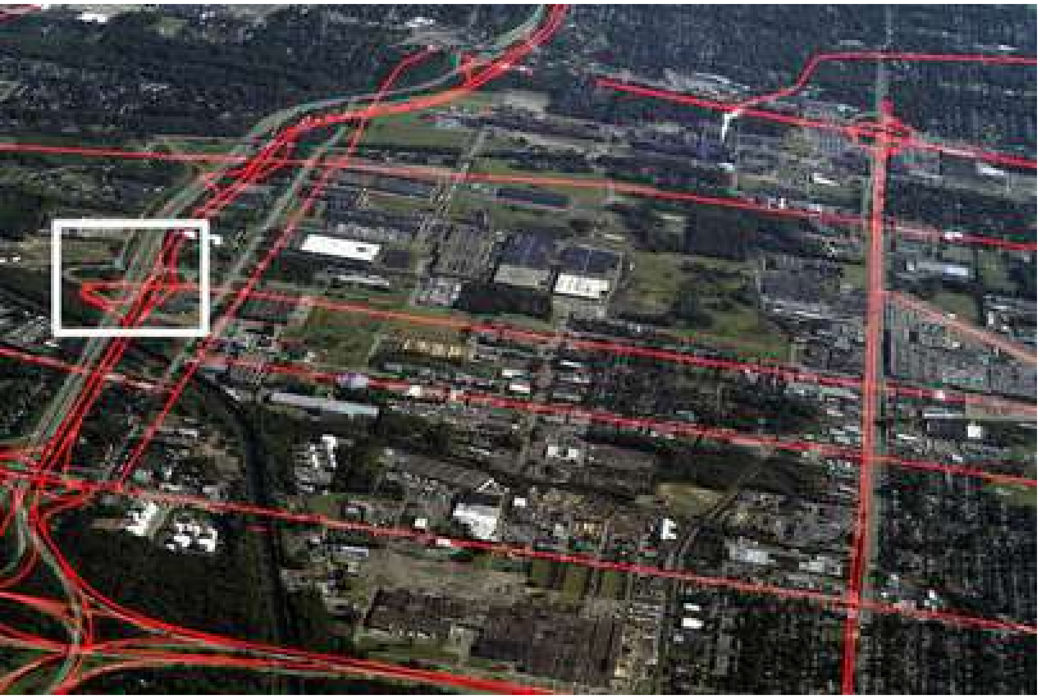
SIFT Based Alignment (SBA)

- Sample SIFT Correspondences

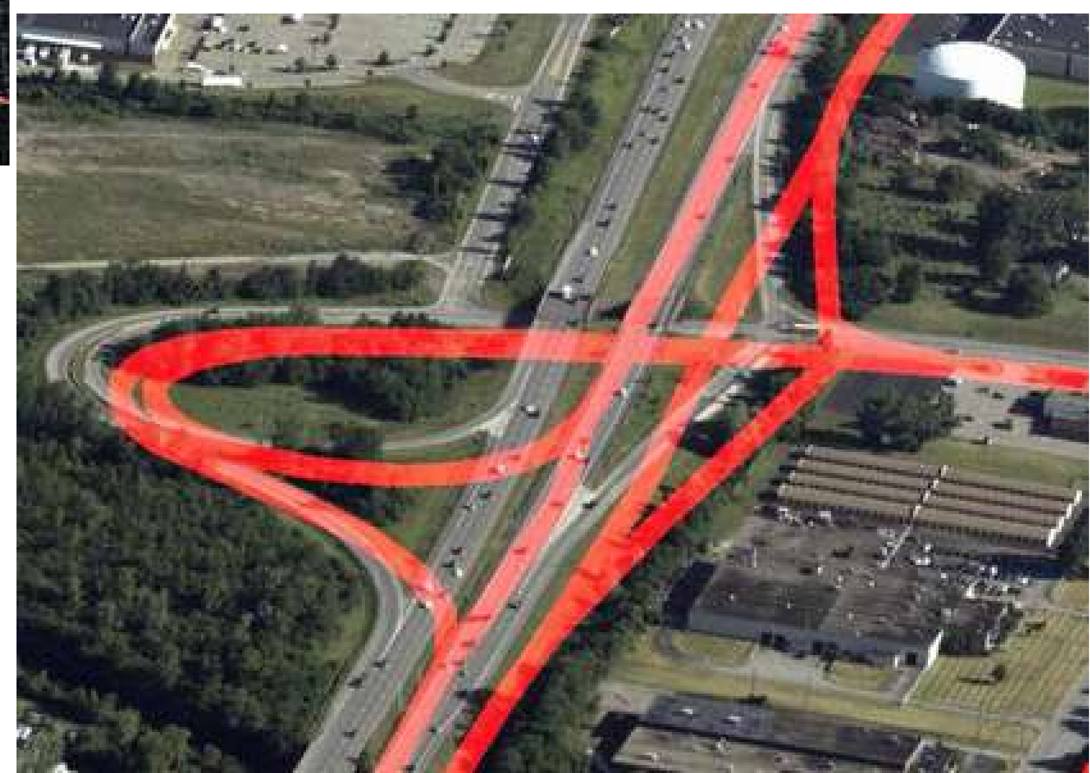
SIFT Based Alignment (SBA)

- SIFT with auxiliary image.

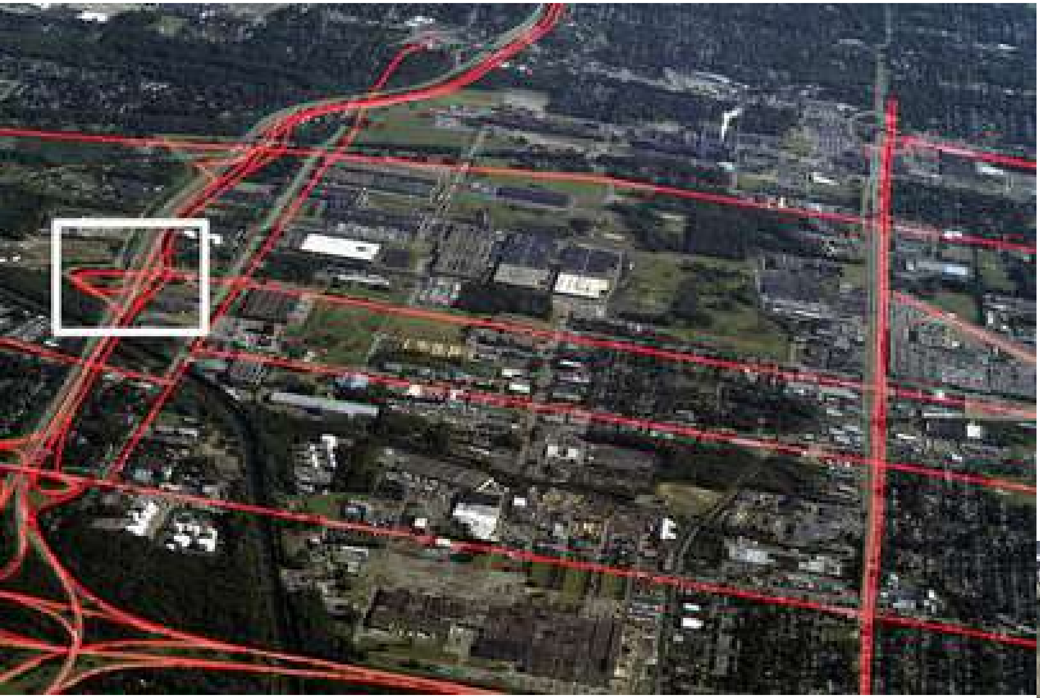
Visual Results (Frame 1)



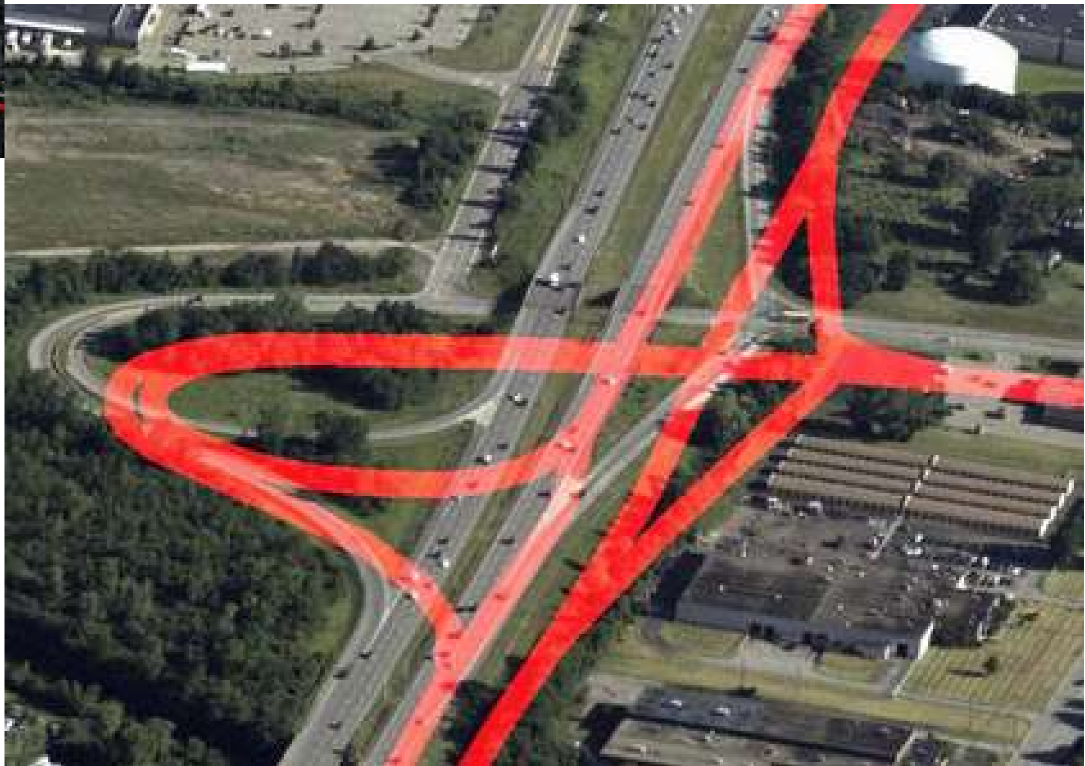
Meta-data Based Alignment (MBA)



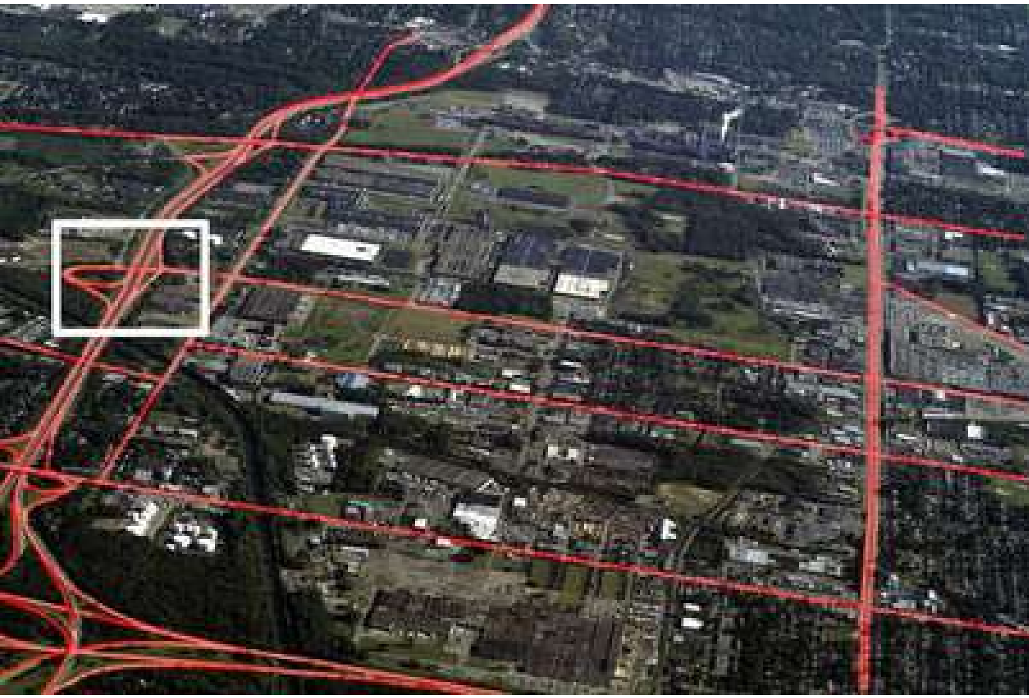
Visual results (Frame 1)



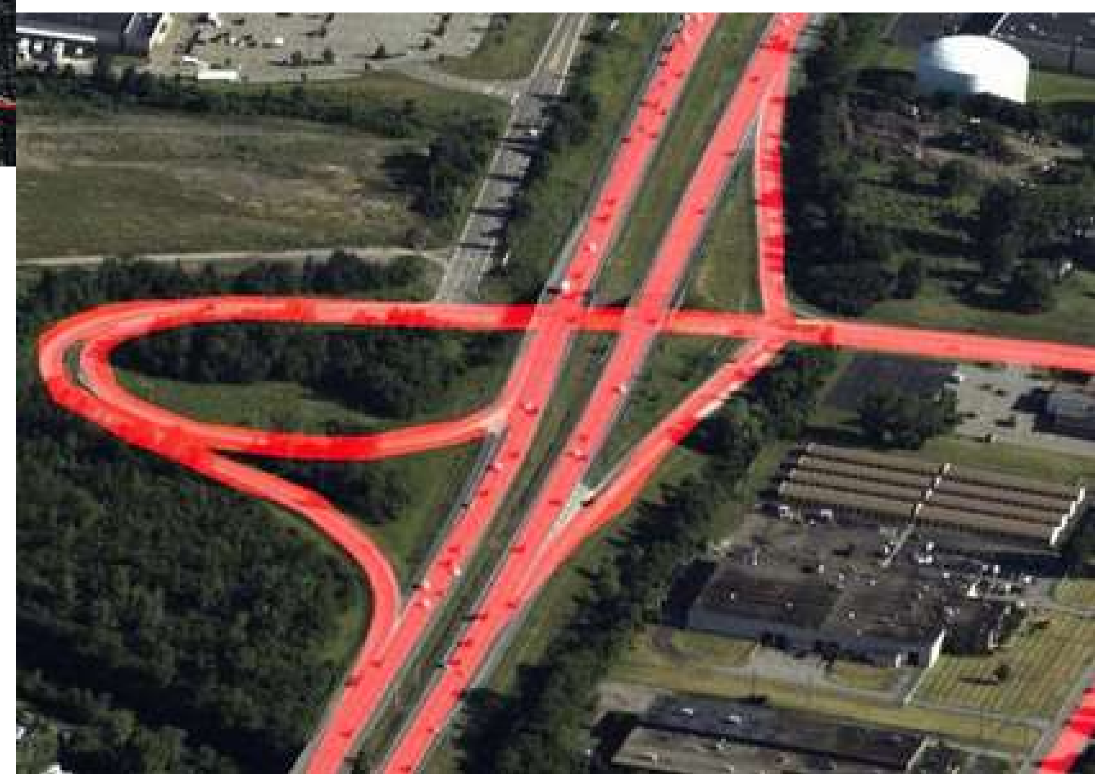
SIFT Based Alignment
(SBA)



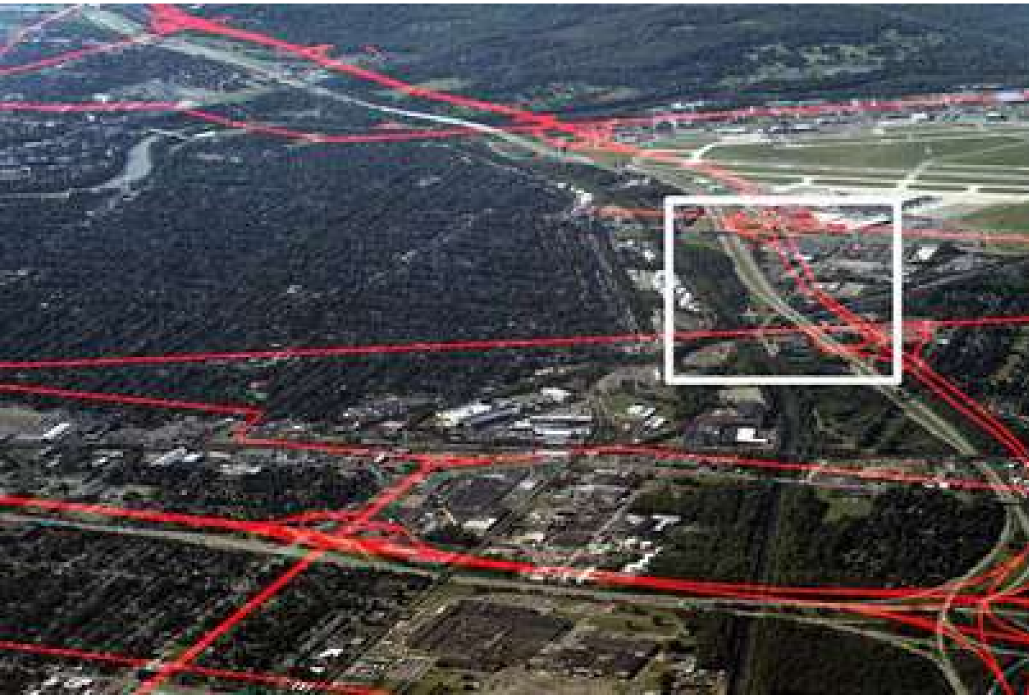
Visual Results (Frame 1)



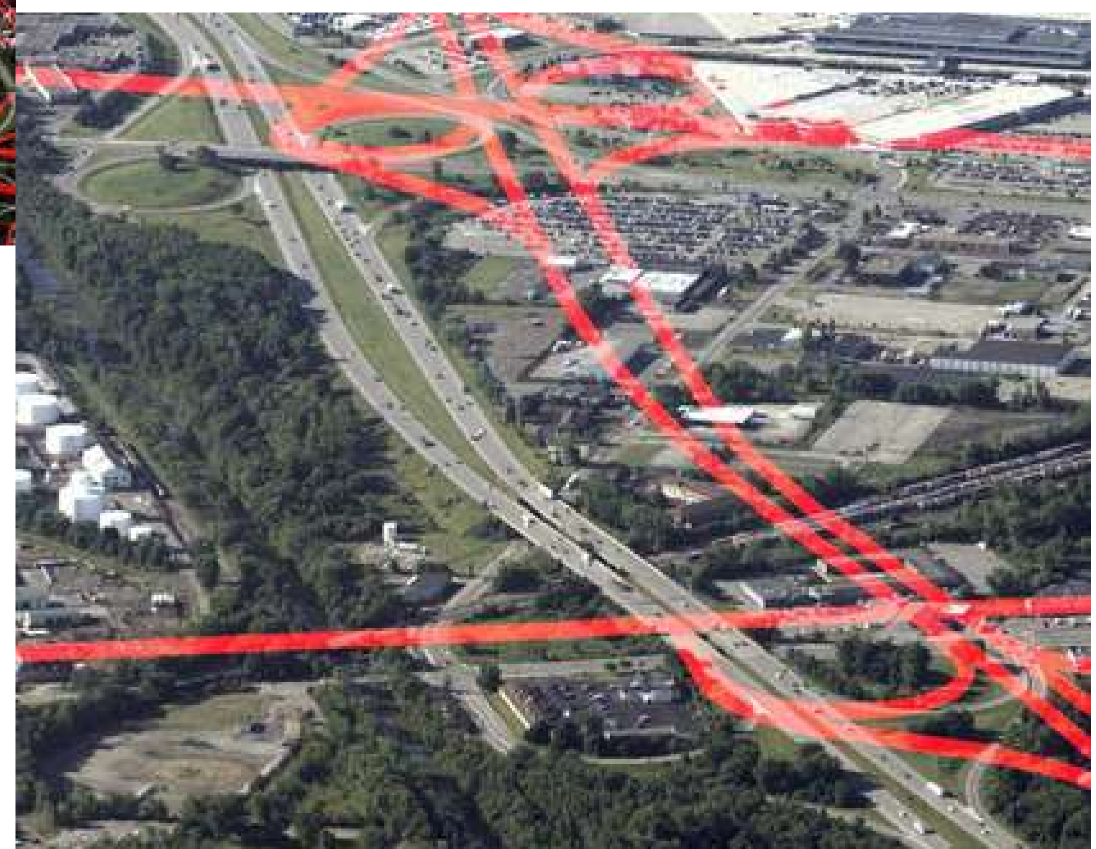
Proposed



Visual Results (Frame 820)

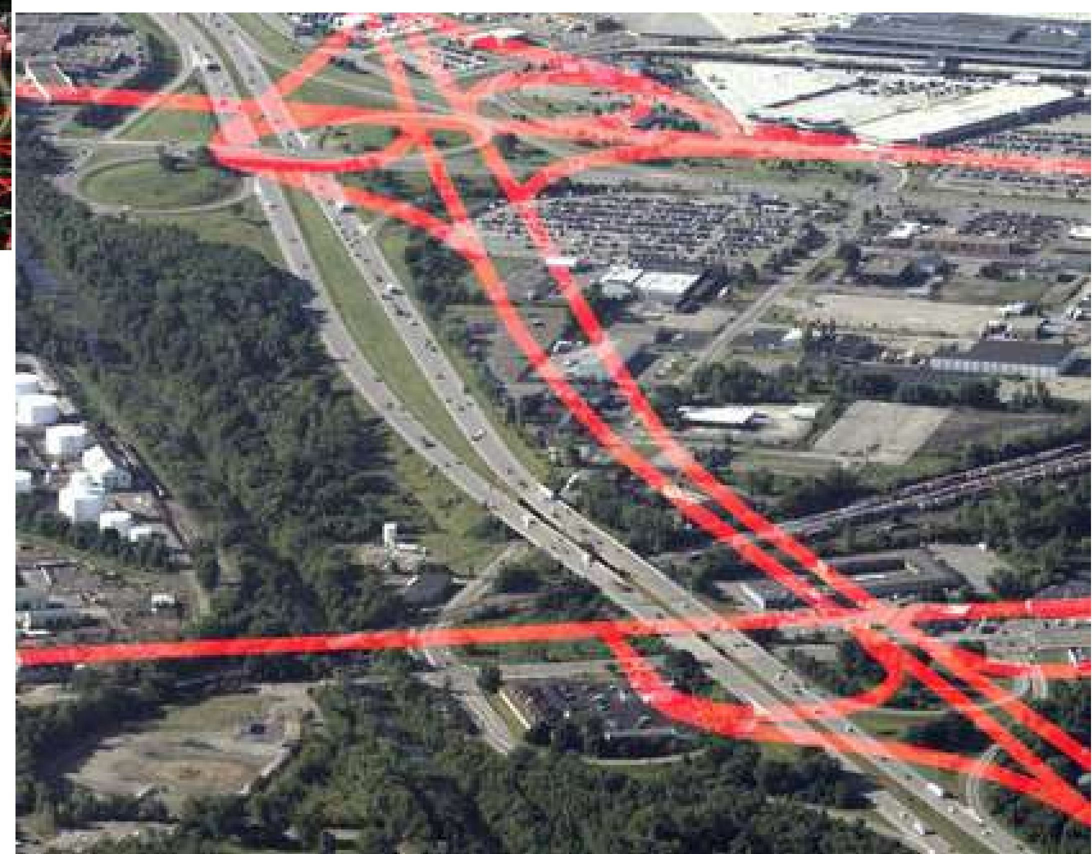


Meta-data Based Alignment (MBA)

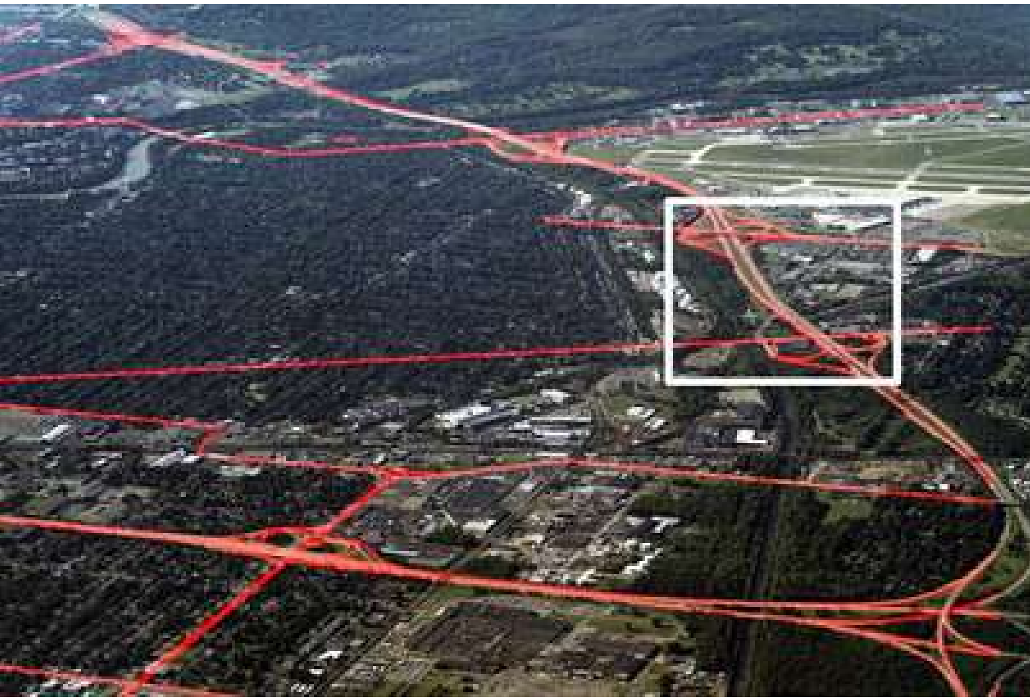


Visual Results (Frame 820)

SIFT Based Alignment
(SBA)



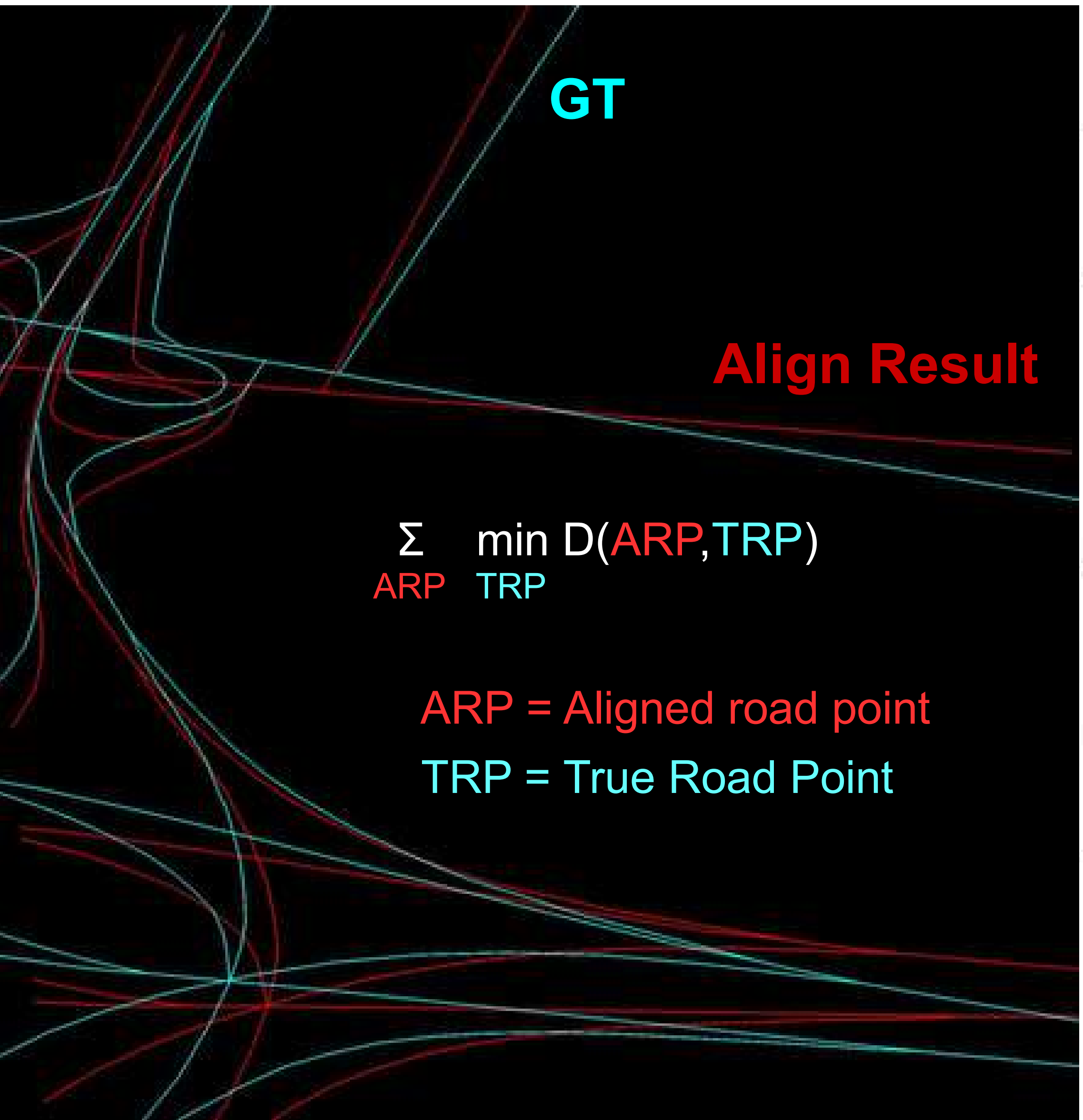
Visual Results (Frame 820)



Proposed

Quantitative Results

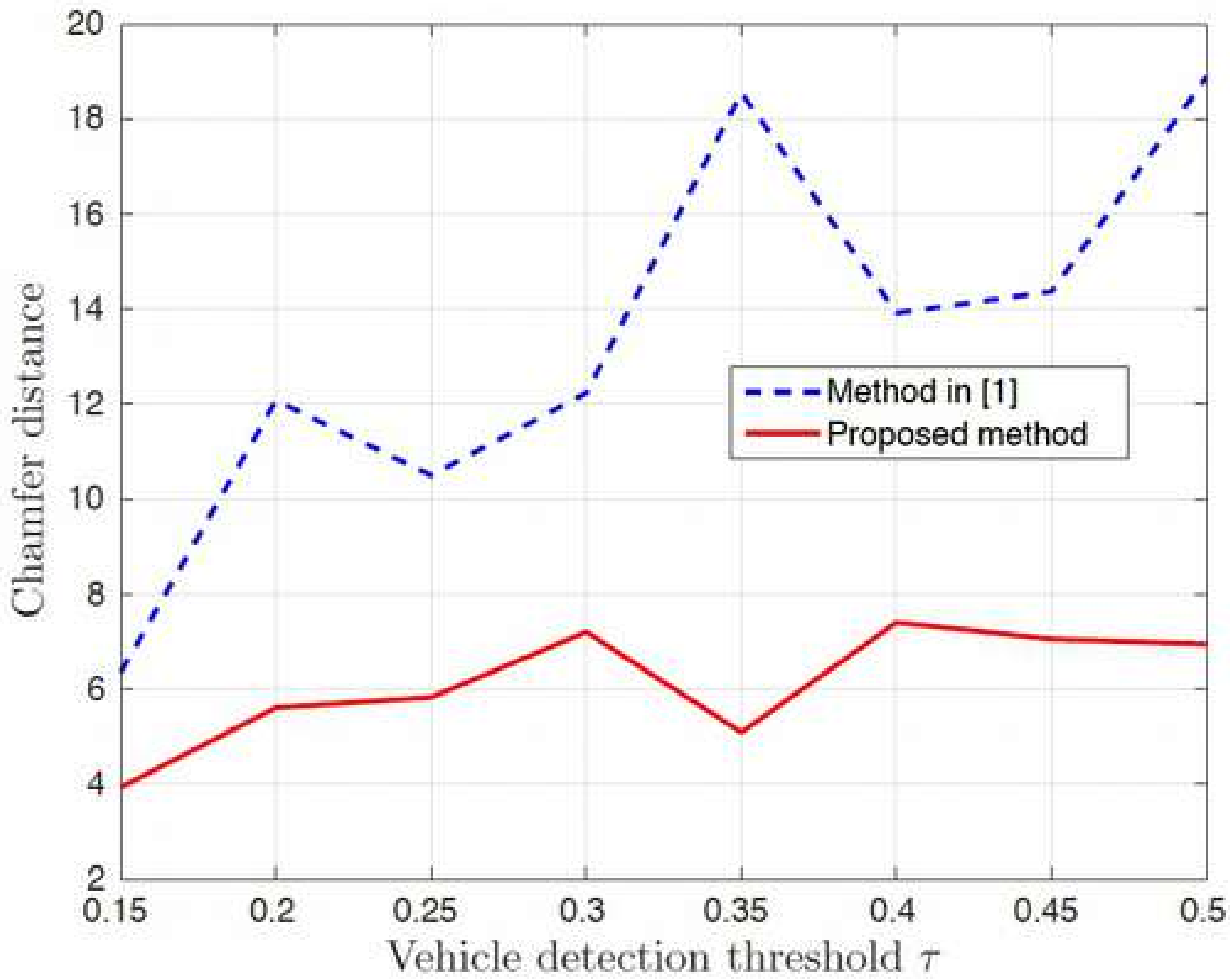
1- Chamfer distance between estimated and ground truth (manually labeled) road network



Dataset	Test area	MBA	SBA	method in [1]	Proposed method
CORVUS (V)	Area 1	28.22	17.1	6.36	**3.95**
	Area 2	122.28	83.09	9.30	**2.07**
	Area 3	36.95	26.49	8.69	**3.45**
	Area 4	87.35	87.29	6.68	**5.21**
CORVUS (IR)	Area 5	450.19	462.76	3.15	**2.13**
	Area 6	104.28	387.84	4.25	**2.14**
	Area 7	179.13	266.85	5.12	**3.11**
	Area 8	81.38	116.37	17.94	**11.34**

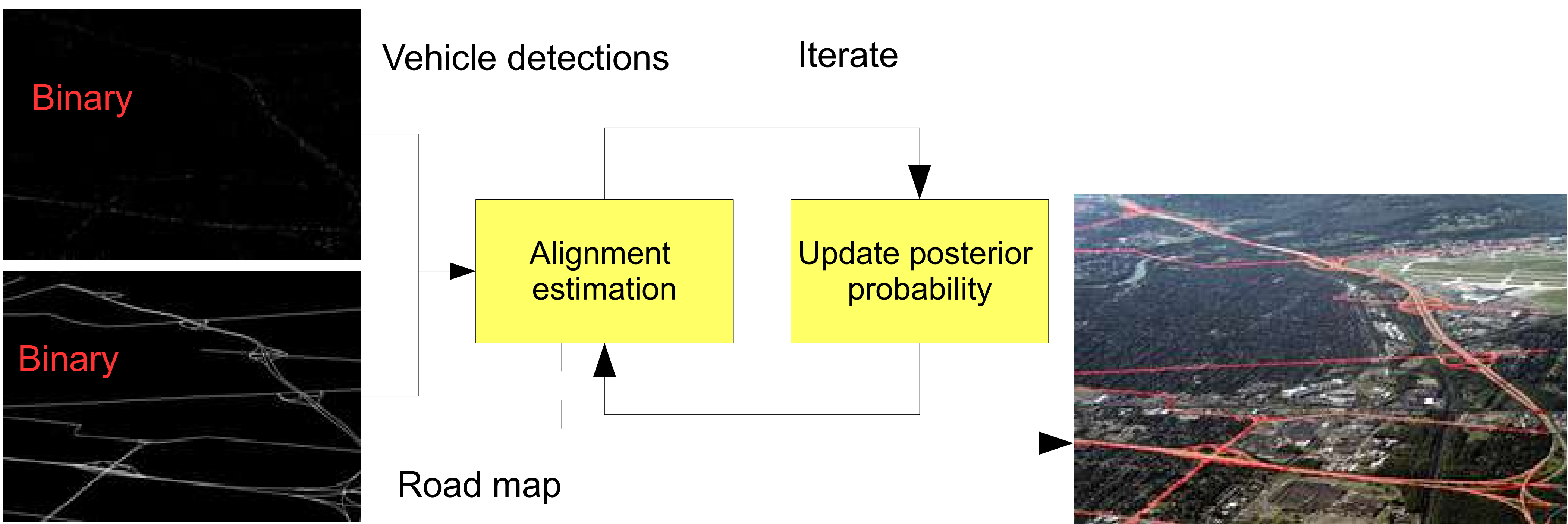
Quantitative Results

2- Robustness to variations in quality of vehicular detections



Proposed Approach : Contributions

- Exploits vehicle detections for efficient vector road map to WAMI alignment using chamfer matching
  - Overcomes problem of different data representations (raster vs. vector) and associated feature matching challenges
  - Directly applies for alternative modalities (e.g. visible vs. night-time infra-red)
- Better for parametric alignment: Road network lies on near planar manifold → coordinate transformation approximated well by homography.
- EM framework provides robustness to variations in quality of vehicular detections.



Agenda

- WAMI Introduction
- Roadmap Alignment Exploiting Vehicle Detections
  - Motivation, Challenges, Problem Statement
  - Prior Approaches and Challenges
  - Proposed Approach Exploiting Vehicle Detections
    - EM formulation
  - Results
- Benefits of EM Formulation

EM Formulation: Details

- Latent variables: Detection j onroad $z_j \in \{0, 1\}$
  - IID Bernoulli with unknown parameter $\gamma = p(z_j = 1)$
- d_j = Min square Euclidean distance of jth detection (Random Variable)
  - Conditioned on $z_j = 1$, exponential with unknown parameter $\lambda = 1/\mathbb{E}[d_j | z_j = 1]$
  - Conditioned on $z_j = 0$, uniform over image area
- Estimate unknown parameters $\boldsymbol{\theta} = \{\boldsymbol{\beta}, \lambda, \gamma\}$, by maximizing likelihood

$$p(\mathbf{d}|\boldsymbol{\theta}) = \prod_{j=1}^{N_v} p(d_j|\boldsymbol{\theta})$$

EM Formulation: Likelihood

- Likelihood

$$
\begin{aligned}
p(d_j|\boldsymbol{\theta}) = \sum_{z_j \in \{0,1\}} & p(d_j, z_j|\boldsymbol{\theta}) \\
= & p(z_j = 1)p(d_j|\boldsymbol{\theta}, z_j = 1) + \\
& p(z_j = 0)p(d_j|\boldsymbol{\theta}, z_j = 0) \\
= & \gamma\lambda e^{-\lambda d_j} + \frac{(1-\gamma)}{M^2},
\end{aligned}
$$

EM Algorithm: E step

- Posterior distribution of latent variables, for current estimate of parameters

$$p_j = p(z_j = 1 | d_j, \boldsymbol{\theta}^t)$$

$$\begin{aligned} p_j &= \frac{p(d_j | z_j = 1, \boldsymbol{\theta}^t) p(z_j = 1 | \boldsymbol{\theta}^t)}{p(d_j | \boldsymbol{\theta}^t)} \\ &= \frac{\gamma \lambda e^{-\lambda d_j}}{\gamma \lambda e^{-\lambda d_j} + \frac{(1-\gamma)}{M^2}}. \end{aligned}$$

EM Algorithm:M Step

- Maximize expectation of complete data log likelihood

$$p(\mathbf{z}, \mathbf{d}|\boldsymbol{\theta}) = \prod_{j=1}^{N_v} p(z_j, d_j|\boldsymbol{\theta})$$

$$p(z_j, d_j|\boldsymbol{\theta}) = \begin{cases} \gamma\lambda e^{-\lambda d_j}, & \text{if } z_j = 1 \\ \\ \frac{(1-\gamma)}{M^2}, & \text{if } z_j = 0 \end{cases}$$

$$\begin{aligned} \mathcal{Q}(\boldsymbol{\theta}, \boldsymbol{\theta}^t) &= \sum_{j=1}^{N_v} \sum_{z_j \in \{0,1\}} p(z_j|d_j, \boldsymbol{\theta}^t) \ \ln\left(p(z_j, d_j|\boldsymbol{\theta})\right) \\ &= \sum_{j=1}^{N_v} p_j \left[\ \ln(\gamma) + \ \ln(\lambda) - \lambda d_j\right] + \\ &\quad (1 - p_j) \left[\ \ln(1-\gamma) - 2\ \ln(M)\right] \end{aligned}$$

M Step (contd.)

- Maximizing parameter values

$$\gamma^* = \frac{\sum_{j=1}^{N_v} p_j}{N_v}$$

$$\lambda^* = \frac{\sum_{j=1}^{N_v} p_j}{\sum_{j=1}^{N_v} p_j d_j}$$

$$\boldsymbol{\beta}^*$$

$$f(\boldsymbol{\beta}) = \sum_{j=1}^{N_v} p_j d_j(\boldsymbol{\beta})$$

References

1. A. Elliethy and G. Sharma, “Automatic Registration of Wide Area Motion Imagery to Vector Road Maps by Exploiting Vehicle Detections,” IEEE Trans. Image Proc., accepted for publication August 2016.

2. A. Elliethy and G. Sharma, “Vector road map registration to oblique wide area motion imagery by exploiting vehicles movements,” in IS&T Electronic Imaging: Video Surveillance and Transportation Imaging Applications. IS&T, Springfield, VA, 2016, San Francisco, CA, 14-18 February 2016.

3. A. Elliethy and G. Sharma, “ A joint approach to vector road map registration and vehicle tracking for wide area motion imagery ,” in Proc. IEEE Intl. Conf. Acoustics Speech and Sig. Proc. , March 2016.

4. W. Song, J. Keller, T. Haithcoat, and C. Davis, “Automated geospatial conflation of vector road maps to high resolution imagery,” IEEE Trans. Image Proc., vol. 18, no. 2, pp. 388–400, Feb 2009.

5. C.-C. Chen, C. A. Knoblock, and C. Shahabi, “Automatically conflating road vector data with orthoimagery,” GeoInformatica, vol. 10, no. 4, pp. 495–530, 2006.

6. Lowe, David G. "Distinctive image features from scale-invariant keypoints." International journal of computer vision 60.2 (2004): 91-110.

7. E. Blasch, G. Seetharaman, S. Suddarth, K. Palaniappan, G. Chen, H. Ling, and A. Basharat, “Summary of methods in wide-area motion imagery (WAMI),” in Proc. SPIE, vol. 9089, 2014, pp. 90 890C–90 890C–10.