Motion Based Vehicle Identification in Car Video

Amirali Jazayeri, Hongyuan Cai,
Jiang Yu Zheng, Mihran Tuceryan

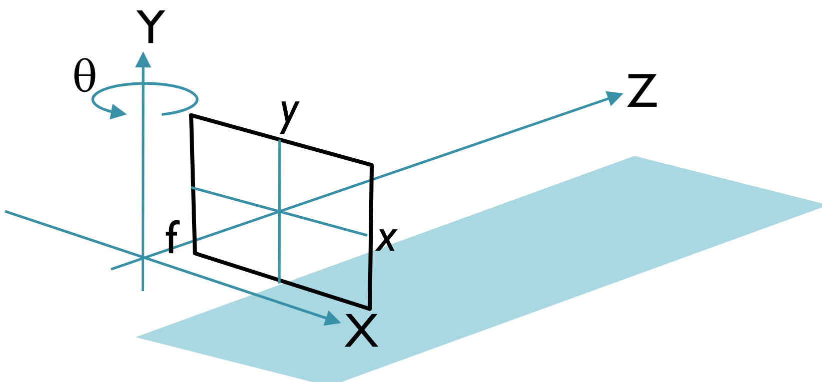
Dept. of Computer Science
Indiana University Purdue University Indianapolis

Outline

- Our goal is to track and identify moving vehicles ahead against static background in car video
- Under ego-motion, dynamic vehicles and background display different motion behaviors
- Profile features in car video and track motion for fast and robust processing
- Describe motion probability regarding image position and image velocity
- Use HMM to identify vehicles from their continuous movement

In-car Video: Assumptions

- Ego-motion (speed and steering) is controllable or readable from a vehicle
- All vehicles run on road, and are mostly moving in the same direction as the camera
- Horizontal image velocity change and vertical scaling happen in both background area and target vehicles
- We separate moving vehicles and static background including stopped vehicles



Challenges

- Variations of vehicles in shape, size, color, type, etc.
- Illumination changes in outdoor environments (day and night, shadow and highlight, etc.)
- Unpredictable occlusion between vehicles
- Cluttered background or no prior knowledge on changing background
- Real-time Performance

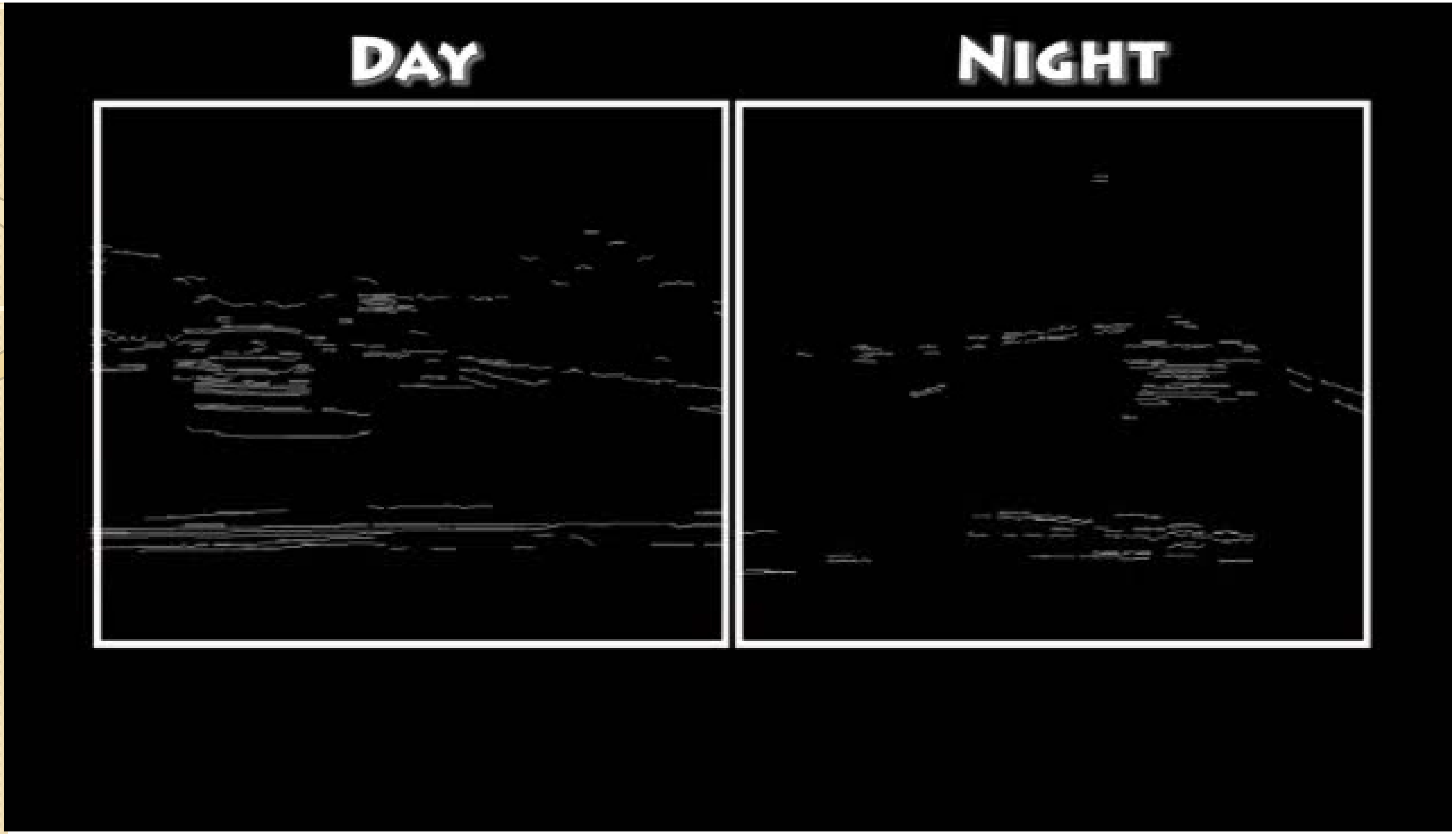
Related Works on Vehicle Detection Based on Shapes and Models

Employ priori knowledge to hypothesize vehicle locations in an image

- Intensity
  - Learning the characteristics of vehicle classes from a set of training images
- Shape cues
  - Symmetry, color, shadow, corners, vertical/horizontal edges, vehicle lights, and texture
- Model
  - Vehicle templates or models with varying degrees of deformability
- Optical flow or motion traces
  - Tracking vehicle

Our Approach Using Image Motion

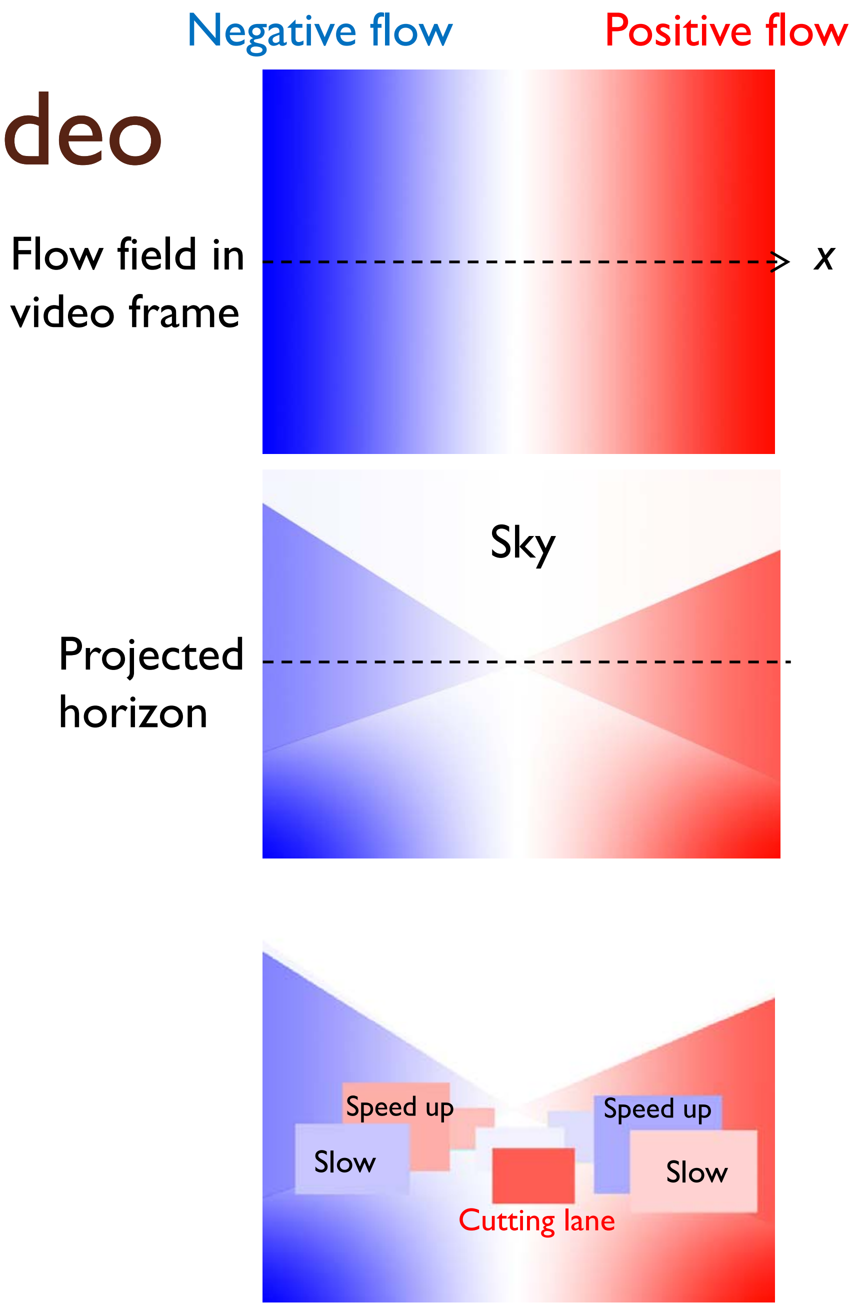
- Ego-motion (forward translation + horizontal rotation) is general for all types of observer vehicle
- Ego-motion generated background motion is determinant in direction and scale
- The vehicle motion against background is invariant to the shape, color, and size of target vehicles



Motion alone tells something

Motion Model in Video

- Ego-motion
  - Dividing flow into positive and negative fields in the video frame
  - Smaller flow at center than at two sides
- Background Motion
  - Distant scenes have lower velocity than close scenes
  - Sky has zero flow
- Vehicle Motion
  - Speed up and slow down
  - Change lane

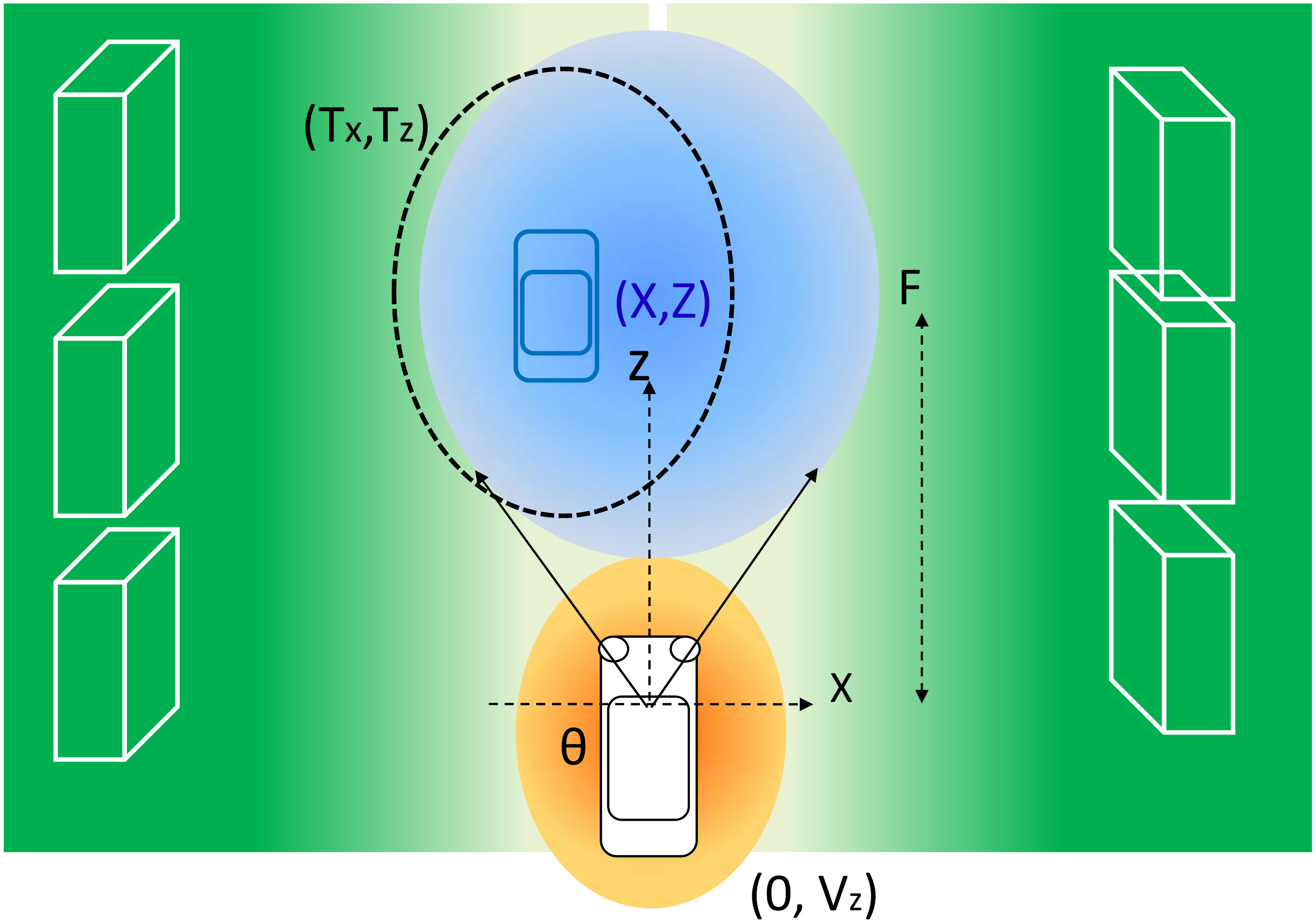


Method

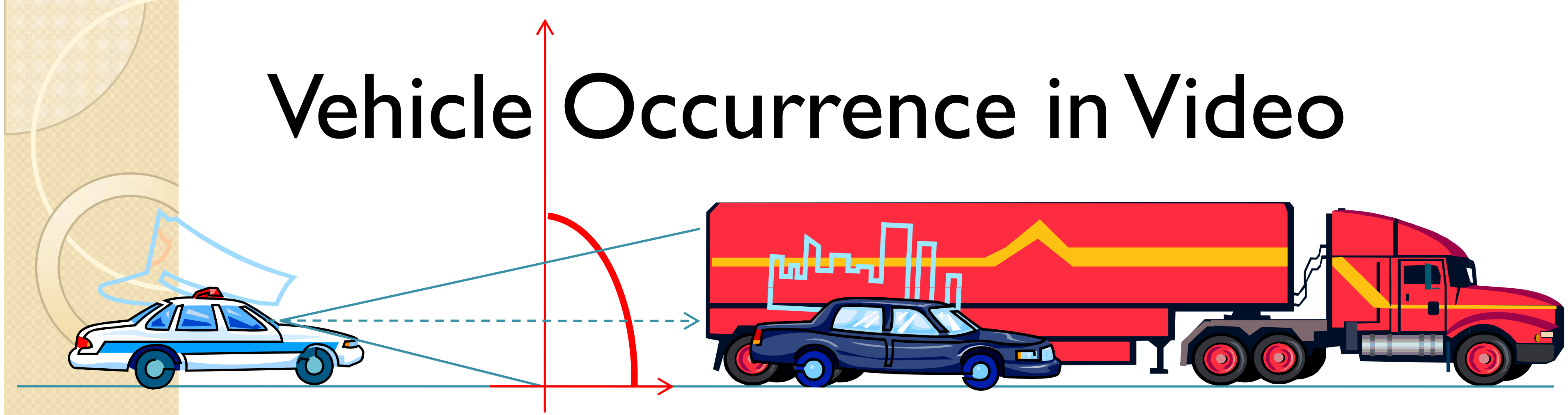
- Using horizontal image velocity only
  - Profiling video vertically and tracking the condensed image
  - This yields variables: (x,v)

    (Image position, Horizontal image velocity)
- Computing likelihood probability distribution P(x,v) for vehicles and background
- Detection using continuous motion behavior
  - Hidden Markov Model (HMM)
  - Results in identity with probability description

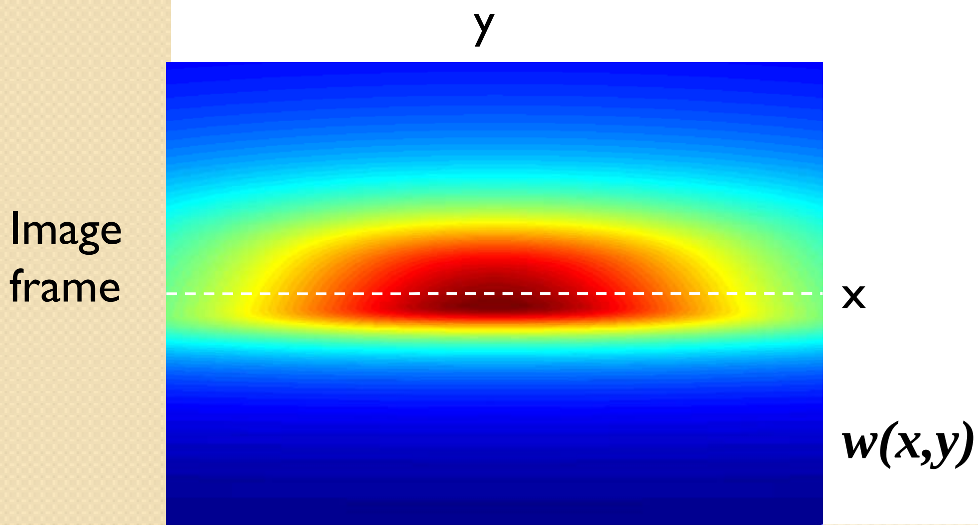
Background and Vehicle Distributions on Road



Vehicle Occurrence in Video



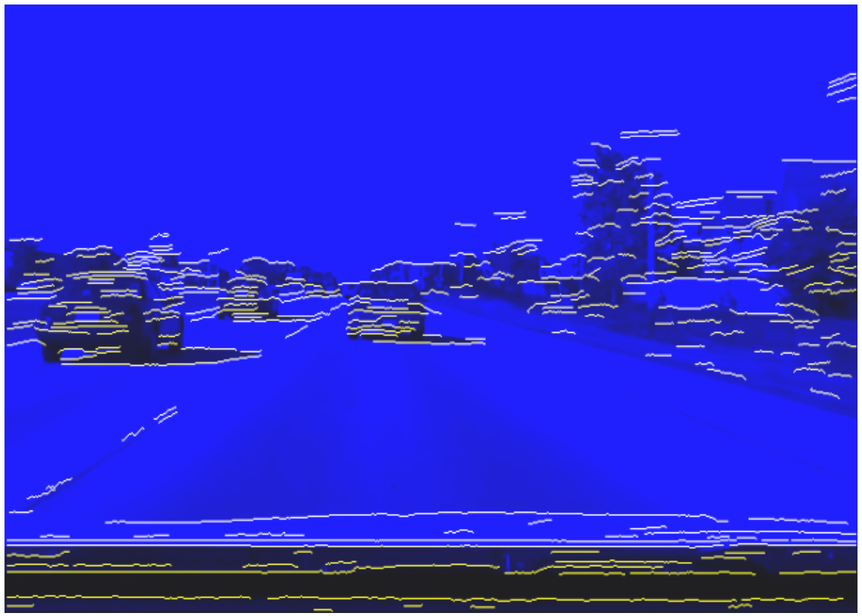
- Vehicles appear most frequently in the image across the projected horizon
- Taking1D profile from image using a mask to speed up tracking



$$T(x) = \sum_{y=-h/2}^{h/2} w(x,y) I(x,y)$$

Vehicle Feature Extraction

- Vehicles typically contain many horizontal edges formed by top and bottom boundaries of a car, license plate, and window edges.
- Other features include corners, intensity peaks, etc.



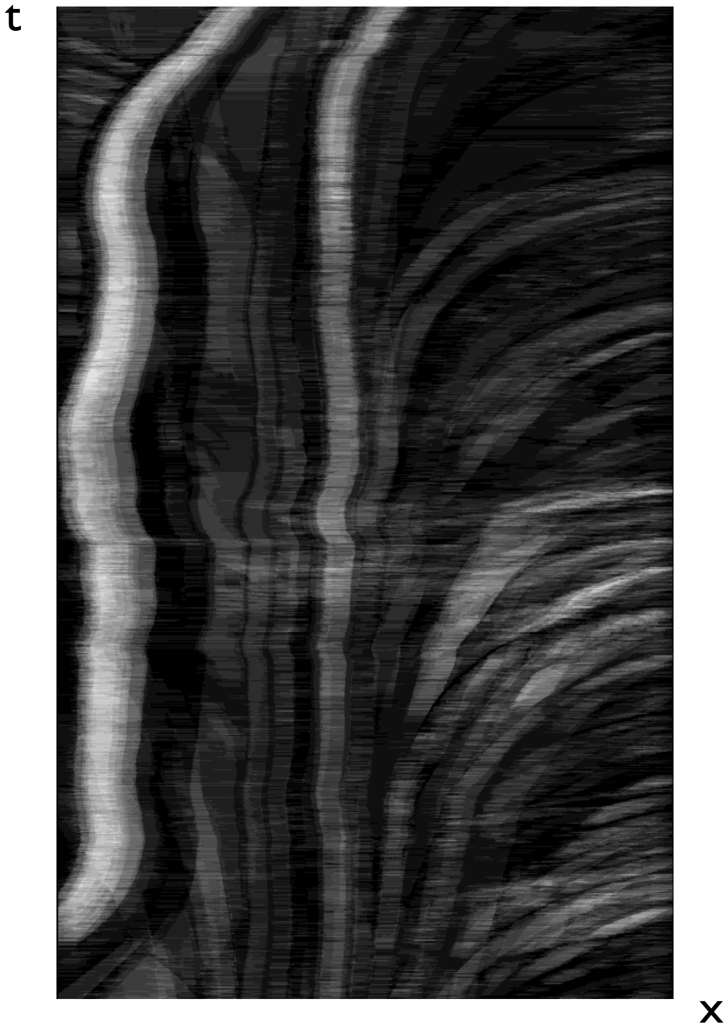
1D Profiles form a Condensed Image

- Vertically profile features extracted in the video frames to generate traces of vehicles in the spatial-temporal image

$$T(x,t) = \sum_{y=-h/2}^{h/2} w(x,y) I(x,y,t)$$

- The condensed image containing feature traces for detecting vehicles
- Reducing the influence of vehicle shaking in pitch and roll

Examples of Condensed Images

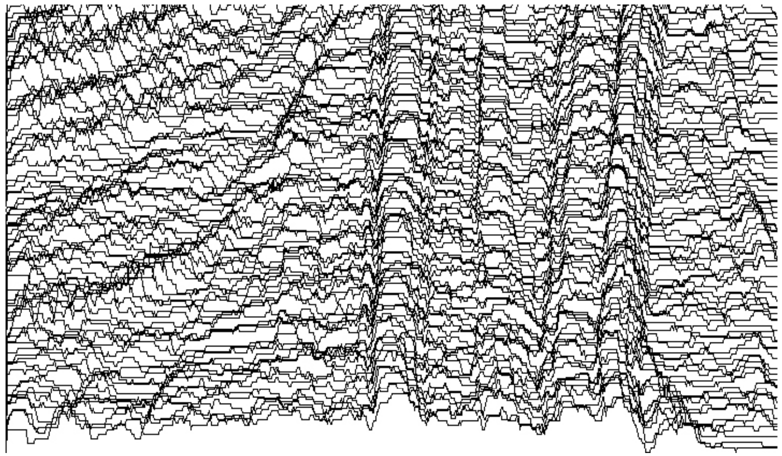


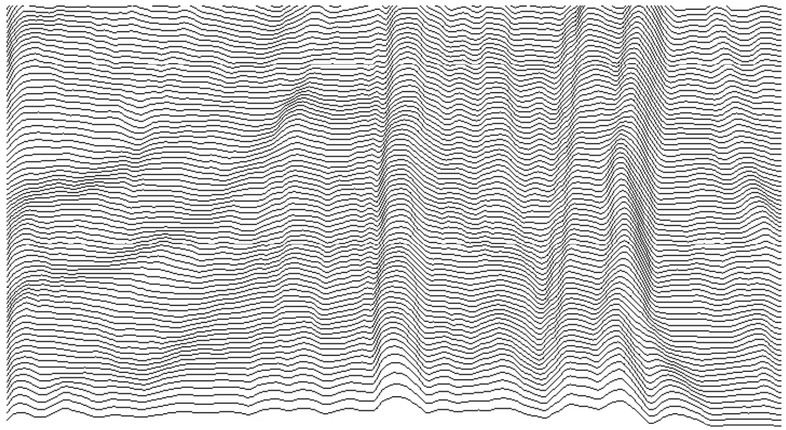
Horizontal Line Profile

Intensity Peak Profile

1-D Profiles Processing

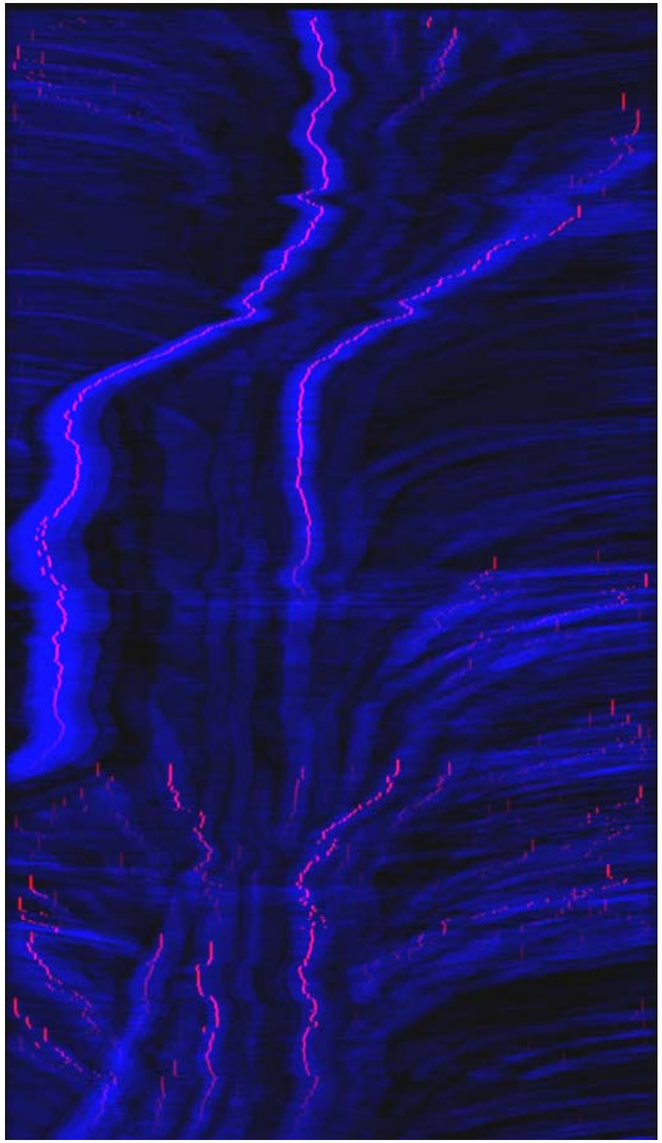
- Due to the presence of horizontal lines in the background scene, the original 1D profile is very noisy
- Pyramid scaling operations are preformed on the profiles to eliminate the noise and extract major feature traces



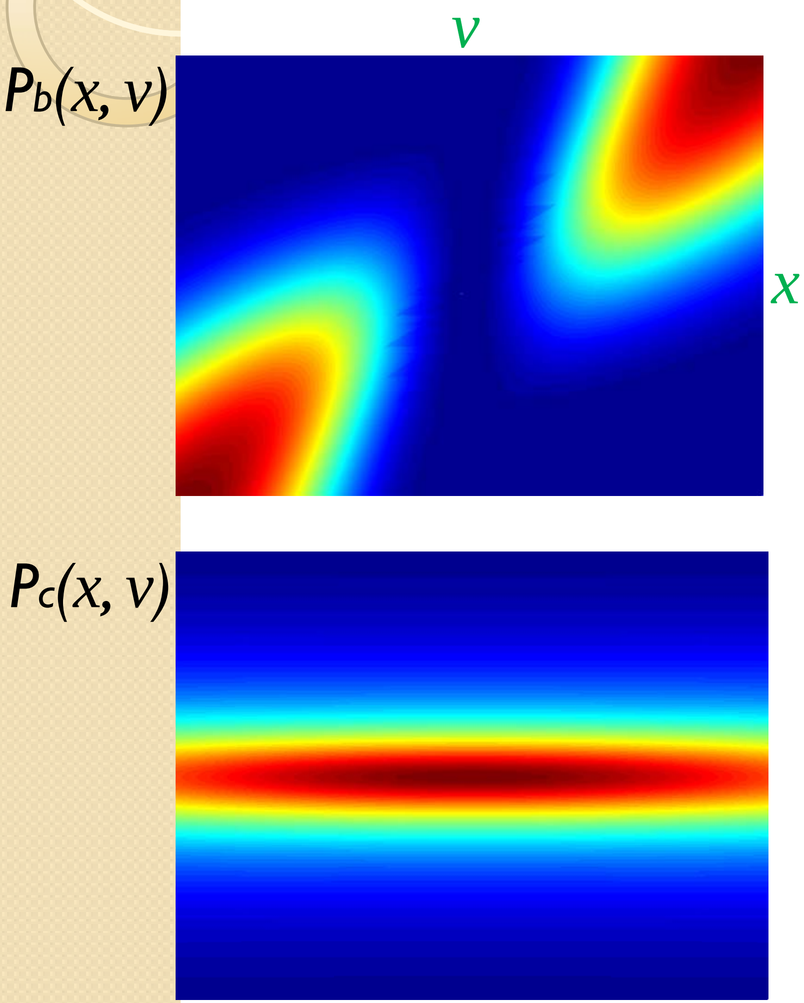


Tracking Traces in Condensed Image

- Tracking traces in the condensed image over time determines (*x*,*v*) sequences
- For unstable traces from line segments, we track the center part of each trace
- Track traces in the condensed image using Kalman filter during the vehicle motion
- Motion continuity is applied in tracking to avoid noise from instantaneous light changes



Likelihood Probability Distribution of Motion in Video

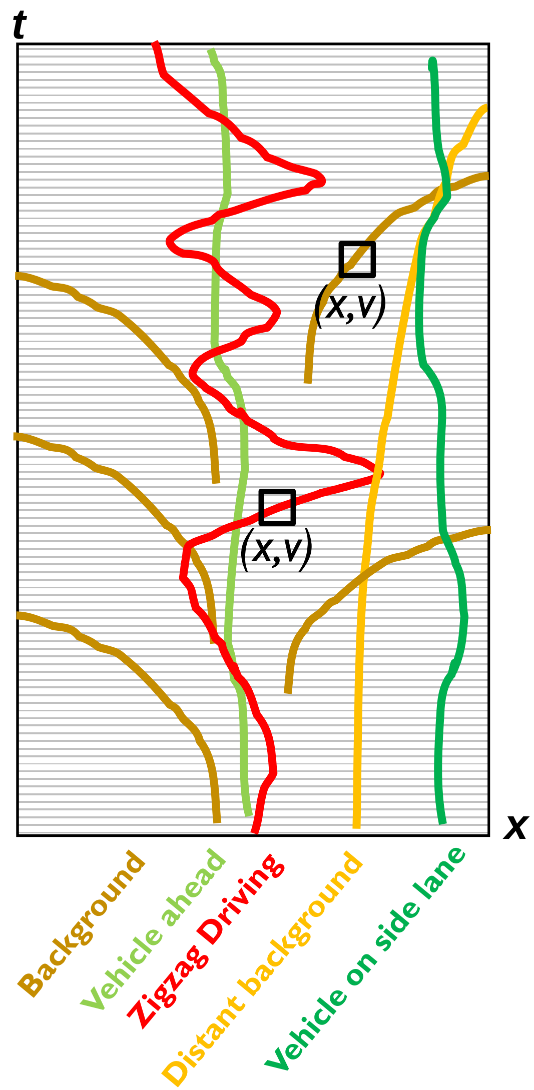


Red = high probability; Blue = low probability

- Joint probability $P_b(x,v)$ for background
  - Including minor steering N(0,5°)
  - From uncertain positions at distance towards certain positions on left and right sides of observer's vehicle
- Joint probability $P_c(x,v)$ for vehicle features
  - Relatively low image velocity when vehicles are confined in the road space

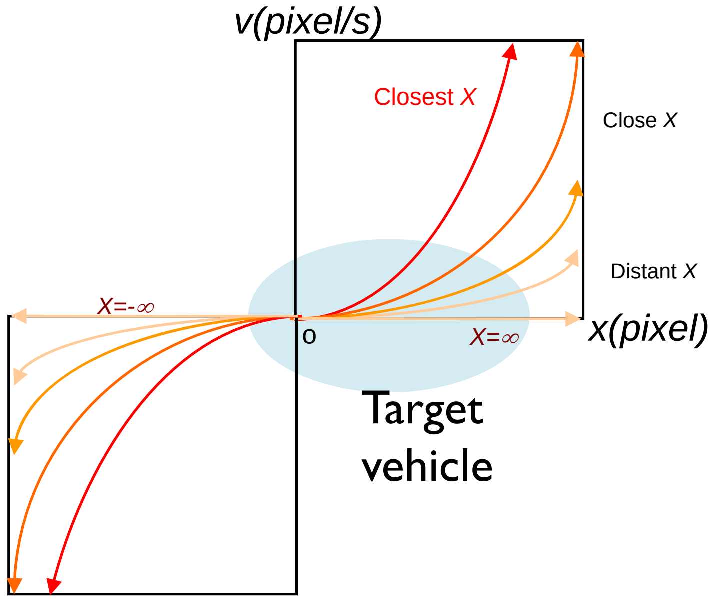
Motion Behaviors along Traces

- Background objects pursue hyperbolic trajectories expanding from the Focus of Expansion.
- The curvature of a background trace is high if it is closer to road and is low if it is further from road.
- Their image velocity is higher on scenes passing by, and is lower at scenes down the street.
- On the other hand, vehicles tracked within the road may stay in the image frame even they drive irregularly in a zigzag way



Identifying Traces during Tracking

- Using a probabilistic formulation to model the motion to avoid using sensitive thresholds in classification
- Hidden Markov Model (HMM) is used to model the continuous process of vehicles and background motion
- Two hidden states describe every trace at any time *t*:
  - State C_t as vehicle
  - State B_t as background
- The observation is a sequence of (*x*(*t*),*v*(*t*)) obtained from tracking trace
  - Image position *x(t)*
  - Horizontal image velocity *v(t)*



$$v(t) = \frac{fXV}{Z^2(t)} = \frac{Vx^2(t)}{fX}$$

Estimate Status of a Trace

- Posterior probabilities $P(C_t \mid x(t),v(t))$ and $P(B_t \mid x(t),v(t))$ are updated by

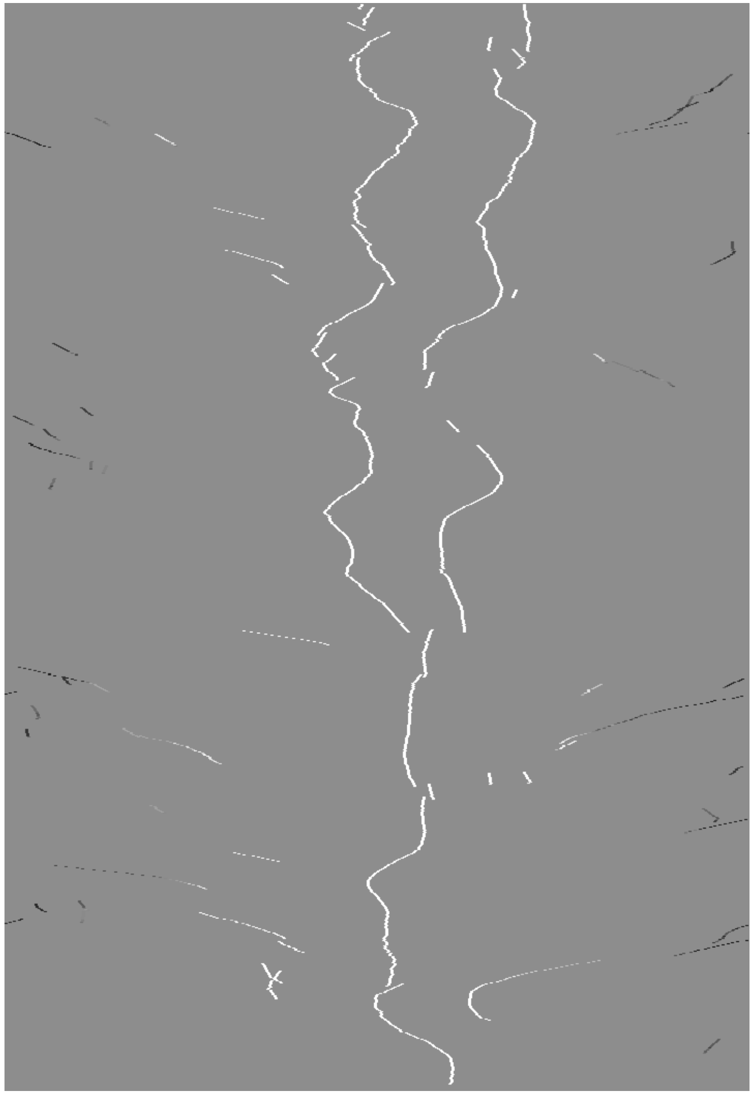
$$P(C_t) = \max[\, P(B_{t-1})P(C_t \mid B_{t-1})p(x(t),v(t)|C_t),$$
$$P(C_{t-1})P(C_t \mid C_{t-1})p(x(t),v(t) \mid C_t)]$$
$$P(B_t) = \max[\, P(B_{t-1})P(B_t \mid B_{t-1})p(x(t),v(t)|B_t),$$
$$P(C_{t-1})P(B_t \mid C_{t-1})p(x(t),v(t)|B_t)]$$

  using Viterbi algorithm
- If $P(C_t) > P(B_t)$, the trace is considered as a car at time t, or as background otherwise
- At any time t, $P(C_t)+P(B_t)=1$, for normalization

$$P(C_t) \leftarrow \frac{P(C_t)}{P(C_t)+P(B_t)} \qquad P(B_t) \leftarrow \frac{P(B_t)}{P(C_t)+P(B_t)}$$

- The identified trace is formally output after it is tracked over a certain duration. Otherwise, such a short trace is removed as noise

Vehicle Identification

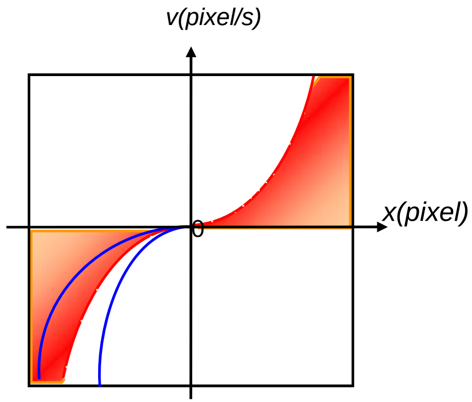


Vehicle Detection Results

Vehicle Detection Results

- The longer the tracked duration, the more certain the identification becomes.
- If a detected vehicle moves too far from the observer car, it will be ignored
- Approaching vehicles on the opposite lane are classified as background
- Turning at a street corner needs another likelihood distribution, but not dealt with here

True-Positive	**86.9**
False-Negative	14.1
True-Negative	85.9
False-Positive	13.2



Opposite lane vehicle

Conclusion

- Detected features and tracked their profiled trajectories in spatial-temporal condensed image
- Introduced a probability model of background and vehicles and computed the likelihood probability distribution of their motions
- Used HMM to estimate the process of location dependent motion for vehicle identification

Questions?

Background PDF

$p(X)$~1-exp(-$X^2/2D^2$)

(x,v) ← V, (X, Z), Ry

N(S,σ)

N(0, 5°)

$$x = \frac{fX}{Z}$$

$$v = \frac{x^2 V}{fX} - \frac{x^2 + f^2}{f} R_y$$

$$p\left(x, v \mid B\right) = \int_{R_y} p(R_y) p(x, v \mid R_y) dR_y \quad = \int_{R_y} p(R_y) p(x, v \quad \left| v = \frac{x^2 V}{fX} - \frac{x^2 + f^2}{f} R_y \right.) dR_y$$

$$= \int_{R_y} \int_{X} p(R_y) p^{(d)}(X) p\left(X\right) p\left(Z, V \quad \mid X\right) dX dR_y \qquad \left(Z(x, v, X), V(x, v, X)\right)$$

$$= \int_{R_y} \int_{X} p(R_y) p^{(d)}(X) p\left(X\right) p\left(Z = \frac{fX}{x} \quad \middle| X\right) \times p\left(V = \left(v + \frac{x^2 + f^2}{f} R_y\right) \frac{fX}{x^2} \quad \middle| X\right) dX dR_y$$

$$= C_{1r} \int_{R_y} \int_{X} e^{\frac{-R_y^2}{2\sigma_r^2}} \frac{1 - e^{\frac{-X^2}{2D^2}}}{|X| + 1} e^{\frac{-\left(\left(v + \frac{x^2 + f^2}{f} R_y\right) \frac{fX}{x^2} - S\right)^2}{2\sigma^2}} \left|\frac{fX}{x^2}\right|^2 dX dR_y$$

Vehicle PDF

N(0, 5°)

(x,v) ← (X,Z), (Tx,Tz), Ry

(X, Z)~N((0,F,) (D, 2F))

(Tx,Tz)~N((0, 0), σ_x, σ_z)

$$v(t) = \frac{fT_x(t) - x(t)T_z(t)}{Z(t)} - \frac{x^2(t) + f^2}{f} R_y(t) = v^{(t)}(t) + v^{(r)}(t)$$

$$p(x,v \mid C) = p(x,v \mid (X,Z,T_x,T_z) \in C) = p\left(x = \frac{fX}{Z}, v = \frac{fT_x - xT_z}{Z} \right) = \int_Z p(Z) p\left(x = \frac{fX}{Z}, v = \frac{fT_x - xT_z}{Z} \mid Z \right) dZ$$

$$= \int_Z p(Z) p\left(X = \frac{xZ}{f}, T_z, T_x = \frac{Zv + xT_z}{f} \mid Z \right) \left| \frac{Z}{f} \right|^2 dZ = \int_Z p(Z) p\left(X = \frac{xZ}{f} \mid Z \right) p\left(T_z, T_x = \frac{Zv + xT_z}{f} \mid Z \right) \left| \frac{Z}{f} \right|^2 dZ$$

$$= \int_Z p(Z) p\left(X = \frac{xZ}{f} \mid Z \right) \left\{ \int_{T_z} p(T_z) p\left(T_x = \frac{Zv + xT_z}{f} \mid T_z, Z \right) dT_z \right\} \left| \frac{Z}{f} \right|^2 dZ$$

$$p(x,v \mid C) = C_2 \int_Z \int_{T_z} e^{\frac{-(Z-F)^2}{2(2F)^2}} e^{\frac{-\left(\frac{xZ}{f}\right)^2}{2D^2}} e^{\frac{-T_z^2}{2\sigma_z^2}} e^{\frac{-\left(\frac{Zv + xT_z}{f}\right)^2}{2\sigma_x^2}} \left| \frac{Z}{f} \right|^2 dT_z dZ$$

Parameter Selection in Probability Computation

D	**Average road width As wide as three lanes**	**6m**
F	**Distance to target Minimum safe distance**	**10m**
σ_F	**Standard deviation of target distance**	**20m**
σ_x	**Standard deviation of relative horizontal speed T_x of target vehicle** **Maximum cutting of three lanes, tolerant for moving on curved path**	**6m/s**
σ_z	**Standard deviation of relative translation speed T_z, T_z is zero if target is pursued**	**10m/s**
S	**Average pursuing speed of observer vehicle**	**50km/h 15m/s**
σ	**Standard deviation of the speed**	**10km/h 5m/s**
f	**Camera focal length Through offline calibration**	**900 pixel**
$\int_z$	**Range for integration From camera position to distance close to infinity**	**0~200m**
$\int_x$	**Range for integration Wider than a road to include all backgrounds in video**	**-50~50m**
$\int_{Tz}$	**Range for relative speed**	**-40~40m/s**
$\int_{Ry}$	**Range of integration**	**-10~10 degree/s**
H	**The maximum height of vehicle, As high as a truck, but mostly for cars**	**4m**
σ_r	**Standard deviation of steering angle of R_y** **From the maximum tuning radius of a vehicle and road curvature.**	**5 degree/s**