Trajectography of an Uncalibrated Stereo Rig in Urban Environments

Nicolas SIMOND, Patrick RIVES

Firstname.Lastname@sophia.inria.fr

ICARE Project Team

Institut National de Recherche en Informatique et Automatique

Sophia-Antipolis, France

Motivation

The use of GPS for precise localization (<0.5m) currently fails in the urban environments.

- Visibility may be blocked between tall building (urban canyons) and depends on current satellites configuration;
- Estimation process may be corrupted by multiple path effects (reflection and diffraction).

Motivation

The use of GPS for precise localization (<0.5m) currently fails in the urban environments.

- Visibility may be blocked between tall building (urban canyons) and depends on current satellites configuration;
- Estimation process may be corrupted by multiple path effects (reflection and diffraction).

$\Rightarrow$ GPS must be hybrided by a dead-reckoning method.

A dead-reckoning method based on vision

A dead-reckoning method based on vision

- The vanishing lines and points result from the projection of sets of parallel lines in the 3D scene;
- The traffic marking belongs to the road plane.

[Simond,N. and Rives,P. - IROS’2004]

A dead-reckoning method based on vision



- The vanishing lines and points result from the projection of sets of parallel lines in the 3D scene;
- The traffic marking belongs to the road plane.

⇒ Segmentation of main planes in the scene (facades, road..)

A dead-reckoning method based on vision

- The vanishing lines and points result from the projection of sets of parallel lines in the 3D scene;
- The traffic marking belongs to the road plane.

⇒ Segmentation of main planes in the scene (facades, road..)

⇒ Estimation of the roadway geometry

A dead-reckoning method based on vision

- The vanishing lines and points result from the projection of sets of parallel lines in the 3D scene;
- The traffic marking belongs to the road plane.

⇒ Segmentation of main planes in the scene (facades, road..)

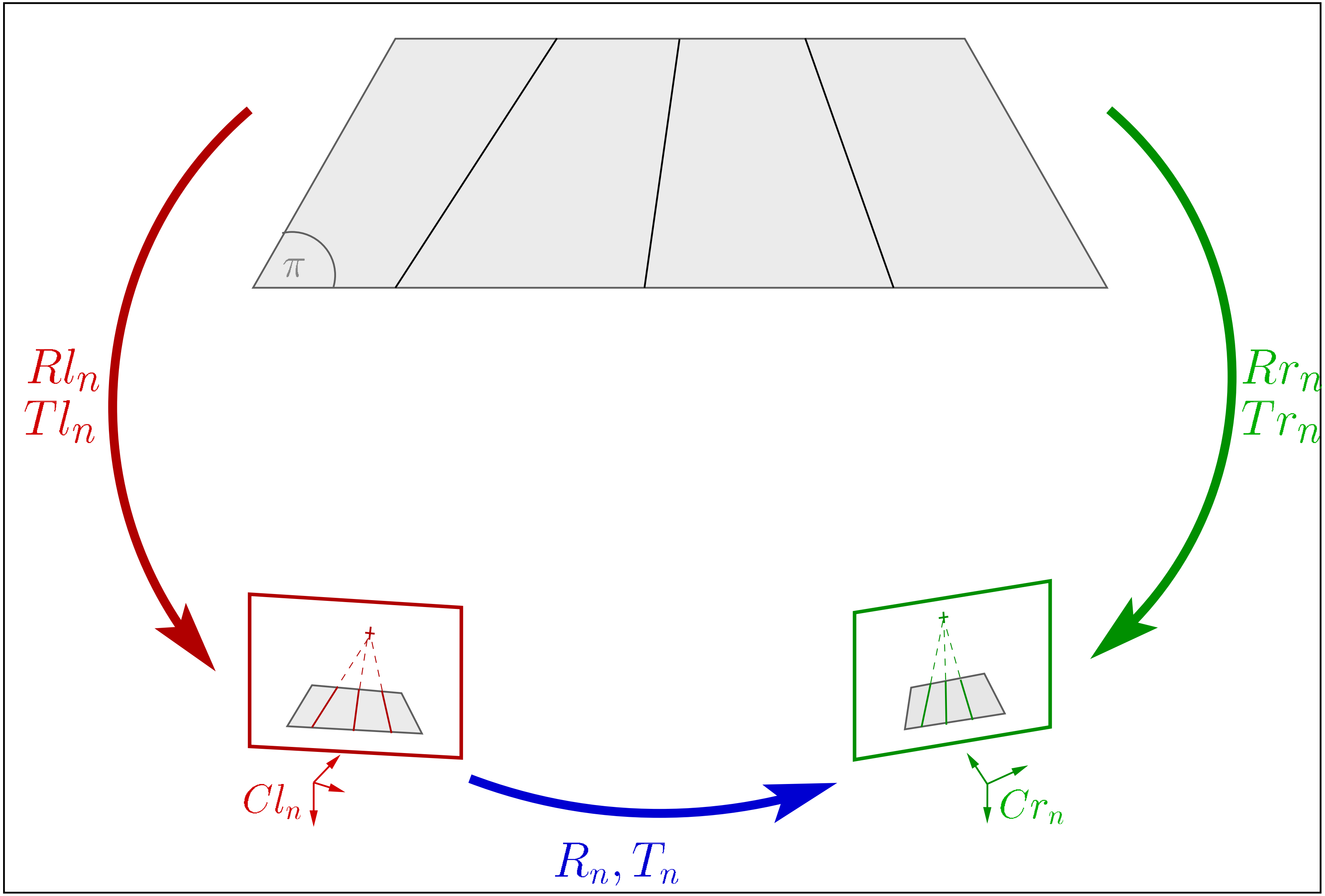
⇒ Estimation of the roadway geometry

⇒ Estimation of the vehicle egomotion

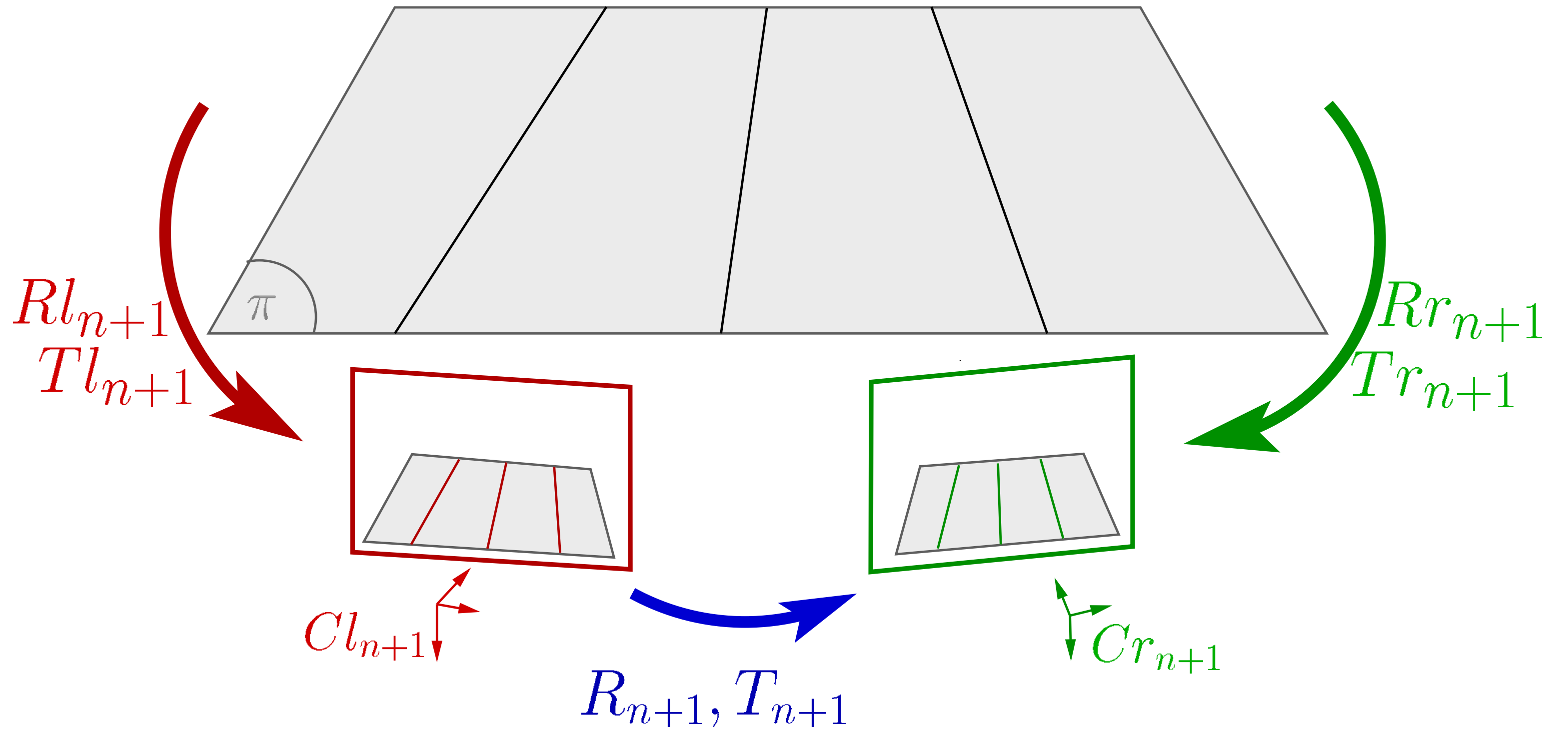
Assumptions

- the road is locally planar with parallel boundaries;
- we use a rigid stereo rig with weakly calibrated cameras;
- the sequence is recorded at video rate.

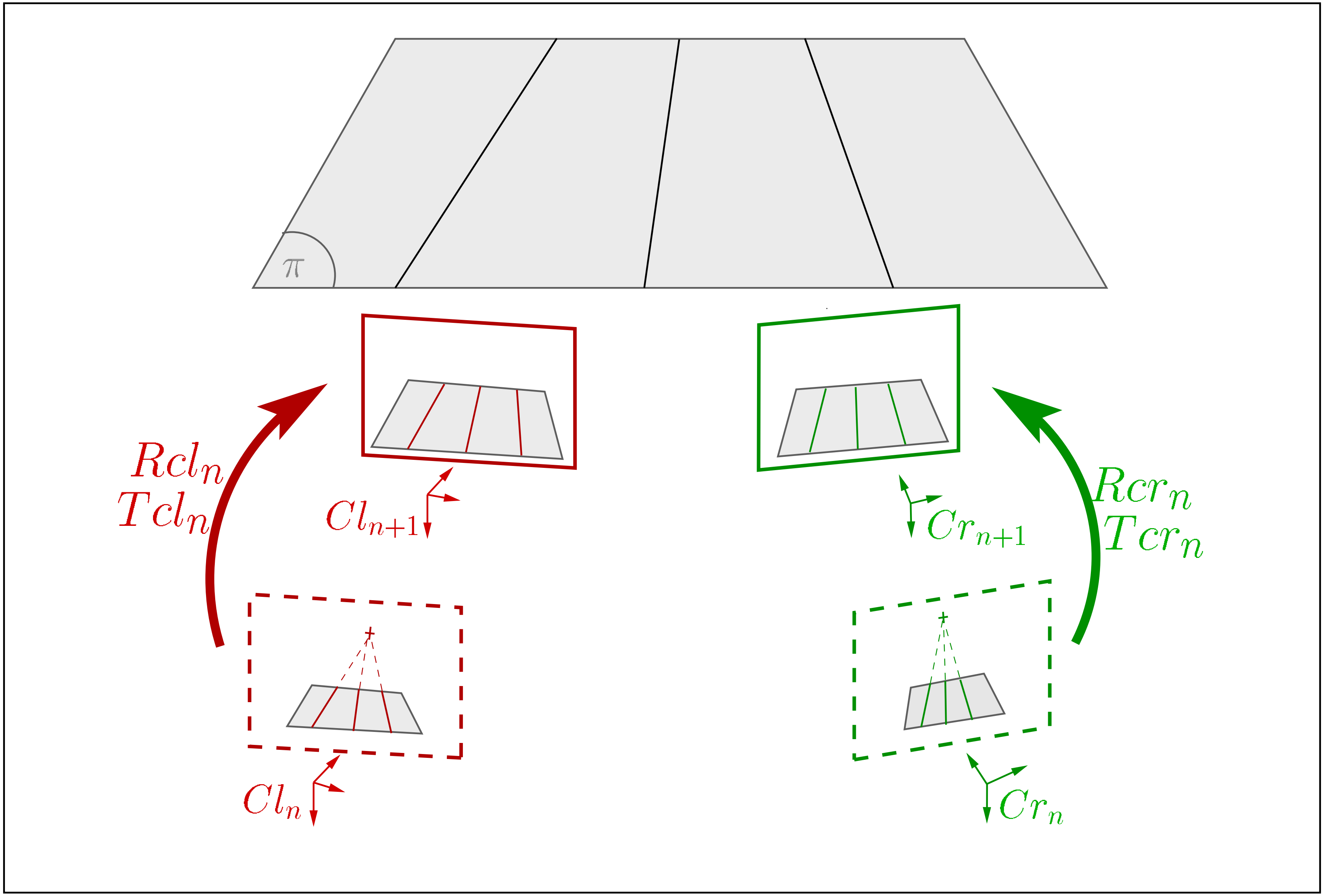
Geometrical modeling



Geometrical modeling



Geometrical modeling



Coplanar Features

Features lying on the π plane have projections in 2 images constrained by the **planar homography** H_π:

for $k \in [1, 2]$,

$$\mathbf{l_k} \propto \mathsf{H}_\mathsf{k}^{-\mathsf{t}}.\mathbf{L}$$

$$\mathbf{p_k} \propto \mathsf{H}_\mathsf{k}.\mathbf{P}$$

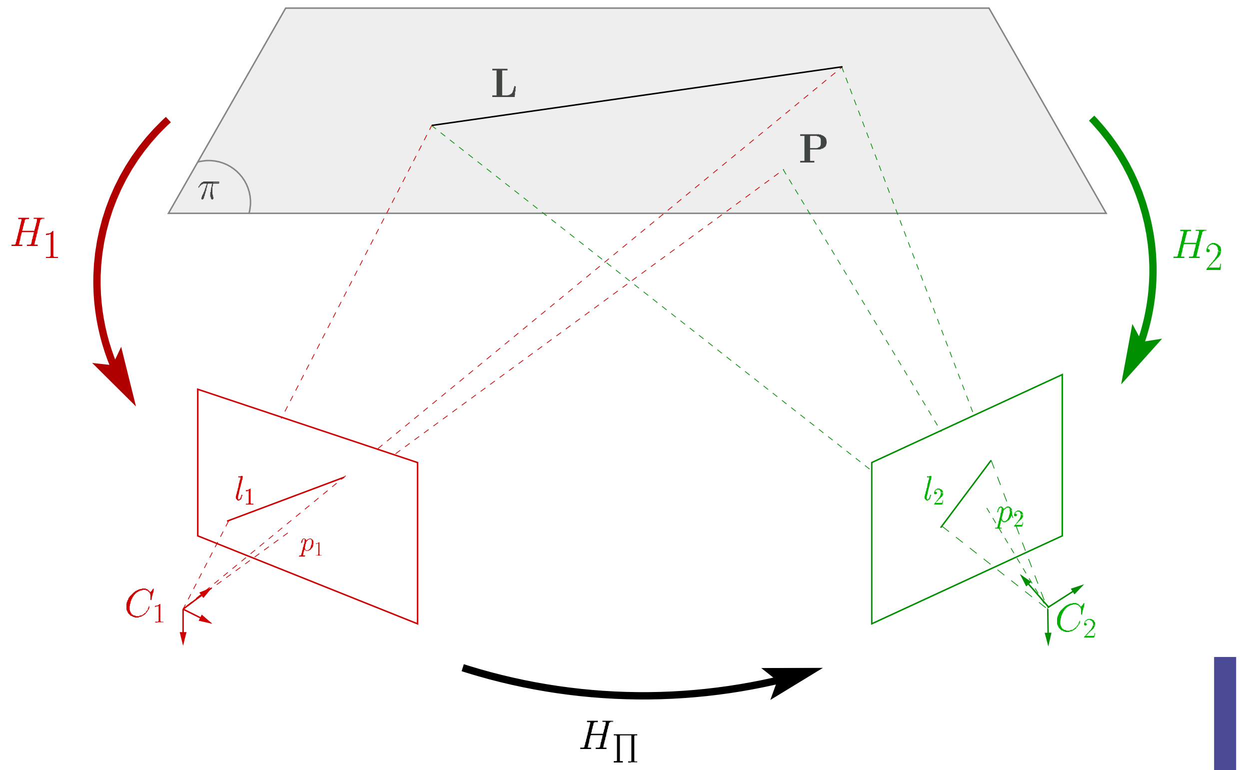
then

$$\mathbf{l_2} \propto \mathsf{H}_\pi^{-\mathsf{t}}.\mathbf{l_1}$$

$$\mathbf{p_2} \propto \mathsf{H}_\pi.\mathbf{p_1}$$

with

$$\mathsf{H}_\pi = \mathsf{H}_2.\mathsf{H}_1^{-1}$$



Algorithm overview

The four-stage algorithm supplies:

1. the segmentation of the roadway in the left and right images;
2. the extraction of coplanar features;
3. the tracking of coplanar features along the sequence;
4. the stereo rig motion.

Algorithm overview

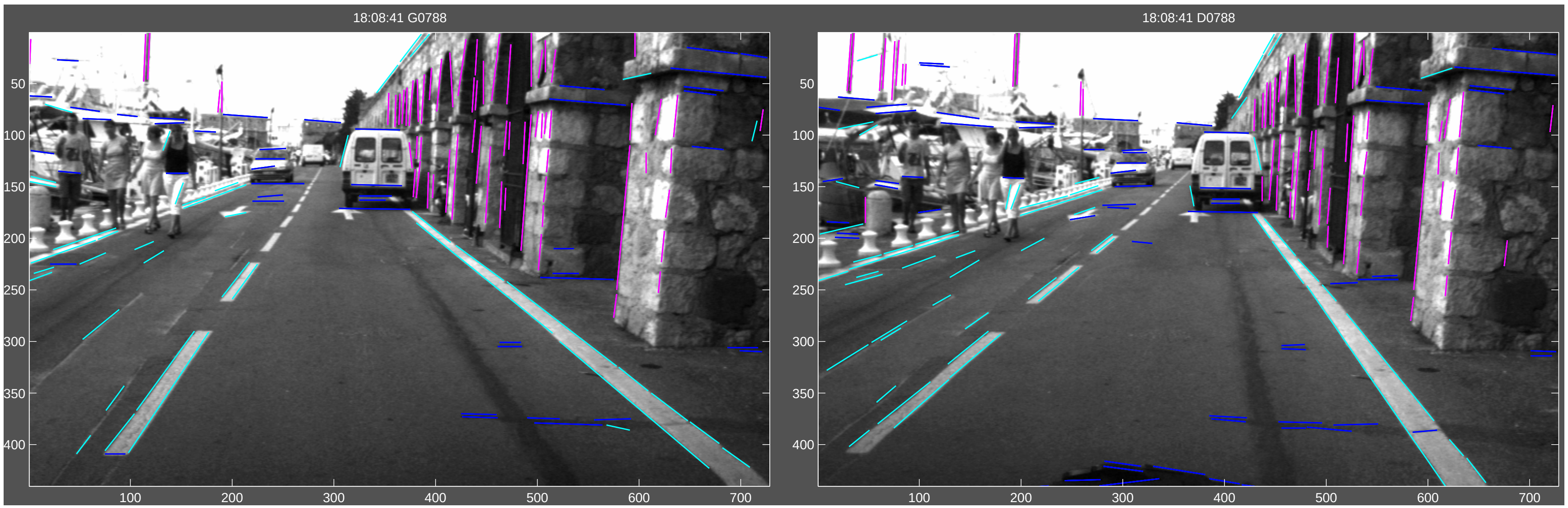
The four-stage algorithm supplies:

1. the segmentation of the roadway in the left and right images:
   - the Vanishing Lines (VLs);
   - the Dominant Vanishing Point (DVP).
2. the extraction of coplanar features;
3. the tracking of coplanar features along the sequence;
4. the stereo rig motion.

Segmentation of the roadway

For the left and the right images :

- a Canny edge detector supplies edges;
- the edges are classified according to the three main directions;
- the DVP corresponding to the roadway is estimated and tracked along the sequence with a Kalman filter;
- the road borders are extracted.



[Simond,N. and Rives,P. - IROS'2004]

Algorithm overview

The four-stage algorithm supplies:

1. the segmentation of the roadway in the left and right images;

2. the extraction of coplanar features:
   - the Points of Interest (PIs);
   - the epipolar constraint;
   - the stationarity of the homography H_{st} between left and right views along the sequence.

3. the tracking of coplanar features along the sequence;

4. the stereo rig motion.

Extraction of coplanar features

- the PIs are detected with a Harris detector in the ROI corresponding to the road;
- the candidates to the matching process satisfy the predicted homography: $\hat{H}^k_{st} = H^{k-1}_{st}$;
- the estimation of H^k_{st} is performed from the set of matches;
- the matches which do not satisfy the homography are rejected.

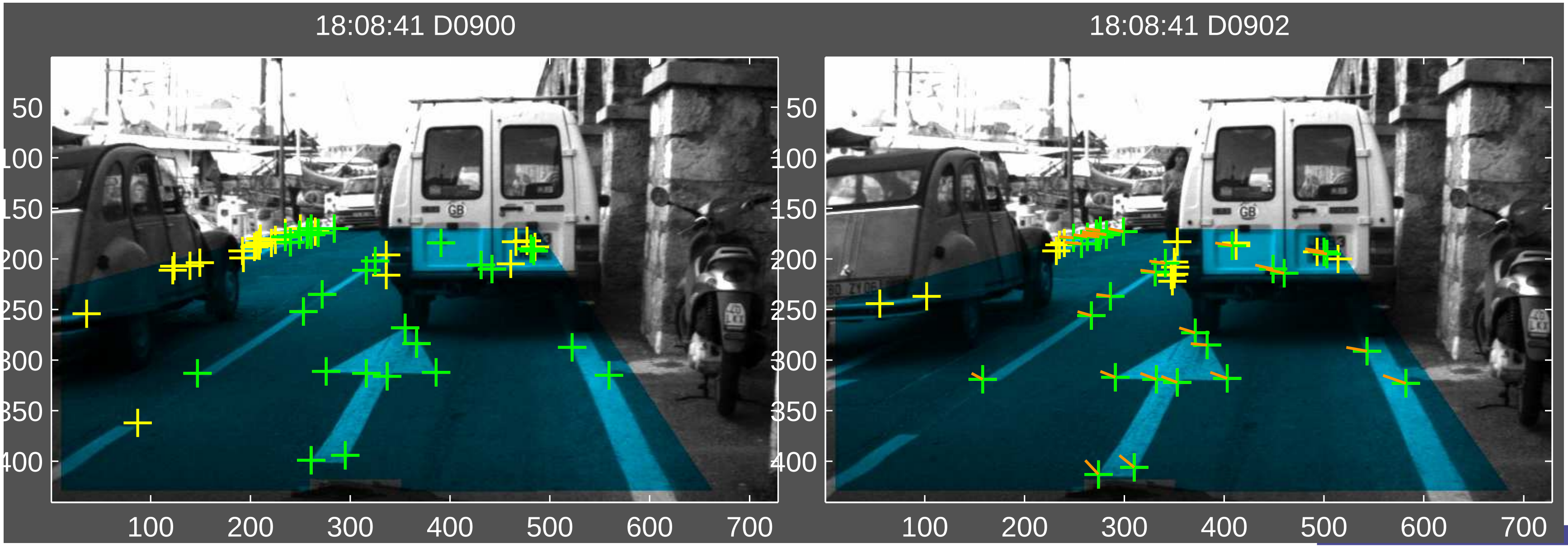
Algorithm overview

The four-stage algorithm supplies:

1. the segmentation of the roadway in the left and right images;
2. the extraction of coplanar features;
3. the tracking of coplanar features along the sequence :
   - assuming a small displacement between two successive images (video rate);
   - the right matched coplanar features verify the homographies H_l^k and H_r^k in the left and right sequence.
4. the stereo rig motion.

Tracking of coplanar features along the sequence

- a coarse prediction of the PIs $\hat{P}_i^{k-1} = P_i^k$;
- the research area for matching the features increases with their v-ordinates;
- the homography computations H_l^k and H_r^k use the same method than H_{st}^k;
- the matched PIs over the ground (obstacles) are rejected.



Algorithm overview

The four-stage algorithm supplies:

1. the segmentation of the roadway in the left and right images;
2. the extraction of coplanar features;
3. the tracking of coplanar features along the sequence;
4. the stereo rig motion:
   - Multi-view constraints between the set of PIs;
   - Smoothing and outliers rejection using the Super-Homography computation.

Super Homography (1)

The concept of Super-Homography [Malis and Cipolla*]:

- the Super-Homography SH generalizes the constraint $\mathsf{H_{ac}} \propto \mathsf{H_{ab}}.\mathsf{H_{bc}}$ to m views:

$$rank(\mathsf{SH^k}) = 3 \quad \forall m \geq 3 \quad \Longleftrightarrow \quad \mathsf{SH^{k}}^2 = m.\mathsf{SH^k}$$

- the Super-Points of Interest (SPIs) $\mathbf{P^k} = [\mathbf{P^k_1}, \cdots, \mathbf{P^k_n}]$ represent the coordinates of each FP in the m views;
- due to the measurement noise, $^0\mathsf{SH^k}$ do **not** verify $rank(^0\mathsf{SH^k}) = 3$.

* *“Multi-view constraints between collineations: application to self-calibration from unknown planar structures”* (ECCV’00)

Super Homography (2)

images	I_l^{k-2}	I_r^{k-2}	I_l^{k-1}	I_r^{k-1}	I_l^{k}	I_r^{k}
I_l^{k-2}	I_3					
I_r^{k-2}	H_{st}^{k-2}	I_3				
I_l^{k-1}	H_{l}^{k-1}	$\mathsf{H}_{l}^{k-1}.\mathsf{H}_{st}^{k-2}$	I_3			
I_r^{k-1}	$\cdots$	H_{r}^{k-1}	H_{st}^{k-1}	I_3		
I_l^{k}	$\cdots$	$\cdots$	H_{l}^{k}	$\mathsf{H}_{l}^{k}.\mathsf{H}_{st}^{k-1}$	I_3	
I_r^{k}	$\cdots$	$\cdots$	$\cdots$	H_{r}^{k}	H_{st}^{k}	I_3

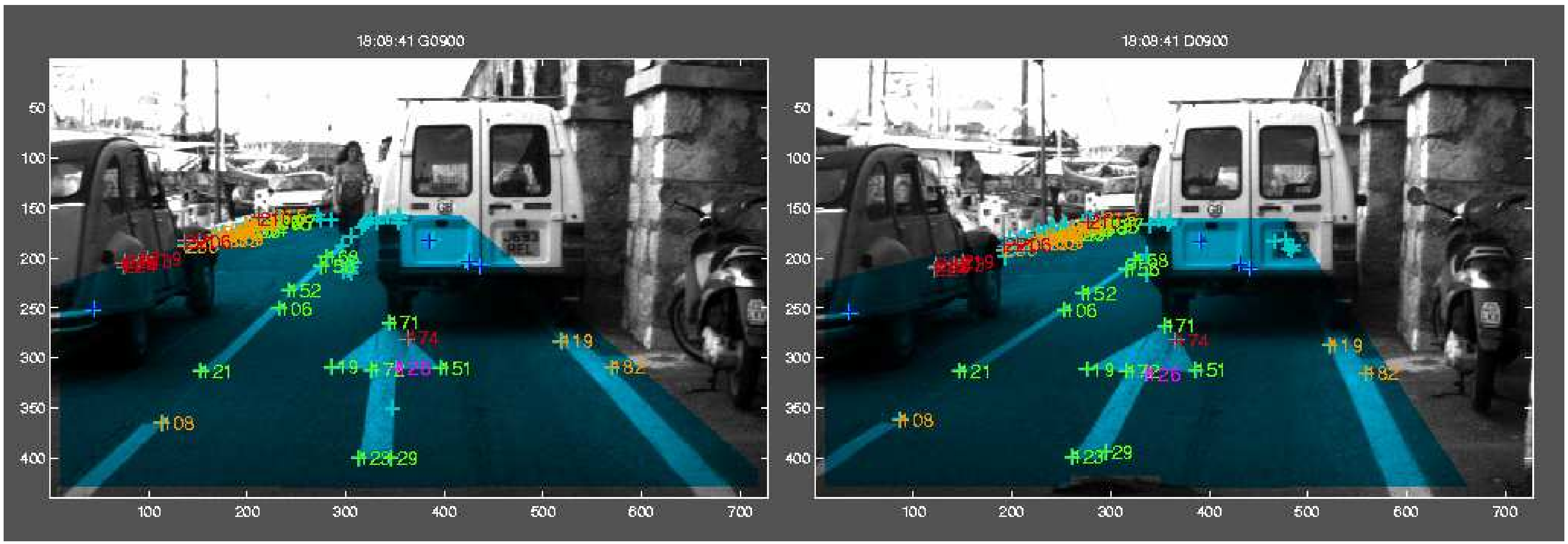
$$\mathbf{P_n^k} = \left[\mathbf{p_{ln}^{k-2}}\ \mathbf{p_{rn}^{k-2}}\ \mathbf{p_{ln}^{k-1}}\ \mathbf{p_{rn}^{k-1}}\ \mathbf{p_{ln}^{k}}\ \mathbf{p_{rn}^{k}}\right]^t$$

[Simond,N. and Rives,P. - IROS'2004]

Vehicle Egomotion

assuming that the height of the cameras is $d_l = d_r = 1m$ and the calibration matrix :

$$K_l = K_r = \begin{bmatrix} 1000 & 0 & 364 \\ 0 & 1000 & 236 \\ 0 & 0 & 1 \end{bmatrix}$$



Future issues

- apply the same approach to estimate the vertical planes;
- determine an affine auto-calibration of the stereo pair, using the structured environment;
- extraction of landmarks for a precise geo-referenced localization based on dGPS and cadastral map knowledge.

Questions ?

contact : Nicolas.Simond@sophia.inria.fr

Estimation of the Super Homography

Computation of SH^k:

- $q=0$;
- computation of ${}^0SH^k$ from SH^{k-1} and H^k_{st}, H^k_l, H^k_r;

while$(rank(SH^k) > 3)$;

  - $q = q+1$;
  - estimation of ${}^qSH^k$ from the SPIs ${}^{(\mathbf{q-1})}\mathbf{P^k}$;
  - computation of the new coordinates:

$$
{}^{\mathbf{q}}\mathbf{P^k} = \frac{1}{m}.{}^qSH^k.{}^{(\mathbf{q-1})}\mathbf{P^k};
$$

end