Trajectography of an Uncalibrated Stereo Rig in Urban Environments

Nicolas Simond, Patrick Rives

▶ To cite this version:

Nicolas Simond, Patrick Rives. Trajectography of an Uncalibrated Stereo Rig in Urban Environments. IEEE RSJ/International conference on Intelligent Robot and System (IROS'04), Sep 2004, Sendaï, Japan. inria-00001045

HAL Id: inria-00001045

https://hal.inria.fr/inria-00001045

Submitted on 20 Jan 2006

Trajectography of an Uncalibrated Stereo Rig in Urban Environments

Nicolas Simond, Patrick Rives

ICARE Project-Team, INRIA

2004 route des Lucioles

06 902 Sophia-Antipolis France

Email:{Firstname.Lastname}@sophia.inria.fr

Abstract— This paper describes an original method to compute the relative motion of an uncalibrated stereo rig in urban environments from features lying on the road. The extraction of significant reliable features on the road remains the critical step of this method. We nevertheless detect them according to the stereo constraints and a priori knowledge on the scene. The motion between two frames of the stereo rig is considered as rigid: the homography computation is enforced by the redundancy of the feature locations in multiple views. The method has been tested on video sequences recorded from a test vehicle that was driven an urban environments. Promising results from these experiments will be presented.

I. INTRODUCTION AND RELATED WORK

Urban environments appear difficult to model due to their inherent complexity linked to their respective topological and dynamical characters. However, the common static elements which structure the city streets are the road, the kerbs and the vertical facades. They all contain planar surfaces that we can easily model with a polygonal approximation. The identification followed by the track of these "natural" landmarks certainly authorizes the trajectography of a vehicle in these complex environments.

Hence, in "urban canyons" where the use of a GPS-based system is unreliable, a vision system is needed as the main localization sensor. Furthermore, this solution is much cheaper as systems using dead-reckoning on long ranges. Nevertheless, the basic drawback of a vision-based system is the need for a clear view that depends on the urban traffics.

Most of vision-based systems dedicated to mobile robots can be splitted in two main methods: gradient-based and feature-based approaches. The method generally depends on the type of application (road or obstacles detection, location of the vehicle between lanes, motion estimation), the scene (unstructured environments, highways, city streets), considering restrictive assumptions (translation motion, static environment or restricted lanes) or the type of vision-sensor (mono/stereo, camera, calibrated or not, black and white or color images). It is commonly assumed that the feature-based method is used in indoor environment or in outdoor application where a model of the environment is available.

Presently, we highlight few recent works which are representative of the last results of mobile robotics in outdoor applications. Stein and Shashua [1] use a direct methods combined with a probability function to compute the ego-motion of a vehicle relative to the road using a single calibrated camera. They effectively obtain accurate results with a night video where the road is illuminated by the headlights of the car. Katsura et al. [2] segment in-line the elements which structure the streets (buildings, trees and sky) to cope with the illumination or seasonal changes a mobile robot have to face when it autonomously moves along a guided recording path.

Otherwise, Turchetto and Manduchi [3] use a stereo rig to estimate the borders of an unstructured road, considering a correlation field which is the scalar product of the gradient of the elevation field and the brightness field. The authors indeed assume that the both should be collinear and have large values on a curb's face near its edge. Near to our method, Okutomi et al [4] locate the ground plane by computing the image projective invariants using a calibrated stereo vision.

Taylor et al. [5] estimate the pose and the orientation of a stereo vision system by considering the lanes with real-time constraints. Se and Bradly [6] describe 3 methods to estimate the pose of cross-walks and stair-cases to perform outdoor navigation. These elements have parallel stripes with constant width which can be discriminate from their edges or corners using projective tools as homography and vanishing lines.

Our goal is the estimation of the trajectory of a vehicle from the data we can extract from an on-bord uncalibrated stereo rig. In a previous article [7], we detail the method we use to extract features on the road plane and justify our assumptions on the scene. The contribution of this article is a method to compute the trajectography of the cameras only taking account the locations of features in multiple views and a priori knowledge of the scene. We hence improve the estimated results by considering the spatial and temporal constraints induced by the motion of the stereo rig during the video sequence.

II. METHODOLOGY

We present in the sequel an overview of the complete method. We detail in the next subsections the main steps: the segmentation of the road plane, the homography computation between two generic views, the super-homography computation and the estimation of the cameras motion.
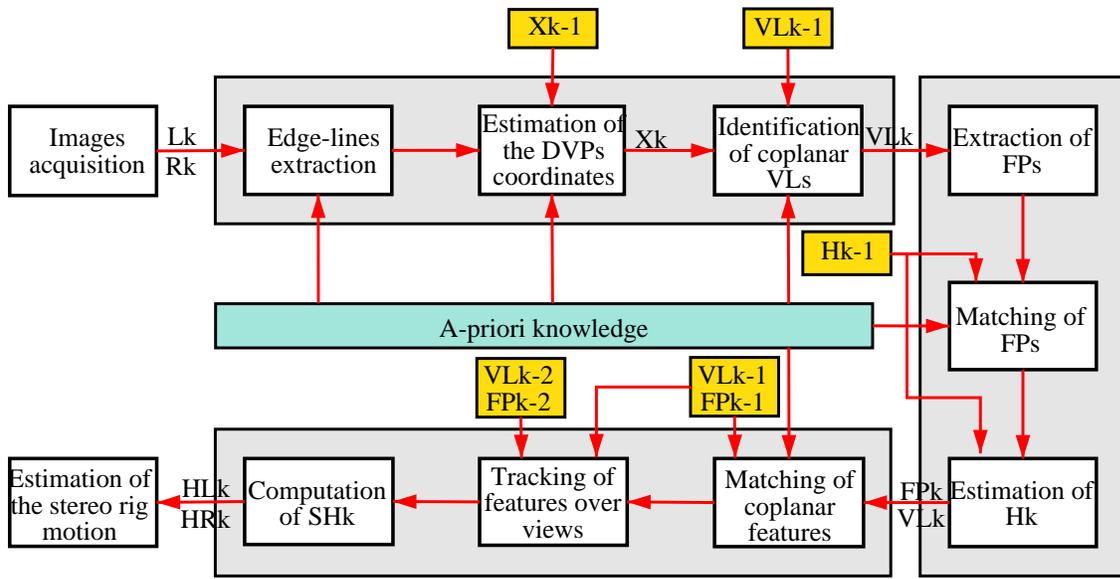
Fig. 1. Schematic diagram of the whole algorithm. The first stage correspond to the segmentation of the road plane in images L_k and R_k. The second stage extract coplanar FPs FP_k by the computation of the homography H_k. The use of VL_k in this operation is not represented. The third stage provides the left and right camera motion HL_k and HR_k according to the list of features we succeed to match between different views.

A. Assumptions and outlines

The city streets contain sets of planar surfaces which can be used to estimate the motion of a mobile robot by computing the homography between two frames. The main difficulty also is to identify planes in the scene. If the road can be locally considered as a plane, we hence possess a reliable link between the frames all along the video sequence. Like most of authors, we assume:

- the road has parallel borders or and may contain lines,
- the camera model is a pinhole camera model,
- the high frame rate of the video sequence allows small motions of the characteristics features in images.

The estimation of the stereo rig motion between two frames requires 3 main steps (see Fig.1). First, the extraction of the road plane in both images by taking account parallel coplanar lines. The second step detects and identify which are the coplanar feature points considering the stereo constraints. The third step tracks and matches the coplanar features between the consecutive frames to estimate the motion of each camera.

B. Segmentation of the road plane in the image

Two sorts of features can be detected from images of urban environments : edge lines and Feature Points, we develop in II-C.1. Edge lines represent the fusion of segments issued from the same contour after a Canny edge detection. Due to the perspective projection, a subset of parallel lines in the 3D scene projects onto the image as a pencil of edge-lines called Vanishing Lines (VLs) which converge to the same point, called Vanishing Point (VP). The VP, which belong to the plane at infinity, is the projection in the image plane of the common intersection of 3D parallel lines.

Among all the VPs we can detect in the image, we only focus on the Dominant Vanishing Point (DVP). The DVP is the intersection of the 3D edge-lines parallel to the road boundaries. In the following, VLs only represent the subset of edge-lines converging to the DVP location.

1) Extraction of the Dominant Vanishing Point: By definition, the DVP location $\mathbf{X}$ is invariant to the translation motion of the camera. Thus, the variation of the DVP coordinates in the image has two different origins : the pan angle when the vehicle turns and the tilt angle when the vehicle speeds up or brakes. The variation on the tilt angle infers moreover to the representation in image of the road plane normal $\mathbf{N_k}$.

In normal conditions of traffic, we assume that the DVP coordinates satisfy a stationary process in both images. The Kalman filter we implement provides a prediction $\widetilde{\mathbf{X}}_k$ of the DVPs locations at each new frame. A subset of edge-lines is also selected in each image which either converge to the DVP prediction or have similar characteristics (orientation, gradient, location) than the VLs used at the last iteration. The observation of $\mathbf{X}_k$ used in the Kalman filter is then recursively computed by solving the weighted least square system formed by the selected edge-lines. The smoothness of the filtered representation of the DVP coordinates in Fig.2 confirm our assumption.

2) Matching of coplanar Vanishing Lines: The matching between VLs in two views splits into two steps. In the first step, we try to find correspondences between VLs detected at the last iteration with the current ones, according to the fact that the variation between two consecutive images of the same camera is really weak. We also segment the road plane area with the extrema corresponding VLs.

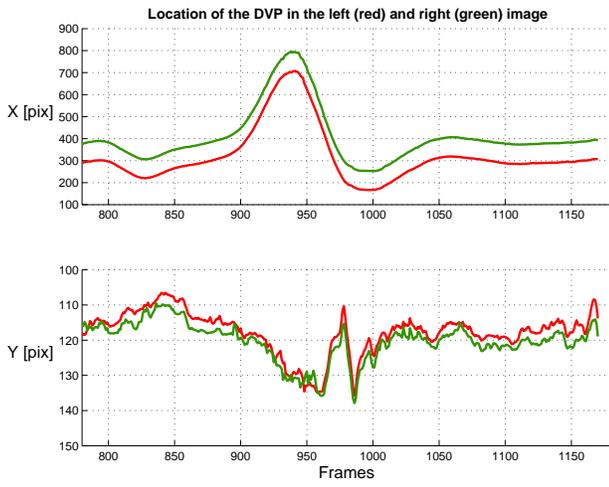In the second step, after the estimation of the current homography H_k, we match the two stereo pencils of VLs

Fig. 2. Chronogram of the left (red) and right (green) DVP coordinates (u_X at the top, v_X at the bottom), along a video sequence recorded on the harbor sequence which consists on an over-pass of a stopped vehicle along a straight road.

considering the angle difference between the projection of the left VLs with the right ones. Only the couples which have a difference less than $5°$ and the same gradient are considered as good matches.

At the first iteration ($k = 1$), we assume that we can extract from L_1 and R_1 two sets of VLs converging to the DVPs $\mathbf{X}_{l1}$ and $\mathbf{X}_{r1}$. We hence estimate the DVP locations $\mathbf{X_1}$ as the coordinates which minimize the mean square weighed distance between most of the edge-lines. The method could be viewed as a Hough transform where the DVP location is the intersection of most of VLs. We then identify the correspondence between the VLs which formed the two pencils of VLs using the invariance of the cross-ratio.

C. Homography between two views

Two types of homography are estimated at each iteration. The first one permit us to determine which are the features lying on the road plane among all those are detected in the stereo images. Indeed, only the features lying on the road plane verify the planar homography H_k between the stereo views. The second type of computation estimates the motion of each camera HL_k and HR_k between two frames, considering the coplanar features we have just extracted.

The use of a rigid stereo rig makes the identification of coplanar features simplest. We indeed possess a prediction of the estimated homography $\widetilde{\mathsf{H}}_k = \mathsf{H}_{k-1}$ which ease the matching operation and the rejection of outliers (see Fig 3). The rigid motion between two cameras fixed on a vehicle only depends on the pan/tilt rotation angles of the vehicle:

$$\mathsf{H}_{k-1} = \mathsf{K_r}.[\mathsf{R} - \frac{\mathbf{T^t}.\overline{\mathbf{N}_{k-1}}}{\mathsf{d}_{k-1}}].\mathsf{K_l}^{-1}$$

$$\mathsf{H}_k = \mathsf{K_r}.[\mathsf{R} - \frac{\mathbf{T^t}.\overline{\mathbf{N}_k}}{\mathsf{d}_k}].\mathsf{K_l}^{-1}$$

with $\mathsf{K_l}, \mathsf{K_r}$ matrix of intrinsic parameters of the left and right cameras and $\mathbf{N^t} = [\overline{\mathbf{N^t}}, d]$ 3D coordinates of the π

plane normal, considering the center of the left camera as the framework origin. R and $\mathbf{T}$ represent the rigid transformation between the camera frames. The variation of the $4th$ coordinate d_k is not significant in our application because the altitude of the left camera from the road plane varies with the tilt angle which is less than $10°$.
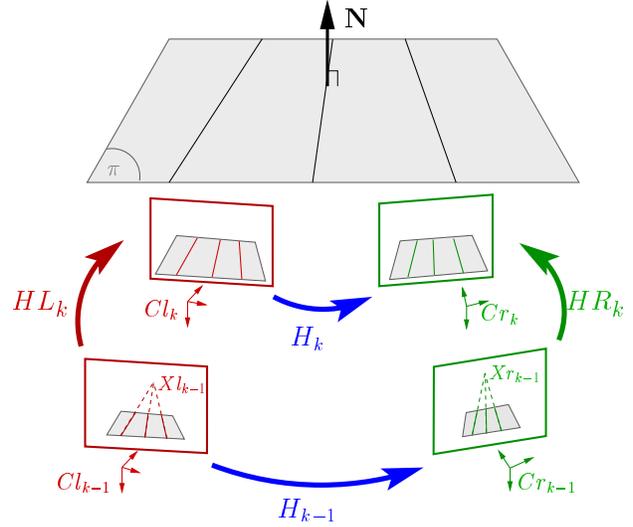

Fig. 3. We consider the homographies H_{k-1} and H_k induced by the π plane between the rigid stereo rig in frame $(k-1)$ and k. Only the pan/tilt rotation motion between the two orientations of the left camera infer to these homographies estimations. This orientation is fully characterized by the representation of the plane normal. The rotation motion between two frames can hence be deduced from the variation between the normalized plane normal N_{k-1} and N_k.

1) The Feature Points: The computation of a homography requires at least 4 independent coplanar features. However, the linear dependency (that is the DVP) between the VLs do not allows a homography computation using only a pencil of corresponding VLs. We then extract with a Harris detector the feature points (FPs) in the region of the image considered as the road area. We only consider those which have an appropriate Harris score [8] upper than an experimentally threshold. This value of threshold allow the detection of reliable FPs between images.

The main features we can detect on the road plane are edges of road markers and their corners when they have acute angles. The Harris filter sometimes detects few cracks on the asphalt, corners of large variations of the pattern like in the pedestrian crossing area or painted arrows.

Unfortunately, other features can be detected on the road plane area: they generally represent obstacles lying on the road plane which do not belong to the road plane. These features have to be rejected before the computation of the homographies.

2) Matching of the Feature Points: The matching of the FPs is the crucial step of the whole algorithm. Indeed, we search couples of corresponding FPs in two views considering a-priori knowledge on the scene. The relative motions between the corresponding FPs are strongly constrained by the plane:

1) the foreground of the image is generally free from dynamic obstacles,
2) the stereo rig is rigid: the corresponding coplanar FPs follow the same motion, whose we know a prediction in image H_k,
3) in case of matching between two stereo images, if the cameras have equivalent intrinsic parameters and parallel image planes, then the relative motion is mainly a lateral translation,
4) the distance between the coordinates of the same FP in two consecutive images increases with the FP ordinate. The DVP location looks stationary whereas FPs on the bottom of images have relative motions higher than $30\,pixels$.

We also establish a list of corresponding candidates for each FP of the both images. We create a sparse matrix Corr $[n_{fp1}, n_{fp2}]$ which is fulfilled with the covariance correlation between $[30 * 30]\,pixels$ regions, centered on each couple of the list. We next reset to 0 each coefficient under a fixed threshold. Two FPs can be considered as corresponding if their covariance correlation is upper than 0.5 between two stereo images and 0.8 between two consecutive images of the same camera. The pairs of corresponding FPs are finally obtained by the SVD decomposition of this appropriate correspondence strength matrix, according to [9].
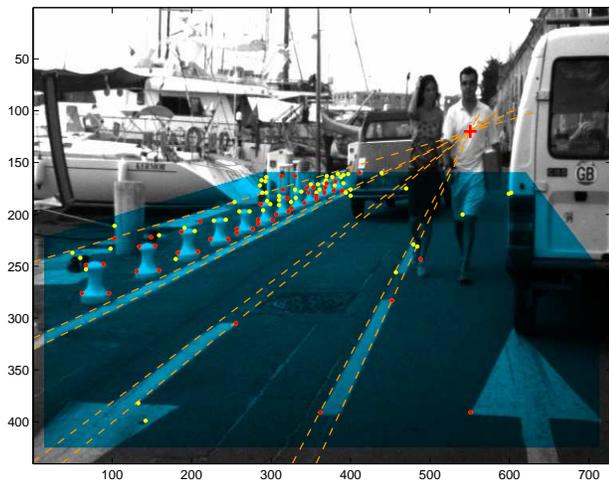

Fig. 4.   Left view of the stereo rig. The yellow FPs are detected in the blue region of interest, limited by the matched VLs (orange dashed lines) and a-priori knowledge. The coplanar FPs highlighted in red, as the DVP (red cross), allows the stereo homography H_k. Most of FPs located at an altitude lower than 20 cm from the road plane are rarely rejected with the left/right homography computation. In this frame, there is not enough true features lying on the road plane to constrained correctly the homography computation although the prediction was satisfying. In the same way, short cameras motion or short period of image acquisition do not permit to distinguish direct features issued from dynamical obstacles to real FPs.

Nevertheless, this operation provides unreliable results on FPs in two cases: the discrimination between FPs is ambiguous due to their proximity or their re-projection distance is lower than the threshold condition like in Fig.4. Therefore, the list of corresponding FPs is updating during the step of homography computation.

3) Estimation of a homography: According to [6], the homography computation requires VLs even if FPs are detected on them to improve the robustness. Indeed, the VLs constraint the plane motion in occluding parts. Nevertheless, our method could not be associated to a basic lane detector because we only assume the road has parallel borders.

The couples of corresponding features lying on the road plane have coordinates which satisfy the linear system issued from the next two equations, where $\mathbf{l}$, $\mathbf{p}$ respectively represent the coordinates of a VL and a FP in two images:

$$\mathbf{l_2} \propto H^{-t}.\mathbf{l_1} \quad \text{and} \quad \mathbf{p_2} \propto H.\mathbf{p_1}$$

The linear system is then composed of two distinct subsystems formed by the coordinates of the VLs and the FPs. It is difficult to optimize the conditioning number of the system matrix due to the two types of coordinates: $[cos(\theta), sin(\theta), -\rho]^t$ for the VLs and $[u, v, 1]^t$ for the FPs. We then estimate iteratively the homography considering the match of VLs as right. We also reject the FPs which do not lie on the road plane. According to II-B.2, as soon as the coplanar FPs are identified, we verify the correspondence between the two pencils of VLs.

We qualify the reliability of the homography computation by considering the variance of the distance between projected FPs. If the variance is over the threshold $\sqrt{2.65}.\sigma$ with $\sigma = 1\,pixel$, we re-estimate the homography with a robust RANSAC (random sample consensus) algorithm. The threshold of correspondence is also equal to $\sqrt{5.99}.\sigma$ with this estimator according to [10]. This method is also used for the first iteration ($k = 1$) when no prediction of H_1 is available.

D. The super-collineation matrix

The computation of homographies depends on the coplanar features we achieve to detect. To improve the quality of each estimation by rejecting some outliers, we introduce spatial and temporal constraints on feature locations by considering several views. The goal of the algorithm is to estimate iteratively the ideal coordinates of the features such as the homographies hence induced verify cross-composition between different views.

1) The method: Malis and Cipolla [11] describe an efficient method to impose the constraints between the homographies computed from a sequence of views of a planar structure. Taking into account multiple views provides a set of constraints between the coordinates of coplanar features in different views. Moreover, it minimizes the effects of errors on the matching step and it reduces numerical instability when the motion between two views is not significant.

Such a method improves the consistence of the current homographies (H_k, HL_k, HR_k) estimation. Nevertheless, we also face a hard compromise between increasing the distance between two views to reduce the numerical instability and keeping a significant number of correspondent features to constrain the computations of homographies.

The method introduces a super-collineation matrix SH_k $[3m * 3m]$ which contains all the [3*3] homographies between the $m(m-1)$ couples of different views. In our case, we use the three stereo couples corresponding to the frames $(k-2), (k-1)$ and k, the number of images is then $m = 6$. The authors demonstrate that $rank(\mathsf{SH}_k) = 3$ whatever $m \geq 3$.

2) The Virtual Feature Points: Taking into account the coplanar VLs characteristics in the SH_k estimation requires an adaptation to the FPs coordinates. We therefore introduce the Virtual Feature Points (VFPs) to use the same representation as the FPs. The VLs effectively do not verify the same homography than the FPs. Furthermore, all the VLs converge to their respective DVP, the constraints due to the VLs locations are always applied when we use the VFPs.

The VFPs are then issued from the intersections of the coplanar pencil of VLs with a line lying on the road plane, defined by two coplanar FPs whose the coordinates are known in the m views. The global feature matrix $\mathbf{F}$ $[3m, n]$ is then composed of the coordinates of the FPs and the VFPs in the m images with $n = n_{FP} + n_{VL}$.

3) Tracking of coplanar features over the sequence: The 3 homography computations ($\mathsf{H}_k, \mathsf{HL}_k, \mathsf{HR}_k$) provide 3 lists of correspondences between the features of the couples of views. We then verify and identify the cross links to fulfilled a dedicated structure where each coplanar feature is labeling. We hence record principally the coordinates, the number of the frame and its type (left or right image). Nevertheless, if no correspondence between the left or right images ($L_{k-1}L_k$ or $R_{k-1}R_k$) makes the link between a couple of features detected at the last frame with their new locations in the current frame, the algorithm regards current couple of features as new.

4) Estimation of the super-collineation matrix: The initialization of the $(3 * 3)$ homography of SH_k^1 between each couple of images is issued from the composition of the previous estimation of homographies at the frame $k, k-1, k-2$, e.g.:

$$\mathsf{SH}_k^1(L_k, R_{k-2}) \propto \mathsf{HL}_k.\mathsf{HL}_{k-1}.\mathsf{H}_{k-2}$$

.

For $q \geq 2$, the estimation of the sub-matrix $\mathsf{SH}_k^q(M_r, M_c)$ is computed from the homography between the coordinates $\mathbf{F}_k^q$ of the n selected features in the images mr and mc:

$$\mathbf{F}_k^q(M_r, :) \propto \mathsf{SH}_k^q(M_r, M_c).\mathbf{F}_k^q(M_c, :)$$

with $M_r = (3mr-2 : 3mr)$ and $M_c = (3mc-2 : 3mc)$ when $\{mr, mc\} \in [1, m]$.

The algorithm copes even some feature coordinates are unknown in few images. The iterative optimization algorithm generally stops after 3 iterations unless the system was ill-conditioned due to errors on feature matching or the detected features do not constraint enough the homographies.

The sub-pixellic coordinates of the n selected features in the L_k and R_k current images are respectively $\mathbf{F}_k^{q+1}(3m-$

$5:3m-3, :)$ and $\mathbf{F}_k^{q+1}(3m-2:3m, :)$ whereas the current homographies ($\mathsf{H}_k, \mathsf{HL}_k, \mathsf{HR}_k$) can be extracted from the final estimation of SH_k:

$$\mathsf{SH}_k = \begin{bmatrix} \mathsf{I}_3 & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \ddots & \cdots & \cdots & \cdots & \cdots \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \cdots & \cdots & \mathsf{HL}_k & \cdots & \mathsf{I}_3 & \mathsf{H}_k^{-1} \\ \cdots & \cdots & \cdots & \mathsf{HR}_k & \mathsf{H}_k & \mathsf{I}_3 \end{bmatrix}$$

E. Motion of the cameras

We have recorded a video sequence (see Fig.5) in the harbor neighborhood of Antibes (France). The vehicle motion is straight with a change of lane to overtake a parked car. The road appears structured. The speed was about 10 m/s and the frame rate was 25 Hz . The pinhole model cameras are uncalibrated. We suppose in the following a generic calibration matrix such that:

$$\mathsf{K}_l = \mathsf{K}_r = \begin{bmatrix} 1000 & 0 & 364 \\ 0 & 1000 & 266 \\ 0 & 0 & 1 \end{bmatrix}$$

We hence compute from each homography the extrinsic matrix $\mathsf{K}_2^{-1}.\mathsf{H}.\mathsf{K}_1$ to finally extract the rotation matrix R_{12}, the vectors associate to translation motion $\mathbf{T}_{12}$ and the road normal $\mathbf{N}$ in the first view. We use a specific function which provide the estimated parameters according to a prediction of $\mathbf{N}$. There is indeed two numerical solutions to the extraction whose one is wrong.

The Fig.6 represents the evolution of the motion between the two cameras of the stereo rig. The results are quite stationary except during the frames 920 and 960 where the detected features do not constraint enough the homography computation. A filtering process certainly allows a constant estimation.
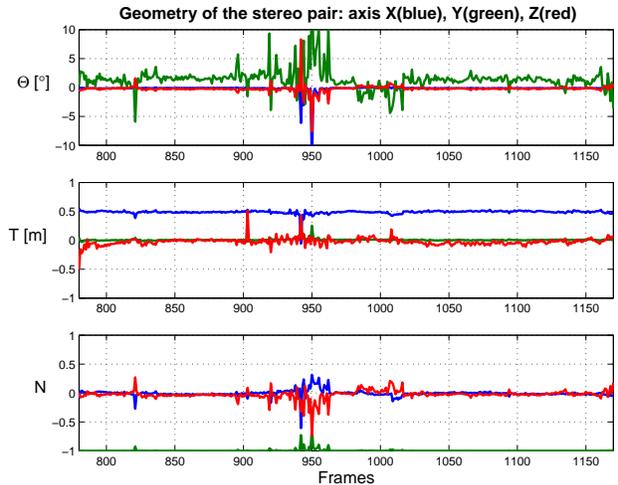

Fig. 6. Chronogram of the evolution of the rigid stereo rig parameters over the sequence. Above: angles of rotation around the axis of left camera. Middle: coordinates of the center of the right camera, the scale factor d is considered equal to $1\,m$. Bottom: normalized coordinates of the road normal, viewed from the left camera.
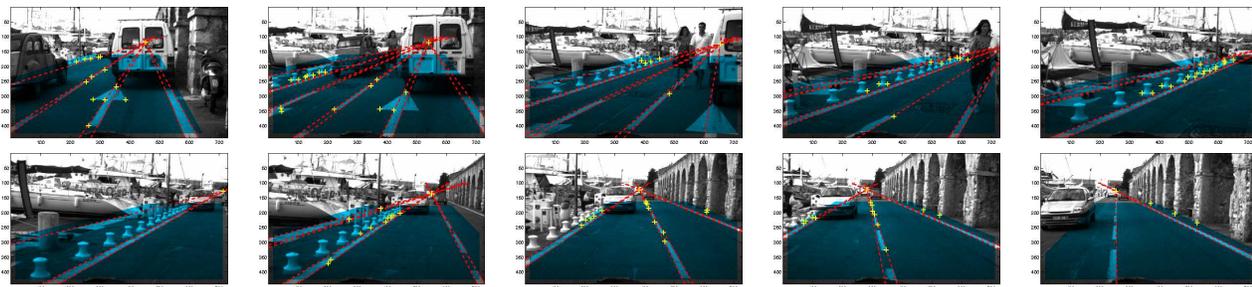
Fig. 5. Top-left, bottom-right, details of right view during the difficult part (900-990) of the sequence every 10 frames. The algorithm extracts on the blue region of interest the coplanar VLs (red dashed lines) and FPs (yellow '+'). The test sequence is a straight motion with an overtaking of a parked vehicle.

The Fig. 7 shows the estimation of the right camera motion between two consecutive frames along the sequence. The motion estimation remains noisy as expected but we nevertheless retrieve the vehicle motion: it drives in straighted lines to stop before the parked car, it overtakes it and continue its normal cruise. In other hand, we stress that the clear view obstruction due to the driving vehicle has no critical effect on the motion estimation between the frames 892 and 915.
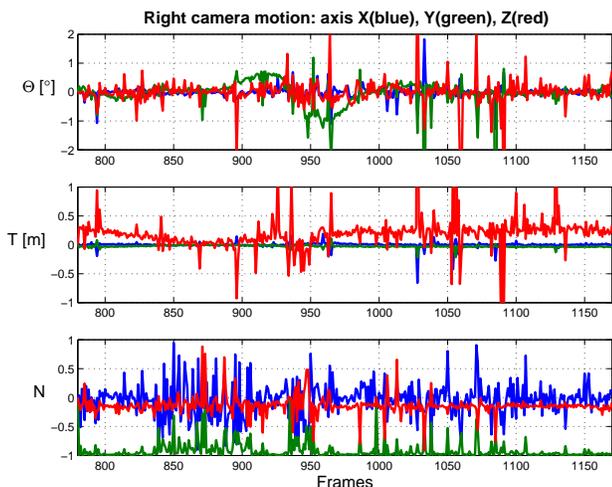

Fig. 7. Chronogram of the motion parameters between two consecutive views of the right camera. Above: angles of rotation around the axis of right camera. Middle: 3D translation coordinates of the center of the right camera, the scale factor d is considered equal to $1\,m$. Bottom: normalized coordinates of the road normal, viewed from the right camera.

III. CONCLUSION AND FUTURE WORKS

We present an original method to estimate the trajectography of a stereo rig in the urban environments. We assume the road as locally planar and detected ocomputationsn it features (points and lines) considering the constraints induced by the stereo vision and a-priori knowledge of the scene. The camera motion is estimated between frames considering only the coplanar features hence detected. The homographies computation are improved by taking account temporal and spatial constraints induced by the feature locations in multiple views.

Our future work consists on filtering the motion data considering the motion of the two cameras and next establish an on-line auto-calibration of the stereo rig to improve the autonomy of the vehicle. In the peculiar environment of the urban canyon, we certainly can extract some new planes from the facades and hence estimate the intrinsic parameters of the stereo rig.

REFERENCES

[1] G. Stein, O. Mano, and A. Shashua, "A robust method for computing vehicle ego-motion," in *IEEE Intelligent Vehicles Symposium (IV'00)*, Dearborn, MI., Oct. 2000.

[2] H. Katsura, J. Miura, M. Hild, and Y. Shirai, "A view-based outdoor navigation using object recognition robust to changes of weather and seasons," in *IEEE RSJ/International conference on Intelligent Robot and System (IROS'03)*, Las Vegas, Nevada, USA, Oct. 27-31 2003, pp. 2974–2979.

[3] R. Turchetto and R. Manduchi, "Visual curb localization for autonomous navigation," in *IEEE RSJ/International conference on Intelligent Robot and System (IROS'03)*, Las Vegas, Nevada, USA, Oct. 27-31 2003, pp. 1336–1342.

[4] M. Okutomi, K. Nakano, J. Maruyama, and T. Hara, "Robust estimation of planar regions for visual navigation using sequential stereo images," in *IEEE International Conference on Robotics and Automation (ICRA'02)*, Washington DC, USA, May 2002, pp. 3321–3327.

[5] C. Taylor, J. Kosecka, R. Blasi, and J. Malik, "A comparative study of vision-based lateral control strategies for autonomous highway driving," *International Journal of Robotics Research*, vol. 18, no. 5, pp. 442–454, may 1999.

[6] S. Se and M. Brady, "Road feature detection and estimation," *Machine Vision and Applications Journal*, vol. 14, no. 3, pp. 157–165, July 2003.

[7] N. Simond and P. Rives, "Homography from a vanishing point in urban scenes," in *IEEE RSJ/International conference on Intelligent Robot and System (IROS'03)*, Las Vegas, Nevada, USA, Oct. 27-31 2003, pp. 1005–1010.

[8] C. Harris and M. Stephens, "A combined corner and edge detector," in *Alvey Vision Conference*, Manchester, UK, Aug. 1988, pp. 147–151.

[9] M. Pilu, "A direct method for stereo correspondence based on singular value decomposition," in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'97)*, San Juan, Puerto Rico, June 17-19 1997, pp. 261–266.

[10] R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*. Cambridge University Press, 2000, pp. 101–112.

[11] E. Malis and R. Cipolla, "Multi-view constraints between collineations: application to self-calibration from unknown planar structures," in *European Conference on Computer Vision (ECCV'00)*, vol. 2, Dublin, EIRE, June 2000, pp. 610–624.