AUTOMATIC LANE PARAMETERS EXTRACTION
IN MOBILE MAPPING SEQUENCES

R. Roncella*, G. Forlani

Dept. of Civil and Environmental Engineering and Architecture, University of Parma, Parma, 43100, Italy
riccardo.roncella@ unipr.it, gianfranco.forlani@unipr.it

Commission V, WG 1

ABSTRACT:

Technical prescriptions for road inventories specify that the road width must be measured with a tolerance of 10 cm and that variations larger than the same threshold should be spotted and recorded.
In the framework of software development for Mobile Mapping, underway at the Department of Civil Engineering of Parma University, a semi-automatic method for the extraction, the tracking along an image sequence and the measurement of the distance between lane markings was developed. The input data is a sequence of stereo images with known interior and exterior orientation. The program starts by interactively defining a ROI for each lane to be extracted; both images are then rectified with respect to a common object plane; by edge detection the markings extracted in each image pair; edge pixel are selected and interpolated with a straight line and the distance between these lines measured. From their differences the true distance and the height difference due to the cross slope are computed; the measurement is repeated at specified intervals and recorded. By predicting the position of the markings in the next image pair, the sequence is scanned as long as the markings can be reliably extracted. When, for a specified number of successive pairs, no marking is found, the operator is requested to restart the process. The first trials on artificial and real image sequences have shown accuracy well within the specifications and the robustness of the extraction.

1. INTRODUCTION

1.1 Mobile Mapping and road inventory

Mobile mapping vehicles (*Schwarz and El-Sheimy*, 1996) are nowadays the most efficient systems to create and maintain road inventory databases. They are equipped with positioning and orientation sensors (GPS, INS, odometers) and data acquisition sensors (such as digital cameras, laser sensors) for road feature measurements). The survey usually concern areas 10 to 20 meters across the road axis and supplies an absolute positioning precision up to 20-30 cm while, in terms of relative precision (i.e. distances between features), accuracies range from a few cm to 10-20 cm. The main informations on road characteristics are acquired by means of the images taken by the vision system. Virtually all vehicles are equipped with several synchronized digital cameras, which cover the road and its neighbourhood. In Italy the number of cameras and their mounting position on the vehicle are often defined by the technical prescriptions of the tender. Nevertheless different approaches may be utilized for road inventory purpose: beside monoscopic systems with every camera pointing in a different direction without image overlapping (in the past, this was the preferred set-up, to limit the image database size), stereoscopic or multi-image systems are more and more employed in this kind of application. Even if synchronization, image storage and hardware cost issues are more demanding in the latter systems, development of improved acquisition hardware and restitution software will soon fade away these issues, while the benefits provided by three-dimensional restitution over monoscopic ones (i.e. higher accuracy, more reliable results, higher flexibility, etc.) will gain importance the more automation in feature extraction is introduced.
However, at the present, processing and restitution in such scenarios are still one of the main issues to take into account in the development, setting and utilization of a mobile mapping vehicle.

2. LANE WIDTH MEASUREMENT

One of the main features to be provided in road inventory applications is the road width: its variation may determine a change in the classification of the road, since the level of service of the road depends on its geometric characteristics; critical spots, were e.g. the roads narrows, should be highlighted to take appropriate decisions (i.e. speed limits, alternative routes, etc.). Since old roads lack a project (the actual geometric characteristics are often unknown) and very often just spot changes were made to the track, a point wise knowledge of the road width is a basic requirement for planning improvements and standardization of the level of service.
With our current vision system features (base length = 1.7 m, focal length = 8 mm, mean expected accuracy in image point extraction = 1 pixel), to obtain road markings position accurately, collimated points must lie in the 10÷15 meters range in front of the vehicle.
In the framework of software development for Mobile Mapping, underway at the Department of Civil Engineering of Parma University, a semi-automatic method for the extraction, the tracking along an image sequence and the measurement of the distance between lane markings has been developed.

3. AUTOMATIC LANE WIDTH EXTRACTION

The procedure, which is largely inspired by a technique developed for the real time extraction of road lane markings in autonomous driven vehicles (*Broggi*, 1999), can be summarized as follow (see figure 1): acquisition and rectification of a stereo image pair; edge extraction on each rectified image; comparison of extracted lines between the two images to

account for the cross slope; computation and recording of the road section length; prediction of the edge position on the next stereo pair. In detail, low level segmentation is performed on the stereo rectified images using high pass filters tailored to vertical edge identification. The highlighted vertical edges are likely to represent road and lane markings. By means of the Hough Transform, a list of straight lines which approximate the edge pixels is obtained.
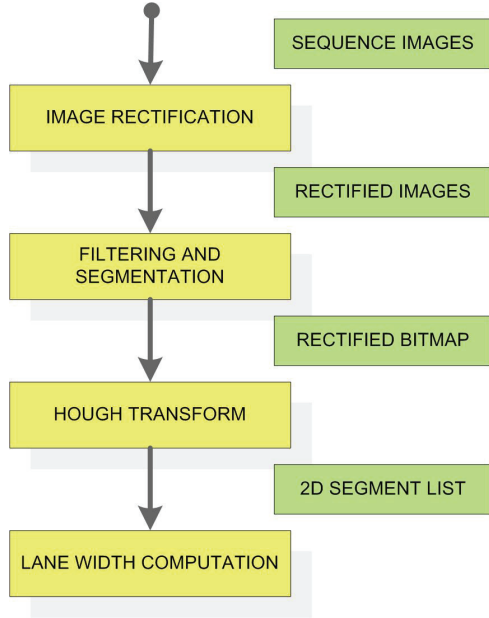

Figure 1. Workflow of the automatic lane extraction algorithm.

Lane marks are grouped in different lists: one for the leftmost border, one for the center and one for the rightmost line. Assuming that in each list at least one line is present, road width measurement could be performed on the rectified stereopair. By measuring the width on a synchronous stereo pair the approach is actually stereoscopic and a full 3d measurement takes place.

In the following sections each step of the algorithm is described and analyzed in detail.

3.1 Image rectification

Due to the perspective effect, the marks on the original images in straight road segments converge to a vanishing point located approximately at the center of the horizon line. This complicates the localization of the road marks since no strict constraint can be applied in the extraction procedure. Actually the lines are not parallel, and even if the vanishing point is located on the horizon line (which basically is fixed on the image, at least if the vehicle is not subject to pitch and roll) its exact position is not known. Thus it is difficult to correctly find out where the road section lies on the images. Furthermore, in this scenario, more ambiguities in road lane extraction arise: all the scene lines convergent to the vanishing point must be considered as possible lane marks. Finally, to limit computational efforts, it could be advisable to process only a small portion of the image around each line. Again, due to the perspective effect in the original image space, it's hard to contour such portion.

On the contrast, rectifying the images with respect to the road mean plane, the computation of the image portions to be

analyzed becomes very simple. Actually it is sufficient to know the homography which maps the image space with the selected road mean plane: an exact knowledge of the lane plane location and orientation is not required. As already underlined, the method is stereo and image rectification only allows a more efficient extraction of the markings. From the position of each perspective center (i.e. the exterior orientation of each frame obtained by the navigation system) subsequent images may be tied together. In straight trajectories road marks are approximately represented by vertical lines, while cross section of the road are basically horizontal in the rectified images.
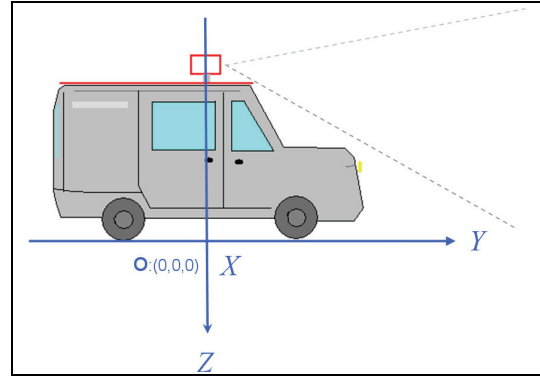

Figure 2. Reference system used in images rectification.

Using just one rectified image is not suitable to obtain accurate measurement of the road width: out of plane effects leads usually to error of 10 – 20 cm due to the imaging geometry. The use of two rectified images, acquired by different positions with adequate base length, allows a full 3d reconstruction of the road features. The homography computation for both cameras can be performed as follows: consider the reference system depicted in figure 2 with its Z axis pointing down in the vertical direction and Y axis approximately parallel to the vehicle motion direction. The reference system is strictly connected to the vehicle and therefore moves according to it. The road plane, in such conditions, is basically parallel to the XY reference plane, and all points on it have null Z coordinate. Suppose that each camera is associated with its corresponding algebraic projector (i.e. the camera matrix) P; in the previously depicted reference system the camera matrix parameters don't change as the vehicle moves. As shown by several authors (*Hartley and Zisserman*, 2000) the homography between object plane (i.e. the road mean plane) and image plane can be easily computed from the camera matrix P:

$$x_i = P_i \cdot X \qquad \begin{vmatrix} x_1^i \\ x_2^i \\ x_3^i \end{vmatrix} = P_i \cdot \begin{vmatrix} X_1 \\ X_2 \\ X_3 \\ X_4 \end{vmatrix} \qquad [1]$$

$$x = \begin{vmatrix} p_1 & p_2 & p_3 & p_4 \end{vmatrix} \begin{vmatrix} X_1 \\ X_2 \\ 0 \\ X_4 \end{vmatrix} = \begin{vmatrix} p_1 & p_2 & p_4 \end{vmatrix} \begin{vmatrix} X_1 \\ X_2 \\ X_4 \end{vmatrix} = H_{3x3} \cdot \begin{vmatrix} X_1 \\ X_2 \\ X_4 \end{vmatrix}$$

where $P_i = P_i(X_L^i, Y_L^i, Z_L^i, \varpi_L^i, \varphi_L^i, \kappa_L^i)$ and p_i is the *i*-th column of the camera matrix P.

The homography matrix for each camera can be computed from the corresponding camera matrix, basically disposing of its third column; a LUT (Look Up Table) mapping the

transformation may be computed, to limit computational efforts at each sequence step.

3.2 Segmentation of the rectified image

Lane markings recognition and extraction require a pre-processing to obtain simplified features onto the images. To this aim filtering and segmentation algorithms are applied in sequence:
1. Using a Sobel mask, edge detection is performed on the rectified images. Since the road markings are expected to be approximately vertical only the horizontal gradients are taken into account in the filtering (noise in the orthogonal direction does not affect the extraction).
2. A median filter removes further spot noise in the data.
3. Eventually a binary filter with adaptive thresholding (the threshold is computed on the basis of previously recognized features) is used to obtain a bitmap image.

Further improvements in the filtering procedure are under development, mainly based on morphological operations (*Broggi*, 1995). Several issues are still to be solved in lane marks highlighting: presence of vegetation or shadows partly covering the marks, low contrast markings, etc.

3.3 Hough Transform

The Hough transform (*Hough*, 1959) is a technique which allow the identification of specific patterns in an image (i.e. line edges). The identification is obtained by mapping the patterns on a parameter space where the recognition is easier. The transformation densifies the information: all image pixels agreeing with the model are converted to a single point in the parametric space. The pattern recognition is performed in the parameter space (where information is concentrated) rather than in image space (where the information is intrinsically dispersed). Thus, the estimation of the model parameters is very robust i.e. it is not affected by noise or blunders in the data set. In our case the parameter space to be used can be generated by any couple of variables which identifies univocally a line: the slope θ and the distance from the origin ρ.

$$x \cos \theta + y \operatorname{sen} \theta = \rho \qquad [2]$$

On the $\theta\rho$ plane, the curve corresponding to a point (x,y) is described by a trigonometric function. Between the two spaces xy and $\theta\rho$ a duality exists: each line in a Cartesian space is mapped to a point in the parameter space while a curve (described by eq. [2]) in $\theta\rho$ corresponds to a point in the xy reference system.

For each edge pixel identified during the filtering step, the correspondent curve is drawn in the parameter space. If the pixels lie on the same line, on the parameter space the different curves will intersect each other at the θ,ρ coordinates which define the line on the xy plane.

An accumulation matrix is allocated representing a discretization of the $\theta\rho$ space. The accumulation matrix is composed by as many rows and columns as the number of intervals respectively for the parameter θ and the parameter ρ.

For each edge pixel the $\theta\rho$ points of the dual curve are computed and the correspondent position on the accumulation matrix is incremented: if the filtered image has been binarized the increment is 1; otherwise the value each cell of the accumulation matrix is incremented by the response value obtained of the image filtering. Once all edge pixel in the image space have been transformed, the accumulation matrix is analyzed seeking for local maxima, i.e. for the parameters $\theta\rho$ of the identified line.

The technique is as simple as effective. Actually not only it is almost not affected by gross errors or noise, but also can jump over short gaps in the edge that may have arisen in the segmentation due e.g. to shadows, poor contrast etc.

Thus, also broken lane markings or gaps in the solid ones do not harm the performance of the algorithm, unlike what may happen using line following algorithms, whose results are strongly affected by the threshold values used.

Both road intersections and Y junctions will normally either cause failure or provide inaccurate results, since there the assumptions underlying the method do not hold. On the other hand, being singular points, junctions require a specific survey; they are overcame interactively by the operator.

3.4 Lane classification strategy

Even if the Hough transform results are remarkably robust in the detection of linear entities, the presence of other linear features on the road section could create ambiguities during the extraction process.
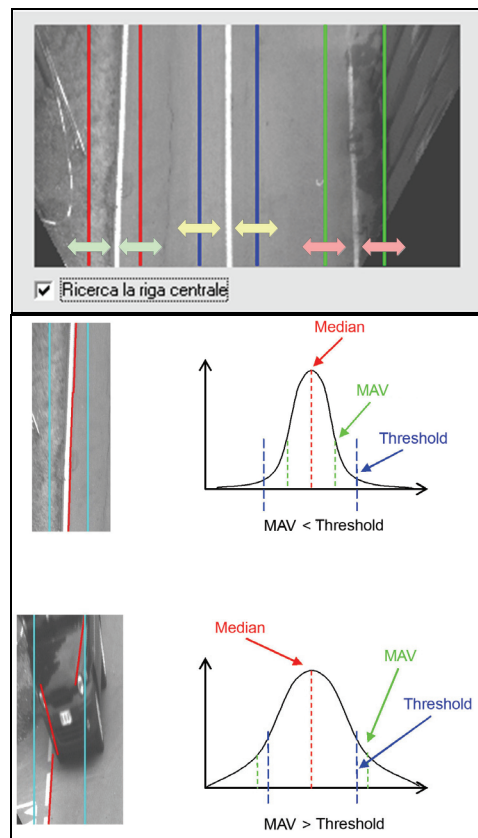

Figure 3. Lane classification strategies: (top) search ranges to limit the analysis region; (down) comparison and validation of the extracted edges.

In this case the road boundaries cannot be estimated with high accuracy. To make the estimation more robust and the extraction more effective, the Hough accumulation space is limited, i.e. only a few regions of the parameter space are considered as valid. Assuming that the edge lines appear on the rectified image approximately as vertical edges, the parameter θ is discretized on a small interval around the value 0: therefore all lines which don't lie almost vertically are discarded. The

interval is selected large enough to get the procedure working even on curved path, where the edges are inclined respect to the vertical line. This interval, moreover, can even be adjusted dynamically on the basis of previously extracted edges: if the road follows a pretty straight path the discretization interval can be narrowed; on the contrary, if edges are inclined, wider regions on the accumulation matrix will be used. Likewise, once the initial position of the edges lines is estimated, the algorithm applies the Hough transform only to a restricted area around the line, changing its amplitude based on the quality of the identification in the preceeding frames. Some critical cases may cause errors or failure: apart from sections where no line can be found in the image (e.g. due to occlusions or when actually the marking is missing), the internal consistency of the lines estimated in terms of θ is controlled. To this aim a post-processing step evaluates the morphological features of the identified lines. For each line list a median line is estimated and, when the absolute median deviation exceeds a certain threshold, the algorithm points out the possible unreliability of the result.

If the road width cannot be estimated for more than a given number of consecutive frames (currently three frames) the processing is suspended; the user may restore the correct position of the search ranges and restart the processing of the last frames or, if appropriate, move on along the sequence and restart from another position. This allows for example to go over a crossroad or to skip a sequence portion were heavy traffic in the opposite lane hide the lane marking.

3.5 Lane width computation

Once both the rectified images are analyzed and lane borders are detected on them, the road section width computation takes place. The stereo restitution might be performed in the original image space. Nevertheless it is also possible to perform the computation directly in the rectified space: consider a generic object point X (let its homogeneous coordinates be $(X_1\ X_2\ X_3\ X_4)$) and its projection on the mean road plane $X_{\bar{3}} = (X_1\quad X_2\quad X_3)$.
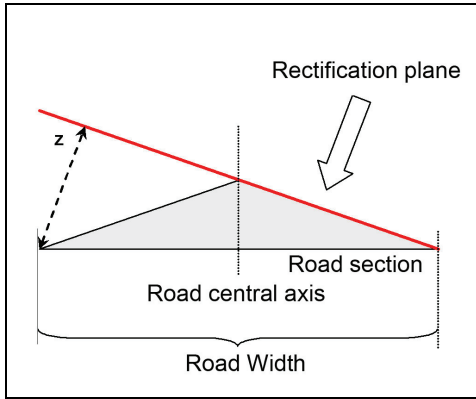

Figure 4. Road section reference rectification plane.

As far as the point X lies on the rectification plane, its mapping on the original image space by means of the homography H is correct. Otherwise, if the point presents out-of-plane components, its reprojection on the image plane is affected by an error strictly dependent on the third column of the camera matrix:

$$
\begin{aligned}
&if\ X_3 = 0 &\Rightarrow\quad& x = H \cdot X_{\bar{3}} \\
&if\ X_3 \neq 0 &\Rightarrow\quad& x = H \cdot X_{\bar{3}} + p_3 \cdot X_3
\end{aligned}
\qquad [3]
$$

In other words, if a point lies on the rectification plane, its mappings on the common rectification space by means of the homographies associated with each camera, basically are coincident; otherwise some sort of parallax occurs.

From the knowledge of such parallax the out-of-plane value of the point can be estimated: let ξ_1 and ξ_2 be the two homogeneous rectified image coordinates respectively on the first and second rectified image; let ξ_{1n} and ξ_{2n} be the corresponding normalized coordinates. Finally let (1) and (2) be a subscript associated with the first and second frame. From eqs. [3] it follows (proof is omitted):

$$
\begin{cases}
\xi_1 = H_{(1)}^{-1}\left(H_{(1)} \cdot X_{\bar{3}} + p_{(1)3} \cdot X_3\right) = X_{\bar{3}} + X_3\left(H_{(1)}^{-1} \cdot p_{(1)3}\right) \\
\xi_2 = H_{(2)}^{-1}\left(H_{(2)} \cdot X_{\bar{3}} + p_{(2)3} \cdot X_3\right) = X_{\bar{3}} + X_3\left(H_{(2)}^{-1} \cdot p_{(2)3}\right)
\end{cases}
\quad [4]
$$

$$
(\xi_{1N} - \xi_{2N}) + \frac{X_3}{X_4}\left(\left(H^{-1}\right)_{(1)}^{3} p_{(1)3}\xi_{1N} - \left(H^{-1}\right)_{(2)}^{3} p_{(2)3}\xi_{2N} + \left(H_{(2)}^{-1}p_{(2)3} - H_{(1)}^{-1}p_{(1)3}\right)\right) = 0 \quad [5]
$$

$$
\Rightarrow\quad (\xi_{2N} - \xi_{1N}) = \frac{X_3}{X_4}\left(A\xi_{1N} - B\xi_{2N} + C\right)
$$

$$
\begin{aligned}
where: \\
A &= \left(H^{-1}\right)_{(1)}^{3} \cdot p_{(1)3}, \\
B &= \left(H^{-1}\right)_{(2)}^{3} \cdot p_{(2)3}, \\
C &= \left(H_{(2)}^{-1} \cdot p_{(2)3} - H_{(1)}^{-1} \cdot p_{(1)3}\right)
\end{aligned}
\qquad [6]
$$

In eq. [6] $(H^{-1})^3$ is the third row of the inverse homography matrix, A, B are constant scalar values, while C is a constant vector, all depending on the camera matrices: note that their computation is required only at the start of the sequence.

From the first scalar equation in [6] it is possible to compute the out-of-plane values, which corresponds (under the assumptions) to the Z coordinate of the point: if $(C_1\ C_2\ C_3)$ are the components of the vector C, we have:

$$
\bar{Z} = \frac{X_3}{X_4} = \Delta Z = \frac{\left(\xi_{2N}^1 - \xi_{1N}^1\right)}{\left(A\xi_{1N}^1 - B\xi_{2N}^1 + C_1\right)}
\qquad [7]
$$

Then, substituting the Z value in the other two scalar equations the planimetric coordinates may be computed:

$$
\left|\begin{matrix}\bar{X} \\ \bar{Y}\end{matrix}\right| = \left|\begin{matrix}\xi_{1N}^1 \\ \xi_{1N}^1\end{matrix}\right| \cdot \left(1 + \bar{Z} \cdot A\right) - \bar{Z} \cdot H_{(1)}^{-1} \cdot p_{(1)3}
\qquad [8]
$$

If the lane markings are not straight (the road path is curved) or the trajectory of the vehicle is not parallel to them, the lines on the rectified images are not vertical: thus, to obtain accurate measurement, homologous point must be selected. Vice versa, if the road marks are vertical the parallax may be computed only by means of the x value of each line. It may be shown that epipolar lines in the rectified space are represented by parallel horizontal lines. Thus homologous points can be selected finding lane border position at the same y image value.

3.6 Curved road sections

Along curves, road markings are no more vertical nor straight in the rectified images. The Hough extraction routine is therefore less likely to correctly recognize the lane boundaries.

Anyway, as far as the road curvature radius is not too small compared to the area to be analysed (usually 4÷6 m), the arc portion to be considered will be still nearly vertical (see figure 5), so the filtering procedure will still get good results in the marking recognition.
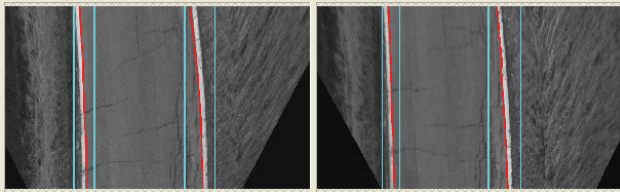

Figure 5. Curved path reconstruction: the circular boundaries are approximated by several straight edges.

The extraction of straight segments with the Hough transform is simply applied to partitioned horizontal areas of the rectified images, in order to substitute the curve by a polyline. Even with high curvatures, the method extracts segments which interpolate the lane marks. Thus, even if the boundaries of lane markings are not correctly extracted (the segments intersects), the measurement error in the lane width computation is basically negligible.

4. RESULTS

Currently the method has been tested, with success, on simulated data and on a small sample of real sequences. A computer generated sequence, where all geometric parameters (camera E.O., lane width, etc) are known and images are not affected by lens distortion or other systematic disturbances (e.g. synchronization errors in stereo imaging, etc). In all instances the algorithm provided accuracy of measurement in the order of 1÷2 cm, i.e. in the same order of the pixel ground resolution.

Figure 6. Computer generated sequences used to check the algorithm reliability.

To verify also its robustness in more realistic conditions, the marking texture was modified to have little contrast with that of the road surface (Figure 6, bottom); gaussian noise was also added to the image grey values. Two different textures were used for the road surface: in the first (Figure 6, top) the performance of the Hough transform in locating the marking edge despite the large amount of high-gradient points in the road surface; in the second (bottom) the performance has been examined under poor contrast and presence of shadows. In both cases results were satisfactory, although in the latter case, wherever the percentage of shadows was high, the width could not be measured because of poor accuracy of the estimated lane markings.

In a second series of tests, real sequences were used and the lane width was measured photogrammetrically using our own-developed program for mobile mapping (Varini et al, 2005); Accuracies, based on previous experiences, are in the range of 3÷4 cm. The comparison between both series of measurements was performed by a spatial analysis, finding sections where differences were relatively homogeneous; results are presented in Table 7 and refer to about 350 m of a countryside road.

The average difference, weighted with the length of the road section, turned out to be about 4 cm, i.e. close to the values of the manual restitution.

Section	Length (m)	Diff. (cm)	Section	Length (m)	Diff. (cm)
1	44	9	17	44	4
2	43	8	18	9	1
3	30	4	19	6	1
4	4	1	21	10	1
5	23	4	22	9	2
6	20	2	23	12	2
7	19	3	25	21	4
8	19	2	26	21	4
9	13	3	27	31	2
13	113	4	28	40	2
14	24	3	29	7	4
15	31	4			
Weight. Ave. (cm)	4.0				
MAX (cm)	8.8				

Table 7. Absolute values of the differences between manual photogrammetric measurement and automatic measurement of lane width.

The width measured by the algorithm is, in most cases, smaller than that of the manual measurements; the reason is that, currently, when resampling the image in the rectification, radial distortion is not taken into account. Since the cameras we use exhibit a strong barrel distortion, the road image is narrower than it should be. Besides, radial distortion in principle affects also the straightness of the markings in truly straight line road sections; however the effect on the accuracy of the extraction with the Hough transform and the prediction on the next images are not too affected.

Figure 8 shows the differences in color scale (left) and the changes of lane width larger than 10 cm (right); in the 380 m processed, thirteen changes occurred.

To evaluate the robustness of the method and the processing time, a sequence of about 1.7 km with 500 image pairs has been processed, along a countryside road; only in 23 pairs (about 4.6% of the total) the algorithm could not perform the

measurement, mainly because of strong shadows (see figure 9). On a P4 machine with 1.3 MB images, the processing time is about 1 s per image pair.
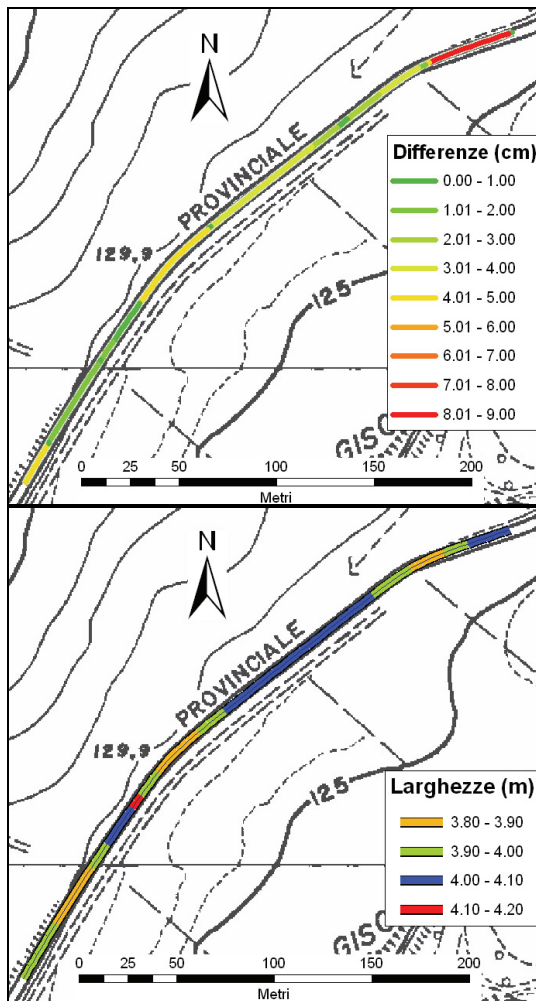

Figure 8. Automatic road width extraction results: (top) differences between automatic and manual maesurements; (bottom) road width variation along the experimental sequence (classes of 10 cm each).

5. SUMMARY, CONCLUSIONS AND PERSPECTIVES

A semi automatic method for the extraction of lane markings and the automatic measurement of lane and road width has been presented. All the informations acquired (lane width and corresponding sets of points on the border) are sent to a GIS system in which further visualization and elaboration of the data can take place.

From a metrological standpoint the results obtained so far on simulated and real data show an accuracy potential within the specifications of the road inventory technical prescriptions. Good agreement was also obtained with the manual restitution. Accuracy improvements are expected by image distortion removal and sub-pixel identification of the road marks; the relevant software has already been developed but not yet integrated in the program.

When such integration will be performed new, more detailed tests on the algorithm performance will take place. In particular much more experience is required to ensure that the proposed method is really efficient in terms of automation: when the software run into ambiguous situation (i.e. road crossings,

disappearance of one or both the lane marks, changes in the number of lanes, etc.) human interaction is still required.
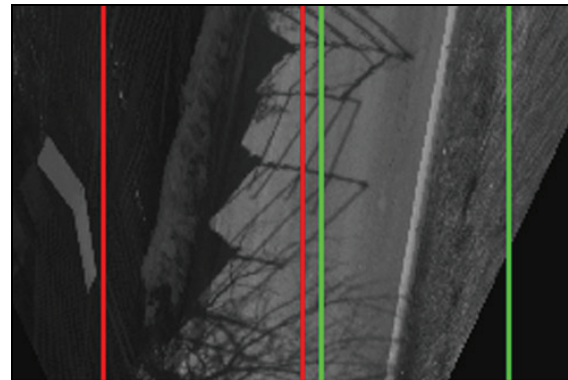

Figure 9. Failure of the extraction: in this frame the left lane marking is not recognised due to strong shadows.

REFERENCES

Schwarz, K.-P., El-Sheimy N. (1996). Kinematic Multi-Sensor Systems for Close Range Digital Imaging. International Archives of Photogrammetry and Remote Sensing, Vol. XXXI, Part B3, Vienna, pp.774-785.

Broggi A., Bertozzi M., Fascioli A., Argo and the Mille Miglia in Automatico, IEEE Intelligent Systems, pp. 55-64, Jan-Feb 1999.

Hartley, R., Zisserman, A., 2000. *Multiple View Geometry in computer vision*. Cambridge University Press, Cambridge, pp. 1-496.

Broggi A.,1995. Robust Real-Time Lane and Road Detection in Critical Shadow Conditions. In Procs. IEEE Intl. Symposium on Computer Vision, pp. 353-358, Coral Gables, Florida, IEEE Computer Society.

Hough P.V.C., Machine Analysis of Bubble Chamber Pictures, International Conference on High Energy Accelerators and Instrumentation, CERN, 1959.

A. Varini, G. Forlani, R. Roncella (2004). Test di precisione di un veicolo rilevatore. Bollettino della SIFET, n.1/2005, pp. 83-99.