AUTOMATIC TRAFFIC SIGNS INVENTORY USING A MOBILE MAPPING SYSTEM

For GIS applications

Sérgio R. MADEIRA, Luísa C. BASTOS, António M. SOUSA, João F. SOBRAL and Luís P. SANTOS

A Mobile Mapping System (MMS) designed to support quick surveys of traffic signs, including automatic sign recognition and exact location, is being developed in a partnership between Geonav, a private enterprise, and the University of Minho. The system acquisition is based on a road survey in a vehicle carrying digital video cameras, and a navigation system based in the integration of DGPS with an Inertial Measurement Unit and other dead reckoning systems. The acquisition system allows the association of a rigorous position and attitude to each digital frame. Furthermore, the pixel locations of every traffic sign visible in the video can be identified automatically through dedicated software, and their absolute geographical coordinates can be extracted. During the acquisition and processing stages, the relevant data is archived in a data base compatible with most CAD/SIG applications. Results of the tests carried out with this system show its effectiveness, with a good percentage of the traffic signs being automatically recognized and geo-referenced.

KEYWORDS

Mobile Mapping System, Analytical Photogrammetry, Inertial Navigation System, Computer Vision, Features Autonomous Detection and Identification, Traffic sign detection, Contour signatures, HSI color space, Correlation.

INTRODUCTION

The concept of Mobile Mapping Systems has seen a significant growth in the early nineties with the outcoming of direct orientation technologies. In this kind of systems position and orientation techniques are used to directly measure the positions and orientation of sensors when in movement either in an aerial or terrestrial platform. The way to achieve such objectives is greatly depending on a proper integration of GPS and Inertial measurements, which may also, mainly in terrestrial applications, be complemented with other sensors such as odometers, digital compasses etc. which may improve the effectiveness of the system.

In the past decade several groups worked in the development of Mobile Mapping Systems. An overview of the main achievements can be found for example in Grejner-Brzezinska, 2001 [16] or El-Sheimy, 1996 [14].

The work presented here is mainly the result of a partnership between a private enterprise working in the mobile mapping area, Geonav, and the University of Minho. The main goal is to exploit the potential of the MMS approach and combine it with image correlation techniques to automatically extract information from digital images. In order to achieve sub-meter positioning accuracy in urban environments it is necessary to use differential GPS data, integrated with data from an Inertial Measurement Unit (IMU). The IMU data allows the system to maintain the precision standards where GPS fails due to signal obstruction. The video camera being used has a 1 mega pixel colour CCD and a frame rate up to 10 frames per second.

Field surveys are done by Geonav, using a modified van carrying a platform where a digital video camera, a GPS and an IMU are mounted and with relative positions very well determined. An odometer was also installed in order to add additional constrains to the navigation solution and to optimize image acquisition rate in a dynamic way.

A crucial requirement is the perfect time synchronization between the instant of frames acquisition and GPS time. This is guaranteed by a proper design of the system hardware and relies on the PPS output from the GPS receiver.

For image manipulation and processing, an algorithm was developed by the University of Minho (Informatics Department). The software allows the automatic extraction of the pixel coordinates of the traffic signs visible in the images. The algorithm is divided in two steps: detection and identification. Detection is achieved by performing image segmentation based on colour information (blue and red regions) in HSI colour space, and shape, through contour signatures.

The use of HSI colour space makes easier the segmentation of images into regions with similar colour, since it allows a better tolerance to changes in lighting conditions compared to other colour models. The candidate traffic sign regions are first selected by analyzing H (Hue) and Chromatic information.

Shape information is then obtained by using contour signatures and compared with predefined shapes after normalization to a specific dimension. A shape classifier validates these regions as being similar to the predefined ones or not. Contour signatures are an effective way to obtain shape information and to compare it, since they have low processing requirements.

The second step, concerning traffic signs identification, is performed using a database of target objects images, structured according to their shape. The approach used is based on grayscale normalized correlation between database images and the possible traffic signs regions. Finally, the developed algorithm returns the traffic signs positions and identification in each image.

After the traffic sign has been identified by the software it extracts every pixel coordinate in the frames where the sign is visible. This information is then used together with the coordinates and attitude of the frames, to estimate the sign location using a photogrammetric triangulation. A least square solution is applied in order to take advantage of the redundant information.

The software also allows the automatic filling of a data base where the surveyed objects are stored with their coordinate and other relevant attributes, in a way that allows an easy export to CAD or CAD/SIG platforms.

In this article we start with a general view about the surveying procedures and other necessary previous calibration steps, next we approach, in some detail, the auto detection of traffic signs events in the surveyed images and take, as last developed aspect, the subject of getting absolute coordinates of the detected objects.

OVERALL DESCRIPTION OF SURVEYING PROCEDURES

There are two kinds of sensors present. In first, the direct geo-referencing sensors, which allow knowing the exterior orientation parameters of the system in different surveying instants. The referred orientation parameters are composed by 3 cartesian or geographic absolute coordinates and 3 rotation angles related to the main axes of the absolute reference system. Here we use a dual frequency GPS receiver with respective antenna, a Litton LN 200 Inertial measurement unit (IMU) and a car odometer. Secondly, the object sensors. These are 2 CCD digital video cameras that acquire video files directly to a laptop computer. The system can be used with one camera only, but with a decrease in robustness related to some measurement aspects.

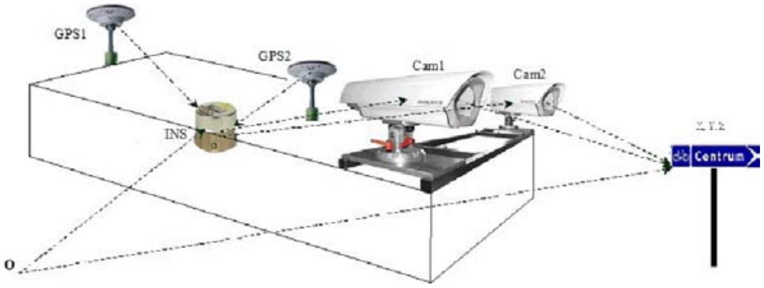

Figure 1: Surveying components and configuration.

The system configuration when in surveying situation is represented in figure 1. To obtain the instantaneous position and orientation of the mobile vehicle is used an integration of the data from the GPS and the Inertial Navigation System (INS). In present case, the integration is achieved through a Kalman filter [17]. A third input to the system is given by a car odometer that reinforces the direct geo-referencing quality. In parallel, a second GPS receiver is allowing one or two digital CCD video cameras to acquire video frames synchronized with GPS time.

It is of note that, mainly in urban areas, the GPS system can not achieve constant rates in the precision and interval of acquired positions due to several signal obstructions. Furthermore, the IMU data (actual attitude and velocity) slowly drift, keeping accuracy values only for a short amount of time. But when GPS data and Inertial Navigation System data are mutually integrated, the geo-referencing system becomes much more effective and with the advantage of enhanced acquisition rate, much beyond the 1 second possible with most GPS receivers. A critical aspect of this integration happens in car stopping periods, when the direct geo-referencing system can drift away by a large amount. So, a third set of integrated geo-dynamic data, given by the car odometer, is of great deal mainly in stop and go situations like in urban areas.

The interaction between the components of the direct geo-referencing system and the image acquisition system is showed in the next diagram.

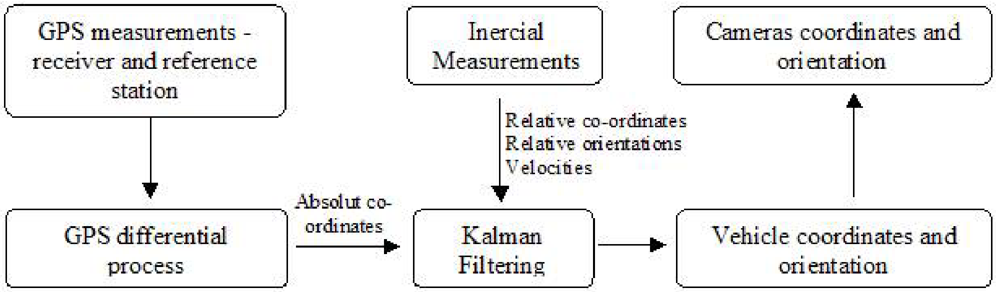

Figure 2: Simplified diagram of geo-referencing acquisition.

The relative positions of the system components remain unchanged during the surveying procedure, and that parameters are settled accurately in a previous step. So, it is defined a vehicle reference system, that is forced to coincide with the INS reference system and all the parameters are measured relative to this. Of major importance is the determination of camera orientations related to the vehicle reference system and the relative orientation between the cameras themselves, in the case of two used. These are determined resorting to photogrammetric techniques lying in the collinearity condition [13].

PREVIOUS CALIBRATION STEPS

The image acquisition system demands for some previous work namely calibration routines.

The characteristics and behavior of the video sensors is a crucial subject in the present work. Besides the necessary requirement of few demands in the process of acquire and storing video files in a laptop computer, what can be done either through a dedicated frame grabber or a firewire protocol, are the overall solidity of the lenses and the possibility of iris and focal length fixing. The lens system itself must keep low the inherent distortion factors. To achieve this last requirement it is necessary to carry out a self-calibration technique [15] to determine the calibration parameters of the video camera lenses. These are the principal point coordinates, focal length, three radial deformation parameters and a scale y coordinate factor. The first step is to obtain several images of an object with well defined points known in a true-size object reference system. These can lie in a plane like in present case, see figure 3. The used pattern was retired of a well known photogrammetric software, the Photomodeler, but the software itself wasn't used.

Figure 3: Image acquisition to obtain calibration parameters of the video camera lenses.

The process is performed independently for each camera and the routines used, to obtain the photo coordinates and to perform the self-calibration itself, based in the collinearity condition [13], are self-made. The initial data is composed by 63 object point coordinates and 63x10 image point coordinates lying in 10 images obtained from different perspectives. There are no required initial approximations, except for the focal distance, and the final calibration parameters are obtained from an iterative process with bundle adjustment.

The determination of the position and mainly the orientation of the cameras relatively to the vehicle system are a critical step in the process because the lenses are not visible outside and the orientations have to be determined with significant precision. To overcome this situation is performed a preliminary step before each survey. The car must be prepared with all the components in their final locations, parked near a wall were some points were measured in the absolute reference system with teodolits, offering confidence at a millimetre level. See figure 4. The cameras must then obtain images with sufficient coordinated points in view. In a photogrammetric process known as space resection [13], the positions and attitudes, i.e., the angular position in the reference system, of the camera(s) are determined and complemented with metric and GPS measurements, with the antennas in the vehicle, to finally obtain the orientation of the camera(s) in the vehicle reference system.

Figure 4: Façade with known absolute coordinates used in orientation procedure.

Figure 5: Preparing for a survey.

AUTOMATIC TRAFFIC SIGNS RECOGNITION

The methodology for automatic localization of traffic signs in images is divided in three steps:

1. **region color detection and segmentation**: Image regions are selected based on their color (blue and/or red), size and localization;

2. **shape detection**: Regions from the previous step are classified into one of a set of system pre-defined shapes - circular, triangular, quadrangular and inverted triangular;

3. **recognition**: For each classified shape in the previous step, the trafic sigh recognition is accomplished by comparing it with pre-defined objects in database, where objects are organized according to their shape.

Automatic traffic signs localization and classification does not require any user assistance. Since the developed system deals with traffic signs detection, a sign database was analyzed and concluded that their color and shape are the most distinguishing features. Traffic signs are organized, in the system, by their most common shapes and colors (Figure 6). The detection problem is thus divided in two parts: color detection – internal feature, and shape classification – external feature.

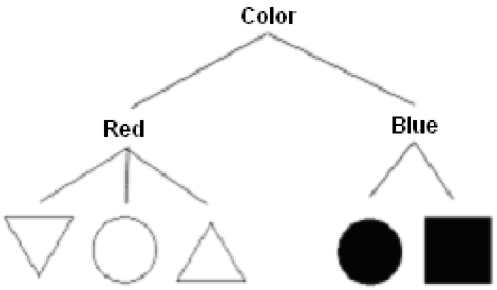

Figure 6: Detection tree

The next three subsections presents in more detail each of these steps for automatic localization of traffic signs.

REGION COLOR DETECTION AND SEGMENTATION

The RGB (R= Red, G= Green, B= Blue) color space is not a good choice for segmentation purposes when dealing with outdoor environment images. On one hand, due to environment reasons, there is no control over the local illumination levels. On the other hand sign colors degenerate due to aging. One other factor for this degeneration is the intrinsic features of RGB components. These components have a strong correlation with each others, rendering the use of RGB color space inadequate in the presence of environment light intensity variations.

Although there are several alternative color spaces, such as Nrgb and L*a*b, this methodology is based on the HSI (Hue, Saturation, Intensity) color space for detection/classification. This colour space represents different information parts in each component [2], [6], [11], and is similar to the way humans perceive colours. Even more important, light conditions information is represented mainly in the I component, which is thus decorrelated with the H component [2]. There are several algorithms proposed to transform RGB space into HSI [8], [4], [12], [7].

Segmentation is specially prone to errors when image regions external to signs are connected to them: whiter (e.g., sky) or darker (e.g., road), being the most usual cases. To overcome this problem the RGB image chromatic component is used instead of Saturation on the segmentation process. The Saturation value quantifies the pixel colour distance to white/grey/black (color purity, which depends on intensity (I)), but it is very sensitive to noise for low values (black); the RGB image chromatic component measures the difference to the equivalent grey level, and is not as sensitive to noise. Figure 7 compares these two components. The chromatic component is calculated as in [5]. Figure 8 presents results after the initial segmentation is performed.

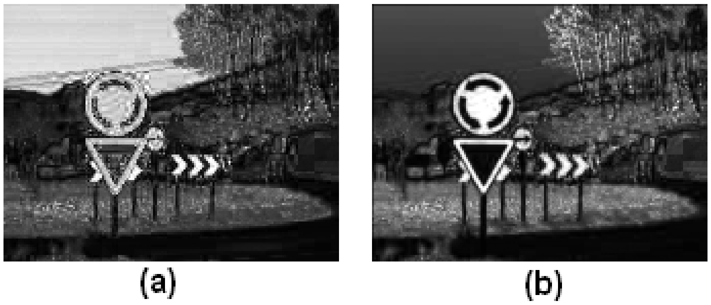

Figure 7: Component (a) saturation (b) chromatic

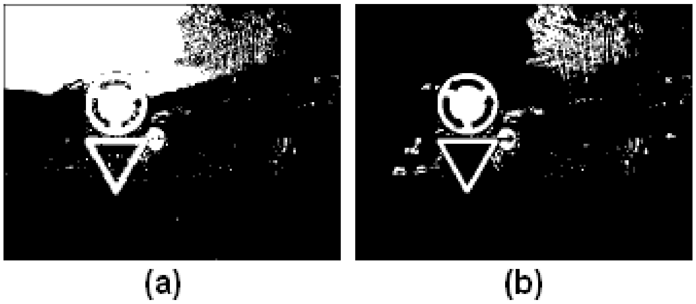

Figure 8: Segmentation (a) Hue/Saturation (b) Hue/Chromatic

After initial image segmentation each image region is identified using a labeling algorithm. Filtering is divided in four steps and consists of region separation and/or elimination of regions that do not have determined requisites (Figure 9.b)):

Figure 9: Example image (a) original and (b) filtering illustration

Step1: Regions connected to image sides are eliminated

Step2: Since the system deals with traffic signs and since we verified sometimes the presence in motorways of two signs in the same support, to avoid possible region unions related to those signs, their division is made as long as the ration Height/Width is higher than some threshold.

Step3: Regions without a minimum area and a minimum Height and Width are eliminated.

Step4: Regions totally inside other regions are eliminated, thus avoiding duplicate references to the same region.

Figure 10 presents the filtering result applied to the previous image.

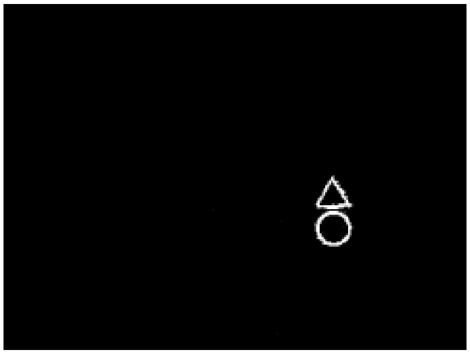

Figure 10: Filtering result

In order to facilitate comparison, recognition and shape classification regions are resized (normalised) in images with a pre-defined height and width (obtained experimentally 64x64 pixels), maintaining their ratio height/width.

SHAPE DETECTION

The developed system is able to detect and recognise 4 traffic sign pre-defined shapes: circular, triangular, quadrangular and inverted triangular. Shapes validation and classification is performed with classifier based contour signatures. This representation is obtained through a 1-D function that indicates the distance from the object's center to its contour as function of $r(\theta)$ angle.

Contour signatures are independent of the object position, but they are rotation and scale dependent. Taking into account that the system analyses objects that are rigid and normally positioned at approximately 0° related to y axes (traffic signs) the system assumes that rotation is not significant and does not affect this methodology. Scale independence is based on the normalisation of $r(\theta)$ values, i.e., resizing detected regions to a pre-defined height and width value.

The system doesn't use all the 360 values obtainable from contour signature but only a small subset. This option allows the analyzed number of points to decrease, while maintaining enough shape information. In the used examples, for each pre-defined shape was obtained a normalized image to 64 x 64 pixels, corresponding to n =90 points of contour signature and a sample step $\Delta\theta$ = 4°. The obtained values were rounded to closest integer.

As an example the main features of triangular contour signature are shown in Figure 11:

- three points with object center maximum distance, standing out from all the others, corresponding to triangle corners (Figure 12);
- three points with object center minimum distance corresponding to middle points between triangle corners.

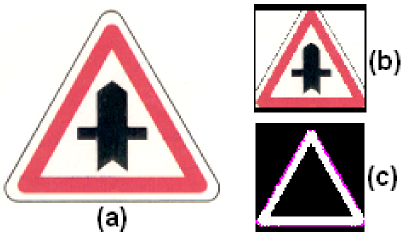

Figure 11: Triangular contour detection (a) model sign, (b) normalised sign and (c) contoured sign

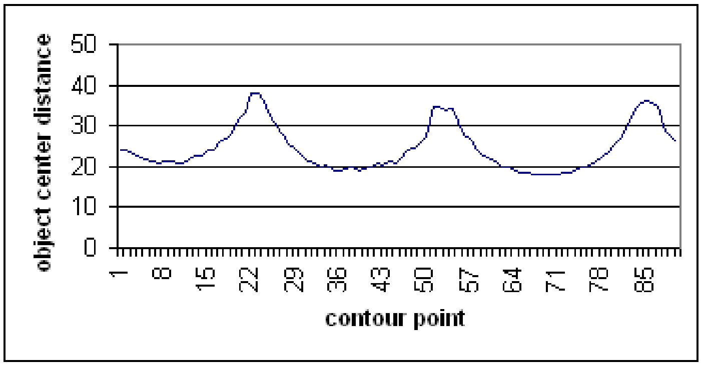

Figure 12: Triangular contour signature

The contour signature is determined for each region obtained from the color detection process, using the already mentioned process. Following this, each signature is compared with the pre-defined shapes signatures and decided if the analysed region is valid (i.e. pre-defined shaped region) or not (i.e. non pre-defined shaped region/noise). This procedure is accomplished calculating for each region signature the average deviation (σ) between it and each of pre-defined signatures.

$$\sigma = \frac{1}{n}\sum_{i=1}^{n}\left\| x_i - y_i \right| - \mu\right|$$

Where,

$\mu = \frac{1}{n}\sum_{i=1}^{n}\left| x_i - y_i \right|$: the arithmetic average of difference between signatures point to point absolute values;

n : signature number of points;

x_i : analysed region signature value in i point;

y_i : pre-defined shape signature value in i point;

Being four the number of pre-defined shapes, there are four average deviations for each region, i.e., $\sigma_{region} = \left\{ \sigma_{circular}, \sigma_{triangular}, \sigma_{quadrangular}, \sigma_{inv\_triang} \right\}$. From these the minimum value is obtained, $\sigma_{\min} = \min(\sigma_{region})$ and associated to region the correspondent shape. Finally it is verified if $\sigma_{\min} \leq k$ constant, where k value, obtained experimentally, represents the maximum average deviation allowed by the system so that the analysed region is considered sufficiently close to the associated shape.

RECOGNITION

There are recurrent difficulties in the recognition process that must be considered, added to others when dealing with traffic signs recognition:

- large number of signs;
- closely/largely different ideograms of same sign [10];
- total/partial sign occlusion;
- ilumination variations;
- scale differences.

The two first problems are the most difficult to solve. In several papers the used sign models are a very restrictive number of all the traffic signs [1], many of which use neural networks. If this kind of system is used in a traffic sign surveying or even used in several countries the neural network training database would be huge.

Taking into account all these considerations the developed system uses for recognition the greyscale normalised correlation, which has been frequently used through the years [5], [9], [3].

$$C = \frac{\sum_{x,y}(I-\overline{I})(T-\overline{T})}{\sqrt{\sum_{x,y}(I-\overline{I})^2\sum_{x,y}(T-\overline{T})^2}}, \text{ where } \overline{I} = \sum \frac{I}{N} \text{ e } \overline{T} = \sum \frac{T}{N}$$

T and I are the model sign and analysed region information, respectively, and N the number of image pixels.

The signs database is classified by their shape, because the detected signs are divided the same way. There are four sign classes: circular, triangular, quadrangular and inverted triangular as mentioned before. This division goal is to lower the database candidates to use for correlation, therefore reducing recognition time.

To avoid correlation errors due to small target sign translations compared with model sign ± 2 target sign image pixels are displaced in x and y axis, being the considered correlation value, for each candidate, the 25 correlation maximum result. Finally the selected candidate is the one obtaining higher correlation value.

RESULTS

The automatic traffic sign localization, obtained good results, although the tests were restricted to four shapes and two colors. The methodology was submitted, to outdoor environment scenes, more specifically 172 images, from which 238 localizable traffic signs were spotted. In this test 90.3% color and shape were correctly detected and from these 82.8% were correctly recognized, although the algorithm used in this recognition phase is only an initial approach.

OBTAINING ABSOLUTE COORDINATES

As a result of the automatic recognition process follows a identification file that contains, for each sign, the frame numbers and image positions of the frames where it occurs, a specific identification of the traffic sign and other attributes. The next task is to add the respective absolute cartesian or geographic coordinates in order to transform the reported file in a geo-referenced database easy to export to CAD or CAD/SIG platforms.

The first step is creating a file that associate a position and attitude to each video frame related to the absolute reference system - the synchronization file. To accomplish this task is necessary previously interpolate the positions and attitudes of the direct geo-referencing system to GPS times equal to those of frames. Then the linear and angular offsets determined in the calibration step are used to obtain the final position/attitude of the frames.

To obtain the absolute coordinates of the objects find by the auto-detection method, again is used the collinearity condition. Since the image positions of one object were identified in at least 2 frames, a set of 4 collinearity equations can be established to obtain the 3 absolute coordinates of the given object. In the simplest case, when only 2 frames are employed, the calculus is analogous to an intersection in 3 dimensions – see figure 6. In this case, due to observation and instrumental errors, the vectors do not intersect, like in the 2D case, and the solution is encountered with a least squares approach. However, in the present work, several images of the same object (traffic sign) can be obtained and any new image position add up 2 new system equations and no new variables. The used approach is to get a large set of well defined photo coordinates of the object and apply the least squares method to the redundant equation system.

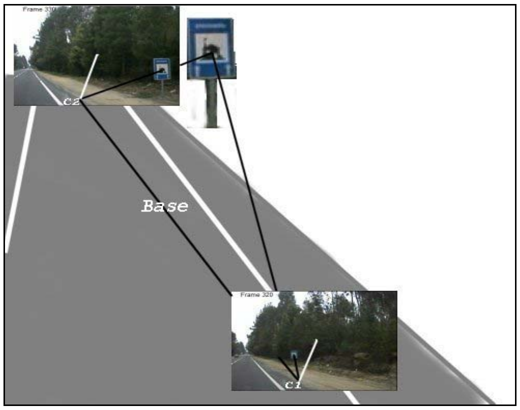

Figure 6: 3D intersection.

A mean of 15 to 20 different frames of an object can be used to a get the respective absolute position. The obvious advantages is adding strength to the computed values and, also, a better estimation of the associated errors, trough the standard deviations of the residuals.

An important subject is the expected accuracy of the computed absolute coordinates. Two factors are in consideration: the linear and angular accuracy given by the system. The direct geo-reference system is usually in operation by Geonav where is used in several field works of diverse nature. The established accuracy points to a linear standard deviation of about 1 meter and sub-minute standard deviation in attitude. Furthermore, it is present an angular error associated to pixel selection that depends on pixel dimension and focal distance. Obviously this error grows with pixel dimension and decrease with focal distance. Considering the mean parameters of the cameras in use we can set this error to a maximum of 2 angular minutes. Another factor is the overall geometry of the intersections. It is known that the quality of an intersection decrease when it runs away from the right angle. The nature of this work means that the usual intersection figure is somewhat different to that ideal, but the great redundancy of the used equations system will allow to attenuate this effect. So, we can conclude that the basic linear error of the direct geo-referencing system will be transported to the final absolute coordinates and another step of degradation will occur due to pixel selection and, mainly, due to a far ideal geometry of the intersection figure.

CONCLUSION

The overall procedures and preparing tasks for directly acquire geo-referenced data, video frames and synchronize them with GPS time are established and performing well. The system configuration allows transferring absolute positions and attitudes to each captured video frame. The GPS/INS integration enables, in most of time, a constant frequency acquisition, and keeps as low as possible the inherent effects of GPS errors and gaps – mainly in urban areas, and IMU drifts.

A significant stage of the proposed work is to automatically detect the occurrence of traffic signs events in the video files. The developed methodology localizes traffic signs based on their color and shape information. The limitation to traffic signs localization is only intended to assure results validation. The approach flexibility allows the inclusion of any type of shape and color information. This fact can have as inconvenient the system performance decrease concerning analysis speed as the recognition objects quantity grows.

The traffic sign localization is only complete if sign position information has associated their identification (recognition). A database was created, with the localisable signs normalised RGB images, separated by shape classes and compared with target sign. However, system performance, in terms of execution time per frame, may decrease with a significant increase of the number of objects and different shapes to detect and identify.

Finally, all the data is processed to obtain the absolute cartesian or geographic coordinates of the identified traffic signs, using photogrammetric triangulation processes, and a general geo-referenced database is constituted.

ACKNOWLEDGMENTS

This work, was partially supported by project "SIGMA – Sistema de Geo-referenciação Móvel assistido por Imagem", funded by "AdI – Agência de Inovação", POSI.

A special acknowledgment to a colleague, Eng. Sérgio Rui Silva, of the Electronics department of Porto University, who was of great help in some technical aspects concerning the use of complex equipment during the field surveys.

REFERENCES

1. Aoyagi, Y. and Asakura, T. A Study on Traffic Sign Recognition in Scene Image Using Genetic Algorithms and Neural Networks. Proc. 22nd IEEE Int. Conf. Industrial Electronics, Control, and Instrumentation, Taipeh, Taiwan, August 1996.

2. Arnoul, P., Viala, M., Guerin, J. and Mergy, M. Traffic signs localisation for highways inventory from a video camera on board a moving collection van. Proc. '96 IEEE Intelligent Vehicles Symp., Tokyo, Japan, September 1996.

3. Betke, M. and Makris, N. Fast Object Recognition in Noisy Images Using Simulated Annealing. Proc. 5th IEEE Int. Conf. Computer Vision, Cambridge, USA, June 1995.

4. Carron, T. and Lambert, P. Color edge detector using jointly hue, saturation and intensity. Proc. 1st IEEE Int. Conf. on Image Processing, Austin, USA, November 1994.

5. Escalera, A., Armingol, J., Pastor, J. and Rodríguez, F. Visual Sign Information Extraction and Identification by Deformable Models for Intelligent Vehicles. IEEE Transactions on Intelligent

Transportation Systems, vol. 5, no. 2, pp. 57-68, June 2004.

6. Hibi, T. Vision based extraction and recognition of road sign region from natural color image, by using HSL and coordinates transformation. Proc. 29th Int. Symp. Automotive Technology and Automation, Florence, Italy, June 1996.

7. Ledley, R., Buas, M. and Golab, T. Fundamentals of True-Color Image Processing. Proc. 10th IEEE Int. Conf. on Pattern Recognition, Atlantic City, USA, June 1990.

8. Logicol S.r.l. Software EasyRGB-PC. RGB-HSL color conversion algorithm, http://www.easyrgb.com/math.html, 2004.

9. Miura, J., Kanda, T., Nakatani, S. And Shirai, Y. An active vision system for on-line traffic sign recognition. IEICE Trans. Inform. Syst., vol. E85-D, no. 11, pp. 1784-1792, 2002.

10. Paclik, P., Novovicova, J., Pudil, P. and Somol, P. Road Sign Classification using the Laplace Kernel Classifier. Proc. 11th Scandinavian Conf. on Image Analysis, Kangerlussuaq, Greenland, 1999.

11. Piccioli, G., Micheli, E., Parodia, P. and Campani, M.. Robust method for road sign detection and recognition. Image Vis. Comput., vol.14, no. 3, pp. 209-223, April 1996.

12. Welch, E., Moorhead, R. and Owens, J. Image processing using the HSI color space. Proc. '91 IEEE Southeastcon, Williamsburg, USA, April 1991.

13. Wolf, P. R. and Dewitt, B. A., Elements of photogrammetry, with applications in GIS. Ch. XI, pp. 233-259, McGraw-Hill, 3rd Edition, 2000.

14. El-Sheimy, N., The development of VISAT – A Mobile Survey System for GIS Applications. PhD Thesis, Calgary University, Canada, 2000.

15. Fraser, C., Digital camera self calibration, ISPRS Journal of Photogrammetry & Remote Sensing, nº2, 1997, pp 149-159.

16. Grejner-Brzezinska, A., Direct Sensor Orientation in Airborne and Land-based Mapping Applications, Report nº 461, Department of Civil and Environmental Engineering and Geodetic Science, Universidade de Ohio, 2001

17. Tomé, P., T. Cunha, S. Cunha and Bastos, L., Evaluation of a DGPS/IMU Integrated Navigation System. ION GPS 2000, Salt Lake City, Utah, USA, 2000.

AUTHORS INFORMATION

Sérgio M. MADEIRA
smadeira@utad.pt
UTAD.

Luísa C. BASTOS
lcbastos@geonav.pt
Geonav

António M. SOUSA
amrs@mail.pt

João F. SOBRAL
jls@di.uminho.pt
Universidade do Minho

Luís P. SANTOS
psantos@di.uminho.pt
Universidade do Minho