A MOBILE SYSTEM FOR VISION BASED ROAD SIGN INVENTORY

Wanda Benesova, Yuriy Lypetskyy, Jean-Philippe Andreu, Lucas Paletta [a,1],
Andreas Jeitler, Evelyn Hödl [b,1]

[a] JOANNEUM RESEARCH Forschungsgesellschaft mbH, Institute of Digital Image Processing, Graz, Austria
(lucas.paletta@joanneum.at)

[b] University of The Applied Sciences FH Joanneum, 8020 Graz, Austria

KEY WORDS: mobile geo-information systems, road sign inventory, visual object recognition

ABSTRACT:

Inventory of road infrastructure represents a key application for integrated mobile mapping systems. The classical approach is to post-process geo-referenced imagery that has been captured from mobile mapping vans while driving within the road. Problems arise then within city centres where dense traffic and parking vehicles may often hinder occlusion free image captures of the objects of interest. We present an alternative solution in terms of a mobile PDA based geo-information system for pedestrian based road sign data collection which is targeted towards complementing the classical approach, particularly in city centres. The novel system consists of a mobile computing unit which is equipped with a GPS sensor, a digital camera, a standard mobile GIS software, and a user interface for minimal user intervention. Imagery for road sign inventory is firstly captured and geo-referenced in the street and then analyzed in a second, off-line processing step. We present a post-processing tool for semi-automated classification of road signs, and in parallel an identification of the sub-plate information, strongly supporting the processing of huge data sets, e.g., for the processing of data from complete city areas. The tool includes advanced image analysis methodology with (i) localization of selective key descriptors and regions of interest, (ii) the fitting of geometrical constraints to the extracted set of interest descriptors, and (iii) the matching of content information from the visual information within the sign plate. Text information on the sub-plate is first detected using a wavelet based texture filter, and then post-processed with segmentation operators and standard OCR software. The system has been successfully applied in a project on a complete inventory within the City of Graz, Austria, processing more than 30.000 road signs. The system achieved high accuracy ($\approx$90%) in the automated image based classification of road signs on more than 20 different road sign classes.

1. INTRODUCTION

Mobile mapping of traffic infrastructure is becoming an increasingly important topic, also in the context of municipality based administration of road signs. Within the scope of the research project 'road sign based city area land register' an efficient method was developed for the acquisition, annotation and position detection of city owned road sign data. The system is based on the components ArcPad (mobile G SHARP-software of the company ESRI) carried out on a handheld device with touch-screen, GPS positioning and an image recognition procedure. The project goal was the creation of a mobile system for the detection and administration of road signs consisting of a Windows CE device, GPS receiver and a digital camera as well as adapted software for a simple integration of the received data into the ArcGIS environment, to be operated by a mobile user in the field. For the geodetic referencing of the digital photos, a time synchronization procedure was used. The classification of the road signs was carried out using an image recognition software developed at the Institute of Digital Image Processing. Using this package, it was possible to catalogue nearly 30,000 road signs in the Grazer city zone within only 12 days.

The mounting of a road sign attends to the responsible street upholder. The legal base for this purpose is given by the streets traffic order. In the sense of an improved administration, the procedure described in this work was developed on behalf of the office of the Graz municipality that is responsible for road infrastructure monitoring and maintenance in cooperation with the city measurement office.

The digitally available data of the land register of the Graz city zone (Gauß-Krüger coordinates, M34) were used as the based input data. These included land boundaries, building edges, fences, streets and sidewalks, as well as the position coordinates of stayers, mast and traffic lights at which road signs can be mounted on. However, the semantic information about the road signs themselves was missing in the Geo-database. An acquisition timeout of two weeks (because of the dynamic of the changes) as well as the use of a cost-effective hardware was needed. A user friendly system for the data acquisition was created by means of a handheld device (Bluetooth capable Pocket PC for the connection to one also Bluetooth capable GPS-receiver) and the mobile GIS-software solution ArcPAD.

For the evaluation of the acquired data, an automatic post-processing procedure was developed. The geodetic referencing of the digital images is based on time synchronization in order to assign to every image its corresponding location of capture. In a subsequent processing step, an image recognition procedure, the classification and corresponding labelling of the road signs was performed, and the corresponding imagery, together with eventually available sub-plate information was geo-referenced. The obtained information can be integrated for further use into the Geo-database.

2. DATA ACQUISITION

The graphic representation of the map on the handheld device serves for the orientation whereby the corresponding section of the map becomes the GPS position. Furthermore, the
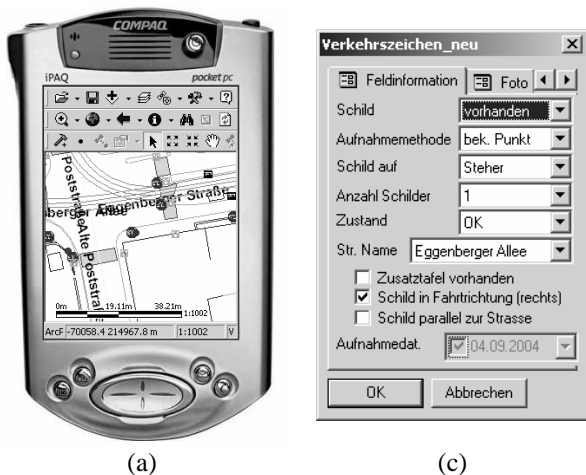
(a)   (c)

Figure 1. (a) ArcPad user interface of Graz map with overlaid road sign locations. (b) Input mask for attributes associated with road sign data collection.

handheld device is responsible for the task of data saving. The position of the road sign on the digital card could be saved by the user either manually or automatically via GPS, together with the basis information input, such as, mounting type and condition of the road sign, and visibility situation due to the current appearance of the street scenery. Subsequently a digital photo of the road sign was taken. The coordinates received by the GPS as well as the system-time were recorded simultaneously with the handheld device. The inaccuracy of the GPS position was corrected using a 'snap to existing point' methodology in the case of road signs with a known position – in this case the relevant point was caught automatically. In addition, the calculation of the trans-formation of the even Gauß-Bartender-coordinates into the ellipsoid length and width on the Bessel Ellipsoid were carried out. These calculations served as a preparation for the Geodetic referencing. The acquired information was stored in one record together with the date and time stamp and parameters of the GPS signal capture.

The software control and the data acquisition of the mobile GIS solution was enhanced by some VBSCRIPT functions and/or catalogue supported input mask forms (with friendly support through the ESRI user community and with special thanks to Craig Greenwald, Bradshaw Consulting service). The visualisation part of the map was also optimized for better orientation of the user.

After initial testing, the complete area of the Graz city zone has been operated with the mobile device by 20 students of the University of the Applied Sciences FH Joanneum, within only 12 days of actual work. The largest part of the city was processed by pedestrians, only for the periphery of the city, where road sign density deteriorated, the data collectors used a motor vehicle. For this purpose, some advanced logistic was needed to achieve a uniform work distribution and a continuous data processing.

3. MOBILE GEO-REFERENCING OF ROAD SIGN IMAGERY

For the point-labelling of the corresponding road sign imagery, the EXIF data within the image code was used where the date and time of the image captured was stored. This time stamp of the capture was compared with the time of point label within a predetermined tolerance range and thus

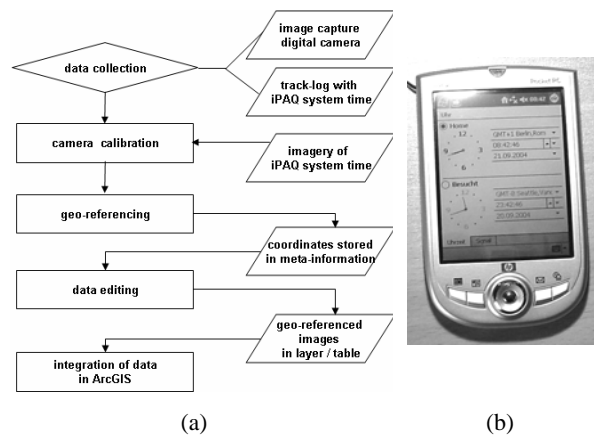(a)   (b)

Figure 2. (a) Data flow for geo-referencing mobile imagery. (b) Image capture for synchronization of time.

the image of the road sign received a specific position information.

The presented application was accomplished with a commercial software. The software PixPoint™ of the Red Hen Systems company is an ArcGIS extension and represents a comfortable tool for image processing including an image data management tool. The software synchronizes each image with the recorded data and labels each digital image with geodetic information. This is done automatically via time synchronization of the EXIF data.

To avoid a potential time difference between the system time of the handheld PC and the camera time, a calibration photo with the current system time of the pocket PCs scould be captured for each series of image captures.

4. IMAGE ANALYSIS FOR ROAD SIGN RECOGNITION

The analysis of the geo-coded images aims at an automated feature description of the concrete complete road sign configuration. On the one hand a detection of the road sign class was performed (warning signs, prohibition signs, instruction signs, and reference signs), on the other hand the sub-plate information plate and the associated text was segmented and classified. The fundamental ambiguity in the visual information under outdoor image capture conditions, for example, in the object recognition of 2D views, generally causes recognition rates below 100%. Therefore a semi-automated procedure was implemented, in which two essential processing steps are to be distinguished:

- **Automated object recognition process.** Detection and identification of the road sign and the sub-plate information were processed. Furthermore, a confidence measure as a means for quality evaluation was computed.

- **Supplementary post-processing of the recognition results.** On the basis of the classification results of the automated recognition system, the results of each road sign classification can be interactively updated by the user. Hence the automated recognition step can rapidly economise the amount of time and quality of the road sign classification - especially for large amounts of image data.
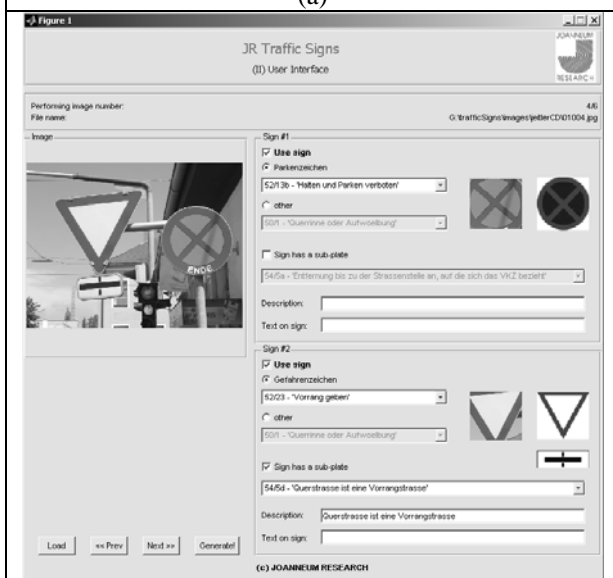
Figure 3. User interface for semi-automated image based road sign classification. (a) User interface for the monitoring of the on-going automated road sign recognition process. (b) User interface for the interactive post-processing update of the results of the image analysis.

The processing steps and the corresponding system evaluation will be described in the detail in the following Sections.

4.1 Categorization of Road Sign Images

The processing of the mobile images started with an automated image feature analysis (Figure 3). Figure 4 illustrates single processing steps for a specific example: (a) original image, (b) regions of interest on the basis of colour based pixel classifications, (c) specific contrast based contours are localized, (d) colour based classification of the contours and contour following [1,8] of the contrast transitions creates locally (e) straight lines and/or (f) curve segment [2], (g) first form categorization results after corresponding post-processing in the form of a global meta-grouping procedure with regard to straight lines cut angles and curve overlapping, (h) original image with detected road sign content.

Two relevant processing steps were, (i) the extraction of the contrast regions and (ii) the grouping of the gradient vectors. The contrast region were extracted using the principle of the bi-modal characterization of the local colour histograms [3]. Typical contrasts were in this way represented by the distributions of pixel training samples in colour attribute space. In this way, the contrast distributions of "red-white ", "red-blue "" and "yellow-white" were extracted, which are typical contrasts in road signs.

The gradient vector of the contrast contour normal serves then as a basis for the grouping of similar local orientations [4]. This leads to the feature generation of the straight lines as well as feature generation of curve segments on the other hand. Straight line and curve segments are then processed by a rule-based routine which derives the position of triangular regions, rectangular regions or circular sign regions. This logical step corresponds to detection and recognition of a basic road sign category: circular signs can be assigned to prohibition signs, triangular signs to the warning signs, etc.

A confidence value serves now for the evaluation of the recognition quality. The proposed approach extracts a confidence value out of every 'region of interest'-based contribution. In the next processing step, these single region contributions were aggregated using the combination rules of the evidence theory [5] into a global value. This value is used for the integrated classification of a road sign category. The confidence value attributed to the colour based categorisation together with the bi-modal contrast extractions, and with the segment classification were also taken into account for a global confidence estimation. Figure 5 shows results of the confidence aggregation regarding (i) the single pixel characterization and (j) the corresponding classification into road sign categories (see above).

4.2 Identification of Individual Road Signs

The extracted sign content was then fed into a pattern classification methodology for road sign identification. Because of the the different kind of appearances of the road signs, different classifiers were to be used. Those signs, the content of which could be well segmented by the background, were processed by a neuronal network (Multi-Layer Perceptron [6]) with 3 processing layers. However, the quality of the segmentation results achieved from the sign categorization was often quite insufficient in order to process patterns containing the complete sign appearance. In this way, we did not receive good recognition results for, e.g., instruction signs. Therefore an alternative classification procedure was selected which is able to process only local, scale- and rotation invariant pattern information. We applied therefore the Scale Invariant Feature Transform (SIFT, [7]) to characterize the road sign content. SIFT features are known to perform with rather high accuracy in the matching despite larger changes in the direction of the viewing angle, brightness and scale. The recognition accuracy achieved with SIFT features was eventually higher than with using the MLP.

4.3 Performance Analysis

A reference image database, which includes 4224 images, was selected for a statistical analysis from more than 20,000 road sign images about the Graz city zone.
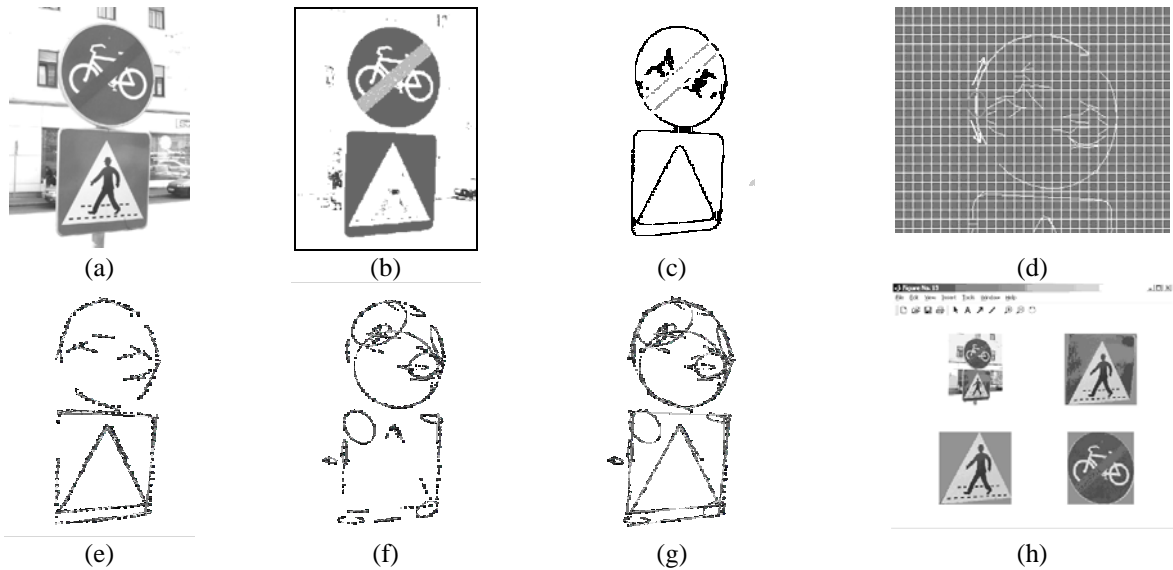
Figure 4: Automated identification of road sign classes. (a) original image, (b) regions of interest on the basis of colour based pixel classifications, (c) specific contrast based contours are localized, (d) colour based classification of the contours and contour following [1] of the contrast transitions creates locally (e) straight lines and/or (f) curve segment [2], (g) first form categorization results after corresponding post-processing in the form of a global meta-grouping procedure with regard to straight lines cut angles and curve overlapping, (h) original image with detected road sign content.
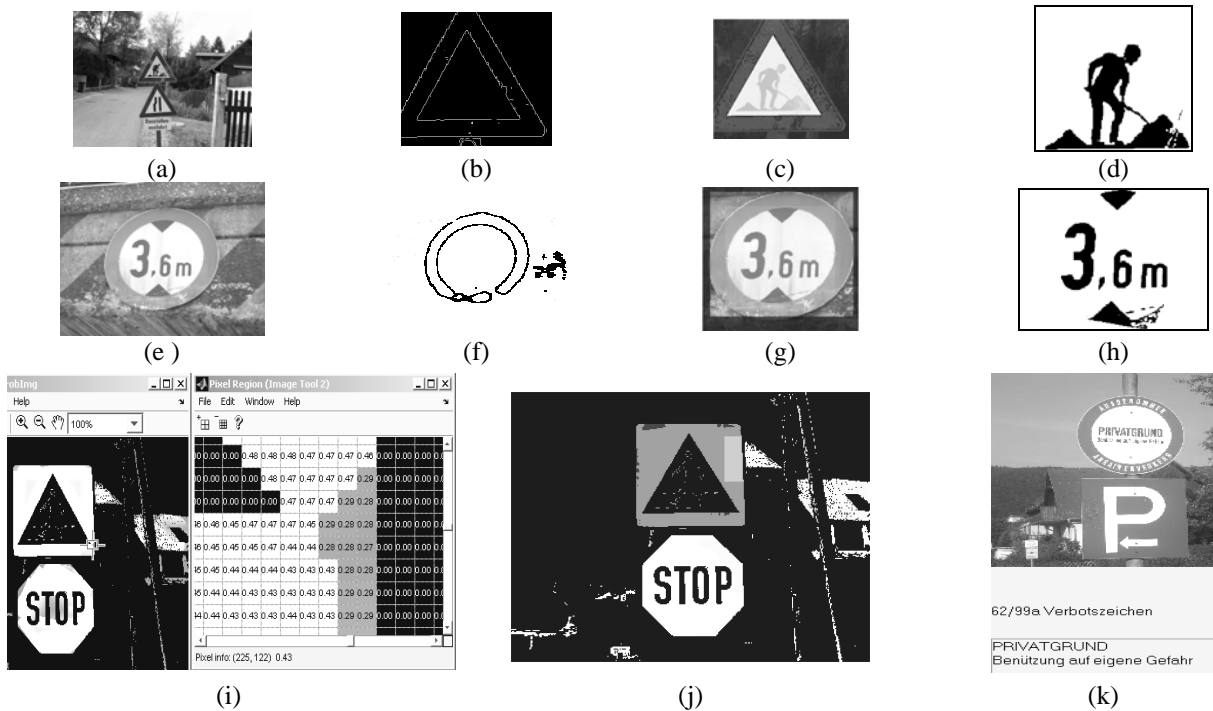

Figure 5. Identification of road signs. (a,e) original images, (b) result of the Radon-transformation for a straight lines recognition, (c) segmentation using colour classification, (d) the extracted sign content is then classified in terms of identifying the concrete road sign using an MLP neural network [6], (f) local contrast regions are characterized through the corresponding centre pixel, (g, h) localization and segmentation of the sign content, (i) local confidence needed for to the road sign categorization, (j) local classification on the basis of the confidence, (k) text recognition with OCR standard software on 'regions of interest'.

Table 1. Statistical evaluation of a reference record of the acquired road signs of Graz.

Category	Total	PT	PF	PT [%]	PF[%]
Warning	206	188	18	91,3	8,7
Prohibition	620	564	56	91,0	9,0
Instruction	452	324	128	71,7	28,3
Crossing	431	389	42	90,3	9,3
Parking	1599	1451	148	90.7	9,3
Preference	123	122	1	99,2	0,8
No entrance	409	397	12	97,1	2,9
Stop	267	267	0	100,0	0,0
Dead end	117	117	0	100,0	0,0
Total	4224	3819	405	90,4	9,6

Tab 1 shows results of the performance analysis in terms of recognition rates with respect to basic road sign categories and selected sign classes. The results present a rather high hit rate for the fully automated image analysis despite of altogether very changing illumination conditions, shadows, partial overlaps and sometimes extreme viewing angles. The methods can be refined in the future by intelligent analysis of local pattern detectors, for example using the classification of local grouping on the local descriptor responses.

4.4 Analysis of Sub-Plate Information

Additional signs are detected by the well-defined image features in the context of the road sign. Thereby a 'region of interest' is segmented, and stored for the subsequent optical character recognition procedure (OCR - standard software; Figure 5k).

5. DATA VISUALISATION IN GIS

After the image data were processed in terms of semi-automated road sign categorization and classification, the results were geo-referenced and stored in the Geo-database. A further module is then responsible for the specific visualization of the road signs. The individual images of the road signs were post-processed by an affine transformation which positions the image plane perpendicular to the viewing axis. In the ArcGIS presentation, the orientation of the road signs was aligned with the viewing axis in the direction of the road axis. This enables, together with the local information, a correct representation of the road sign images and symbols in the ArcGIS.

The complete image and road sign attribute data were then uploaded to the mobile road sign inventory system. This mobile system can then be used in the field, right at the spot, for monitoring and updating the data.

6. CONCLUSIONS

This work presented the methodology and the design of a complete solution for a mobile road sign inventory system, including a post-processing step on mobile imagery about on-trip collected data. Road sign categorization and classification was performed in a semi-automated manner in order to guarantee the robust operability of the system. Finally, we have demonstrated that the presented methodology was able to perform road sign identification with a sufficiently high quality, i.e., with recognition rates around 90%, and confidence values which enable to provide a quality control with respect to the complete process. Furthermore, we presented – to the knowledge of the authors – for the first time results with respect to the complete inventory of road signs of a whole city area.
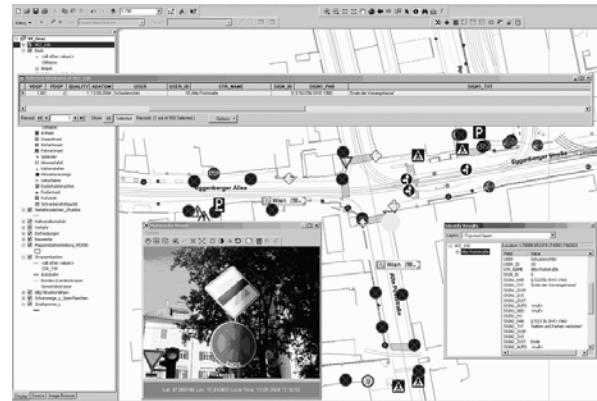

Figure 6. Screen Shot of the road sign land register, desktop GIS application – ArcMap.

Figure 7. Edit forms and/or query forms of the mobile GIS application

7. REFERENCES

[1] Geladi, P. & Grahn, H., *Multivariate Image Analysis*, John Wiley & Sons, NY, 1966.

[2] Fitzgibbon, A. W.and Pilu, M. and Fisher, R.B. (1999). Direct least-squares fitting of ellipses, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 21(5):476–480.

[3] Matas, J., Koubaroulis, D. & Kittler, J., Colour Image Retrieval and Object Recognition Using the Multimodal Neighbourhood Signature, *Proc. ECCV 2000*, pp. 48-64, 2000.

[4] Wyszecki, G. & Stiles, W.S., *Colour Science*, John Wiley & Sons, NY, 2000.

[5] Shafer, G.A., *Mathematical Theory of Evidence*, Princeton University Press, Princeton, NJ:, 1976.

[6] Haykins, S., *Neural Networks*, Prentice Hall, NY, 1999.

[7] Lowe, D., Distinctive image features from scale-invariant keypoints, *International Journal of Computer Vision*, 60, 2 (2004), pp. 91-110.

[8] Bracewell, R.N., *Two-Dimensional Imaging*, Englewood Cliffs, NJ, Prentice Hall, 1995, pp. 505-537.