Paper ID TP0787

Computer vision and automated traffic sign updates from commercial fleets

Markus Melander[1*], Petri Hienonen[2]
1.Vionice ltd, markus.melander@vionice.fi, Finland
2.Vionice Ltd, petri.hienonen@vionice.fi, Finland

Abstract
In some European countries, the aims to maintain the up to date traffic sign database are taken seriously. Finland is among those European countries that are participating in joint project called TN-ITS led by ERTICO. How to up-to-date public databases and the data delivery to navigation service providers are increasingly relevant challenges to solve in the era of connected navigation and rise of automated vehicles. Finnish Transport Agency (FTA) executed a pilot project to solve the challenge for updating the traffic sign chances with computer vision in real time. To map the changes the postal delivery vehicles were equipped with mobile phones with a state-of-the-art computer vision application. The aim was to get update information about traffic signs and get them mapped without manual field works. In presented solution verification possibility from geospatial images is one of the key factors to ensure updating process. Results of the project are providing an excellent showcase that presented system is feasible to automate traffic sign database management. This paper is written from technical and process perspective. It is trying to reveal the practical steps how city or road administration would be able to deliver and maintain high quality traffic sign database with simple technical setup.

Keywords:
Traffic signs, computer vision, TN-ITS, navigation

Pilot arrangements and online video data collection

In the pilot, Postal vehicles were equipped with a smartphone based video recording software to collect video every day during a month in September 2016. The fleet size of six vehicles were selected from Posti, Postal service provider. This is because the up-to-date process was to be demonstrated. For technical demonstration there would have not been a need to equip that many vehicle but the sufficient data handling processes were planned in the way that there is no need for extra investment and the setup was considered to be enough to illustrate the capability of the system in limited pilot. For Vionice is was important to highlight some of the issues that might be present when implementing a possible national scale solution. The general overview of the pilot is presented in Figure 1.

The equipment used in pilot´s data collection were a mobile phone, memory card, mobile phone holder, and data roaming contract. During driving, videos were automatically uploaded to the servers where computer vision processing was used to extract turn restriction signs and speed signs.

The selected testing area was chosen from region of City of Kouvola (Figure 2) which includes both urban and semi-urban scenes (examples in figure 3). In figure 4 is represented an example of same scenes that are recorded multiple times during the constant surveillance of the same routes. The change management if demonstrated by querying a specific traffic sign against its previous detections. The figure 4 presents same sign, as found by automation in different days.

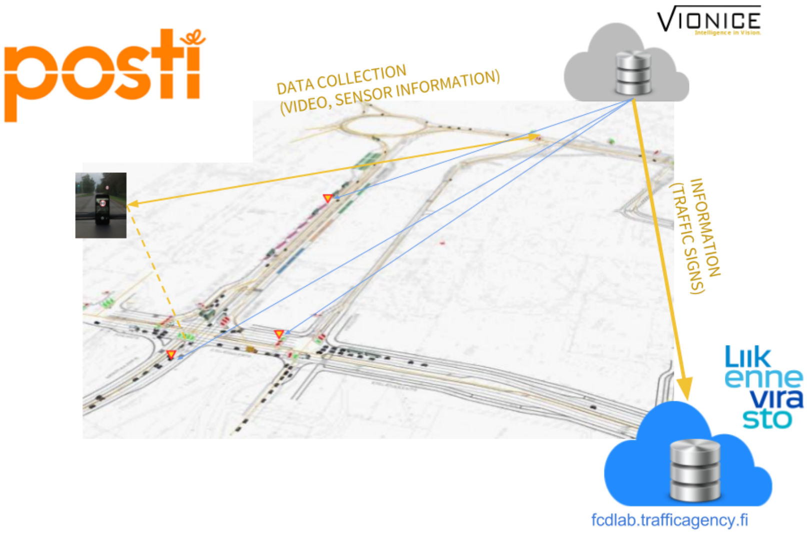

Figure 1. Project process model

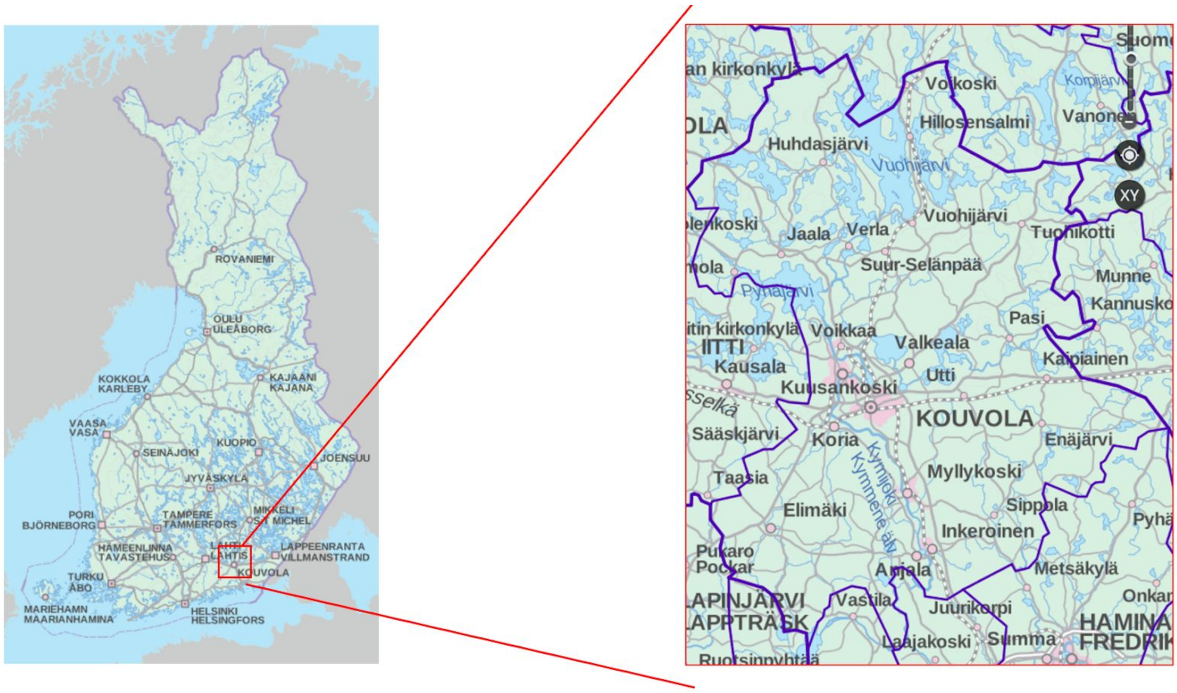

Figure 2. Pilot area is located in Kouvola region

(a) (b)

Figure 3. Two common views of the video material: a) Example of non urban scene; b) Example of urban scene.

(a) (b) (c)

Figure 4. Detections of same traffic sign from different days (drives). The detection times are as follows: a) 2016-09-20_12:43:02:338 b) 2016-09-26_17:45:08:687 c) 2016-09-16:17:48:07:581.

During the project, more than two months, 7400 hours of video was recorded and processed through computer vision pipeline extracting traffic signs. In a fraction of the material the global positioning system (GPS) accuracy was worse than anticipated based on the previous experience, but due to the large overlap between different drives not so good material could be discarded and only material with good GPS signal was used for traffic sign extraction and 3D model generation.

Posti’s fleet covers quite trustworthy the whole network, because the mail is to be delivered to all the houses. Figure 5 shows how the road network were covered in this pilot. The stronger the line in Figure, more times the same route was covered. In this pilot the camera was directed forwards, but with the same effort camera could be installed to either face backwards to cover the road to both directions. One notice made during the project was that phone might not be an ideal solution for data collection in large scale operational process. The discussion was mainly raised because interaction required with the device from the Posti’s workers. After the pilot some additional features have been developed according to lessons learnt; geospatial fencing to avoid recording on private properties and automated switch-on of the recording.

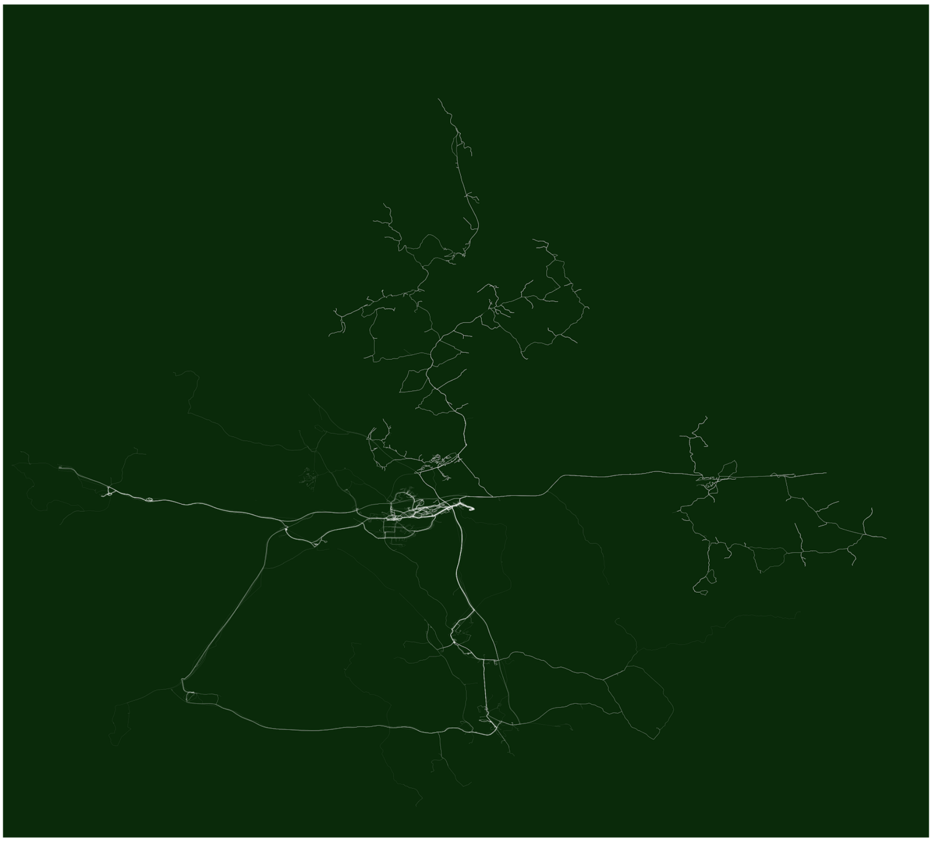

Figure 5. Image of the all driven routes.

Results of the computer vision system

Total 1186 individual speed, turn restriction and weight limitation traffic signs were detected during the pilot. Because the driving took place in same routes multiple times, all of the signs accepted and evaluated to be right by the algorithms were detected more than once. This was in order to verify the correctness of detection. After detection and classification the signs within succeeding drives were combined using clustering. As more signs are discovered the computer vision can be trained to recognize signs in more difficult environments and each discovered mistake made by computer vision can be used to improve the system. As the signs are collected each time that a camera drives next to a sign, the inventory is made more accurate.

Essential technical perspective in pilot was the usage of video for automated detection of signs. The video based analysis is better if compared to image analysis for two main reasons: 1) From video the generation of 3D model is possible. Signs are mapped more accurate while 3D point cloud is used, 2) Video provides multiple frames for detection and this increases dramatically the detection confidence of each individual signs. (Figure 6)

The overall accuracy of the computer vision itself is on par with human performance, every sign that can be reasonably detected is detected and classified correctly. Because this was the first highway inventory in Finland, there was no previous material of some of the signs adapting the computer vision to a new environment was required.

In FTA pilot there were not available sufficient ground truth data of signs. In similar pilot in Norway the detections accuracy of system was more than 98%. The result is excellent and reason for good result is based on algorithms visualised in image 6.

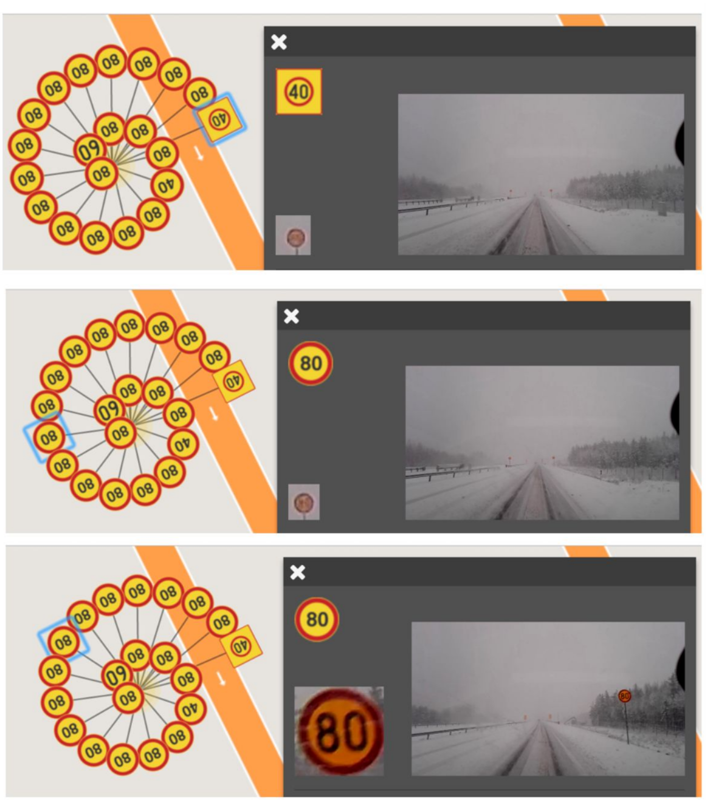

Figure 6. Illustration how all the signs are detected of multiple video frames

Change detection and sign database management

As a proof-of-concept, one of the goals of the project was set to detect changes in the signs as the roads were traversed. In total there were three automatically detected changes in speed signs all relating to the repairs or maintenance of the roads. The signs discovered could easily be verified by humans, categorized automatically (for example because there is a road repair sign the speed limit is temporary). While thinking usage of produced data there is most likely desire for man made verification. The system used have integrated image service of each detection. This provides location and detection information to be provided for human verification (Figure 7). The change management is generally easier to implement if material from wider temporal window is available. Automation always makes some mistakes (even if the number is lower than a human doing the task) and statistical analysis provides safety against (rare) mistakes.

Figure 6. Automated alerts were provided from detected changes.

Figure 7. All the signs are visualised to map with relevant attributes and metadata.

Why video data is suitable ?

All of the video material should have corresponding (the timing error should be synchronized to millisecond class) GPS data and gyroscope data for the correction. For more advanced computer vision methods video material is necessary and allow better analysis and less dead zones where there is no material available.

As the sensors in the future develop more sophisticated sensing becomes available for inventories. Some of the possible sensing devices (cost < 600€) for national scale that could improve inventories use include panorama cameras, directly depth sensing sensors (such as lidar or other time-of-flight sensors) and more accurate RTK GPS devices. To include sensors into existing processes there is also a need for easy to use solutions. The phones will provide most up to date and cost efficient solutions due to larger volumes, and for example the phone's camera often corresponds to a very expensive computer vision camera.

According to results in presented project there is no need for more expensive sensors or sensor fusion. The data accuracy and detection confidence is good enough for maintenance process to national ITS databases as well as for navigation map making process where there is linear referencing in use. [1]

What sensor data is required at minimum to create an inventory

As far as can be told the minimum for a quality (accuracy error < 2m) inventories require normal phone quality GPS and recorded video. Video is necessary to use more advanced computer vision methods for significantly better results. The one thing that has to be highlighted that all material will have to be in same time format and the timestamps will be accurate and synced up to 1/100 fraction of a second for the information to be useful.

Data collection can be improved by adding extra sensors to the mix. Two possible additional sensors for the inventory that provide most benefits are:

1. RTK GPS
2. Panorama camera

Table 1. Requirements for the collected video material to enable automatic traffic signs inventory.

SENSOR	REQUIREMENTS
GPS	At least 1/s. Improved GPS provides easier to utilize data, but does not necessary improve results.
Video	At least 10 frames/s to prevent dead spaces between images and to use more advanced computer vision methods.
Gyro, Accelerometer, Compass (optional, though very useful)	Can be used to improve GPS by performing dead reckoning (estimation of the movement based on the sensor information).

Automatic data production

In general the system in a rough level consists of three parts that can be evaluated separately:

1. Quality of the recorded data from the streets.
2. Using computer vision to extract information.
3. Combining information into a register format, detecting changes and informing about updates.

These three parts define the accuracy of the final system and output data. The current computer vision recognition produces nearly perfect results (system have been evaluated in Norway by Vegvesen 98% of the signs are detected)[2] in comparison to a human performed inventory from the same images. One problem is that the computer vision is somewhat moving target, in next six months a new iteration of the inventory pipeline is reached that either improves results or lowers the quality of data required.

1

2

Computer vision in general can be used as a tool to extract useful information from sensor data, such as images. Computer vision is a general term for a subfield of computer science and currently at least few thousand researchers are working worldwide task related to improving computer vision capabilities. In the recent years the improvements in the camera technology, processing power of computers and machine learning algorithms have enabled deployment of computer vision to tasks that previously have required human participation.

Change management of traffic sign information

Combining detections or data from different sources is a challenging task, especially when considering of informing of a missing sign. The question is what is the right threshold to remove signs if the sign has not been seen in the previous passes. For example, there could be some occluding objects such as cars or snow on top of the traffic sign, and different processes most likely will require different reaction times. Basically there are two different problems:

1. Combining the data within one drive (so that the same traffic sign does not multiply) and between multiple drives so that the detections of a same sign can be correctly assigned into the existing data has been studied in the research project. The first problem is solved directly in the computer vision side of the things by tracking each individual detection frame by frame.
2. The second problem is more relevant in the perspective of centralized database since multiple contractors will eventually start to send overlapping data. The best existing and robust solution basically requires a human operator. Even though it is not completely sure yet whose responsibility in the process is this, but what we can say is that the current solution is hard to scale up and after all it is a technical problem where we saw an opportunity to use our expertise.

We have learned that this is actually a very crucial step in producing the most valuable information of what have been changed with respect to some reference data. This leads us to the situation where the update process is technically possible. Commercial map makers are already providing online maps and this is why the process to provide change information is valuable. In coming era of automated vehicle it is most likely essential.

Verification process and user interface

A good precautions until the automation can be raised to a level of full automation is to verify the data with a human. A such process already exists naturally in those countries or map makers that do change manually point data to referenced linear information. A current video/image (example in figure 8) of the change situation can be shown to the person validating the change to remove any possibility of introducing new errors to the official data. If this process and the statistics are recorded it is later possible to automate the validation process itself. In first place the right solution is to visualize the detected change to user interface from WFS or like interface.

Figure 8: Example of up-to-date visual images for validating

Recommendations how the data collection could be organized

In this chapter the discussion on the options for the data gathering process is discussed. Basically there are two different models for data gathering: 1) distributed where multiple parties contribute to the data collection and 2) one commercial provider that is used in this pilot. Table 2. summarizes distributed and fleet based data collection. As a option crowdsourcing might be a cost effective option but some commercial contractor is recommended to ensure that all the routes are covered on certain intervals.

Table 2. Distributed vs centralized data collection models.

	Distributed data collection (for example municipalities)	**Large fleet based data collection (for example Posti)**
Data handling and contracts	Require separate contracts	Single contract
Training of personnel	Requires separate person(s) to handle the operations	Large scale HR practices to train people
Equipment	If the collection happens more sparsely the equipment can be more readily available smartphone.	In large scale operations, according the Posti, should be fixed to make the collection easier for everyday operations. The costs are similar as with smartphone, but the equipment can be made more reliable.
Costs	If the usage of data recording could be done as a service for FTA or mandated by a suggestion to the municipalities the collection costs would be low. If standardized equipment would be needed cost might increase.	If the data collection can be done within normal processes and the benefits could be found to the fleet owner the cost could be relatively low. Separate drives possibly be dedicated if some parts of the network are never driven. Camera can be set to watch backwards and forwards on separate days.
Coverage	Full coverage could be achieved using crowdsourced, business fleets or combination. Service provider with attitude and right kind of funding can make this happen. Still this should be tested but assumed to be rather easy to organize.	The normal operations can cover up to 70% of the road network. Every mailbox where there is permanent residence is visited. Easy to organize.
Updates	Getting stable update rates might be difficult to coordinate	Updates are easy to coordinate

Conclusion

The need for up-to-date processes and technologies to traffic sign data management is needed. The cost-efficient methods are valued because the correctness is not critical. In this pilot project the system that is capable to produce desired data and map it in high accuracy is presented. The processes to implement video data collection process along the streets is more the problem than the technical solutions, which is proven to work well.

Data privacy is always a concern while images and video is captured and processed. The computer vision based anonymisation is a solution to ensure that no

As the technology is here some pilots and concepts for the crown sourced or fleet based proof-of-concept like projects would be needed. As well as the discussion among the cities and road administrations how seriously they are going to react to the problem of miss information in navigation. The problem is assumed to be more and more topical in times of connected navigation and automated vehicles.

References

1. Borzacchiello, M. T., Boguslawski, R., Pignatelli, F. (2016), "Improving accuracy in road safety data exchange for navigation systems". Available at: http://publications.jrc.ec.europa.eu/repository/bitstream/JRC104569/jrc104569_d%2021%20tp%20final%20report%20-%20v1.7%20pubsy.pdf