

Automatic number plate recognition for the observance of travel behavior

Autoren / Authors:

Markus Friedrich

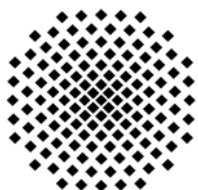
Prokop Jehlicka

Johannes Schlaich

Lehrstuhl für Verkehrsplanung und Verkehrsleittechnik, Universität Stuttgart
vorname.nachname@isv.uni-stuttgart.de

Veröffentlicht in / Published in:

Friedrich, M., Jehlicka, P., Schlaich, J. (2008): Automatic number plate recognition for the observance of travel behavior, 8th International Conference on Survey Methods in Transport: Harmonisation and Data Comparability, Mai 2008, Annecy, Frankreich.



Universität Stuttgart
Institut für Straßen- und Verkehrswesen
Lehrstuhl für Verkehrsplanung und Verkehrsleittechnik
www.isv.uni-stuttgart.de/isv/

AUTOMATIC NUMBER PLATE RECOGNITION FOR THE OBSERVANCE OF TRAVEL BEHAVIOUR

**8th International Conference on Survey Methods in Transport,
France, May 25-31,2008**

AUTHORS

Prof. Dr.-Ing. Markus Friedrich

Dipl.-Inf. Prokop Jehlicka

Dipl.-Ing. Johannes Schlaich MBA (USQ)

Editor address: Universität Stuttgart
 Lehrstuhl für Verkehrsplanung und Verkehrsleittechnik,
 Seidenstraße 36
 D-70174 Stuttgart
 Phone: +49 (711) 685-82480
 Email: forename.surname@isv.uni-stuttgart.de

ABSTRACT

Systems for Automated Number Plate Recognition (ANPR) can provide a valuable data source for transport planning and engineering. Multiple tasks can be solved with ANPR-systems located at one or more survey points. The paper describes some characteristics of ANPR-systems and presents five applications of the systems: vehicle classification, travel time measurements, through traffic surveys, route choice observations and estimation of O-D matrices.

KEYWORDS

Automated number plate recognition, route choice, travel time, through traffic.

1 INTRODUCTION

Classical stationary detectors such as induction loops can only measure volumes and local speed. They do not permit to measure travel times for longer distances or to survey the route choice behaviour of drivers. New survey methods such as Floating Car Data (FCD) allow these measurements. However, due to high installation and data transfer costs the critical percentage of FCD-equipped vehicles can often not be reached.

An apparent method to solve this problem and to identify vehicles at different locations is to observe the number plate, which is unique for every vehicle. Manual number plate surveys have a long tradition in transport planning. They are primarily used to determine the origin, destination and through traffic of an area. Manual recordings of number plates, however, are costly and error prone as human surveyors are overburdened by this task.

In recent years Automatic Number Plate Recognition systems (ANPR) have become a reliable and affordable state-of-the-art technique. They are used by the police for the enforcement of traffic rules or for crime detection. In London ANPR-systems are applied to register the vehicles entering the congestion charging zone.

The paper first describes some characteristics of ANPR-systems and reports on observed detection rates. Then it presents five applications using ANPR-systems: vehicle classification, travel time measurements, through traffic surveys, route choice observations and estimation of O-D matrices.

2 AUTOMATIC NUMBER PLATE RECOGNITION

2.1 System description

ANPR-systems for the recording of number plates of vehicles normally consist of two components. Firstly, a camera that detects passing vehicles and continuously sends the images to a computer. Secondly, software that recognises number plates with its characters and stores them in a database. Some websites of ANPR-System contractors are listed in the Annexure.

Camera

The Camera itself consists of an infrared detecting camera, a general optical colour detecting camera and an infrared light emitting array of LEDs. The LED array beams infrared light in the direction of the infrared camera, which then captures the light reflected by the white background of the number plates of passing by vehicles, which appears white on the image. The non-reflecting colour of the characters and the vehicle's surface appear black. Direct sunlight enhances the infrared reflection, the LED array however is bright enough to recognise number plates in absolute darkness. The focal length of the infrared camera is adjusted to detect an overall width of one lane. The colour camera with a lesser focal length

generates images for overall view and alignment of the whole camera body. Both images are sent in intervals of 300ms to a Computer, where the installed software processes them.



Figure 1: Two ANPR-systems set up on a bridge.

ANPR-systems can be either set up as shown in Figure 1 on a bridge construction over a carriageway or on the hard shoulder of a carriageway. In the latter case it is not possible to detect traffic on two lanes because the further lane will not be recognized optimally due to shadowing effects.

SOFTWARE

As soon as the software recognizes an image with a number plate, the full string is identified by an optical character recognition algorithm and is checked for international plate syntax to determine the country of origin.

Afterwards the number plate string is saved in conjunction with a timestamp and the ambience image of the colour-camera to a database. Figure 2 shows the operational sequence of the whole recognition process.

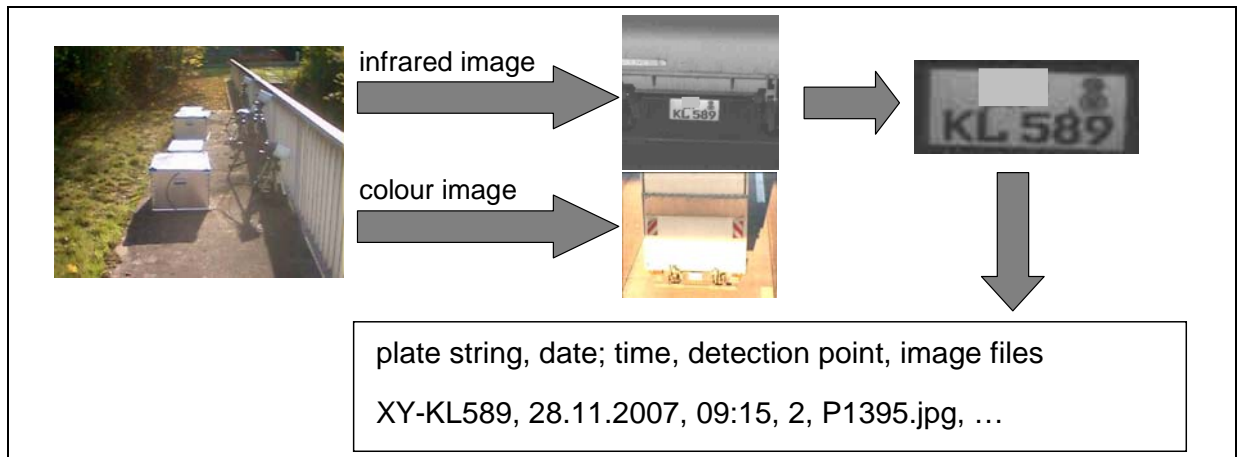


Figure 2: Schematic operational sequence of a number plate recognition.

2.2 Rate of detection

This chapter reports on an analysis of the detection rate based on measurements on German roads with mostly German number plates with black letters on white reflecting background.

Examining the detection results a deviation between recognized number plates and passed by vehicles is noticed. The detection rate depends on environmental factors, which can not be influenced. These factors include precipitation, angle and intensity of the sun, shading on the pavement as well as dirty or deformed number plates. However the detection rate also depends on controllable factors such as the angle of the camera relative to the horizon and the resulting distance between camera and number plates.

Figure 3 shows the detection rate over time on two days with similar weather conditions and traffic states. The noticeable likeness of the days leads to the conclusion that the rate of detection depends on the time, that results in changing angle of solar radiation and shading on the pavement caused by trees or structures on the roadside.

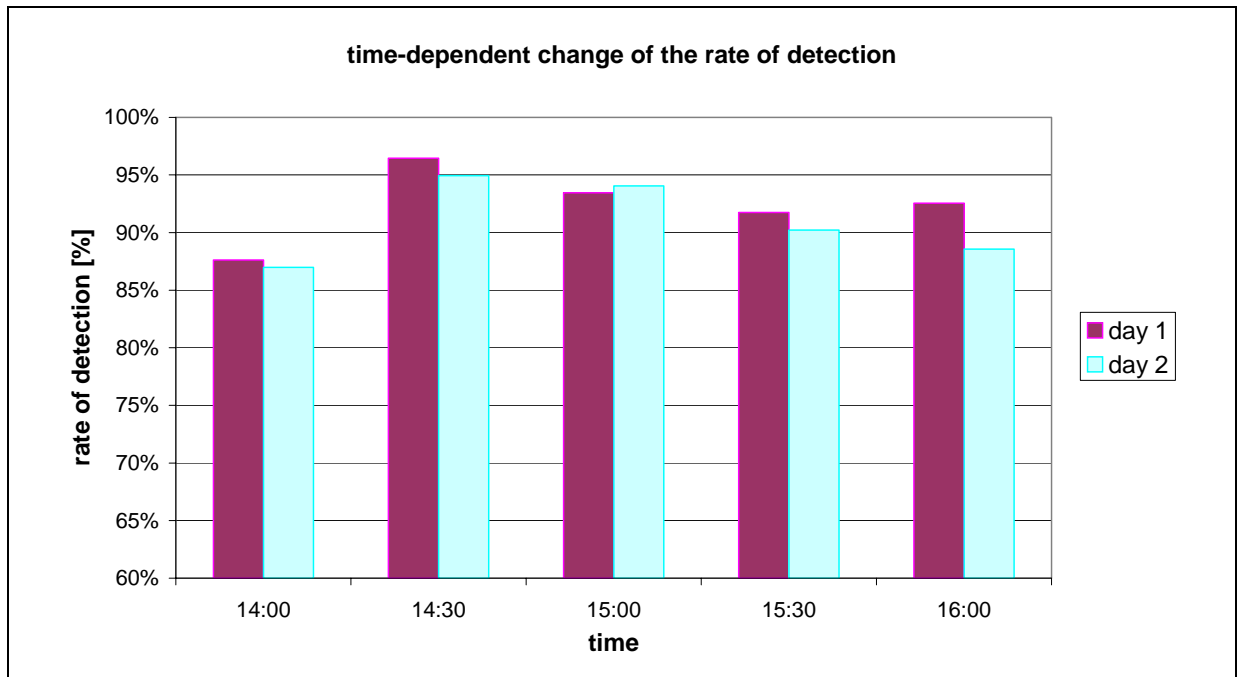


Figure 3: Time dependent detection rates for two days.

Figure 4 presents the change of detection rate dependent on the angle of the camera's field of view with reference to the pavement. Thereby, 100% equals the best result, which was achieved at an camera angle of 40° and 45°. The rate of detection decreases with lower angles as well as with higher angles. At a low angle the distance between camera and number plate increases. As the illumination of the infrared LEDs decreases by the square length of the distance, the number plate reflects not enough light to allow recognition. On the other side a high angle results in a high skew of the number plate, so that the number plate string is not recognized correctly. Also a high angle can result in the plate being concealed by superstructures of the vehicle (e.g. a truck trailer).

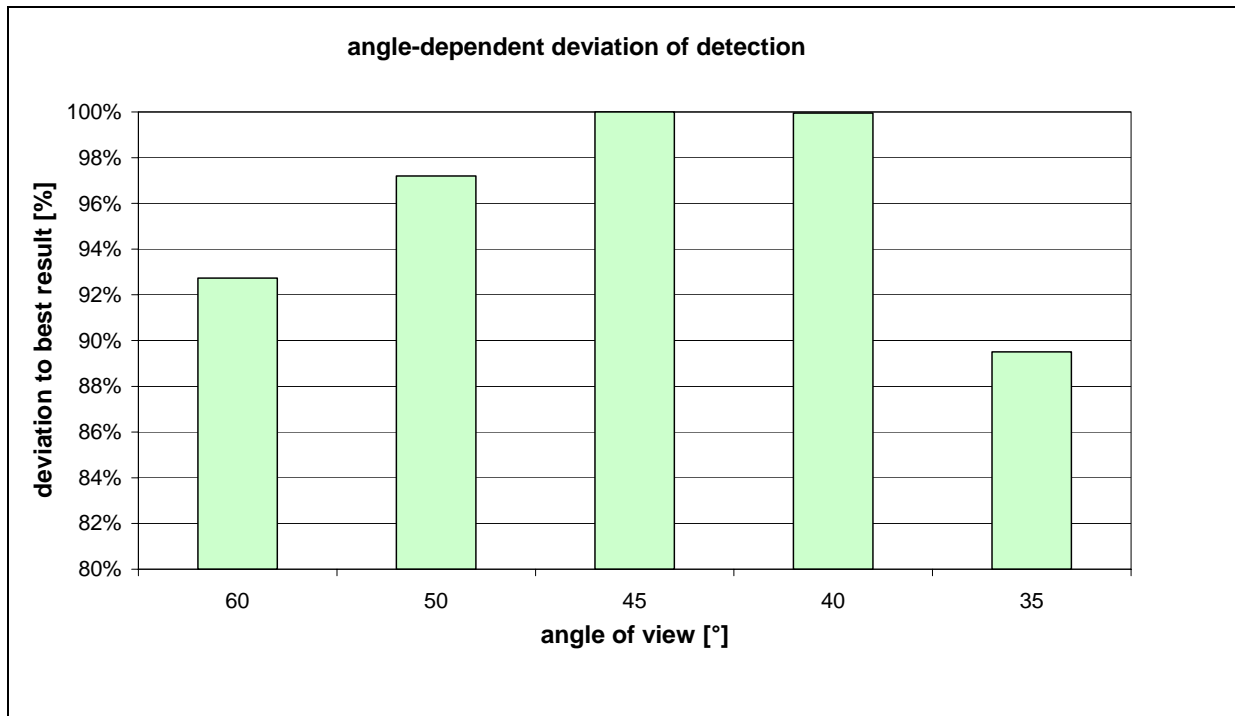


Figure 4: Angle dependent detection rate.

In summary one can expect that a properly set up ANPR-System recognises at least 80% of all passing by vehicles, but in most cases more than 90%. The measurements show that the systems are capable of recording high speed vehicles on motorways without significant loss. The systems also records heavy traffic volumes with observed 2,300 vehicles per hour

2.3 Matching number plate strings

To derive travel times or routes it is necessary to compare the number plate strings recorded at various ANPR-locations. This requires a string comparison procedure.

Due to data protection regulations it is usually necessary to encode the number plate string. An invertible function for encryption can be enabled in most ANPR-systems, which then saves an unambiguous text string instead of the plain text number plate string. With this encrypted string it is possible to identify the same number plate string on two different locations without knowing the decoded number plate string by merely comparing the encrypted strings. This can be achieved by a simple database query.

When number plate strings are recorded by human surveyors a common error results from twisting the sequence of characters, e.g. instead of “S-XY 123” the string “S-YX 123” is recorded. This type of error can not occur with ANPR-systems. Using ANPR-systems two types of errors may occur in the number plate recognition procedure:

- With a polluted number plate the procedure can falsely identify one or more characters, e.g. 0 instead of O, H instead of II or B instead of 8.

- Vehicles changing lanes may cause a state where only a portion of the number plate string is recognized and stored in the database.

If such a number plate string is now encrypted, it can not be identified at other locations. If the string is not encrypted, it is possible to expand the comparison from a single equality check to a calculation of similarity by checking each character in a recognized string with its adjacent character on the other string. Then the distance-value between two strings is the amount of unequal characters plus the difference between the lengths of the strings. Test series showed that strings with a difference-value of 1 still can be regarded as identical. A difference-value of 2 in contrast predominantly indicates different number plate strings. The computation time for such an extended comparison is approximately 30 times longer than a simple comparison.

When processing strings recorded at multiple ANPR-locations, a number plate comparison must be arranged between all relevant pairs of locations. In doing so a number plate string can be recognized at multiple locations, if by chance some similar number plates are poorly recognized. To obtain a reproducible and highly accurate result, all relations have to be processed multiple times. A threshold is defined that defines the limit up to which two number plate strings are noted as identical. At first, the threshold is set to the minimum, so that only equal strings are defined as identical. In each subsequent iteration the threshold is raised and applied to a new round of comparison of those strings that have not been identified yet.

3 APPLIANCE OF ANPR-SYSTEMS

With ANPR-systems it is possible to solve a multiple set of tasks. In this chapter five types of applications are presented. Table 1 lists these applications and discusses which applications permit the encrypting of the number plate strings.

Application	Encrypting of the number plate strings possible?
Vehicle classification	No. Authentic string required to identify vehicle class.
Travel time measurements	Yes. Encrypting does not significantly reduce the sample size of observed vehicles.
Determination of through traffic volume	Yes, but it reduces the quality of results as this application requires the complete sample of vehicles. If not all vehicles are detected the detected vehicle volumes must be projected using additional count data.
Analysis of route choice behaviour	Yes, but it reduces the quality of results as this application requires the complete sample of vehicles. If not all vehicles are detected the detected vehicle volumes must be projected using additional count data.
Estimation of O-D matrices from the area code	Yes, if only the second part of the string is encrypted and not the area code.

Table 1: Appliance of ANPR-systems

There are multiple applications, ANPR-systems are already used for, e.g. electronic tolling of zones and roads or law enforcement. These applications will not be presented in this paper.

3.1 Vehicle classification

The ANPR-systems can not differ between the types of vehicles when detecting a number plate. If nevertheless there is the need to analyse the composition of a detected number plate vehicle fleet like a differentiation of passenger car and HGV or the portions of emission classes of all detected vehicles, it is possible to do so by sending a plain text query of the number plate strings to the German Federal Motor Transport Authority (Kraftfahrt-Bundesamt). Data protection regulations require sending larger groups of number plate strings to the authority, e.g. all vehicles registered in one hour or on one route. The result is a classification for each group containing the portions of queried attributes in each group. Figure 5 shows a classification of HGVs in emission classes in two successive years.

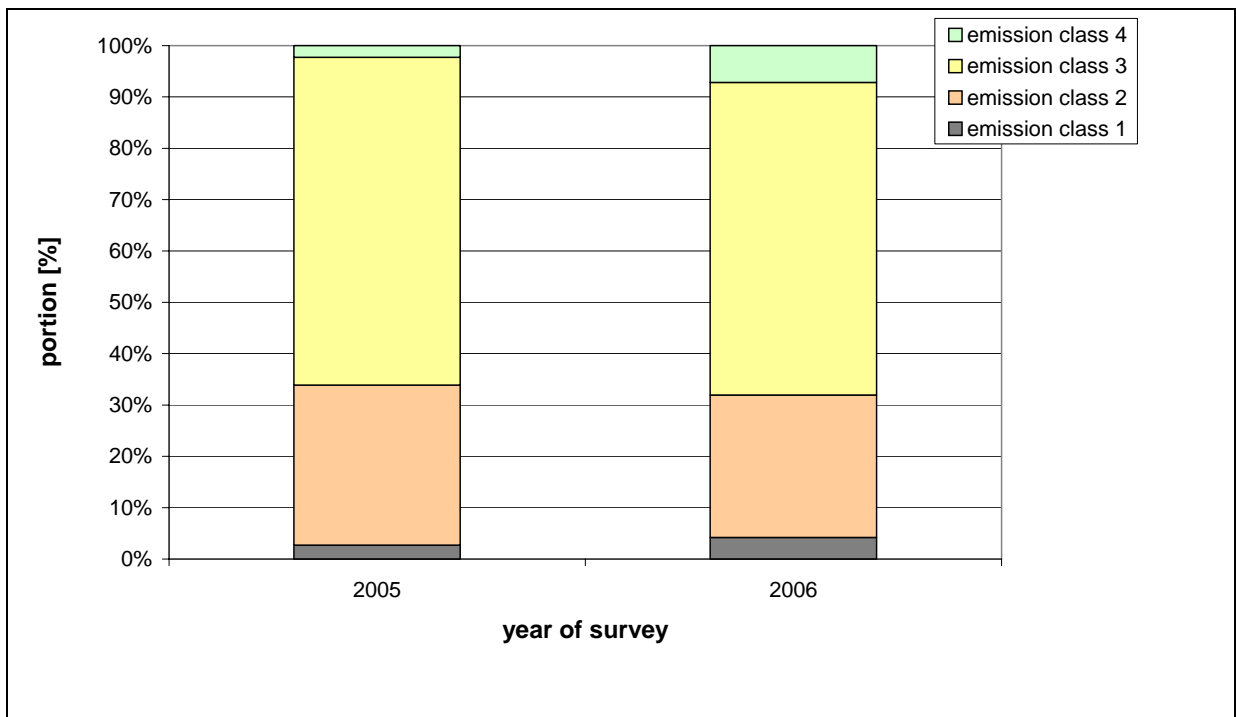


Figure 5: Proportion of HGV by emission class in the years 2005 and 2006 at the same location and at the same day of the week.

3.2 Travel time measurement

Measurement of travel time with conventional techniques is insufficient or only feasible with high costs. Measurements of velocity on stationary detectors allow only an estimation of travel time. A more accurate measurement of travel time is possible by tracking vehicles on

the whole route where the travel time is to be measured. In doing so or accessing Floating Car Data only a small portion of motor vehicles is observed.

When setting up ANPR-systems at a minimum of two positions of one route, travel times of all travelling through vehicles can be measured with high accuracy. Figure 6 shows the travel times of single vehicles displayed as points. Between 8:00 and 11:00 o'clock the travel time raises due to congestion on the standard route. A differentiation between the two vehicle groups can be noticed:

- Vehicles on the standard route: These vehicles have a raised traffic time due to congestion.
- Vehicles in subordinate net: These vehicles noticed the congestion in time and circumnavigated it in the subordinate net. Their increased travel time was clearly lower than that of vehicles on the standard route.

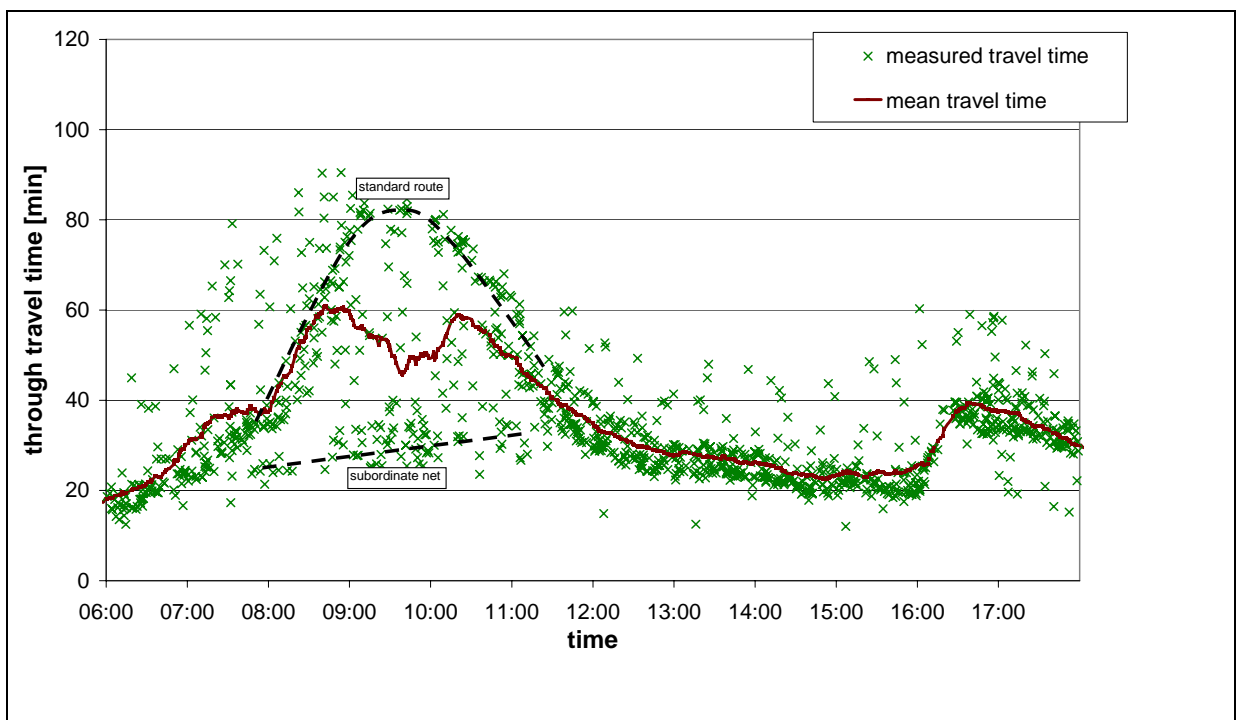


Figure 6: Travel time through a corridor.

By detecting vehicles on a route including an intersection with traffic lights, the amount of stops and the lost time on the intersection can be assigned to every single vehicle when detecting its travel time.

Travel times detected on highways can be compared to broadcasted travel news. Figure 7 shows a linear regression analysis between the length of the broadcasted congestion and the measured travel time. The measured values are based on multiple surveys on one route of a motorway on working days.

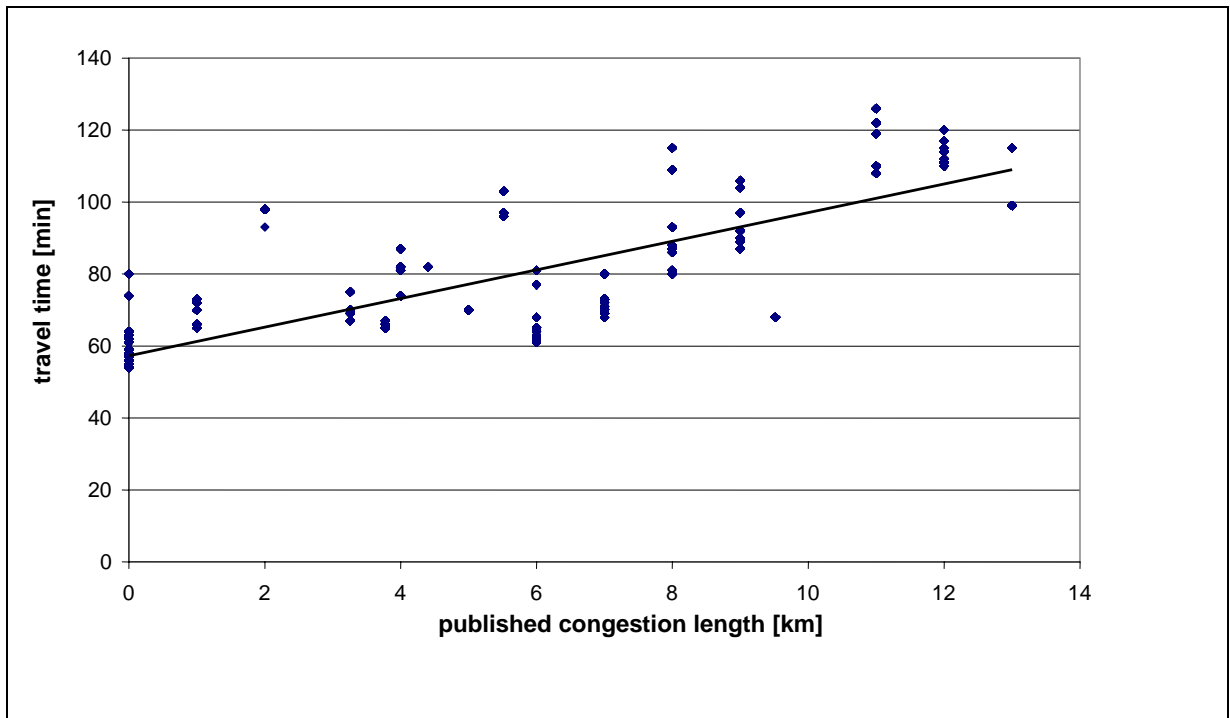


Figure 7: Regression analysis between broadcasted congestion and measured travel time.

3.3 Determination of through-traffic volume

With ANPR-systems vehicles entering a study area on one measuring point of a cordon node and exiting on another measuring point, can be detected. In contrast to travel time measurement it is essential to correctly register as many vehicles as possible. Because of the very computing intensive calculation of similarity a filter is needed to reduce the compared number plate pairs. For each pair of measure points a minimum and maximum travel time is defined. Only number plate pairs are considered whose travel time would be between the defined bounds.

A vehicle detected on two measure points is only assigned to through traffic, if one can assume that the vehicle transited the area without stopovers. Vehicles with a stopover are called „fractional through traffic“. Since there is no possibility to differ real through traffic from fractional through traffic based only on their travel time, an algorithm is applied to filter the through traffic with a significantly higher travel time than the majority of vehicles at the same moment. This algorithm consists of two steps.

- Determining of the mean through traffic travel time: For each through travelling vehicle the mean travel time of the surrounding vehicles is identified. Therefore all through travelling vehicles are considered that pass the entering measuring point within the interval of 15 min before and after the examined vehicle. From these retrieved vehicles the fastest 50% are taken to calculate the mean through travel time.

- Branching off the real through traffic: All vehicles, whose travel time undercuts twice the value of its determined mean through traffic travel time, are counted as real through traffic.

The differentiation between through traffic and fractional through traffic at one measure point is shown in Figure 8.

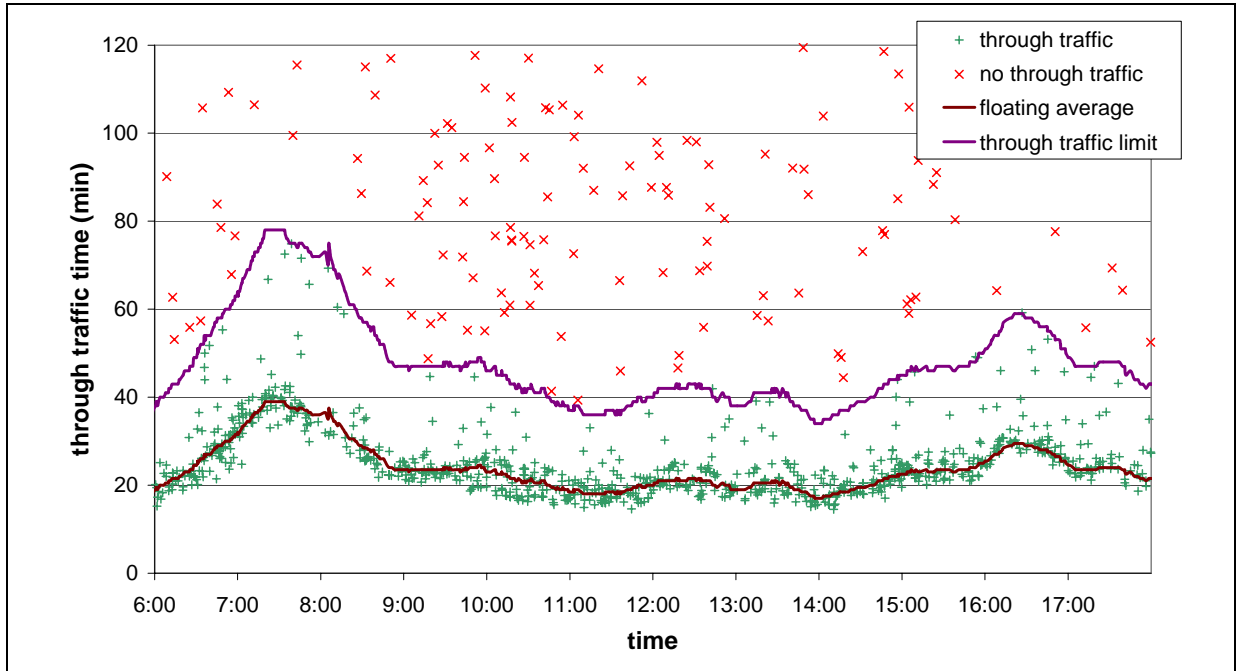


Figure 8: Differentiation of through traffic.

3.4 Route choice behaviour

Today, route choice behaviour in road networks can hardly be monitored for a large sample. Monitoring the route choice behaviour of travellers e.g. with GPS is costly and can only be realised with small samples (e.g. Wermuth/Sommer/Wulff, 2004). Interviews or laboratory experiments (e.g. Schreckenberg, 2001) lack the realistic context and are costly and time-consuming.

An observation of the actual route choice behaviour with ANPR-systems in the road networks can overcome some of the limitations, as the real route choice behaviour of a large sample is observed. The number of required ANPR-locations depends on the road network. Figure 9 shows an example of a road network, which requires three ANPR-locations for an observation of two alternative routes.

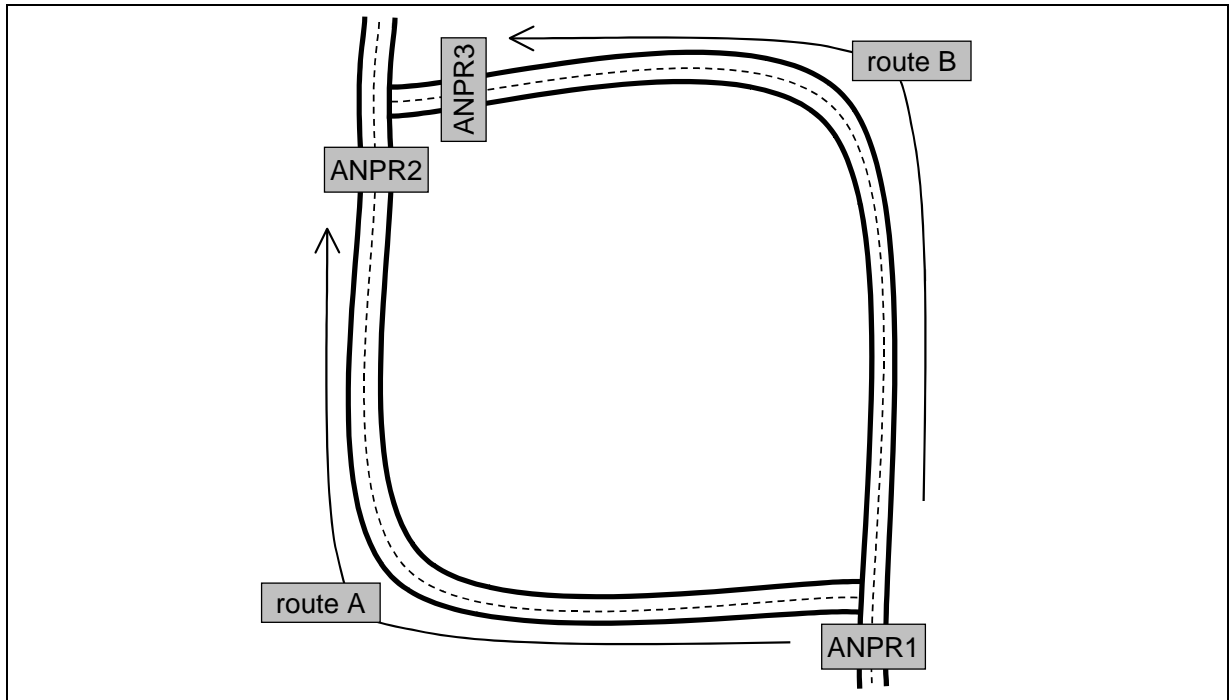


Figure 9: Observation of route choice behaviour with three ANPR-locations

There are, however, some disadvantages in observing the route choice behaviour with ANPR-systems. These are:

- There is no information about the driver and the trip purpose.
- Cars, which are detected at two ANPR-locations can be either through-traffic or traffic with a destination between the two ANPR-locations (e.g. courier service). The differentiation can only be done based on average travel times, which may be wrong in some cases (see chapter 3.3 for details).
- Travellers, who take a detour in order to avoid a traffic jam near an ANPR-location, will falsify the results.

Depending on the network and the traffic load, it is possible to collect a large sample size within one day, which allows comprehensive research on route choice behaviour. With statistical methods such as the maximum-likelihood estimation the relevance of possible factors for the route choice decision can be proven or rejected and finally the weighting of each factor in the utility function can be determined. Table 2 shows some potential influence factors for route choice behaviour:

Static	Dynamic
Trip length	Broadcasted travel news
Trip time in the unloaded network	Typical trip time for the time of the day
Static traffic sign	Variable message signs
Number of turning	Construction sights

Table 2: Potential influence factors for route choice behaviour

This allows the calibration of transportation models and better traffic control in operation centres. For an extensive examination of route choice behaviour see Schlaich/Friedrich (2008).

3.5 Estimation of O-D matrices based on the area code

The first one to three characters of a German licence plate number, separated by a hyphen, indicate the administrative district, where the vehicle is registered. The licence plate number “S – AB 123” for example stands for the City of Stuttgart with the area code “S”. In total more than 400 different area codes are used in Germany. Figure 10 shows a section of the German autobahn network in the South-West of Germany with the area codes.

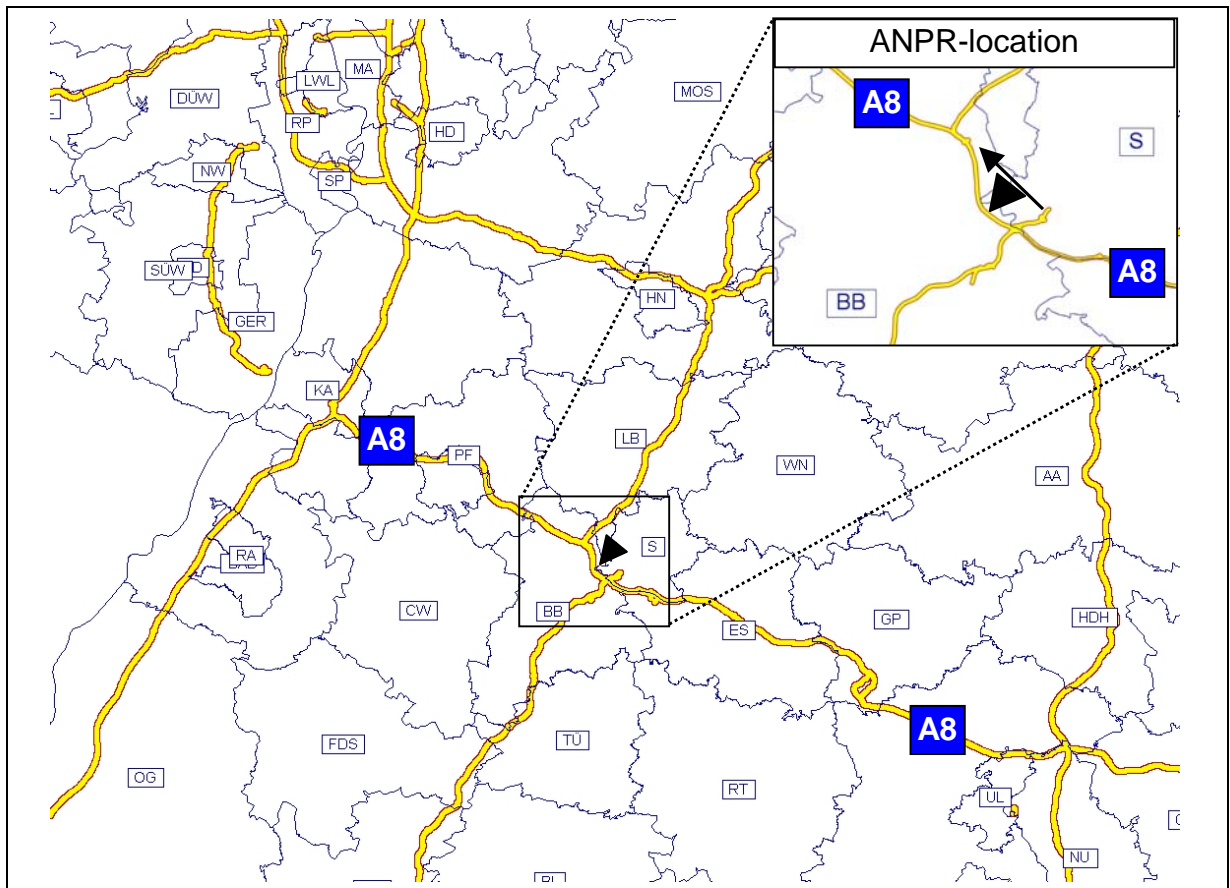


Figure 10: Autobahn network and area codes in South-West Germany with one ANPR-location.

Analysing the area code of a specific vehicle at a given cross section allows to estimate either the origin or the destination of this vehicle. Depending on its area code a vehicle travelling on the autobahn A8 south of Stuttgart from East to West can provide one of the following information:

- *Origin of the journey:* A vehicle with an area code located East of Stuttgart, e.g. UL (Ulm) or M (Munich), most likely is on its outward journey.
- *Destination of the journey:* A vehicle with an area code located West of Stuttgart, e.g. KA (Karlsruhe) or HD (Heidelberg), most likely is on its homeward journey.
- *Origin or destination of the journey:* For a vehicle with the area code of a nearby area, e.g. S (Stuttgart), it is not always possible to assume whether the vehicle is on a outward or homeward journey.

Whether the location of an area code is an origin or a destination of a journey can be determined by performing a shortest path search from the ANPR-location point to the centroid of all areas. Figure 11 illustrates this procedure.

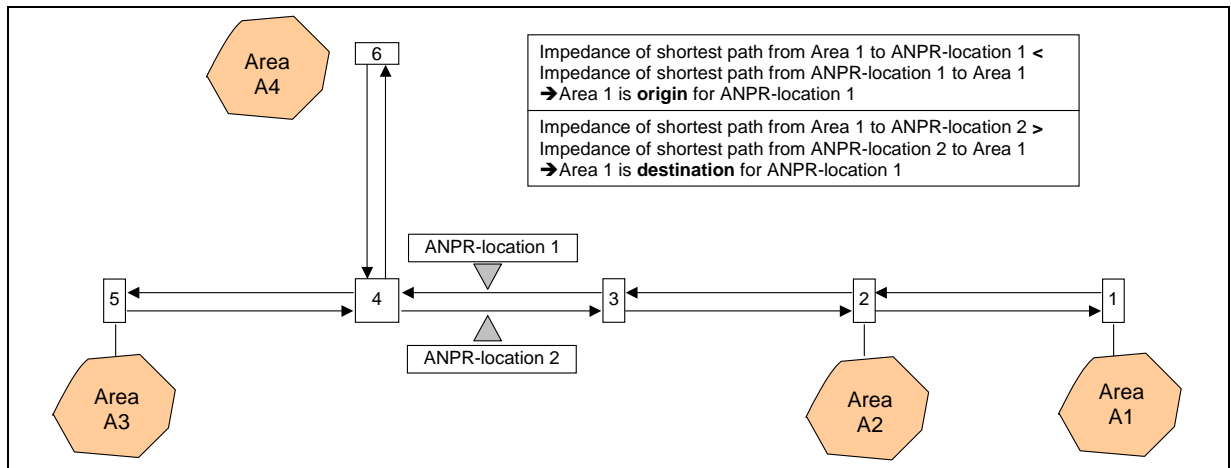


Figure 11: Rules for identifying origins and destinations for an ANPR-location.

Assuming that the destinations of vehicles with an identified origin are proportionally distributed to the destinations of vehicles with an identified destination and vice versa it is possible to derive an O-D matrix for any cross section. Figure 12 shows the origins and destinations for the ANPR-location described in Figure 10. The size of the circle and of the link bars indicates the volume.

It is obvious that such O-D matrices can only be generated for cross sections with a reasonable volume of long distance trips. The resulting matrices can be used to validate travel demand forecast models by comparing the survey results with the results of a selected link analysis.

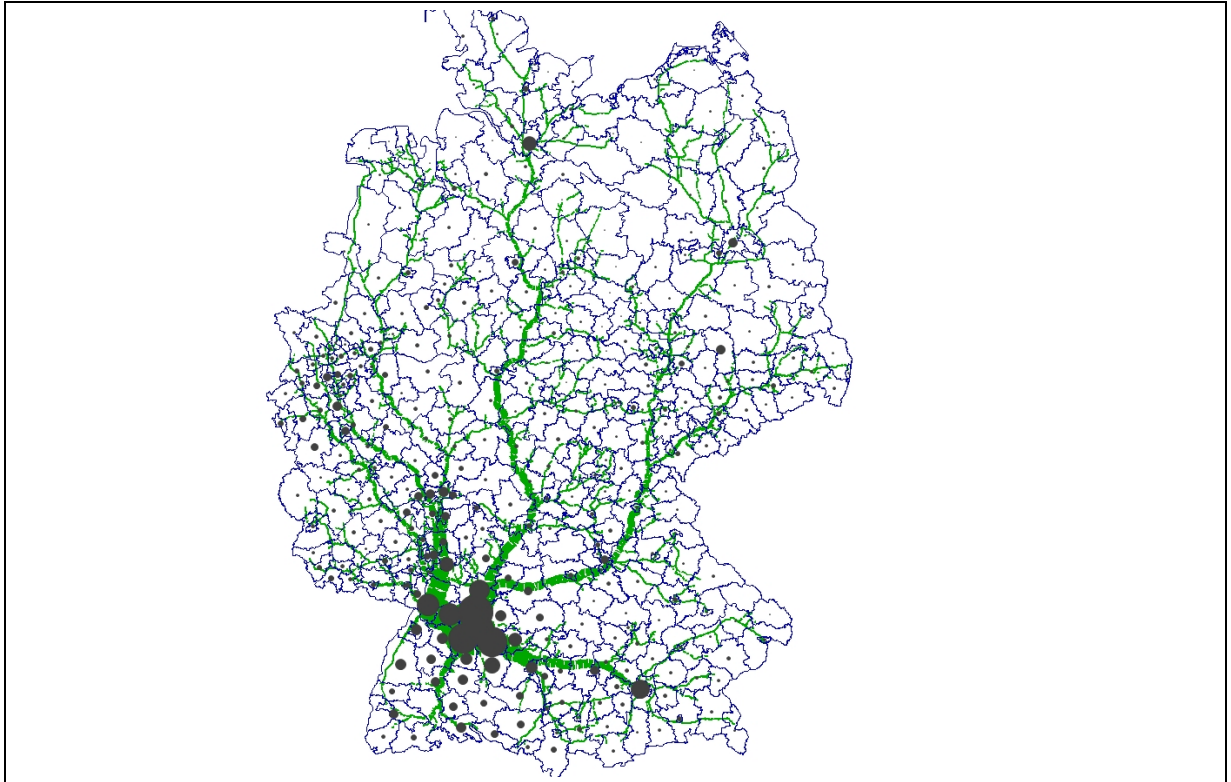


Figure 12: Estimated origins and destinations of vehicle trips traversing an ANPR-location on the A8 near Stuttgart.

4 CONCLUSION

The examples presented in this paper demonstrate the potential of ANPR-systems for transport planning, traffic engineering and traffic operation. The systems have overcome the experimental stage and provide a new data source bridging the gap between stationary and mobile detectors. The authors have applied the systems as mobile systems for one or two day surveys. For operators of urban roads or motorways a permanent installation would significantly improve the traffic state and incident detection. This can be used to inform the drivers and to optimise traffic control systems. At the same time the systems would provide valuable information on travel patterns. The investment costs for a single lane ANPR-System add up to approximately 20,000 €. This appears to be a high investment. Compared to infrastructure costs or the costs for traffic signals the investment, however, is relatively minor. With an investment of about 1 million € an average European City could detect travel times on arterial roads at a much better level than with other currently available methods. Experience in some places, e.g. in Munich (Grüber and Röhr, 2007), show that operators are starting to exploit this new data source.

5 REFERENCES

Grüber, B. and T. Röhr (2007). *Reisezeitmessungen setzen sich durch*. Straßenverkehrstechnik, 5.2007, 264-267.

Schreckenberger, M., R. Selten, T. Chmura, T. Pitz and J. Wahle (2001). *Experiments on Day-to-Day Route Choice*. Proceedings on 18th Dresden Conference on Traffic and Transportation Sciences, Dresden, Germany.

Wermuth, M., C. Sommer and S. Wulff (2004). *Erhebung der individuellen Routenwahl zur Weiterentwicklung von Umlegungsmodellen*. Research Report FE-Nr. 01.158/2002/CGB by order of 'Bundesministeriums für Verkehr, Bau- und Wohnungswesen', Braunschweig/Berlin, Germany.

Schlaich, J. and M. Friedrich. (2008). *Routenwahl in Autobahnnetzen*, forthcoming publication in 'Straßenverkehrstechnik', 2008.

ANPR-System contractors

CAT Traffic: <http://www.cat-traffic.de/>

Hi-Tech Solutions: <http://www.htsol.com/>

Messtechnik Mehl: <http://www.messtechnik-mehl.de>

PIPS Technology: <http://www.pipstechnology.com>